RADIO AT ULTRA-HIGH FREQUENCIES

RCA INSTITUTES TECHNICAL PRESS

RADIO AT ULTRA-HIGH FREQUENCIES

Technical Papers by RCA Engineers on Propagation, Transmission, Relaying, Measurement, and Reception Above 30 Mc.

Edited by

LEWIS M. CLEMENT

RALPH R. BEAL

DR. H. H. BEVERAGE

ROBERT S. BURNAP

DR. ALFRED N. GOLDSMITH

CHARLES W. HORN

APRIL, 1940

Published by

RCA INSTITUTES TECHNICAL PRESS

A Department of RCA Institutes, Inc.

75 VARICK STREET, NEW YORK, N. Y.

Printed in U.S.A.

PREFACE

ELECTROMAGNETIC waves in the ultra-high-frequency domain have found more numerous applications and rapidly increasing usefulness in recent times. The construction of vacuum tubes within which ultra-high-frequency energy may be generated, modulated, amplified, or demodulated, the technique of related and appropriate circuit designs and elements, and the development of effective radiating and receiving antenna systems have all shown great progress. A number of services have been attracted to the ultra-high-frequency channels including such diverse branches of radio as sound broadcasting, aircraft communication, television broadcasting, police communications, facsimile broadcasting, and radio relay systems. It is evident that a new and major branch of radio engineering is here in progress of development.

Attractive possibilities and considerable promise are shown by the u-h-f communication systems. Their advantageous features include extremely compact antenna systems or, alternatively, systems producing a concentration of energy with a resultant which is equivalent to a gain in transmitter power; compact and lightweight transmitting and receiving apparatus; greatly reduced levels of natural interference which permit comparatively low powers to be used for effective medium-range reception with low noise levels; a relatively high degree of constancy of the received signal irrespective of secular, diurnal, and atmospheric changes; and a comparative absence of the sky wave with consequent fairly definite limitation both of the service range and of the usual interference area. Certain of these characteristic advantages become increasingly evident in the higher frequency portions of the u-h-f domain and in the "micro-wave", sub-meter, or centimeter-wave region.

A further significant development, finding one of its first practical applications in the u-h-f field, is frequency modulation of the carrier wave. There exist numerous possible variations in the controlling parameters of the transmission, in the design and utilization of the transmitting equipment, and in the construction of the corresponding receivers. For this reason, as well as the high-quality communication possibilities of frequency modulation, this type of intelligence transmission has been the subject of engineering study for some time and increasingly attracts the attention of research and development workers in the u-h-f field.

It is therefore timely to present in convenient and assembled form a group of contributions to ultra-high-frequency radio engineering, including studies of frequency modulation, which originated in the laboratories of the Radio Corporation of America.

In the preparation of this book, a problem of selection was encountered. The large number of available and technically meritorious papers would in the aggregate have required a number of pages exceeding those which could conveniently be included in this volume. Accordingly it became necessary to choose certain of the papers for presentation in this book in summary form rather than in full. It must be emphasized that adherence to a subject-classification plan with suitably distributed emphasis determined the selection of papers to be published in full, and that the presentation of a paper in summary form in no wise indicates any lack of technical or literary merit or importance. The technical papers have been arranged according to the following general classification: Ultra-High Frequencies below 300 Mc, with the sub-divisions of Transmitting Methods and Equipment, Propagation and Relaying, Measurements, and Reception; and Ultra-High Frequencies above 300 Mc.

The RCA Institutes Technical Press gratefully acknowledges the courtesy of authorization of re-publication which has been extended to it by certain organizations in whose publications some of the papers in this book have appeared. These organizations include The Institute of Radio Engineers, McGraw-Hill Publishing Company, and the Radio Club of America.

The appreciation of the RCA Institutes Technical Press is also extended to those authors who have written original papers specially for the present volume. The comprehensive and relevant data concerning ultra-high-frequency phenomena here assembled are offered with the thought that they may speed developments in this field of radio.

Contents

PAPERS PUBLISHED IN FULL

Part I — Ultra-High Frequencies Below 300 Mc

Transmitting Methods and Equipment

		PAGE
Simple Television Antennas	P. S. Carter	1
Television Transmitting Antenna for Empire State Building	Nils E. Lindenblad	19
A Turnstile Antenna for Use at Ultra-High Frequencies	G. H. Brown	41
Frequency Control by Low-Power Factor Line Circuits	C. W. Hansell and P. S. Carter	53
A Cathode-Ray Frequency Modulation Generator	R. E. Shelby	76
Carrier and Side Frequency Relations with Multi-Tone Frequency or Phase Modulation	M. G. Crosby	86

Propagation and Relaying

A Study of U-H-F Wide-Band Propagation Characteristics	R. W. George	90
Ultra-High Frequency Propagation	M. Katzin	107
Frequency-Modulation Propagation Characteristics	M. G. Crosby	143
Frequency-Modulation Noise Characteristics	M. G. Crosby	160
The Service Range of Frequency Modulation	M. G. Crosby	203
Practical Application of an U-H-F Radio-Relay Circuit	J. E. Smith, F. H. Kroger and R. W. George	226
U-H-F Equipment for Relay Broadcasting	W. A. R. Brown	242

Measurement

Wide-Band Variable-Frequency Testing Transmitters	G. L. Usselman.	255
Field Strength Measuring Equipment for Wide-Band U-H-F Transmission	R. W. George	262
A New Method for Measurement of Ultra-High-Frequency Impedance	S. W. Seeley and W. S. Barden	272
A Survey of Ultra-High-Frequency Measurements	L. S. Nergaard	294

Reception

Vacuum Tubes of Small Dimensions for Use at Extremely High Frequencies	B. J. Thompson and G. M. Rose, Jr.	334
Simple Antennas and Receiver Input Circuits for Ultra-High-Frequencies	R. S. Holmes and A. H. Turner	349

Part II—Ultra-High Frequencies Above 300 Mc

Magnetron Oscillators for the Generation of Frequencies Between 300 and 600 Mc	G. R. Kilgore	360
An Ultra-High-Frequency Power Amplifier of Novel Design	A. V. Haeff	378
Development of Transmitters for Frequencies Above 300 Mc	Nils E. Lindenblad	386
Transmission of 9-Cm Electromagnetic Waves	I. Wolff and E. G. Linder	421

PAPERS PUBLISHED IN SUMMARY FORM

Part I — Ultra-High Frequencies Below 300 Mc

Transmitting Methods and Equipment

PAGE

Frequency Assignments for Television ..... E. W. Engstrom and C. M. Burrill 430
Television Transmitters Operating at High Power and Ultra-High Frequencies ..... J. W. Conklin and H. E. Gihring 430
The Requirements and Performance of a New Ultra-High-Frequency Power Tube ..... W. G. Wagener 431
Effect of Electron Transit Time on Efficiency of a Power Amplifier ..... A. V. Haeff 431
On the Optimum Length for Transmission Lines Used as Circuit Elements ..... B. Salzberg 431

Propagation and Relaying

A Study of the Propagation of Wavelengths Between 3 and 8 Meters ..... L. F. Jones 432
An Urban Field-Strength Survey at 30 and 100 Mc ..... R. S. Holmes and A. H. Turner 433
Observations on Sky-Wave Transmission on Frequencies Above 40 Mc ..... DeWitt R. Goddard 433
Notes on Random Fading of 50-Mc Signals Over Non-Optical Paths ..... K. G. MacLean and G. S. Wickizer 434
Application of Abelian Finite Group Theory to Electro-Magnetic Refraction ..... Russell A. Whiteman 434

Measurement

The Use of Gas-Filled Lamps as High-Dissipation, High-Frequency Resistors, Especially for Power Measurements ..... E. G. Linder 435
Measurements of Admittances at Ultra-High Frequencies ..... J. M. Miller and B. Salzberg 435
Electrical Measurements at Wavelengths Less than Two Meters ..... L. S. Nergaard 436
Radio-Frequency Generator for Television Receiver Testing A. H. Turner 437

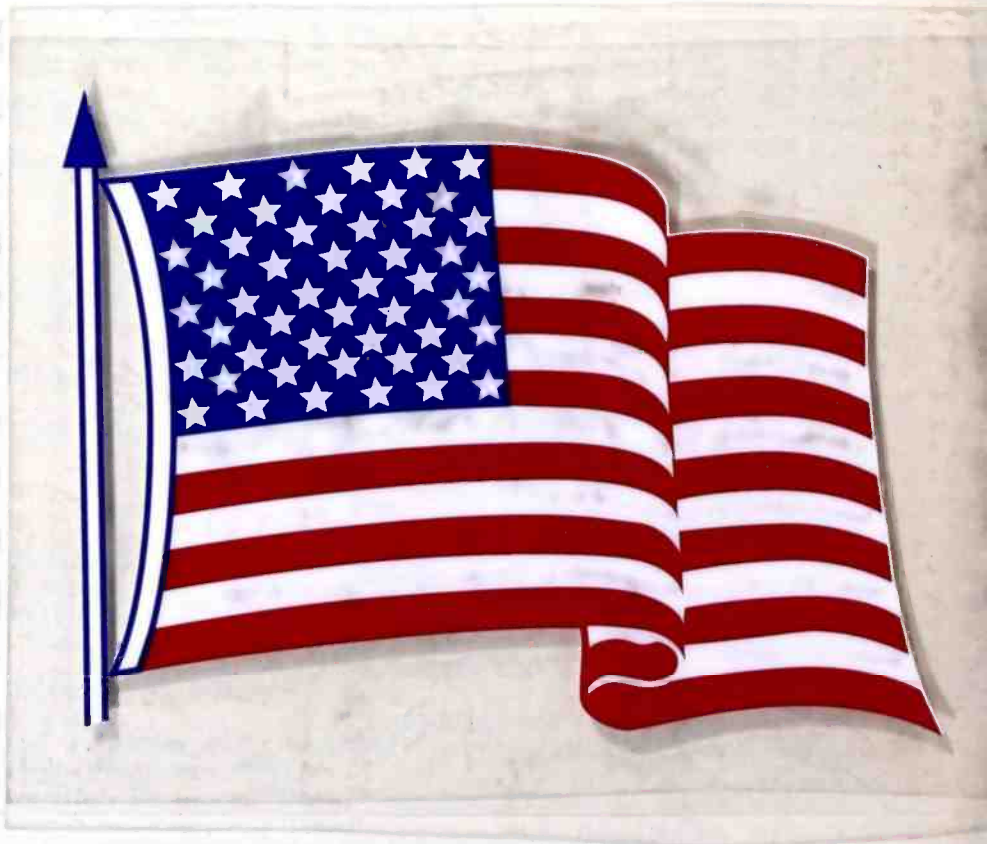
Reception

Input Resistance of Vacuum Tubes as Ultra-High-Frequency Amplifiers ..... W. R. Ferris 437
Analysis of the Effects of Space Charge on Grid Impedance ..... D. O. North 438
Multi-tube Oscillators for the Ultra-High Frequencies ..... P. D. Zottu 439
Recent Developments in Miniature Tubes ..... B. Salzberg and D. G. Burnside 440
Effects of Space Charge in the Grid-Anode Region of Vacuum Tubes ..... B. Salzberg and A. V. Haeff 441
An Electron Oscillator with Plane Electrodes ..... B. J. Thompson and P. D. Zottu 442
The Developmental Problems and Operating Characteristics of Two New Ultra-High-Frequency Triodes ..... W. G. Wagener 442
A Push-Pull Ultra-High Frequency Beam Tetrode ..... A. K. Wing 443
New Television-Amplifier Receiving Tubes ..... A. P. Kauzmann 443
Design and Use of "Acorn" Tubes for Ultra-High Frequencies B. Salzberg 444
Review of Ultra-High-Frequency Vacuum-Tube Problems B. J. Thompson 444
Construction and Alignment of the Television Receiver C. C. Shumard 445

Part II—Ultra-High Frequencies Above 300 Mc

Application of Frequencies Above 30,000 kc to Communication Problems ..... H. H. Beverage, H. O. Peterson and C. W. Hansell 445
Notes on Propagation at a Wavelength of 73 Centimeters ..... Bertram Trevor and R. W. George 446
Description and Characteristics of the End-Plate Magnetron E. G. Linder 447
The Magnetron as a High-Frequency Generator ..... G. R. Kilgore 447
Micro-Waves in NBC Remote Pickups ..... R. M. Morris 448

PART I



PART I

SIMPLE TELEVISION ANTENNAS*

By

P. S. Carter

Engineering Department, R.C.A. Communications, Inc.

Summary—The frequency band widths demanded by high-definition television have considerable range when considered in relation to resonant circuits. The transmitting antenna and transmission-line systems must, therefore, meet stringent requirements if multiple images or ghosts in the received picture are to be avoided.

Before discussing the characteristics of particular antenna systems, the transmitting and receiving antenna problems are considered. The input impedance of a transmission line, even when loaded with a resonant circuit having a Q as low as 2, undergoes considerable variation with frequency within the transmission band. If the television receiver is designed to present a pure resistance to its transmission line, equal to the characteristic impedance of the latter throughout the transmission frequency band, the receiving antenna requirements are not difficult to meet.

The measured impedance-frequency characteristic of a half-wave dipole of large diameter conductors, when compared with that obtained for a similar antenna of small diameter conductors, shows the advantages of the former.

A method of impedance matching has been devised whereby the usual narrowing of the useful frequency band caused by impedance transformation is overcome.

The "folded dipole" antenna and combinations of these units are superior to ordinary dipoles for television purposes. Measurements indicate that ground and other reflecting surfaces considerably affect the impedance-frequency characteristics of antennas.

The use of a type of antenna called the "double cone" or "hour glass" antenna results in a very flat impedance-frequency characteristic at the input terminals of a transmission line over a wide range of frequency. By properly proportioning the dimensions of this antenna its impedance can be made to match the characteristic impedance of all practical open-wire transmission lines. The current and electric field distributions along the surfaces of the conical conductor have been measured. The theory of this antenna is briefly considered.

Curves of the characteristics of the systems discussed are included. For comparison purposes the measurements of line reflection vs. frequency for the several antenna systems considered are shown by a family of curves in a single figure.

THE frequency band demanded by high-definition television transmission has a width of a few megacycles. Its ratio to the carrier frequency is of the order of 10 per cent to 15 per cent, a very substantial range when considered with respect to resonant circuits. The transmission lines used in conjunction with television antennas, whether for reception or transmission, rarely have lengths equivalent to less than two or three wave lengths.

* Presented at the National Convention of the Institute of Radio Engineers, San Francisco, June 27-30, 1939.

Reprinted from *RCA Review*, October, 1939.

Unless the impedance of a transmitting antenna perfectly matches the characteristic impedance of its associated transmission line at all frequencies within the transmitted band, displaced images or ghosts may appear on the screen of the receiver. Ghosts may also be caused by imperfect matching between the input impedance of the receiver and the characteristic impedance of its transmission line. In addition to the ghost phenomena the pictures may be more or less distorted by poor impedance matching between the transmitter and its transmission line or between the receiving antenna and its transmission line. It would thus appear that the impedances at both terminals of both the transmitter and receiver transmission lines should be perfectly matched throughout the frequency band. Since practical transmitters

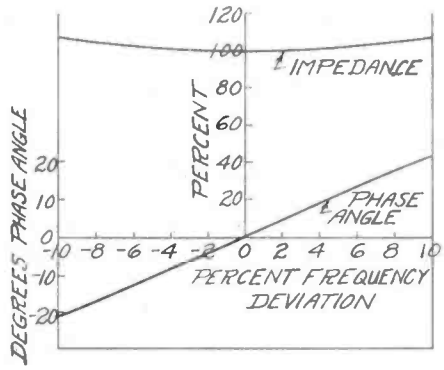


Fig. 1

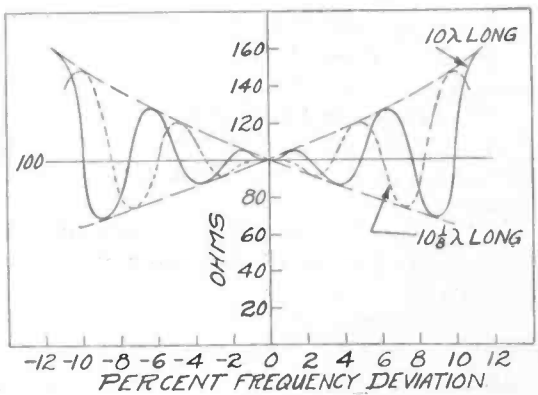


Fig. 2

and receivers are not electrically equivalent in all respects it seems desirable to consider each end of the space circuit as a separate problem before investigating particular antenna systems.

The Transmitting-Antenna Problem

Unless the transmitting antenna presents a pure resistance load to its associated transmission line and in addition this resistance is equal to the characteristic impedance of the line at all frequencies within the transmission band, the input impedance of the system, as seen by the transmitter, will vary with frequency. The phenomena involved and the approximate rules governing them can best be explained by a simple though somewhat idealized combination. Let us assume that a transmission line ten wave lengths long feeds an antenna having an impedance-frequency characteristic like that of a simple series-tuned circuit. Also, let us suppose the circuit to have a resistance of one hundred ohms, a reactance of two hundred ohms, or Q of two, and that the characteristic impedance of the transmission line is one hundred ohms. Figure 1 shows the impedance characteristic of

the load circuit by itself. The magnitude of the impedance is nearly constant through a frequency band of 20 per cent of the resonant frequency, and, at first sight the antenna represented by this circuit might appear ideal for television transmission. Figure 2 shows the input impedance of the line terminated with this antenna, which is far from constant.

The large oscillation in the curve of Figure 2 is due to reflection caused by the reactive component of the load impedance since the magnitude of the impedance is nearly constant over a considerable range of frequency. Charts have been developed by the writer by which the magnitude and phase angle of the reflected wave and the complex input impedance of the transmission line may be rapidly determined

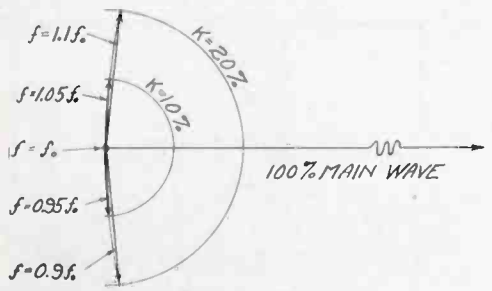


Fig. 3

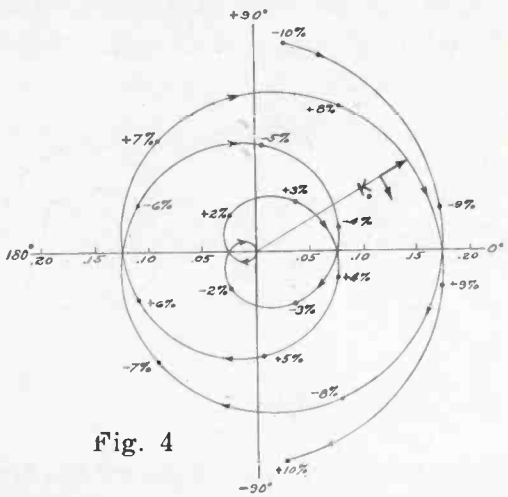


Fig. 4

for all conditions of load impedance.[1] Curves like that of Figure 2 are easily computed from the charts. The dashed curve of Figure 2 shows the effect of increasing the length of the line one-eighth wave length. It will be noted that this curve is symmetrical with respect to the resonant frequency as well as the solid curve for a length of 10 wave lengths. Such symmetry exists for lengths equal to any integral multiple of an eighth wave length. An increase of the length to 10¼ wave lengths results in an impedance curve approximately equivalent to the solid curve turned upside down. In order to avoid confusion only two curves are shown in the figure but the limits of the oscillation of the input impedance for any length of line are the two envelope curves indicated. Figure 3 shows the variation in the reflected wave vector at the load for the conditions assumed, as the frequency is increased from 90 per cent to 110 per cent of the resonant frequency. It will be noted that the reflected-wave vector leads or lags the main-wave vector

[1] Charts for Transmission-Line Measurements and Computations by P. S. Carter, *RCA Review*, January 1939.

by nearly 90 degrees for frequencies above or below resonance respectively.

Let us consider the relation between the input impedance of the line and the reflection coefficient. Line length in terms of wave length is directly proportional to frequency and therefore a 10 per cent increase in frequency increases the length of our 10 wave length line by one wave length. As we proceed from the antenna toward the transmitter a vector representing the main wave is rotated in a leading direction through an angle of 10 times 360 degrees at the carrier frequency, while the vector representing the reflected wave is rotated in the lagging direction by the same angle. At the input end of the line, the relation between the main wave and reflected-wave vectors is the same as though we had rotated the reflected-wave vector through an angle of 20 times 360 degrees without rotating the main

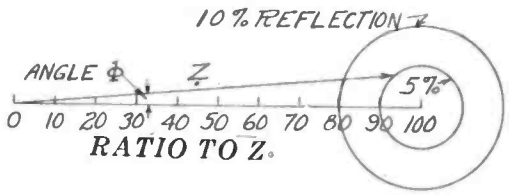


Fig. 5

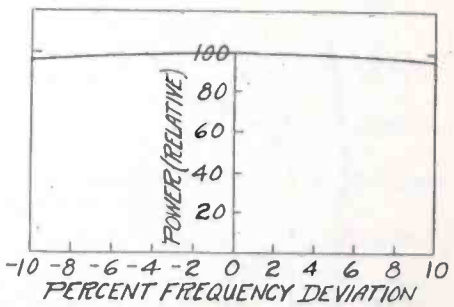


Fig. 6

wave vector. Figure 4 shows the trace of the terminus of the reflected-wave vector at the input terminal of our system when we assume the main vector fixed and the frequency is increased from 90 per cent to 110 per cent of the resonant frequency. The resulting input impedance has already been shown in Figure 2.

If the reflected-wave vector were of a constant magnitude of 5 per cent throughout the band the input impedance would oscillate between 90 and 110 ohms as the frequency varies. The reason for this is that when the frequency is of such a value as to bring the reflected and main-voltage vectors into phase the vectors representing the main and reflected current waves are in phase opposition. The resultant voltage is then 105 per cent of the main-wave voltage and the resultant current 95 per cent of the magnitude of the main-current wave vector. The input impedance is therefore $(105/95) \times 100 = 110$ ohms, approximately. When the frequency is such that the main- and reflected-voltage vectors are in phase opposition the above relations are reversed and the input impedance becomes approximately 90 ohms.

If we agree on a maximum allowable variation in the input terminal impedance, we may lay down tolerance rules. Let us assume the allow-

able limits of this impedance to be 90 per cent and 110 per cent of the characteristic impedance. The reflection coefficient must then be less than 5 per cent throughout the frequency band. If the magnitude of the antenna impedance remains substantially constant and equal to the characteristic impedance of the line the reactance component, whether capacitive or inductive, must not be greater than 10 per cent of the resistive component. In most practical systems both components vary with frequency and the terminus of the load impedance vector must lie within the circle as shown in Figure 5.

The transmission lines for use with television transmitting installations usually can be designed to have efficiencies of better than 90 per cent; and the consideration of power losses may, therefore, be neglected without seriously affecting the conclusions resulting from the present discussion.

If the transmitter is either very loosely coupled to its transmission line so as to act, in effect, as a constant-voltage generator, or very

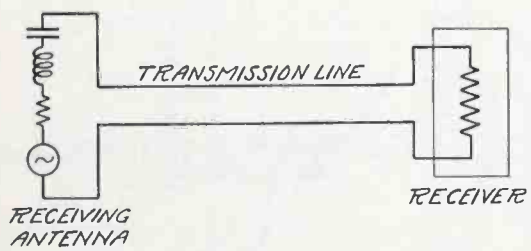


Fig. 7

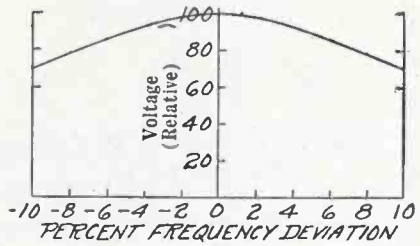


Fig. 8

tightly coupled so that it becomes approximately equivalent to a constant-current generator, the power radiated from the antenna would have the same variations as the impedance curve of Figure 2. The power output under either assumption would be very low. If the transmitter could be represented by a voltage in series with a pure resistance equal to the characteristic impedance of the line the power output of our assumed antenna system would be nearly constant over the frequency band in spite of the wide variations in the input terminal impedance of the transmission line. Under this assumption the power output would then be as shown in Figure 6. However, such an equivalent circuit representation for an actual transmitter is far from justified. An alternator in series with a series-tuned circuit having a low value of resistance relative to the characteristic impedance of the transmission line is a more nearly true equivalent. The variation in radiated power will then be considerable. We, therefore, conclude that in a practical television transmission system the input terminal impedance of the transmission line feeding the antenna should show little variation with frequency within the transmission band.

The Receiving-Antenna Problem

Within regions where the field strength is high a television receiver may be so designed as to have a substantially constant input resistance equal to the characteristic impedance of its associated transmission line within the television transmission band. The receiving antenna is equivalent to a generator in series with a complex impedance. Both the resistance and reactive components of this equivalent impedance vary with frequency but as a first approximation the antenna may be represented by an alternator together with a series tuned circuit, the equivalent circuit for the complete system being as shown in Figure 7.

Let us assume the receiver to act as a pure resistance and the

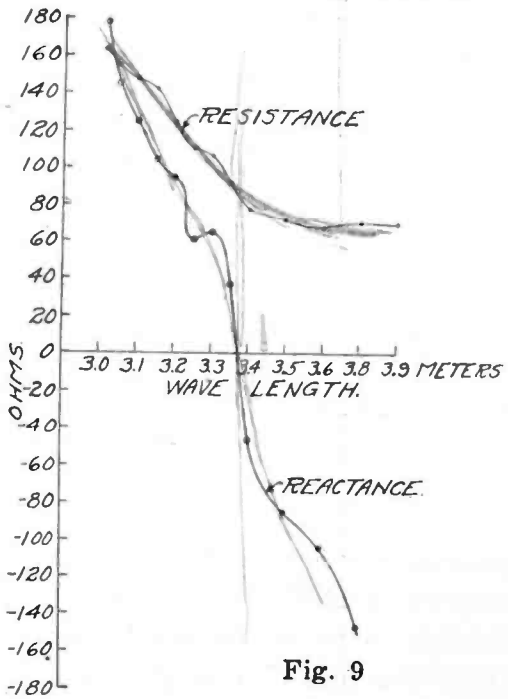


Fig. 9

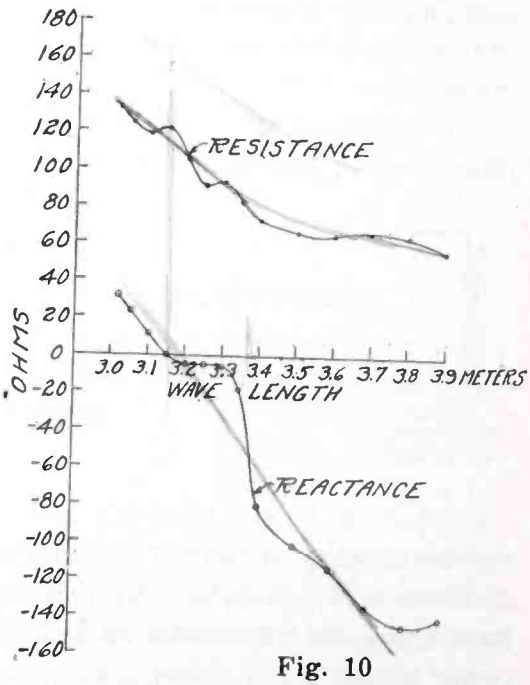


Fig. 10

antenna to be the same as that discussed in the transmitting problem where the equivalent circuit has a Q of 2. If the transmission line is designed to match the resistance of the antenna, the power in the receiver is the same as that shown in Figure 6 for the transmitting condition where the transmitter is an alternator in series with a pure resistance equal to the characteristic impedance of the line. Thus we see that, due to the fact that a receiver may be designed to have a constant resistance over a considerable band, an antenna and transmission line system which is far from satisfactory when used with a transmitter may serve as an excellent receiving system.

Figure 8 shows the voltage developed across the input resistance of a receiver when the receiving antenna has an equivalent Q of 10, and the receiver is assumed to be matched to the transmission line. This curve is true regardless of whether the characteristic impedance of

the transmission line matches the antenna resistance or not. However, maximum signal strength is, of course, obtained when these impedances are matched. This figure, therefore, indicates that an antenna with the relatively high Q of 10 may be satisfactory for reception purposes.

For locations where the signal strength is relatively low, receiver amplification cannot well be sacrificed in order to obtain a wide-band constant-resistance input characteristic. A good match between the antenna and transmission line impedances is then desirable for two reasons,—(1) Under these conditions the receiver circuits are quite similar, insofar as the transmission line is concerned, to the transmitter circuits previously discussed. Multiple images may be caused by impulses which, after having been reflected by the receiver, are

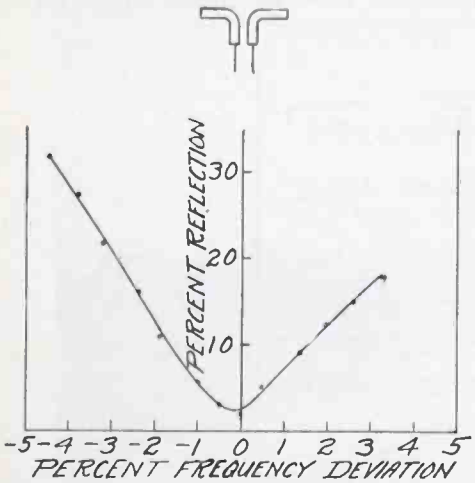


Fig. 11

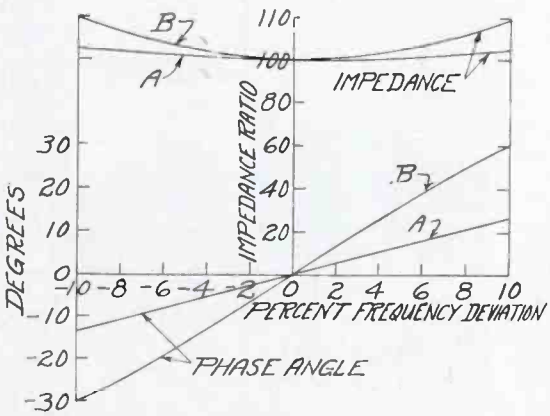


Fig. 12

again reflected at the antenna and returned to the receiver at a later time corresponding to the time required for the wave to make one or more round trips to the antenna, i.e. roughly one-tenth microsecond for each fifty feet of line. (a) Noise generated in the receiver itself becomes comparable with signal strength and the maximum energy should be drained from the antenna to overcome this noise. This can only be accomplished when the line and antenna impedances are equal.

Some transmission lines in use with receivers have efficiencies lower than 10 per cent even when their lengths are not greater than 100 feet. The use of an efficient line together with a proper matching of antenna and line impedances may, in some instances, give an improvement of 100 to 1 in received power or 10 to 1 in received signal voltage.

Half-Wave Dipole Antennas

Figure 9 shows the impedance vs. frequency characteristic of a simple horizontal half-wave dipole of No. 10 wire when fed by an

open, No. 10 wire, transmission line. The height above ground was 180 cm or a little under a half wave length at the resonant frequency. The ripple in these curves is without doubt due to the wave reflected from the side of a nearby sheet-metal building, since the frequency between ripple crests corresponds quite accurately with the time of travel from the antenna to the building and back again.

Figure 10 is a similar characteristic for a dipole of 1½-inch copper pipe positioned at the same height and fed by the same line. By comparing these two figures the effect of the increase in the diameter is seen. The reactance of the pipe antenna is less than 10 per cent of the resistance over a band width of nearly 7 per cent but within this range the resistance changes from 80 to 120 ohms even if we average out the ripples. The larger-diameter antenna is very much superior

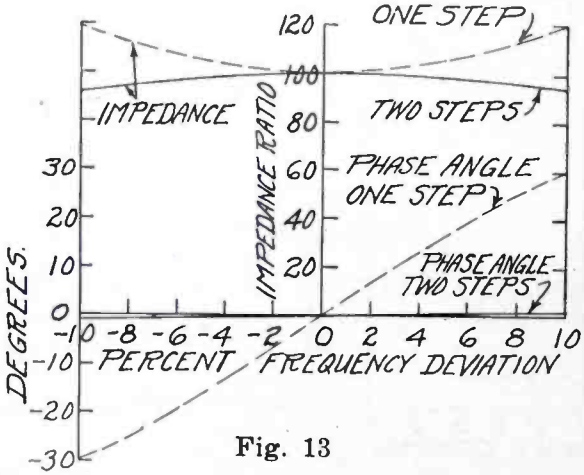


Fig. 13

to the No. 10 wire dipole because of the improved reactance characteristic of the former even though the improvement in the resistance variation is not very great.

The 1½-inch pipe dipole might be satisfactory for some transmitting installations if the band width were not further narrowed by impedance-matching circuits necessary to transform this impedance up to the value of the characteristic impedance of a practical line.

Figure 11 shows the reflection coefficient vs. frequency for a 1½-inch pipe dipole and quarter-wave impedance matching system in which the pipes are bent so that a single section of pipe serves both as a leg of the dipole and as one conductor of the impedance matching circuit. The band width at which the reflection is less than 5 per cent is only 1.5 per cent.

Before proceeding further it would appear well to consider briefly the effect of impedance matching circuits upon the impedance-frequency characteristics of antenna systems. Let us assume a quarter-wave line section loaded with a pure resistance. When this line is used

for transforming in the impedance ratio of 2 to 1, the resulting input terminal impedance characteristic is that shown by curves *A* of Figure 12. When used for an impedance transformation ratio of 16, the characteristic is as shown in curves *B* of the same figure.

It is apparent from these curves that the detrimental effect of quarter-wave impedance matching circuits is in proportion to the magnitude of the impedance transformation ratio. This law generally holds for all types of impedance matching circuits. It will be observed that the quarter-wave line circuit gives a symmetrical input impedance characteristic. Some types of circuits, such as circuits formed with shunt reactances at particular positions along the transmission line

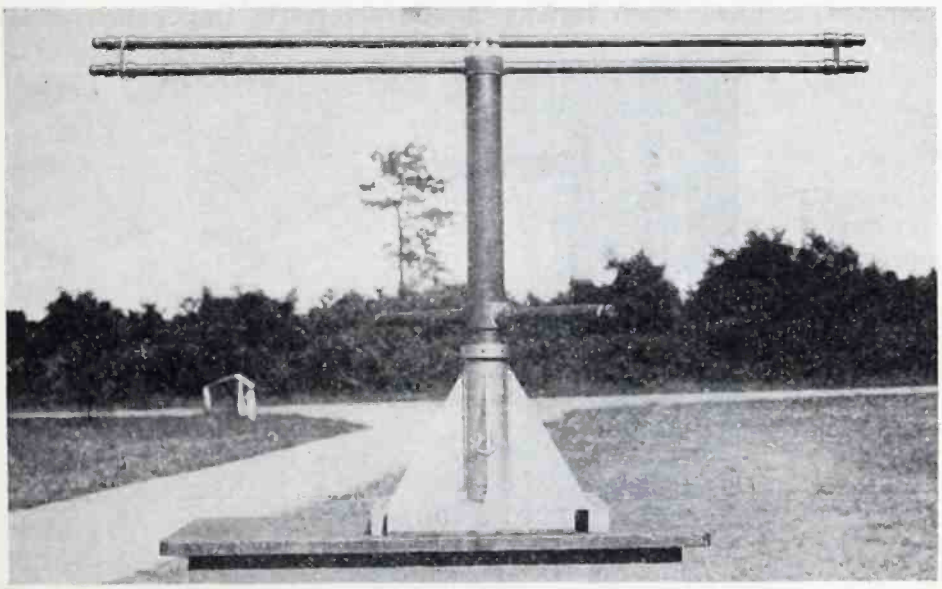


Fig. 14—Dipole for 81-86 Mc frequency sweeping transmitter used in television propagation survey.

produce unsymmetrical impedance-frequency characteristics. We shall not take the time to consider other types of circuits or the impedance matching problem in detail.

The difficulties due to an impedance transformation may be almost entirely overcome by dividing the transformation into two or a greater *even* number of steps. As an illustration let us assume that we use two quarter-wave line sections in series to transform a load impedance from $16R$ to R. If we make the characteristic impedance of the line section near the load $8R$ and the remaining section $2R$ the first section transforms from $16R$ to $4R$ and the second section from $4R$ to R. Figure 13 shows the input impedance characteristic when such a system is used; the curve for the same transformation in one step being shown in broken lines. The reactance present when a single step is used has been practically eliminated by the use of two steps.

The Folded Dipole

Figure 14 is a photograph of a transmitting type of antenna which we call "the folded dipole". It was developed in order to do away with impedance matching circuits and their detrimental effects upon the frequency band width and also to obtain a mechanically strong structure. This particular unit was installed at the 85th floor of the Empire State Building (see Figure 15) and has been in use for several months in order to provide a temporary system while changes were being made in the main antenna and transmission line. The impedance-frequency characteristic of this antenna as measured on its associated transmission line is shown in Figure 16. The transmission line was a pair of concentric lines, each having a characteristic impedance of 120

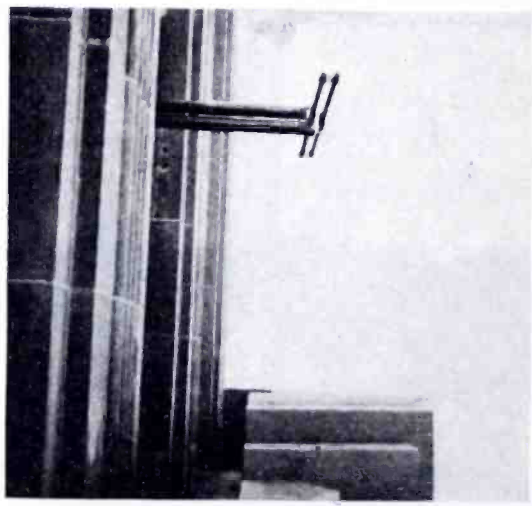


Fig. 15—Folded dipole.

ohms. The antenna perfectly matched the line at 46.5 megacycles, the carrier frequency in use at that time. Other units for survey tests in New York City were located near the top of the tower of the Empire State Building and are shown in Figure 17 and Figure 18.

This type of antenna consists of two closely spaced half-wave dipoles connected together at their ends. One of the dipoles is broken at its center where it is fed from a balanced transmission line. The instantaneous currents in both units are in the same direction in space while both are flowing toward a nodal point at one extremity of the radiating structure. The current distribution does not differ greatly from that for an ordinary half-wave dipole and is approximately sinusoidal.

Since the two radiators are very closely spaced in terms of wave length the radiation pattern is essentially the same as the pattern for an ordinary single half-wave dipole. The total power radiated per total loop current squared, or radiation resistance, is therefore about 73

ohms. However, if the diameters of the two radiators are equal, this same radiated power is equivalent to a radiation resistance with respect to the current in one branch of four times 73 ohms or 292 ohms. The latter value of resistance is that which is seen by the transmission line at its terminals. This type of antenna thus serves the double purpose of a radiator and impedance matching transformer.

When three radiators of equal diameters are arranged in accordance with this method, a transformation ratio of nine is obtained. Any desired ratio of transformation may be obtained by the use of two or more radiators of unequal diameters. In such an arrangement of two units wherein the smaller-diameter cylinder is fed from the

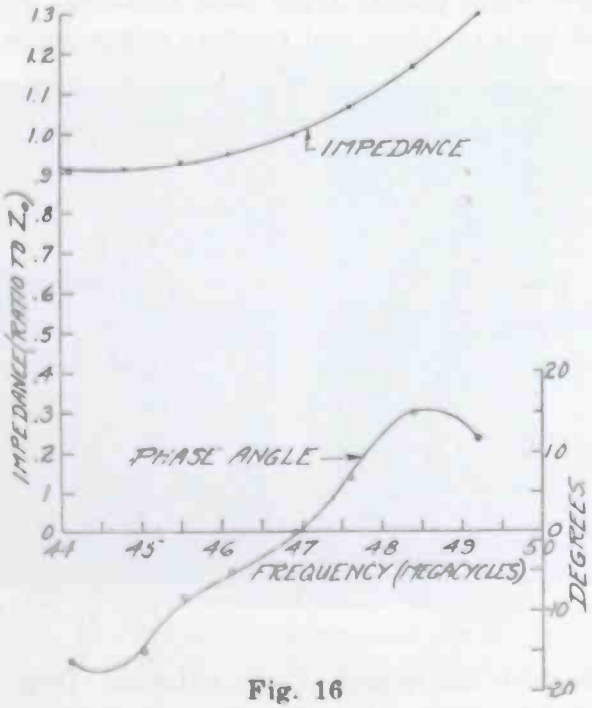


Fig. 16

transmission line the transformation ratio is greater than four since the greater of the two currents flows along the larger pipe.

When a single folded dipole is used with a long transmission line the input impedance-frequency characteristic does not quite meet the most rigid transmitting requirements. However, when two such units are arranged in criss-cross fashion and fed in quarter-phase relation by means of two branch feeder lines differing in length by a quarter-wave length the impedance characteristic presented to the main line at its junction with the branch feeders is quite satisfactory for present band widths. Figure 19 is a photograph of such a unit. The impedance characteristic for such an arrangement is shown in Figure 20. The resulting small reactance component is due to the impedance-inverting property of a quarter-wave line section.

The same advantages insofar as band width alone is concerned may be obtained from an array of four of these units in a square when one pair of units is fed in quarter-phase relation with the other pair in order to provide an antenna for the radiation of horizontally polarized waves. However, unless the square is small in terms of wave length the horizontal radiation pattern is far from circular.

Impedance and Height

The impedance characteristic of an antenna, when located near the ground or other reflecting surfaces, may differ considerably from the characteristic when remote from such surfaces. Figure 21 shows the variation of both resistive and reactive components of the imped-

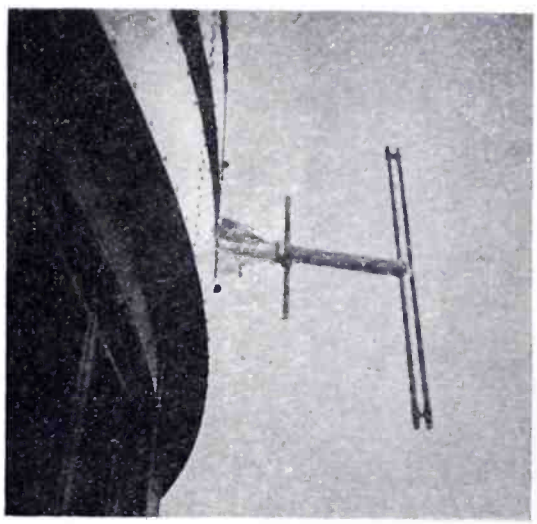


Fig. 17—Folded dipole.

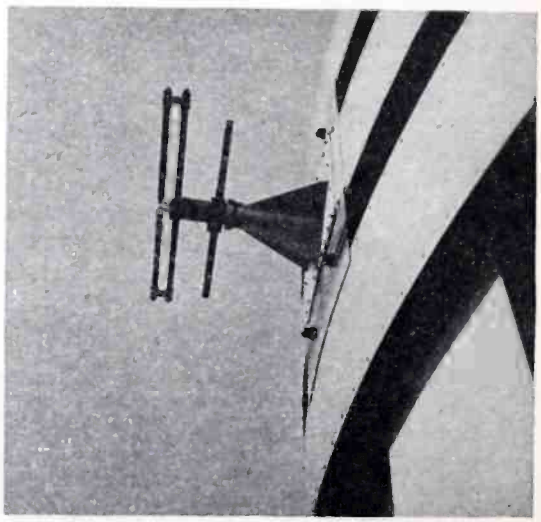


Fig. 18—Folded dipole.

ance of a dipole with the height of the antenna. Data were taken at several different frequencies, but the curve shown corresponds to the frequency at which the average reactance is approximately zero. The measurements were made at Rocky Point, New York, where the soil is mostly quartz sand. The dielectric constant of the soil is around 9 and conductivity is extremely low. In a previous paper[2] the writer has shown a method by which both the resistance and reactance components of the radiation impedance of an antenna above a perfectly conducting plane may be determined. The theoretical curves obtained by this method, assuming a perfectly conducting ground, are shown in dotted lines in Figure 21 for comparison purposes. It will be noted that the variation in impedance for the antenna above the extremely poor ground at Rocky Point is nearly as great as when a perfectly

[2] Circuit Relations in Radiating Systems, *Proc. I.R.E.*, Vol. 20, No. 6, June 1932.

conducting ground is assumed and that the approximate heights for the maximums may be obtained by adding about 0.1 wave length to the actual height and assuming ideal ground.

The effect of a metal building, at some distance from the antenna, upon the impedance characteristic has already been shown in Figure 9. These two examples of the effect of reflecting surfaces should serve to explain why it is seldom possible to predict the impedance characteristic of an antenna system accurately.

The Double-Cone or Hour-Glass Antenna

Figure 22 is a photograph of an arrangement of two copper cones together with an open two-wire transmission line. The impedance-

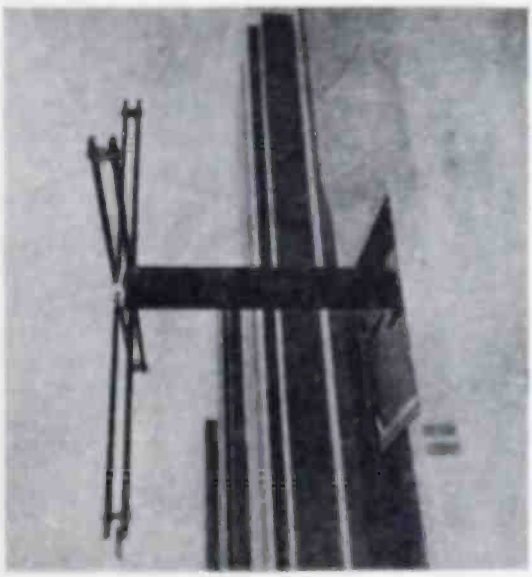


Fig. 19—Folded dipole antenna for study of propagation characteristics of circular polarization.

frequency characteristic is shown in Figure 23. It will be noted that the magnitude of the impedance curve is nearly constant and the phase angle less than 10 degrees within a frequency band of nearly 20 per cent. This characteristic is far superior to that of any of the antennas so far discussed.

The impedance characteristics of a number of double-cone antennas of sheet copper have been measured in order to determine the laws governing the relationship between length, angle of revolution of the cones, and impedance presented to the transmission lines. Tests were also made with several sizes of wire-cage cones. When twelve or more wires were used for the cage with these wires converged into a small solid metal cone to form the apices of the large cones, the impedance characteristics were found to be nearly as good as those for the sheet cones.

Figure 24 shows the effective impedance as a function of angle of revolution of the cones. The length of each cone taken along the surface from apex to rim should be about 0.365 wave length at the mid-frequency regardless of angle of revolution, at least for the impedance values usually encountered in practice.

Figure 25 shows the measured current distribution along the surface of a double-cone antenna. Strictly, the curve represents the product of the magnetic field intensity near the conducting conical surface multiplied by the circumference of the cone at the position of measurement but according to the fundamental laws of Maxwell this product is equivalent to the relative total current flowing along

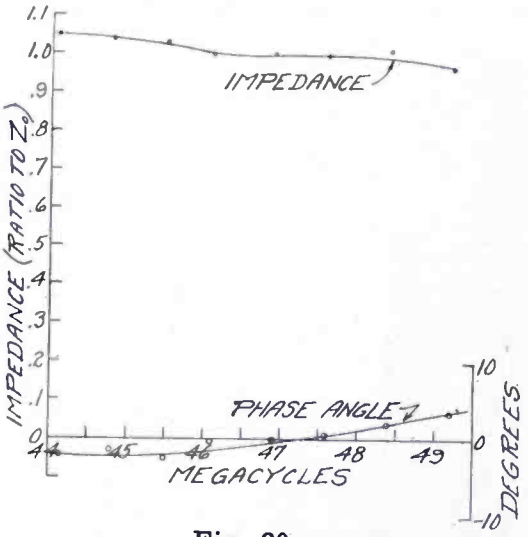


Fig. 20

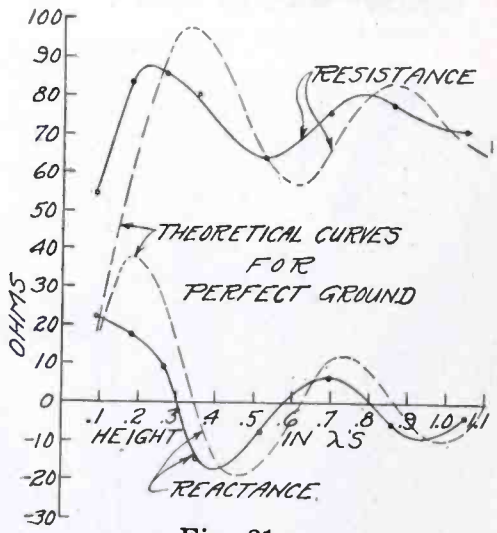


Fig. 21

the surface of the cone. It will be noted that the current distribution departs greatly from the usually assumed sine-wave form and varies little with distance for positions near the apex.

The distribution of the voltages between the surfaces of the conical sheets is also shown in Figure 25. In order to interpret these two distribution curves and obtain an understanding of the nature of the phenomena involved it is necessary to consider the field configurations of the electromagnetic waves in the vicinity of such conducting surfaces.

Radio engineers are familiar with the principles of radio-frequency transmission lines but usually consider the phenomena in terms of inductance and capacitance visualized as coils and condensers. In order to obtain a clear understanding of electromagnetic wave propagation we must take the more fundamental viewpoint of Maxwell, Faraday, and Heaviside. The energy is carried by the space surrounding the conductors while the currents flowing in the surfaces of the conductors of a two-wire transmission line are of secondary importance only. The

small portion of energy which penetrates the wires serves no useful purpose and is wasted in heat. In any wave front, which for a transmission line is a plane cutting the wires at right angles, the lines of electric force are arcs of circles terminating on the conductors while the lines of magnetic force are a complete system of circles surrounding one or the other of the conductors. From the inverse-distance relations governing the electric and magnetic field strengths it should be obvious that by far the larger portion of the energy of the electro-

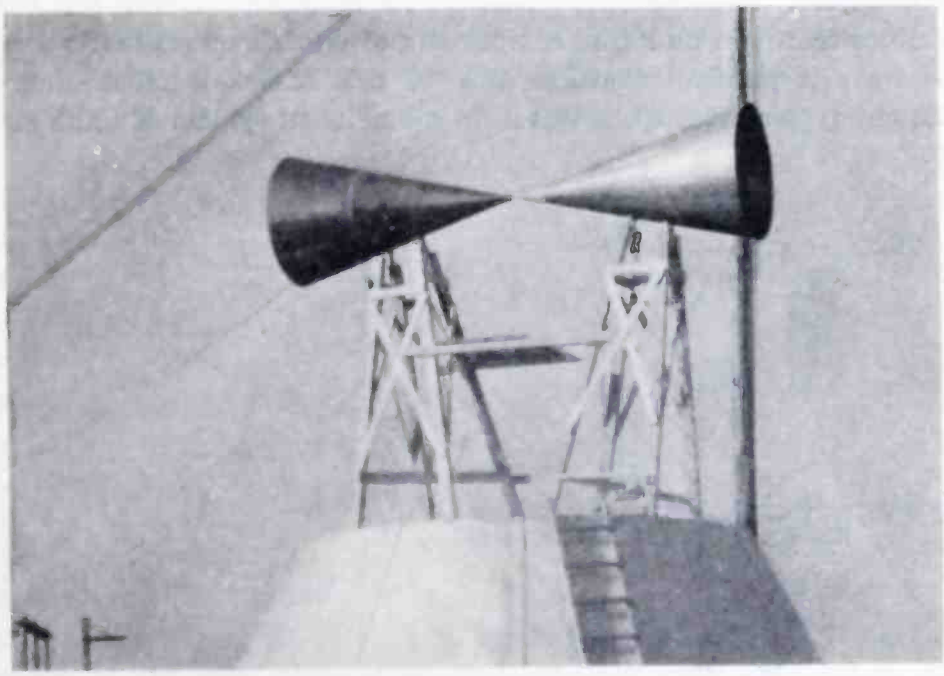


Fig. 22—Experimental double-cone antenna.

magnetic wave is carried by the fields in close proximity to the conductors. The conductors thus serve as excellent guides allowing very little energy to escape into free space.

The purpose of a television antenna system is the exact opposite from that of a transmission line. The conductors forming the antenna must continue to influence the energy which has been guided to their vicinity by the transmission line but, instead of guiding it to a fixed location, they must help the wave to expand into free space without any sudden disturbances in its electromagnetic field configuration. Such sudden disturbances result in reflected-wave energy which, in general, makes itself known in terms of reactance at the junction of the transmission line and the antenna.

Let us consider the electromagnetic wave propagated between two conical conducting surfaces extending indefinitely and arranged in hour-glass relationship. This wave is the simplest possible type of

spherical wave. Figure 26 shows the distribution of electric field at a particular instant of time. These lines simply slide along the conductors as time progresses. The number of lines per inch represent an attempt to show the variation of intensity with radial distance. The variation in thickness is a rough indication of the variation in strength with latitude when the common axis of the two cones is considered as the polar axis for all spherical wave fronts. The lines of magnetic force are latitude lines on any wave front and are not shown in the figure.

Since there are no radial components of electric or magnetic force, this wave is purely transverse and for this reason is much simpler than any type of free space wave. In a consistent system of units such

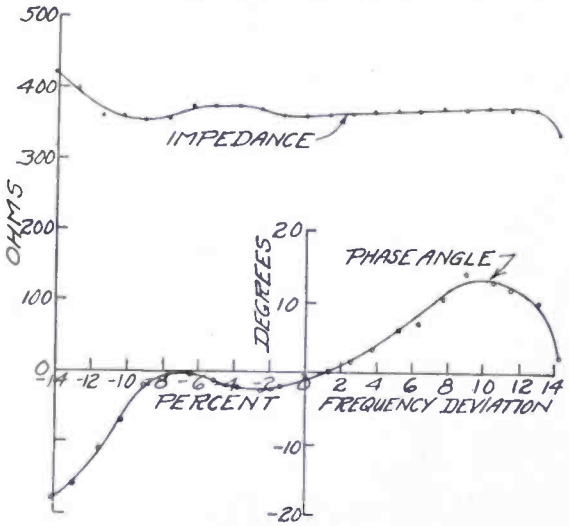


Fig. 23

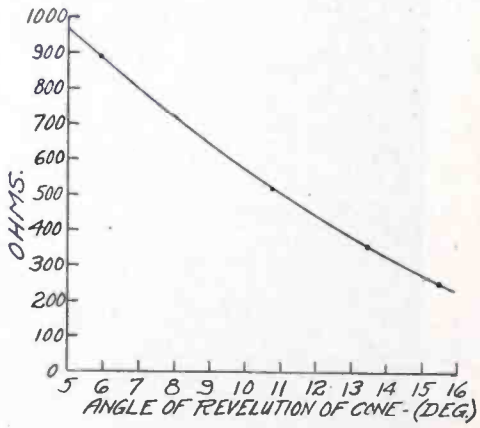


Fig. 24

as the Heaviside-Lorentz system, the electric and magnetic forces are everywhere equal in magnitude. These intensities are inversely proportional to the distance from the apex and inversely proportional to the sine of the co-latitude angle on any spherical wave front. The voltage from one conducting surface to the other in any wave front is a constant independent of distance of travel. The same is true of the total surface current along the conductors. If we define the wave impedance or characteristic impedance in the same manner as for an ordinary transmission line, i.e. as the ratio of voltage to current in a travelling wave, we see that the characteristic impedance of the expanding wave between this system of two cones is a constant, independent of distance. Its value is 120 $\log_e$ (cot $\alpha/2$) where α is the angle of revolution of the cones. This constancy of the wave impedance is a very important property of this type of wave.

The preceding statements summarize the results of a mathematical analysis of spherical wave propagation on conical surfaces. Those

familiar with harmonic analysis may be interested to know that the conical conductors are responsible for a spherical wave derived from the Legendre function of the second kind. In most treatments of wave functions this solution is not discussed, probably because the corresponding type of wave cannot exist without conical surfaces and the investigators were concerned almost entirely with free space waves which must be derived from the Legendre function of the first kind.

We are now ready to consider the wave phenomena in connection with the actual arrangement of two finite cones. As a wave front expands while the wave travels outwardly from the apex, the energy does not remain heavily concentrated near the surfaces of the conductors as in the case of an ordinary transmission line, but is fairly

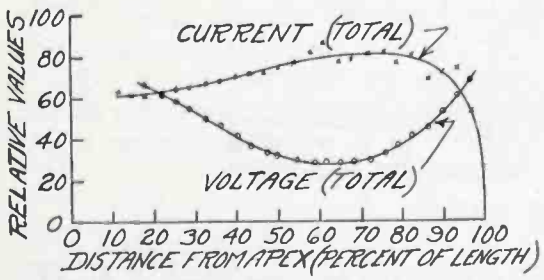


Fig. 25

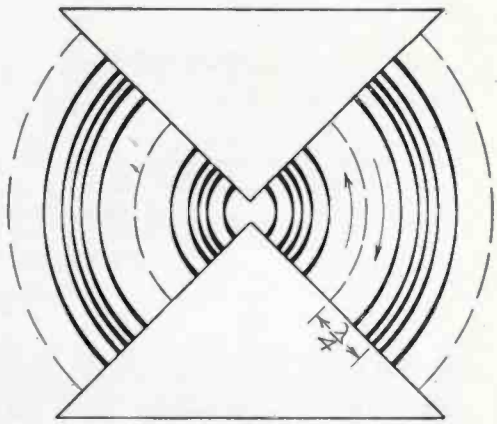


Fig. 26

evenly distributed throughout a wave-front surface. Although the energy immediately adjacent to the conducting surfaces must be completely reflected when the wave front reaches the outer edges of the cones the larger portion of the total energy in the wave front is free to continue to progress into free space. Speaking rather loosely, the wave has by this time expanded sufficiently "to gain a good hold on the surrounding ether." Since the characteristic impedance of this antenna is independent of distance from the apex, the only reflection which can take place on the antenna itself is that which occurs at the rims of the cones. The reflected-wave energy reaching the apices is quite small in comparison with that of the outgoing wave and for this reason the antenna presents a substantially pure resistance to the feeder line over a considerable range of frequency.

Summary and Conclusions

For reception at locations where the signal strength is high an ordinary half-wave dipole of large conductors fed from a convenient type of balanced transmission line should generally prove satisfactory

if the transmission line is loaded, at the receiver, with a pure resistance equal to its characteristic impedance. For remote locations a better system is a folded dipole matched to an efficient transmission line.

For transmitting a horizontally polarized wave, a pair of folded dipoles arranged in turnstile fashion and fed in quarter-phase relation,

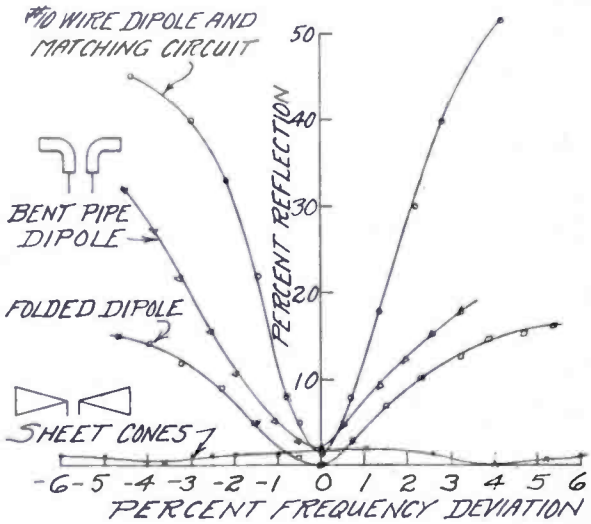


Fig. 27

would seem to satisfy most requirements at the present time. A double-cone antenna, together with an open-wire transmission line, satisfies the most rigid requirements with regard to band width but is not without serious mechanical disadvantages.

For comparison purposes curves of the coefficient of reflection versus frequency as measured on the transmission line feeding several of the various antennas which have been discussed are shown together in Figure 27.

TELEVISION TRANSMITTING ANTENNA FOR EMPIRE STATE BUILDING

By

NILS E. LINDENBLAD

Transmitter Laboratory, R.C.A. Communications, Inc., Rocky Point, N. Y.

Summary—The theory for a new approach to the design of wide-band radiators for vision antennas is set forth and demonstrated in its application both to models and to the television antenna recently installed on top of the Empire State Building in New York City. A band width was obtained which is several times that required to meet television standards in the United States.

The design principles of a sound antenna capable of being located close to the vision antenna without causing mutual interference, are described and demonstrated.

The paper also deals with important transmission line phenomena and circuit problems, particularly as they occur in connection with this antenna.

A short description of mechanical and auxiliary details is included.

INTRODUCTION

HIGH-DEFINITION television has forced the development of circuit features which only recently seemed unattainable. Providing transmitting antennas, of the broadcasting type, having constant characteristics over a frequency band sufficiently wide to accommodate the full band spectrum of high-definition television, becomes increasingly difficult with decreasing carrier frequency.

The recently installed antenna on top of the Empire State Building in New York City represents a successful attempt to remove these difficulties from the frequency regions with which television is concerned. This antenna, Figure 1, consists of two separate, independent radiator systems for vision and sound transmission supported by a common column. The radiators at the middle of the column constitute the vision antenna and the top radiators constitute the sound antenna.

In the vision antenna is represented the application of a new approach to wide-band antenna development by which the radiating elements themselves, uncombined and unaided by any form of compensating circuit, present input impedances which are resistive and constant over a frequency band considerably wider than that required by present television standards.

Rational carrier-frequency allocation[1] and certain receiver requirements call for closely adjacent carrier frequencies for vision and sound

Reprinted from *RCA Review*, April, 1939.

transmission. In spite of this close frequency spacing, it was found possible to design the vision and sound radiator systems so that negligible coupling effect was obtained even when the two systems are located no further apart from each other than the distance of a half wave, Figure 1. The arrangement represents a clean and efficient method of preventing mutual interference between vision and sound transmitters.

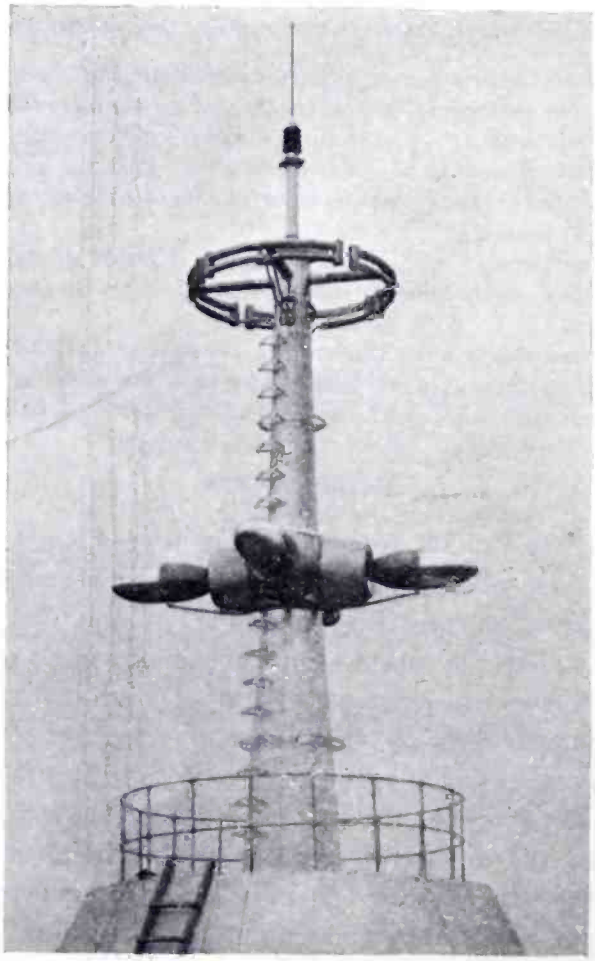


Fig. 1—Television antenna for Empire State Building, as erected for test at Rocky Point Laboratory.

While the radiation from each of the radiating systems is horizontally polarized, the vertical directivity of each approximates that of a vertical dipole. Propagation tests have fairly well established the superiority of horizontal polarization[2] which for this reason has been recommended by RMA as standard.

General Principles—Vision Antenna

It is a well-known principle that the total impedance of circuits having inductance and capacitance, becomes independent of frequency

if these components are associated with resistive loads of values equal to the square root of the inductance-capacity quotient, Figure 2. An equivalent version of the circuit in Figure 2b is shown in Figure 3a. By composing a radiator from coaxial transmission-line sections as shown in Figure 3b, it is possible to make it consist of an inductive component *A* and a capacitive component *B* of such relative lengths that they carry equal portions of the radiation load. By properly proportioning the transverse dimensions of radiator components *A* and *B*, the square root of their inductance-capacitance quotient may be made equal to the radiation resistance of each component. The characteristics of the circuit in Figure 3a will thus be simulated. However,

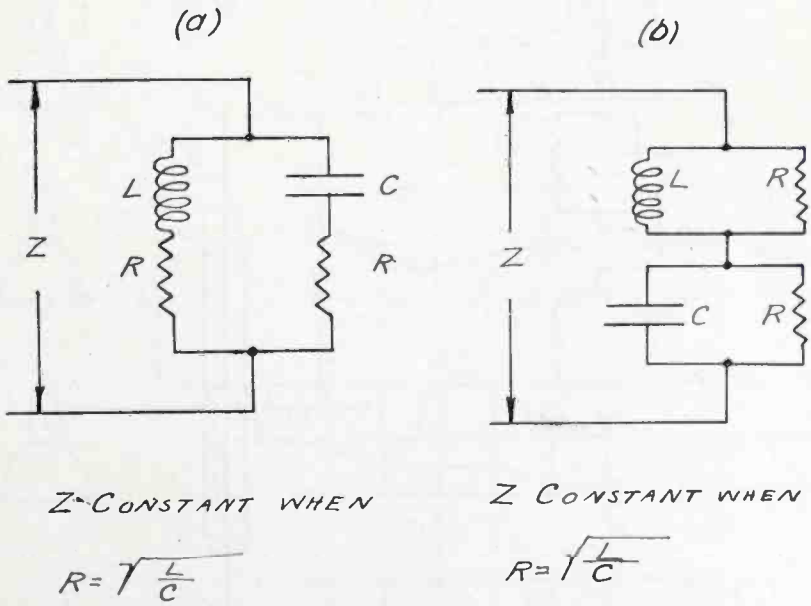


Fig. 2—Circuits for constant impedance at variable frequency.

due to the fact that the electric and magnetic high-frequency fields of the radiator in Figure 3b are distributed and thus dependent upon frequency, it is evident that inductance, capacitance and radiation resistance all become variables with frequency. The precise requirement for the ideal constant-impedance circuit—as in Figure 3a which calls for fixed values of inductance, capacitance, and resistance—is, therefore, not fulfilled. An input impedance independent of frequency could, therefore, not be expected. It was, however, considered reasonable to expect appreciable benefit when the proportions of a composite radiator are such that the requirements for the constant-impedance circuit are fulfilled at mid-spectrum frequency.

Proportioning of the inductive and capacitive radiator components required much experimenting. Steps in this experimental evolution of proportions toward optimum dimensions may be followed by the samples shown in Figure 4. The effects of shape as well as of dimen-

sions were studied. A generally elliptic shape of all the radiator surfaces, Figures 4 and 5, seemed to permit the fullest utilization of the constant-impedance principle. Gradual expansion of the transmission line into the radiator also appeared as the only satisfactory method of connection. It is known that expansion of a radiator away from its feed point is an aid to impedance constancy at variable frequency. This coincidence is no doubt an aid to the characteristics of the composite radiator.

As may be concluded from the band-width curves shown in Figure 4, the band width increases rapidly with transverse radiator dimensions

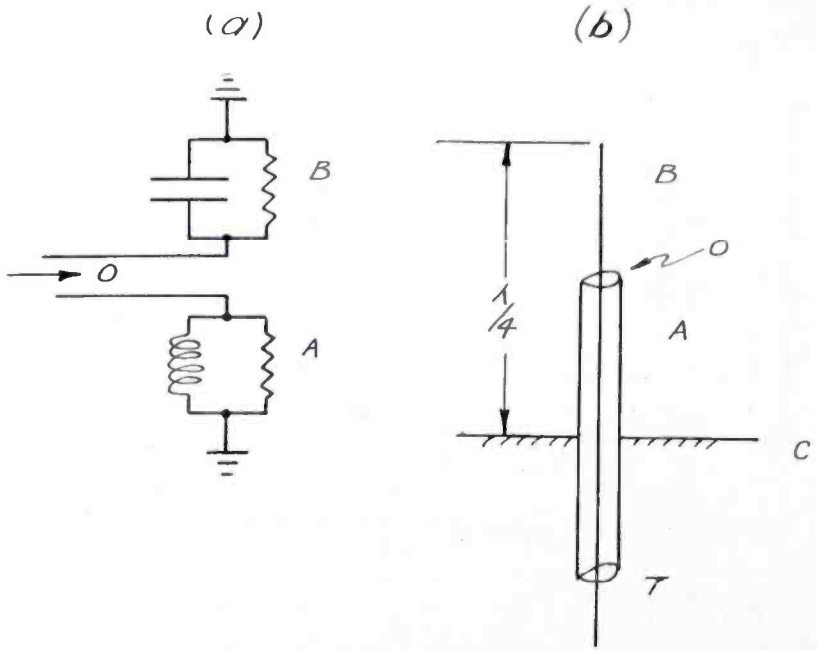


Fig. 3—Application of constant-impedance circuit to composite radiator.

as proper proportions are approached. When these dimensions are exceeded the band width again diminishes. This fact seems to indicate that the theory set forth in connection with Figure 3 represents the basic principle upon which the performance of the composite radiator depends.

The elliptically shaped extension of the center line conductor, which represents the capacitive component *B* of the composite radiator Figure 5, will hereafter be referred to as the "ellipsoid". The expanding portion of the outer conductor of the transmission line, which has elliptic curvature, will be referred to as the "throat". The surface which forms the continuation of the throat into the radiator, constitutes the inductive component *A* of the composite radiator. The curving of this surface is also elliptic and this radiator component will be referred to as the "collar".

In order that the protruding portion of the ellipsoid and the collar might be equally loaded by radiation, their relative lengths, Figure 5a, should be as 7 to 5. The input impedance, governed by this ratio, is in the order of 110 ohms. The best ratio between major and minor axis of the ellipsoid was found to be 15 to 6. The best ratio between

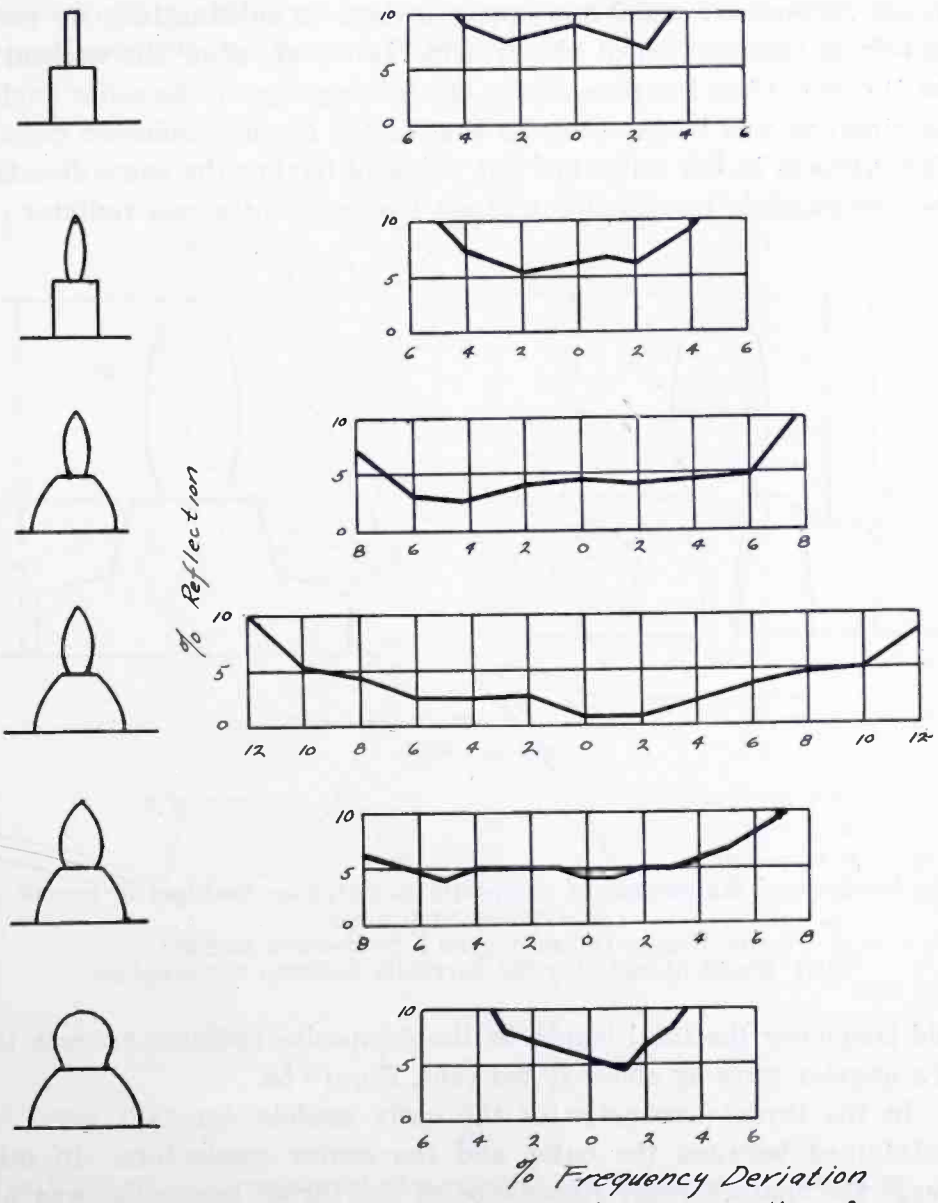


Fig. 4—Experimental steps in determining optimum proportions for maximum band width of composite radiator.

the mean outside diameter of the collar and the minor axis of the ellipsoid appeared to be 3 to 2.

The band width, within which reflection of waves from the radiator back over the transmission line was less than 5 per cent, with a composite radiator Figure 5a, vertically mounted from a horizontal metallic

surface, was 20 per cent, Figure 4. This result was obtained with small scale models designed for a mean frequency of 150 Mc.

As can be understood from Figures 3 and 5a, the transmission line is in effect connected in series with the two components of the composite radiator. Since the currents on the center conductor and the throat surface are equal and opposite there is substantially no radiation from this portion of the system. However, after the current on the throat surface has passed over the joining edge to the collar surface its direction will be the same as that of the center conductor current. The currents in the collar and the ellipsoid having the same direction, the two radiator components will act like one continuous radiator. At

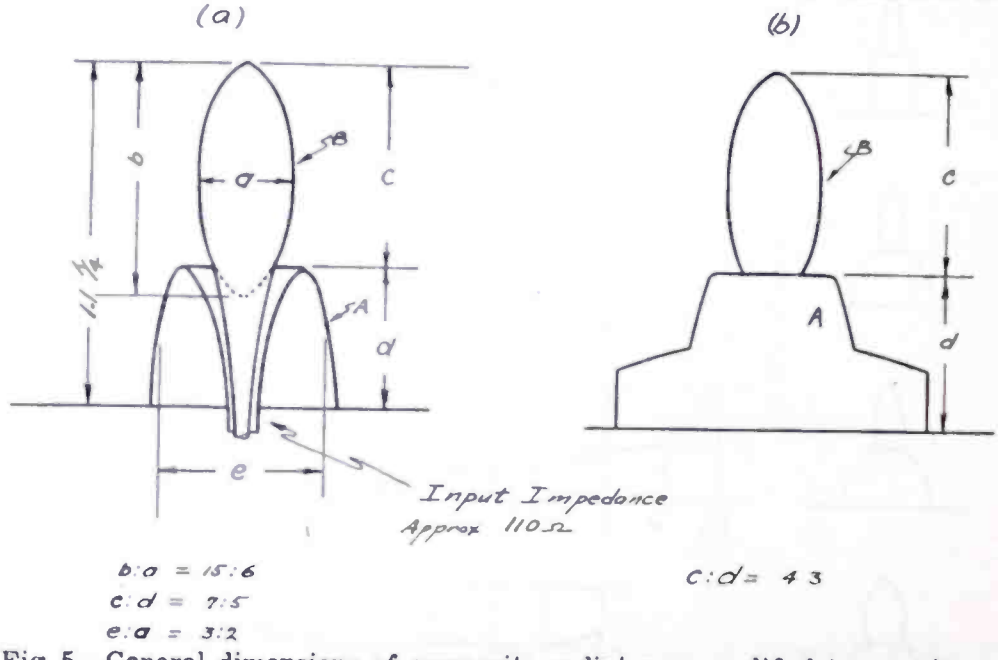


Fig. 5—General dimensions of composite radiator as modified to permit use of supporting bracket.
(a) Single radiator over a conductive surface.
(b) Model of radiator for turnstile antenna combination.

mid-frequency the total length of the composite radiator exceeds that of a quarter wave by about 10 per cent, Figure 5a.

In the throat expansion of the early models, constant ratio was maintained between the outer and the center conductors. In other words the characteristic impedance of the throat expansion was held constant. For practical purposes, however, such as mechanical strength and to obtain immunity against lightning, some form of metallic bracket, Figure 7, between ellipsoid and collar seemed indispensable. The capacitance required to balance the inductive shunt introduced by the bracket was most effectively obtained by gradual decrease of the characteristic impedance of the throat section in the direction of expansion.

In the development of the bracket only experimental methods were practical. There were many variables such as self-inductance, mutual inductance, size, length, and position which all had to be coordinated to secure the maximum mechanical strength obtainable without sacrifice of band width. By not compromising on the band width, the dimensions of the bracket were limited, but fortunately these dimensions provided adequate mechanical safety factor. The bracket as finally

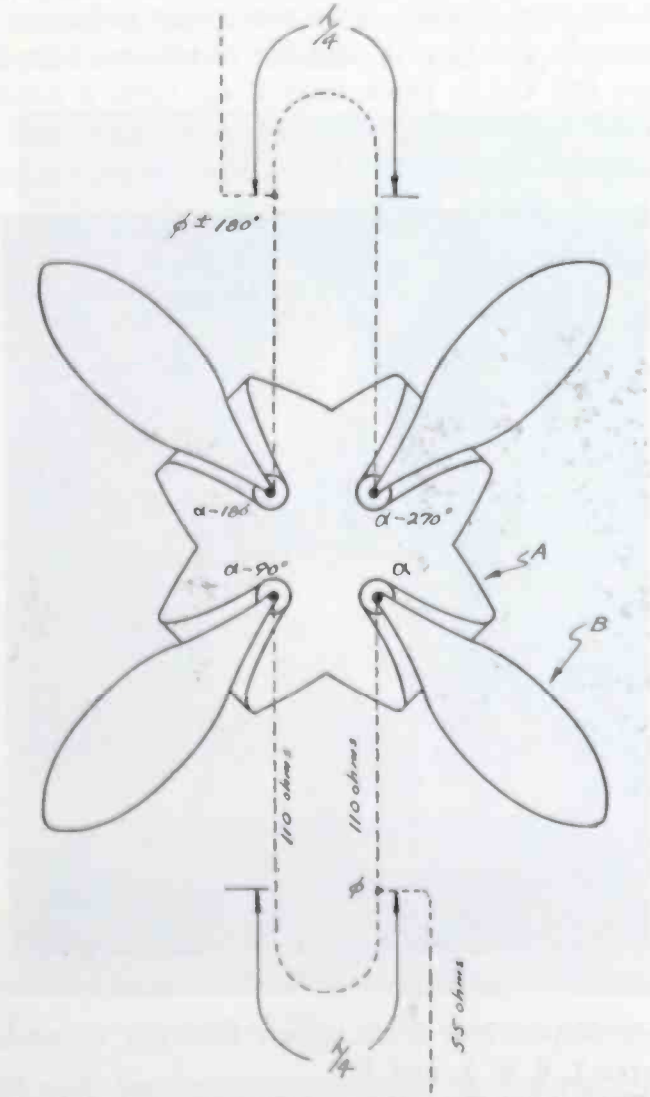


Fig. 6—General arrangement of composite radiator turnstile antenna.

developed for the full scale antenna is shown in Figures 1, 12, 15, and 17.

For broadcasting horizontally uniform, vertically polarized radiation, a single vertical composite unit as shown in Figure 5a is sufficient. However, since horizontal polarization was required, several horizontally located composite units had to be combined to produce horizontally uniform radiation, or the equivalent. A simple method of

combining, which was adopted, is to arrange four composite radiators into a "turnstile" antenna. A turnstile antenna, in its simplest form, consists of two dipoles crossed at right angles carrying high-frequency currents in phase quadrature.[3] Although the radiation in the plane of each dipole is not uniform, the combination of the two, by virtue of their phase and geometric quadrature, results in a rotating electromagnetic field which produces the equivalent of uniform radiation in the plane of the two dipoles. Two composite radiators of the type shown in Figure 5a pointing in opposite directions, joined at the base of their collars and fed in phase opposition, form a composite dipole. Two such dipoles may then be combined in the horizontal plane into a composite turnstile antenna. The four collars in such a turnstile form

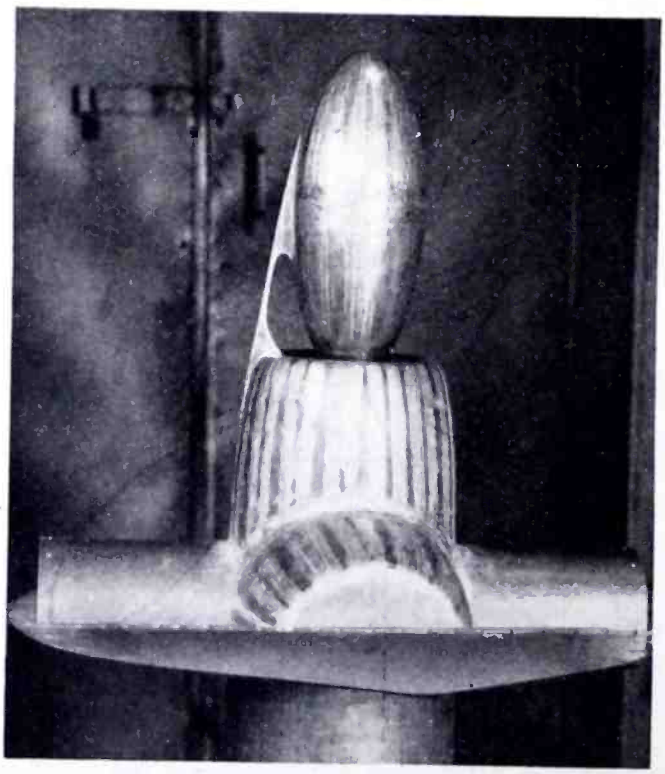


Fig. 7—Model of radiator for 100-150 Mc turnstile antenna combination.

a central cross-shaped hub from which the four ellipsoids protrude radially, Figures 1, 6, 8, 9, and 12.

Before building a complete 100-150 Mc turnstile model, Figures 8 and 9, a half section of such a model as shown in Figure 7 was made in order to expedite determination of the new relative collar length resulting from the alterations in the collar characteristic when four collars are combined into a hub. The new length ratio between the protruding portion of the ellipsoid and the collar became 4 to 3 as shown in Figure 5b.

A complete 100-150 Mc turnstile model was then built and tested,

Figures 8 and 9. From the foregoing discussion it is clear that the four composite radiators of this turnstile were so connected that the currents they carried were equal and in progressive phase quadrature, Figure 6. That means that the currents in any adjacent pair of radiators were in phase quadrature while the currents in any opposite pair were in phase opposition.

The input resistance of each radiator in the turnstile was 110 ohms. The feeders to the radiators were, therefore, designed to have a characteristic impedance of 110 ohms. In this way the characteristic

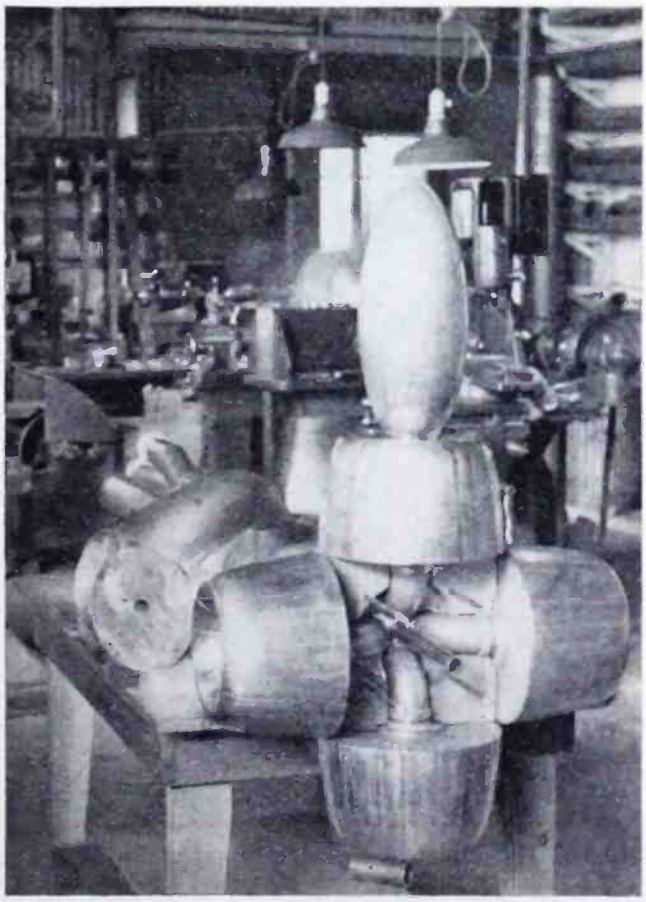


Fig. 8—Shop view of 100-150 Mc turnstile antenna.

impedance of each feeder was "matched" by the radiator, and high-frequency waves arriving over the feeder were radiated into space substantially without reflection back toward the transmitter.

Since the waves on a reflection-free line travel only in one direction, the phase of the high-frequency current it carries becomes a linear function of the distance along the line from any point of reference. The phase between the currents at points a quarter wave apart thus becomes 90 degrees. Progressive phase quadrature between the currents of the four radiators might then have been obtained by

providing four separate matched feeders progressively differing in length by a quarter wave.

Since, in a progressive phase-quadrature system, the currents in opposite radiators are in phase opposition, Figure 6, it was more practical and economical to provide two equal 55-ohm main feeders coupled in phase opposition to the transmitter. Each was split into two 110-ohm branches near the antenna to feed a pair of adjacent radiators. These 110-ohm branches were made to differ in length by a quarter wave. In this way, the feeder branches of two opposite radiators could be

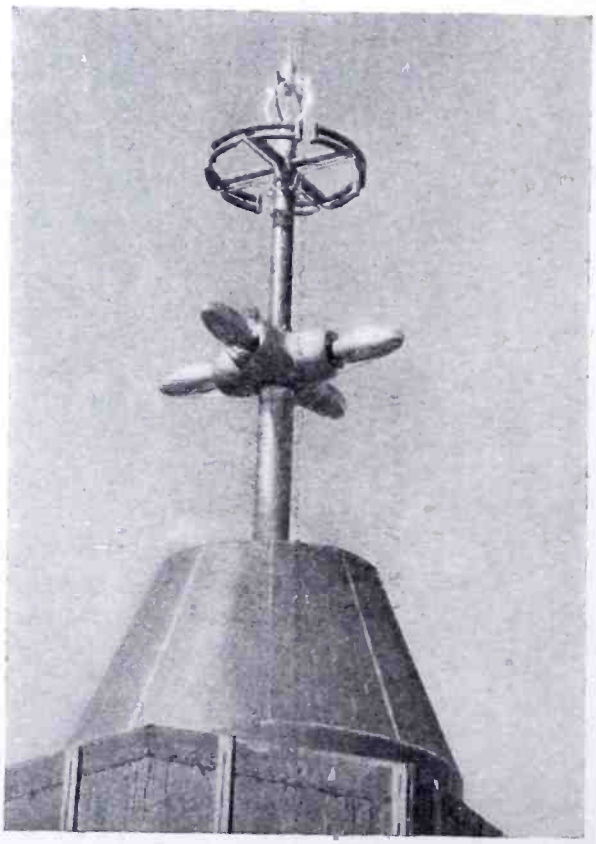


Fig. 9—Complete 100-150 Mc model of combined vision and sound antennas.

made equal in length if they were connected to main feeders of opposite phase. Progressive phase quadrature was thus obtained by using only two different lengths of branch feeders.

When two equal high-frequency loads, such as an adjacent pair of radiators in a turnstile, are supplied with power from a common branch junction through sections of transmission lines, a quarter wave different in length, the input impedance to such a system, as a whole, is more constant over a wide frequency band than the input impedance into each branch line.

If equal reflections are produced by the two radiators, the input impedances of the branch lines, at their junction, become reciprocals

with respect to the characteristic impedance of the feeders. These reciprocal impedances are brought into parallel combination at the branch junction. If the difference between the reciprocal values is moderate, the resultant parallel impedance is very little different from the characteristic impedance of the main feeder, which has been designed to match the combined characteristic impedance of the branch feeders. That is, if the reflections on the feeder branches are moderate, these reflections will nearly compensate each other at the junction. When the branches differ in length by an odd multiple of a quarter

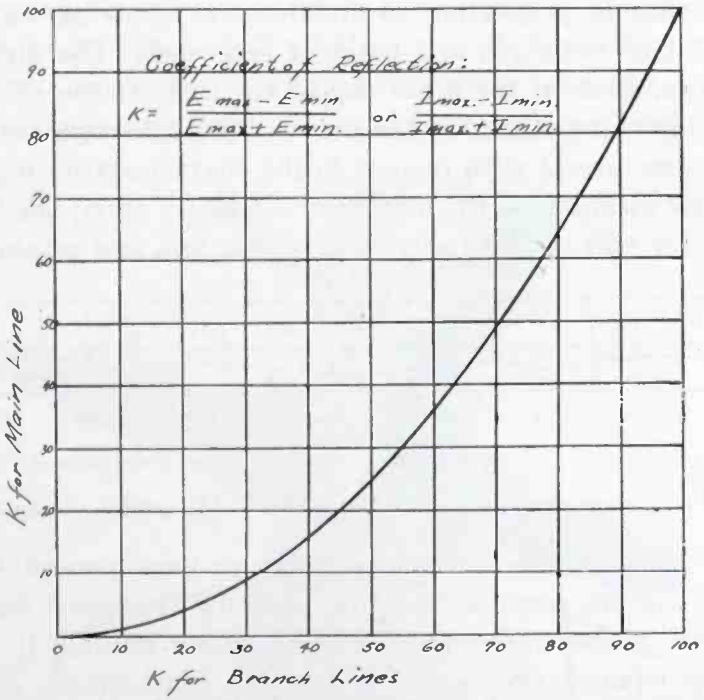


Fig. 10—Curve indicating relation between coefficients of reflection of main feeder and parallel quarter wave, differential length, branch feeders.

wave and the reflections are equal, the coefficient of reflection on the main line becomes equal to the square of the coefficient of reflection on each branch. This relationship between coefficient of reflection on branches and main feeder is shown by the curve in Figure 10. It is evident from the curve that the ratio of compensation is large for low reflections and decreases toward unity as the reflection on the branches approaches 100 per cent. Thus, while the compensation is tenfold at 10 per cent reflection on the branches, there is no compensation when the reflection is 100 per cent.

It should be clear that the compensation becomes less and less effective as departure is made from midfrequency since the difference in length between the branches is then no longer a quarter wave. This is an important consideration in the case of circuits for wide frequency bands as required for television. The composite radiator possesses the

important virtue that its input impedance departs from non-reflecting values very slowly with departure from midfrequency. This type of radiator is, therefore, capable of taking fullest advantage of the compensation intrinsic in phase-quadrature combinations.

It may be realized that antenna loads which closely match the main lines throughout the operating frequency band, become an increasing necessity as the distance between transmitter and antenna becomes large in terms of wave length. A standing wave, resulting from a certain mismatching, will shift in its location in proportion to frequency and also in proportion to distance, in terms of wave length, between point of reflection and point of reference. The biggest shift will, therefore, occur at the input end of the line. If the shift is large enough the input impedance of the line will pass through both a maximum and its reciprocal with respect to the characteristic impedance of the line. This means that for a certain frequency shift, the longer the line, the faster will be the recurrence of maxima and minima. For a

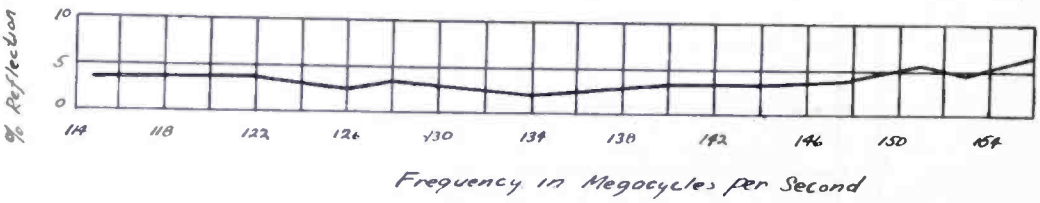


Fig. 11—Frequency-response curve of 100-150 Mc model vision antenna.

long line it is, therefore, entirely possible to have several impedance maxima and minima occur within the operating frequency band. Since the impedance passes through both maxima and minima it is evident that the input-impedance variation assumes considerable proportions even if the mismatch at the load end is small. If the line is sufficiently short, such large variations as from a maximum to a minimum will not occur within the operating band. The impedance variation will, therefore, be smaller.

It is of theoretical interest to note that impedance variations at the transmitter end of a line will not necessarily cause variations in the power output of a transmitter. If the internal impedance of the transmitter matches the characteristic impedance of the line, the power output at a fixed reflection will not vary with frequency. In practice, however, variable impedance is a serious factor.

The 100-150 Mc model antenna, Figures 8 and 9, was found to possess a band width well in excess of 30 per cent, Figure 11, which was over 50 per cent more than the band width of a single composite radiator unit. The "band width" is defined as the frequency spread within which reflection of waves back on the main feeders is less than 5 per cent. Coefficient of reflection is the quotient of the difference

and the sum of maximum and minimum voltage or current on a line, Figure 10. Percentage reflection is obtained by multiplying this quotient by 100.

The important dimensions, which had been determined with the model, were then reproduced to scale in a full-size turnstile antenna, having a midfrequency of about 45 Mc, which was constructed for installation on top of the Empire State Building, Figure 12. When the antenna was tested it was found to have the truly remarkable band width of over 60 per cent, Figure 13. This is six times the frequency spread of a single television channel for vision and sound. The band is from 6 to 10 times that obtainable with heretofore conventional designs combined with complicated correction networks.

The discrepancy in band width between model and full-size antenna

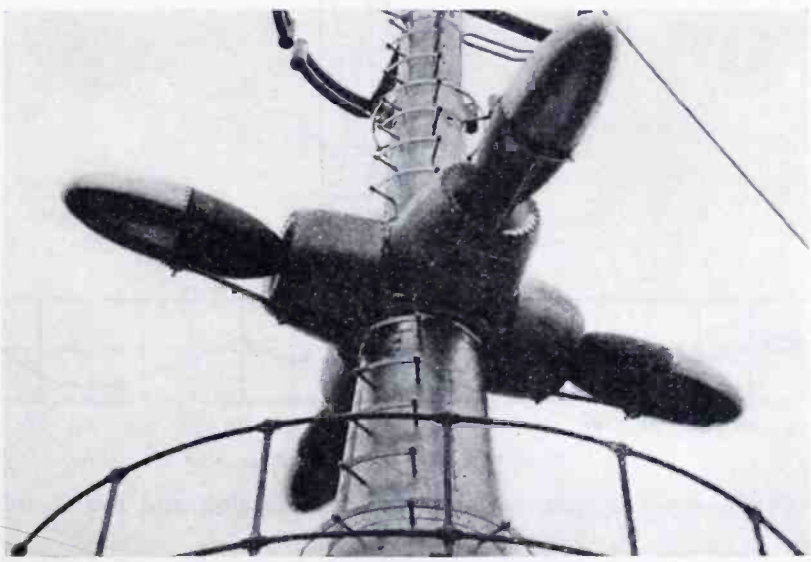


Fig. 12—Empire State Building vision turnstile radiator system.

must be attributed to greater accuracy in construction and perhaps also to better testing equipment. Up to the present, no opportunity has been available to study the band width of such large, single composite units as used in the antenna for the Empire State Building installation. It is possible that the 20 per cent band width obtained with 150-Mc models of single composite radiator units may be exceeded under the more accurate conditions possible with larger models.

When the full-scale antenna was set up and tested at Rocky Point, Figure 1, it was found possible to provide a metal rail around the testing dome on which it was mounted, by allowing the reflection to rise to a 6 per cent peak at 54 Mc. This slight marring of the band characteristic was, of course, of no consequence since this peak lies outside the operating band. After using an even more substantial rail

and adding a non-symmetrically located ladder step near the radiators when installing the antenna at the Empire State Building, it was found that this peak shifted to 53 Mc and rose to a value of about 8 per cent.

Although a single turnstile antenna has less vertical directivity than a vertical dipole the difference was made up for by locating it a half wave above the roof of the dome.

It may be of interest to note that in the transmission lines used in connection with the full-scale antenna, the suspension insulators of the center conductor were located in pairs with a quarter-wave spacing between the insulators of a pair. In this way, use was made of the phase-quadrature principle for the purpose of reducing insulator effects.

Likewise, it may be interesting to note that the 55-ohm main feeders, which are several hundred feet long, were so made that in them

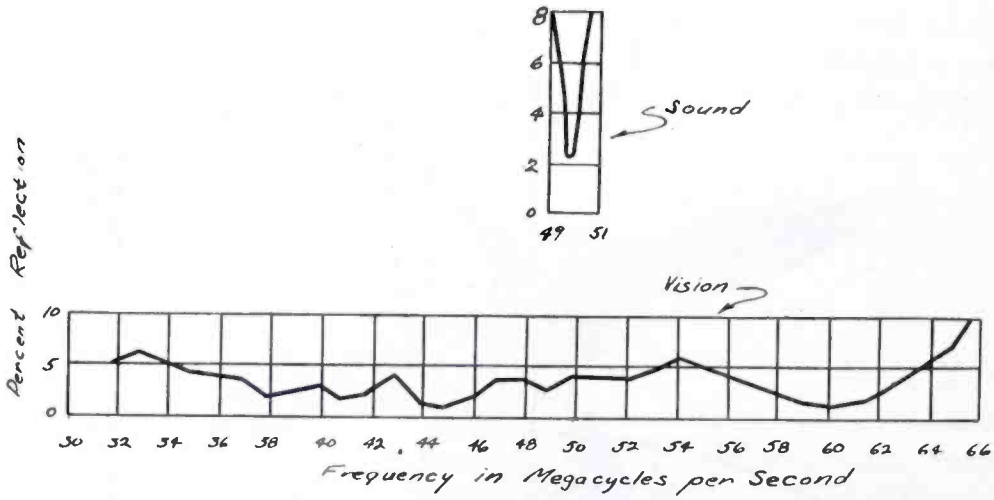


Fig. 13—Frequency-response curve of full-scale vision and sound antennas.

is represented an economic compromise between lines having a characteristic impedance of 77 ohms for securing maximum power-transfer efficiency and lines having a characteristic impedance of 30 ohms for securing maximum power-carrying ability for any given diameter of outer conductor.[4] The diameters of the outer and inner conductors of the main feeders were made 2.5 inches and 1 inch respectively.

Sound Antenna

In choosing the general design principles for an antenna for sound transmission to be mounted on the same structure as the vision antenna, minimum coupling between antennas is found to be the chief consideration.

Low coupling may be obtained in two different ways. Since the vision antenna produces a rotating field, the sound antenna could be built on the same principle, but be connected for phase rotation in the

opposite direction. However, it may be readily understood that such a system would provide minimum coupling for only one frequency, at which the phase quadrature in both the sound and the vision antenna would be perfect. Since, in addition, it is desirable that each antenna have perfect phase quadrature at midfrequency, the condition cannot

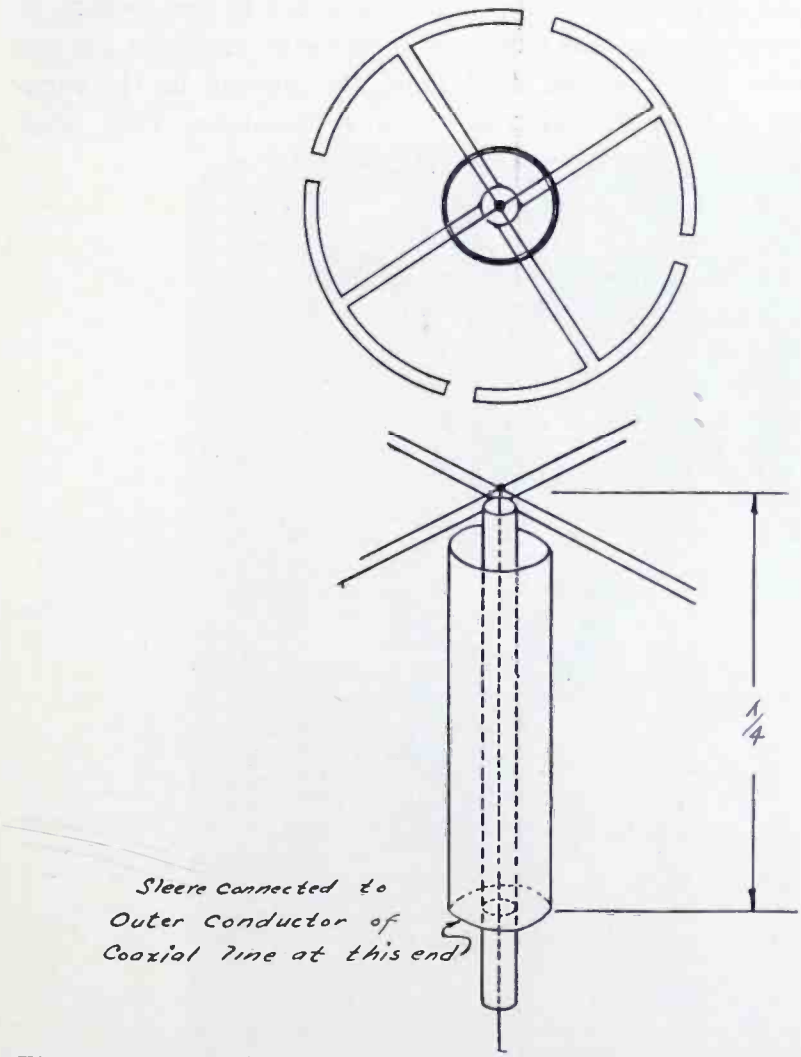


Fig. 14—Schematic diagram of sound antenna connections including line balance converter.

readily be fulfilled. Fortunately other means were available by which it was possible to furnish a design which provided low coupling for all operating frequencies.

A circular loop antenna of uniform characteristics and distribution, if located in a plane parallel to the horizontal plane of a turnstile antenna and on a common, central, vertical axis, would have no coupling to the vision antenna at any frequency. The diametrically located radiators of the vision antenna could not induce voltages giving rise to circulating currents in the loop and vice versa.

A simple way of providing the approximate equivalent of a loop antenna is to locate a number of dipoles in a circle. Cancelling effect between the pick-up of the segments or dipoles is in this case obtained at the point of connection to the common transmission line feeding the segments in parallel. This may best be understood by referring to Figure 14. The common feed line is so connected to the feeders of the individual radiators that the currents in opposite radiators become opposite. Currents, which on the other hand are induced by the turnstile antenna, are of same direction in opposite radiators. They must

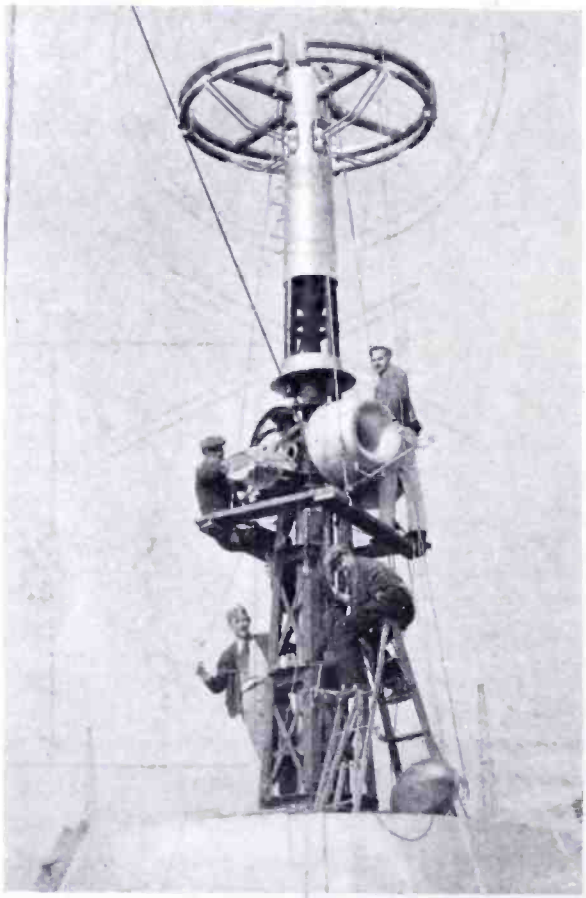


Fig. 15—Television antenna for Empire State Building during erection for test at Rocky Point Laboratory.

therefore balance each other without entering the common feed line. In the present case, the antenna was made up of four folded dipoles bent into circular segments, Figures 1, 9, 15, and 19. Among the dimensions of folded radiators which, at a given frequency, result in resistive input impedance, only the two smallest dimensions are of any interest in this case. At the larger of these dimensions the distance between the folding points of the radiator is approximately a half wave. The folding points coincide with maximum potential and the currents in the parallel conductors flow in the same direction. The

distance between the folding points at the smaller dimension is only about a quarter of a wave. The input terminals are at maximum potential and the currents in the parallel conductors flow in opposite directions. A ring antenna of this latter type need only be about half the size of that required by the first type. The smaller dimensions further reduce the possibility of undesirable mutual effects between the sound and the vision antennas and reduce the mechanical problems. For these reasons, the small-type folded dipole was chosen.

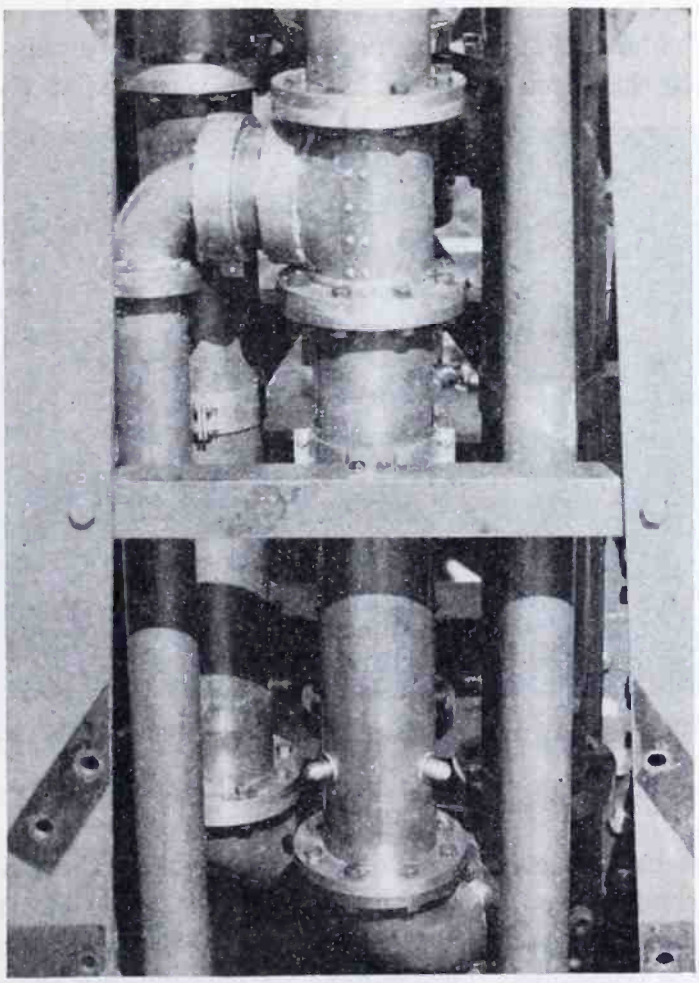


Fig. 16—Main line, junction and branch feeders inside the central supporting structure of the full-scale antenna. Note the quarter-wave branch feeder loops.

In the small-type folded dipole, the current folds with the conductor and tapers toward the input terminals where it becomes a minimum. On account of the voltage distribution thus obtained in this radiator, the capacitance between the parallel branches provides a path for circulating currents which lower the power factor. The resulting increase in frequency selectivity in the sound antenna further decreases the possibility of energy transfer from the vision antenna into the sound transmitter.

The input impedance of a single folded dipole, operating as described, would normally provide very high input impedance. However, when combining four such radiators into a ring-shaped antenna, by properly spacing adjacent folding points, it was found possible to influence the characteristics of the radiator so that the input impedance of each radiator in the combination was reduced to 220 ohms. This impedance was desirable since by parallel connection an impedance of 55 ohms could be obtained, without impedance transformation, which would equal the characteristic impedance of a coaxial line similar to the main feeders of the vision antenna. The four dipoles were then connected by balanced open feeder pairs, of 220-ohm characteristic

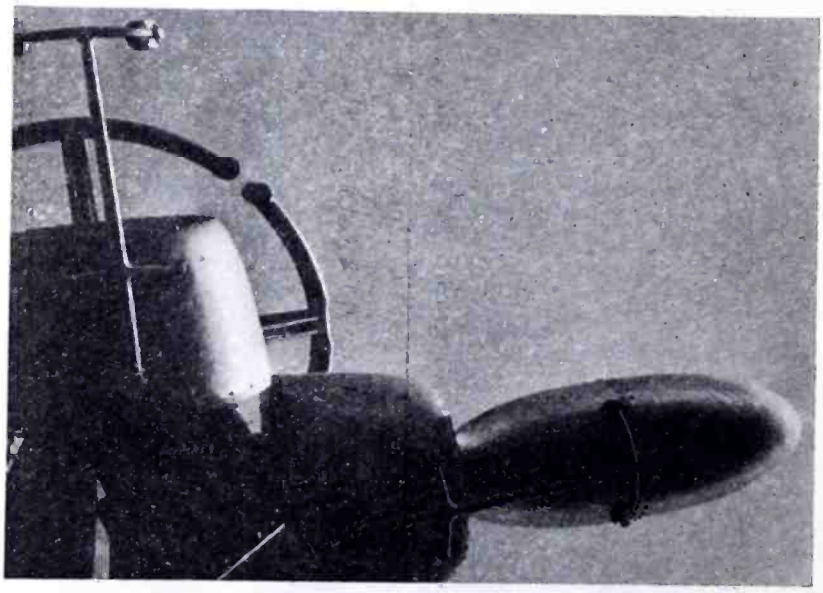


Fig. 17—Position of supporting bracket with respect to collar and ellipsoid.

impedance, through a balance converter to a 55-ohm coaxial line, Figure 14.

The line balance converter, Figure 14, involves the principle that the outer conductor of a coaxial line must be made electrically free from its surroundings in order not to destroy the balance of a system to which it is connected. This is done by surrounding the end section of the outer conductor with a concentric sleeve and by connecting the two together at a point a quarter wave from the end of the line, Figure 14. The impedance between the end of the outer conductor, of the coaxial line, and the sleeve which is bonded to the surrounding supporting structure has by this procedure been made very high. The current flowing on the inside of the outer conductor of the coaxial line can, therefore, become the continuity of the current in the connected load circuit. There is no longer any shunting effect to divert some of this current from its desired path.

This line balance converter, which was developed some time previous to the design of this antenna has found many varied uses and has made possible a number of coaxial-line combinations not otherwise possible.

Models of the sound antenna designed for 150 Mc were made and studied. The final model was located a half wave above the 100-150 Mc vision antenna model and the two antennas were supported by a common column, Figure 9.

The full-scale sound antenna, similarly mounted with respect to the full-scale vision antenna is shown in Figure 1. Its frequency response is shown in Figure 13. The power transfer from one antenna

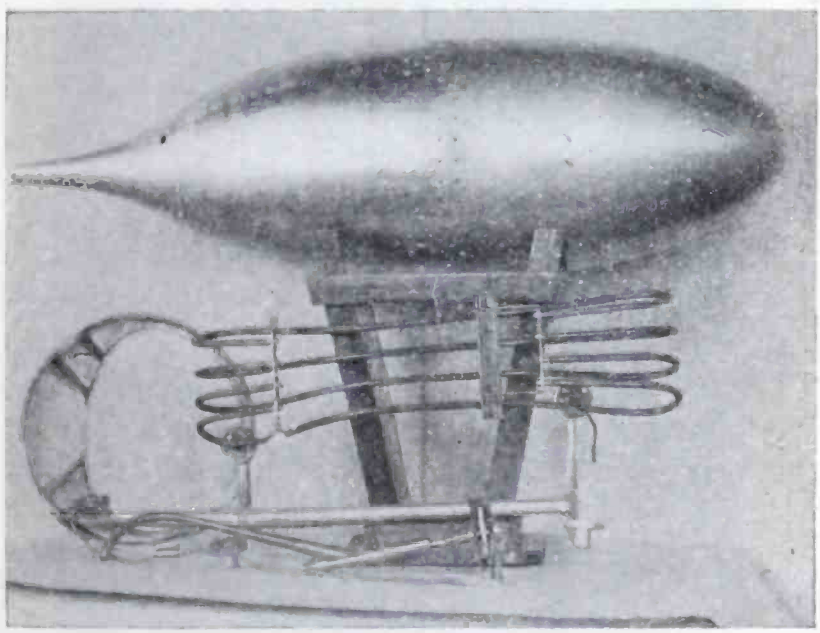


Fig. 18—Ellipsoid and associated parts.

to the other was found to be less than one part in a million of the input power.

Mechanical Design, Ice Removal, Etc.

The mechanical design of the antenna provides for a safety factor of 5 at a wind load of 40 lbs. per square foot, which corresponds to a true wind velocity of about 130 miles per hour. The weight of the total structure amounts to approximately 6,000 lbs.

The central supporting structure consists of a square lattice tower, Figure 15, which contains the necessary feeders, quarter-wave loops, Figures 6 and 16, and the balance converter, Figure 14. This tower, which is tapered, is covered with a circularly bent sheet metal skin, Figures 1 and 15. The four collars of the vision antenna at the middle of the tower have internal steel construction which is bonded to the

tower itself, Figures 15 and 20. Figures 1 and 15 show how the collar and throat surfaces, which are made of spun sheet metal, are fitted over this frame. Attached to the inside framework of the collars are the supporting brackets for the ellipsoids. These brackets, which are shown in Figures 1, 15, and 17, are made of thick-walled Shelby steel tubing. The components of the brackets are welded together. The brackets are slightly tapered and have a bore through which runs the connection to the heating units located within the ellipsoids for ice removal. The ellipsoids are made in two halves joined together at the

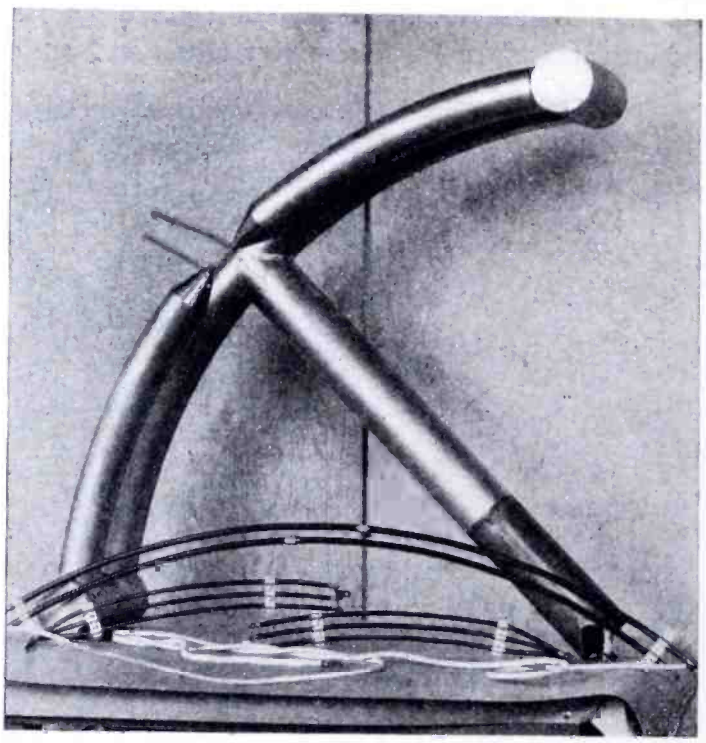


Fig. 19—Folded dipole unit of sound antenna with its associated parts.

middle by an aluminum ring, Figure 18, which is also bolted to the two ends of the bracket bridle, Figure 17.

The anchor-shaped or circular segments of the sound antenna, Figures 1, 15, and 19, are also made of steel tubing with their various components welded into one unit. The center T-member of each radiator slides into the supporting cross which is integral with the top of the tower, Figures 1 and 15. The sound antenna radiators are made up of different sizes of tubing. The different sizes are 3, 4, and 5 inches, Figure 19.

Calrod heating units, bent into convenient shapes, are installed in the radiator sections of the sound antenna, Figure 19, as well as in the ellipsoids, Figure 18, and the collar members of the vision antenna, Figure 20. The heaters of each radiator section are connected by individual leads to a distributing panel in the building dome under the

antenna. Each section of the sound antenna, each ellipsoid, and each corresponding collar require normally about 1.5 kw of heating power for ice removal. Since there are four of each of these elements, each

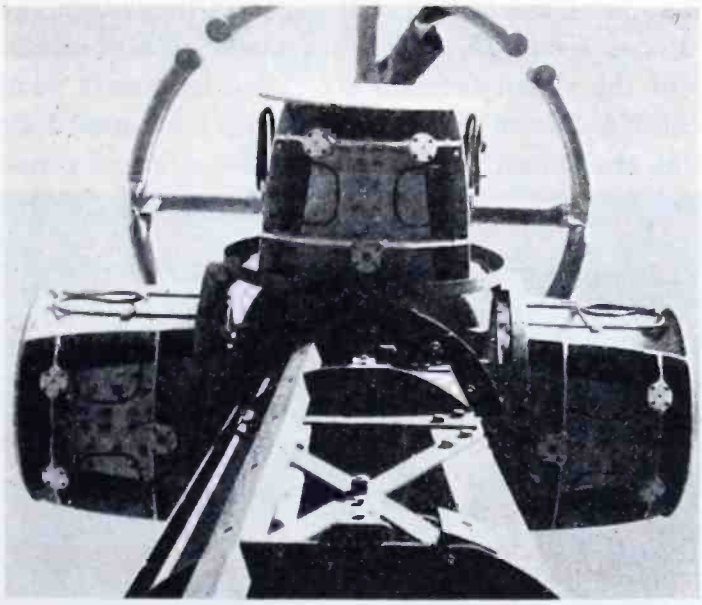


Fig. 20—Internal structure of collars. Note the Calrod heating units.

group of four will normally require 6 kw. There are three such groups which are then connected in three phase. Voltage tapping is arranged

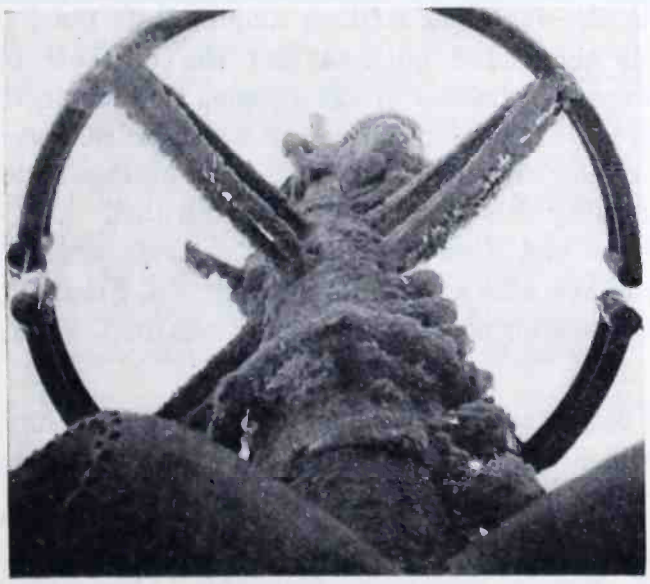


Fig. 21—Ice formation on antenna at Empire State Building. Note the effect from the application of heat to the radiators.

to suit conditions so that the total three-phase power may be varied up to a maximum value of 27 kw. The photograph, Figure 21, shows ice conditions recently encountered. Application of heat keeps the radiators clear of ice and the electrical characteristics of the antenna remain unchanged.

The insulation throughout the transmission-line system is quartz. Three radially placed quartz rods form a suspension point for the center conductor. An exception to the use of quartz is the pressure-sealing insulators of the lines and the lead-in insulators for the sound antenna, Figures 1 and 15, which are made of high-grade porcelain. The throats of the vision antenna are closed by quartz windows which in turn are shielded from water creepage by mica guard rings located further out on the throat surface and on the ellipsoid connector.

Fig. 22—Lightning striking Empire State Building.

Except for the sound antenna and the brackets of the vision antenna, all exposed surfaces of the antenna and supporting structure are chromium-plated copper. There are no electrical reasons prohibiting the use of non-corrosive metals such as stainless steel when and if the cost of using this material can be tolerated.

Wind-velocity and wind-direction instruments are mounted at the top of the structure, also a lightning pick-up rod, Figure 1, to facilitate the lightning research carried on at this location[5], Figure 22.

Acknowledgment

The work described in this paper would have been impossible without the wholehearted cooperative spirit, power of imagination and skill of the entire personnel of the laboratory and shop.

Bibliography

[1] Albert F. Murray, "Television Standards," *Communications*, December, 1938.

[2] R. W. George, "A Study of U-H-F Wide-Band Propagation Characteristics," *Proc. I.R.E.*, January, 1939, Vol. 27, No. 1. (See page 90).

[3] George H. Brown, "The Turnstile Antenna," *Electronics*, April, 1936. (See page 41).

[4] Clarence W. Hansell and Philip S. Carter, "Frequency Control by Low Power-Factor Line Circuits," *Proc. I.R.E.*, April, 1936, Vol. 24, No. 4. (See page 53).

[5] K. B. McEachron, "Lightning to the Empire State Building," *Journal of the Franklin Institute*, February, 1939, Vol. 227, No. 2.

A TURNSTILE ANTENNA FOR USE AT ULTRA-HIGH FREQUENCIES

By

GEORGE H. BROWN

RCA Manufacturing Co., Inc., Camden, N. J.

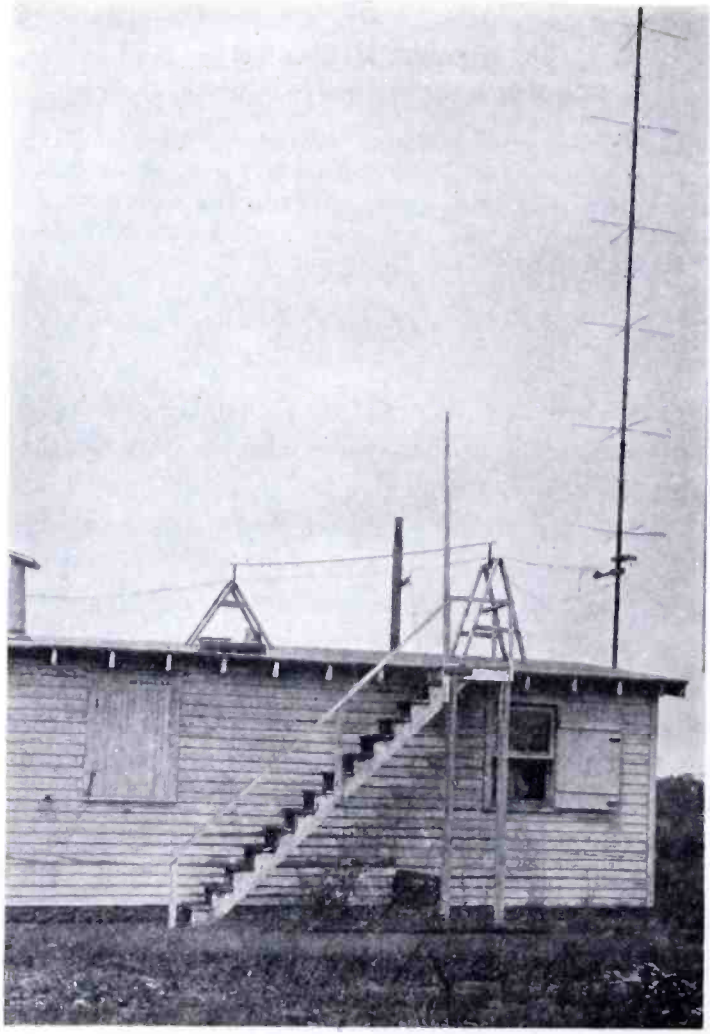
Summary—The Turnstile antenna provides a circular horizontal radiation pattern, with a concentration of energy in the vertical plane. This paper describes the antenna, its theoretical aspects, the experimental tests made on the antenna, as well as the procedure and the circuits necessary in its operation.

ANTENNAS operated at short wave lengths can have dimensions of the order of several wave lengths without becoming unwieldy in actual physical size. A great many arrangements can be used which direct the radiated energy in some one particular direction. This concentration thus yields a greater field strength than does, for example, a single half-wave antenna operated at the same power. If the signal strength in the remaining directions is of no consequence, the arrangement of a few wires in a directional array has accomplished the same result as building a more powerful transmitter. Now that the ultra-high frequencies are being used for broadcast purposes, the directional arrays are not always desirable. If the antenna is located in the heart of a city, it is desirable to radiate equal signals in all directions in a horizontal plane. It is still possible to rob energy from the high angles and concentrate it near the horizon. An investigation was undertaken to develop an antenna for ultra-high frequency use which embodied the following features.

1. The antenna should give a circularly symmetrical radiation pattern.
2. The antenna should concentrate the energy in the vertical plane so that the signal strength toward the horizon for a given power input will be considerably greater than that obtained from a single half-wave vertical antenna with the same input power.
3. The antenna must be structurally possible where high winds occur and should preferably be a rather simple structure not liable to damage easily.
4. If possible, the antenna should be supported by a single mast.

Reprinted from *Electronics*, April, 1936.

One system which immediately suggests itself is the "Franklin" antenna which consists of half-wave elements placed vertically, one above the other, and connected with phase shifting devices so that the currents in all the elements are in phase. This antenna fulfills the first two conditions, but it seems very difficult to meet the last two conditions.



Experimental model of the Turnstile antenna.

Theoretical Development

Arrangements which use horizontal elements usually do not fulfill the first condition stated, since each horizontal element yields a "Figure 8" as the horizontal pattern.

The antenna which was finally developed uses horizontal elements and fulfills the four conditions outlined.

Let us first examine the action of a horizontal half-wave antenna in free space. In the horizontal plane which passes through the an-

tenna, the field strength is horizontally polarized. The horizontal pattern has the shape of a "Figure 8" with the maximum intensity occurring in a direction normal to the axis of the antenna (Figure 1).

Suppose that another half-wave antenna is placed parallel to the first, and one-half wave length above the first antenna (Figure 2). The antennas are both excited so that the currents in each antenna are equal and in phase. This arrangement still yields a "Figure 8" pattern in the horizontal plane, but the magnitude of the horizontal pattern has increased since energy has been robbed from high angles and sent out horizontally.

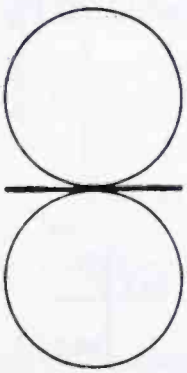


Fig. 1 — Pattern of horizontal half-wave antenna.

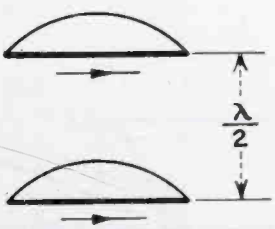


Fig. 2 — Two half-wave antennas ½ λ apart.

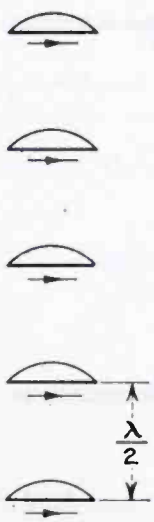


Fig. 3—Half-wave array.

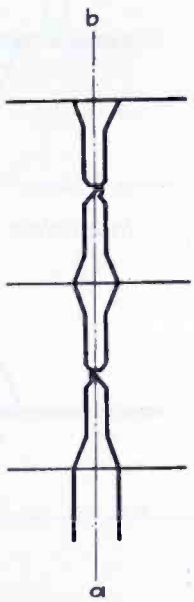


Fig. 4—Feed line.

If still more elements are placed in the array, Figure 3, each one-half wave from its neighbor and excited all in phase, the horizontal signal is still further increased, but the horizontal pattern still remains "Figure 8" in shape. For the time being, we will ignore this latter fact and consider means of constructing the arrangement of Figure 3.

Suppose that the elements are supported in space in some fashion. Then the elements can be excited in the proper phase and current magnitude by means of a single two-wire transmission line transposed once between each pair of elements. The half-wave length of transmission line gives a phase reversal of voltage along the line so that the single transposition returns the voltages on adjacent elements to the in-phase condition.

In Figure 4, the line *a-b* lies in a neutral plane with respect to the antenna elements and the transmission line. Thus if this line were a wire or piece of metal, there would be no voltage induced in it due to the radiating system. This fact makes it possible to replace the line *a-b* by a metal shaft or flag pole, thus affording a supporting structure for the system. Each half-wave antenna, instead of running through the pole, can consist of two quarter-wave rods screwed into opposite sides of the pole. The transmission lines, instead of having an abrupt transposition, twist continuously around the pole. It is possible to do this if supporting insulators are placed on the pole midway between the elements (Figure 5). This arrangement now fulfills all the condi-

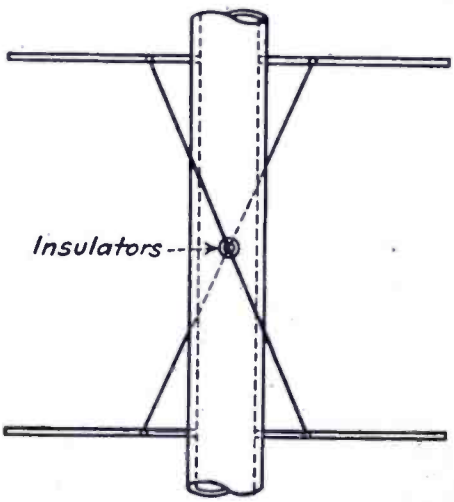


Fig. 5—Antenna wires and feed lines.

Fig. 6 — Looking down on antenna.

tions imposed except that of the circularly symmetrical horizontal pattern.

On our flag pole, let us put a second system of radiators and transmission line identical with the first, but so placed that the two sets of radiators are at right angles and corresponding elements are at the same level on the pole. Thus with two sets of identical elements on the pole, we have two separate transmission lines coming down the pole to the transmitter. These two transmission lines are so fed, with equal power into each line, so that the currents in one set of radiators are in time quadrature with the currents in the other set which is at right angles in space with the first set. Figure 6 shows a view of the antenna looking down from the top. Then the field in the horizontal plane due to Set No. 1 is

$$F_1 = I \sin(\omega t) \sin\theta \qquad (1)$$

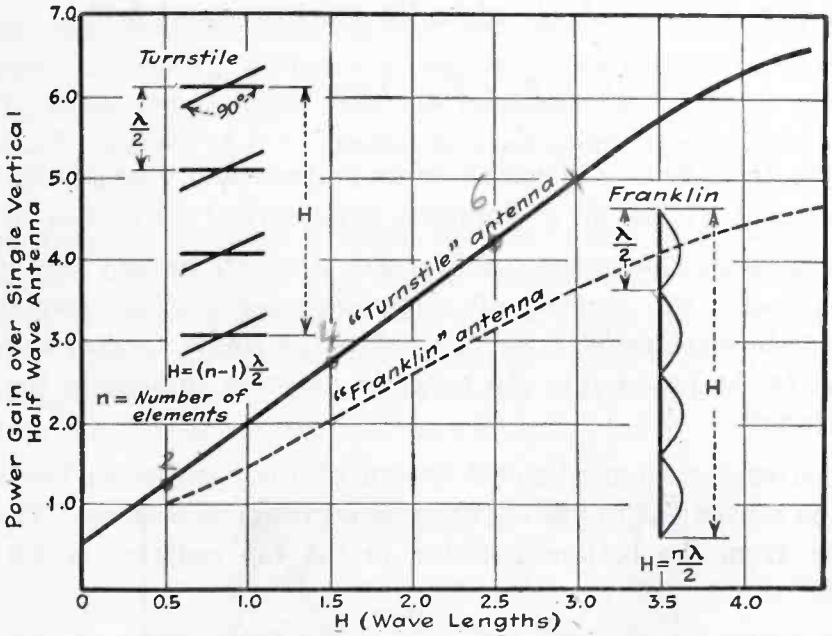


Fig. 7—Theoretical power gain of Turnstile and Franklin antennas.

where θ is the angle indicated on Figure 6.

The field due to Set No. 2 is

$$F_2 = I \cos(\omega t) \cos\theta \qquad (2)$$

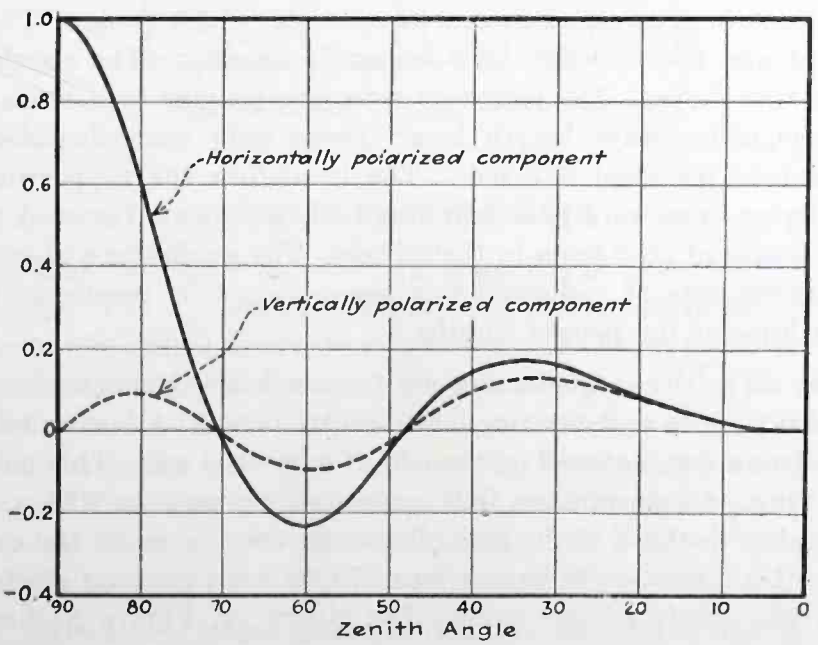


Fig. 8—Vertical radiation of 6-element Turnstile antenna.

The sum of (1) and (2) gives the total resultant field.

$$F_t = F_1 + F_2 = I \cos(\omega t - \theta) \tag{3}$$

Thus the total field is constant in magnitude and changes in phase as θ changes, giving us a circularly symmetrical horizontal pattern.

As mentioned previously, the signal strength toward the horizon increases with the number of antenna elements. The ordinates of Figure 7 show the ratio of the power into the single vertical half-wave antenna to the power into the array in question to achieve the same field strength.

Experience with mechanical design of these antennas has shown that it is convenient to use six antenna elements in each set. Then the distance from the bottom radiator to the top radiator is 2.5 wave lengths.

The vertical radiation characteristic of this antenna is made up of a horizontally polarized component and a vertically polarized component. The vertically polarized component becomes zero in the horizontal plane. Figure 8 shows these two components as a function of the angle measured from the zenith, when the antenna consists of six elements per set.

While this antenna layout looked very good on paper, it was necessary to verify the results experimentally. The experimental method was also used to determine certain optimum dimensions. Accordingly, a model was built to operate on a wave length of 3.0 meters. The flag pole used was 42 feet long and 3 inches in diameter. The six-element antenna was chosen. The radiators were one-quarter inch brass rods, each one-quarter wave length long. These rods were threaded and screwed into the steel flag pole. The insulators for supporting the transmission lines were porcelain stand-off insulators, fastened to the pole by means of stud bolts in their bases. The quadrature phase relation between sets of radiators was accomplished by means of transmission lines of the proper lengths.

After all critical adjustments were made, a check was made of the horizontal pattern to determine how circular it was. A horizontal half-wave antenna was mounted on the end of a bamboo pole. This pole was 26 feet long. A transmission line connected this antenna with a detector placed at the base of the pole. Readings were taken on the circumference of a circle whose radius was 175 ft. with the axis of the flag pole as the origin of the circle. The circles on Figure 9 show the results of this test.

Next, the elements pointing east and west were disconnected to determine the expected "Figure 8" pattern. The crosses on Figure 9 show the measured results. This test indicates the necessity of using two sets of elements if it is desired to send equal signals in all directions.

A measure of the field strength was made at a fixed point. Then the flag pole was replaced by a single vertical half-wave antenna excited with the same power. The field strength from this arrangement was slightly less than one-half that obtained with the array, indicating a power gain for the array of approximately four to one. From Figure 7, we find the theoretical figure to be 4.27 to 1.

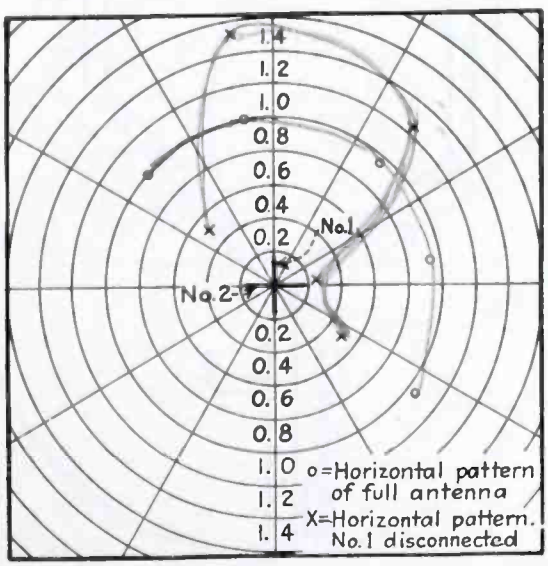


Fig. 9—Result of field measurements.

Constructional Details

The first full scale antenna was constructed for operation at 45 megacycles. The supporting pole extended 70 feet above the roof. The radiators were nickel-steel tubes, copper plated. These tubes were made with sufficient wall thickness to allow a slight taper. This taper is supposed to avoid possible fracture due to vibration of the tubes. The stand-off insulators are 8 inches in length. The transmission lines are made up of No. 8 hard drawn wire. The antenna elements are placed slightly less than one-half wave length apart so that the transmission line length between elements is exactly one-half wave.

To achieve the proper phase shifts, the antenna is fed by an arrangement of transmission lines as shown in Figure 10. If the lines

have a characteristic impedance of 500 ohms, the following dimensions will hold *approximately*.

1. The distance from *c* or *d* to the lowest antenna element is 0.355 wave lengths (plus any integral number of half-wave lengths desired).

2. The length of *c* (or *d*) from line connection to shorting bar is 0.085 wave lengths.

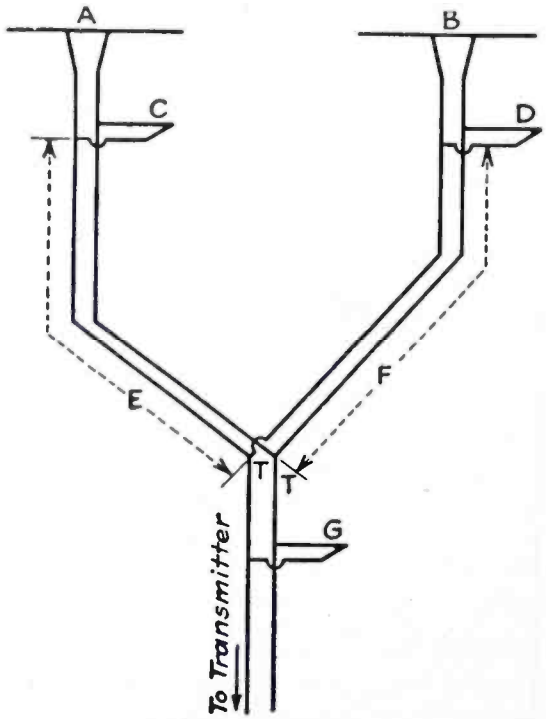


Fig. 10—Transmission line arrangement.

3. *e* is any convenient length.

4. *f* equals *e* plus one-quarter wave length.

5. The distance from *T-T* to point of connection of *g* is 0.4 wave lengths.

6. The length of *g* is 0.15 wave lengths.

Four strain insulators are placed at the top of the pole to neutralize the pull on the top elements due to the transmission lines.

The transmission lines are connected to the antenna elements by means of clamps placed 0.06 wave lengths from the surface of the supporting pole.

Small metal pads are fabricated to the pole to insure firm horizontal mounting for the insulators and radiating rods.

In designing a particular antenna, consideration should be given the highest recorded wind velocity, the possibility of the formation of sleet on the antenna, and the possibility of excessive corrosion due to proximity to salt water.

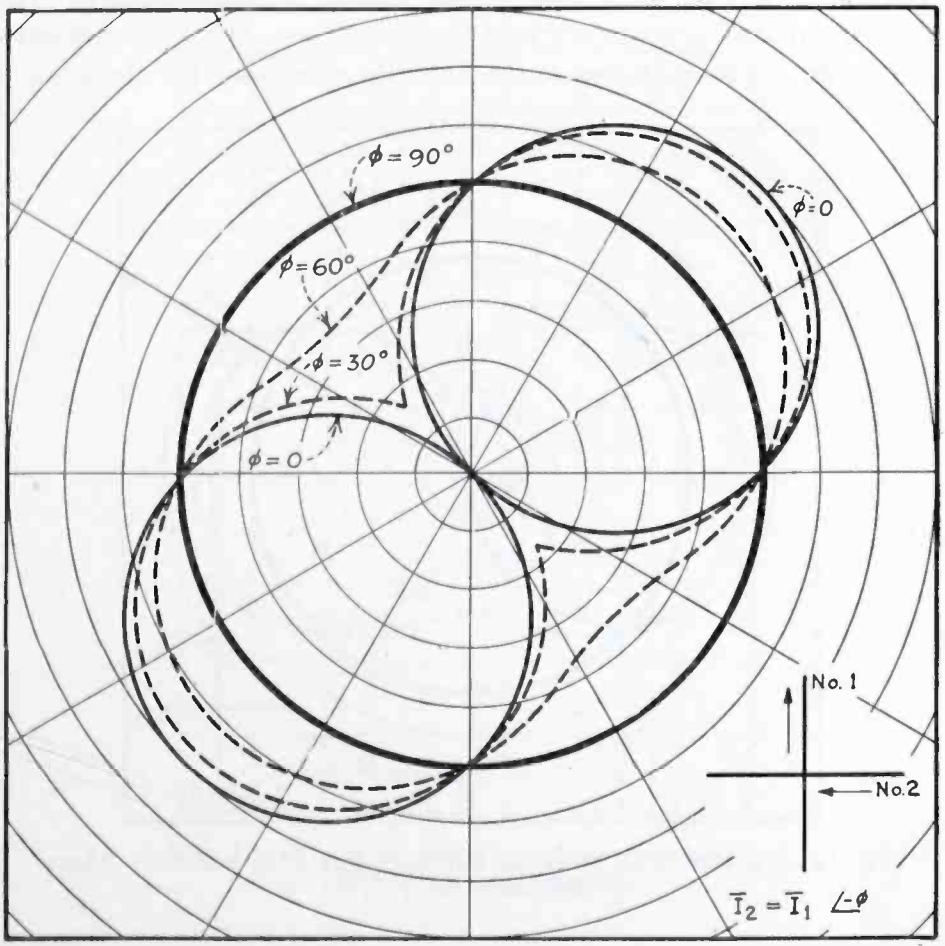


Fig. 11—Horizontal radiation pattern—currents of equal magnitude but varying in phase.

Factors Affecting Horizontal Radiation Pattern

The design of the turnstile antenna has been based on the premise that it is most desirable to have a circularly symmetrical horizontal radiation pattern. It is, however, conceivable that such an antenna might be located in the heart of a city which is oblong in shape. In this event, it would be desirable to have a horizontal pattern which is elongated. This may be accomplished by controlling the phase relation

points are naturally scattered since the measurements were made through the city.

The theoretical curves show that the field intensity drops off as the inverse square power of the distance, and thus is a straight line with a slope of —2.0 when plotted on log paper. By means of the theory of least squares, an analysis was made of the experimental points for all measurements made at one mile or more. The best straight line on log paper was found to have the equation

$$E = 170/r^{1.98}$$

where E is the field strength in millivolts per meter and r is the distance in miles. Curve *E*, Figure 13, shows this equation.

When the 45 megacycle antenna was placed in operation, another striking effect was noticed. Observers reported that, in districts where signals from a single half-wave antenna had fluctuated as much as ten to one due to changes in field distribution due to moving automobiles and possibly elevator cables, the signal from the turnstile only shifted between limits whose ratio was two to one. This effect is probably due to the fact that the transmitting antenna is spread through a space two and one-half wave lengths long, thus giving "diversity" effect.

FREQUENCY CONTROL BY LOW POWER FACTOR LINE CIRCUITS*

By

CLARENCE W. HANSELL AND PHILIP S. CARTER

(R.C.A. Communications, Inc., Rocky Point, L. I., New York)

Summary—*This paper points out the advantages of concentric conductor lines as low power factor or high Q resonant circuits for controlling the frequency of very high frequency oscillators. The electrical characteristics of lines of various dimensions at various frequencies is given. Several forms of temperature compensated lines are described. Oscillator circuits, circuit combinations, and precautions for obtaining stable transmitter frequencies are suggested. Photographs of typical line controlled transmitters are included. The results obtained with line control indicate that the method has great potential usefulness comparable with the usefulness of piezoelectric crystal control.*

INTRODUCTION

IN A previous paper[1] it was pointed out that sections of transmission line could be used to stabilize the frequencies of transmitters. It was suggested that the lines, serving as very low power factor circuits, might be used advantageously to replace piezoelectric crystals in transmitters operated with relatively high power output and at very high frequencies. Since publication of the previous paper much progress has been made in the design of lines for frequency control and in their application to transmittters operated with output frequencies ranging from about seven to 500 megacycles.

The form of line best suited for frequency control is one made up of two concentric conductors, with the outer conductor completely enclosing the inner one. This form of line is relatively easy to construct and is completely shielded. The power factor of the line as a resonant circuit is not increased by radiation or coupling to surrounding objects and circuits. If desired, the outer conductor may be utilized as a means of mounting and support for tubes and other circuit elements. Copper is one of the most satisfactory materials from which to construct the line for ordinary applications but aluminum or aluminum alloys may be used where weight is an important consideration.

---

* Decimal classification: R355.6. Presented before I.R.E. Tenth Annual Convention, Detroit, Mich., July 1, 1935.

[1] Conklin, Finch, and Hansell, "New Methods of Frequency Control Employing Long Lines," *Proc. I.R.E.*, vol. 10, pp. 1918-1930; November, (1931).

Reprinted from *Proc. I.R.E.*, April, 1936.

points are naturally scattered since the measurements were made through the city.

The theoretical curves show that the field intensity drops off as the inverse square power of the distance, and thus is a straight line with a slope of —2.0 when plotted on log paper. By means of the theory of least squares, an analysis was made of the experimental points for all measurements made at one mile or more. The best straight line on log paper was found to have the equation

$$E = 170/r^{1.98}$$

where E is the field strength in millivolts per meter and r is the distance in miles. Curve E, Figure 13, shows this equation.

When the 45 megacycle antenna was placed in operation, another striking effect was noticed. Observers reported that, in districts where signals from a single half-wave antenna had fluctuated as much as ten to one due to changes in field distribution due to moving automobiles and possibly elevator cables, the signal from the turnstile only shifted between limits whose ratio was two to one. This effect is probably due to the fact that the transmitting antenna is spread through a space two and one-half wave lengths long, thus giving "diversity" effect.

FREQUENCY CONTROL BY LOW POWER FACTOR LINE CIRCUITS*

By

CLARENCE W. HANSELL AND PHILIP S. CARTER

(R.C.A. Communications, Inc., Rocky Point, L. I., New York)

Summary—*This paper points out the advantages of concentric conductor lines as low power factor or high Q resonant circuits for controlling the frequency of very high frequency oscillators. The electrical characteristics of lines of various dimensions at various frequencies is given. Several forms of temperature compensated lines are described. Oscillator circuits, circuit combinations, and precautions for obtaining stable transmitter frequencies are suggested. Photographs of typical line controlled transmitters are included. The results obtained with line control indicate that the method has great potential usefulness comparable with the usefulness of piezoelectric crystal control.*

INTRODUCTION

IN A previous paper[1] it was pointed out that sections of transmission line could be used to stabilize the frequencies of transmitters. It was suggested that the lines, serving as very low power factor circuits, might be used advantageously to replace piezoelectric crystals in transmitters operated with relatively high power output and at very high frequencies. Since publication of the previous paper much progress has been made in the design of lines for frequency control and in their application to transmittters operated with output frequencies ranging from about seven to 500 megacycles.

The form of line best suited for frequency control is one made up of two concentric conductors, with the outer conductor completely enclosing the inner one. This form of line is relatively easy to construct and is completely shielded. The power factor of the line as a resonant circuit is not increased by radiation or coupling to surrounding objects and circuits. If desired, the outer conductor may be utilized as a means of mounting and support for tubes and other circuit elements. Copper is one of the most satisfactory materials from which to construct the line for ordinary applications but aluminum or aluminum alloys may be used where weight is an important consideration.

---

* Decimal classification: R355.6. Presented before I.R.E. Tenth Annual Convention, Detroit, Mich., July 1, 1935.

[1] Conklin, Finch, and Hansell, "New Methods of Frequency Control Employing Long Lines," *Proc. I.R.E.*, vol. 10, pp. 1918-1930; November, (1931).

Reprinted from *Proc. I.R.E.*, April, 1936.

The degree to which a line can be made the predominating element in determining the frequency of an oscillator is proportional to the amount of oscillatory energy which may be maintained in it with a given amount of power. Therefore the quality or figure of merit for the line may be taken as the ratio of oscillatory energy to power loss. This ratio is popularly known as the Q of the line.

The effective length of line may be near any number of quarter wave lengths. Making the line longer than a quarter or a half wave will not improve its Q, or sharpness of tuning, but it will increase the amount of oscillatory energy which can be stored in the line without flashover or excessive temperature rise. Usually it is preferable to obtain the desired storage rating by using large diameters rather than a length exceeding a quarter or half wave.

Characteristics of Lines

The mathematical determination of the characteristics of concentric conductor lines gives the results listed in the following table of formulas, the derivation of some of which will be given in an appendix:

Symbols

a = radius of outside surface of inner conductor in centimeters

b = radius of inside surface of outer conductor in centimeters

f = frequency

I = current in line at point of maximum current

λ = wavelength in meters

Formulas

Inductance, $L = 2 \times (10)^{-7} \log_e b/a$ henrys per meter.

Capacity, $C = (10)^{-9}/(18 \log_e b/a)$ farads per meter.

Characteristic impedance, $Z = 60 \log_e b/a$ ohms.

Resistance, $R = 41.6(10)^{-7} \times \sqrt{f}(1/a + 1/b)$ ohms per meter for a line constructed of copper.

Attenuation constant $\alpha = R/2Z$.

Power loss in a tuned line, $W = I^2R\lambda/8$ watts per quarter wave of line for a line constructed of copper.

Oscillatory energy, $VA = \pi fLI^2\lambda/4 = I^2\lambda/16\pi fC$ per quarter wave.

Figure of merit of a tuned copper line, $Q = VA/W = 2\pi fL/R = 1/2\pi fCR$.

Maximum figure of merit obtainable with a copper line, for a given value of b, is $Q_{max} = 1/(6.86(10)^{-4}\sqrt{\lambda}/b) = 1460\ b/\sqrt{\lambda}$.

Ratio b/a giving maximum Q, for a given value of b is 3.6. (See footnotes 2, 3, and 4.)

Value of maximum voltage gradient $= E/(a \log_e b/a)$.

Ratio b/a giving smallest maximum voltage gradient for a given maximum voltage and a given value of b is 2.72.

Ratio b/a giving minimum voltage gradient for a given oscillatory energy and a given value of b is 1.65.

Assuming equal thickness of inside and outside conductor the greatest oscillatory energy storage is obtainable per pound of copper when $b/a = 4.68$.

To obtain maximum impedance with a section of line, having a given value of b, (for example, for use as the equivalent of an insulator) the ratio of b/a should be 9.18.

Maxims

There are certain maxims made apparent by the study of characteristics of concentric conductor lines which are very useful in rapid interpolation of line characteristics.

1. The Q of a line is inversely proportional to the square root of the resistivity of the material used in it.

2. The Q of a line is proportional to the square root of the frequency and inversely proportional to the square root of the wavelength.

3. The Q of a line is proportional to the diameter of the conductors so long as the ratio of diameters is constant.

4. The maximum allowable oscillatory energy in a line is substantially proportional to the square of the diameters so long as the ratio of diameters is constant.[5]

Examples

Figure 1 shows the value of the figure of merit, Q for various frequencies and diameters of the outer conductor of concentric conductor copper lines, assuming a ratio of diameters of 3.6. The values of Q for other materials and frequencies may be determined readily with the aid of maxims 1 and 2.

---

[2] C. S. Franklin, British Patent No. 284,005 and corresponding U. S. Patent No. 1,937,559.

[3] Sterba and Feldman, "Transmission Lines for Short-Wave Radio Systems," *Proc. I.R.E.*, vol. 20, pp. 1163-1202; July, (1932).

[4] F. E. Terman, "Resonant Lines in Radio Circuits," *Elec. Eng.*, vol. 53 pp. 1046-1061; July, (1934).

[5] For more exact laws of variation in flashover voltage gradient and energy storage of lines of different dimensions see "Dielectric Phenomena in High Voltage Engineering," published by McGraw-Hill, and other publications of F. W. Peek, Jr.

At 60 megacycles the minimum length of inner conductor to tune would be about 125 centimeters (49 inches). From mechanical considerations a reasonable diameter of outer conductor might be taken as 60 centimeters (24 inches). The inner conductor would be 6.5 inches in diameter. The over-all dimensions required for the finished line would be about $24 \times 24 \times 72$ inches. A line of this size would have a Q of about 20,000. Only ten watts of input power would be required to maintain an oscillatory energy of 200 kilovolt-amperes in this line.

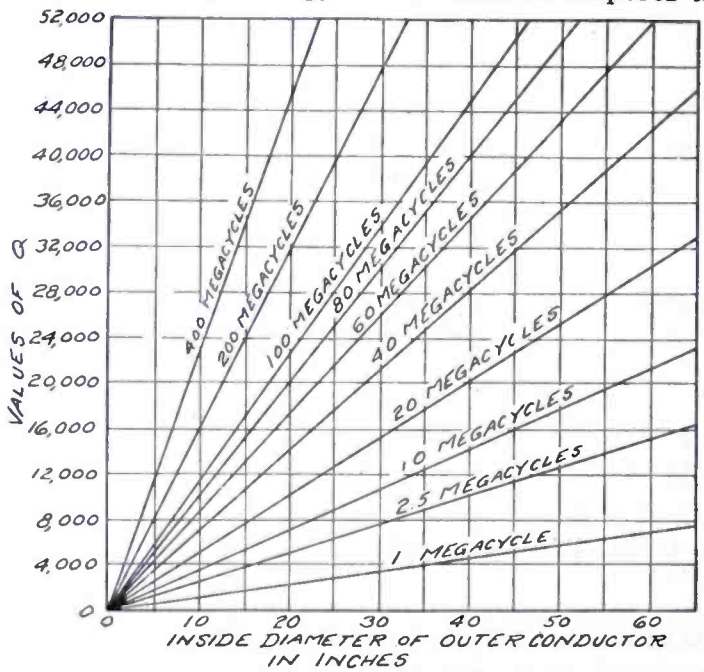


Fig. 1—Values of Q (inverse of power factor) for concentric conductor copper lines of various diameters at various frequencies, assuming a ratio of diameters of 3.6.

Temperature Coefficient

Tests and theory indicate that lines made up of straight tubular conductors have a temperature coefficient of frequency variation corresponding fairly closely to the mechanical temperature coefficient of linear expansion for the material of which the line is made. So long as both conductors have the same temperature, the ratio of their diameters and, therefore, the electrical constants per unit of length do not change with temperature. To a reasonable degree the change in frequency with change in temperature can be considered as due only to change in length.

In practice, lines used for frequency control are also subject to frequency variations due to unequal heating of inner and outer conductors. This effect is most evident in lines used with relatively large power dissipation and low frequencies and causes a temporary frequency drift while the line is warming up. The effect can be made

small by using large dimensions and heavy material of good heat conductivity in the line. It is not very important in oscillators operated at 50,000 kilocycles or higher with power levels obtainable from commercially available tubes.

The approximate temperature coefficient of linear expansion, the resistivities and heat conductivities for materials which may be used advantageously in the construction of lines are as follows:

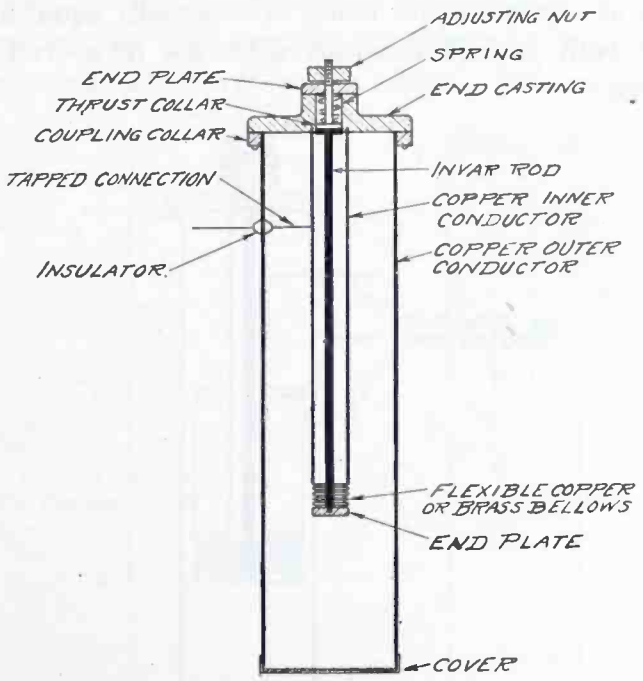


Fig. 2—Cross section of concentric conductor line one-quarter wave long having frequency adjustment and temperature compensation by means of flexible metal bellows and invar rod.

Material	Temperature Coefficient Parts per Million	Electrical Resistivity Ohm/cm³ × (10)⁶	Heat Conductivity Calories/cm³/°C
Copper	16.8	1.7	0.9
Aluminum	23.1–25.5	2.8	0.5
Brass	19	6.4–8.4	0.2–0.26
Invar (1st category)	0.8 or less	80	0.025

Reducing the Temperature Coefficient

One of the simplest and most effective means for reducing the temperature coefficient of frequency variation is to hold the effective length of the inner conductor constant regardless of temperature. This can be done in the manner shown in Figure 2, where a small portion of the inner conductor is made in the form of a flexible metal bellows, and the inner conductor, including the bellows, is held constant in length

by means of a rod of material, such as invar, which has a very low expansion coefficient. This construction is also particularly well adapted to making exact adjustments of frequency by adjusting the free length of the invar rod to stretch or compress the flexible bellows.

Figure 3 shows a form of line constructed with two sizes of inner conductor in such a way that the over-all length of line required to tune to a given frequency is greatly reduced. The lengths of each of the two sizes of conductor are made substantially equal and the over-all length of both is held constant with the invar rod and flexible bellows system.

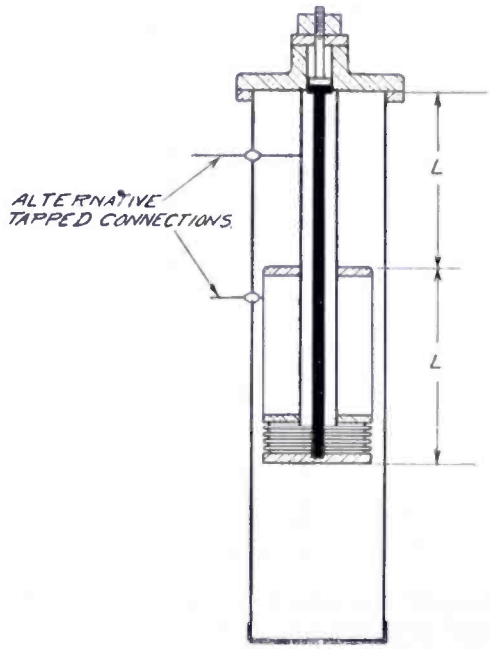
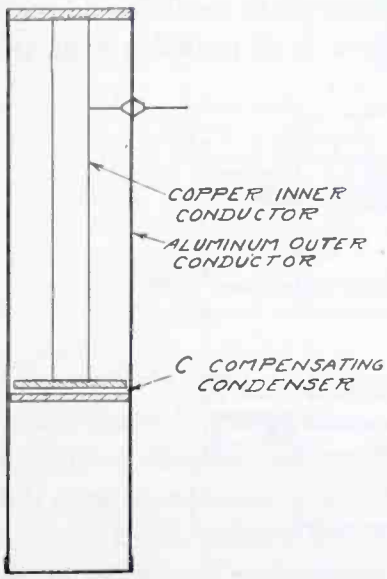


Fig. 3—Cross section of concentric conductor line having frequency adjustment and temperature compensation combined with shortening of over-all length by use of two diameters of inner conductor.

In such a line the smaller conductor and outer pipe form an effective inductance while the larger conductor and outer pipe form an effective capacity. The inductance and capacity are each very nearly proportional to the length of the respective conductors. Since the over-all length of the two inner conductors is constant, and the two are equal in length, any elongation or contraction of the smaller conductor, due to change in temperature, causes an equal and opposite percentage change in the larger conductor. Thus changes in temperature vary the inductance and capacity of the circuit equally and oppositely and there is little if any change in natural frequency. If lines are to be used which are physically shorter than a quarter wave, the general arrangement of Figure 3 is a satisfactory form of construction.

Another method for reducing the temperature coefficient and at the same time shortening the line is indicated in Figure 4. This form of line makes use of the difference in temperature coefficient of expansion of copper and aluminum to vary the capacity of *C* in a direction tending to compensate for variation in length of the inner conductor. If the line temperature rises the aluminum expands more than the copper and so increases the spacing of the plates at *C*. This decreases the capacity at *C* and tends to increase the resonant frequency of the line, compensating the tendency for the frequency to decrease as the copper inner conductor increases in length. This arrangement also



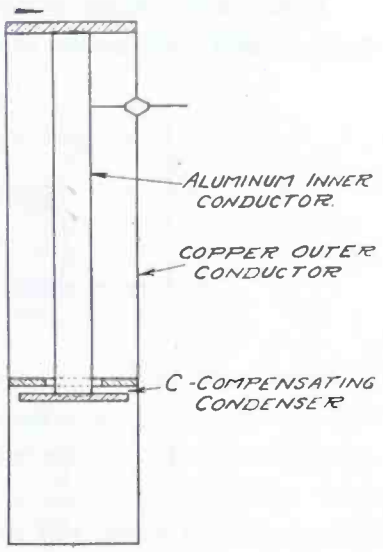


Fig. 4—Cross section of temperature compensated line employing copper inner conductor and aluminum outer conductor.

Fig. 5—Cross section of temperature compensated line employing aluminum inner conductor and copper outer conductor.

tends to compensate for the higher temperature rise of the inner conductor due to losses in the line. Figure 5 is an inverted variation of the arrangement shown in Figure 4. Both these arrangements require careful design, construction, and adjustment and are not so convenient to use as the lines of Figures 2 and 3.

Figures 2, 3, 4, and 5 show quarter-wave lines which are most obviously applicable to single tube oscillators. Of course, the lines may also be made a half-wave long, equivalent to two quarter-wave sections in series, in which case they are most readily applicable to push-pull oscillators. To save over-all length the half-wave system may be bent into the shape of a U as shown in Figure 6.

Figure 6 shows a means for compensating the effects of temperature variation by using the relatively large variation in volume of

some liquids, such as lubricating oil, with change in temperature to operate a condenser plate. The oil is contained in a pipe system exposed to the same temperature variations as the line and operates the condenser plate through a flexible metal bellows and a lever arm. By suitably locating the oil piping and varying its insulation, together with proper adjustment of the condenser plate and bellows it is possible to obtain quite close compensation for temperature variations due to changes in ambient temperature and due to line losses. However, considerable skill and patience is required in making the design and adjustment.

Eventually, when materials of low temperature coefficient become more readily available in commercial shapes it is probable that they

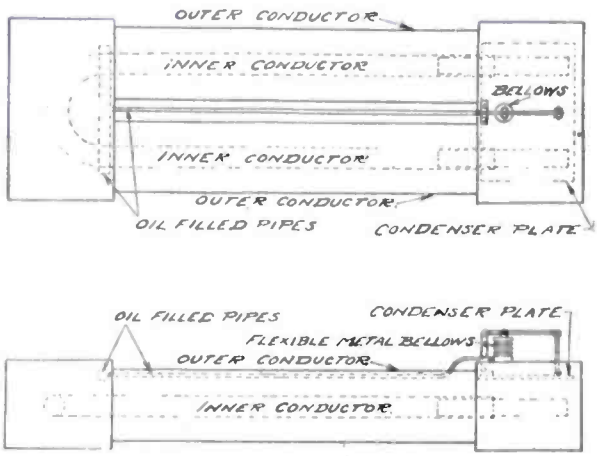


Fig. 6—Half-wave line with temperature compensation by means of condenser plate controlled by expansion and contraction of a liquid.

will be used to obtain lines having low temperature coefficients without compensating arrangements. Lines constructed of these materials will require silver or copper plating of the conducting surfaces.

Internal stresses in the material of the lines, which may cause progressive changes in dimensions due to temperature cycling, should be avoided by care in manufacturing or by suitable heat treatment. This source of frequency variation is most troublesome in newly constructed lines and will usually decrease with age.

Circuits for Line Controlled Oscillators

In general, lines may be used for frequency control, and will operate in a manner similar to piezoelectric crystals. They differ from crystals chiefly in their ability to control high power oscillators and in their ability to perform satisfactorily at frequencies far higher than can be reached by crystals. As we go to higher frequencies, crystals

gradually become less useful in stabilizing oscillators while lines increase in their stabilizing ability.

Figure 7 shows a line controlled oscillator similar to one often used for crystal oscillators. In operating this circuit it is best to adjust the regeneration control condenser to give about the smallest feedback

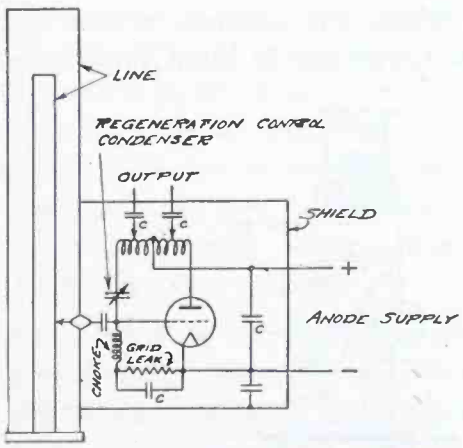


Fig. 7—Single tube line controlled oscillator circuit.

from plate to grid circuit which can be used to make the oscillator work efficiently. Any excess feedback reduces the ability of the line to stabilize the frequency. The circuit will function with the regeneration control condenser set either above or below the capacity value required for a balance but one adjustment or the other will be prefer-

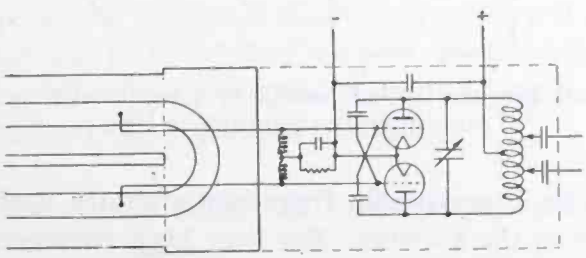


Fig. 8—Push-pull line controlled oscillator circuit with half-wave line.

able depending upon the ratio of effective resistance in anode and grid circuits and the frequency. The resistance ratio, electron time lag at very high frequencies and feed-back adjustment are all factors which should be taken into account to determine which adjustment will give the best phase relation between anode and grid radio-frequency voltages.

Figure 8 is a push-pull oscillator circuit similar in principal and operation to the single tube circuit of Figure 7.

Figure 9 is a circuit suitable for stabilizing the frequency of a push-pull oscillator by means of a quarter-wave line. In this circuit the grids of the two tubes are inductively coupled to the line by means of coupling loops of opposite polarity.

Line Controlled Transmitter Combinations

In some cases, where the antenna system is made mechanically rigid and free from variations in input impedance due to weather, it

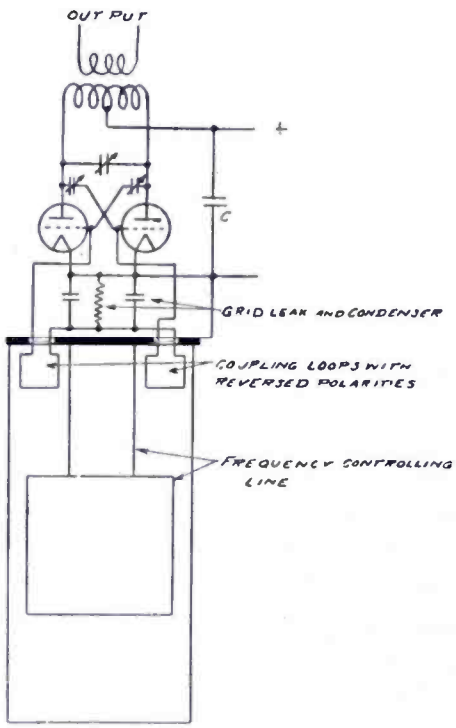


Fig. 9—Push-pull line controlled oscillator circuit with quarter-wave line and inductive coupling to line.

is possible to obtain acceptable frequency stability with the oscillator coupled directly to the antenna. For very high frequency transmitters of limited range, in places where interference is not a problem, single stage transmitters will often be entirely satisfactory and can be recommended. The interisland radiotelephone system of the Mutual Telephone Company, in Hawaii, where single stage line controlled transmitters have been in regular commercial operation since 1931, is an excellent example.[1,6]

In most cases, however, it is desirable or necessary to interpose one or more stages of amplifier between the line controlled oscillator and

[6] Beverage, Peterson, and Hansell, "Application of Frequencies Above 30,000 Kilocycles to Communication Problems," *Proc. I.R.E.*, vol. 19, pp. 1313-1333; August, (1931).

the antenna. For most ordinary requirements a single radio-frequency amplifier between oscillator and antenna will suffice if the amplifier is carefully neutralized and shielded to prevent feedback to the oscillator.

For very accurate frequency control, such as will be required in the future to make maximum use of very high frequencies, two stages of amplifier following the oscillator are recommended. At frequencies near the upper limit for the tubes it is desirable to operate the master oscillator at a half or a third of the final output frequency and to follow it with a frequency multiplying amplifier and a power amplifier. The use of frequency multiplication permits the oscillator to work at a lower frequency where the tubes are more efficient and very greatly reduces the probability of frequency changes due to variable radio-frequency feedback from the later stages.[7]

When line control is used to control a relatively low-frequency transmitter it will sometimes be desirable to operate the oscillator at two or three times the output frequency and to follow it with a controlled oscillator and power amplifier, both operated at the output frequency. This greatly reduces the dimensions of line, and improves its frequency holding ability. By careful design and adjustment the controlled oscillator can be made to serve as an uncoupling link between the oscillator and power amplifier with an effectiveness about equal to that obtained from a frequency multiplier.

The controlled oscillator should have weak regeneration at the output frequency so that the grid voltage at the output frequency and the grid voltage at the harmonic input frequency will not be greatly different. It may also be noted that the controlled oscillator should not be amplitude modulated by any great amount. Amplitude modulation should be applied only to the power amplifier.

Precautions for Reducing Undesired Modulations

In general, line controlled oscillators require the same precautions for obtaining a pure continuous wave output as are required for crystal oscillators. Ripples in the direct power voltages and alternating cathode heating current all tend to produce undesired amplitude phase, and frequency modulations of the output. Because of the line these undesired modulations will be far less than they would be in a simple oscillator but they will always be present.

The most obvious means for reducing these modulations are to use very smooth direct voltages and direct-current cathode heating. Both

---

[7] Hallborg, Briggs, and Hansell, "Short-wave Commercial Long-distance Communication," *Proc. I.R.E.*, vol. 15, pp. 467-500; June, (1927).

of these expedients are undesirable from the standpoint of cost, simplicity, and reliability but may be necessary in some cases. In other cases, satisfactory results may be obtained while using alternating-current cathode heating and impure direct-current anode supply by taking a few simple precautions.

Amplitude modulations introduced in all but the last stage, may be kept small by using sufficient excitation in the later stages to produce limiting. If the last stage is amplitude modulated by means of the Heising constant current modulating system this stage is automatically provided with very good anode supply smoothing by the modulation choke.

Elimination of undesired phase and frequency modulations requires holding constant tube impedances and the use of circuits tending to minimize the effect of changing impedances upon the tuning of the circuits. Small grid current and the use of grid leak and cathode return resistor biasing, particularly in the oscillator and next succeeding amplifier, assist materially in holding constant effective grid impedances, provided the resistances have sufficiently small parallel dielectric capacity to prevent appreciable phase lag in bias variation in response to radio-frequency amplitude variations. Also the effective series radio-frequency reactance from the anodes of one stage to the grids of a succeeding stage must be made a minimum. Leads from the output circuit of one stage to the input circuit or grids of a succeeding stage should be large and extremely short or else equal to a half wave or multiples of a half wave long. Leads which are near a quarter wave or multiples of a quarter wave in length should be carefully avoided.

It is interesting to note that any phase or frequency modulation noise introduced in an early stage of a transmitter will be increased in proportion to the amount of frequency multiplication used after that stage. A crystal controlled transmitter having an output of 100,000 kilocycles would probably start out with an oscillator frequency of about 3125 kilocycles. One degree of phase modulation in the output of the crystal oscillator would then appear as thirty-two degrees in the output of the transmitter and produce side frequency energy equivalent to that obtained with about sixty per cent amplitude modulation. With line control the oscillator may be operated at the output frequency so that one degree of phase modulation in the oscillator will appear as one degree in the transmitter output.

Mechanical vibration of the line and circuits must be prevented. This requires rigid construction of the line and all coils, condensers, leads, etc. Poor workmanship must be avoided. If the equipment is to be operated near rotating machinery or other sources of vibration

it is desirable that the whole radio-frequency system be hung on springs and rubber shock absorbers so adjusted that the rubber is subjected to little initial stress.

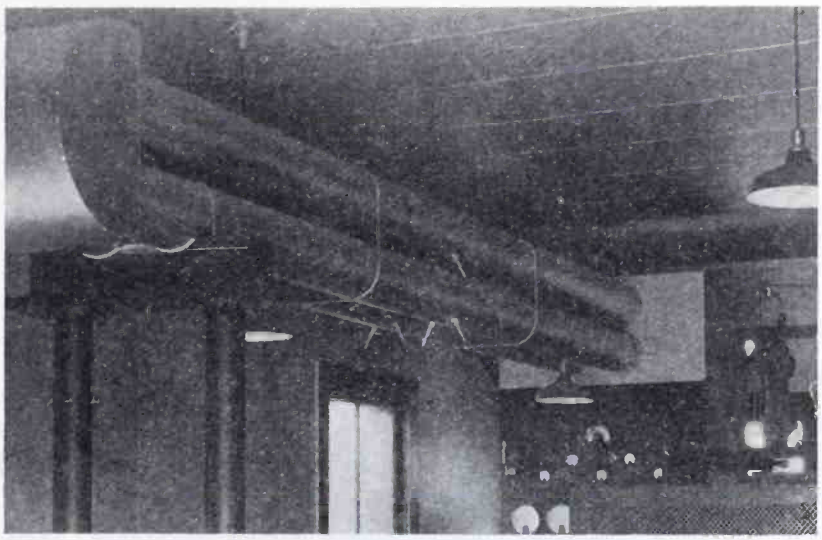


Fig. 10—Line controlled master oscillator for commercial transmitter WQO-6725 kilocycles.

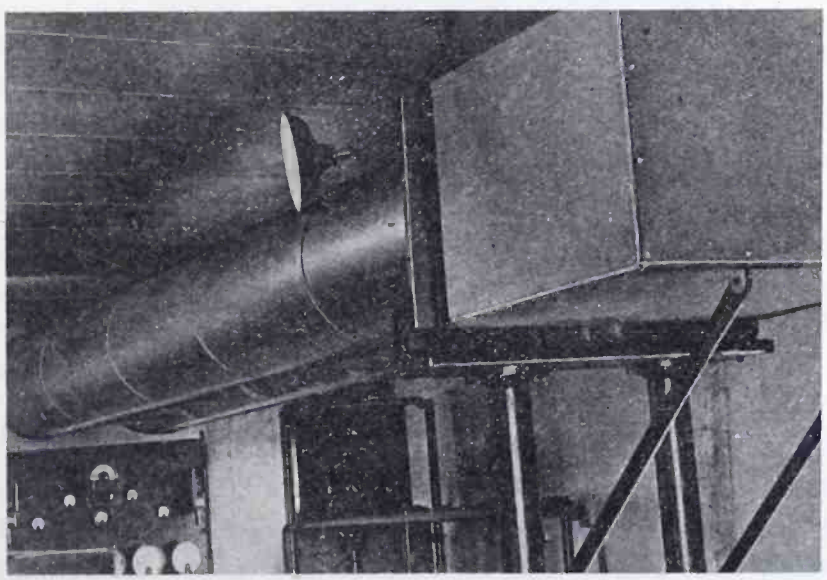


Fig. 11—Line controlled master oscillator for commercial transmitter WHR-13,420 kilocycles.

Examples of Line Controlled Transmitters

Transmitters WQO and WHR

Figures 10 and 11 show the general construction and mounting of the lines used to control the frequencies of transmitters WQO—6725

kilocycles and WHR—13,420 kilocycles, respectively, at the Rocky Point, New York, station. These two transmitters each have two RCA-846 tubes in a push-pull master oscillator, employing the circuit shown in Figure 8, and two UV-858 tubes in a push-pull power amplifier. The master oscillators are located inside the boxes at the ends of the lines, above the power amplifier units. Keying is accomplished by grid-bias control of the master oscillators. Both transmitters are supplied with anode power from a common rectifier. The two transmitters are operated alternately for day and night service but can be operated simultaneously at reduced power. Both are started, stopped, and keyed by remote control over the control circuit from New York, about sixty miles away. Application of keying energy to the control circuits in New York automatically starts up either transmitter and removal of the keying for several minutes automatically shuts it down.

These two line installations are each a half wave long and employ the arrangement illustrated in Figure 6 for compensating the effect of temperature variations upon the frequency. The line for WQO—6725 kilocycles has an outer pipe eight inches inside diameter, #10 Stubs gauge copper and an inner pipe of two-inch iron pipe size copper. The line is built up in four sections joined together at the ends into one half-wave oscillator circuit.

The line for WHR—13,420 kilocycles has an outer pipe twenty inches diameter, #12 Stubs gauge copper and an inner pipe of five-inch iron pipe size copper. It is built up in two sections bent into one U-shaped oscillator circuit.

The RCA Central Frequency Bureau, at the Riverhead, New York, receiving station, makes frequent routine checks on the frequencies of all transmitters operated by R.C.A. Communications, Inc., as well as those of most other transmitters engaged in long-distance service. These routine checks have been utilized to determine the relative effectiveness of line control in maintaining the frequencies of WQO and WHR.

Taking the period of May 1, 1934, to April 30, 1935, as representative, these checks show maximum variations of WQO and WHR to be 0.01 and 0.015 per cent, respectively. The average of maximum variations reported in any one week were WQO—0.0052 and WHR—0.0056 per cent. The average of weekly maximum variations reported for a representative group of fifty crystal controlled transmitters in the same period was 0.0115 per cent.

Figure 12 is a photograph of W2XHG—25,700 kilocycles, 100 to 150 watts, located on the roof of the RCA Building, New York City. This transmitter as shown in the photograph had a line controlled

master oscillator followed by a modulated power amplifier, utilizing RCA-852 tubes. The line was constructed in the manner illustrated in Figure 3. For propagation survey purposes the transmitter has been adjusted successively for a range of frequencies from 25,700 kilocycles upwards. The antennas usually used have been horizontal dipoles at

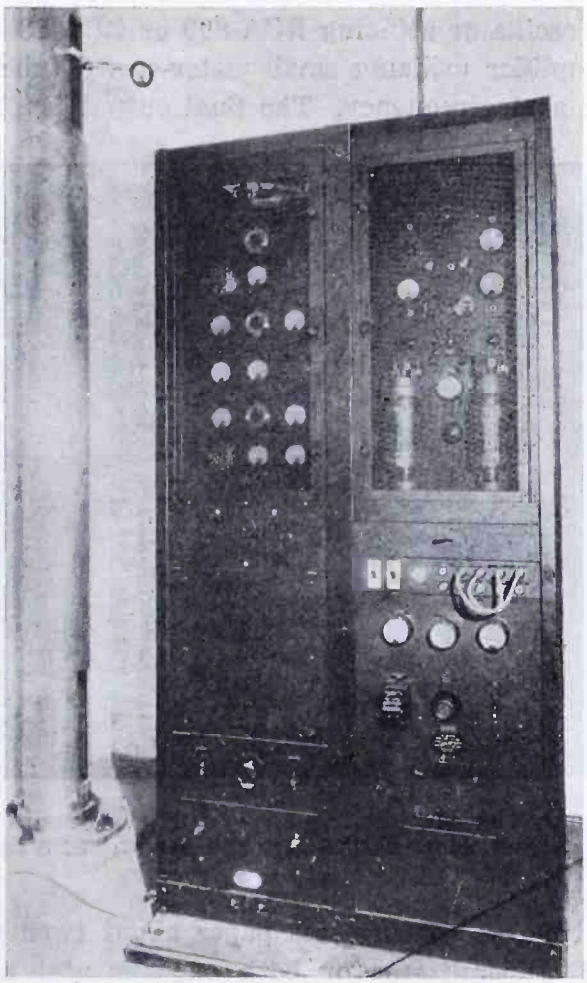


Fig. 12—Line controlled transmitter on roof of RCA Building, 30 Rockefeller Plaza, New York City, operated at 25.7 megacycles or higher.

the top of a mast which extends from the transmitter room to about twenty-eight feet above the roof. The base of the mast may be seen at the left of the transmitter. The transmission line to the antennas was run inside the mast.

Figure 13 is a photograph of W2XBN—91,800 kilocycles, 100 watts, located on the top floor of the Continental Bank Building at 30 Broad Street, New York City. This transmitter utilizes RCA-852 tubes and has a line controlled oscillator, similar to Figures 2 and 7, at 45,900 kilocycles followed by a frequency doubler and an amplifier

system. It provides a multiplex radio control circuit between the central traffic office of R.C.A. Communications, Inc., at 64 Broad Street, New York City, and the transoceanic station at New Brunswick, N. J. A directive antenna on the roof provides a power gain of about eight to one.

Figure 14 is a photograph of W2XS—200,000 kilocycles designed with a master oscillator utilizing RCA-800 or RCA-834 tubes followed by a power amplifier utilizing small water-cooled tubes. Both stages operate at the same frequency. The final output is about 250 watts.

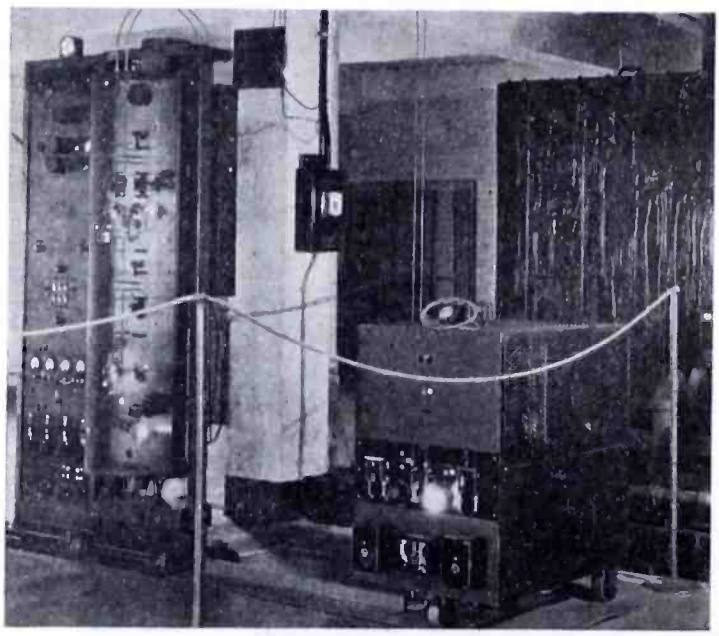


Fig. 13—Line controlled transmitter on roof of Continental Bank Building, 30 Broad Street, New York City, operated at 91.8 megacycles.

Up to the time of writing this paper about twenty-five or thirty line controlled transmitters for both experimental and commercial service have been built. At the lowest frequency, 6725 kilocycles, the power output was about thirty kilowatts. At the highest frequency, 450,000 kilocycles, the power output has ranged up to 110 watts.

Conclusion

Our experience during the past six years with line control of oscillator frequencies indicates that the method has great potential usefulness. It seems almost certain that it will provide a frequency stabilizing device for use at frequencies above about 20,000 kilocycles of as great practical value as piezoelectric crystals have been for use at lower frequencies.

APPENDIX

Radio-Frequency Resistance

In order to illustrate the principles involved in the phenomena of skin effect we shall first consider a very simple case. Assume a tubular line in which the inside conductor consists of two thin cubes a and b, differing in radius by da connected in parallel and in which the single tube outside conductor has the radius C. (See Figure 15.)

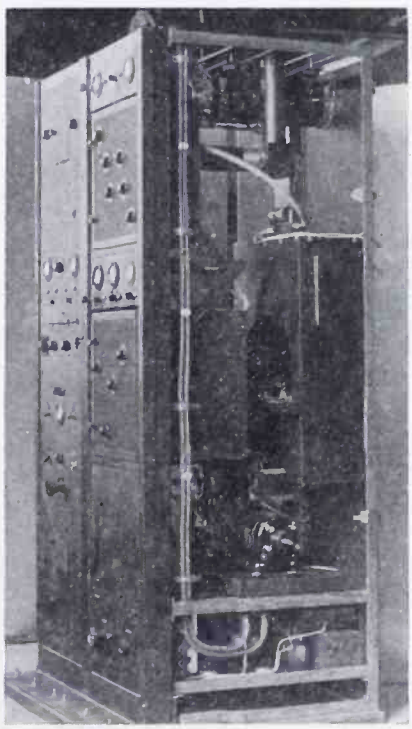


Fig. 14—Development model line controlled transmitter for operation at 200 megacycles.

Faraday's law states that the line integral of electric force around any closed circuit is equal to the negative rate of change of magnetic flux through the circuit, in accordance with the right-hand screw rule. Taking a path of unit length on conductors a and b in a plane cutting them longitudinally, we have $E_a - E_b = d/dt(\Phi_{ab})$. The magnetic force H between a and b is equal to $2i_a/a$. If the current is sinusoidal and represented by the real part of $I_a\epsilon^{j\omega t}$, $d/dt(i_a) = j\omega i_a$ and $E_a - E_b = -j\omega(2i_a/a)da$, as the area of a unit length is da. If R is the resistance of both a and b we have from Ohm's law:

$$E_a = i_a R \text{ and } E_b = i_b R.$$

Hence $R(i_a - i_b) = -i_a(2j\omega(da/a))$. $2da/a$ is the inductance per unit length of a circuit consisting of tubes a and b alone. Calling this L_{ab} we have:

$$R(i_a - i_b) = -j\omega L_{ab} i_a$$

or,

$$i_a/i_b = \frac{R}{R + j\omega L_{ab}}.$$

From this relation it is apparent that if either the resistance approaches zero or the frequency approaches infinity, the current will all flow in the tube of larger radius. Also it should be noted that the

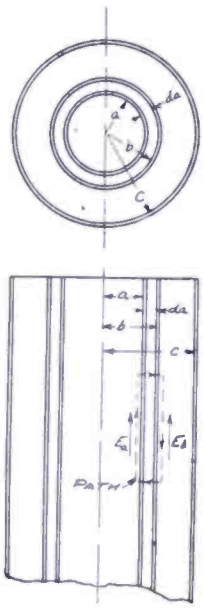


Fig. 15

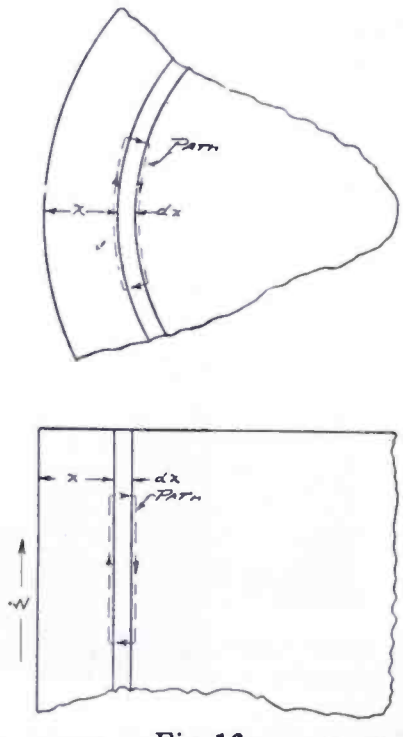


Fig. 16

radius C of the return conductor has no effect upon the distribution of the current between a and b.

Let us now consider the actual problem of current distribution within a solid conductor. For frequencies of the order of megacycles we know the penetration of current to be very small. If we assume this at the start so that we can neglect a change in resistance of successive thin cylinders as the radius is decreased, the problem of determination of effective resistance is greatly simplified. We shall proceed upon this assumption. Call the distance from the surface of the conductor X and the distance along the axis Z. (See Figure 16.)

Let u = the current density = the real part of $Ue^{j\omega t}$ so that $du/dt = j\omega u$.

Take a cylinder of thickness dx at a distance X from the surface and draw a circuit of unit peripheral length. The total current within this circuit is udx, the magnetic force at X is H and at $X + dx$ is $H + (\delta H/\delta x)dx$. The line integral around the circuit is then $H - (H + (\delta H/\delta x)dx) = -(\delta H/\delta x)dx$.

This by Ampere's law is equal to 4π times the total current included.

Hence $4\pi u dx = -(\delta H/\delta x)dx$ or $4\pi u = -\delta H/\delta x$.

Now take a circuit in a longitudinal plane. Since the current in the first case was assumed flowing into the paper the line integral of electric force around a circuit of unit length is

$$-E + \left(E + \frac{\delta E}{\delta x}dx\right) = \frac{\delta E}{\delta x}dx.$$

This by Faraday's law is equal to the negative rate of change of magnetic induction included by the circuit, or $\mu d/dt(Hdx)$.

Hence,

$$\frac{\delta E}{\delta x}dx = -\mu\frac{dH}{dt}dx \text{ where } \mu = \text{permeability}$$

or,

$$\frac{\delta E}{\delta x} = -\mu\frac{dH}{dt}.$$

Differentiating with respect to X and substituting for $\delta H/\delta x$ we get

$$\frac{\delta^2 E}{\delta x^2} = -\mu\frac{d}{dt}\left(\frac{\delta H}{\delta x}\right) = 4\pi\mu\frac{du}{dt} = 4\pi\mu j\omega u.$$

However, from Ohm's law $E = \rho u$ where ρ is the resistivity. Hence $\delta^2 u/\delta x^2 = j(4\pi\omega u/\rho)$. The solution of this differential equation is

$$u = u_0\varepsilon^{-(\alpha + j\alpha)x}$$

where,

$$\alpha = \sqrt{2\pi\omega\mu/\rho}.$$

The total current per unit peripheral length is then

$$i_0 = u_0\int_0^\infty \varepsilon^{-\alpha(1+j)x} = \frac{u_0}{\alpha(1+j)} = \frac{u_0}{\sqrt{2\alpha}}\varepsilon^{-j(\pi/4)}.$$

The average heat loss per unit area and unit length is $(U^2/2)\rho dx$ or $(U_0^2/2)(\varepsilon^{-\alpha x})^2 dx$.

The total heat loss per unit width is

$$(U_0^2\rho/2)\int_0^\infty \varepsilon^{-2\alpha x}dx = U_0^2\rho/4\alpha = \rho(i_0^2/\alpha).$$

Hence the effective resistance $= \rho\sqrt{2\pi\mu\omega/\rho} = \sqrt{2\pi\mu\rho\omega}$ per unit width.

For a radius r the resistance is $R = \sqrt{2\pi\mu\rho\pi}/2\pi r = \sqrt{\mu f\rho}/r$ electromagnetic units.

In practical units $R = \sqrt{\mu f\rho/r} \times 10^{-9}$ ohms per centimeter length for r in centimeters.

For copper $\rho = 1724$ electromagnetic units, $\mu = 1$ and $R = (41.5/r)\sqrt{f} \times 10^{-9}$ ohms per centimeter of length.

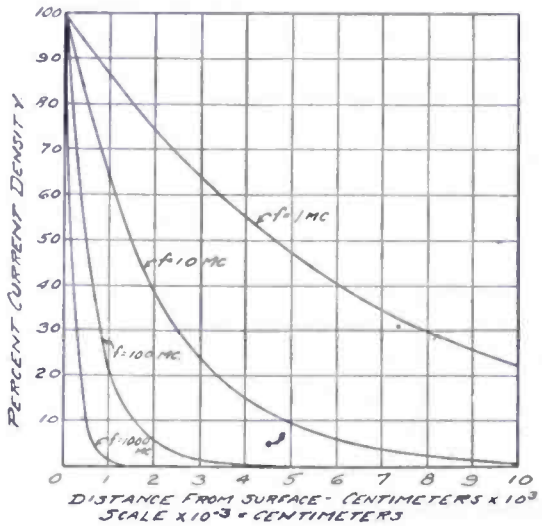


Fig. 17

The relation for current distribution shows that the current density falls off exponentially as we proceed away from the surface into the wire and the phase shifts through an angle of ninety degrees in proceeding from the conductor surface to an internal position where the value of the current is negligible. The phase angle of the total current lags that of the surface current by forty-five degrees. We may consider this phase-shift effect to be due to an internal self-inductance. It may be of interest to note the manner in which the current distributes itself near the surface of a conductor at different frequencies. This is shown for several frequencies in Figure 17. It may be noted that for a frequency of ten megacycles the current density becomes ten per cent of its value at the surface at a distance of five-thousandths of a centimeter. It is apparent from the foregoing reasoning that the same relations hold for both the internal and ex-

ternal conductors of a concentric tube line. Hence for the total resistance of a unit length of concentric copper tube line we have $R = 41.5 \times 10^{-9} \sqrt{f}(1/a + 1/b)$ ohms per centimeter where a and b are the radii of the tubes in centimeters.

Power Factor

A tuned transmission line short-circuited at the far end an odd number of quarter waves long has a distribution of voltage given approximately by $E = E_0 \cos 2\pi x/\lambda$ and a current distribution given by $i = (E_0/Z_0) \sin 2\pi x/\lambda$ where X is the distance, Z_0 the characteristic impedance $= \sqrt{L/C}$, and E_0 the input voltage. For open-circuited lines any integral number of half waves long the distribution of voltage and current is given by $E = I_0 Z_0 \sin 2\pi x/\lambda$ and $I = I_0 \cos 2\pi x/\lambda$.

The power dissipated in a section of length dx is $i^2 R dx$ where R is the resistance per centimeter.

The total power is then $W = \int_0^l I_0^2 R \sin^2(2\pi x/\lambda)\, dx = (1/2) I_0^2 R l.$

Hence the effective resistance is $R/2$. The displacement current in a length dx is $E\omega c dx$ and the voltamperes $(VA) = E^2 \omega c dx$ where c is the capacity per centimeter of length.

The total volt-amperage is then

$$VA = \int_0^l \omega c E_0^2 \sin \frac{2\pi x}{\lambda}\, dx = \frac{\omega c E_0^2 l}{2} = \frac{\pi E_0^2 l}{z_0 \lambda} = \frac{I_0 l}{\omega c}.$$

For all practical purposes the effective power factor $= \dfrac{\text{watts}}{VA}$

$$= \frac{\frac{1}{2} I_0^2 R l}{\frac{1}{2} E_0^2 \omega c l} = \frac{R}{Z_0^2 \omega c}.$$

But $Z_0 = 60 \log_\epsilon b/a$

and $C = \dfrac{1}{2 \log_\epsilon b/a} \times \dfrac{1}{9} \times 10^{-11}$ farads per centimeter

and the power factor $= PF = \dfrac{R}{2 \times 10^{-9} \omega \log_\epsilon b/a}$ where R is in ohms per centimeter length but $R = \sqrt{\rho f}(1/a + 1/b) \times 10^{-9}$ assuming non-magnetic conductors.

$$\text{Hence } PF = \left(\frac{1}{a} + \frac{1}{b}\right) \frac{\sqrt{f\rho}}{2\omega \log_\epsilon b/a}$$

and for copper $PF = \frac{3.3}{\sqrt{f}}\left(\frac{1}{a}+\frac{1}{b}\right)\frac{1}{\log_\epsilon b/a}$ for a and b in centimeters.

The power factor may be expressed in terms of the attenuation constant in which it becomes $PF = \alpha\lambda/\pi$ where α is taken for the same unit of length as λ.

When b/a is made the best ratio, 3.6, for minimum loss, the power factor becomes $PF = \frac{1}{4\pi a}\sqrt{\frac{\mu\rho}{f}}$ or $Q = \frac{1}{PF} = 4\pi a\sqrt{\frac{f}{\mu\rho}}$ and for copper $PF = 1.905 \times 10^{-5}\frac{\sqrt{\lambda}}{a}$, for a and λ in centimeters, and

$$Q = 5.25 \times 10^{-4}\frac{a}{\sqrt{\lambda}}.$$

Voltage Gradient

The maximum electric force or voltage gradient E for a concentric tube line is given by $E_a = \frac{V}{\log_\epsilon b/a}\times\frac{1}{a}$ volts per centimeter where V is the voltage. For a given voltage and given outside tube diameter this becomes a minimum when $b/a = \epsilon \cong 2.72$. However, this is not usually of interest. The important consideration is that of obtaining a maximum of volt-amperage for a given gradient. In terms of volt-amperage the gradient E becomes

$$E = \frac{60}{\pi l/\lambda}/a(\log_\epsilon b/a)^{1/2}.$$

Minimizing this expression we find $b/a = \epsilon^{1/2} \cong 1.65$.

For this ratio the gradient becomes

$$E = \frac{10.2}{b}\sqrt{\frac{VA}{l/\lambda}} \text{ volts per centimeter.}$$

When the line is designed for a minimum loss ($b/a = 3.6$) the gradient is thirty-six per cent greater than the value given above.

Input Impedance

The input impedance Z_0 of a quarter-wave line short-circuited at its far end or of a half-wave line open at its far end is

$$Z_i = \frac{Z_0}{\tanh \alpha l} \cong \frac{Z_0}{\alpha l}$$

where α is the attenuation factor.

However, $\alpha = R/2Z_0$ and therefore

$$Z_i = \frac{2Z_0^2}{Rl}$$

which in terms of b and a becomes

$$Z_i = \frac{2.60^2 \ (\log_\epsilon b/a)^2}{41.5 \times 10^{-9} \sqrt{f} \ (1/a + 1/b) \ l}$$

Maximizing this expression we find Z_i to be a maximum when $b/a = \epsilon^{2\,(1+a/b)}$ or $b/a = 9.18$.

When the ratio is 9.18 the input impedance becomes

$$Z_i = \frac{8.4 \times 10^{10}}{\sqrt{fl}} \times b \text{ ohms, for } b \text{ and } l \text{ in centimeters}$$

$$= 11.2b \sqrt{f} \quad \text{for a one-quarter-wave-length line (closed)}$$

or $$5.6b \sqrt{f} \quad \text{for a one-half-wave-length line (open).}$$

and for copper $PF = \dfrac{3.3}{\sqrt{f}}\left(\dfrac{1}{a}+\dfrac{1}{b}\right)\dfrac{1}{\log_\epsilon b/a}$ for a and b in centimeters.

The power factor may be expressed in terms of the attenuation constant in which it becomes $PF = \alpha\lambda/\pi$ where α is taken for the same unit of length as λ.

When b/a is made the best ratio, 3.6, for minimum loss, the power factor becomes $PF = \dfrac{1}{4\pi a}\sqrt{\dfrac{\mu\rho}{f}}$ or $Q = \dfrac{1}{PF} = 4\pi a\sqrt{\dfrac{f}{\mu\rho}}$ and for copper $PF = 1.905 \times 10^{-5}\dfrac{\sqrt{\lambda}}{a}$, for a and λ in centimeters, and

$$Q = 5.25 \times 10^{-4}\frac{a}{\sqrt{\lambda}}.$$

Voltage Gradient

The maximum electric force or voltage gradient E for a concentric tube line is given by $E_a = \dfrac{V}{\log_\epsilon b/a} \times \dfrac{1}{a}$ volts per centimeter where V is the voltage. For a given voltage and given outside tube diameter this becomes a minimum when $b/a = \varepsilon \cong 2.72$. However, this is not usually of interest. The important consideration is that of obtaining a maximum of volt-amperage for a given gradient. In terms of volt-amperage the gradient E becomes

$$E = \frac{60}{\pi l/\lambda}/a\,(\log_\epsilon b/a)^{1/2}.$$

Minimizing this expression we find $b/a = \varepsilon^{1/2} \cong 1.65$.

For this ratio the gradient becomes

$$E = \frac{10.2}{b}\sqrt{\frac{VA}{l/\lambda}} \text{ volts per centimeter.}$$

When the line is designed for a minimum loss ($b/a = 3.6$) the gradient is thirty-six per cent greater than the value given above.

Input Impedance

The input impedance Z_0 of a quarter-wave line short-circuited at its far end or of a half-wave line open at its far end is

$$Z_i = \frac{Z_0}{\tanh \alpha l} \cong \frac{Z_0}{\alpha l}$$

where α is the attenuation factor.

However, $\alpha = R/2Z_0$ and therefore

$$Z_i = \frac{2Z_0^2}{Rl}$$

which in terms of b and a becomes

$$Z_i = \frac{2.60^2 \ (\log_\epsilon b/a)^2}{41.5 \times 10^{-9} \sqrt{f} \ (1/a + 1/b) \ l}$$

Maximizing this expression we find Z_i to be a maximum when $b/a = \epsilon^{2\,(1+a/b)}$ or $b/a = 9.18$.

When the ratio is 9.18 the input impedance becomes

$$Z_i = \frac{8.4 \times 10^{10}}{\sqrt{fl}} \times b \text{ ohms, for } b \text{ and } l \text{ in centimeters}$$

$$= 11.2b \sqrt{f} \quad \text{for a one-quarter-wave-length line (closed)}$$

or $$5.6b \sqrt{f} \quad \text{for a one-half-wave-length line (open).}$$

A CATHODE-RAY FREQUENCY-MODULATION GENERATOR

By

R. E. SHELBY

Engineering Staff, National Broadcasting Company, New York

Summary—A new device is described for phase or frequency modulating a carrier wave obtained from a source of constant frequency and phase. It consists essentially of an electron gun, a target anode of special design and a means for deflecting the cathode-ray beam. Deflection of the beam is by means of two carrier voltages derived from a common source, but differing in phase by 90° and applied in such a way that the electron stream traces a circular path on the target anode. When amplitude modulation is applied to the deflection voltages, a phase-modulated signal is obtained in the circuit associated with the target anode. Frequency modulation may be obtained by proper pre-distortion of the audio-frequency signal prior to modulation.

A U-H-F radio transmitter set-up used in testing such a device is described.

THE device to be described is proposed mainly as a means of frequency or phase modulating a carrier signal obtained from an oscillator of constant frequency and phase, such as a crystal-controlled oscillator. However, various embodiments of the underlying principle may be used for other purposes.

The majority of the known methods for generating frequency or phase modulation involve variation of the frequency or phase of the master oscillator. There is one well-known method of producing phase modulation of a carrier obtained from a constant source which involves the addition of two voltages of the same constant frequency having a constant phase difference of 90°, one voltage having constant amplitude and the other varying in amplitude with the modulation signal.

This method is limited to maximum phase shifts of the order of 30°, which means that in order to obtain a large shift at the final carrier frequency, it is necessary to employ frequency multiplication of large ratio. The method to be proposed here does not have this limitation. Theoretically it should give distortionless phase shift of many times 360°.

Structurally the device consists essentially of an electron gun, two sets of electrostatic deflecting plates, and a target anode of special design, enclosed in an evacuated container of suitable size and shape. A typical arrangement is shown schematically in Figure 1. It will be seen that the device illustrated is similar to a conventional cathode-ray oscillograph tube. The only fundamental difference is the novel design

Reprinted from *Electronics*, Feb. 1940.

of the target anode. Figure 2 illustrates one form which the target anode may have. It consists of two (or more) metallic plates with curved edges, upon which the electrons from the electron gun impinge. In order to obtain phase modulation in which the angular shift is a linear function of the modulation amplitude the edges of the plates have a curvature given by the polar equation

$$r = a\,\Theta \qquad (1)$$

which defines an Archimedian spiral.

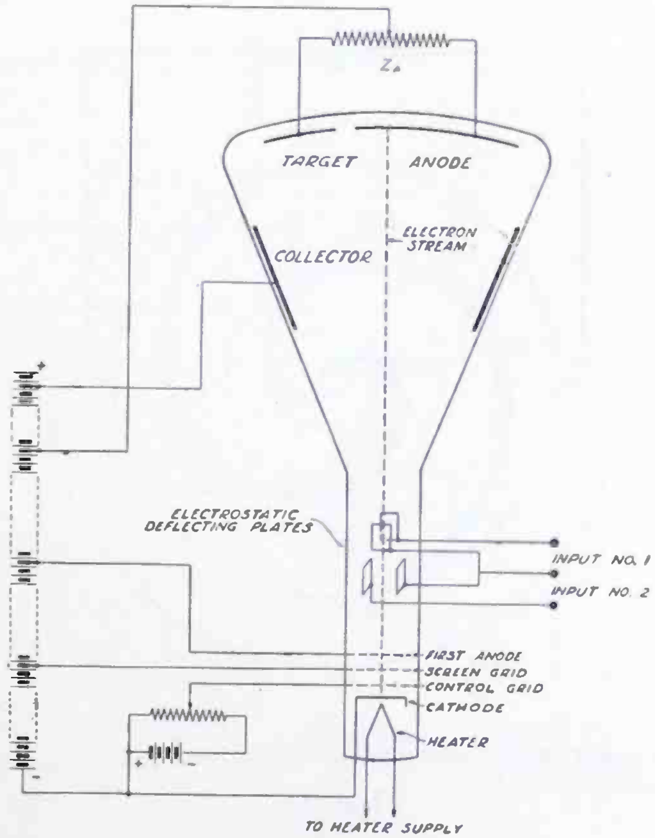
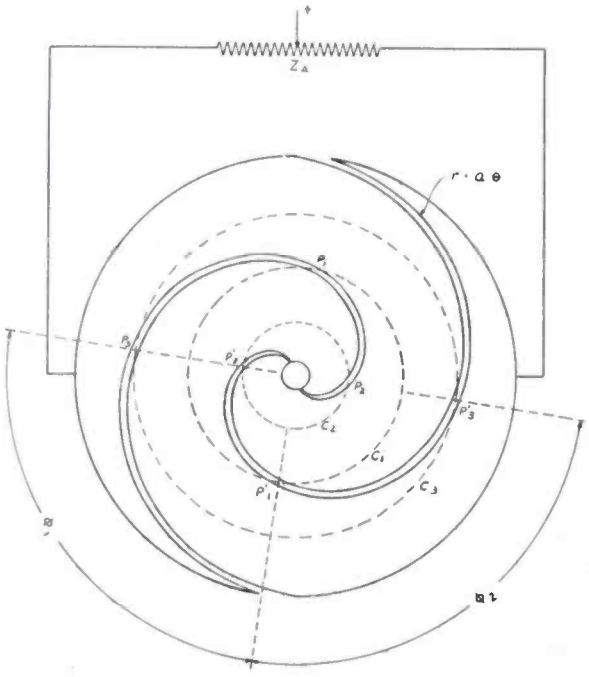


Fig. 1.

In operation the electron stream is deflected in such a way that it traces out a circle on the target anode, the diameter of the circle being a variable which is directly proportional to the instantaneous value of the modulation signal. This involves no new concepts for anyone familiar with cathode-ray oscillographs. When controlled in this way the electron stream produces a phase (or frequency) modulated wave upon striking an anode having the special design just described. Operation of the system will now be described more fully.

The electron gun is controlled and focused by adjusting the d-c potentials applied to the cathode, control grid, screen grid, and first

anode, just as in the case of oscillograph tubes and kinescopes. The electron stream is sharply focused on the target anode and when there is no voltage applied to the electrostatic deflecting plates it is adjusted to strike the exact geometrical center of the target anode. The inputs, which consist of two carrier waves of equal frequency and almost equal amplitude, but differing in phase by 90° and having the same amplitude modulation, are applied to the two sets of electrostatic deflecting plates. Figure 3 illustrates one means of supplying the inputs. The output appears in the target anode circuit, across the impedance Z_A.



NOTE: DOTTED LINES ARE NOT STRUCTURAL LINES

Fig. 2.

If Z_A is a pure resistance the output voltage appearing across it will be a flat-topped wave. By using a target anode composed of a larger number of curved sections the fundamental frequency of the flat-topped output wave may be made any desired multiple of the input frequency. Thus it is seen that, if desired, frequency multiplication may be obtained during the process of converting amplitude modulation into frequency or phase modulation.

For a more detailed explanation of the way in which amplitude modulation is translated into phase modulation, we refer now to Figure 4. First with no modulation the phase-shifting network and amplitude controls are adjusted so that the electron beam describes a circle on

the target anode. If the master amplitude control is adjusted so that this circle is of the size designated by c_1 (Figure 2) then the voltage appearing across Z_A will be as shown in Figure 4(a). Note that the electron stream passes from one segment of the target to the other at points p_1 and p'_1. If now the two deflecting voltages are decreased 50 per cent by changing the master amplitude control, all other controls being left the same, the locus of the end of the electron stream will be c_2 and the voltage across Z_A will be as shown in Figure 4(b). The electron stream now passes from one segment of the target to the other

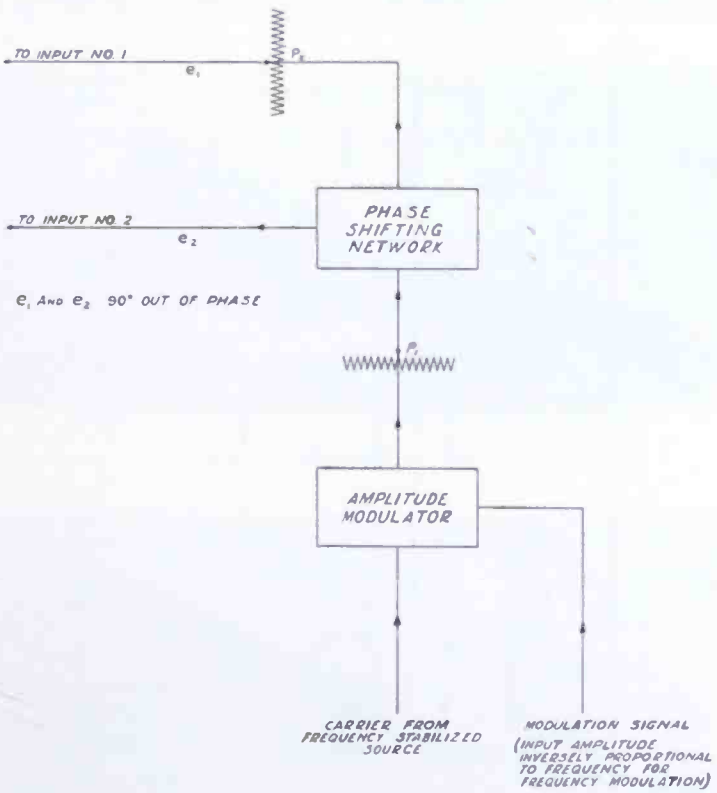


Fig. 3.

at points p_2 and p_2'. Likewise, if the master amplitude control is adjusted to give voltages 50 per cent greater than those which gave the locus c_1 then c_2 will be the new locus, the electron stream will pass from one target segment to the other at points p_3 and p_3' and the voltage appearing across Z_A will be as shown in Figure 4(c). Now if the master amplitude control is reset so that the electron circle falls on c_1 and a fifty per cent modulation is then applied to the carrier in the manner indicated by Figure 3 the locus of the end-point of the electron stream will expand and contract between the limits c_2 and c_3 and the output wave will shift in phase between the limits indicated by Figure 4(b) and (c).

The amount of maximum phase shift in the device described above will be determined by the curvature of the target anode boundary—that is, it will depend upon the value of a in the equation $r = a\,\Theta$. For the target illustrated in Figure 2 the phase shift is plus and minus approximately 90° when the input is amplitude modulated 50 per cent. Amplitude modulation of 75 per cent on the input will give shift of plus and minus 135°, etc.

It should be noted that this device is fundamentally a generator of phase modulation—not frequency modulation. However, it is well known that any phase modulator may be made to produce the equiva-

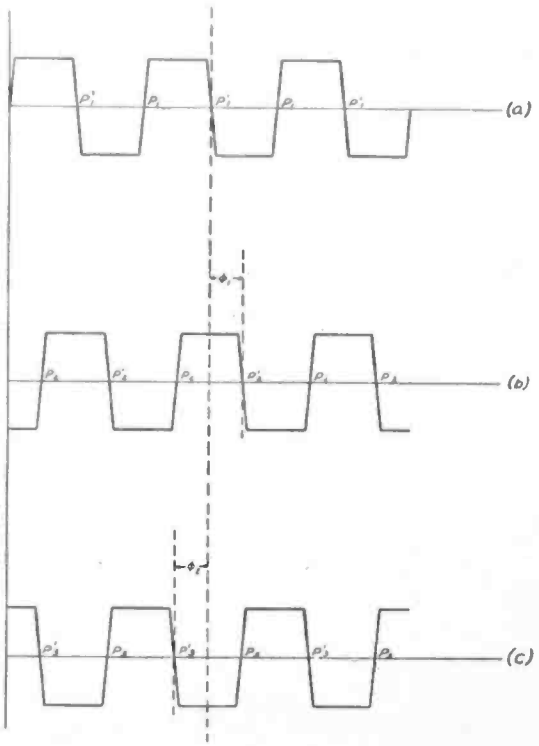


Fig. 4.

lent of frequency modulation by means of a network in the audio input having a characteristic that is inversely proportional to frequency. It will be understood that whenever the cathode-ray modulator is referred to as a generator of frequency modulation, the use of such a network is implied.

Electromagnetic deflection of the electron stream may be utilized instead of electrostatic deflection as described above.

Variation of the anode voltages will cause the sensitivity of the electron stream to vary so that for constant voltages on the deflecting plates the size of the circle described on the final anode by the electron stream will vary as the anode voltages vary, thereby producing a form

of phase modulation. This means that the d-c voltages supplied to the electron gun must be well-filtered and free from fluctuations.

Auxiliary electrodes may be added to the modulator tube for control or monitoring purposes. For example a fluorescent screen may be provided beyond the target anode, so that the electron stream will produce a pattern upon it when it passes between segments of the anode or beyond the outer edges of the anode plates. Additional electrodes, located in the same plane as the final anode, but electrically separate from it, may be used for adjusting the modulator and also for indicating overmodulation. Such an electrode, of small area, located

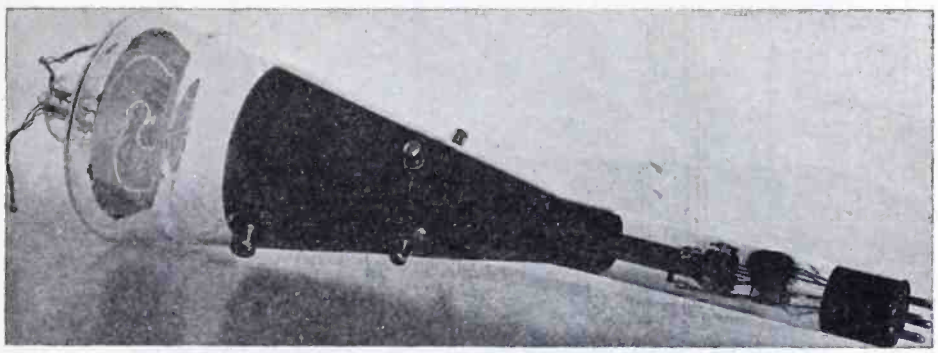


Fig. 5.

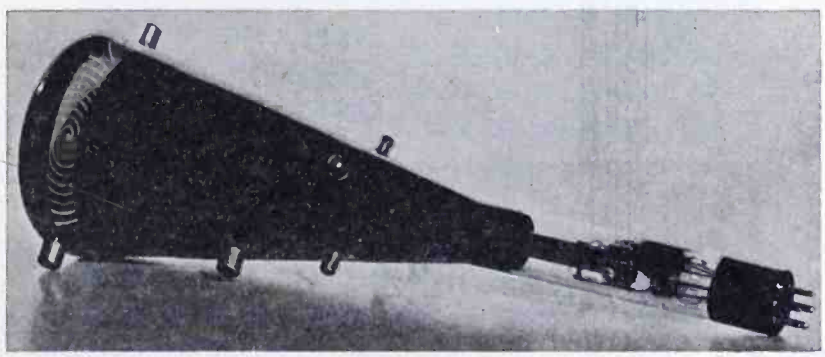


Fig. 6.

at the geometrical center of the anode (where the anode plates are cut way) is useful in centering the electron stream. A narrow annular ring around the outside of the main target facilitates adjustment of the phase-shifting network to obtain circular deflection of the electron stream. Many other auxiliaries are possible.

The description so far has related to the production of modulation in which the phase shift is directly proportional to the modulating voltage, but the system is quite flexible in this respect. By use of a properly shaped target, the phase shift may be made any reasonable

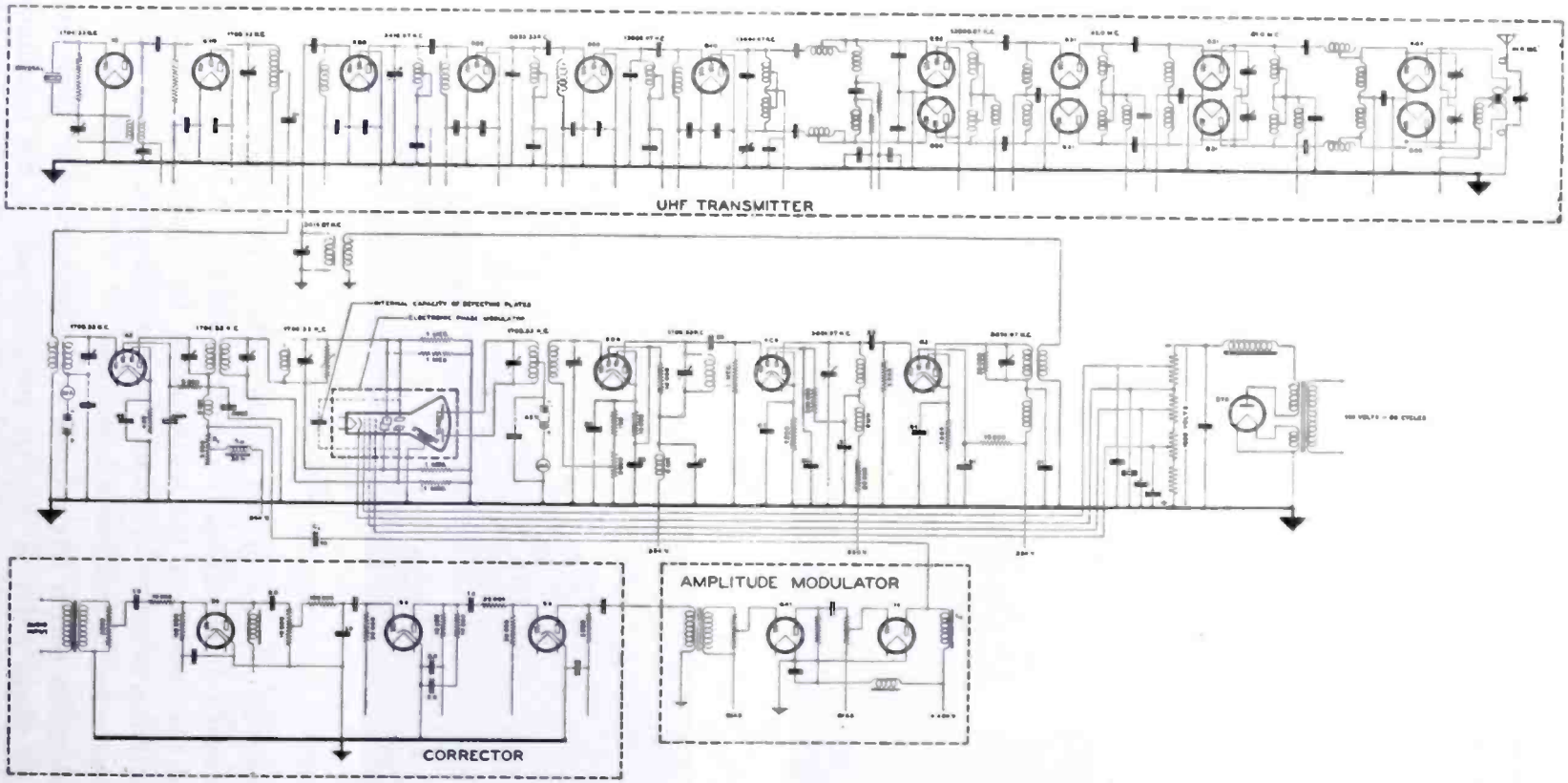


Fig. 7—Schematic circuit of the transmitter set-up employed to test the frequency modulator tube.

function of the modulating voltage, this being determined by the curvature of the edges of the target plates and the nature of the path traced on the target by the electron stream.

The amplitude of the flat-topped output voltage wave may be varied independently of the frequency (or phase) modulation by varying the d-c potential applied to the control grid. If it is desired to amplitude modulate the output wave in addition to or instead of phase modulating it in the manner described, this may be done readily by applying the amplitude modulation signal to the control grid, provided the electron

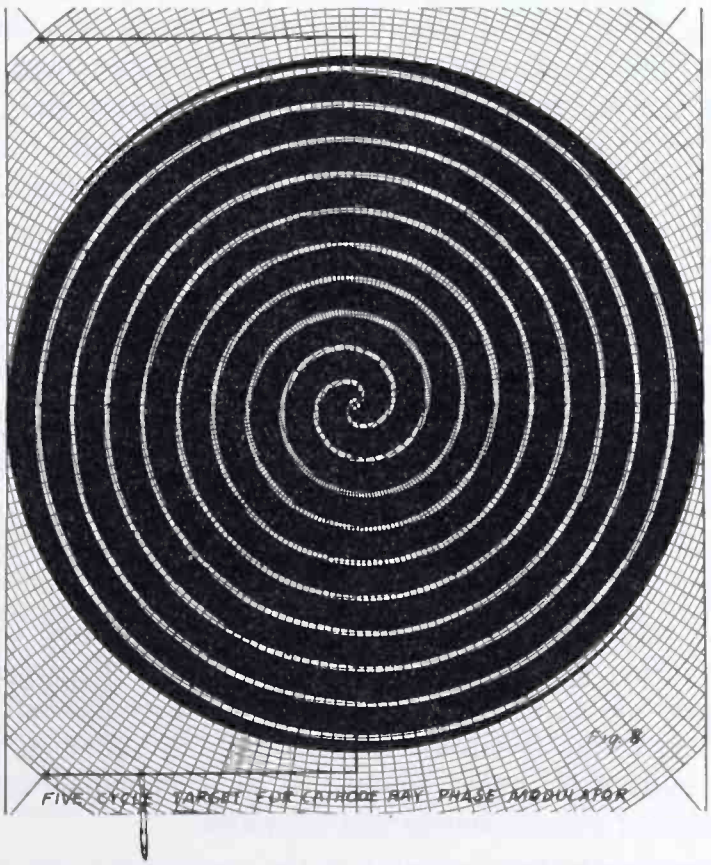


Fig. 8.

gun is so operated that the rate of electron emission from the gun is a linear function of control grid voltage over the operating range.

The photograph of Figure 5 shows the general appearance of an early model of this device constructed by the Radiotron Division of the RCA Manufacturing Co. It consists of a standard type cathode-ray oscilloscope tube with a spiral target of the type specified above sealed into the large end in place of the usual fluorescent screen. The target was constructed by applying a coating of platinum to a sheet of mica of the kind used for Iconoscope mosaics and scribing the separating

line in the platinum surface to give the proper shape to the two electrodes. The target is mounted by wire supports which are sealed into the large end of the tube. Electrical connections to the two electrodes of the target are brought out through this same seal.

The spiral of this target was designed to give a maximum over-all phase shift of 360°—that is to say, for this particular target the value of a in the equation $r = a\,\Theta$ is $\frac{1}{\pi}$ inches per radian (the diameter of the target is approximately 4 inches). In order to facilitate adjustment and to judge operating performance 4 concentric circles and 16 radial lines were marked on the surface of the target with willemite. The entire surface of the target was also covered with a very thin coating of willemite. This results in the electron beam tracing out a visible path on the target, which is helpful in testing the tube.

In constructing this tube the Radiotron engineers made provision for utilizing secondary emission from the target anode to provide additional output. This was done by providing a separate collector for secondary electrons in the form of a coating on the inner wall of the tube.

Figure 6 shows a photograph of a later model. The general design is the same except that the target anode consists of a conducting coating deposited directly upon the inner surface of the large end of the tube and the spiral in this case contains five complete cycles. The configuration of this target is shown by Figure 8.

Figure 7 shows a schematic of the transmitter set-up employed to test the frequency modulator tube. That portion of the circuit enclosed by the upper dotted rectangle represents a conventional r-f transmitter. It was of the crystal-controlled type with buffer amplifiers and frequency multiplier stages driving a final amplifier delivering a carrier output of approximately two kilowatts.

The crystal-controlled oscillator in the transmitter was used as the primary source of stabilized r-f voltage, the amplifier chain in the transmitter being broken between the buffer amplifier and the first doubler stages for insertion of the cathode-ray phase modulator and associated apparatus, as shown.

To obtain the 90° phase relationship between the voltages on the two sets of deflecting plates of the cathode-ray tube, a phase-splitting network consisting of a capacitive reactance in series with a pure resistance was employed, one set of deflecting plates being connected across the capacitive reactance and the other set across the resistance. In this set-up the internal capacity between one pair of the deflecting

plates was utilized as the capacitive reactance. A parallel tuned circuit was bridged across the other set of deflecting plates to permit tuning out the reactance of the variable resistor and also that due to the capacity between these deflecting plates, thus obtaining a purely resistive impedance. The variable resistor in this circuit was used to adjust the relative amplitudes of the voltages on the two sets of deflecting plates, and the amplitudes of the two deflecting voltages were adjusted simultaneously by means of the variable resistor R_1 in series with the d-c plate supply to the modulated amplifier.

The output circuit connected to the two halves of the target anode was tuned to the fundamental crystal frequency and coupled inductively to the grid of an amplifier which in turn fed a frequency doubler. The succeeding stage fed a shielded r-f transmission line, the other end of which was coupled to the 860 stage in the transmitter which ordinarily operated as a doubler, but which was used as a straight amplifier in this case. From that point on the various stages of the transmitter were operated in their usual manner.

The transmitter was located in the Empire State Building and reception tests were made with a frequency modulation receiver located in the development laboratory in Radio City. These tests showed that the cathode-ray frequency modulator performed as predicted. When properly shielded against stray magnetic fields and provided with well-filtered d-c potentials, it introduced no measurable distortion and was perfectly stable in its operation.

In conclusion acknowledgment is made of the cooperation of the RCA Radiotron Laboratories in constructing and supplying the tubes used in these tests and of the helpful suggestions offered by Radiotron and NBC engineers.

The author also wishes to acknowledge with deep appreciation the encouragement and cooperation extended by Professor E. H. Armstrong of Columbia University during the early tests on this device.

CARRIER AND SIDE-FREQUENCY RELATIONS WITH MULTI-TONE FREQUENCY OR PHASE MODULATION

By

MURRAY G. CROSBY

R.C.A. Communications, Inc., Riverhead, N. Y.

Summary—*The equation for the carrier and side frequencies of a frequency or phase-modulated wave is resolved for the case of two applied modulating tones. It is shown that when more than one modulating tone is applied, the amplitude of the carrier is proportional to the product of the zero-order Bessel Functions of all of the modulation indexes involved. The amplitudes of the side frequencies are proportional to the products of Bessel Functions equal in number to the number of tones applied and having orders respectively equal to the orders of the frequencies involved in the side frequency. Beat side frequencies are produced which have higher-order amplitude and do not appreciably widen the band width occupied by frequency modulation.*

PREVIOUS analyses of the carrier and side-frequency relations in frequency and phase modulation have considered either the case of a single tone as applied modulation,[1,2,3] or the case of telegraph transmission[2] where the Fourier resolution shows a tone and its harmonics to be present. It is the purpose of this paper to develop the equations for the carrier and side frequencies so that their characteristics will be known when more than one modulating tone is applied.

In amplitude modulation the addition of a second modulating tone merely produces an additional pair of side frequencies which are displaced from the carrier frequency by the frequency of the tone. These side frequencies have amplitudes which are linearly proportional to the carrier amplitude and the depth of modulation at which the tone is applied (assuming a linear modulator). However, in frequency and phase modulation this linear relation between the depth of modulation and the amplitudes of the side frequencies does not exist and a separate development is required for the multi-tone case.

---

[1] John R. Carson, "Notes on the Theory of Modulation," *Proc. I.R.E.*, Vol. 10, pp. 57-64, Feb. 1922.

[2] Balth. Van der Pol, "Frequency Modulation," *Proc. I.R.E.*, Vol. 18, pp. 1194-1205, July 1930.

[3] Hans Roder, "Amplitude, Phase, and Frequency Modulation," *Proc. I.R.E.*, Vol. 19, pp. 2145-2176, Dec. 1931.

Reprinted from *RCA Review*, July, 1938.

In the case of frequency modulation the instantaneous frequency for the two-tone case is given by:

$$f = F_c + F_{dp} \sin pt + F_{dq} \sin qt \tag{1}$$

where F_c = the carrier frequency, F_{dp} = frequency deviation applied by the tone having angular velocity p and F_{dq} = frequency deviation applied by the tone having angular velocity q.

Since the instantaneous frequency is the rate of change of phase, (1) must be integrated to find the phase angle.

$$\omega t = 2\pi \int_0^t (F_c + F_{dp} \sin pt + F_{dq} \sin qt)\, dt$$

$$= \omega_c t - \frac{F_{dp}}{F_{mp}} \cos pt - \frac{F_{dq}}{F_{mq}} \sin qt \tag{2}$$

where F_{mp} and F_{mq} = the modulation frequencies of the tones and $\omega_c = 2\pi F_c$.

If, for simplification $F_{dp}/F_{mp} = P$ and $F_{dq}/F_{mq} = Q$, the frequency modulated wave is:

$$e = E \sin (\omega_c t - P \cos pt - Q \cos qt) \tag{3}$$

Equation (3) could also represent a phase-modulated wave in which $P = \phi_p$ and $Q = \phi_q$, where ϕ_p and ϕ_q are the phase deviations applied by the tones.

By applying the addition formula for the sine to Equation (3), the following is obtained.

$$\begin{aligned} e = E \sin \omega_c t\, [&\cos (P \cos pt) \cos (Q \cos qt) - \sin (P \cos pt) \\ &\sin (Q \cos qt)\,] - \cos \omega_c t\, [\sin (P \cos pt) \cos (Q \cos qt) \\ &+ \cos (P \cos pt) \sin (Q \cos qt)\,] \end{aligned} \tag{4}$$

Substituting the following Bessel Function expansions in Equation (4),

$$\cos (x \cos \phi) = J_0(x) - 2J_2(x) \cos 2\phi + 2J_4(x) \cos 4\phi - \ldots\ldots \tag{5}$$

$$\sin (x \cos \phi) = 2J_1(x) \cos \phi - 2J_3(x) \cos 3\phi + 2J_5(x) \cos 5\phi - \ldots \tag{6}$$

obtaining the products of the series, applying the addition formulas for the sine and cosine, and rearranging gives:

$$\begin{aligned} e = E\, [&J_0(P)\, J_0(Q) \sin \omega_c t - J_1(P)\, J_0(Q) \cos (\omega \pm p)\, t \\ &- J_2(P)\, J_0(Q) \sin (\omega \pm 2p)\, t \end{aligned}$$

$$
\begin{aligned}
&+ J_3(P)\, J_0(Q) \cos(\omega \pm 3p)\, t \\
&+ J_4(P)\, J_0(Q) \sin(\omega \pm 4p)\, t \\
&- \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&- J_0(P)\, J_1(Q) \cos(\omega \pm q)\, t \\
&- J_0(P)\, J_2(Q) \sin(\omega \pm 2q)\, t \\
&+ J_0(P)\, J_3(Q) \cos(\omega \pm 3q)\, t \\
&+ J_0(P)\, J_4(Q) \sin(\omega \pm 4q)\, t \\
&- \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&- J_1(P)\, J_1(Q) [\sin(\omega + p \pm q)\, t + \sin(\omega - p \pm q)\, t] \\
&+ J_1(P)\, J_2(Q) [\sin(\omega + p \pm 2q)\, t + \sin(\omega - p \pm 2q)\, t] \\
&+ J_1(P)\, J_3(Q) [\sin(\omega + p \pm 3q)\, t + \sin(\omega - p \pm 3q)\, t] \\
&- \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&+ J_2(P)\, J_1(Q) [\sin(\omega + 2p \pm q)\, t + \sin(\omega - 2p \pm q)\, t] \\
&+ J_2(P)\, J_2(Q) [\sin(\omega + 2p \pm 2q)\, t + \sin(\omega - 2p \pm 2q)\, t] \\
&- J_2(P)\, J_3(Q) [\sin(\omega + 2p \pm 3q)\, t + \sin(\omega - 2p \pm 3q)\, t] \\
&- \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&+ J_3(P)\, J_1(Q) [\sin(\omega + 3p \pm q)\, t + \sin(\omega - 3p \pm q)\, t] \\
&- J_3(P)\, J_2(Q) [\sin(\omega + 3p \pm 2q)\, t + \sin(\omega - 3p \pm 2q)\, t] \\
&- J_3(P)\, J_3(Q) [\sin(\omega + 3p \pm 3q)\, t + \sin(\omega - 3p \pm 3q)\, t] \\
&+ J_3(P)\, J_4(Q) [\sin(\omega + 3p \pm 4q)\, t + \sin(\omega - 3p \pm 3q)\, t] \\
&+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&- J_4(P)\, J_1(Q) [\sin(\omega + 4p \pm q)\, t + \sin(\omega - 4p \pm q)\, t] \\
&- J_4(P)\, J_2(Q) [\sin(\omega + 4p \pm 2q)\, t + \sin(\omega - 4p \pm 2q)\, t] \\
&+ J_4(P)\, J_3(Q) [\sin(\omega + 4p \pm 3q)\, t + \sin(\omega - 4p \pm 3q)\, t] \\
&+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&- J_5(P)\, J_1(Q) [\sin(\omega + 5p \pm q)\, t + \sin(\omega - 5p \pm q)\, t] \\
&+ J_5(P)\, J_2(Q) [\sin(\omega + 5p \pm 2q)\, t + \sin(\omega - 5p \pm 2q)\, t] \\
&+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}
\tag{7}
$$

From (7) it can be seen that the amplitude of the carrier is proportional to the product of the zero-order Bessel Functions of the two quantities P and Q which would be the modulation index (F_{dp}/F_{mp} or F_{dq}/F_{mq}) for frequency modulation, or the deviation in radians (ϕ_p or ϕ_q) for phase modulation. The side frequencies present are not only those which would be present with one tone on at a time, but include beat side frequencies which are produced in a manner similar to the way in which beat frequencies are produced in a detector. The amplitudes of these beat side frequencies are equal to the products of Bessel Functions having orders equal to orders of the tone frequencies involved in the side frequency.

A similar resolution for the three-tone case shows that the carrier is proportional to the product $J_0(P)\ J_0(Q)\ J_0(R)$ where R is the modulation index of the third modulating tone. The beat side frequencies are produced in the same manner with more possible combinations, but with their amplitudes proportional to three Bessel Functions the

orders of which depend upon the orders of the tone frequencies involved in the side frequency. Thus it may be assumed that as the number of tones increases, the number of Bessel Functions in the amplitude coefficients increase in the same accordance while the number of side frequencies increases at a very high rate due to the great number of beat combinations which are possible.

Since side frequencies are produced which are displaced from the carrier by the sum of all the tone frequencies involved, it might be assumed that the band width occupied by frequency modulation is rather alarming. However, investigation of the amplitudes of these beat side frequencies shows that their amplitudes are rather small as far as out-of-band interference is concerned and that the major portion of the side bands remain in a channel approximately equal to twice the frequency deviation or the maximum modulation frequency whichever is the greatest. The reason for this is that when two tones are applied the frequency deviation is divided between the two tones so that the modulation indexes are made lower than the index which would exist with a single tone. The Bessel Functions of these indexes are also low so that the resultant amplitude coefficient is made quite small since it is the product of two rather low quantities. For instance, take the case of a frequency modulation system with a deviation ratio[4] of unity and a maximum modulation frequency of ten kilocycles. If an eight-kilocycle tone were applied at full deviation (ten kilocycles) the first-order side frequency would have an amplitude of 0.51 and the second order 0.17. Now if the deviation were equally divided between an eight-kilocycle tone and a nine-kilocycle tone, their first-order side frequencies would have amplitudes of 0.298 and 0.268, respectively, the second-order side frequencies 0.048 and 0.038, respectively, while the beat side frequency corresponding to the sum of the two-tone frequencies would have an amplitude of 0.08. Thus, dividing the deviation between the two tones reduces the second-order side-band out-of-channel interference and produces a beat side frequency having an amplitude which is less than that of the single-tone second-order side frequency. Hence, it can be seen that the greater the number of frequencies present in the modulating wave, the more will the wave be confined to its channel. This is especially true for the case of program or voice modulation where the lower modulating frequencies have the highest amplitudes. It is also especially true for the case of the frequency-modulation systems with the higher deviation ratios.

[4] The deviation ratio referred to here is the ratio between the maximum frequency deviation the system is capable of and the maximum modulation frequency for which the system is designed. This factor might be called the "system" deviation ratio. For further information regarding the deviation ratio see the following reference: Murray G. Crosby, "Frequency Modulation Noise Characteristics," *Proc. I.R.E.*, Vol. 25, pp. 472-514, April, 1937.

A STUDY OF ULTRA-HIGH-FREQUENCY WIDE-BAND PROPAGATION CHARACTERISTICS*

By

R. W. George

R.C.A. Communications, Inc., Riverhead, N. Y.

Summary—Signals reflected from buildings and other large objects introduce distortion in the received signal because of their relative time delay and phase relations. This distortion is especially evident in the form of blurred and multiple images in television reception. Data on the relative merits in this respect, of vertically and horizontally polarized waves transmitted from the Empire State Building in New York City, were obtained at the two frequency ranges of 81 to 86 megacycles and 140 to 145 megacycles. Some data using circular polarization at the lower frequency range were also obtained.

The effects of indirect-path signals were indicated on recorded curves showing field strength versus frequency. The methods and equipment used to record these data at a number of representative receiving locations are briefly described.

A minimum of indirect-path signal interference was found to be generally had with horizontal polarization at both signal-frequency ranges. In this respect, circular polarization was found to be slightly preferable to vertical polarization. Horizontal polarization also gave somewhat greater average field strength.

Miscellaneous data and observations are described, including sample propagation-characteristic curves. In conclusion, some relations between direct- and indirect-path signals and propagation path are discussed.

ONE problem in television is to obtain received pictures free from secondary images caused by time-delayed signal components propagated from the transmitter along paths of different lengths. It is an object of this paper to show something of the nature of this problem, especially in and around such urban areas as New York City.

As an introduction to this subject,[1] P. S. Carter and G. S. Wickizer found that the use of some antenna directivity and horizontal polarization would be most effective for a 177-megacycle television circuit between the RCA Building and the Empire State Building. In the present paper the methods used and results of a similar but more extensive survey made to determine the relative differences in indirect-

* Decimal classification: R113.7. Presented before I.R.E. Thirteenth Annual Convention, New York, N. Y., June 18, 1938.

[1] P. S. Carter and G. S. Wickizer, "Ultra-High-Frequency Transmission Between the RCA Building and the Empire State Building in New York City," *Proc. I.R.E.*, vol. 24, pp. 1082-1094; August, (1936).

Reprinted from *Proc. I.R.E.*, January, 1939.

path propagation between horizontally and vertically polarized waves are described.

The field strength for a given frequency is the vector sum of all the signal components present. At another frequency, the signal components may be of the same intensity but will have different phase relations which result in a different value of field strength. A series of field-strength measurements plotted with corresponding frequencies will show a sinusoidal variation of field strength with frequency in a simple case of the combination of two signal components propagated over paths of different lengths.

It can be shown that if f_1 and f_2 are frequencies between which one signal component has gone through 360 degrees phase shift with respect to the other, or in other words, between which frequencies a complete cycle of field-strength variation is had, the difference in path lengths d and the accompanying time delay t of the longer-path signal are

difference in path length,

$$d = \frac{3 \times 10^8}{f_2 - f_1} \text{ meters} \tag{1}$$

time delay,

$$t = \frac{1}{f_2 - f_1} \text{ seconds.} \tag{2}$$

Equations (1) and (2) may be derived as follows:

N is the number of wavelengths of path length d at the frequency f_1.

$(N + 1)$ is the number of wavelengths of path length d at the frequency f_2.

λ_1 is the wavelength corresponding to f_1.

λ_2 is the wavelength corresponding to f_2.

$$d = N\lambda_1 = (N + 1)\,\lambda_2. \tag{3}$$

Eliminating N from (1)

$$d = \left(\frac{d}{\lambda_1} + 1\right)\lambda_2$$

$$d\left(1 - \frac{\lambda_2}{\lambda_1}\right) = \lambda_2$$

$$d = \frac{\lambda_2 \lambda_1}{\lambda_1 - \lambda_2}$$

$$d = \frac{\frac{c^2}{f_2 f_1}}{\frac{c}{f_1} - \frac{c}{f_2}}$$

where c, the velocity of light, $= 3 \times 10^8$ meters per second

$$\lambda_1 = \frac{c}{f_1} \text{ and } \lambda_2 = \frac{c}{f_2}$$

$$d = \frac{c}{f_2 - f_1} = \frac{3 \times 10^8}{f_2 - f_1} \text{ meters} \tag{1}$$

$$t = \frac{d}{c} = \frac{1}{f_2 - f_1} \text{ seconds.} \tag{2}$$

Thus, measurements over a frequency range of 5 megacycles $(f_2 - f_1)$, are ample to indicate signal-arrival-time delays of 0.2 microsecond and upwards. Other indirect-propagation-path characteristics are indicated by such measurements and will be discussed later.

Measuring Equipment and Methods

A frequency range of 81 to 86 megacycles was chosen to avoid interference with existing local radio circuits. Tests were also conducted at the higher-frequency range of 140 to 145 megacycles. The transmitting systems were built and operated by members of the engineering department of R.C.A. Communications, Inc., at Rocky Point, Long Island, New York.

Each transmitting system supplied substantially constant radiated power as a mechanical system varied the frequency at a constant rate of change from one extreme to the other in one-half minute and likewise back again in one-half minute. Frequency control was maintained by the use of resonant concentric circuits.[2] Variation was obtained in the 81- to 86-megacycle transmitter by changing the length of the frequency control element with a motor-driven cam. In the 140- to

[2] C. W. Hansell and P. S. Carter, "Frequency Control by Low Power Factor Line Circuits," *Proc. I.R.E.*, vol. 24, pp. 597-619; April, (1936). (See Page 53.)

145-megacycle transmitter, frequency variation was produced by means of a small motor-driven variable condenser. The radiated power was about 750 watts at the lower frequency and 68 watts at the higher frequency.

Both transmitters were installed about 1200 feet above the ground near the top of the cylindrical steel tower on the Empire State Building. Short transmission lines were connected to the respective special doublet antennas. These antennas were each mounted about one-fourth wave from the tower on a copper pipe containing the two-wire trans-

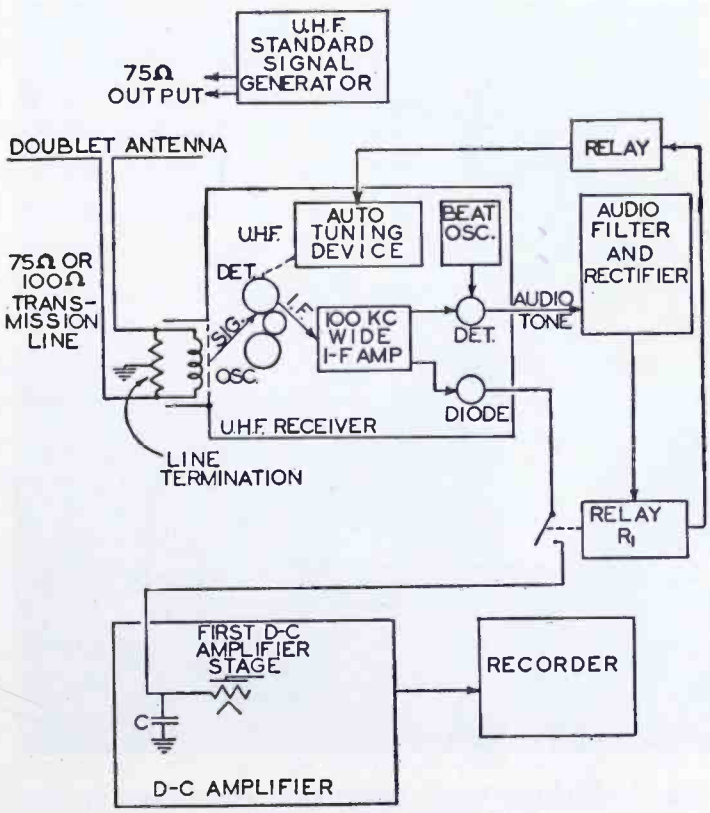


Fig. 1—Block diagram of the field-strength-measuring system.

mission line and were arranged to be turned manually to transmit either vertically or horizontally polarized waves. Both frequencies were in turn radiated from approximately the north and south sides of the tower.

The 81- to 86-megacycle antenna on the south side was made of two doublets crossed at their centers, each with a separate transmission line to the transmitter. By this arrangement, in addition to horizontal or vertical polarization, it was possible to obtain circular polarization by establishing a 90-degree phase difference in the transmission lines before connecting them in parallel at the transmitter.

The object of the field-strength-measuring system was to make a series of recorded measurements as the transmitter frequency varied from one extreme to the other. The block diagram, Fig. 1, indicates the functions which made this system somewhat unusual. A half-wave doublet antenna fed a low-impedance transmission line which was properly terminated at the receiver. The receiver, of conventional design, was suitably modified to meet the present requirements. The ultra-high-frequency detector input and heterodyne-oscillator circuits were ganged together by gears and adjusted to track over the desired

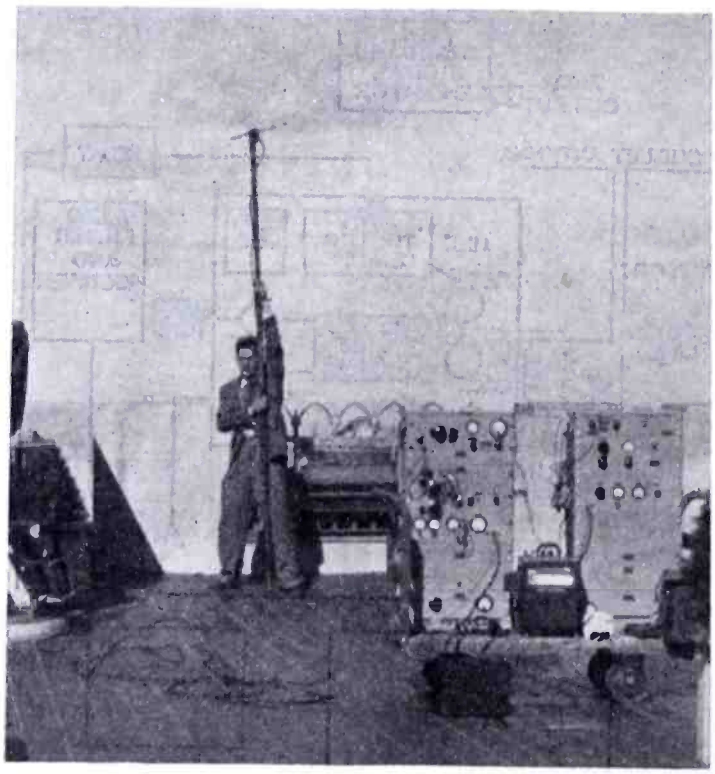


Fig. 2—Field-strength-measuring equipment on the roof of the RCA Building.

signal-frequency range, 81 to 86 megacycles or 140 to 145 megacycles.

In operation, the signal frequency, changing at the rate of 166 kilocycles per second, was in the pass-band of the receiver about 0.3 second before reaching the mid-band of the intermediate-frequency amplifier, and produced, by means of the diode detector, an output voltage corresponding to the value of the input signal voltage. In order to record the signal level, at this instant, use was made of the audio-frequency beat note produced by combination of the intermediate frequency and a fixed heterodyne oscillator in a separate detector. This audio-frequency beat-note output, converted to direct current, caused the relay R_1 to connect the diode voltage to the grid of the first

direct-current amplifier stage A, thus charging the condenser C to the value of the diode voltage, and operating the recorder through the subsequent direct-current power amplifier. After about 0.1 second, the audio-frequency output was removed as the signal frequency increased and passed out of the audio-frequency filter, causing the relay R_1 to disconnect the diode from the condenser C. The condenser C maintained the voltage on the grid of the direct-current amplifier A, holding the recorder at its last reading until the next measurement was made. During subsequent measurements, the condenser C was charged or discharged according to the new signal level.

A series of measurements was started by tuning the receiver to the lowest frequency reached by the signal. When the signal swept into the receiver, a measurement was made, causing the recorder to show abruptly the starting signal level for the forthcoming curve. When the measurement was completed, the relay R_1 returning to normal operated an electro-magnetic device which quickly set the receiver tuning up to a frequency just higher than the increasing signal frequency. As each measurement was completed, this operation was repeated until at the end of one-half minute, the highest frequency was reached. The result was that about 70 evenly-spaced measurements were recorded showing a substantially complete field-strength-versus-frequency curve. The recorder chart was driven at a constant rate, thus spreading the 5-megacycle range uniformly on the chart.

The method of calibrating the record was to substitute a standard-signal generator having a 75-ohm output, for the doublet antenna. This established the voltage in the antenna, and from the known effective height of the doublet, the field strength was determined. The necessarily frequent calibrations were obtained by connecting the signal generator directly to the receiver and taking into account the transmission-line loss. The standard-signal generator was also used to check the response of the measuring system. The over-all response was constant with fixed voltage input over the desired frequency range and the output was directly proportional to the input. The response of the doublet antennas over the 5-megacycle range was known to be substantially constant.

A final check of the flatness of response of the entire measuring system was obtained by recording field strength versus frequency at a receiving location where one would expect, from consideration of the terrain, that a constant field strength should be obtained. This location, about 13 miles from the transmitter, was well away from possible reflecting objects in a clear flat field with several miles of unobstructed ground in the direction of New York. One major in-

direct-path signal existed at this location due to reflection from the ground but its path was so little different in length from the direct-signal path that negligible changes in resultant field strength were had over the 5-megacycle range.

Duplicate measurements were often made, all of which showed good agreement when propagation conditions and the antenna position were unchanged.

The signal generator, receiver, direct-current amplifiers, and voltage-regulating power-supply units were mounted in two portable racks and installed in a ¾-ton panel truck in which measurements were made wherever possible. Otherwise, the equipment was removed to the desired receiving locations. Power was usually obtained from near-by 110-volt alternating-current outlets.

Receiving antenna locations were generally chosen which might be considered suitable for a television antenna. In some cases, such as on the roof of the RCA Building, Figure 2, measurements were made at several antenna positions.

In order to make comparisons, the antenna was always placed in the same position at each location. All locations were within 22 miles from the transmitter except one which was 42 miles away.

Circular polarization was received with an antenna arrangement similar to that used at the transmitter.

Summary of Data: Horizontal and Vertical Polarization

As expected, at a number of random receiving locations, the measurements showed a large variety of direct- and indirect-path-signal combinations. The relative strength of the indirect signals is an indication of the intensity of expected interference or secondary television images. Thus, to show the relative merits of horizontal and vertical polarization with respect to minimum indirect-signal interference, comparisons of the ratio of maximum field strength to minimum field strength, obtained from representative data for each location and polarization are shown in Figure 3. From this figure it is seen that most of the points are above the value of 1.0, indicating that at most of the locations, indirect-path-propagation efficiency was better for vertical polarization than for horizontal polarization.

The geometric means of all of the maximum to minimum field-strength ratios for each polarization and frequency range are shown in Table I. These data indicate that the indirect interfering signals

TABLE I
GEOMETRIC MEANS OF THE RATIOS, MAXIMUM FIELD STRENGTH TO MINIMUM FIELD STRENGTH

	81 to 86 megacycles	140 to 145 megacycles
Horizontal polarization	1.86	2.12
Vertical polarization	2.97	3.38

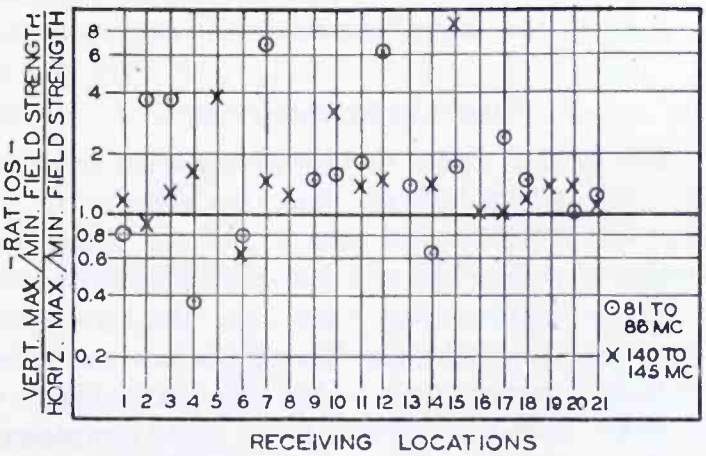


Fig. 3—Comparisons between vertical and horizontal polarization of maximum to minimum field-strength ratios obtained at each receiving location and signal-frequency range.

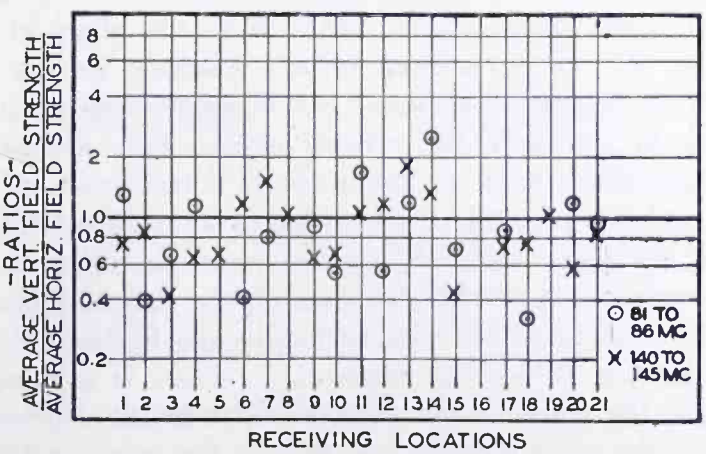


Fig. 4—Comparisons between vertical and horizontal polarization of average field strengths obtained at each receiving location and signal-frequency range.

were from 10 to 20 per cent stronger at the higher frequency and were strongest with vertical polarization at both frequencies.

Average field strengths were determined for each location from the recorded data. In Figure 4 are shown the ratios of average field strength for vertically polarized waves to average field strength for

horizontally polarized waves. The geometric mean of all these ratios for the 81- to 86-megacycle range is 0.83, and for the 140- to 145-megacycle range 0.82. This summary indicates that in general horizontally polarized waves were received about 20 per cent stronger than were vertically polarized waves.

The geometric mean of all the average field strengths using horizontal polarization was for the 81- to 86-megacycle range, 88.5 per cent of that obtained at the 140- to 145-megacycle range. This comparison is based on the same transmitted power at both frequency ranges.

Circular Polarization

It was predicted that under certain propagation conditions involving a direct path and an indirect path, the rotation of a circularly polarized wave would be reversed upon reflection. Thus, the receiving antenna adjusted to receive the direct circularly polarized wave, would not respond to the indirect-path wave. For the circularly polarized wave to be reversed upon reflection, it would be necessary for the wave component polarized perpendicular to the plane of incidence to undergo a 180-degree phase shift with respect to the wave component polarized in the plane of incidence. In general, this phase relation may be established under the following conditions.[3]

For a wave polarized perpendicular to the plane of incidence, the reflected wave is always substantially 180 degrees out of phase with the incident wave. For a wave polarized in the plane of incidence, the phase of the wave reflected from a medium having negligible conductivity, is the same as that of the incident wave when the angle of incidence is less than the critical angle. This critical angle is known as Brewster's angle or the angle of polarization in connection with light, and is a function of the dielectric constant of the reflecting medium. If the stone walls of buildings may be regarded as having negligible conductivity at high frequencies, the desired phase relations may thus be established between vertically and horizontally polarized signal components. When the angle of incidence is greater than the critical angle, the phase of the reflected wave is 180 degrees different from that of the incident wave and thus is the same as the phase of a reflected wave polarized perpendicular to the plane of incidence.

In order for a receiving antenna to respond only to the direct-path signal, it would be necessary for the received indirect-path vertically and horizontally polarized wave components to be equal as well as 180 degrees out of phase. Practically, this condition could not occur

[3] Bertram Trevor and P. S. Carter, "Notes on Propagation of Waves Below Ten Meters in Length," *Proc. I.R.E.*, vol. 21, pp. 387-426; March, (1933).

in New York City because of the presence of several indirect paths, each involving different coefficients of reflection for vertically and horizontally polarized waves, and the angles of incidence usually being too large to give the necessary 180-degree phase difference between vertically and horizontally polarized waves after reflection.

Measurements using circular polarization were made at three types of locations using the 81- to 86-megacycle transmitter. These locations were: south of Newark Airport in the previously mentioned clear area, on the roof of 75 Varick Street Building, and from a north side window on the 26th floor of the Woolworth Building. In order to obtain relatively accurate information, measurements of both hori-

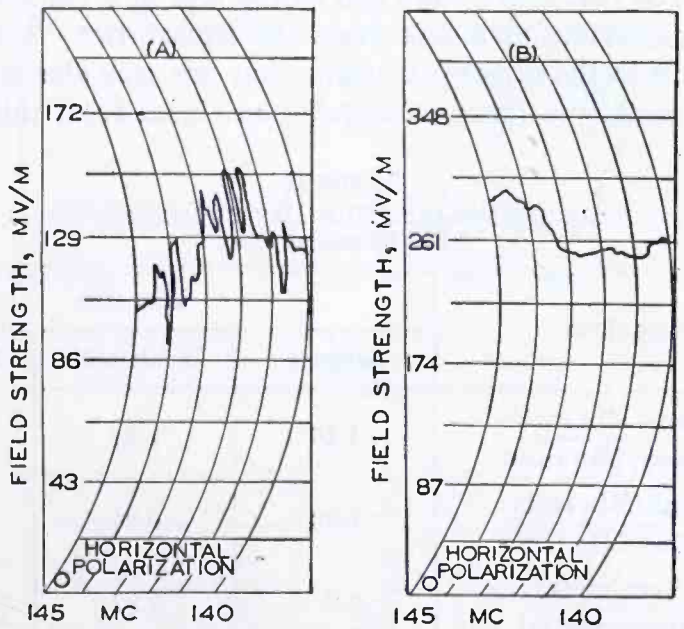


Fig. 5—Propagation-characteristic curves obtained on the roof of a 10-story building about three quarters of a mile south of the transmitter. For curve *B* the antenna had a relatively clear exposure to the transmitter, and for curve *A* its exposure was partially obstructed by a near-by ventilator.

zontal and vertical polarization were made consecutively for direct comparison.

The summary of these measurements in Table II shows circular polarization to be in general less desirable than horizontal and possibly somewhat more desirable than vertical polarization.

Miscellaneous Data and Observations

Some data were obtained on the average field strengths of the horizontally polarized component received for vertically polarized wave transmission and conversely. These indicated that horizontally polar-

ized wave transmission had an average vertical component of from 25 to 30 per cent of its horizontal field strength, and that for vertically polarized transmission, the horizontal component was somewhat less, on the order of 20 per cent of the corresponding vertical-field strength.

A change in receiving-antenna position by a small distance of one or two feet usually altered the shape of the recorded propagation-characteristic curve. Very large differences in field strength as well as shape of the curve were had at locations having near-by obstructions or reflecting objects.

The importance of obtaining the strongest possible direct signal was well illustrated under the following conditions. Two antenna positions, about 30 feet apart, were had on the roof of a ten-story building about three-quarters of a mile from the transmitter. Both positions had exposure to the direct-path signal but one was almost obstructed by a large metal ventilator. The reflected signal from this ventilator

TABLE II
CIRCULAR POLARIZATION DATA COMPARISONS
81 to 86 megacycles

Comparison	Location		
	Airport	Woolworth	Varick St.
$\frac{\text{Vert. Max/Min ratio}}{\text{Horiz. Max/Min ratio}}$	1.16	1.24	1.37
$\frac{\text{Circ. Max/Min ratio}}{\text{Horiz. Max/Min ratio}}$	1.02	0.94	1.11
$\frac{\text{Avg. Vert. mv/m}}{\text{Avg. Horiz. mv/m}}$	0.8	0.97	1.48
$\frac{\text{Avg. Circ. mv/m}}{\text{Avg. Horiz. mv/m}}$	1.0	0.91	1.16

combined with the direct signal at the antenna to produce a weakened resultant signal. Other indirect signals not so much affected by the ventilator, if at all, were relatively strong with the result that they greatly modified the received signal. This was in contrast with the results obtained at the better exposed antenna where the direct-path signal was stronger. Propagation-characteristic curves obtained at each antenna position are shown in Figure 5.

The additional sample propagation-characteristic curves shown in Figures 6, 7, 9, 10, and 11 will give an idea of the nature and range of the field-strength variations encountered. The calibrations shown are corrected for one kilowatt radiated. Usually there may be only one or two indirect paths of major importance and, if desired, their

corresponding time delay t and difference in path length d can be roughly determined. Apparently, a complete analysis of the indirect

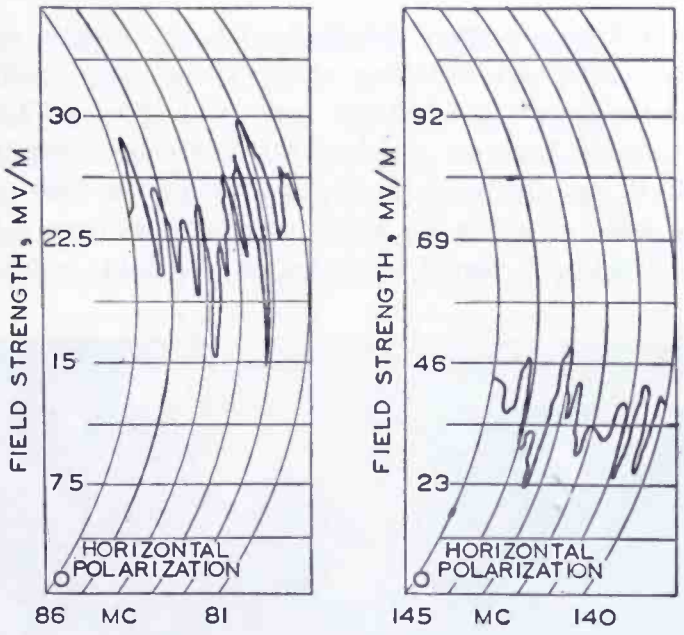


Fig. 6—Propagation-characteristic curves obtained on the roof of a 10-story apartment building about three miles north of the transmitter.

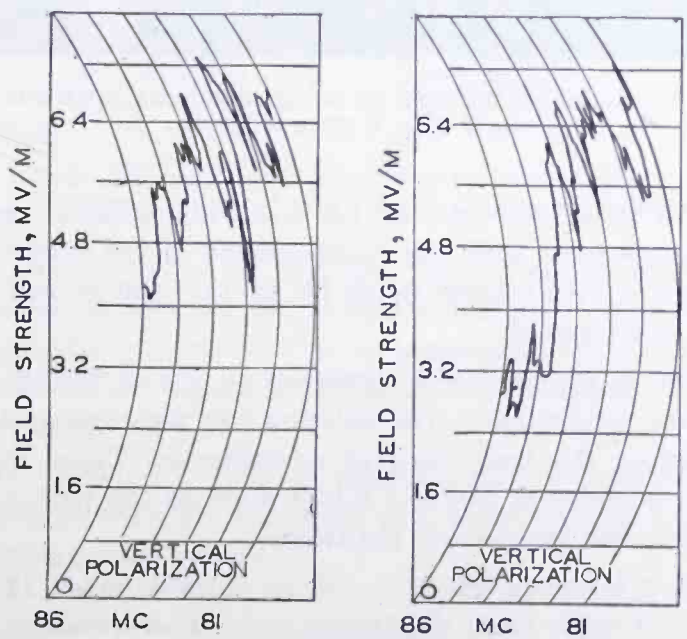


Fig. 7—Propagation-characteristic curves obtained at a residence on Staten Island about 12 miles from the transmitter. The two curves are for antenna positions about seven feet apart.

paths would be difficult and have little value because the problems presented at each proposed receiving location must, in practice, be solved individually.

The data in Figure 6 were obtained with an antenna on the roof of a 10-story apartment building about three miles north of the Empire State Building. A television receiver having a 12-inch Kinescope was available here on which a fairly strong secondary image was observed. It was displaced about one fourth of an inch corresponding to a time delay of about 1.5 microseconds. The time delay of one prominent indirect-path signal indicated on the data, is found by the

Fig. 8—Receiving antenna on a portable mast, with which data in Fig. 9 were obtained.

use of (2), to be on the order of 1.5 to 1.8 microseconds which is in substantial agreement with the displacement of the observed image. The difference in path lengths would be, by (1), 450 meters for a time delay of 1.5 microseconds.

The curves in Figure 7 were obtained on Staten Island about 12.5 miles from the transmitter. The antenna was mounted on the 22-foot portable mast in the back yard of a residence. These curves were taken within a period of two and a half minutes but with a difference of about seven feet in antenna positions.

In Figure 8 is shown one setup of the portable mast, 17 feet high, at a location 14 miles from the transmitter. The crossarm at the top of the mast supported the transmission line and the antenna. The crossarm could be rotated through 90 degrees by means of ropes, thus controlling the polarization of the antenna. The position of the antenna

could be changed on a circle of 5-foot radius by simply rotating the mast. Representative data obtained at this location are shown in Figure 9.

The unusually strong indirect-path signals indicated in Figure 10, were recorded with the antenna on a parapet, 12 stories above and adjacent to Fifth Avenue. The tower of the Empire State Building was just visible between the intervening buildings and was about one mile south. The two curves show considerable differences between the indirect paths involved with horizontally and vertically polarized waves. The uniformity of the field-strength variation curve for hori-

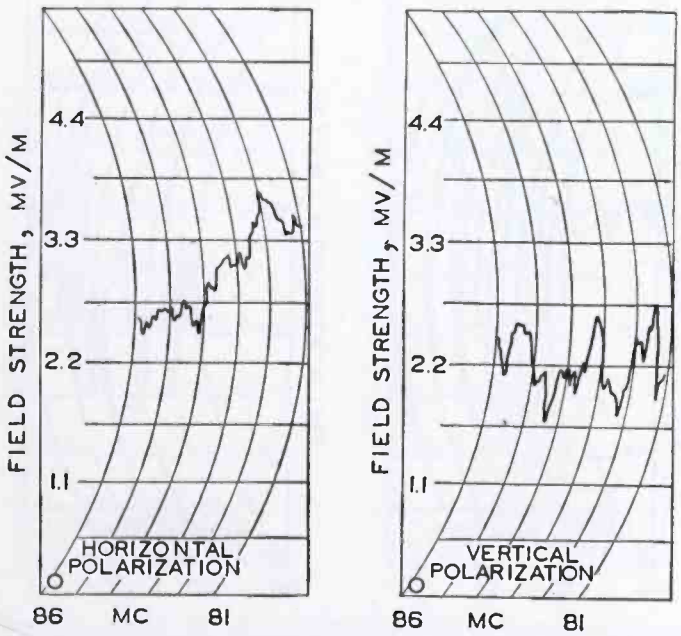


Fig. 9—Propagation-characteristic curves obtained at a residence about 14 miles north of the transmitter.

zontal polarization indicates that only one major interfering indirect wave was received, and this might be eliminated by the use of a simple directive-antenna system. The more complex combination of indirect waves obtained for vertical polarization would apparently not be so easily controlled.

The data in Figure 11 were recorded in the clear area south of Newark Airport with the object of testing the flatness of response of the entire measuring system.

Conclusions

One conclusion substantiated by the data in this paper is that vertically polarized waves are propagated more efficiently along indirect paths than are horizontally polarized waves. The relatively large

differences observed between direct- and indirect-path lengths show that the reflecting objects involved must generally have been the vertical walls of buildings rather than their roofs or ground. Consideration of the general properties of such imperfect reflectors having small values of conductivity, will also lead to the above conclusion. It is known[3] that for all but very large or small angles of incidence, waves polarized perpendicular to the plane of incidence (referred to as horizontal polarization with respect to the earth's surface) are reflected more efficiently than are waves polarized in the plane of incidence.

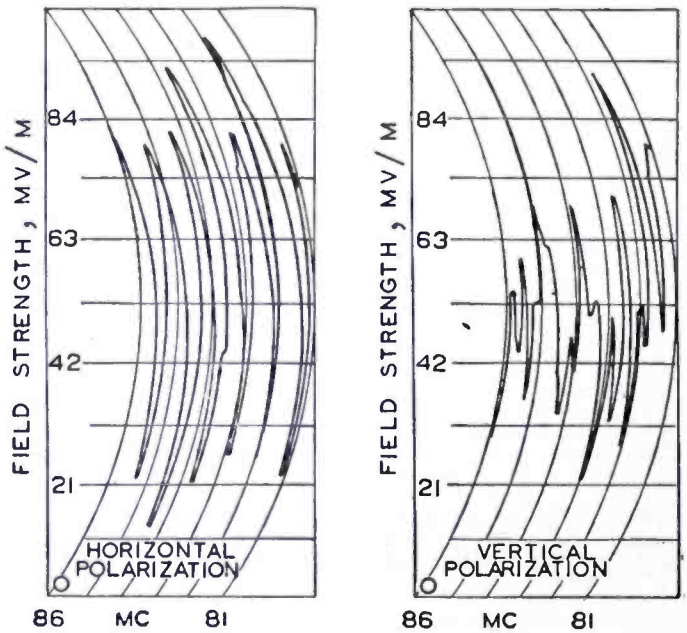


Fig. 10—Propagation-characteristic curves obtained about one mile north of the transmitter on Fifth Ave.

Thus, vertically polarized waves are polarized perpendicular to the plane of incidence with reference to the vertical walls of buildings and consequently reflected therefrom most efficiently.

In some cases, indirect signals will arrive from a reflecting point at one side or behind the receiving antenna and can be reduced or eliminated by the use of relatively simple directive antenna systems.

If the indirect-path carrier is combined in phase with the direct-path carrier, the result will be an increase in carrier strength with no change in the values of the side bands of the respective carriers. Thus, the depth of modulation is reduced for both the direct- and indirect-path side bands. The time delay of the indirect-path side bands ordinarily will not be long enough to affect the synchronizing impulses but will produce a displaced image. This image will be of

the same polarity as that produced by the direct-path side bands. Both images would be complete except for the complex interference between each set of side bands.

If the indirect-path carrier is combined out of phase with the direct-path carrier, the result will be a reduced carrier with no change in the values of the side bands of the respective carriers. A secondary image produced by the indirect-path side bands in this case will be reversed in polarity with respect to the image produced by the direct-path side bands. It is highly probable that the resulting weakened

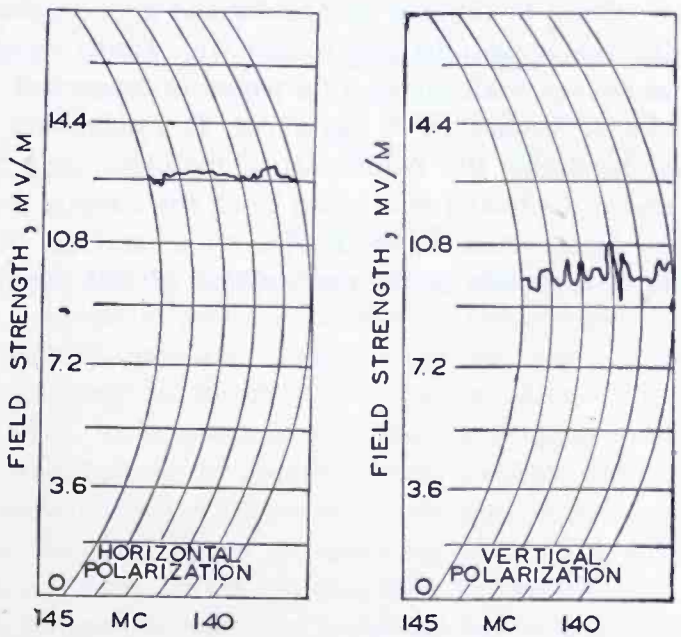


Fig. 11—Substantially constant field-strength-versus-frequency data obtained south of Newark Airport.

carrier would be overmodulated at times by either or both sets of side bands.

It is not within the purpose of this paper to discuss further the complex interference conditions resulting from the combination of direct- and indirect-path, wide-band modulated waves. In closing it may be mentioned that some phase modulation is one product resulting from direct- and indirect-path modulated waves combined with the respective carriers having intermediate phase relations.

No conclusive data are available to show the stability of indirect-propagation-path characteristics under varying weather conditions. However, it is apparent that heavy rain may cause appreciable changes in the constants of certain reflectors.

The study of indirect-propagation-path characteristics at ultra-high frequencies is obviously a large and fruitful field for investiga-

tion. Propagation by reflection from the ionosphere may present some occasional problems in television broadcasting at the lower frequencies but this phenomenon was outside the scope of the investigations here reported. Future investigations with widely different types of transmitter locations or propagation areas may give somewhat different results; however, it is not expected that vertical polarization will generally be found preferable to horizontal polarization.

ACKNOWLEDGMENT

The author wishes to express his thanks and appreciation to members of the RCA organizations and others who kindly co-operated in conducting the survey with which this paper is concerned. Credit is due in particular to Messrs. P. S. Carter, N. E. Lindenblad, and G. L. Usselman who produced the transmitting facilities, and to Messrs. R. K. Gallup and L. L. Young who stood by at the Empire State Building during the tests. Messrs. H. H. Beverage and H. O. Peterson rendered valuable criticism in the preparation of this paper.

ULTRA-HIGH-FREQUENCY PROPAGATION

By

M. Katzin

R.C.A. Communications, Inc., Riverhead, N. Y.

Summary—This paper presents a short summary of the subject of ultra-high-frequency propagation, with theoretical and experimental results presented side by side, to show the general agreement between them. The quasi-optical behavior of u-h-f waves is exhibited in the phenomena of reflection, diffraction, and refraction. The reflecting properties of the ground, and their application in explaining propagation over plane earth, are shown. The presence of multipath reflections in urban areas and their effect on wide-band transmission is discussed. Diffraction and refraction are discussed, the variable nature of refraction giving rise to fading experienced beyond the horizon. Coverage data of the Empire State television transmitter are used to illustrate the practical application of many of the points discussed in the paper.

Introduction

With the development of improved means for generating and controlling power at ultra-high frequencies, have come many applications and proposals for putting these frequencies to work, particularly in the field of communication. To mention but a few, television, local broadcasting, police and other mobile communication, and applications in aviation radio indicate the growing importance of communication by means of ultra-high-frequency waves, and, therefore, the importance of knowing the characteristics of their propagation. Many investigations, both theoretical and experimental, have been made in an endeavor to determine these characteristics, with the result that most of the more important features of the problem appear to be known. Now that television has finally rounded the corner in this country, it seems appropriate to present a short summary of the subject of ultra-high-frequency propagation.

During the past decade, this subject has been actively prosecuted on many fronts through numerous experimental and theoretical researches. This activity has built up a considerable body of knowledge and an extensive literature. Obviously, in any survey of the subject such as presented in this paper it will be possible to aim only at the high spots, reference being made to the bibliography for many of the interesting and important details.

Ultra-high-frequency propagation is governed by the same general laws that apply to ground-wave transmission on lower frequencies, the chief distinction being in order of magnitude. This difference in order of magnitude, however, is a very important one, for many of the

Reprinted from *Proc. Radio Club of America*, September, 1939.

phenomena, which at much lower frequencies are usually of negligible effect, become of controlling importance at the ultra-high frequencies. Thus, buildings, trees, and even irregularities of the terrain produce distortions of the field in their neighborhood. At the ultra-high frequencies the dimensions of such objects and irregularities may become comparable with or greater than the wavelength so that they produce well-defined reflections. A further point is that with wavelengths of the order of feet, very small path differences between separate components or rays are sufficient to result in appreciable phase differences. At lower frequencies, say in the broadcast frequency range, the wavelength is relatively so large that such effects are usually largely negligible. We see from these few examples that the shortness of the wave can be responsible for many phenomena which are not encountered at lower frequencies.

The early propagation experiments, most of them of a qualitative nature, soon established the quasi-optical behavior of these waves. Hence it is possible to borrow some of the concepts from optical theory and apply them to ultra-high-frequency propagation. The analogues of optical reflection, refraction, and diffraction are encountered in propagation at the ultra-high frequencies, so that by referring to the results of optics, many of the phenomena encountered in u-h-f work can be explained.

In this paper, we shall start from the free-space field radiated by an antenna, and proceed to show that the effect of locating the antenna above ground is similar in many respects to the optical concept of reflection. This leads to an examination of the reflecting properties of the ground, and the application of these properties in explaining propagation over plane earth will be shown. The presence of multiple reflections in urban areas will be pointed out and its effect on wide-band transmission discussed. From the case of plane earth there follows a consideration of the effect of the earth's curvature. This leads to a discussion of the phenomena of diffraction and refraction, the variable nature of the latter giving rise to fading experienced beyond the horizon. Finally, a discussion of coverage data of the Empire State television transmitter will be used to illustrate the practical application of many of the points discussed previously. Throughout, the plan will be to present theory and experiment side by side, to show the general agreement between them.

Propagation Over Plane Earth—(Level Terrain)

The radiation field in the equatorial plane of a linear antenna in free space is given by the familiar equation

$$\xi_0 = \frac{60\pi HI}{\lambda d}, \tag{1}$$

where H—effective height of antenna,

I—antenna current $= \sqrt{W/R}$, $\begin{cases} W\text{—radiated power,} \\ R\text{—antenna radiation resistance,} \end{cases}$

λ = wavelength,

d = distance from antenna.

For a half-wave doublet, $H = \lambda/\pi$, $R = 73.3$ ohms, so that (1) becomes

$$\xi_0 = \frac{7\sqrt{W}}{d}. \tag{2}$$

If such an antenna is placed in the vicinity of the ground, or some other plane reflecting service, then, in first approximation, the effect of the ground on the field at any receiving point in space is equivalent to a ray regularly reflected from the ground, and the resulting field at the receiver may be considered as due to the combination of the direct and reflected rays. The total distance traversed by the reflected ray is greater than that by the direct ray, so that a phase difference between them results. In addition, a phase shift, as well as a reduction in amplitude, takes place when a ray is reflected at the ground. This can be expressed by means of a complex reflection coefficient,

$$\text{Reflection Coefficient} = Ke^{j\alpha}.$$

K being the reduction in amplitude, and α the phase shift on reflection. The total phase difference between the direct and reflected rays is thus the sum of the phase difference due to the difference in path length and the phase shift on reflection.

Denoting the path difference between the reflected and direct rays by $\triangle$, and the corresponding phase lag by $\delta = 2\pi\triangle/\lambda$, the total phase lag is $\Phi = \delta - \alpha$, and the total field at the receiver is given by

$$\xi = \xi_0 \cdot \sqrt{(1-K)^2 + 4K\cos^2(\Phi/2)}. \tag{3}$$

Since the reflection coefficient of the ground enters into the expression for the total field, it is of interest to examine its properties.

The magnitude, K, of the reflection coefficient depends on the dielectric constant and conductivity of the reflecting medium, and the angle of incidence of the rays, and, in addition, different relations hold for horizontal and vertical polarization. The simplest case to consider is that of a pure dielectric, for which the phase shift on reflection is

either 0 or 180°, so that the reflection coefficient is real. For horizontal polarization, the wave is always reversed on reflection from a pure dielectric (i.e., $\propto = 180°$) while the magnitude of the reflected wave decreases from a value of $K = 1$ at grazing incidence to a value at perpendicular incidence which depends on the dielectric constant. For small angles of the ray to the reflecting surface, the reflected wave suffers practically no attenuation on reflection, so that K differs inappreciably from unity.

For vertical polarization, however, the behavior is quite a bit different. Here, again, for grazing angles the reflected wave is reversed

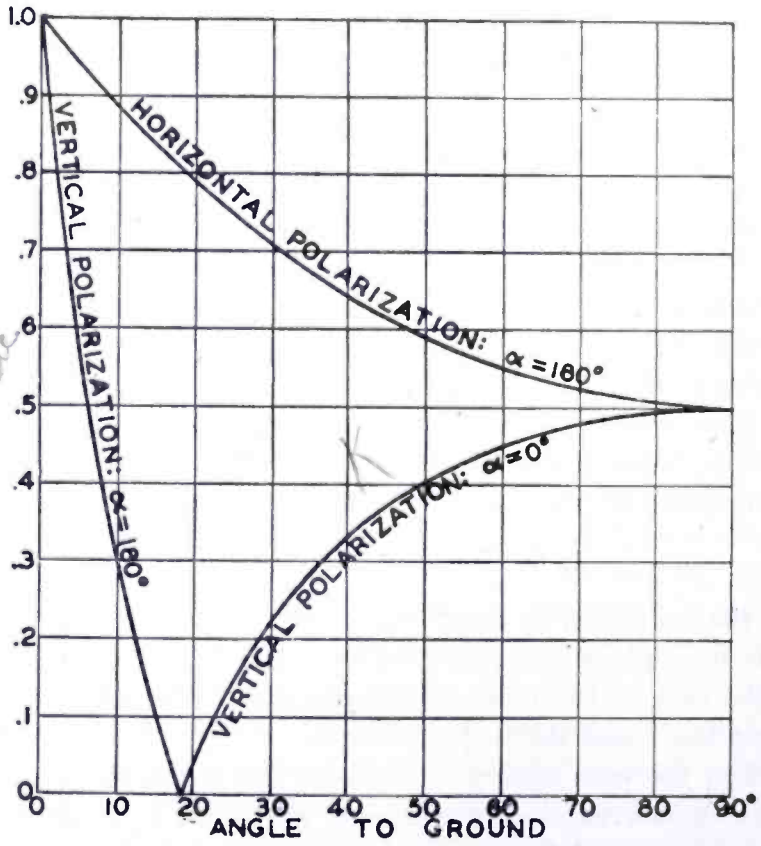


Fig. 1—Reflection Coefficient of Ground Having $\varepsilon = 9$, $\sigma = 0$.

in phase without reduction of amplitude ($\propto = 180°$, $K = 1$) for grazing angles, but for increasing angles its amplitude decreases rapidly and becomes zero at a certain angle whose cotangent is equal to the square-root of the dielectric constant. Above this angle there is no change in phase on reflection ($\propto = 0$), while the amplitude of the reflected wave increases steadily to a value at perpendicular incidence which is the same as that for horizontal polarization. Figure 1 shows the reflection coefficient for both vertical and horizontal polarizations

for a ground of zero conductivity and a dielectric constant of 9, representing Long Island ground, for ultra-high frequencies. In this case, the angle at which no reflection takes place for vertical polarization is about 18°.5.

When the conductivity of the reflecting medium is not negligible, the relations are more involved. The phase shift on reflection is other than zero or 180°, in general (complex reflection coefficient). For horizontal polarization, the phase angle of the reflection coefficient is always in the second quadrant, but, for most practical cases, the effect

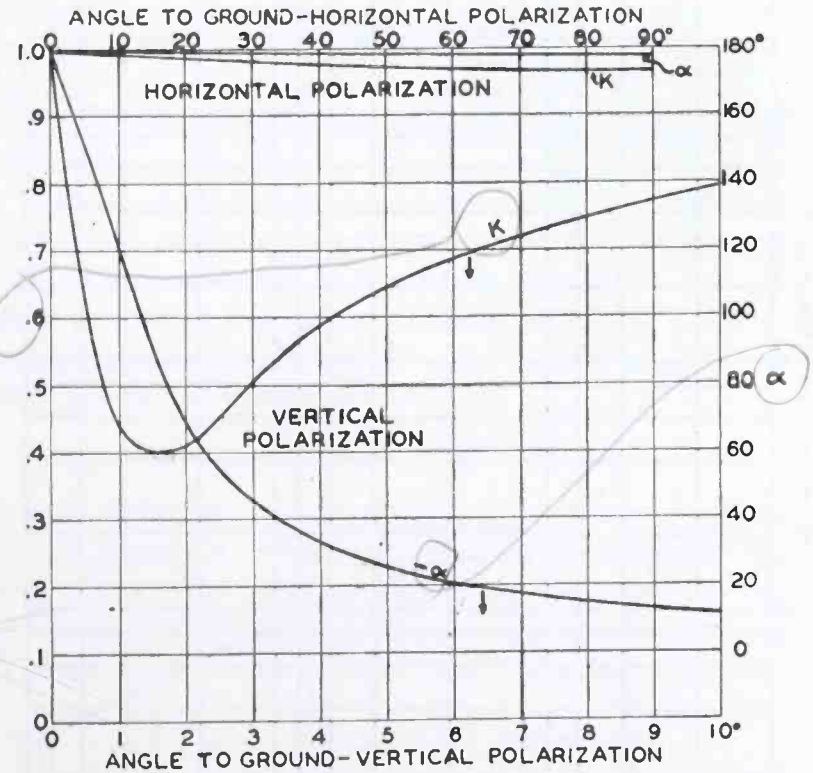


Fig. 2—Reflection Coefficient of Sea Water at 50 Mc/s.
$\varepsilon = 80$, $\sigma = 4.10^{-11}$ e.m.u.

of conductivity is quite negligible. For vertical polarization, on the other hand, the phase angle of the reflection coefficient is in the third or fourth quadrants. Corresponding to the perfect dielectric case, the reflected wave is reversed in phase without reduction of amplitude for zero angle, and the reflected wave decreases in amplitude rapidly for increasing angles. Instead of passing through zero, however, it reaches a finite minimum value, and thereafter increases once more. At the same time, the phase shift on reflection, considered as a lag, decreases from 180° at zero angle to zero at vertical incidence, passing through 90° at the angle for which the amplitude of the reflected wave is practically a minimum. For a given dielectric constant, the effect of

increasing conductivity is to lower the angle at which the amplitude of the reflected wave is a minimum. The reflection coefficients for sea water for a frequency of 50 megacycles are shown in Figure 2.

It is well to bear in mind the above difference in behavior of the reflection coefficient between horizontal and vertical polarization, for

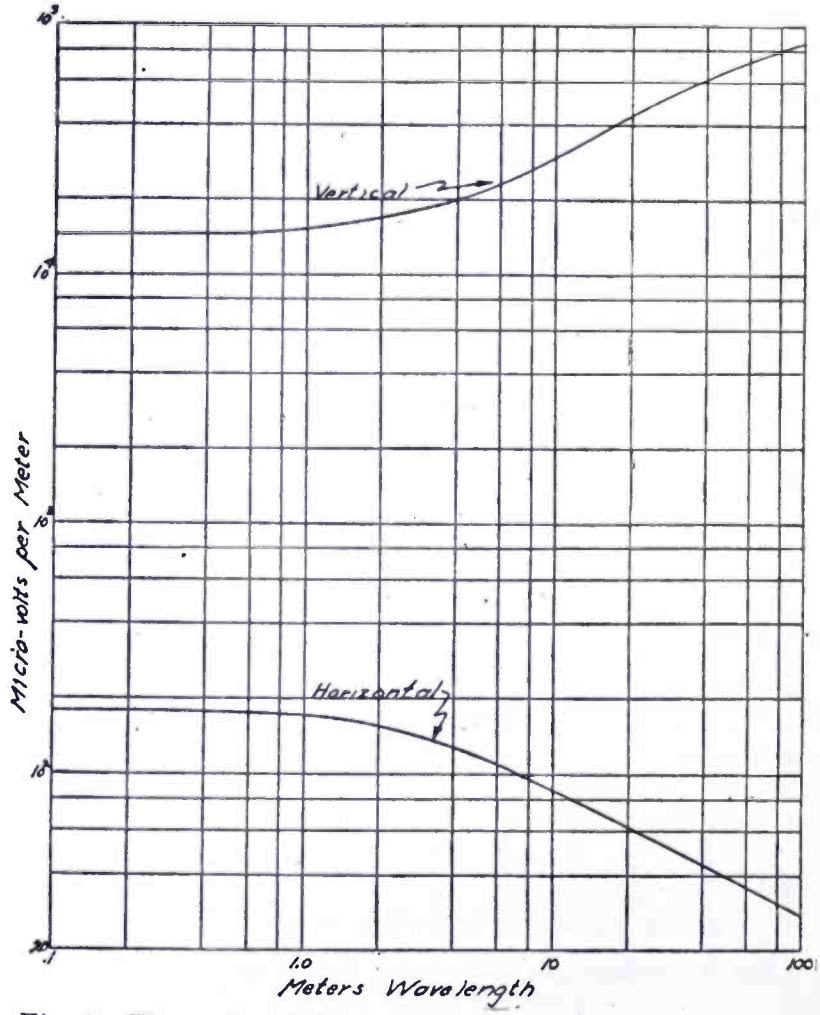


Fig. 3—Theoretical field strength vs. wavelength over salt water at a distance of 1 kilometer from a dipole 8 meters high radiating 1 watt for vertical and horizontal polarization. Receiving antenna height = 0.

it is largely responsible for the difference in behavior of propagation for these two polarizations over mediums of good conductivity, such as sea water. To illustrate this, Figure 3, taken from Trevor and Carter[5*], shows the theoretical variation with frequency of received field strength for low antennas, for both vertical and horizontal polarization. At the higher frequencies, where the curves become horizontal, the dielectric current predominates over the conductivity current, and

* Numbers refer to bibliography at end.

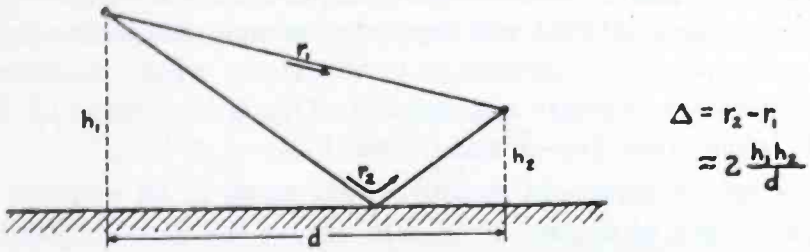


Fig. 4

the sea water "ground" behaves like a pure dielectric, the limiting ratio of vertical to horizontal polarization field strength then being equal to the dielectric constant, in this case 80. The phase shift on reflection for vertical polarization departs from 180° as the frequency is lowered, giving increased field strengths with vertical polarization. With horizontal polarization, on the other hand, there is no appreciable change in phase of the reflected ray with frequency, but the magnitude of the reflection coefficient approaches unity more closely as the frequency is lowered, resulting in lower field strengths.

Instead of writing the reflection coefficient as $Ke^{j\alpha}$, it is convenient for some purposes to write it as $-Ke^{j\Theta}$ where $\Theta = \alpha \pm \pi$, so that Φ becomes $\delta - \Theta$. (3) then becomes

$$\xi = \xi_0 \sqrt{(1-K)^2 + 4K \sin^2(\Phi/2)}. \qquad (3a)$$

The path difference, $\triangle$, depends on the geometry of the circuit, as shown in Figure 4, being a function of the antenna heights and their separation. For distances not large compared to the antenna heights, the path difference may amount to a number of wave-lengths. In such cases, as the height of the receiving antenna, or the distance,

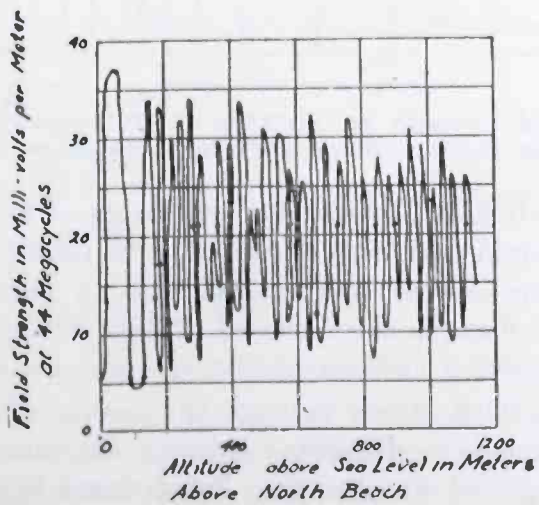


Fig. 5—Field strength vs. altitude at North Beach, 9.6 kilometers from Empire State 44-megacycle transmitter. Radiation 2 kilowatts.

is changed, the path difference will change through a number of wavelengths, and the total field will fluctuate through maximums and minimums, corresponding respectively to conditions where the direct and reflected rays are in-phase and opposed. This is illustrated by Figures 5 and 6, taken from Trevor and Carter[5].

These figures represent measurements made in an airplane on the radiation from a 44-megacycle transmitter on the Empire State Building, radiating vertically polarized waves, with a power of 2 kilowatts. Figure 5 shows the field received at North Beach, a distance of about

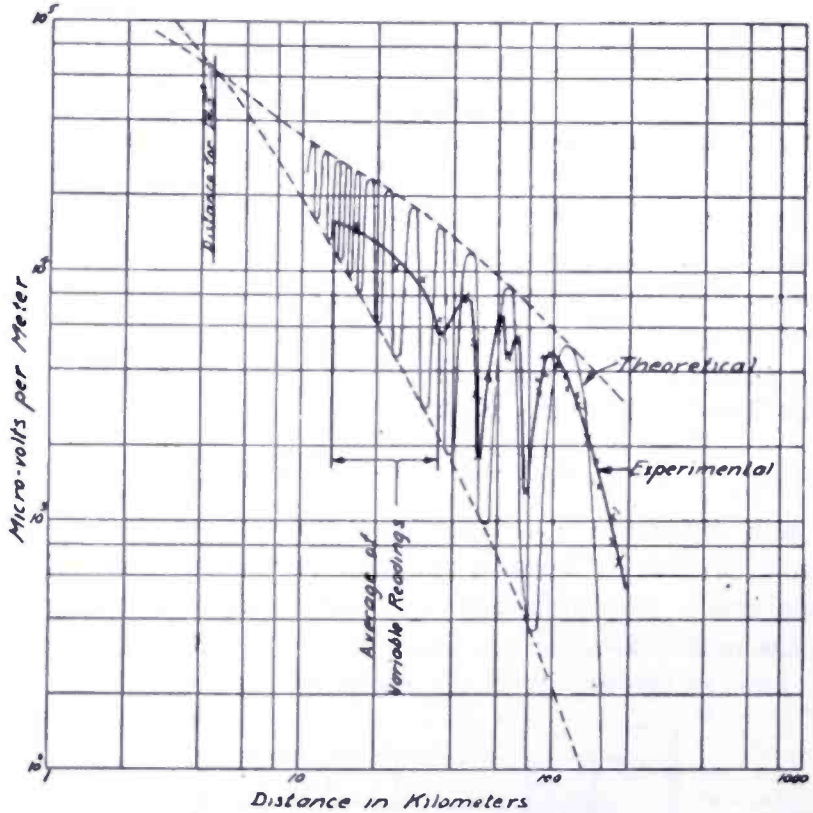


Fig. 6—Field strength vs. distance at 1200 meters altitude—Empire State 44-megacycle transmitter. Radiation 2 kilowatts.

6 miles, as a function of altitude. The variations of field strength with height were so rapid that it was impossible to record them manually, hence a free-hand sketch was resorted to, to portray the effects observed. Figure 6 shows the variation of field strength with distance at a constant receiving antenna height, in this case about 4000 feet.

The previous illustrations indicate the nature of the phenomena observed when the path difference between the direct and reflected rays is of the order of wavelengths. For antenna heights small compared to their separation, which is the more usual case, the angle of the reflected ray is very small, and it is permissible to put $K = 1$, while

Φ will be small, except for cases like sea water. If the path difference is large enough to neglect Θ in comparison with Δ, the total received field becomes

$$\xi = \xi_o 2 \sin \frac{\delta}{2} = \xi_o 2 \sin \frac{2\pi h_1 h_2}{\lambda d} \tag{4}$$

If the antennas are low enough, the sine factor may be replaced by its argument, giving the familiar expression

$$\xi = \xi_o \frac{4\pi h_1 h_2}{\lambda d}$$

$$= 240\pi^2 HI \frac{h_1 h_2}{\lambda^2 d^2} \tag{5}$$

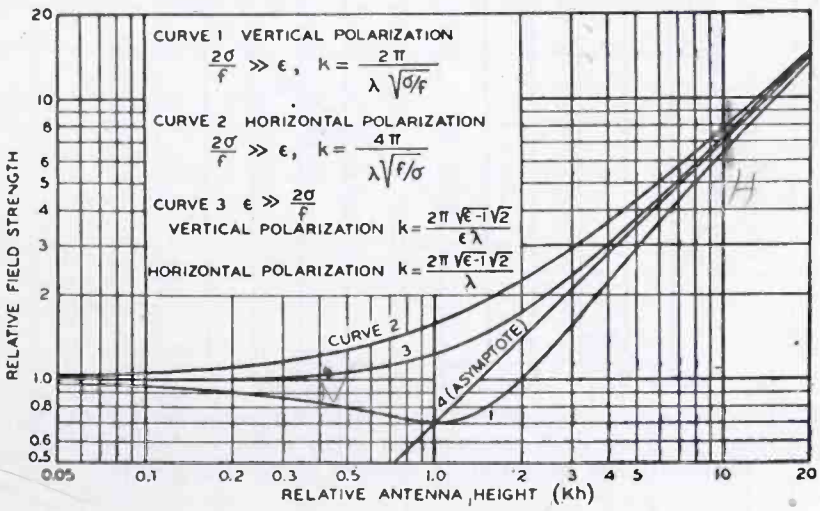


Fig. 7—Variation of received field strength with antenna height.

For a half-wave doublet, this is

$$\xi = 88\sqrt{W}\frac{h_1 h_2}{\lambda d^2}, \text{ (watts, meters, volts/meter)}$$

$$= 0.01052\sqrt{W}\frac{h_1 h_2 f}{d^2}\ \mu v/m, \tag{6}$$

where, in the latter form, h_1 and h_2 are in feet, d in miles, f in megacycles, and W in watts. This is the familiar inverse-square law for propagation over a plane earth. Under the conditions for which it is valid, the received field varies inversely as the square of the distance, directly as the antenna heights, and directly as the frequency. It may

be well to point out the conditions to which use of the above relations are restricted. (4) holds only for antenna heights that are great enough to insure that the path increment yields a phase angle much greater than Θ, and for angles small enough so that K is practically unity. (5) suffers the additional restriction that the path increment must be small enough to permit the sine term to be replaced by its argument, which means, practically, that the path increment should be less than about 1/6 wavelength. An idea of the nature and magnitude of the deviations from the simple form of (5) can be obtained from Figure 7, taken from a paper by Burrows[22]. This figure shows the deviations at low antenna heights from the linear increase of field strength with height predicted by (5), for the two extreme cases of a good conductor (conduction current much larger than dielectric cur-

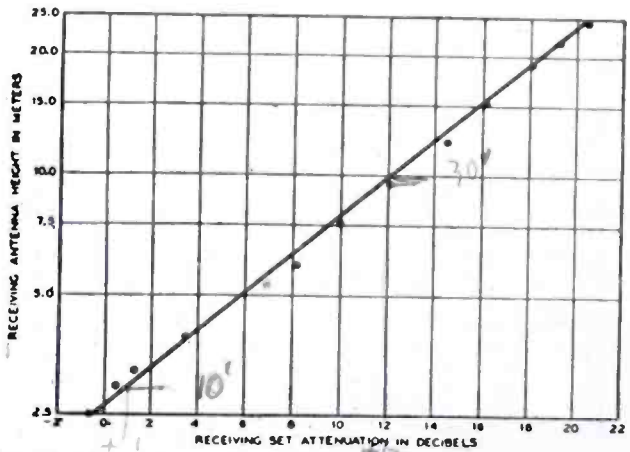


Fig. 8—Variation of received field with antenna height for 26.3 kilometer path on 34.6 megacycles with a horizontal half-wave transmitting antenna 24 meters above the ground.

rent), and a pure dielectric (dielectric current much larger than conduction current). The pure dielectric case applies sufficiently well to ultra-high-frequency propagation over land, while the pure conductor case applies to propagation over sea water. For the dielectric case (land), the linear relation of field to height holds for heights above about one-quarter wavelength for horizontal polarization, and for vertical polarization above about two wavelengths.

The "relative field strength" given by the ordinate of Figure 7 is the ratio of the field strength at any height to that at zero height. The abscissa gives the antenna height to a scale (the factor k) which depends on the ground constants and the polarization. Curves 1 and 2, for vertical and horizontal polarization, respectively, apply to transmission over a good conductor, while curve 3 applies to the case of a pure dielectric. Curve 1, for vertical polarization, shows an initial decrease of field strength with increase of height to a minimum at a

height corresponding to an abscissa value of 1.0, whereas curve 2, for horizontal polarization, indicates a steady increase of field strength with height. The height (abscissa) scales, however, are different for the two cases (different values of k). If curve 2 were plotted to the same abscissa scale as curve 1, curve 4 would result. It is seen that the great advantage of vertical over horizontal polarization in propagation over a good conductor such as sea water holds only for small heights, and that above the height corresponding to 1.0 on the abscissa

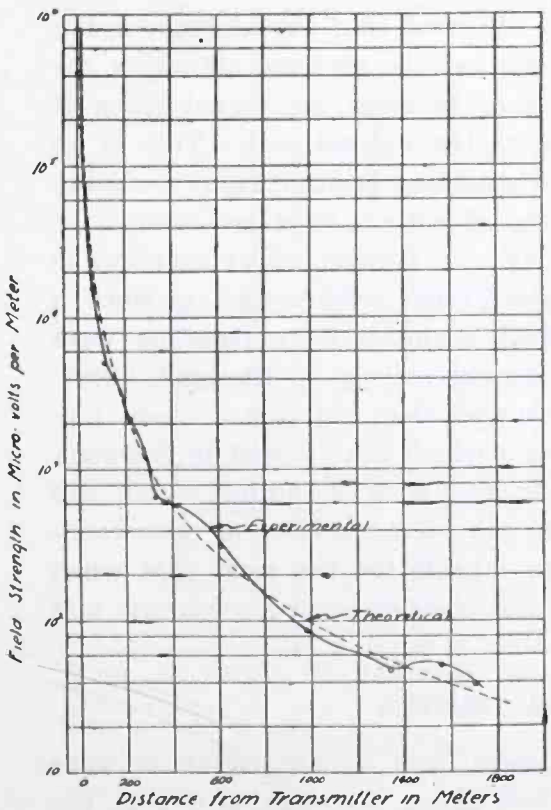


Fig. 9—Field strength vs. distance, vertical polarization, 41.4 megacycles. Transmitting and receiving antennas 2.9 and 3.1 meters above Long Island ground.

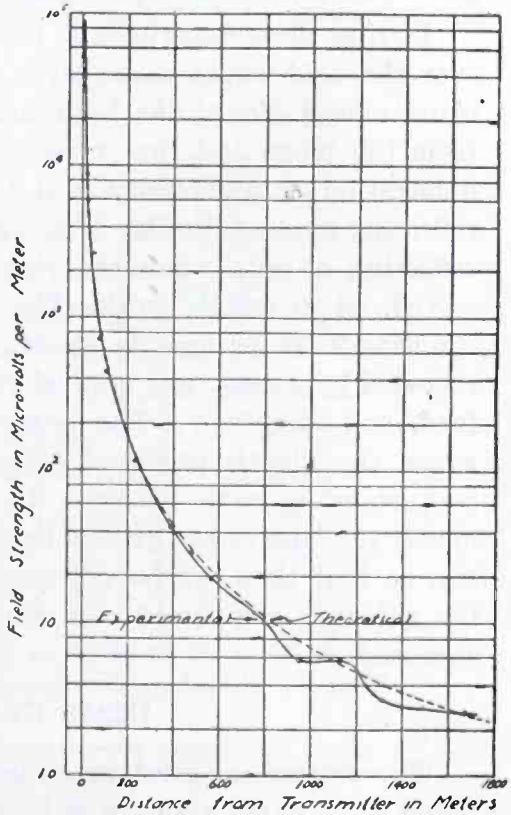


Fig. 10—Field strength vs. distance, horizontal polarization, 41.4 megacycles. Transmitting and receiving antennas 2.9 and 1.6 meters above Long Island ground.

of Figure 7, horizontal polarization actually gives the greater fields. For a frequency of 50 megacycles, this height is 84 feet.

An experimental check of the linear height-field strength relation is shown in Figure 8, taken from a paper by Burrows, Decino, and Hunt[17]. The best straight line through the log-log plot of the experimental points shows a slope very close to unity. Numerous other measurements have substantiated the linear relation predicted from (5).

Experiments conducted by Trevor and Carter[5] over the flat ground at Suffolk Airport near Riverhead yielded results showing confirmation of the theoretical inverse-square variation of field strength with distance. The results for a frequency of 41.4 megacycles are shown in Figures 9 and 10, for vertical and horizontal polarization, respectively. The dotted curves in these figures represent an inverse-square-of-distance variation. Similar results were obtained at 61 megacycles. It is seen that there is not much difference between the results on horizontal and vertical polarization.

During these experiments, it was observed that any airplane flying over the field would cause large variations in received signal as the plane moved along, due to interference between the signal reflected from the plane and that received over the normal path. This is an illustration of interference at a fixed receiving location from a moving reflecting object. Similar interference effects are produced from fixed reflecting objects when the receiving (or transmitting) location is moved, as in mobile work. The interference effects then produce a standing-wave pattern in space. Such standing-wave patterns were reported by Jones[4], and studied rather extensively by Englund, Crawford, and Mumford[7]. The latter showed that individual trees, guy wires, etc., clearly produced reflections which contributed to the complex standing-wave pattern. In general, such standing waves are mostly prevalent near ground level in suburban locations or open country, so that they can be expected to subside for the most part when the antennas are placed well above the level of irregular objects and surfaces.

Urban Characteristics

The theoretical relations which have been considered above were based on an earth assumed to be a level plane. Actually, of course, the character of the ground usually differs considerably from such an ideal assumption. This is particularly true in urban locations, especially in the larger metropolitan areas with many high buildings. Buildings present surfaces which are many wavelengths in extent for ultra-high frequencies, so that well-defined reflections can be expected from them. As a result, transmission between two points in an urban area may take place over a great many paths. Since the path increments between these components may be a number of wavelengths, a slight change in location of the receiving antenna may change the phase relations between the various received components sufficiently to result in a large alteration of the resultant field. In fact, the superposition of the waves which are reflected back and forth between buildings sets up a very complicated standing-wave pattern in space. Obviously, then, the

field intensity may fluctuate up and down widely as the receiving antenna is moved around. This was pointed out very forcefully by the experiments reported by Jones[4].

The standing-wave patterns which result from the superposition of the many components bouncing around building areas are dependent on the phase and amplitude relations between these components. Since for a given set of paths the phase relations depend on the frequency, entirely different standing-wave patterns usually result for frequencies differing even only moderately. Such behavior is of importance in

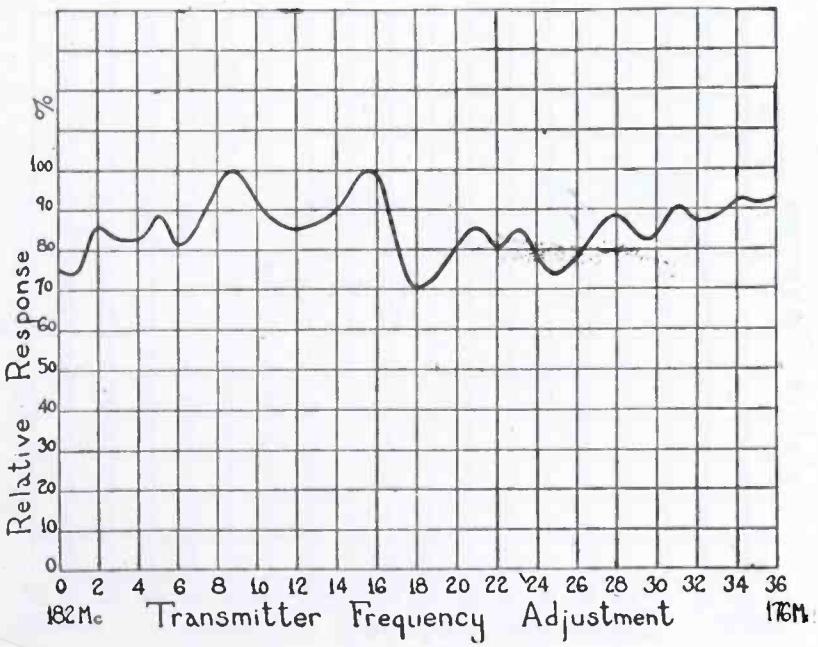


Fig. 11—Response curve with half-wave dipoles at transmitter and receiver. Horizontal polarization, transmitter at fourteenth floor.

wide-band services such as television, for the effect is equivalent to a distortion of the frequency characteristic of the system.

The simple case of two components can be used to illustrate the principles involved in this form of frequency distortion. For a given difference in path between these two components the phase difference between them will be directly proportional to the frequency. At a certain frequency, say f_1, the two components will be in phase and a maximum resultant field will be produced. If the frequency is increased by an amount which makes the phase angle due to the path difference between the two components increase by 180°, say to f_2, the two components will oppose, producing a minimum resultant field. The frequency interval $f_2 - f_1$, from maximum to minimum, or vice versa,

is inversely proportional to the path increment. If this path increment is large enough, therefore, the received field will vary considerably over a band such as used for television transmission. The relations become much more complicated when a multiplicity of components is present.

Such effects have been investigated by Carter and Wickizer[19], and by R. W. George[35]. Carter and Wickizer investigated the frequency characteristic of a circuit transmitting from the RCA Building to a receiver on the 85th floor of the Empire State Building in New York

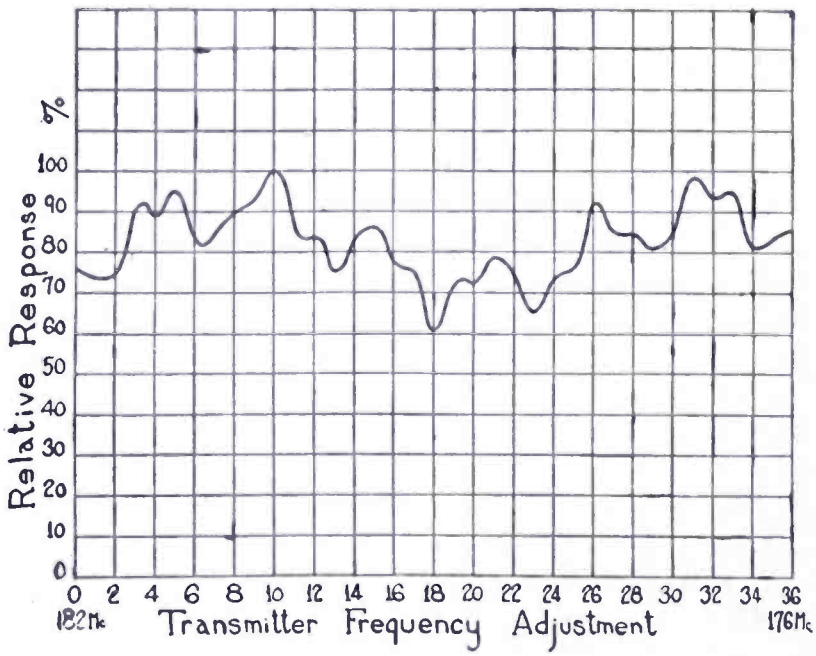


Fig. 12—Curve constructed from vector diagrams.

City. Figure 11 shows the frequency characteristic obtained with the transmitter on the 14th floor of the RCA Building, horizontal half-wave doublets being used for both transmitting and receiving. To explain the apparently irregular spacing of the peaks and dips in this characteristic, an artificial characteristic was constructed by assuming four components to be present, having various path lengths and amplitudes. Such a derived characteristic is shown in Figure 12. It will be noted that this curve bears certain similarities to Figure 11, indicating that the experimental characteristic, Figure 11, is the resultant of a direct ray and several indirect rays having considerably longer paths.

For this particular circuit, such large path differences would be expected only from rays arriving at relatively wide angles, so that the use of directive antennas should reduce the variations in response over

the frequency band. As a check on this, Figure 13 shows the experimental characteristic obtained with directive receiving and transmitting antennas; the maximum variations were reduced to less than ± 1 decibel.

The above characteristics were obtained using horizontal polarization. A comparison of the performance obtained with vertical polarization is afforded by Figure 14, which was obtained with the same antennas used for Figure 11, but arranged vertically. It will be observed that the amplitudes of the variations are much greater than

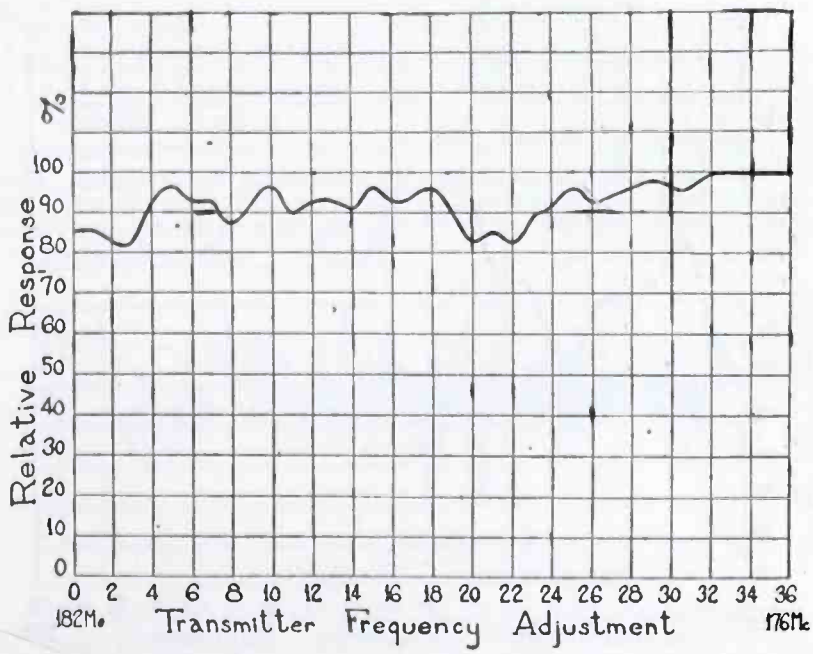


Fig. 13—Response curve with directive transmitting and receiving antennas. Horizontal polarization, transmitter at fourteenth floor.

for the same setup using horizontal polarization, Figure 11. This indicates that the amplitudes of the indirect rays are greater with vertical than with horizontal polarization. A reasonable explanation for this may be deduced from the magnitude of the reflection coefficients for horizontal and vertical polarization given in Figure 1. It will be seen that the magnitude of the reflection coefficient for horizontal polarization (electric field parallel to the reflecting surface) is always greater than that for vertical polarization (electric field perpendicular to the reflecting surface), except for the limiting cases of grazing and perpendicular incidence. Since we are concerned with reflecting surfaces which are predominantly vertical, instead of horizontal ground to which Figure 1 applies, the roles of horizontal and vertical polarization

are interchanged, so that transmission from a vertical antenna corresponds to horizontal polarization with respect to the vertical buildings, and vice versa. Therefore, it is to be expected that vertical antennas would result in reflected ray components of higher amplitude, on the average, than would horizontal antennas.

Similar conclusions were reached by George[35], who made measurements of the received frequency characteristic over the bands from 81-86 megacycles, and 140-145 megacycles at a number of locations in the New York area. Both horizontal and vertical polarizations were used. Figure 15 shows a mass plot which compares the maximum-to-

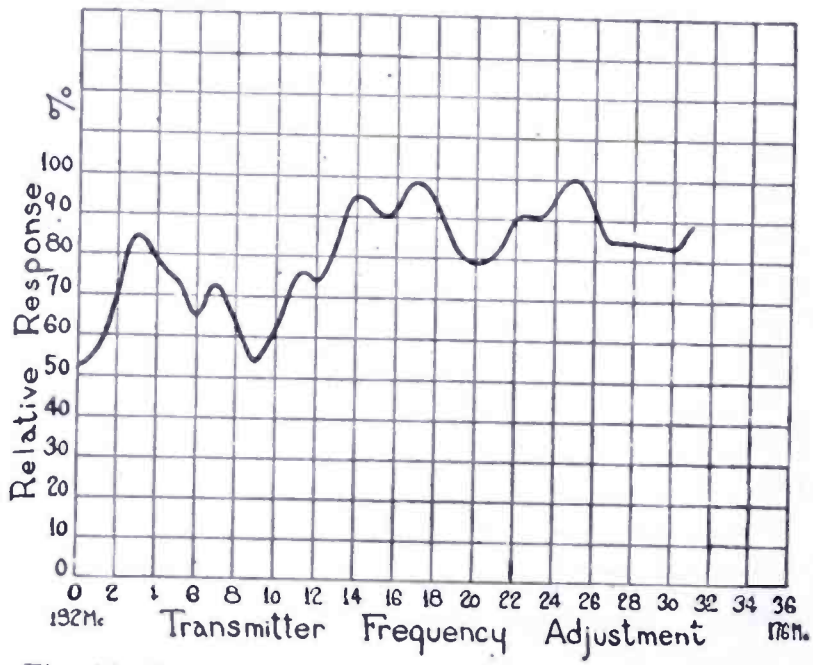


Fig. 14—Response curve with half-wave dipoles at transmitter and receiver. Vertical polarization, transmitter at fourteenth floor.

minimum ratios measured over the five-megacycle band for vertical and horizontal polarization. It is seen that for most of the test locations, vertical polarization resulted in larger variations over the band. In addition, for any one polarization, the mean maximum-to-minimum ratio was greater over the 140-145 megacycle band than over the band 80-85 megacycles.

Another interesting point that emerged from George's study is that, on the average, horizontal polarization produced an average field about 2 decibels stronger than vertical polarization, for both frequency bands. This is shown by Figure 16.

The above experiments represent measurements at a number of fixed points. A more complete picture of the standing-wave variations

can be obtained by mobile recording. This was done in a mobile survey of the Boston area by Burrows, Hunt, and Decino[11], and of the New York area by Wickizer[38]. Sample portions of the records taken in

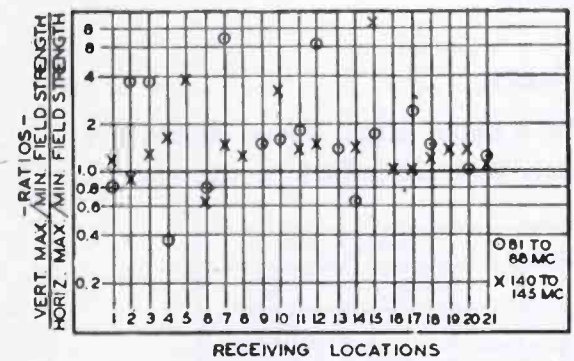


Fig. 15 — Comparisons between vertical and horizontal polarization of maximum to minimum field-strength ratios obtained at each receiving location and signal-frequency range.

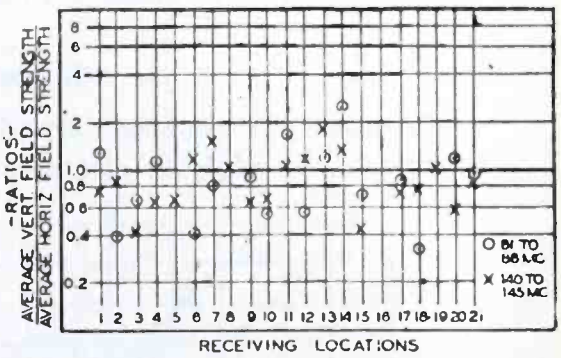
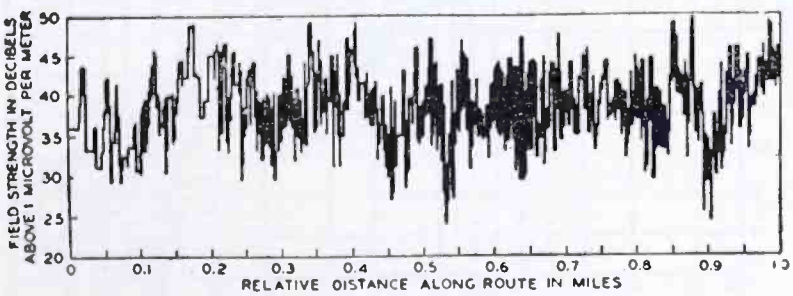


Fig. 16 — Comparisons between vertical and horizontal polarization of average field strengths obtained at each receiving location and signal-frequency range.



Portion of record showing the large field strength variations as recorded while driving through the business district of Boston at a distance of about 1.5 miles from the transmitter.

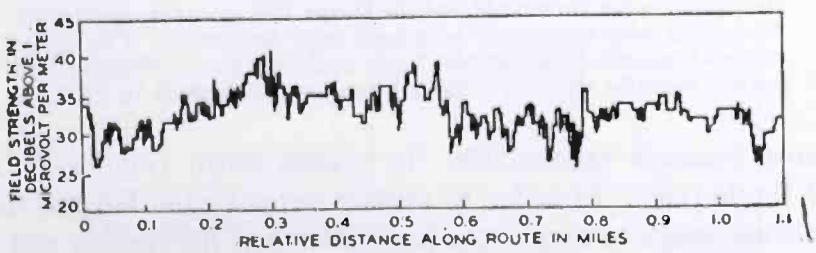


Fig. 17—Portion of record showing the small variations of field strength while driving through residential section of Boston at a distance of about 5 miles from the transmitter.

Boston are shown in Figure 17, which shows the much larger variations in field strength in the business district than in the residential district. A mass plot of the measurements is shown in Figure 18, for which values averaged over short intervals of distance have been used. It is seen that most of the points lie within ± 10 decibels of the mean

curve having an inverse-square-of-distance slope. The dashed line is the theoretical inverse-square-of-distance curve for level terrain plotted from (5). It thus appears that the effect of irregular terrain is to lower the mean average field by about 10 decibels and to superimpose variations of about ± 10 decibels.

Curved Earth

Equation (4) has been derived for a plane earth, hence should not be expected to apply at distances such that the effect of the earth's

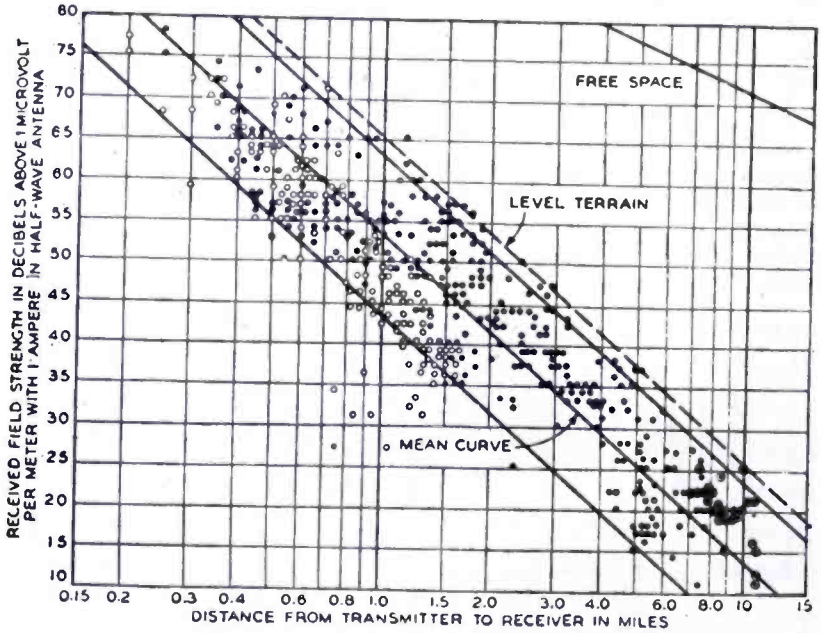


Fig. 18—Mass plot of field intensities measured at various distances from the transmitter at Berkeley and Stuart Streets in Boston. The values corresponding to distances less than two miles represent field strengths averaged over one-tenth mile intervals, while those for greater distances represent averages over one-half mile intervals. The open circles indicate fields in the high building area. Residential points outside the city limits have been enclosed in circles.

curvature becomes appreciable. It breaks down completely at and beyond the horizon. In order to explain propagation beyond the horizon, it is necessary to invoke the mechanisms of diffraction and refraction, both familiar from optical theory.

As is well known, it is not correct to assume that waves travel strictly in straight lines. Rectilinear propagation is a result of wave interference between components propagated over paths at either side of the rectilinear path. If an obstacle is interposed in the path of some of these components, part of the destructive interference taking place among them will be destroyed, so that radiation into the "dark" region

behind the obstacle can take place. This mechanism supplies a diffracted field which, for a given frequency, depends on the geometry of the circuit, and hence a steady field.

Refraction in the atmosphere is due to its varying index of refrac-

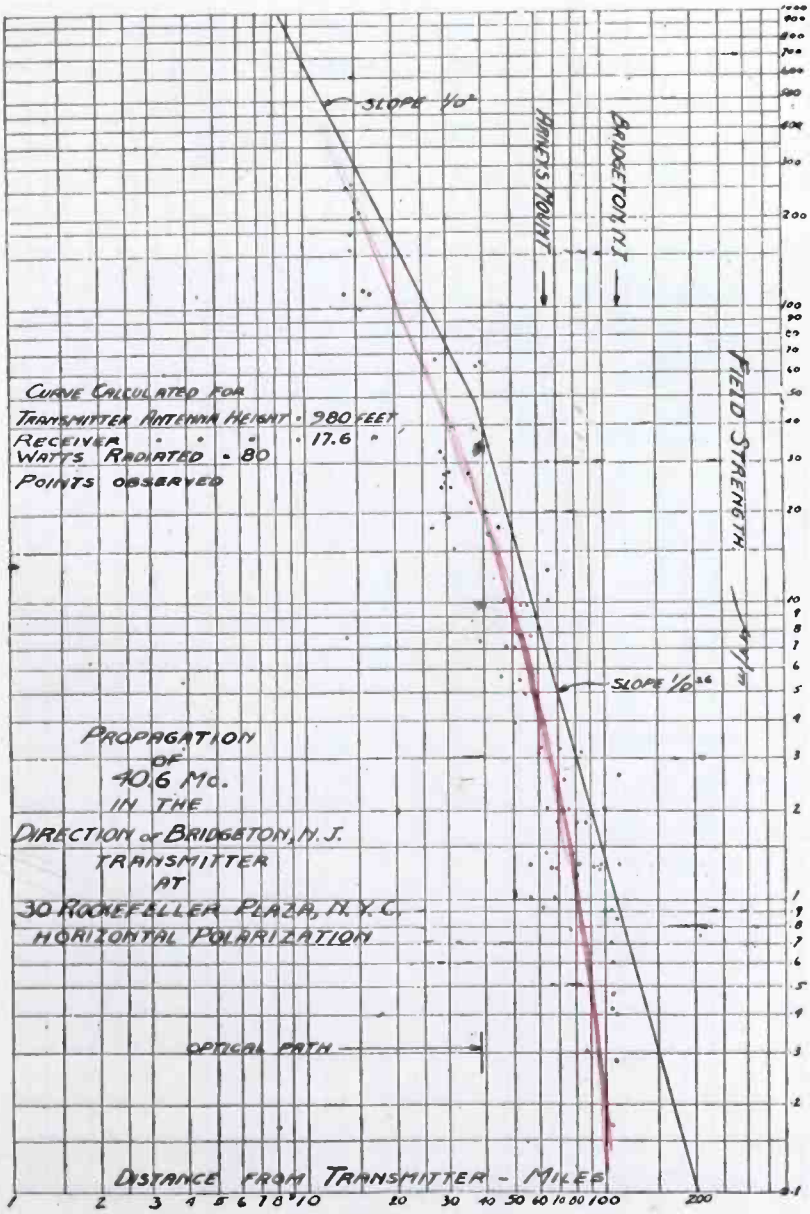


Fig. 19

tion with height. The index of refraction decreases with height, so that, for wave fronts propagated nearly horizontally, the phase velocity of the upper portions of the wave front is higher than that of the lower portions, resulting in bending of the wave front downward. This downward bending of the wave front compensates in part for the

curvature of the earth, resulting in increased field strengths at distances beyond the horizon. The index of refraction of the atmosphere varies with its temperature, relative humidity, and movement of air masses, so that the rise in field strength due to refraction is subject

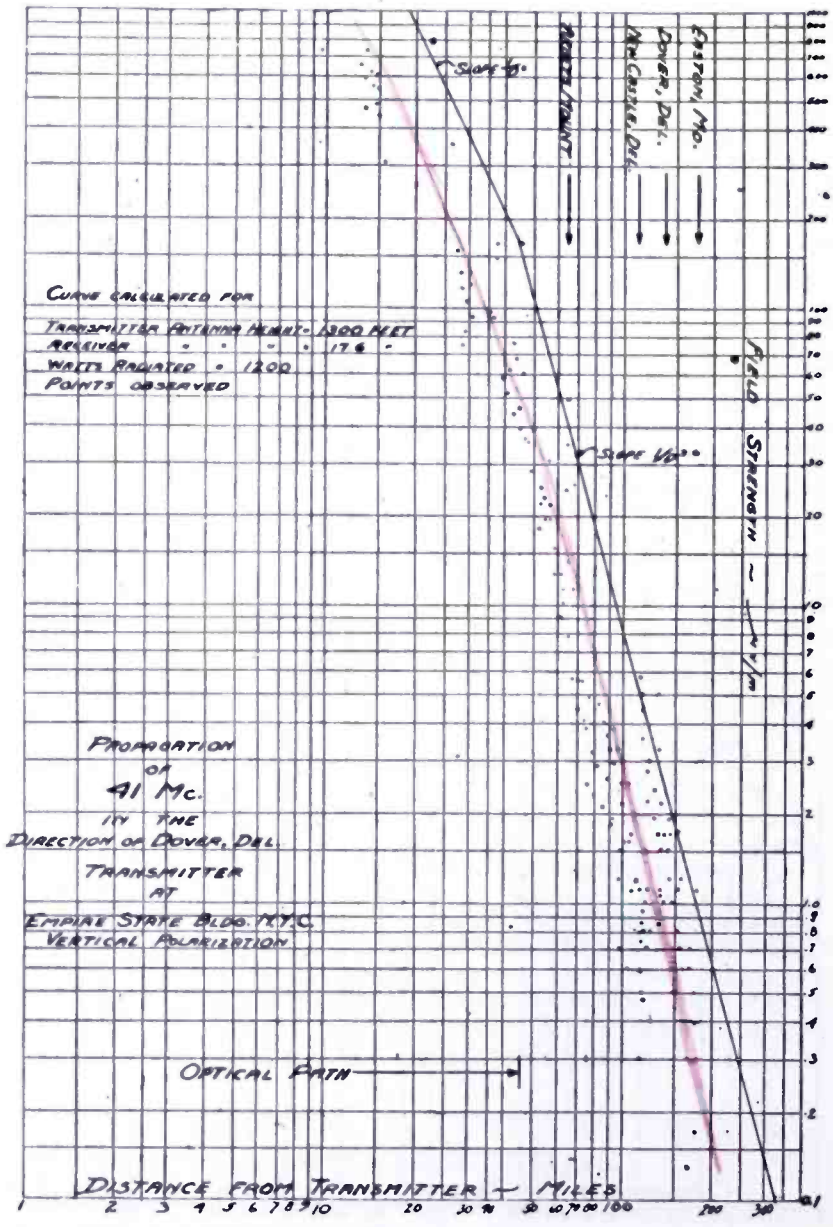


Fig. 20

to corresponding variations. The slow fading up and down of ultra-high-frequency signals over longer paths is attributable to these variations in index of refraction. In addition, it has been demonstrated recently that reflections from atmospheric irregularities[34], for example air-mass boundaries, contributes to the field beyond the horizon, and

is probably responsible for the more rapid types of fading experienced. This may be considered as a limiting case of refraction, much as refraction in the ionosphere is spoken of as "reflection."

Diffraction

A rigorous solution of the diffraction of radio waves around the earth is exceedingly complicated. A general solution, in the form of an infinite harmonic series, was given in 1910 by Poincare for vertical polarization, but this result was of little use for practical purposes in the radio case, since at frequencies of about 40 megacycles it has been

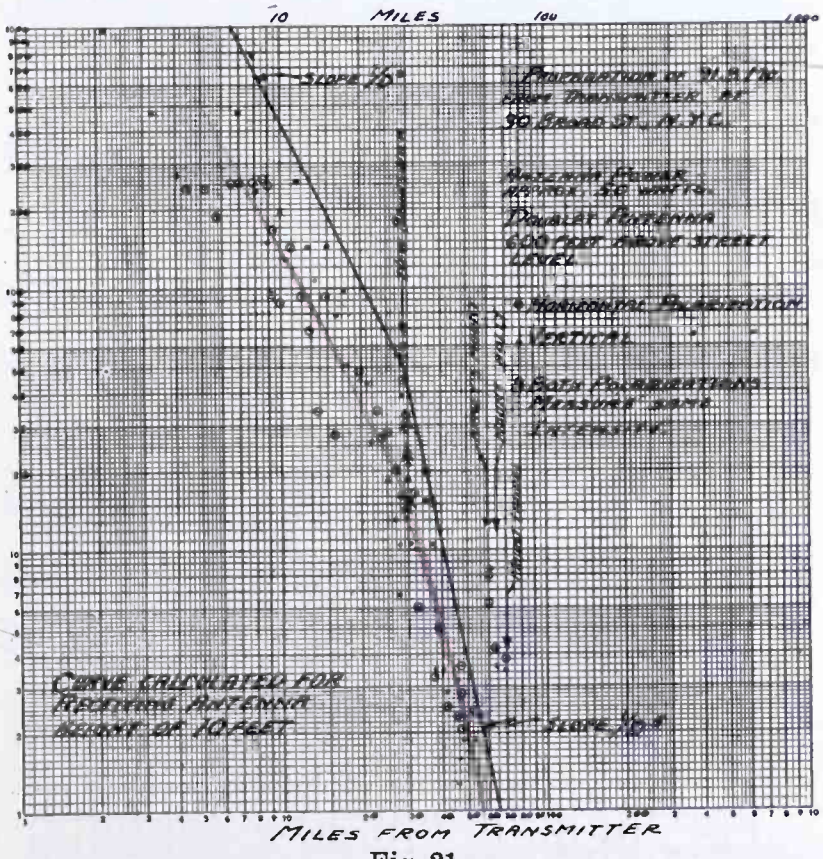


Fig. 21

estimated[33] that the principal contribution to the numerical result would be given by about a thousand terms either side of the six-millionth term. Watson[1] was the first to attain a solution that was adaptable to numerical computation, but his analysis was limited to earths that behaved like a perfect conductor. The extension to the general case including the effect of the dielectric properties of the ground was indicated by T. L. Eckersley[3] in 1932. The complete treatment of the problem, taking account of finite antenna heights, was recently worked out by van der Pol and Bremmer[27,32,33], Wwedensky[16],

and by Eckersley and Millington[31]. These theories proceed from a solution of Maxwell's equations with appropriate boundary conditions. All of them treat only the case of vertical polarization.

Epstein[10] on the other hand, attempted to assess the effect of diffraction on the field beyond the horizon by applying Huyghens' principle, treating the earth as a perfectly absorbing screen. He showed

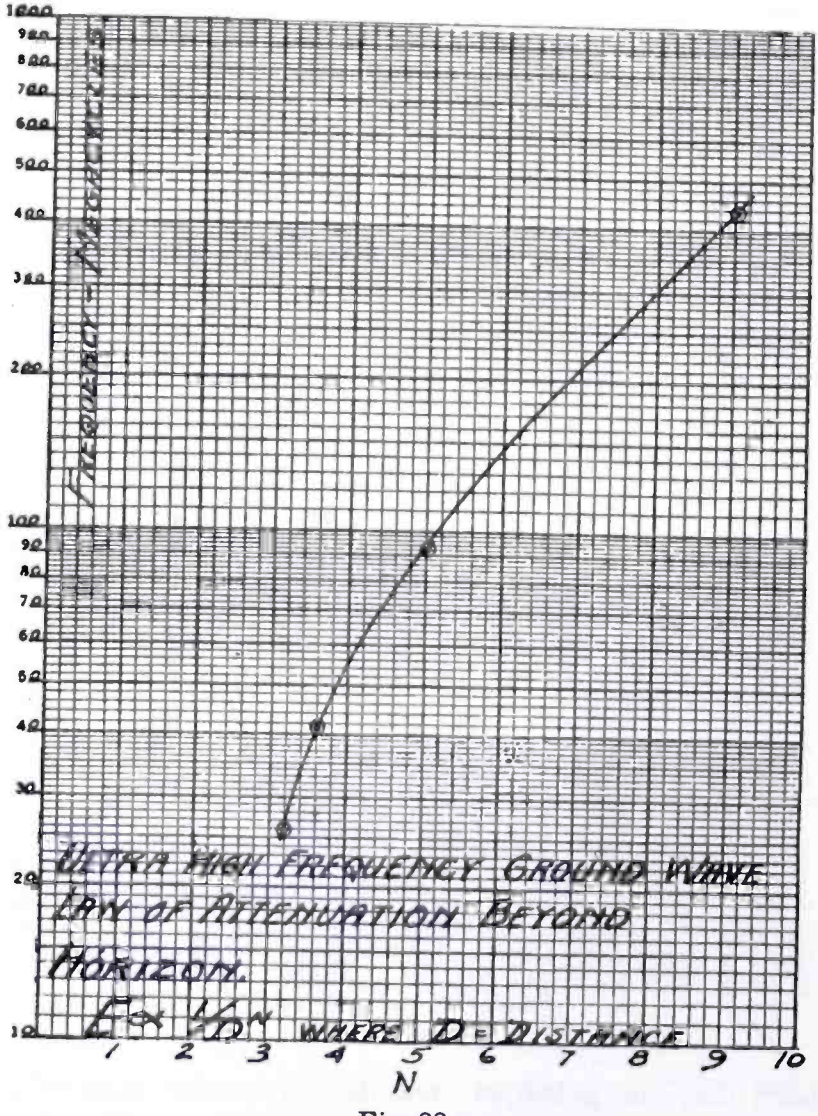


Fig. 22

that the screening effect could be approximated by replacing the earth by a straight edge at the intersection of the planes tangent to the earth from transmitter and receiver. Burrows, Decino, and Hunt[17] showed that it was not possible to overlook the reflected ray components by treating the earth as a perfect absorber, and then proceeded to correct Epstein's method by including the effect of ground reflection. By assuming low antennas, at equal heights, they arrived at the con-

clusion that the field beyond the horizon could be calculated by applying to (4) a correction factor due to diffraction attenuation. This correction factor is the amplitude of the Cornu spiral evolved in optical diffraction theory. For large distances, it varies inversely as the three-halves power of distance, so that, according to their result, the total field should vary inversely as the seven-halves power of distance at large distances, irrespective of frequency.

From a study of experimental data on propagation extending beyond the horizon, Beverage[21] was led to an empirical relation for the variation of field strength with distance. Some of the data that illustrate his method are shown in Figures 19, 20, and 21. Within the

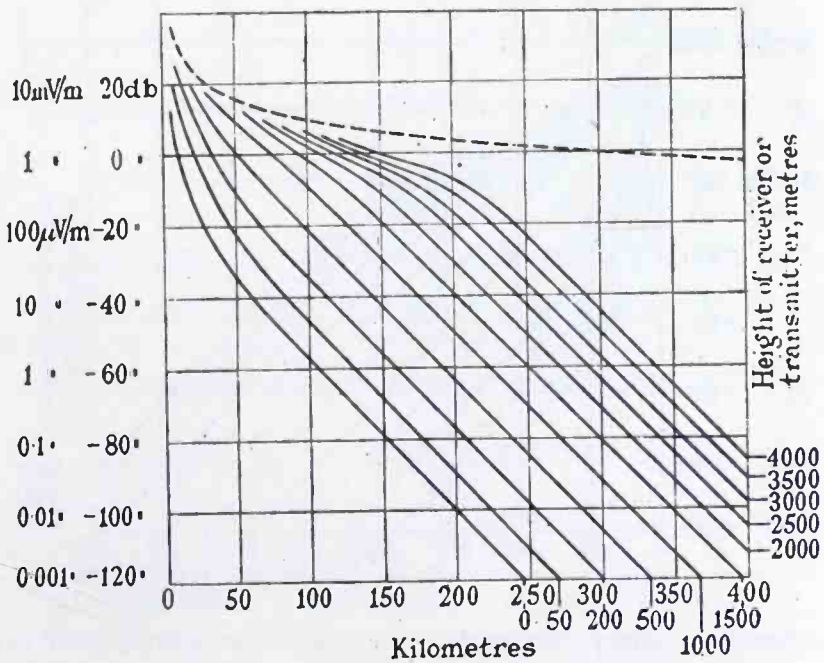


Fig. 23—Results obtained over sea: $\lambda = 6$m, $k = 80$, $\sigma = 10^{-11}$, 1 kK radiated.

horizon of the transmitter, a straight line is drawn on the log-log plot with an inverse-square of distance slope, corresponding to values obtained from (5). At the horizon distance, a second line is drawn for distances beyond the horizon with a slope to fit the data. From data taken at a number of frequencies, Beverage found that the slope of the line drawn to fit the data beyond the horizon increased with frequency. By plotting this slope against frequency, Beverage obtained Figure 22.

This relation, of course, is purely empirical, and hence may not be justified on theoretical grounds. For one thing, the method indicates that at distances beyond the horizon the field strength varies as a

negative power of distance, so that a straight line is obtained when plotted on log-log paper. On the other hand, the theoretical works of Wwedensky[16], van der Pol and Bremmer[33], and Eckersley and Millington[31] all indicate that at such distances the relation of field strength to distance is of exponential form, so that a straight line is obtained by plotting field strength to a logarithmic scale, and distance to a linear scale. This is also the type of relation obtained by von Handel and Pfister[18], although their slopes (exponents of e) are based on Watson's[1] analysis of perfect conductivity.

The theoretical works mentioned above are all in substantial agreement, hence it will be sufficient to consider only one of them. For this

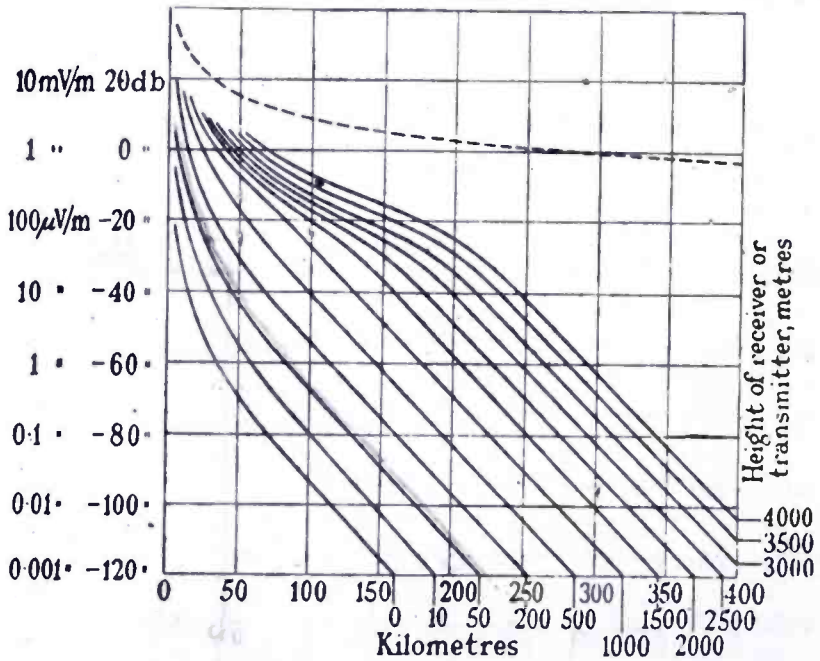


Fig. 24—Results obtained over land: $\lambda = 6$m, $k = 5$, $\sigma = 10^{-13}$, 1 kW radiated.

we select the treatment of Eckersley and Millington[31], since they have pushed their analysis to the most complete numerical results[26].

As mentioned previously, the theoretical solution is obtained in the form of an infinite series. Eckersley's curves for 6-meter (50-megacycle) propagation over sea and land are shown in Figures 23 and 24, respectively. At distances beyond the horizon, the first term of the series is predominant, so that calculation is simplified. In this region it is found that the curves of field strength vs. distance for various heights are parallel, so that at any distance the variation of field strength with height is independent of distance and the earth constants (provided, of course, that we remain in the region where the first term of the series predominates). These curves show the received field strength with the receiver (transmitter) on the ground and the

transmitter (receiver) at the height indicated on the curve, for a radiated power of one kilowatt. To apply these curves to the general case where both transmitter and receiver are elevated, a factor is obtained from Figures 25 or 26 to correct for the height above ground

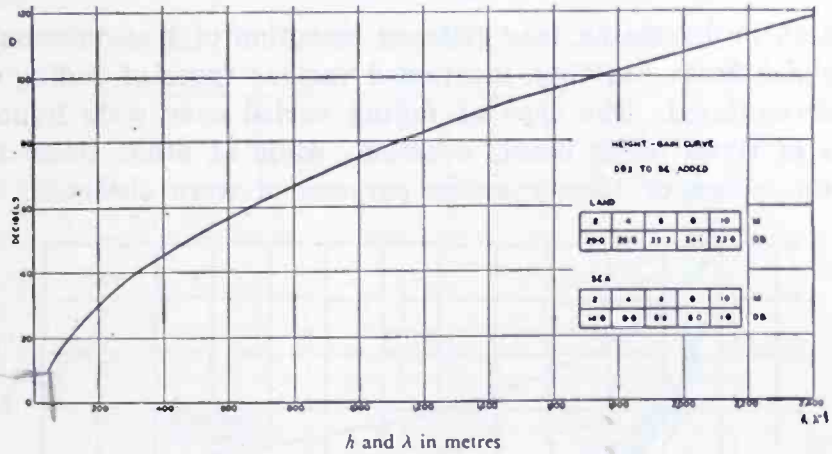


Fig. 25

of the terminal assumed to be on the ground in Figures 23 and 24. Above a certain height, h_o (indicated in Figure 26), the variation of field strength with height is independent of the earth constants, as shown in Figure 25. Below h_o, however, the variation of field strength

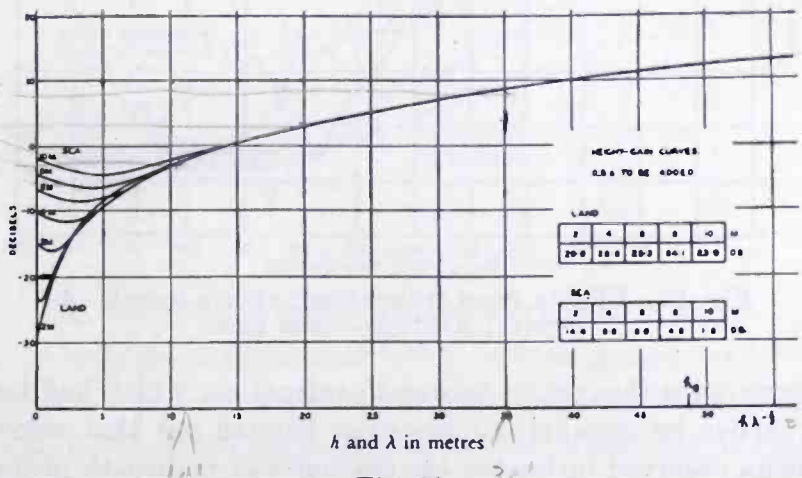


Fig. 26

with height depends on the earth constants and the frequency. This is indicated by the fanning out of the single curve into a family of curves for various frequencies and types of ground.

The above theories, and their numerical results given in the form of curves, yield the value of the field strength beyond the horizon due

to diffraction around the earth's curvature. In addition to diffraction, however, refraction is a mechanism which contributes to the field, especially beyond the horizon.

Refraction

Jones[4], in discussing long-distance reception of transmission from the Empire State Building, mentioned various types of fading which were encountered. The type of fading varied over wide limits, the signals at times being nearly constant, while at other times fading rates up to ten or twenty cycles per second were observed. These

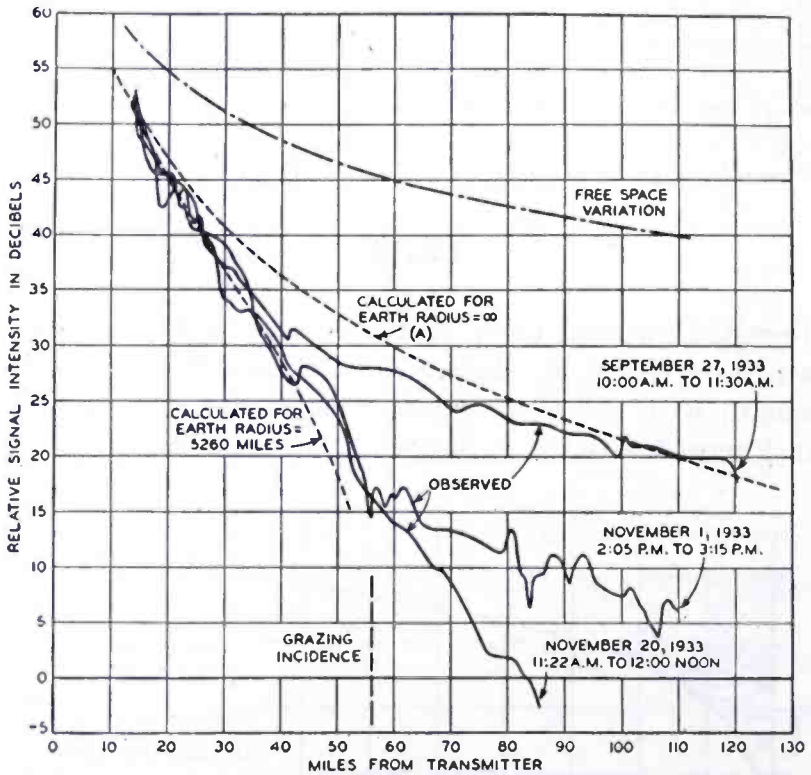


Fig. 27—Flights from transmitter. Wave-length—4.6 meters; Altitude—1000 feet.

phenomena were thought to be due to refraction, which had been suggested earlier by Jouaust[2]. Jones even pointed out that some of the phenomena observed intimated the existence of multipath propagation, which has recently been demonstrated by Englund, Crawford, and Mumford[34]. Jouaust[2], in his early work, pointed out the similarity of refraction of radio waves to that of light waves, due to the varying index of refraction of air with height. He mentioned the great influence of the water vapor content of the atmosphere, and pointed out that optical experiments had demonstrated a seasonal effect, the refraction being greater in summer than in winter. Schelleng, Burrows,

and Ferrell[6] showed that refraction was equivalent to an increase in the earth's radius, and used an increase to four-thirds the actual radius to represent average conditions. Englund, Crawford, and Mumford[15], in a series of airplane flights over ocean water, showed that the effect of refraction varied over a wide range with season.

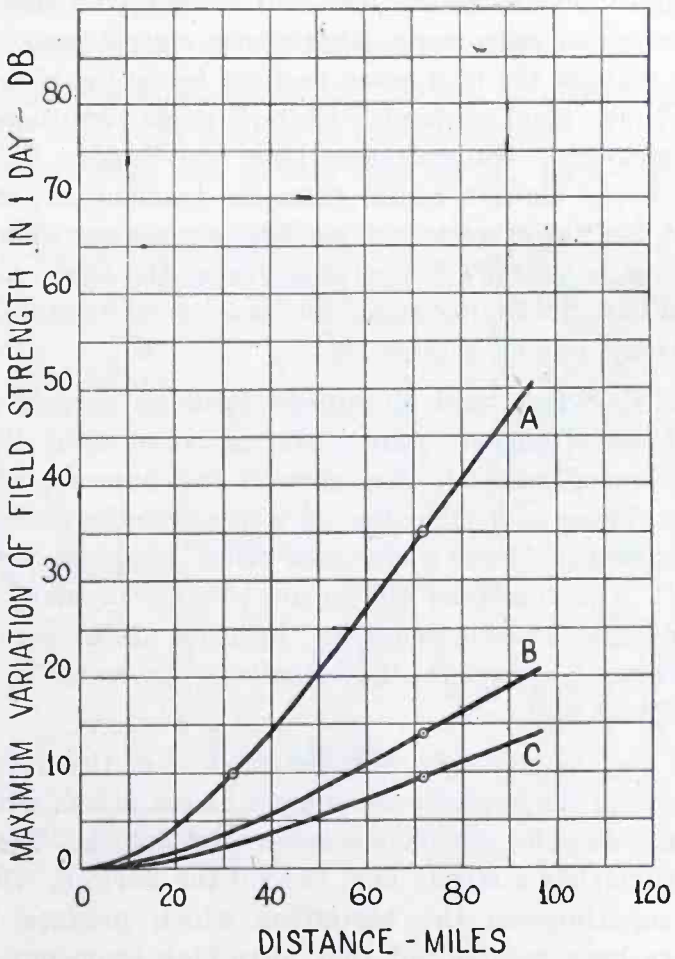


Fig. 28—Field strength variation at various distances.

Curve A—Maximum variation observed in one day.

Curve B—Variation neglecting highest and lowest 5% of time.

Curve C—Variation neglecting highest and lowest 10% of time.

Figure 27, taken from their paper, shows that the received signal at great distances from the transmitter progressively decreased as winter approached. Variation of water vapor content of the air was suggested as a plausible explanation.

Ross Hull[14,25] showed the dependence of the variation of received signal levels over long paths on air-mass conditions, and showed that

strong signals resulted from the presence of temperature inversions in the lower atmosphere. He considered that the rapid types of fading were caused by turbulence in the atmosphere. The general experience is that fading is more pronounced, and signals stronger, at night.

To obtain statistical information on the variations of field strength due to fading, Burrows, Decino, and Hunt[30] analyzed a year's recording of 150-megacycle signals over a slightly non-optical path. They found that fading reduced the field seven decibels below the average one per cent of the time. MacLean and Wickizer[37] made simultaneous recordings of 50-megacycle transmissions from the Empire State Building at three locations along a radial from the transmitter, one near the line of sight, and the other two at successively greater distances. They found that the variation of field strength at the three locations was random, and that the fading range increased with distance. The latter point is brought out by Figure 28.

Englund, Crawford, and Mumford[34] made an extensive investigation of ultra-high-frequency fading over an ocean path. By employing a frequency-sweep method, they proved the presence of multipath propagation. These multiple paths were traced to the normal "ground" ray plus grazing-incidence reflections from air-mass boundaries, or "sky-waves." The combined effects are similar in many respects to fading phenomena experienced near the edge of the service area of broadcast stations, although the so-called "sky-waves" originate in entirely different manners.

Thus, it has been amply demonstrated that the received fields, especially beyond the horizon, are subject to variations with time, due to refraction and reflection in the lower atmosphere. The process of diffraction furnishes a steady field beyond the horizon, while variable refraction superimposes the variations which produce the fading. Some writers have pointed out that ultra-high-frequency signals are received over distances much greater than would be the case if diffraction acted alone, and that, therefore, the effect of diffraction is of no practical importance. This cannot be true, however, as shown by the following argument. It has been shown that the effect of refraction is to increase the effective earth's radius, and therefore, to push the effective horizon of a transmitter farther out. If refraction were the only mechanism of importance, then, for given refraction conditions there should be a distance at which the field strength suddenly sinks to zero, this distance being the effective horizon due to refraction. This does not happen, of course, so that the view we must adopt is that refraction and diffraction are mutually-assisting mechanisms. Refraction may be considered to set up a new horizon for the trans-

mitter, while diffraction carries the signal beyond this new horizon into the "shadow" region.

Due to the variable nature of refraction, however, it is difficult to obtain an experimental check on the magnitude of diffraction attenuation as predicted by theory. A further complication is that the irregular nature of practical types of terrain differs widely from the smooth spherical surface on which the theoretical works are based.

To conclude this summary of the subject of ultra-high-frequency propagation, it may be of interest to show how some of the phenomena discussed above affect the service area of an ultra-high-frequency

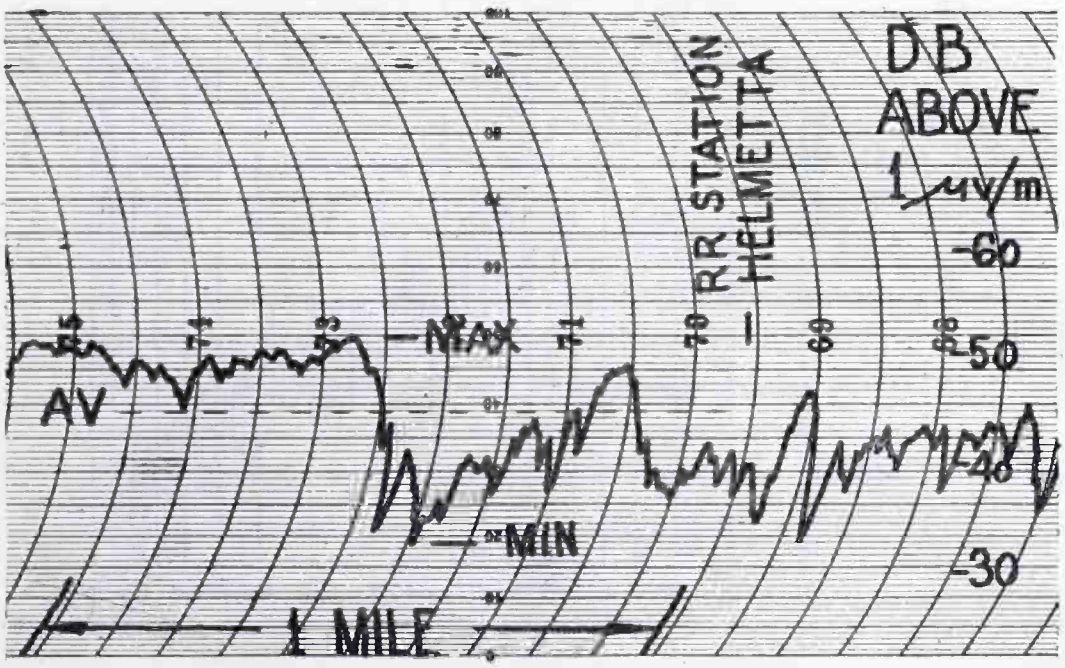


Fig. 29

broadcast station. For this purpose the results of the extensive survey made by Wickizer[38] of the 52.75-megacycle audio channel of the NBC television transmitter on top of the Empire State Building will be used. In this survey, continuous recordings of the signal along a number of radials from the transmitter were made in an automobile. The records showed the usual large variation in amplitude of the received field within short distances caused by buildings and other local objects. A short sample of one of the records, taken in a suburban location, is shown in Figure 29. The records were divided into short sections and the maximum, minimum, and average values for each section determined, as shown on Figure 29. In this way, the data were transferred to the form shown in Figure 30, which shows the

summary of the record taken between New York City and Camden, New Jersey. The extremities of the vertical lines represent maximum and minimum values for each section of the record, while the short horizontal bars through the vertical lines indicate the average values. From the average values, curves such as shown in Figure 31, which corresponds to the values in Figure 30, were then drawn. A set of such curves, showing the variation of average field strengths with distance in various directions from the transmitter, provided data from which equi-signal contours could be drawn. The field-strength

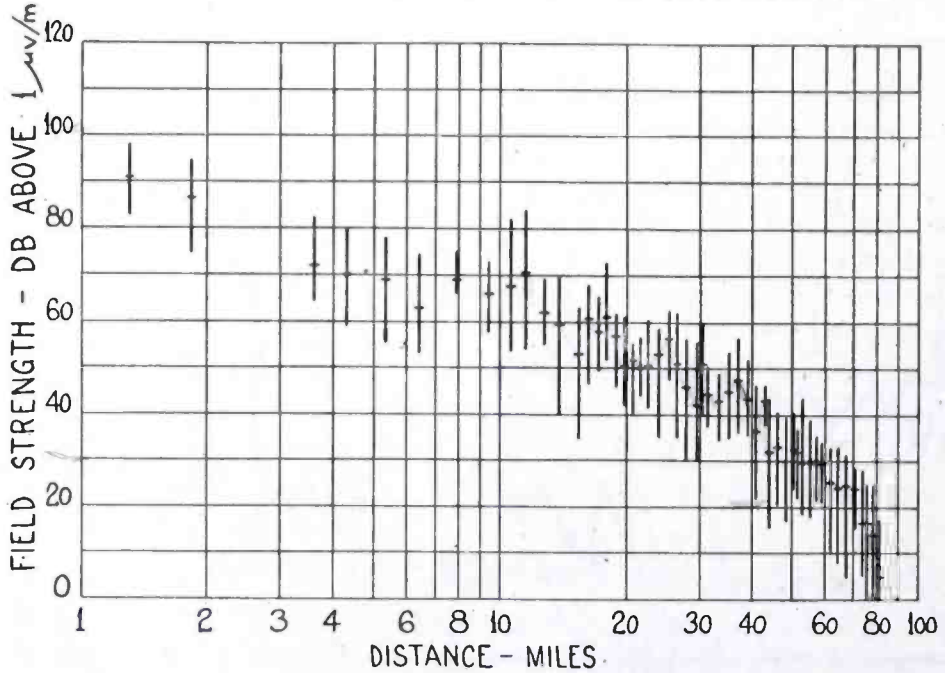


Fig. 30

maps obtained in this manner are shown in Figures 32 and 33, in which the numbers affixed to the contours represent average field strengths in decibels above one microvolt per meter, for a receiving antenna 30 feet high.

These contours represent average field-strength values as measured by the survey car along highways. They do not necessarily represent the field strength that would be obtained at a particular receiving installation. An idea of the order of magnitude of the deviations to be expected from the average values represented by the contours can be obtained from data giving the mean deviations from the average field strengths that occurred. Wickizer found that the mean deviations covered a range of about 20 decibels, roughly 10 decibels above and

below the average values. It will be recalled that Burrows, Hunt, and Decino found a similar range in their Boston survey. As a result of these variations, then, an antenna located on one of the contours might receive a field as little as the next lower contour, or as great as the next higher contour.

Near the horizon, and beyond, refraction in the atmosphere causes a variation in field strength, or fading, which increases at greater distances from the transmitter. From the work of MacLean and

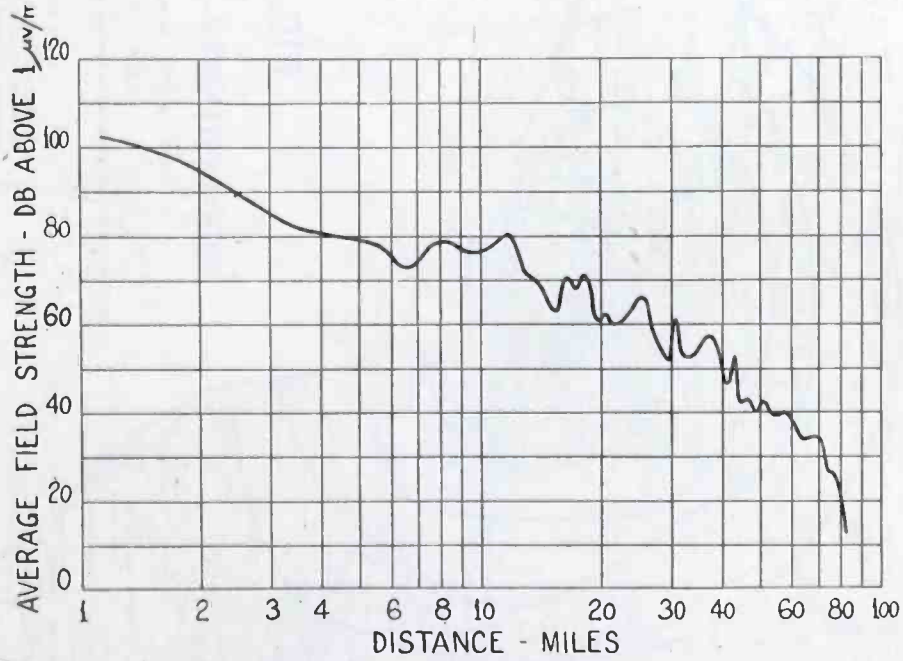


Fig. 31

Wickizer we obtained Figure 28, showing the increase in fading range with distance from the Empire State transmitter. This figure may be used to estimate the additional variation of received field due to fading. For example, the maximum variation, or fading range, at thirty-two miles, which is near the 60-decibel contour was 10 decibels, or the difference between adjacent contours. If the extreme maximum and minimum values are discarded, the variation becomes of considerably smaller range, as shown by curves B and C of Figure 28. These represent the fading ranges which were exceeded less than five per cent and ten per cent of the time, respectively.

Wickizer combined the data showing the variations due to irregular terrain, and those due to fading, and obtained Figure 34, showing the approximate limits of the variations from the field strengths indi-

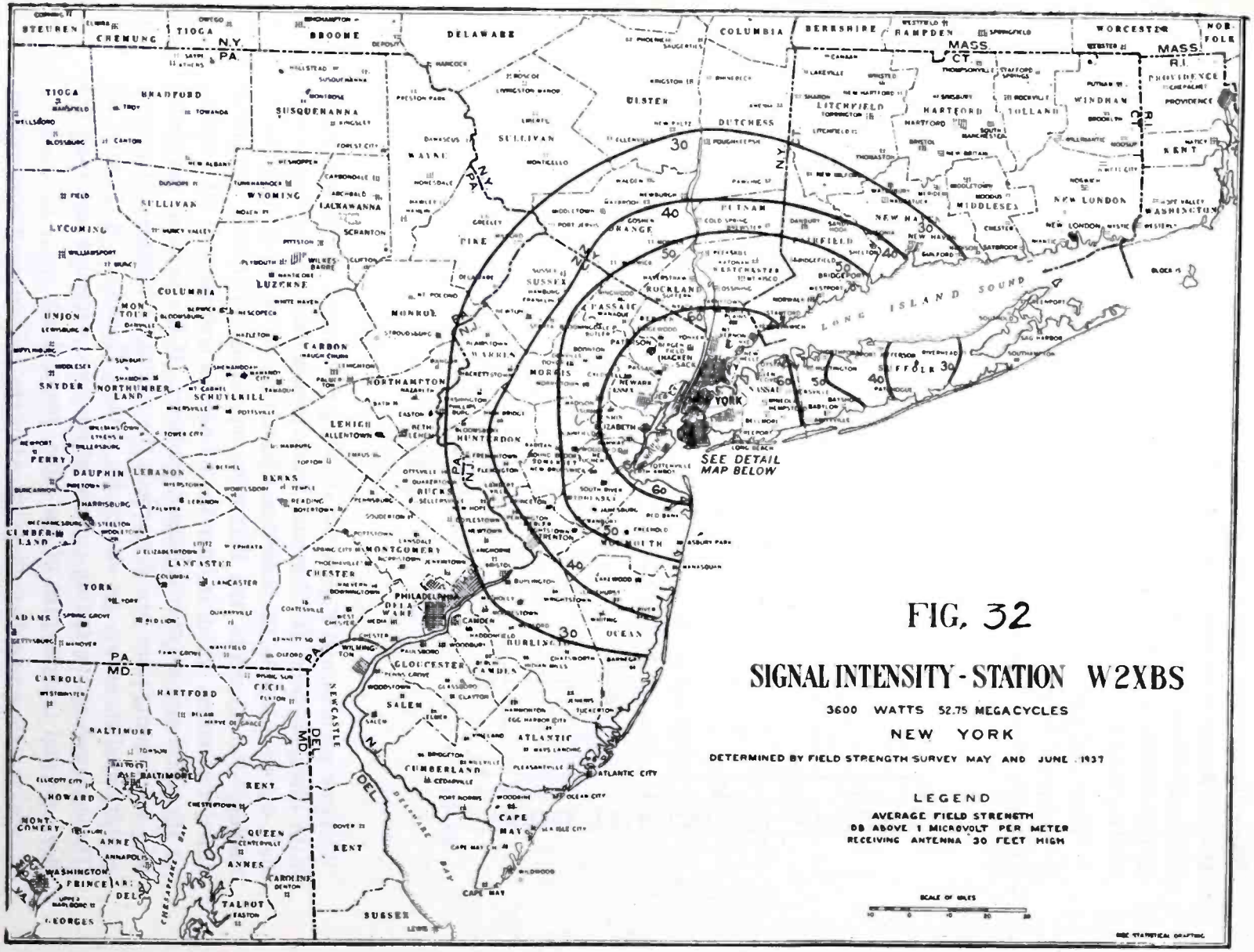


FIG. 32

SIGNAL INTENSITY - STATION W2XBS

3600 WATTS 52.75 MEGACYCLES

NEW YORK

DETERMINED BY FIELD STRENGTH SURVEY MAY AND JUNE 1937

LEGEND

AVERAGE FIELD STRENGTH
DB ABOVE 1 MICROVOLT PER METER
RECEIVING ANTENNA 30 FEET HIGH

SCALE OF MILES

FCC STATISTICAL DRAFTING

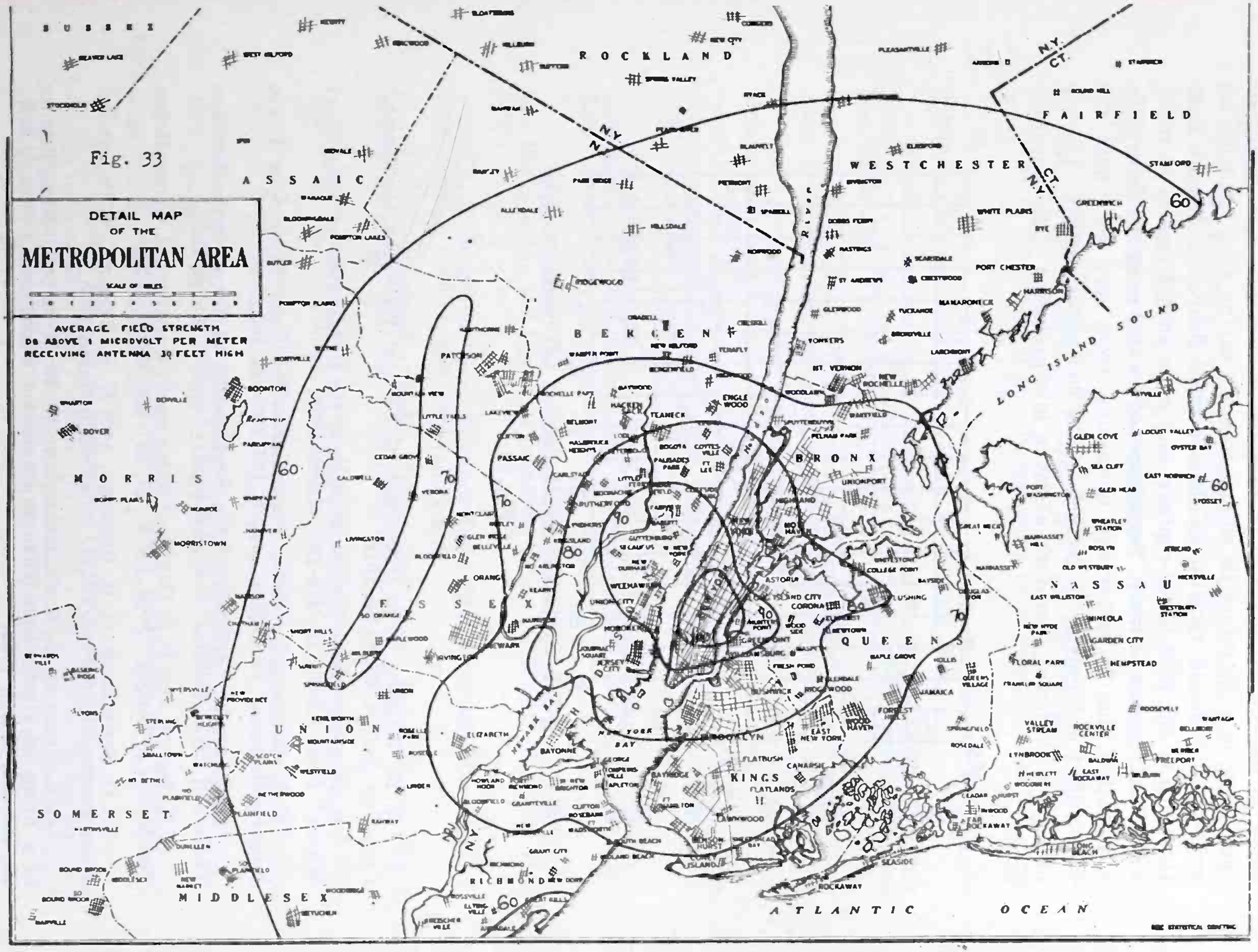


Fig. 33
DETAIL MAP
OF THE
METROPOLITAN AREA
AVERAGE FIELD STRENGTH
DB ABOVE 1 MICROVOLT PER METER
RECEIVING ANTENNA 30 FEET HIGH
SUSSEX
ROCKLAND
FAIRFIELD
WESTCHESTER
PASSAIC
BERGEN
BRONX
MORRIS
ESSEX
QUEENS
NASSAU
UNION
KINGS
SOMERSET
RICHMOND
MIDDLESEX
LONG ISLAND SOUND
ATLANTIC OCEAN
N.Y.
CT.
N.J.
60
70
80
90

Fig. 33

cated on the contour maps to be expected in practice. The results of the survey may therefore be summarized by saying that residents on any contour should receive, as their normal field strength, at least the value of the next lower contour, and some should receive as much as the next higher contour. Superimposed on this normal field will be a variation, due to refraction, increasing from zero very near the transmitter up to about 15 decibels at the 30-decibel contour. The use of Figure 34 should, therefore, assist in interpretation of the coverage maps.

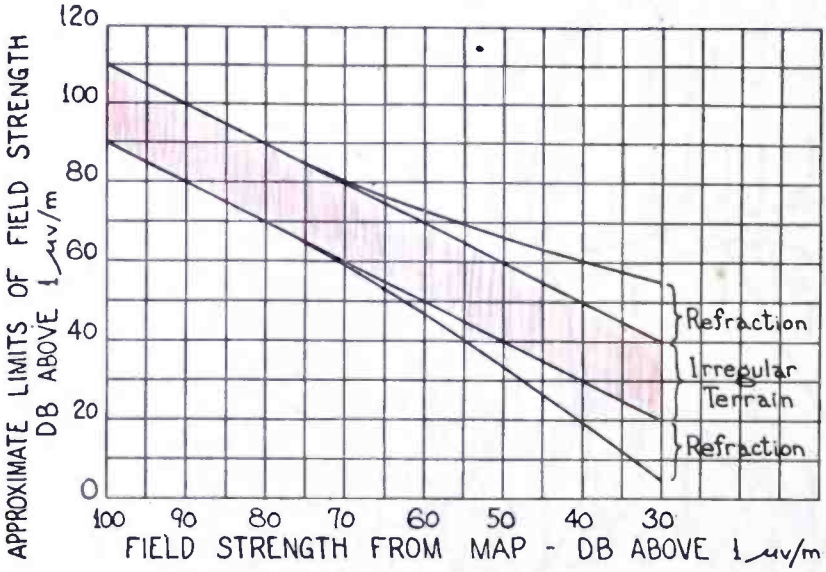


Fig. 34

Bibliography

[1] G. N. Watson, "The Diffraction of Electric Waves by the Earth," *Proc. Roy. Soc.*, A, 95, 83-99, 1918.

[2] R. Jouaust, "Some Details Relating to the Propagation of Very Short Waves," *Proc. I.R.E.*, 19, 479-488, March 1931.

[3] T. L. Eckersley, "Radio Transmission Problems Treated by Phase Integral Methods," *Proc. Roy. Soc.*, A, 136, 499-527, 1932.

[4] L. F. Jones, "A Study of the Propagation of Wavelengths Between Three and Eight Meters," *Proc. I.R.E.*, 21, 349-386, March 1933.

[5] B. Trevor and P. S. Carter, "Notes on Propagation of Waves Below Ten Meters in Length," *Proc. I.R.E.*, 21, 387-426, March 1933.

[6] J. C. Schelleng, C. R. Burrows, and E. B. Ferrell, "Ultra-Short-Wave Propagation," *Proc. I.R.E.*, 21, 427-463, March 1933.

[7] C. R. Englund, A. B. Crawford, and W. W. Mumford, "Some Results of a Study of Ultra-Short-Wave Transmission Phenomena," *Proc. I.R.E.*, 21, 464-492, March 1933.

[8] C. B. Feldman, "Optical Behavior of the Ground for Short Radio Waves," *Proc. I.R.E.*, 21, 764-801, June 1933.

[9] W. H. WISE, "Note on Dipole Radiation Theory," *Physics*, 4, 354-358, Oct. 1933.

[10] P. S. EPSTEIN, "On the Bending of Electromagnetic Micro-Waves Below the Horizon," *Proc. Nat. Acad. Sci.*, 21, 62-68, Jan. 1935.

[11] C. R. BURROWS, L. E. HUNT, AND A. DECINO, "Mobile Urban Ultra-Short-Wave Transmission Characteristics," *Elec. Eng.*, 54, 115-124, Jan. 1935.

[12] B. TREVOR AND R. W. GEORGE, "Notes on Propagation at a Wavelength of Seventy-three Centimeters," *Proc. I.R.E.*, 23, 461-469, May 1935.

[13] C. R. BURROWS, "Radio Propagation Over Spherical Earth," *Proc. I.R.E.*, 23, 470-480, May 1935.

[14] ROSS HULL, "Air-mass Conditions and the Bending of Ultra-High-Frequency Waves," *QST*, 19, pp. 13-18, 74-75, June 1935.

[15] C. R. ENGLUND, A. B. CRAWFORD, AND W. W. MUMFORD, "Further Results of a Study of Ultra-Short-Wave Transmission Phenomena," *Bell Sys. Tech. Jour.*, 14, 369-387, July 1935.

[16] B. WWEDENSKY, "The Diffractive Propagation of Radio Waves," *Tech. Phys. U.S.S.R.*, 2, 624-639, 1935; 3, 915-925, 1936; 4, 579-591, 1937.

[17] C. R. BURROWS, A. DECINO, AND L. E. HUNT, "Ultra-Short-Wave Propagation Over Land," *Proc. I.R.E.*, 23, 1507-1535, Dec. 1935.

[18] P. VON HANDEL AND W. PFISTER, "Ultra-Short-Wave Propagation Along the Curved Earth's Surface," *Hochfreq. u. Elek.*, 47, 182-190, June 1936; *Proc. I.R.E.*, 24, 346-363, March 1937.

[19] P. S. CARTER AND G. S. WICKIZER, "Ultra High Frequency Transmission Between the RCA Building and the Empire State Building in New York City," *Proc. I.R.E.*, 24, 1082-1094, Aug. 1936.

[20] K. A. NORTON, "The Propagation of Radio Waves Over the Surface of the Earth and in the Upper Atmosphere," *Proc. I.R.E.*, 24, 1367-1387, Oct. 1936.

[21] H. H. BEVERAGE, "Some Notes on Ultra High Frequency Propagation, *RCA Review*, 1, 76-87, Jan. 1937.

[22] C. R. BURROWS, "Radio Propagation Over Plane Earth—Field Strength Curves," *Bell Sys. Tech. Jour.*, 16, 45-75, Jan. 1937.

[23] W. H. WISE, "The Physical Reality of Zenneck's Surface Wave," *Bell Sys. Tech. Jour.*, 16, 35-44, Jan. 1937.

[24] C. R. BURROWS, "The Surface Wave in Radio Propagation Over Plane Earth," *Proc. I.R.E.*, 25, 219-229, Feb. 1937.

[25] ROSS HULL, "Air Wave Bending of Ultra-High-Frequency Waves," *QST*, 21, pp. 16-18, 76, 78, 80, 82, May 1937.

[26] T. L. ECKERSLEY, "Ultra-Short-Wave Refraction and Diffraction," *Jour. I.E.E.*, 80, 286-304, 1937.

[27] B. VAN DER POL AND H. BREMMER, "The Diffraction of Electromagnetic Waves from an Electrical Point Source Round a Finitely Conducting Sphere, with Application to Radiotelegraphy and the Theory of the Rainbow," *Phil. Mag.*, 24, 141-176, June 1937; 24, 825-864, Supp. Nov. 1937.

[28] K. A. NORTON, "Physical Reality of Space and Surface Waves in the Radiation Field of Radio Antennas," *Proc. I.R.E.*, 25, 1192-1202, Sept. 1937.

[29] K. A. NORTON, "The Propagation of Radio Waves Over the Surface of the Earth and in the Upper Atmosphere." Part II—"The Propagation From Vertical, Horizontal, and Loop Antennas Over a Plane Earth of Finite Conductivity," *Proc. I.R.E.*, 25, 1203-1236, Sept. 1937.

[30] C. R. BURROWS, A. DECINO, AND L. E. HUNT, "Stability of Two-Meter Waves," *Proc. I.R.E.*, 26, 516-528, May 1938.

[31] T. L. ECKERSLEY AND G. MILLINGTON, "Application of the Phase Integral Method to the Analysis of the Diffraction and Refraction of Wireless Waves Round the Earth," *Phil. Trans. Roy. Soc.*, A, 237, 273-309, June 10, 1938.

[32] B. VAN DER POL AND H. BREMMER, "The Propagation of Radio Waves Over a Finitely Conducting Spherical Earth," *Phil. Mag.*, 25, 817-834, Supp. June 1938.

[33] B. VAN DER POL AND H. BREMMER, "Results of a Theory of the Propagation of Electromagnetic Waves Over a Sphere of Finite Conductivity," *Hochfreq. u. Elek.*, 51, 181-188, June 1938.

[34] C. R. ENGLUND, A. B. CRAWFORD, AND W. W. MUMFORD, "Ultra-Short-Wave Transmission and Atmospheric Irregularities," *Bell Sys. Tech. Jour.*, 17, 489-519, Oct. 1938.

[35] R. W. GEORGE, "A Study of Ultra High Frequency Wide Band Propagation Characteristics," *Proc. I.R.E.*, 27, 28-35, Jan. 1939. (See Page 90.)

[36] R. L. SMITH-ROSE AND A. C. STRICKLAND, "Ultra-Short-Wave Propagation," *Wireless Eng.*, 16, 111-120, March 1939.

[37] K. G. MACLEAN AND G. S. WICKIZER, "Notes on the Random Fading of 50-Megacycle Signals Over Non-optical Paths," presented at U.R.S.I./I.R.E. Meeting, Washington, D. C., April 29, 1938; *Proc. I.R.E.*, 27, 501-506, Aug. 1939.

[38] G. S. WICKIZER, "Field Strength Survey, 52.75 Megacycles from Empire State Building," presented at U.R.S.I./I.R.E. Meeting, Washington, D. C., April 28, 1939.

March 1933 I.R.E.

FREQUENCY MODULATION PROPAGATION CHARACTERISTICS*

By

MURRAY G. CROSBY

R.C.A. Communications, Inc., Riverhead, N. Y.

Summary—*Early work on frequency modulation is described wherein the propagation characteristics of frequency modulation were determined for frequencies between 9000 and 18,000 kilocycles. Oscilloscopic wave form and aural program observations, taken on a circuit between California and New York, showed that frequency modulation is much more distorted by the effects of multipath transmission than is amplitude modulation. The distortion is greatest at the lower modulation frequencies and higher depths of modulation where the side frequencies are most numerous.*

Oscilloscopic observations of the Lissajou figures formed by placing the outputs of receivers connected to spaced antennas on opposite oscilloscope plates showed that the diversity characteristics of frequency modulation are similar to those of amplitude modulation. That is, the detected outputs of the receivers tend to remain in phase for the lower modulation frequencies and become more phase random as the modulation frequency is increased. However, this tendency is almost obliterated on the lower modulation frequencies of frequency modulation by the presence of unequal harmonic distortion in the two receiver outputs.

Theory is given analyzing the distortion encountered in a two-path transmission medium under various path amplitude and phase relation conditions. The theory explains phenomena observed in the tests and points out the extreme distortion that can be encountered.

INTRODUCTION

IN THE annals of radio, most of the work done on frequency modulation has been of a theoretical nature[1,2,3] and very little has been offered as a result of any experimental development.

Considerable theory has also been presented in the consideration of

---

* Decimal classification: R148. Presented as part of a paper on "Propagation and Characteristics of Frequency Modulated Waves," before I.R.E. New York meeting, January 8, 1936.

[1] John R. Carson, "Notes on the Theory of Modulation," *Proc. I.R.E.*, vol. 10, pp. 57-64; February, (1922).

[2] Balth van der Pol, "Frequency Modulation," *Proc. I.R.E.*, vol. 18, pp. 1194-1205; July, (1930).

[3] Hans Roder, "Amplitude, Phase, and Frequency Modulation," *Proc. I.R.E.*, vol. 19, pp. 2145-2176; December, (1931).

Reprinted from *Proc. I.R.E.*, June, 1936.

this type of modulation as a defect of amplitude modulation.[3,4,5,6,7] A more recent contribution[8] includes an experimental investigation checking theory on the detection products produced in a frequency modulation receiver. It is the purpose of this paper to report on development work undertaken by the engineering department of R.C.A. Communications, Inc., wherein the propagation characteristics of frequency modulation were determined for a circuit between California and New York.

Early work done on frequency modulation within the Radio Corporation was that of H. O. Peterson[9] in which a laboratory frequency modulation circuit was set up. Later, in an attempt to reduce fading by frequency diversity, one of the telegraph transmitters on the Argentina and Brazil circuits was frequency modulated. Finally a definite development program was undertaken consisting of the development of a frequency modulation transmitter at the Rocky Point transmitting research and design laboratories and the development of a suitable receiver at the Riverhead receiving research and design laboratories.

After the transmitter and receiver were developed to a sufficient stage to permit a long-distance test to determine the propagation characteristics of frequency modulation, the transmitter was shipped to Bolinas, California, where it was operated by J. W. Conklin who had assisted in its development. The transmissions were observed at the Riverhead station on the receivers developed by the author. The tests were carried out in 1931 during the months from March to June, inclusive.

Propagation Tests

In the tests carried on between Bolinas and Riverhead, two major problems presented themselves for solution. First, how would fre-

---

[4] R. Bown, D. K. Martin, and R. K. Potter, "Some Studies in Radio Broadcast Transmission," *Proc. I.R.E.*, vol. 14, pp. 57-131; February, (1926).

[5] R. K. Potter, "Transmission Characteristics of a Short-Wave Telephone Circuit," *Proc. I.R.E.*, vol. 18, pp. 633-648; April, (1930).

[6] J. C. Schelleng, "Some Problems in Short-Wave Telephone Transmission," *Proc. I.R.E.*, vol. 18, pp. 933-937; June, (1930).

[7] T. L. Eckersley, "Frequency Modulation and Distortion," *Experimental Wireless and The Wireless Engineer*, vol. 7, pp. 482-484; September, (1930). It is interesting to note that in this article Eckersley makes the following prediction from his theory which agrees with the results obtained in the work described in this paper: "It will easily be seen that this results in most appalling distortion, and renders futile any attempt to use even pure frequency modulation in any transmission where appreciable echo delays (of the order of two or three milli-seconds) are present."

[8] J. G. Chaffee, "The Detection of Frequency Modulated Waves," *Proc. I.R.E.*, vol. 23, pp. 517-540; May, (1935).

[9] H. O. Peterson, U. S. Patent No. 1,789,371.

quency modulation withstand the ravages of fading. Second, could the detected outputs of two frequency-modulation receivers, fed by spaced antennas, be added directly[10] in a diversity[11] receiving system; that is, would the detected audio outputs remain in phase so that they could be combined directly, or would some kind of diversity antenna choosing device, as is in present use, be necessary.

In order to solve the problem of the effect of fading on frequency modulation, oscilloscopic wave form and aural program observations were made on the tone and program modulation after it had traversed the radio circuit from California to New York. In this manner the instantaneous harmonic distortion could be observed.

The diversity problem was answered by applying the outputs of two receivers, fed by spaced antennas, to opposite oscilloscope plates. Hence, by proper interpretation of the Lissajous figures obtained, the relative phases and amplitudes of the two outputs could be determined.

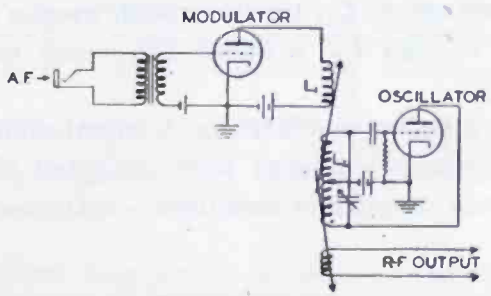


Fig. 1—Circuit diagram of frequency modulated master oscillator.

The Transmitter

To obtain the frequency modulation, a circuit of the type shown in Figure 1 was employed. This type of modulator was used as the master oscillator to replace the crystal oscillator normally used in a standard R.C.A. Communications telegraph transmitter. Special precautions were taken to insure constant element voltages for mean frequency stability. For the same reason, a one-hour warm-up period was observed before each transmission to allow the tube temperatures to settle and prevent drifting. By modulating at master oscillator frequency a frequency multiplication of from eight to sixteen was employed depending upon the radiated frequency. Hence, the frequency deviation, with modulation, was multiplied by a factor of from eight to sixteen and the frequency deviation applied to the oscillator tube was therefore, one eighth to one sixteenth of that radiated.

---

[10] C. W. Hansell, U. S. Patent No. 1,803,504.

[11] H. O. Peterson, H. H. Beverage, and J. B. Moore, "Diversity Telephone Receiving System of R.C.A. Communications, Inc.," *Proc. I.R.E.*, vol. 19, pp. 562-584; April, (1931).

The operation of the modulator of Figure 1 depends upon the variation of the effective inductance of the oscillator tuned circuit inductance, L_2, by virtue of the variation of the modulator tube plate resistance connected across coil L_1 which is coupled to the oscillator inductance, L_2. The modulator tube is biased at such a point that as the modulation is applied at its grid, its plate resistance will vary in

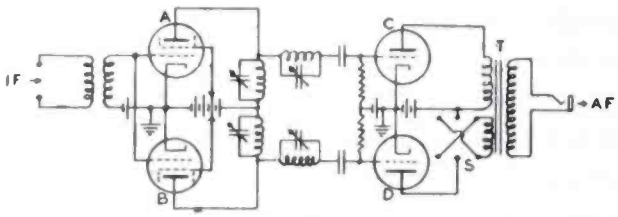


Fig. 2—Circuit diagram of frequency modulation receiver showing the conversion filters and detectors.

accordance with the modulation. By a proper choice of the modulator bias and the inductance of L_1, together with proper adjustment of the coupling between L_1 and L_2, a linear frequency modulation may be obtained.

During the frequency modulation transmission the transmitter amplifiers and frequency doublers were operated class C throughout so as to limit off any amplitude modulation introduced in the modulator or otherwise.

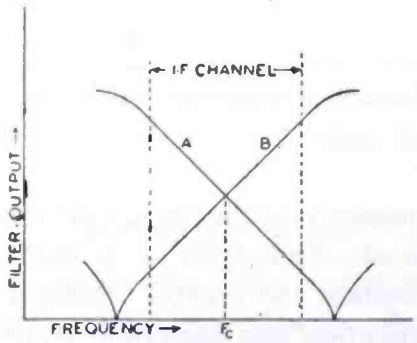


Fig. 3—Characteristics of the conversion filters used to convert frequency modulation into amplitude modulation for detection.

An amplitude modulating stage was include in the transmitter so that the transmitter could be adjusted for amplitude modulation and direct comparison could be made between frequency modulation and amplitude modulation.

The Receiver

A schematic diagram of the frequency modulation receiver is shown in Figure 2. Energy at the superheterodyne intermediate frequency is fed to the two coupling tubes, *A* and *B*, having the conversion filters for converting the frequency modulation into amplitude

modulation in their plate circuits. The characteristics of these two filters are given by the curves *A* and *B* in Figure 3. The carrier frequency F_c, is tuned to the middle of the linear portion of the sloping characteristic formed by the two tuned circuits connected as shown. By adjusting the tuned circuits so that the points of maximum and minimum output fall outside of the intermediate-frequency channel, only the linear portion of the characteristics are utilized. The outputs of these filters are fed to the detectors *C* and *D* of Figure 2 whose audio outputs are combined by the transformer *T* in push-pull or parallel combination, depending upon the switch *S*. With the switch in the push-pull position, frequency modulation may be received and amplitude modulation balanced out. The balance obtainable on amplitude modulation, however, is complete only for the condition of no frequency modulation present. As soon as the carrier is frequency

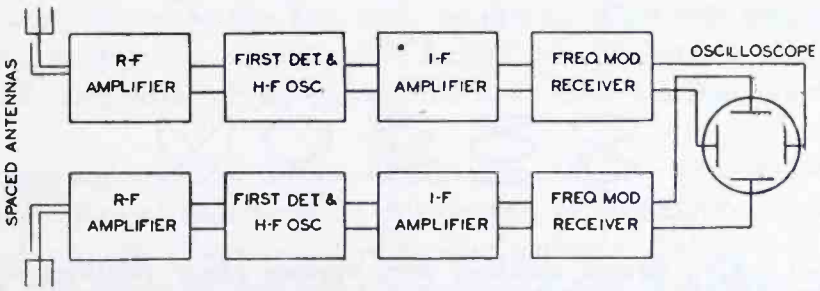


Fig. 4—Block diagram of receiver arrangement used in determining the phase relation between the detected outputs from receivers connected to spaced antennas.

deviated from its unmodulated position, the carriers fed to the detectors are no longer balanced due to the fact that the sloping filters have opposite slopes. Hence the balance is thrown off by an amount proportional to the frequency deviation from the unmodulated carrier position, and the resultant balance is an average between the balanced and unbalanced conditions. With the switch in the push-pull position, second harmonic square-law detector distortion is completely balanced out. With the switch in the parallel position amplitude modulation may be reecived and frequency modulation balanced out.

The intermediate-frequency amplifier of the receiver was adapted from broadcast components and had a twelve-kilocycle pass band. Consequently the maximum frequency deviation was limted to one half of this intermediate frequency channel or six kilocycles. An experimental harmonic analysis of the receiver showed that it was capable of this amount of deviation without serious distortion.

An amplitude limiter was arranged so that it could be switched in and out of the intermediate-frequency circuits. This limiter consisted

of a separate multistage intermediate-frequency amplifier feeding a tube with lowered element voltages so that the tube would be heavily overloaded and its output would be constant regardless of the changes of the input. Thus amplitude modulation could be removed. This method of removing amplitude modulation has an advantage over the balance obtained with the opposite sloping conversion filters in that the amplitude modulation is completely removed in the presence of frequency modulation as well as in the unmodulated condition.

For the diversity experiments, an arrangement as shown in Figure 4 was used. Two complete frequency modulation receivers were fed by two harmonic wire antennas spaced 450 feet apart. The detected

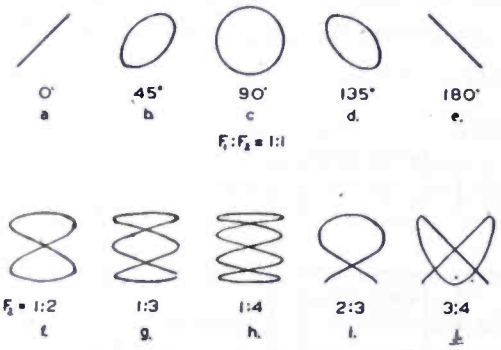


Fig. 5—Lissajou figures obtained with various phase relations and frequency ratios between voltages applied to opposite oscilloscope plates.

outputs of the tone modulation were fed to the opposite oscilloscope plates. Hence, by noting whether the oscilloscope diagram was a line in the second and fourth quadrants, a circle, or a line in the first and third quadrants, the relative phase relation of the two detected outputs could be determined as being in phase, ninety degrees out of phase, or 180 degrees out of phase, respectively. (See Figure 5a, b, c, d, and e.)

For the wave form observations, the output of a single receiver was applied to one set of oscilloscope plates and a timing voltage to the other. Quality distortion was also checked by the aural observation of music which was transmitted by the frequency and amplitude modulation transmitters.

Wave Form Data

Day and night observations were made using music and tone modulation on frequencies of 18,020, 13,690, and 9010 kilocycles. The general conclusions derived were that frequency modulation is far more susceptible to the effects of selective fading than amplitude modulation. On music, the bass and lower notes were distorted and jumbled

together while the violin and the other higher solo instruments came through fairly well. The intelligibility of speech was very low; in some cases a signal which gave fair intelligibility on amplitude modulation was practically unintelligible on frequency modulation. The distortion seemed to be proportional to the depth of modulation, since a sort of blasting effect was observed on the modulation peaks, and transmissions with lowered deviation came through better than those with higher deviation. The distortion also seemed to be greatest during fading minimums at which time "bursts" of distortion would appear; this condition was also present to a lesser extent on amplitude modulation.

The effect due to limiting seemed to be the indication that the distortion was caused by a change in the instantaneous frequency deviation of the wave and not due to an added amplitude modulation. This was evidenced by the fact that the limiter effectively removed all amplitude modulation introduced by fading, but the extreme distortion remained with little or no perceptible difference attributable to the limiter.

The general result of the use of different radiation frequencies over the circuit was to indicate that the quality of modulation was most impaired on the frequencies where the selective fading was most marked. Of the three frequencies observed, 18,020 kilocycles showed the least amount of selective fading, 9010 the greatest, and 13,690 a mean between the other two. Consequently, the quality of frequency modulation on 18,020 kilocycles was only ocasionally inferior to that on amplitude modulation. On 9010 kilocycles there was scarcely an interval of good quality on the frequency modulation, and the amplitude modulation was also considerably distorted. 13,690 kilocycles was considered a representative frequency for the observations, since it presented periods of extremely poor quality along with periods of fairly good quality. Consequently, the oscillograms were taken on the 13,680-kilocycle radiations.

Figure 6 gives the oscillograms showing the effects of distortion fading on wave form. From a comparison of the 300- and 1000-cycle oscillograms taken on frequency modulation, it is apparent that the amount of distortion on 300 cycles is much greater than that on 1000 cycles. At the lower modulation frequencies on frequency modulation, the harmonics were comparable to and stronger than the fundamental for a large percentage of the time. As the modulation frequency was increased, the amount of distortion decreased until at modulation frequencies above about 3000 cycles, little or no harmonic distortion was encountered. One of the reasons for this absence of harmonic distortion on the higher frequencies is probably the fact that the pass band

of the receiver intermediate-frequency amplifier was twelve kilocycles and that of the audio amplifier six kilocycles. Hence, for modulation frequencies above 3000, the side bands above the first were eliminated by the intermediate-frequency filter and the harmonics were eliminated by the audio system. The reason for the excessive distortion at the lower modulation frequencies of frequency modulation is no doubt due to the effects of multipath transmission in disturbing the phase relations between the many side frequencies produced by these lower modulation frequencies.

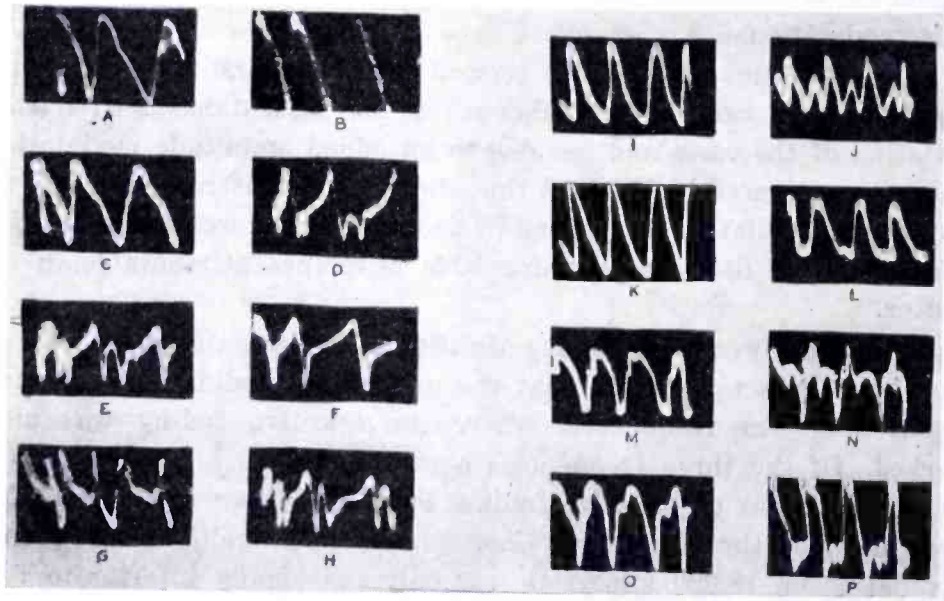


Fig. 6—Frequency and amplitude modulation wave form oscillograms showing the distortion due to fading.

A, B, and C. Amplitude modulation, 300 cycles.
D, E, F, G, and H. Frequency modulation, 300 cycles.
I, J, and K. Amplitude modulation, 1000 cycles.
L, M, N, O, and P. Frequency modulation, 1000 cycles.

Figure 6J shows a wave form taken on amplitude modulation when aural observations indicated that the carrier and not the side bands had faded. The presence of second harmonic is very evident. It has been conjectured that frequency modulation might be less susceptible to the effects of carrier fading because of the fact that the modulated energy is distributed among the side bands to a greater extent; consequently, the detected output is not as dependent for its value upon a single frequency in the transmission medium as is amplitude modulation. The observations did indicate that with frequency modulation, the output was more constant; the quality of this output, however, prevented its constancy from being of any value for the purpose for which it was intended. This observation tends to indicate the advan-

tage of systems wherein telegraph transmitters are frequency modulated to effect a frequency diversity to reduce fading.

Diversity Data

In general the diversity experiments indicated that on both frequency and amplitude modulation the lower modulation frequencies produced outputs from the two receivers which remained in phase. As the modulation frequency was increased, the tendency toward phase differences became very marked until at a 500-cycle modulation fre-

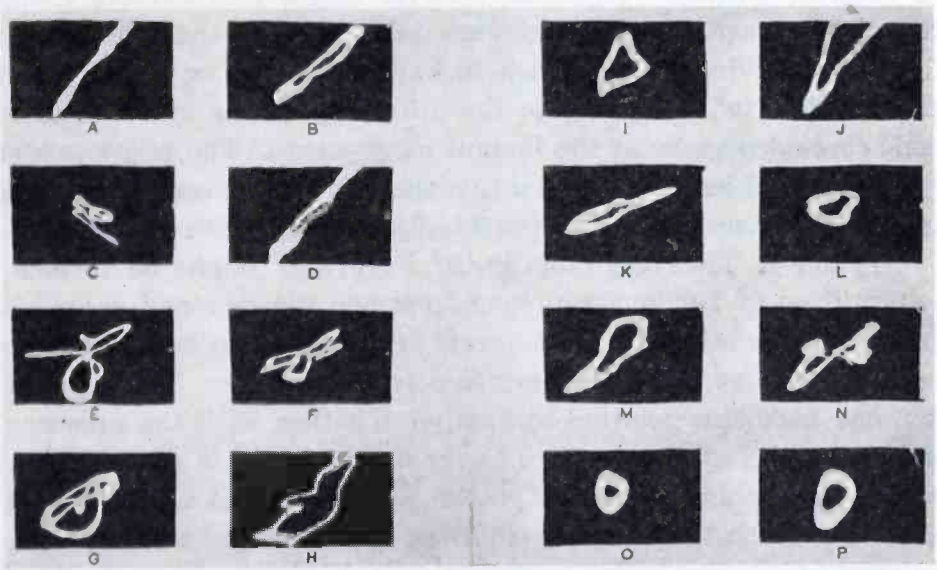


Fig. 7—Oscillograms showing the phase characteristics obtained on frequency and amplitude modulation using the arrangement of Fig. 4.

A, B, and D. Amplitude modulation, 300 cycles.
C, E, F, G, and H. Frequency modulation, 300 cycles.
I and J. Amplitude modulation, 1000 cycles.
K, L, M, and N. Frequency modulation, 1000 cycles.
O and P. Frequency modulation, 5000 cycles.

quency the phase relation was usually random, varying occasionally a full 360 degrees.

The effect of radiation frequency on the phase relations between the two receiver outputs proved to be similar to the effect on wave form. That is, the phase distortion became less as the radiated frequency was increased. Thus on the 18,020-kilocycle radiations there were frequent periods when no phase distortion could be observed in the audio range. On the 13,690-kilocycle radiations the distortion was usually less than ninety degrees for modulation frequencies below 1000 cycles, while for modulation frequencies from 1000 to 5000 cycles marked distortions were apparent. On the 9010-kilocycle radiations, the figures showed random phase variations on all modulation fre-

quencies most of the time. The oscillograms were taken on 13,690 kilocycles with a frequency deviation of about 2000 cycles.

Figure 7 shows the Lissajou figures obtained in the diversity setup. The diagrams for amplitude modulation given in Figure 7A, B, and D for a modulation frequency of 300 cycles show how the two receiver outputs remain in phase on a low-frequency tone. The diagrams obtained on frequency modulation with a 300-cycle modulation frequency, given in Figure 7C, E, F, G, and H, obtain their grotesque distortions from the presence of different harmonic distortions in the two receiver outputs. Thus one receiver might be receiving a signal producing a pure fundamental tone only, whereas the other receiver might be receiving a signal rich in harmonics; consequently, the diagram is distorted according to the difference in the amounts of harmonic content present at the instant of exposure. The effect produced by this unequal harmonic content on the two sets of oscilloscope plates can be seen from a study of the Lissajou figures given in Figure 5f, g, h, i, and j. Thus the diagram of Figure 5f might be formed by the condition of fundamental tone from one receiver and second harmonic from the other, or the diagram of Figure 5i by second harmonic from one receiver and third harmonic from the other. The rapid shift from one harmonic relation to another together with the presence of complex instead of simple wave forms on the two sets of plates makes possible erratic and grotesque patterns.[12] This effect was apparent to such a degree on frequency modulation that difficulty was experienced in determining the phase tendency of the fundamental tones. The effect was only occasionally noticeable on amplitude modulation.

The diagrams given in Figure 7 I and J for a modulation frequency of 1000 cycles, show how the higher modulation frequencies on amplitude modulation are marked by phase differences which are not present on the lower tones. The higher tones on frequency modulation given in Figure 7 K, L, M, and N for a modulation frequency of 1000 cycles, and in Figure 7 O and P for a modulation frequency of 5000 cycles, also indicate marked phase differences and only a slight indication of different harmonic distortions. Amplitude and frequency modulation showed about the same amount of phase distortion on the higher modulation frequencies.

Theory

It is well known that the cause of fading distortion, taking place in the range of frequencies used in these tests, is due to multipath trans-

[12] A more complete set of these Lissajou figures may be found in the book, "High-Frequency Measurements," by A. Hund, pp. 71-75; McGraw-Hill, (1933).

mission in which the signal is conveyed to the receiving antenna by means of more than one path. These paths consist of refractions or reflections from the conducting layers of the ionosphere. Since these conducting layers vary in height and since the signal may follow a course consisting of various numbers of ricochets between the ionosphere and the earth, the signal traveling over the longer path is given a time lag with respect to that traveling over the shorter path. Hence the frequency modulated signal traveling over an earlier path may be expressed by

$$e = E_0 \sin\left(\omega t + \frac{F_d}{F_m}\cos pt\right) \tag{1}$$

where $\omega = 2\pi \times F_c$, F_c = carrier frequency, F_d = peak frequency deviation due to modulation, F_m = modulation frequency, and $p = 2\pi \times F_m$. The signal traveling over a later path may be expressed by

$$e = E_1 \sin\left\{\omega t + \beta + \frac{F_d}{F_m}\cos\,(pt + \alpha)\right\} \tag{2}$$

where β and α take into account the time delay imparted to the carrier wave and the modulation, respectively.

Considering only a two-path case first, the two waves given by (1) and (2) may be combined into a single resultant wave by means of the cosine law resulting in

$$e = \sqrt{E_0^2+E_1^2+2E_0E_1\cos\left\{\frac{F_d}{F_m}\cos(pt+\alpha)+\beta-\frac{F_d}{F_m}\cos pt\right\}}$$

$$\sin\left[\omega t+\frac{F_d}{F_m}\cos pt+\tan^{-1}\frac{\sin\left\{\frac{F_d}{F_m}\cos(pt+\alpha)+\beta-\frac{F_d}{F_m}\cos pt\right\}}{\frac{E_0}{E_1}+\cos\left\{\frac{F_d}{F_m}\cos(pt+\alpha)+\beta-\frac{F_d}{F_m}\cos pt\right\}}\right] \tag{3}$$

Applying the sum and difference formula for the cosine, calling $E_0/E_1 = R$, and $2F_d/F_m \sin \alpha/2 = Z$, gives

$$e = \sqrt{E_0^2+E_1^2+2E_0E_1\cos\left\{z\sin\left(pt+\frac{\alpha}{2}\right)-\beta\right\}}$$

$$\sin\left[\omega t + \frac{F_d}{F_m}\cos pt - \tan^{-1}\frac{\sin\left\{z\sin\left(pt+\frac{\alpha}{2}\right)-\beta\right\}}{R+\cos\left\{z\sin\left(pt+\frac{\alpha}{2}\right)-\beta\right\}}\right]. \quad (4)$$

Since the receiver either limits off or balances out the amplitude modulation on the signal, the amplitude term of (4) is reduced to a constant voltage and the only part requiring investigation is the angle or phase of the wave. To obtain the effective frequency deviation, the instantaneous frequency, or the rate of change of phase of (4) must be determined. This is given by

$$\frac{d\left[\omega t + \frac{F_d}{F_m}\cos pt - \tan^{-1}\frac{\sin\left\{z\sin\left(pt+\frac{\alpha}{2}\right)-\beta\right\}}{R+\cos\left\{z\sin\left(pt+\frac{\alpha}{2}\right)-\beta\right\}}\right]}{\frac{2\pi}{dt}} = f\,(\text{cycles})$$

$$= F_c + F_d\left[\cos pt - \frac{2\sin\frac{\alpha}{2}\cos\left(pt+\frac{\alpha}{2}\right)}{\frac{R+\cos\left\{z\sin\left(pt+\frac{\alpha}{2}\right)-\beta\right\}}{\frac{1}{R}+\cos\left\{z\sin\left(pt+\frac{\alpha}{2}\right)-\beta\right\}}+1}\right]. \quad (5)$$

When $R = 1$, (5) reduces to:

$$f = F_c + F_d\left[\cos pt - \sin\frac{\alpha}{2}\cos\left(pt+\frac{\alpha}{2}\right)\right] \quad (6)$$

which, by application of the cosine law gives

$$f = F_c + F_d \sqrt{1 + \sin^2 \frac{\alpha}{2} - \sin \alpha \cos \left[pt - \tan^{-1} \frac{\sin^2 \frac{\alpha}{2}}{1 - \frac{\sin \alpha}{2}} \right]} \qquad (7)$$

However when $R = 1$, the amplitude term of (4) must be considered since there is a possibility of the wave being 100 per cent amplitude modulated so that the limiter would be unable to hold its output constant and a signal modulated by noise would result. For the case where the amplitude does not go to zero, that is, for conditions in which the quantity $(2\ F_d/F_m \sin \alpha/2 - \beta)$ does not pass through the points π, 3π, 5π, etc., the instantaneous frequency of the wave applied to the frequency modulation sloping filters is given by (7). From this equation it can be seen that the resultant distortion under these conditions is merely a change in the effective frequency deviation of the modulation by a factor depending on α, and the addition of a phase angle depending upon α to the modulation frequency. Thus, provided the amplitude does not go to zero, no harmonic distortion is encountered.

When R is large compared to one, (5) becomes:

$$f = F_c + F_d \cos pt - \frac{2 \sin \frac{\alpha}{2} F_d}{R} \cos \left(pt + \frac{\alpha}{2} \right) \cos \left\{ z \sin \left(pt + \frac{\alpha}{2} \right) - \beta \right\}. \qquad (8)$$

By an application of the addition formulas for the sine and cosine, together with the Bessel function expansions for cos $(x \sin \phi)$ and sin $(x \sin \phi)$ and the recurrence formulas for $J_n(x)$, (8) may be transformed into the following:

$$f = F_c + F_d \cos pt - \frac{2F_m \cos \frac{\alpha}{2}}{R} \left[\cos \beta \left\{ \sum_{n=0}^{\infty} (2n+1) J_{2n+1}(z) \cos (2n+1) \left(pt + \frac{\alpha}{2} \right) \right\} + \sin \beta \left\{ \sum_{n=1}^{\infty} 2n J_{2n} \sin 2n \left(pt + \frac{\alpha}{2} \right) \right\} \right]. \qquad (9)$$

Thus when β is zero, a distortion consisting of fundamental and odd harmonics is added to the frequency deviation originally present.

These harmonics are proportional to the ratio between the weaker and the stronger of the signals from the two paths, the modulation frequency, and $nJ_n(2F_d/F_m \sin \alpha/2)$ where n is the order of the harmonic. When β is ninety degrees, the distortion consists of even harmonics proportional to the same factors.

By far the most destructive distortion occurs when R is not large compared to unity, or is not equal to unity, but is slightly greater than unity. Figure 8 shows wave forms calculated from the part of (5) in the square brackets. These wave forms are for a two-path case in which the ratio between the amplitudes arriving over the two paths is 1.2:1. Figures 8(A), (B), and (C) show the effect of various phase differences between the two radio-frequency carriers with a constant phase difference of ninety degrees between the modulation frequencies modulating the two carriers and a constant ratio of F_d/F_m of ten. It is quite apparent that the distortion is greatest when the radio-frequency carriers are 180 degrees out of phase. This checks the observations that the distortion seemed to be greatest during the fading minimums. With the two radio-frequency carriers out of phase, the resultant unmodulated amplitude would be at a minimum, producing a fade in the signal. Figures 8(C) and (D) show the effect of the value of F_d/F_m on the wave form. The increase of the distortion with increase in F_d/F_m agrees with the observations that the lower modulation frequencies were the most distorted and that the distortion seemed to be proportional to the depth of modulation. Thus for modulation frequencies greater than one half the maximum modulation frequency where F_d/F_m can not be greater than two, a wave form similar to Figure 8(D) would be obtained. For the lower modulation frequencies, for instance, 200 cycles, the wave form of Figure 8(D) would exist for a deviation of 282.4 cycles and that of Figure 8(C) for a deviation of 2000 cycles; hence the distortion increases with increase in depth of modulation.

From a study of (5), it can be seen that the effect of a change in the value of the phase difference, α, between the modulation frequencies on the two paths is approximately the same as a change in the value of F_d/F_m since sin $\alpha/2$ enters in the quantity Z in the same manner as does the quantity F_d/F_m. However it can also be seen that the distortion is zero when α is equal to 0, 2π, 4π, etc. Thus the most severe distortion occurs when $\alpha = \pi$ radians or 180 degrees. Since $\alpha = 2\pi DF_m$, where $D =$ time delay between the two paths, for a given time delay sin $\alpha/2$ will go through maximums and minimums as the modulation frequency is varied. Hence bands of maximum and minimum distortion would be expected throughout the modulation frequency band.

By an application of the same theoretical method, the distortion due to three or more paths could be determined. The addition of a wave arriving over the third path to the resultant given by (4) would add another arc-tangent term to the phase of the resultant in which R would be equal to the ratio between the resultant amplitude of the first two paths and the amplitude of the third. This would add distortion

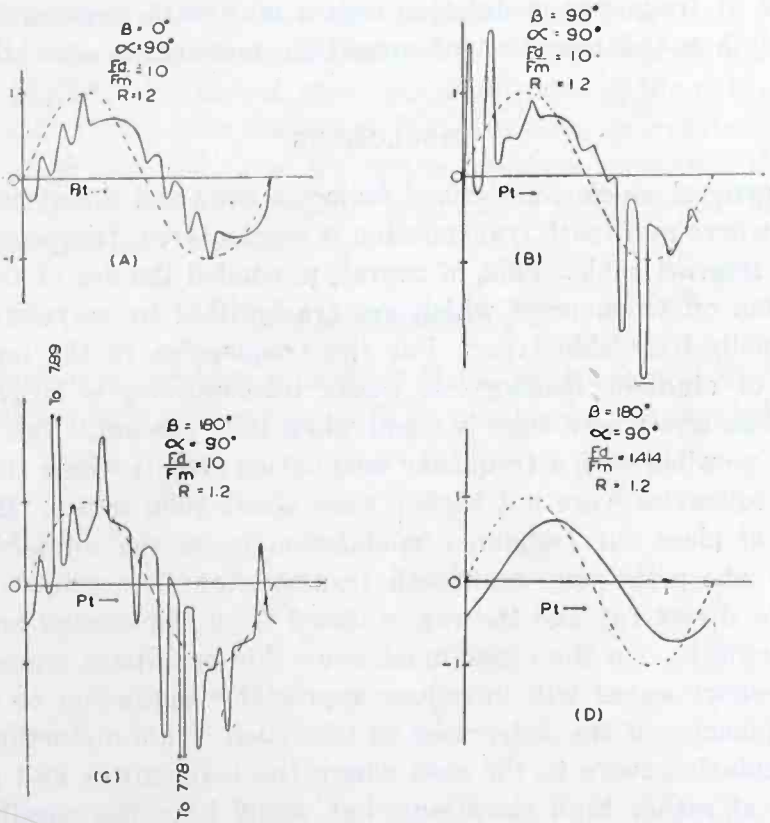


Fig. 8—Calculated wave forms for frequency modulation transmitted over a two-path medium plotted from (6). Ratio between path amplitudes, $R = 1.2$. F_d = frequency deviation. F_m = modulation frequency. α = phase difference between modulation frequencies. β = phase difference between carrier frequencies. The dotted line curve is a pure sine wave (representing no distortion) for comparison purposes.

which would result in a wave form undoubtedly more distorted than the two-path case. However, the degree of distortion would be proportional to the same parameters as the two-path case since the distortion term added by each path is similar to that produced by the first two paths.

The distortion encountered by amplitude modulation in multipath transmission has been theoretically considered in previous literature.[13]

[13] Charles B. Aiken, "Theory of the Detection of Two Modulated Waves by a Linear Rectifier," *Proc. I.R.E.*, vol. 21, pp. 601-629; April, (1933). The "Case of Identical Modulating Frequencies" considered on page 616 covers exactly the same situation as the two-path multipath transmission case.

The general results of this theory are that the distortion consists of a change of the amplitude of the fundamental modulation frequency together with the introduction of second harmonic distortion. However, the higher order harmonic distortion so prone to appear in frequency modulation, and depending upon the ratio F_d/F_m, is not present. It is this extreme high order harmonic distortion which makes the reception of frequency modulation over a multipath medium far more distorted than the reception of amplitude modulation over the same medium.

Conclusions

The general conclusion derived from the tests and theory is that on circuits where multipath transmission is encountered, frequency modulation is impracticable. This, of course, precluded the use of frequency modulation on frequencies which are transmitted by refraction from the Kennelly-Heaviside layer. For the frequencies in the immediate vicinity of eighteen megacycles, where transmission is largely by a single path, and where echo is small when it is present,[14] fair success would be possible with a frequency modulation circuit where the modulation frequencies were not higher than about 5000 cycles. However, the logical place for frequency modulation is on the ultra-high frequencies where the only multipath transmission that exists[15] is that due to the direct ray and the ray reflected from the ground and other near-by objects. On the other hand, even this multipath transmission on ultra-short waves will introduce appreciable distortion to modulation frequencies of the order used in television. This distortion would be particularly severe in the case where the transmitter and receiver are both at rather high elevations, but would have the possibility of elimination by the use of directivity to discriminate against one of the rays.

The specific conclusions of the tests may be summarized as follows: The distortion is most severe on the lower modulation frequencies and on the lower radiation frequencies. The distortion was equally severe with or without limiting. The diversity tests showed a tendency for the two receivers to stay in phase at the lower modulation frequencies, but this tendency was almost obliterated by unequal harmonic distortion on the two receivers. The higher modulation frequencies and the lower radiation frequencies showed the most random phase characteristics.

---

[14] T. L. Eckersley, "Multiple Signals in Short-Wave Transmission," *Proc. I.R.E.*, vol. 18, pp. 106-122; January, (1930).

[15] Bertram Trevor and P. S. Carter, "Notes on Propagation of Waves Below Ten Meters in Length," *Proc. I.R.E.*, vol. 21, pp. 387-426; March, (1933).

The specific conclusions of the theory concerning the two-path case may be summarized as follows: When the two paths are exactly equal in amplitude, provided conditions are such that the resultant amplitude does not go to zero, the effect is a change in the effective frequency deviation with no harmonic distortion. When the strength of one path is great compared to the other, odd harmonics are introduced when the phase difference between the carriers is zero and even harmonics are introduced when the phase difference is ninety degrees. The most destructive distortion occurs when the amplitudes of the paths are not quite equal. Under this condition the distortion is greatest for the higher values of F_d/F_m and for phase relations such that the phase differences between the modulation frequencies and between the carrier frequencies are both 180 degrees.

Acknowledgment

The author is indebted to Messrs. H. H. Beverage and H. O. Peterson under whose guidance the work of this paper was carried out.

FREQUENCY MODULATION NOISE CHARACTERISTICS*

By

MURRAY G. CROSBY

R.C.A. Communications, Inc., Riverhead, L. I., New York

Summary—Theory and experimental data are given which show the improvements in signal-noise ratio effected by frequency modulation over amplitude modulation. It is shown that above a certain carrier-noise ratio in the frequency modulation receiver which is called the "improvement threshold," the frequency modulation signal-noise ratio is greater than the amplitude modulation signal-noise ratio by a factor equal to the product of a constant and the deviation ratio (the deviation ratio is equal to the ratio between the maximum frequency deviation and the audio modulation band width). The constant depends upon the type of noise, being slightly greater for impulse than for fluctuation noise. In frequency modulation systems with high deviation ratios, a higher carrier level is required to reach the improvement threshold than is required in systems with low deviation ratios; this carrier level is higher for impulse than for fluctuation noise. At carrier-noise ratios below the improvement threshold, the peak signal-noise ratio characteristics of the frequency modulation receiver are approximately the same as those of the amplitude modulation receiver, but the energy content of the frequency modulation noise is reduced.

An effect which is called "frequency limiting" is pointed out in which the peak value of the noise is limited to a value not greater than the peak value of the signal. With impulse noise this phenomenon effects a noise suppression in a manner similar to that in the recent circuits for reducing impulse noise which is stronger than the carrier in amplitude modulation reception.

When the power gain obtainable in certain types of transmitters by the use of frequency modulation is taken into account, the frequency modulation improvement factors are increased and the improvement threshold is lowered with respect to the carrier-noise ratio existing in a reference amplitude modulation system.

INTRODUCTION

IN A previously published paper,[1] the propagation characteristics of frequency modulation were considered. Prior to, and during these propagation tests, signal-noise ratio improvements effected by frequency modulation were observed. These observations were made at an early stage of the development work and were investigated by experimental and theoretical methods.

---

* Decimal classification: R148 × R270. Presented as part of a paper on "Propagation and Characteristics of Frequency Modulated Waves," before I.R.E. New York meeting, January 8, 1936. Revised and presented in full before Chicago Section, September 11, 1936.

[1] Murray G. Crosby, "Frequency Modulation Propagation Characteristics," *Proc. I.R.E.*, vol. 24, pp. 898-913; June, (1936). (See Page 143.)

Reprinted from *Proc. I.R.E.*, April, 1937.

It is the purpose of this paper to consider that phase of the frequency modulation development work by R.C.A. Communications, Inc., in which the signal-noise characteristics of frequency modulation are studied. The theory and experimental work consider the known systems of frequency modulation including that independently developed by E. H. Armstrong.[2]

Table of Symbols

C = carrier peak voltage.

C/N = theory: Ratio between peak voltage of carrier and instantaneous peak voltage of the noise in the frequency modulation receiver. Experiment: Ratio between peak voltage of carrier and *maximum* instantaneous peak voltage of the noise.

C/n = ratio between the peak voltage of the carrier and the peak voltage of the noise component.

C_a/N_a = carrier-noise ratio in the amplitude modulation receiver.

F_a = maximum audio frequency of modulation band.

F_c = carrier frequency.

F_d = peak frequency deviation due to applied modulation.

F_{dn} = peak frequency deviation of the noise.

F_1 = intermediate-frequency channel width.

F_m = modulation frequency.

F_n = frequency of noise resultant or component.

F_d/F_a = deviation ratio.

K = slope filter conversion efficiency.

M = modulation factor of the amplitude modulated carrier.

M_f = modulation factor at the output of the sloping filter.

M_{fn} = modulation factor at the output of the sloping filter when noise modulates the carrier.

N = instantaneous peak voltage of the noise.

n = peak voltage of the noise component.

N_a = noise peak or root-mean-square voltage at amplitude modulation receiver output.

N_f = noise peak or root-mean-square voltage at frequency modulation receiver output.

$p = 2\pi F_m$.

S_a = signal peak or root-mean-square voltage at amplitude modulation receiver output.

S_f = signal peak or root-mean-square voltage at frequency modulation receiver output.

$\omega = 2\pi F_c$.

---

[2] Edwin H. Armstrong, "A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation," *Proc. I.R.E.*, vol. 24, pp. 689-740; May, (1936).

$\omega_n = 2\pi F_n$.

$\omega_{na} = (\omega - \omega_n) = 2\pi (F_c - F_n) = 2\pi F_{na}$.

$Z = C/n + n/C$.

$\phi(t)$ = phase variation of noise resultant as a function of time.

THEORY

In the following analysis, frequency modulation is studied by comparing it with the familiar system of amplitude modulation. In order to do this, the characteristics of frequency modulation reception are analyzed so as to make possible the calculation of the signal-noise ratio improvement effected by frequency modulation over amplitude modulation at various carrier-noise ratios.[3] The amplitude modulation standard of comparison consists of a double side-band system having the same audio modulation band as the frequency modulation system and producing the same carrier at the receiver. Differences in transmitter power gain due to frequency modulation are then considered separately. The frequency modulation reception process is analyzed

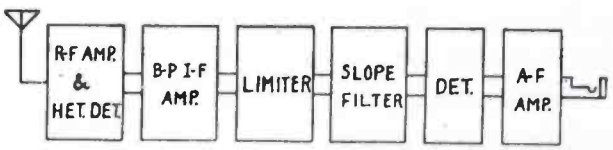


Fig. 1—Block diagram of a frequency modulation receiver.

by first considering the components of the receiver and the manner in which they convert the frequency modulated signal and noise spectrum into an output signal-noise ratio.

The Frequency Modulation Receiver

The customary circuit arrangement used for the reception of frequency modulation is shown in the block diagram of Figure 1. The intermediate-frequency output of a superheterodyne receiver is fed through a limiter to a slope filter or conversion circuit which converts the frequency modulation into amplitude modulation. This amplitude modulation is then detected in the conventional amplitude modulation manner. The audio-frequency amplifier is designed to amplify only the modulation frequencies; hence it acts as a low-pass filter which rejects noise frequencies higher than the maximum modulation frequency.

---

[3] Throughout this paper, carrier-noise ratio will refer to the ratio measured at the output of the intermediate-frequency channel. Signal-noise ratio will refer to that measured at the output of the receiver and will depend upon the depth of modulation as well as upon the carrier strength.

The purpose of the limiter is to remove unwanted amplitude modulation so that only the frequency modulation component of the signal will be received. It may take the form of an overloaded amplifier tube whose output cannot rise above a certain level regardless of the input. Care must also be exercised to insure that the output of the overloaded amplifier does not fall off as the input is increased since this would introduce amplitude modulation of reverse phase, but of equally undesirable character.

The main requirement of the conversion circuit for converting the frequency modulation into amplitude modulation is that it slope linearly from a low value of output at one side of the intermediate-frequency channel to a high value at the other side of the channel. To do this, an off-tuned resonant circuit or a portion of the characteristic of one of the many forms of wave filters may be utilized. The ideal slope

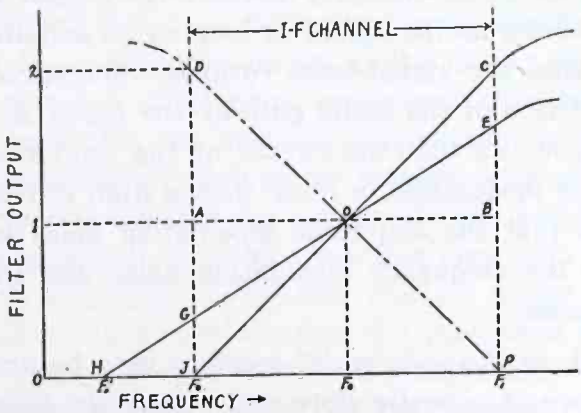


Fig. 2—Ideal sloping filter characteristics.

filter would be one which gave zero output at one side of the channel, an output of one voltage unit at carrier frequency, and an output of two units at the other side of the channel. Such a characteristic is given by the curve *JOC* of Figure 2. From this curve it is easily seen that if the frequency is swung between the limits F_0 and F_1, about the mean frequency F_c, the output of the filter will have an amplitude modulation factor of unity. The modulation factor for a frequency deviation, F_d, will be given by

$$M_f = \frac{F_d}{(F_1 - F_c)} = \frac{F_d}{(F_c - F_0)} = \frac{2F_d}{F_i} \qquad (1)$$

where F_i = intermediate-frequency channel width.

When the converting filter departs from the ideal characteristic in the manner of the filter of curve *HGOE* of Figure 2, the modulation factor produced by a given frequency deviation is reduced by a factor

equal to the ratio between the distances AG and AJ or BE and BC. A convenient term for this reduction factor of the filter is "conversion efficiency" of the filter. Taking into account this conversion efficiency, the modulation factor for a frequency deviation F_d becomes

$$M_f = \frac{2KF_d}{F_1} \tag{2}$$

where $K = AG/AJ = BE/BC =$ conversion efficiency of the filter.

A low conversion efficiency may be used as long as the degree of limiting is high enough to reduce the amplitude modulation well below the level of the converted frequency modulation. This is true since lowering the conversion efficiency reduces the output of the noise in the same proportion as the signal as long as no amplitude modulation is present. Hence the signal-noise *ratio* is unimpaired and the only effect is a reduction of the audio gain by the factor K. If insufficient limiting is applied so that the output of the limiter contains appreciable amplitude modulation, a filter with a high conversion efficiency is desirable so that the amplitude modulation noise will not become comparable to the frequency modulation noise and thereby increase the resultant noise.

A push-pull, or "back-to-back" receiver may be arranged by providing two filters of opposite slope and separately detecting and combining the detected outputs in push-pull so as to combine the audio outputs in phase. Another slope filter having a characteristic as shown by the dot-dash line DOP in Figure 2 would then be required.

A further type of receiver in which amplitude modulation is also balanced out may be arranged by making one of the slope filter circuits of the above-mentioned back-to-back type of receiver a flat-top circuit for the detection of amplitude modulation only. The sloping filter channel then detects both frequency and amplitude modulation; the flat-top channel detects only amplitude modulation. When these two detected outputs are combined in push-pull, the amplitude modulation is balanced out and the frequency modulation is received. This type of detection, as well as that in which opposite slope filters are used, has the limitation that the balance is partially destroyed as modulation is applied. However, if a limiter is used, the amplitude modulation is sufficiently reduced before the energy reaches the slope filters; consequently, for purposes of removing amplitude modulation, the balancing feature is not a necessity.

Noise Spectrum Analysis

The first step in the procedure to be followed here in determining the noise characteristics of the frequency modulation receiver will be to determine mathematically the fidelity with which the noise is transmitted from the radio-frequency branch, in which it originates, to the measuring instrument as a function of frequency. To do this, the waves present at the receiver input will be assumed to be the frequency modulated carrier and the spectrum of noise voltages. This wave and spectrum will be combined into a single resultant whose amplitude and phase are functions of the constants of the component waves. The resultant will then be "mathematically" passed through the limiter to remove the amplitude modulation. From a determination of the instantaneous frequency of the resultant, the peak frequency deviation effected by the noise will be found. A single noise component of arbitrary frequency will then be substituted for the resultant of the noise spectrum, and the modulation factor at the output of the converting filter will be obtained. This noise component will then be varied in frequency to determine the over-all transmission of the receiver in terms of the modulation factor at the sloping filter output. The area under the curve representing the square of this over-all transmission will then be determined. By comparing this area with the corresponding area for an amplitude modulation receiver under equivalent conditions, and taking into consideration the pass band of the intermediate- and audio-frequency channels, a comparison will be obtained between the average noise powers, or the average root-mean-square noise voltages from the two receivers.[4]

The peak voltage characteristics of the two receivers will be compared for fluctuation noise by a correlation of known crest factors with the root-mean-square characteristics. (Crest factor = ratio between the peak and root-mean-square voltages.) The peak voltage characteristics of impulse noise will be determined by a separate consideration of the effect of the frequency modulation over-all transmissions upon the peak voltage of this type of noise.

After a comparison between the noise output voltages from the frequency and amplitude modulation receivers has been obtained, the respective signal output voltages will be taken into consideration so that the improvement in signal-noise ratio may be determined.

---

[4] Stuart Ballantine, "Fluctuation Noise in Radio Receivers," *Proc. I.R.E.*, vol. 18, pp. 1377-1387; August, (1930). In this paper, Ballantine shows that the average value of the square of the noise voltage " . . . is proportional to the area under the curve representing the square of the over-all transimpedance (or of the transmission) from the radio-frequency branch in which the disturbance originates to the measuring instrument as a function of frequency. . . . "

In the process of determining the over-all transmission of the noise, the frequency modulated wave may be expressed by

$$e_s = C \sin \{\omega t + (F_d/F_m) \cos pt\} \tag{3}$$

where C = carrier peak voltage, $\omega = 2\pi F_c$, F_c = carrier frequency, F_d = applied frequency deviation, $p = 2\pi F_m$, and F_m = modulation frequency. The noise spectrum may be expressed by its resultant,[5]

$$e_n = N \sin (\omega_n t + \phi(t)) \tag{4}$$

where N = instantaneous peak voltage of the noise (a function of time). $\phi(t)$ takes into account the fact that the noise resultant is phase modulated, as would be the case with the resultant of a spectrum of many noise voltages. $\omega_n = 2\pi F_n$, F_n = frequency of the noise resultant.

The signal voltage given by (3) and the noise voltage given by (4) may be combined by vector addition to give

$$e = \sqrt{C^2 + N^2 + 2CN \cos\left\{(\omega - \omega_n)t - \phi(t) + \frac{F_d}{F_m}\cos pt\right\}}$$

$$\sin\left[\omega t + \frac{F_d}{F_m}\cos pt + \tan^{-1}\frac{\sin\left\{(\omega-\omega_n)t - \phi(t) + \frac{F_d}{F_m}\cos pt\right\}}{\frac{C}{N} + \cos\left\{(\omega-\omega_n)t - \phi(t) + \frac{F_d}{F_m}\cos pt\right\}}\right]. \tag{5}$$

When the resultant wave given by (5) is passed through the limiter in the frequency modulation receiver, the amplitude modulation is removed. Hence the amplitude term is reduced to a constant and the only part of consequence is the phase angle of the wave. The rate of change of this phase angle, or its first derivative, is the instantaneous frequency of the wave. Taking the first derivative and dividing by 2π to change from radians per second to cycles per second gives

$$\frac{d\left[\omega t + \frac{F_d}{F_m}\cos pt + \tan^{-1}\frac{\sin\left\{\omega_{na}t - \phi(t) + \frac{F_d}{F_m}\cos pt\right\}}{\frac{C}{N} + \cos\left\{\omega_{na}t - \phi(t) + \frac{F_d}{F_m}\cos pt\right\}}\right]}{dt} \times \frac{1}{2\pi}$$

[5] John R. Carson, "The Reduction of Atmospheric Disturbances," *Proc. I.R.E.*, vol. 16, pp. 967-975; July, (1928).

$$= f = F_c - F_d \sin pt - \frac{\left(F_{na} - \frac{1}{2\pi} \frac{d\phi(t)}{dt} - F_d \sin pt \right)}{\frac{C}{N} + \cos \left\{ \omega_{na} t - \phi(t) + \frac{F_d}{F_m} \cos pt \right\}} + 1 \qquad (6)$$

$$\frac{N}{C} + \cos \left\{ \omega_{na} t - \phi(t) + \frac{F_d}{F_n} \cos pt \right\}$$

in which $\omega_{na} = (\omega - \omega_n) = 2\pi (F_c - F_n) = 2\pi F_{na}$.

Equation (6) gives the instantaneous frequency of the resultant wave consisting of the signal wave and the noise resultant voltage. From this equation the signal and noise frequency deviations may be obtained. In order to determine the over-all transmission with respect to the various components in the noise spectrum, a single component of noise, with constant amplitude and variable frequency, will be substituted for the resultant noise voltage given by (4). This makes N equal to n, which is not a function of time, and $\phi(t)$ equal to zero. Making these changes in (6) gives

$$f = F_c - F_d \sin pt - \frac{(F_{na} - F_d \sin pt)}{\frac{C}{n} + \cos \left\{ \omega_{na} t + \frac{F_d}{F_m} \cos pt \right\}} + 1 \qquad (7)$$

$$\frac{n}{C} + \cos \left\{ \omega_{na} t + \frac{F_d}{F_m} \cos pt \right\}$$

The equations for the instantaneous frequency, given by (6) and (7), show the manner in which the noise combines with the incoming carrier to produce a frequency modulation of the carrier. From these equations the frequency deviations of the noise may be determined, and from the frequency deviations the modulation factor at the output of the sloping filter may be found. Hence the over-all transmission may be obtained in terms of the modulation factor at the output of the sloping filter for a given carrier-noise ratio. When the carrier-noise ratio is high, (6) and (7) simplify so that calculations are fairly easy. When the carrier-noise ratio is low, the equations become involved to a degree which discourages quantitative calculations.

High Carrier-Noise Ratios

When C/n is large compared to unity, and the applied modulation on the frequency modulated wave is reduced to zero ($F_d = 0$), (7) reduces to

$$f = F_c - \frac{n^2}{C^2} F_{na} - \frac{n}{C} F_{na} \cos \omega_{na} t. \tag{8}$$

From (8) the effective peak frequency deviation of a single noise component of the spectrum is

$$F_{dn} = \frac{nF_{na}}{C}\left(\frac{n}{C} + 1\right). \tag{9}$$

But, since n/C is negligible compared to unity,

$$F_{dn} = \frac{nF_{na}}{C}. \tag{10}$$

When this value of frequency deviation is substituted in (1) to find the modulation factor[6] of the energy at the output of the sloping filter, the following results:

$$M_{fn} = \frac{n}{C} \times \frac{2F_{na}}{F_i}. \tag{11}$$

Equation (11) shows that the modulation factor of the noise is inversely proportional to the carrier-noise ratio and directly proportional to the ratio between the noise audio frequency and one half the intermediate-frequency channel width. When this equation is plotted with the noise audio frequency, F_{na}, as a variable and the modulation factor as the ordinate, the audio spectrum obtained for the detector output is like that of the triangular spectrum OBA in Figure 3. Such a spectrum would be produced by varying F_n through the range between the upper and lower cutoff frequencies of the intermediate-frequency channel. The noise amplitude would be greatest at a noise audio frequency equal to one half the intermediate-frequency channel width. At this noise audio frequency, the ratio $2F_{na}/F_i$ is equal to unity and the modulation factor becomes equal to n/C. If the detector output is passed through an audio system having a cutoff frequency F_a, the maximum frequency of the audio channel governs the maximum

[6] The ideal filter is used in this case since the use of a filter with a low conversion efficiency would merely require the addition of audio gain to put the frequency modulation receiver on an equivalent basis with the amplitude modulation receiver. The audio gain necessary would be equal to the reciprocal of the conversion efficiency of the sloping filter.

amplitude of the spectrum. The maximum amplitude of the detector output is therefore reduced by the ratio $F_i/2:F_a$. This can be seen by a comparison of the spectrum *OBA* for the detector output and the spectrum *ODH* for the audio channel output.

When the amplitude modulation reception process is analyzed with a carrier and noise spectrum present at the receiver input, the modulation factor of the energy fed to the detector is found to be equal to the reciprocal of the carrier-noise ratio for all of the noise frequencies in the spectrum. That is to say, the receiver transmission for amplitude modulation will be constant for all of the frequencies in the spectrum. Normally the upper cutoff frequency of the audio amplifier is equal to one half the intermediate-frequency channel width ($F_a = F_i/2$). Consequently the audio spectrum of the amplitude modulated noise fed to the detector will be the same as that at the receiver output and will be as portrayed by the rectangle *OCEH*.

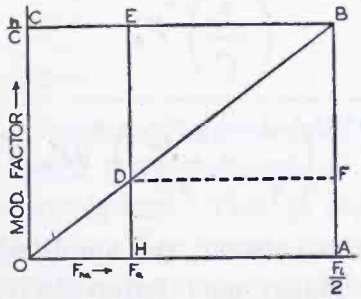


Fig. 3—Amplitude and frequency modulation receiver noise spectra. *OBA* = frequency modulation detector output. *ODH* = frequency modulation receiver output. *OCEH* = amplitude modulation receiver output.

The spectra of Figure 3 show the manner in which the frequency modulation receiver produces a greater signal-noise ratio than the amplitude modulation receiver. The noise at the output of the detector of the frequency modulation receiver consists of frequencies which extend out to an audio frequency equal to one half the intermediate-frequency channel width, and the amplitudes of these components are proportional to their audio frequency. Hence in passing through the audio channel the noise is reduced not only in range of frequencies, but also in amplitude. On the other hand, the components of the signal wave are properly disposed to produce detected signal frequencies which fit into the audio channel. In the case of the amplitude modulation receiver, the amplitude of the components at the output of the audio channel is the same as that at the output of the detector since the spectrum is rectangular. Thus the frequency modulation signal-noise ratio is greater than the amplitude modulation signal-noise ratio by a factor which depends upon the relative magnitudes of the spectra

OCEH and *ODH*. The magnitudes which are of concern are the root-mean-square and peak values of the voltage due to the spectra.

Root-Mean-Square Noise Considerations

The average noise power or average root-mean square voltage ratio between the rectangular amplitude modulation spectrum *OCEH* and the triangular frequency modulation spectrum *ODH*, of Figure 3, may be found by a comparison of the squared-ordinate areas of the two spectra. Thus,

$$\frac{W_a}{W_f} = \frac{\text{area } OCEH \text{ (ordinates)}^2}{\text{area } ODH \text{ (ordinates)}^2}$$

$$= \frac{\left(\frac{n}{C}\right)^2 F_a}{\int_0^{F_a} \left(\frac{n}{C} \times \frac{2F_{na}}{F_i}\right)^2 dF_{na}} = 3\left(\frac{F_i}{2F_a}\right)^2 \qquad (12)$$

where W_a/W_f is the ratio between the amplitude modulation average noise power and the frequency modulation average noise power at the receiver outputs. The root-mean-square noise voltage ratio will be

$$\frac{N_a}{N_f} \text{ (r-m-s fluctuation)} = \sqrt{\frac{W_a}{W_f}} = \sqrt{3}\,\frac{F_i}{2F_a}. \qquad (13)$$

Equation (13) gives the root-mean-square noise voltage ratio for equal carriers applied to the two receivers. The modulation factor of the frequency modulated signal due to the applied frequency deviation, F_d, is, from (3), equal to $2F_d/F_i$. The modulation factor of the amplitude modulated signal may be designated by M and has a maximum value of 1.0. Thus the ratio between the two signals will be given by

$$\frac{S_a}{S_f} \text{ (peak or r-m-s values)} = \frac{F_i M}{2F_d} = \frac{F_i}{2F_d} \text{ (for } M = 1.0). \qquad (14)$$

Dividing (13) by (14), to find the ratio between the signal-noise ratios at the outputs of the two receivers, gives

$$\frac{S_f/N_f}{S_a/N_a} \text{ (r-m-s values)} = \sqrt{3}\,\frac{F_d}{F_a}. \qquad (15)$$

It is apparent that the ratio between the frequency deviation and the audio channel, F_d/F_a, is an important factor in determining the

signal-noise ratio gain effected by frequency modulation. A convenient term for this ratio is the "deviation ratio" and it will be designated as such hereinafter.

Equation (15) gives the factor by which the amplitude modulation root-mean-square signal-noise ratio is multiplied in order to find the equivalent frequency modulation signal-noise ratio. Since this factor is used so frequently hereinafter, it will be designated by the word "improvement." The improvement given by (15) has been developed under the assumption of zero applied frequency deviation (no modulation) and a carrier which is strong compared to the noise. However, as will be shown later, as long as the carrier is strong compared to the noise, this equation also hold true for the case where modulation is present.

Peak Noise Considerations

In the ultimate application of signal-noise ratios, peak voltages are of prime importance since it is the peak of the noise voltage which seems to produce the annoyance. This is especially true in the case of impulse noise such as ignition where the crest factor of the noise may be very high. Thus the energy content of a short duration impulse might be very small in comparison with the energy content of the signal, but the peak voltage of the impulse might exceed the signal peak voltage and become very annoying. The degree of this annoyance would of course depend upon the type of service and will not be gone into here. In view of this importance of peak noise considerations, the final judgment in the comparison between the systems of frequency and amplitude modulation treated here will be based upon peak signal-noise ratios.

When the peak voltage or current ratio of the frequency and amplitude modulation spectra is to be determined, the characteristics of the different types of noise must be taken into consideration. There seem to be two general types of noise which require consideration. The first of these is fluctuation noise, such as thermal agitation and shot effect, which is characterized by a random relation between the various frequencies in the spectrum. The second is impulse noise, such as ignition or any other type of noise having a spectrum produced by a sudden rise of voltage, which is characterized by an orderly phase and amplitude relation between the individual frequencies in the spectrum.

Experimental data taken by the author have shown that the fluctuation noise crest factor is constant, independent of band width, when the carrier is strong compared to the noise. Thus the peak voltage of fluctuation noise varies with band width in the same manner that root-

mean-square voltage does, namely, as the square-root of the band width ratio. Consequently, for the strong-carrier condition, the peak voltage characteristics of fluctuation noise may be determined by applying the experimentally determined crest factor to the root-mean-square characteristics. Hence, in the case of fluctuation noise, (15) applies for peak noise improvement as well as for average root-mean-square noise improvement.

Impulse Noise Characteristics

A simple way of visualizing the manner in which impulse noise produces its peak radio-frequency voltage is to consider the case of a recurrent impulse. It is well known that a recurrent impulse, such as square-wave-form dots, may be expressed by a Fourier series which consists of a fundamental and an infinite array of harmonics. The amplitudes of these harmonics are inversely proportional to their frequencies. The components of the single impulse will be similar to those of the recurrent impulse since the single impulse may be considered as a recurrent impulse with a very low rate of recurrence. The part of this impulse spectrum that is received on a radio receiver is a small band of the very high order harmonics. Since the frequency difference between the highest and lowest frequencies of this band is small compared to the mid-frequency of the band, all of the frequencies received are of practically equal amplitude. These harmonics are so related to each other by virtue of their relation to a common fundamental that they are all in phase at the instant the impulse starts or stops. Hence, for the interval at the start or stop of the impulse, all of the voltages in the band add up arithmetically and the peak voltage of the combination is directly proportional to the number of individual voltages. Since the individual voltages of the spectrum are equally spaced throughout the band, the number of voltages included in a given band is proportional to the band width. Consequently, the peak voltage of the resultant of the components in the spectrum is directly proportional to the band width. Thus impulse noise varies, not as the square root of the band width, as fluctuation noise does, but directly as the band width.[7] Since the voltages in the spectrum add arithmetically, their peak amplitude is proportional to their average ordinate as well as proportional to the band width. This makes the peak voltage of impulse noise, not proportional to the square root of the ratio between

[7] The fact that the peak voltage of impulse noise varies directly with the band width was first pointed out to the author by V. D. Landon of the RCA Manufacturing Company. The results of his work were later presented by him as a paper entitled "A Study of Noise Characteristics," before the Eleventh Annual Convention, Cleveland, Ohio, May 13, 1936; published in the *Proc. I.R.E.*, vol. 24, pp. 1514-1521; November, (1936).

the *squared-ordinates areas*, as is the case with root-mean-square noise, but proportional to the ratio between the *areas* of the two spectra. Hence, (referring to Figure 3)

$$\frac{N_a}{N_f}\,(\text{peak values, impulse}) = \frac{\text{area } OCEH}{\text{area } ODH} = \frac{(n/C) \times F_a}{F_a \times \frac{1}{2} \times \frac{2F_a}{F_i} \times \frac{n}{C}} = \frac{F_i}{F_a}. \quad (16)$$

Dividing (16) by (14) to obtain the ratio between the frequency and amplitude modulation output signal-noise ratios gives

$$\frac{S_f/N_f}{S_a/N_a}\,(\text{peak values, impulse}) = 2\frac{F_d}{F_a}. \quad (17)$$

Equation (17) shows that the frequency modulation peak voltage improvement with respect to impulse noise is equal to twice the deviation ratio or about 1.16 times more improvement than is produced on fluctuation noise. The peak power gain would be equal to the square of the peak voltage gain or four times the square of the deviation ratio.

Low Carrier-Noise Ratios

When the expression for the instantaneous frequency of the wave modulated by the noise component and signal, given by (7), is resolved into its components by the use of the binomial theorem, the following is the result:

$$f = F_c - F_d \sin pt - \frac{(F_{na} - F_d \sin pt)}{Z}\left[\frac{n}{C} - \left(1 - \frac{2n}{ZC}\right)\left\{\frac{K_0}{Z} - K_1 \cos(\omega_{na}t - F_d \sin pt) + K_2 \cos 2(\omega_{na}t - F_d \sin pt) + K_3 \cos 3(\omega_{na}t - F_d \sin pt) \cdots\right\}\right] \quad (18)$$

in which $Z = \frac{C}{n} + \frac{n}{C}$ and

$$K_0 = K_1 = \left(1 + \frac{3}{Z^2} + \frac{10}{Z^4} + \frac{35}{Z^6} + \frac{126}{Z^8} + \frac{462}{Z^{10}} + \frac{1716}{Z^{12}} + \cdots\right) \quad (19)$$

$$K_2 = \left(\frac{1}{Z} - \frac{4}{Z^3} - \frac{15}{Z^5} - \frac{56}{Z^7} - \frac{210}{Z^9} - \frac{792}{Z^{11}} - \frac{3003}{Z^{13}} - \cdots \right) \tag{20}$$

$$K_3 = \left(\frac{1}{Z^2} + \frac{5}{Z^4} + \frac{21}{Z^6} + \frac{84}{Z^8} + \frac{330}{Z^{10}} + \frac{1287}{Z^{12}} + \frac{5005}{Z^{14}} + \cdots \right). \tag{21}$$

Additional terms of the series of (19), (20), and (21), as well as higher order series, may be found with the aid of a table of binomial coefficients.

Equation (18) shows that, as the carrier-noise ratio approaches unity, the effective signal-noise ratio at the receiver output is no

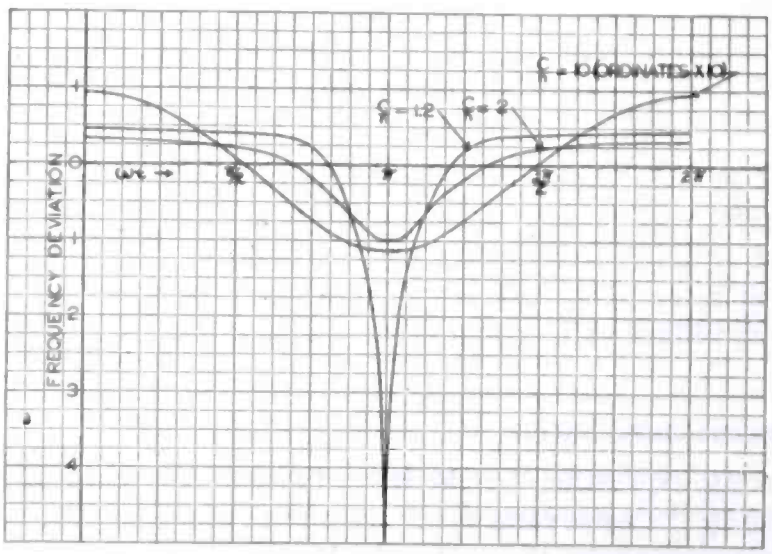


Fig. 4—Calculated wave forms showing the distortion produced on the instantaneous frequency deviation of the wave composed of the combination of the carrier and a single noise component. C/n = ratio between the peak voltage of the carrier and the peak voltage of the noise component.

longer directly proportional to the carrier-noise ratio. The effective frequency deviation produced by the noise has harmonics introduced and a constant frequency shift added. The effect of the harmonics and constant shift is to make the wave form of a single noise component very peaked and of the nature of an impulse. Because of the selectivity of the audio channel, none of the harmonics are present for the noise frequencies in the upper half of the audio spectrum. As the frequency of the noise voltage is lowered, more and more harmonics are passed by the audio channel and as a consequence, the peak frequency deviation due to the noise is increased. This can be more easily understood from the following calculation of the wave form produced by the instantaneous frequency deviation of the single noise component.

The curves of Figure 4 have been calculated from (7) and show how the instantaneous frequency deviation varies with time or the phase angle of the wave. A wave with the instantaneous frequency given by these curves would produce voltages in the output of the detector of the frequency modulation receiver which are proportional to the frequency deviations. It can be seen from these curves that, as the carrier-noise ratio approaches unity, the wave form becomes more and more peaked. The harmonics which enter in to make up this peaked wave form are given by (18) and are completely present for all noise frequencies only in the absence of audio selectivity.

In the presence of audio selectivity, the condition portrayed by (18) is approached as the audio frequency of the noise approaches zero. Thus the wave form of the noise is sinusoidal at a noise frequency high enough to have its harmonics eliminated by the audio selectivity, but becomes more peaked as the frequency is made lower so that more harmonics are included. This effect tends to increase the peak voltage of the low-frequency noise voltages which have a large number of harmonics present. Thus, as the carrier-noise ratio approaches unity, the triangular audio spectrum is distorted by an increase in the amplitude of the lower noise frequencies.

The above gives a qualitative and partially quantitative description of the noise spectrum which results at the lower carrier-noise ratios. Further development would undoubtedly make possible the exact calculation of the peak and root-mean-square signal-noise ratio at the receiver output when the carrier-noise ratio at the receiver input is close to unity, but, because of the laborious nature of the calculations involved in evaluating the terms of (18), and pressure of other work, the author is relying upon experimental determinations for these data.

Noise Crest Factor Characteristics

The crest factor characteristics of the noise can be studied to an approximate extent by a study of (6). This equation portrays the resultant peak frequency deviation of the wave at the output of the limiter. From it, the crest factor characteristics of the output of the detector may be determined since in the frequency modulation receiver frequency deviations are linearly converted into detector output voltages. However, the crest factor characteristics of the receiver output are different from those at the detector output due to the effect of the selectivity of the audio channel. This is especially true in the case of the frequency modulation receiver with a deviation ratio greater than unity, that is, where the audio channel is less than one half the intermediate-frequency channel. Consequently, in order to obtain the final

results, the effect of the application of the audio selectivity must be applied to the results determined from a study of (6).

From the curves of Figure 4, it can be seen that the peak frequency deviation of the wave given by (7) occurs at a phase angle equal to 180 degrees. From the similarity of (6) and (7), it can be seen that the peak frequency deviation of (6) would also occur at a phase angle of 180 degrees. At this phase angle the noise peak frequency deviation from (6) is

$$f_{dn}(\text{peak}) = \frac{\left(F_{na} - \frac{1}{2\pi}\frac{d\phi(t)}{dt} - F_d \sin pt\right)}{\frac{(C/N)-1}{(N/C)-1}+1}$$

$$= \frac{-\left(F_{na} - \frac{1}{2\pi}\frac{d\phi(t)}{dt} - F_d \sin pt\right)}{(C/N)-1}. \qquad (22)$$

Equation (22) shows that the peak frequency deviation of the noise, for any value of carrier-noise ratio, C/N, is proportional to the noise instantaneous audio frequency given by the quantity

$\left(F_{na} - \frac{1}{2\pi}\frac{d\phi(t)}{dt} - F_d \sin pt\right)$, and to the quantity $1/\{(C/N)-1)\}$.

C/N is the resultant instantaneous peak carrier-noise ratio which is present in the output of the frequency modulation intermediate-frequency channel. It is apparent that when this carrier-noise ratio is high, the peak frequency deviation of the noise is proportional to N/C. When the carrier-noise ratio is equal to unity, the peak frequency deviation becomes infinite and it is evident that the frequency modulation improvement, which is based on a high carrier-noise ratio, would be lost at this point. The term "improvement threshold" will be employed hereinafter to designate this point below which the frequency modulation improvement is lost and above which the improvement is realized. Theoretically this term would refer to the condition where the instantaneous peak voltage of the noise is equal to the peak voltage of the carrier. However, in the practical case, where only *maximum* peak values of the noise are measured, the improvement threshold will refer

to the condition of equality of the *maximum* instantaneous peak voltage of the noise and the peak voltage of the carrier.

As the experimental characteristics will show, this increase in peak frequency deviation of the noise is manifested in an increase in crest factor of the noise. The crest factor cannot rise to infinity, however, due to the limitations imposed by the upper and lower cutoff frequencies of the intermediate-frequency channel. This selectivity limits the peak frequency deviation of the resultant of the noise and applied modulation to a value not greater than one half the intermediate-frequency channel width. Hence, in the absence of applied frequency modulation, the peak voltage of the noise at the detector output may rise to a value equal to the peak voltage due to the applied frequency modulation with maximum frequency deviation. In the presence of the applied frequency modulation, the total peak frequency deviation is limited so that the noise peaks depress the signal, that is, they punch holes in the signal, but do not rise above it. Thus a phenomenon which might be termed "frequency limiting" takes place. This frequency limiting limits frequency deviations in the same manner that amplitude limiting limits amplitude deviations. The resulting effect is the same as though an amplitude limiter were placed at the detector output to limit the output so that the peak voltage of the noise or signal, or their resultant, cannot rise above a voltage corresponding to that produced by the signal alone at full modulation.

Since the frequency limiting limits the noise so that its maximum amplitude cannot rise above the maximum amplitude of the applied modulation, a noise suppression effect is present which is similar to that effected by the recent noise suppression circuits[8,9] used for reducing impulse noise which is stronger than the amplitude modulated carrier being received. The result of such limiting is a considerable reduction of the annoyance produced by an intermittent noise, such as ignition, where the duration of the impulses is short and the rate of recurrence is low. With such noise, the depression of the signal for the duration of the impulse reduces the presence of the signal for only a small percentage of the time; the resultant effect is a considerable improvement over the condition where the peaks of the noise are stronger than the signal. On the other hand, for steady noise such as fluctuation noise, as the carrier-noise ratio is made less than unity, the

---

[8] Leland E. Thompson, "A Detector Circuit for Reducing Noise Interference in C. W. Reception," *QST*, vol. 19, p. 38; April, (1935). A similar circuit for telephony reception is described by the same author in *QST*, vol. 20, pp. 44-45; February, (1936).

[9] James J. Lamb, "A Noise-Silencing I.-F. Circuit for Superhet Receivers," *QST*, vol. 20, pp. 11-14; February, (1936).

signal is depressed more and more of the time so that it is gradually smothered in the noise.

When the effect of the audio selectivity is considered in conjunction with the frequency limiting, it is found that the noise suppression effect is somewhat improved for the case of a deviation ratio greater than unity. The reason for this is as follows: The frequency limiting holds the peak voltage of the noise at the output of the detector so that it cannot rise above the maximum value of the signal. However, in passing through the audio channel, the noise is still further reduced by elimination of higher frequency components whereas the signal passes through without reduction. Consequently the over-all limiting effect is such that the noise is limited to a value which is *less* than the maximum value of the signal. The amount that it is less depends upon the difference between the noise spectra existing at the output of the detector and the output of the audio selectivity.

Experimental determinations, which will be shown later, point out that as unity carrier-noise ratio is approached, the frequency modulation audio noise spectrum changes from its triangular shape to a somewhat rectangular shape. Hence the noise spectrum at the output of the detector when frequency limiting is taking place would be approximately as given by *OCBA* of Figure 3. When the audio selectivity is applied, the spectrum would be reduced to *OCEH* and the band width of the noise would be reduced by a ratio equal to the deviation ratio. This would reduce the peak voltage of fluctuation noise by a ratio equal to the square root of the deviation ratio and that of impulse noise by a ratio equal to the deviation ratio. Thus, the resultant effect of the frequency limiting is that the fluctuation noise output is limited to a value equal to the maximum peak voltage of the signal divided by the square root of the deviation ratio. The corresponding value of impulse noise is limited to a value equal to the maximum peak voltage of the signal divided by the deviation ratio. Consequently, with fluctuation noise, when the noise and signal are measured in the absence of each other, the signal-noise ratio cannot go below a value equal to the square root of the deviation ratio; the corresponding signal-noise ratio impulse noise cannot go below a value equal to the deviation ratio. However, these minimum signal-noise ratios are only those which exist when the noise is measured in the absence of the applied frequency modulation. When the applied modulation and the noise are simultaneously present, the noise causes the signal to be depressed. When this depressed signal, with its depression caused by noise composed of a wide band of frequencies, is passed through the audio selectivity, the degree of depression is reduced. The amount of the reduction will be different from the two kinds of noise. The determination of the

actual magnitude of this reduction of the signal depression, as effected by the audio selectivity, will be left for experimental evaluation.

In comparing frequency modulation systems with different deviation ratios at the *low carrier-noise ratios,* the wider intermediate-frequency channel necessary for the high deviation ratio receiver gives that receiver a disadvantage with respect to the low deviation ratio receiver. Since this wider channel accepts more noise than the narrower intermediate-frequency channel of the low deviation ratio receiver, when equal carriers are fed to both such receivers equality of carrier and noise occurs at a higher carrier level in the high deviation ratio receiver. As a result, a higher carrier voltage is required to reach the improvement threshold in the case of the high deviation ratio system. Thus at certain low carrier levels, the carrier-noise ratio could be above the improvement threshold in the low deviation ratio system, but below in the high deviation ratio system; at this carrier level the low deviation ratio system would be capable of producing a better output signal-noise ratio than the high deviation ratio system.

The difference between the improvement thresholds of receivers with different deviation ratios may be investigated by a determination of the carrier-noise ratio which exists in the reference amplitude modulation receiver when the improvement threshold exists in the frequency modulation receiver. This carrier-noise ratio may be found by a consideration of the relative band widths of the intermediate-frequency channels of the receivers. Thus, when the deviation ratio is unity, and the intermediate-frequency channel of the frequency modulation receiver is of the same width as that of the amplitude modulation receiver,[10] the two receivers would have the same carrier-noise ratio in the intermediate-frequency channels. When the deviation ratio is greater than unity, and the intermediate-frequency channel of the frequency modulation receiver is broader than that of the amplitude modulation receiver, the carrier-noise ratio in the frequency modula-

[10] In order to assume that the frequency modulation receiver with a deviation ratio of unity has the same intermediate-frequency channel width as the corresponding amplitude modulation receiver, the assumption would also have to be made that the peak frequency deviation due to the applied frequency modulation is equal to one half the intermediate-frequency channel width. In the ideal receiver with a square-topped selectivity characteristic, this amount of frequency deviation would produce considerable out-of-channel interference and would introduce distortion in the form of a reduction of the amplitudes of the higher modulation of frequencies during the intervals of high peak frequency deviation. However, under actual conditions, where the corners of the selectivity characteristic are rounded, it has been found that the frequency deviation may be made practically equal to one half the normal selectivity used in amplitude modulation practice without serious distortion. Receivers with high deviation ratios are less susceptible to this distortion due to the natural distribution of the side bands for the high values of F_d/F_m which are encountered with such receivers.

tion receiver is less than that in the amplitude modulation receiver. For the case of fluctuation noise, where the peak values vary as the square root of the ratio between the two band widths concerned, the carrier-noise ratio in the frequency modulation intermediate-frequency channel would be less than that in the amplitude modulation intermediate-frequency channel by a ratio equal to the square root of the deviation ratio. Thus, when equal carrier voltage is fed to both receivers,

$$C_a/N_a = (C/N)\sqrt{F_d/F_a} \text{ (fluctuation noise, peak or r-m-s values)} \quad (23)$$

in which C_a/N_a = carrier-noise ratio in the amplitude modulation intermediate-frequency channel and C/N = corresponding ratio in the frequency modulation intermediate-frequency channel.

In the case of impulse noise, where the peak values of the noise vary directly with the band-width ratio, the carrier-noise ratios in the two receivers are related by

$$C_a/N_a = \frac{C}{N}\frac{F_d}{F_a} \text{ (impulse noise, peak values).} \quad (24)$$

From (23), it can be seen that, with fluctuation noise, a carrier-noise ratio equal to the square root of the deviation ratio would exist in the amplitude modulation intermediate-frequency channel when the carrier-noise ratio is at the improvement threshold ($C/N = 1$) in the frequency modulation intermediate-frequency channel. Likewise, from (24), with impulse noise, the frequency modulation improvement threshold occurs at a peak carrier-noise ratio in the amplitude modulation intermediate-frequency channel which is equal to the deviation ratio.

Effect of Application of the Modulation

For the condition of a carrier which is strong compared to the noise, the equation for the instantaneous frequency of the wave modulated by the noise and signal, given by (7), may be reduced to the following:

$$f = F_c - F_d \sin pt - \frac{n^2}{C^2}(F_{na} - F_d \sin pt)$$

$$- \frac{n}{C}(F_{na} - F_d \sin pt) \cos\left\{\omega_{na}t + \frac{F_d}{F_m}\cos pt\right\}. \quad (25)$$

By neglecting the inconsequential term proportional to n^2/C^2, applying the sine and cosine addition formulas, the Bessel function expansions,

and the Bessel function recurrence formulas, (25) may be resolved into

$$f = F_c - (n/C)\Bigg[J_0\left(\frac{F_d}{F_m}\right) F_{na}\cos\omega_{na}t$$

$$- J_1\left(\frac{F_d}{F_m}\right)\{(F_{na}+F_m)\sin(\omega_{na}t+pt)+(F_{na}-F_m)\sin(\omega_{na}t-pt)\}$$

$$- J_2\left(\frac{F_d}{F_m}\right)\{(F_{na}+2F_m)\cos(\omega_{na}t+2pt)+(F_{na}-2F_m)\cos(\omega_{na}t-2pt)\}$$

$$+ J_3\left(\frac{F_d}{F_m}\right)\{(F_{na}+3F_m)\sin(\omega_{na}t+3pt)+(F_{na}-3F_m)\sin(\omega_{na}t-3pt)\}$$

$$+ J_4\cdots\Bigg]. \qquad (26)$$

This resolution shows that the application of frequency modulation to the carrier divides the over-all transmission of the receiver into components due to the carrier and each side frequency. The amplitudes of these components are proportional to the frequency difference between the noise voltage and the side frequency producing the component. The frequency of the audio noise voltage in each one of these component spectra is equal to the difference between the side frequency and the noise radio frequency. Thus the application of the modulation changes the noise from a single triangular spectrum due to the carrier, into a summation of triangular spectra due to the carrier and side frequencies. In the absence of selectivity, the total root-mean-square noise would be unchanged by the application of the modulation since the root-mean-square summation of the frequency modulation carrier and side frequencies is constant; hence the root-mean-square summation of noise spectra whose amplitudes are proportional to the strength of the carrier and side frequencies would be constant. However, since selectivity is present, the noise is reduced somewhat. This can be seen by considering the noise spectrum associated with one of the higher side frequencies. The noise spectrum associated with this side frequency, which acts as a "carrier" for its noise spectrum, is curtailed at the high-frequency end by the upper cutoff of the intermediate-frequency channel. The region of noise below the side frequency is correspondingly increased in range, but yields high-frequency noise voltages which are eliminated by the audio-frequency selectivity. Consequently when modulation is applied, the noise is slightly reduced.

The amount of this reduction may be calculated by a root-mean-square summation of the individual noise spectra due to the carrier and side frequencies. For the case of a deviation ratio of unity, an actual summation of the various spectra for full applied modulation has shown the root-mean-square reduction to be between two and three decibels depending upon the audio frequency of the noise. The same sort of summations also shows that the reduction becomes less as the deviation ratio is increased.

The weak-carrier root-mean-square noise characteristics in the presence of applied frequency modulation do not lend themselves to such straightforward calculation as the corresponding strong-carrier characteristics and will not be gone into here. The same can be said for the peak-noise characteristics in the presence of applied frequency modulation.

Transmitter Frequency Modulation Power Gain

The above considerations, which are based upon the equivalent conditions of equal carrier amplitude at the input of the amplitude and frequency modulation receivers, do not take into account the power gain effected by the use of frequency modulation at the transmitter. Since the power in a frequency modulated wave is constant, the radio-frequency amplifier tubes in the transmitter may be operated in the class C condition instead of the class B condition as is required for a low level modulated amplitude modulation system. In changing from the class B to the class C condition, the output voltage of the amplifier may be doubled. Consequently a four-to-one power gain may be realized by the use of frequency modulation when the amplitude modulation transmitter uses low-level modulation. On the other hand, when the amplitude modulation transmitter uses high level modulation—that is, when the final amplifier stage is modulated, the power gain is not so great. However, for the purpose of showing the effect of a transmitter power gain, the amplitude modulation transmitter will be assumed to be modulated at low levels.

As this paper is in the final stages of preparation, systems of amplifying amplitude modulation have been announced wherein plate efficiencies of linear amplifiers have been increased practically to equal the class C efficiencies.[11,12] Since these systems are not in general use as yet, it will suffice to say that such improvements in amplitude modula-

[11] W. H. Doherty, "A New High Efficiency Power Amplifier for Modulated Waves," presented before Eleventh Annual Convention, Cleveland, Ohio, May 13, (1936); published in *Proc. I.R.E.*, vol. 24, pp. 1163-1182; September, (1936).

[12] J. N. A. Hawkins, "A New, High-Efficiency Linear Amplifier," *Radio*, no. 209, pp. 8-14, 74-76; May, (1936).

tion transmission will tend to remove the frequency modulation transmitter gain in accordance with these improvements. Hence the overall frequency modulation gain will more nearly approach that due to the receiver[13] alone.

With a four-to-one power gain at the transmitter, a frequency modulation system would deliver twice the carrier voltage to its receiver that an amplitude modulation system would with the same transmitter

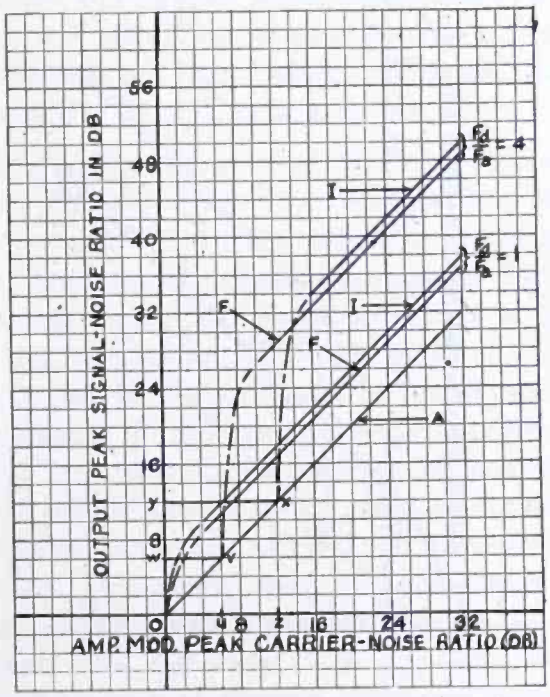


Fig. 5—Theoretical signal-noise ratio characteristics of frequency and amplitude modulation without the transmitter gain taken into account. Curve A = amplitude modulation receiver. The curves marked with I and F show the characteristics of the frequency modulation receivers for impulse and fluctuation noise, respectively. F_d/F_a = deviation ratio.

output stage. Hence (15) and (17), and (23) and (24) become, respectively,

$$\frac{S_f/N_f}{S_a/N_a} \text{ (peak values, fluctuation noise)} = 2\sqrt{3}F_d/F_a \tag{27}$$

$$\frac{S_f/N_f}{S_a/N_a} \text{ (peak values, impulse noise)} = 4F_d/F_a \tag{28}$$

$$C_a/N_a = (C/2N)\sqrt{F_d/F_a} \text{ (fluctuation noise, r-m-s or peak values)} \tag{29}$$

[13] The receiver and transmitter gain are mentioned rather loosely when they are separated in this way. However, it will be understood that the receiver gain could not be realized without providing a transmitter to match the requirements of the receiver.

$$C_a/N_a = (C/2N)\ F_d/F_a \text{ (impulse noise, peak values).} \qquad (30)$$

These equations show that this increase in carrier fed to the frequency modulation receiver not only increases the frequency modulation improvement, but also lowers the carrier-noise ratio received on the amplitude modulation receiver when the improvement threshold exists in the frequency modulation receiver.

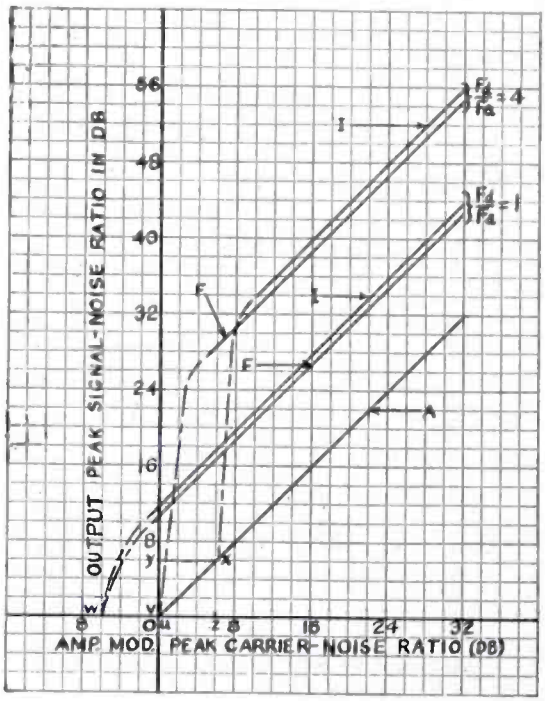


Fig. 6—Theoretical signal-noise ratio characteristics of frequency and amplitude modulation receivers with the transmitter gain taken into account.

Theoretical Conclusions

The curves of Figure 5 and 6 summarize the theoretical conclusions by means of an example in which receivers with deviation ratios of four and one are compared with each other and with an amplitude modulation receiver at various carrier-noise ratios. Figure 5 shows the receiver gain only, whereas Figure 6 takes into consideration a transmitter power gain of four to one. The curves are plotted with peak carrier-noise ratio in the amplitude modulation selectivity channel as a standard of comparison. Thus the curve for the amplitude modulation receiver is a straight line with a slope of forty-five degrees. The curves for the frequency modulation receivers show output signal-noise ratios which are greater or less than those obtained from the amplitude modulation receiver depending upon the carrier-noise ratio.

For Figure 5, (15) and (17) were used to obtain the strong-carrier frequency modulation improvement factors. Hence the frequency mod-

ulation output signal-noise ratios were obtained by multiplying the amplitude modulation signal-noise ratios by the frequency modulation improvement factors. The carrier-noise ratios which exist in the amplitude modulation receiver when the improvement threshold exists in the frequency modulation receiver were determined by substituting a value of unity carrier-noise ratio in (23) and (24). The improvement thresholds are designated in both Figures 5 and 6 by the points *u* and *z* for fluctuation and impulse noise, respectively. Since the theory does not permit actual calculation of signal-noise ratios in the region between high ratios and the improvement threshold, that part of the curves has been sketched in with a dashed line.

The part of the impulse-noise curve, for the deviation ratio of four represented by the line *x-y* shows the characteristic which would be obtained if the noise and signal were measured in the absence of each other. Because of frequency limiting, the noise it limited to equality with the signal at the output of the detector and is then reduced in peak voltage by the audio selectivity. The amount of this reduction for impulse noise would be a ratio equal to the deviation ratio or, in this case, twelve decibels. In the case of fluctuation noise, the reduction of the noise, which is present in the absence of modulation, would be equal to the square root of the deviation ratio, or six decibels, and the corresponding curve is shown by the line *v-w*. However, these lines do not portray the actual signal-noise ratio characteristics since the noise depresses the signal when the carrier-noise ratios go below the improvement threshold. In the case of fluctuation noise this signal depression causes the signal to become smothered in the noise as the carrier-noise ratio is lowered below the improvement threshold. On the other hand, with impulse noise such as ignition, where the pulses are short and relatively infrequent, carrier-noise ratios below the improvement threshold will present an output signal which is depressed by the noise impulses, but which is quite usable due to the small percentage of time that the impulse exists.

The curves of Figure 6, which take into account the frequency modulation transmitter gain, utilize (27) and (28) to obtain the frequency modulation improvements at the high carrier-noise ratios. These curves assume a carrier at the frequency modulation receiver inputs which is twice the strength of that present at the amplitude modulation receiver input. The frequency modulation improvements are therefore increased by six decibels and the improvement thresholds occur at signal-noise ratios in the amplitude receiver which are six decibels below the corresponding ratios for the case where the transmitter gain is not taken into account.

Further conclusions of the theory are as follows: For the high carrier-noise ratios, the application of modulation does not increase the root-mean-square value of the noise above its unmodulated value. Also, in the case of the low deviation ratio receivers, the root-mean-square value of the noise will be slightly reduced as the modulation is applied.

Experiment

In the experimental work it was desired to obtain a set of data from which curves could be plotted showing the frequency modulation characteristics in the same manner as the theoretical curves of Figure 5. To do this it was necessary to have an amplitude modulation reference system and frequency modulation receivers with deviation

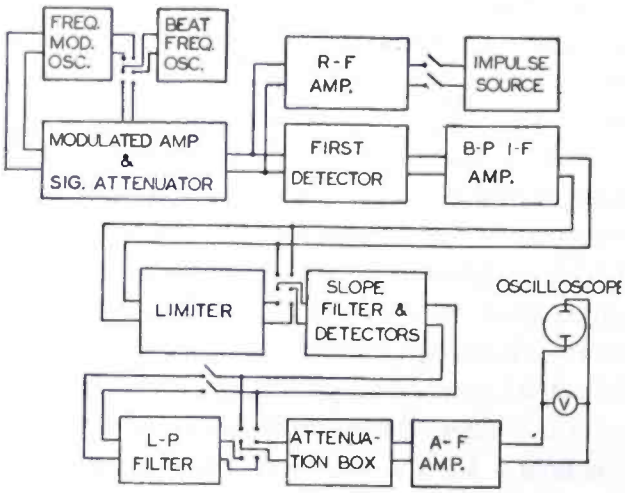


Fig. 7—Block diagram of experimental setup.

ratios of unity and greater than unity. Equal carrier voltages and noise spectra could then be fed to these receivers and the output signal-noise ratios measured while the carrier-noise ratio was varied. Since it was not convenient to measure the carrier-noise ratio at intermediate frequency, the output signal-noise ratios of the amplitude modulation receiver were measured instead and were plotted as abscissas in place of the carrier-noise ratios. This gives an abscissa scale which is practically the same as that which would be obtained by plotting carrier-noise ratios. The validity of this last statement was checked by measuring the linearity with which the output signal-noise ratio of the amplitude modulation receiver varied from high to low values as the carrier-noise ratio was varied by attenuating the carrier in known amounts in the presence of a constant noise. At the very low root-mean-square ratios the inclusion of the beats between the individual noise frequencies in the spectrum increases the apparent value of the

root-mean-square resultant of the noise voltages about two or three decibels. Thus, except for this small error at the low root-mean-square carrier-noise ratios, the amplitude modulation signal-noise ratio can be assumed equal to the carrier-noise ratio.

The block diagram of Figure 7 shows the arrangement of apparatus used in obtaining the experimental data. The frequency modulated oscillator employed a circuit which was similar to that used in the previously mentioned propagation tests.[1] The modulated amplifier consisted of a signal generator which was capable of being amplitude modulated, but whose master oscillator energy was supplied from the frequency modulated oscillator. Thus a signal generator was available

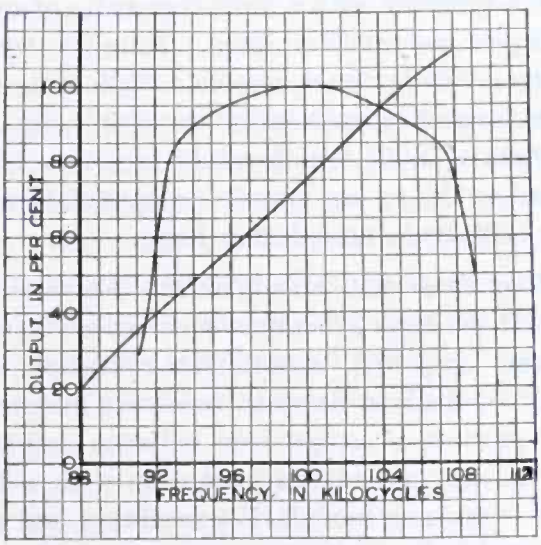


Fig 8—Band-pass characteristic of receiver intermediate-frequency amplifier, and characteristic sloping filter.

which was capable of being either frequency or amplitude modulated. A two-stage radio-frequency amplifier, tuned to the carrier frequency, but with no signal at its input, was used as the source of fluctuation noise. For the impulse noise measurements, the radio-frequency output of a square-wave multivibrator was fed to the input of this radio-frequency amplifier.

In order to make available frequency modulation receivers with different deviation ratios, a method was devised which made possible the use of a single intermediate-frequency channel and detection system for all receivers. The method consisted in the insertion of a low-pass filter in the audio output of the receiver so as to reduce the width of the audio channel and thereby increase the deviation ratio of the receiver. This procedure is not that which might be normally followed since to increase the deviation ratio, the audio channel would normally be left constant and the intermediate-frequency channel increased.

However, since it is only the *ratio* between the intermediate- and audio-frequency channels which governs the frequency modulation improvement, such an expedient is permissible for the purpose of the experiments.

The band-pass filter of the receiver intermediate-frequency amplifier was adapted from broadcast components and gave an output which was about one decibel down at 6500 cycles off from mid-band frequency. (See Figure 8.) Hence maximum frequency deviation was limited to 6500 cycles. The audio channel of the receiver cut off at 6500 cycles and the low-pass filter cut off at 1600 cycles. Thus the following four different types of receivers were available: Number one, a frequency modulation receiver with a deviation ratio of unity which would receive a 6500-cycle modulation band. Number two, an amplitude modulation receiver which would receive a 6500-cycle modulation band. Number three, a frequency modulation receiver with a deviation ratio of about four (6500 ÷ 1600) which would receive a 1600-cycle modulation band. Number four, an amplitude modulation receiver which would receive a 1600-cycle modulation band.

With these four receivers, a comparison between number two and number one would produce a comparison between amplitude modulation reception and frequency modulation reception with a deviation ratio of unity. A comparison between receivers number four and number three would produce a comparison between amplitude modulation reception and frequency modulation reception with a deviation ratio of four. Thus both frequency modulation receivers had as a standard of comparison an amplitude modulation receiver with an audio channel equal to that of the frequency modulation receiver.

The limiter of the frequency modulation receiver consisted of four stages of intermediate-frequency amplification arranged alternately to amplify and limit. The sloping filter detectors utilized the same circuit as used in the propagation tests[1] except that only one sloping filter was used in conjunction with a flat-top circuit as described in the theoretical section of this paper. Thus a balanced detector type of receiver was available which would also receive amplitude modulation by switching off the frequency modulation detector and receiving the detected output of the flat-top circuit. The characteristic of the sloping filter is shown in Figure 8.

The output of the detectors was fed to a switching system which connected either to a low-pass filter and attenuator or directly to the attenuator. The output of the attenuator passed to an audio-frequency amplifier having an upper cut-off frequency of 6500 cycles. The indicating instruments were connected to the amplifier output terminals.

For the root-mean-square fluctuation noise measurements, a copper-oxide-rectifier type meter was used.[14] A cathode-ray oscilloscope was used for all peak voltage measurements.

In the procedure used to obtain the data, the carrier-noise ratio was varied over a wide range of values and the receiver output signal-noise ratios were measured at each value of carrier-noise ratio. To do this, the output of the noise source was held constant while the carrier was varied by means of the signal generator attenuator. The output peak signal-noise ratios were obtained by first measuring the peak voltage of the tone output with the noise source shut off and then measuring the peak voltage of the noise with the tone shut off. The *maximum* peak voltage of the noise was read for its peak voltage. The root-mean-square signal-noise ratios were measured by reading the root-mean-square voltage of the tone in the presence of the noise and then reading the voltage of the noise alone. The signal was then separated from the noise by equating the measured signal-plus-noise voltage to $\sqrt{S^2 + N^2}$, substituting the measured noise voltage for N, and solving for the signal, S. In these measurements, a 1000-cycle tone was used to modulate at fifty per cent the amplitude modulator or to produce one-half frequency deviation (3250 cycles) on the frequency modulator. The output signal-noise ratios were corrected to a 100 per cent, or full modulation, basis by multiplying them by two. The radio frequency used was ten megacycles.

Fluctuation Noise Characteristics

The curves of Figure 9 show the fluctuation noise characteristics, in which peak signal-noise ratios were measured. These curves check the theoretical curves of Figure 5 as nearly as such measurements can be expected to check. With the deviation ratio of four (low-pass filter in), the theoretical strong-carrier improvement should be $4 \times 1.73 = 6.9$ or 16.8 decibels; the measured improvement from Figure 9 is about 14 decibels. With the deviation ratio of unity (low-pass filter out), the measured improvement was about 3.5 decibels as compared to the 4.76-decibel theoretical figure. The full frequency modulation improvement is seen to be obtained down to carrier-noise ratios about two or three decibels above the improvement threshold (equality of peak carrier and

[14] In the preliminary measurements, a thermocouple meter was connected in parallel with the copper-oxide-rectifier meter in order to be sure that no particular condition of the fluctuation noise wave form would cause the rectifier meter to deviate from its property of reading root-mean-square values on this type of noise. It was found that the rectifier type of instrument could be relied upon to indicate correctly so that the remainder of the measurements of root-mean-square fluctuation noise were made using the more convenient rectifier type of instrument.

noise). The fact that the frequency modulation improvement threshold occurs at a higher carrier-noise ratio in the case of the receiver with a deviation ratio of four than in the case of the receiver with a deviation ratio of unity, also checks the theoretical predictions. In this case of fluctuation noise, the improvement threshold for the receiver with the deviation ratio of four should occur at a carrier-noise ratio in the amplitude modulation intermediate-frequency channel which was twice the corresponding ratio for the receiver with a deviation ratio of unity. The curves show these two points to be about seven decibels apart or within one decibel of the theoretical figure of six decibels.

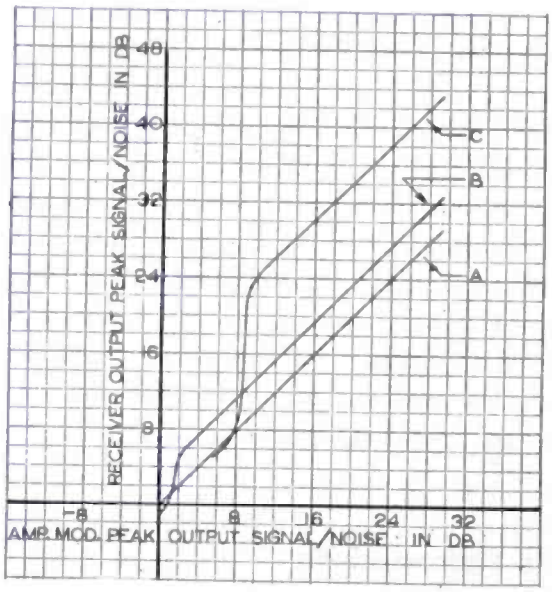


Fig. 9—Measured peak signal-noise ratio characteristics for fluctuation noise. Curve A = amplitude modulation receiver. Curve B = frequency modulation receiver with deviation ratio equal to unity. Curve C = frequency modulation receiver with deviation ratio equal to four.

The data for the curves of Figure 9 were obtained by measuring the peak value of the noise alone and signal alone and taking the ratio of these two values as the signal-noise ratio. Hence the signal depressing effect, occurring for carrier-noise ratios below the improvement threshold, does not show up on the curves. In order to obtain an approximate idea as to the order of magnitude of this effect, observations were made in which the carrier-noise ratio was lowered below the improvement threshold while the tone modulation output (100 per cent modulation in the case of the amplitude modulation observation and full frequency deviation in the case of the frequency modulation observation) was being monitored by ear and oscilloscope observation. It was found that the fluctuating nature of the instantaneous peak voltage of the fluctuation noise had considerable bearing upon the effects

observed. Due to the fact that the instantaneous value of the peak voltage is sometimes far below the maximum instantaneous value, frequency modulation improvement is obtained to reduce still further the peak voltage of these intervals of noise having instantaneous peak voltages lower than the maximum value. This effect seems to produce a signal at the output of the frequency modulation receiver which sounds "cleaner," but which has the same maximum peak voltage characteristics as the corresponding amplitude modulation receiver. Thus, as far as maximum peak voltage of the noise is concerned, the frequency

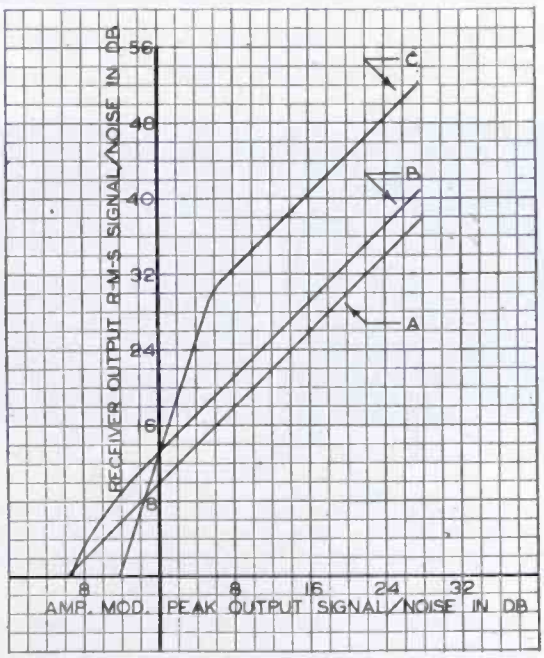


Fig. 10—Measured root-mean-square signal-noise ratio characteristics for fluctuation noise. Curve A = amplitude modulation receiver. Curve B = frequency modulation receiver with deviation ratio equal to unity. Curve C = frequency modulation receiver with deviation ratio equal to four.

modulation receiver produces about the same output as the amplitude modulation receiver for carrier-noise ratios below the improvement threshold. The reduction of the peak voltage of the noise during the intervals of lower instantaneous peak value reduces the energy content of the noise in the output; hence some idea of the magnitude of this effect can be obtained from the root-mean-square characteristics of the noise.

The curves of Figure 10 are similar to those of Figure 9 except that the root-mean-square signal-noise ratios are plotted as ordinates. Since the crest factor of the signal is three decibels and that of fluctuation noise is about thirteen decibels (as later curves will show), the root-mean-square signal-noise ratios are ten decibels higher than the cor-

responding peak ratios. It can be seen that the root-mean-square characteristics differ from the peak characteristics in the range of carrier-noise ratios below the improvement threshold; above the improvement threshold, the characteristics are similar.

Since the root-mean-square and peak signal-noise ratios display different characteristics below the improvement threshold, it is quite evident that the crest factor of the noise changes as the carrier-noise ratio is lowered below this point. The crest factor can be obtained from the curves of Figure 9 and 10 as follows: By adding three decibels to the ordinates of Figure 10 they will be converted to peak signal to

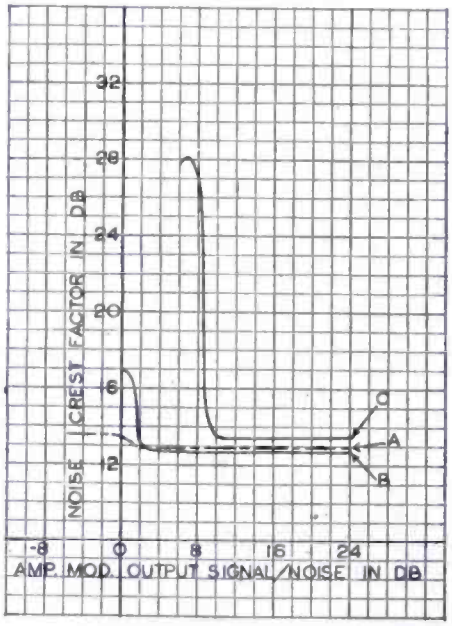


Fig. 11—Crest factor characteristics of frequency and amplitude modulation receivers. Curve A = amplitude modulation receiver with 6500-cycle audio channel. Curve B = frequency modulation receiver with deviation ratio equal to unity. Curve C = frequency modulation receiver with deviation ratio equal to four.

root-mean-square noise ratios. Hence by subtracting from these ratios the corresponding ordinates of Figure 9, the crest factor of the noise is obtained. The results of such a procedure are shown in Figure 11.

In the case of the frequency modulation receiver with a deviation ratio of four, Figure 11 shows that the crest factor increases by about 14.5 decibels at the improvement threshold. Hence the frequency modulation improvement, which is about fourteen decibels by measurement and sixteen by calculation, is counteracted by an increase in crest factor. This same situation exists in the case of the receiver with a deviation ratio of unity. Here the increase in crest factor is about four decibels; the measured frequency modulation improvement is about 3.5 decibels and the calculated value 4.76 decibels.

Curve *A* of Figure 11 shows the crest factor characteristics of the amplitude modulation receiver. It is seen that this crest factor is about equal to that for frequency modulation above the improvement threshold. The average value of the crest factor for both amplitude and frequency modulation in this region is thirteen decibels or about 4.5 to one. This value checks previous measurements of crest factor where a slide-back vacuum tube voltmeter was used in place of an oscilloscope to measure the peak voltage and a thermocouple was used to measure the root-mean-square values.

The point where the crest factor of the noise increases, which occurs at the frequency modulation improvement threshold, has a rather distinctive sound to the ear. When fluctuation noise is being observed,

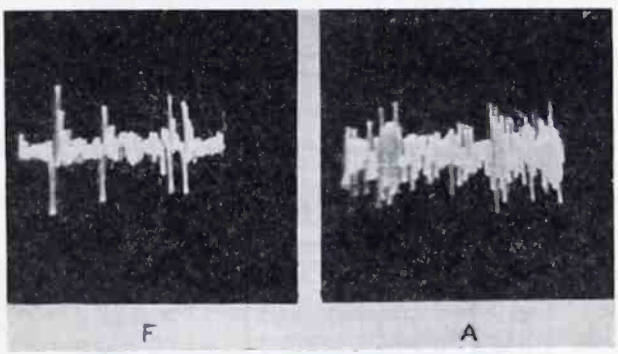


Fig. 12—Wave form of fluctuation noise output at unity carrier-noise ratio in the frequency modulation receiver. *F* = frequency modulation receiver. *A* = amplitude modulation receiver.

as this point is approached the quality of the hiss takes on a more intermittent character, somewhat like that of ignition. This point has been termed by the author the "sputter point," and since it coincides with the improvement threshold it is a good indicator for locating the improvement threshold. It is caused by the fact that the fluctuation noise voltage has a highly variable instantaneous peak voltage so that there are certain intervals during which the instantaneous peak voltage of the noise is higher than it is during other intervals. Consequently, as the maximum peak value of the noise approaches the peak value of the signal, the higher instantaneous peaks will have their crest factor increased to a greater degree than the lower instantaneous peaks. Figure 12 shows oscillograms taken on the fluctuation noise output of the frequency and amplitude modulation receivers with the 1600-cycle low-pass filter in the audio circuit and with the signal-noise ratio adjusted to the sputter point. These oscillograms also tend to show how the frequency modulation signal would sound "cleaner" than the amplitude modulation signal when the carrier-noise ratio is below the improvement threshold.

Data were also taken to show the fluctuation noise characteristics as frequency modulation is applied. These data were taken by inserting low-pass or high-pass filters in the audio system and then applying a modulation frequency to the frequency modulated oscillator which would fall outside the pass band of the filters. The low-pass filter cut off at 1600 cycles so that modulating frequencies higher than 1600 cycles were applied. The output of the filter contained only noise in the range from zero to 1600 cycles and the change of noise versus frequency deviation of the applied modulation could be measured. The high-pass filter also cut off at 1600 cycles so that measurements of the

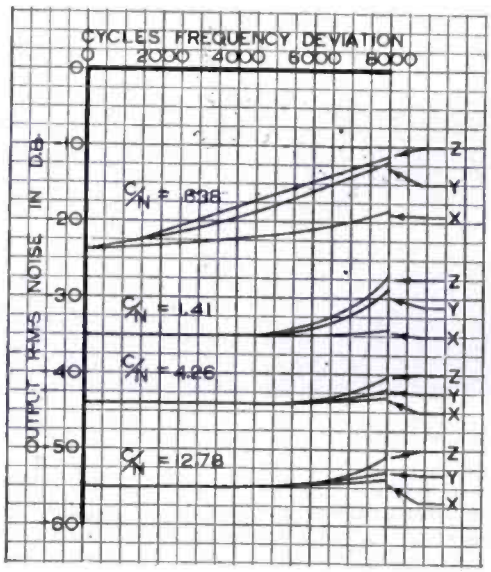


Fig. 13—Variation of frequency modulation receiver output noise as frequency modulation is applied. 1600-cycle low-pass filter in audio output. Modulation frequency: for curve $X = 6000$ cycles, $Y = 3000$ cycles, and $Z = 2000$ cycles. C/N = peak carrier-noise ratio in the output of intermediate-frequency channel.

noise in the range from 1600 to 6500 cycles were made while applying modulation frequencies below 1600 cycles. In the case of the high-pass filter, the harmonics of the modulating frequencies appeared at the filter output in addition to the noise. Consequently, a separate measurement of the harmonics in the absence of the noise was made so that the noise could be separated from the harmonics by the quadrature relations. The results with the low-pass filter are shown in Figure 13. The results with the high-pass filter are shown in Figure 14.

The curves of Figure 13 are representative of a system with a deviation ratio of four. They point out the fact that when the peak carrier-noise ratio in the frequency modulation intermediate-frequency channel is greater than unity, the root-mean-square noise is substantially unchanged due to the application of modulation. The one curve

for a carrier-noise ratio less than unity shows a gradual increase of the noise, which would effect a decrease of the signal-noise ratio as the modulation is applied; this increase in the noise is displayed to a greater extent on the lower modulation frequency of 2000 cycles than on the higher modulation frequencies of 3000 and 6000 cycles.

Figure 14 is approximately representative of a receiver with a deviation ratio of unity. This is because the range of noise frequencies from zero to 1600 cycles, which were eliminated by the high-pass filter, were a small part of the total range extending out to 6500 cycles. At the highest carrier-noise ratio, the noise is decreased as the

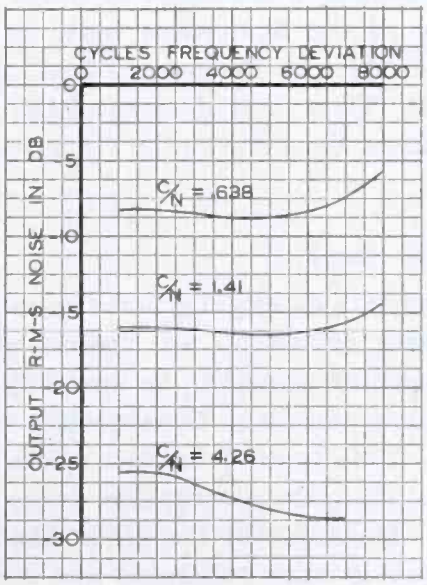


Fig. 14—Variation of frequency modulation receiver output noise as frequency modulation is applied. 1600-cycle high-pass filter in audio output. Modulation frequency = 1000 cycles.

modulation is applied. This is in accordance with the deductions of the theory in the section *Effect of Application of Modulation.* As the carrier-noise ratio is lowered this tendency is eliminated.

Data similar to that for Figure 13, with the low-pass filter in the audio circuit, were taken measuring the output *peak* voltage of the noise. The characteristics obtained were identical to those obtained with root-mean-square measurements.

Since the harmonics of the tone present in the output of the high-pass filter could not readily be separated from the noise for the peak voltage measurements, the high-pass filter data were taken by root-mean-square measurements only.

Measurements were also made to determine how much the audio selectivity reduced the degree of signal depression present at the output of the detector of the frequency modulation receiver. The carrier-

noise ratio was set so that the maximum peak voltage of the fluctuation noise was equal to the peak voltage of the carrier. At this carrier-noise ratio the maximum noise peaks depressed the signal down to zero at the output of the detector. At the output of the 1600-cycle low-pass filter, the maximum noise peaks depressed the signal five decibels. Thus, without the audio selectivity, the signal was depressed by an amount equal to its total amplitude; with the audio selectivity, the signal was depressed to five decibels below full amplitude or down to an amplitude of 56 per cent. Hence the reduction of the depth of the signal depression was from a 100 per cent depression to a depression of $(100 - 56) = 44$ per cent or a reduction of about seven decibels. The theoretical reduction of the fluctuation noise in the absence of the modulation would be equal to the square root of the ratio of band widths or six decibels. Thus the reduction of the signal depression is, for all practical purposes, the same as the reduction in the peak voltage of the noise alone.

Impulse Noise Measurements

The first measurements on impulse noise were made using an automobile ignition system driven by an electric motor. However the output from this generator proved to be unsteady and did not allow a reasonable measurement accuracy. Consequently a square-wave multivibrator was set up. This type of impulse noise generator proved to be even more stable than the fluctuation noise source and allowed accurate data to be obtained. On the other hand, the output of the receiver being fed by this noise generator was not as steady as would be expected. In the absence of the carrier the output was steady, but as the carrier was introduced the output peak voltage started to fluctuate. Apparently the phase relation between the components of the noise spectrum and the carrier varies in such a manner as to form a resultant wave which varies between amplitude modulation and phase or frequency modulation. Hence the output of a receiver which is adjusted to receive either type of modulation alone will fluctuate depending upon the probability considerations of the phase of combination of the carrier and noise voltages.

The preliminary impulse noise measurements were made on an amplitude modulation receiver by comparing the peak voltage ratio between the two available band widths of 6500 and 1600 cycles. The 6500-cycle channel was fed to one set of oscilloscope plates and the 1600-cycle channel to the other. Thus, when the peak voltages at the outputs of the two channels were equal the oscilloscope diagram took a symmetrical shape somewhat like a plus sign. The two channel levels were equalized by means of a tone. Hence, when the noise voltage was

substituted for the tone, the amount of attenuation that had to be inserted in the wider band to produce a symmetrical diagram on the oscilloscope was taken as the ratio of the peak voltages of the two band widths. In this manner a series of readings was taken which definitely proved that the peak voltage ratio of the two band widths was proportional to the band width ratio. These readings were taken on both the ignition system noise generator and the multivibrator generator. As a check, readings on fluctuation noise were also taken which showed that the peak voltage of fluctuation noise varies as the square root of the band width.

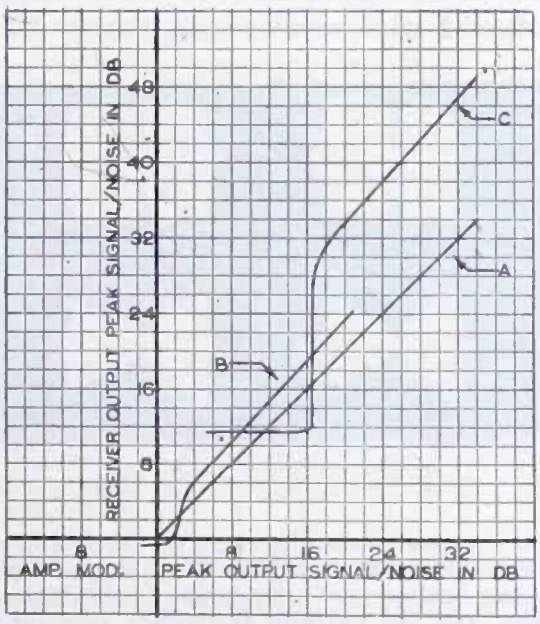


Fig. 15—Measured peak signal-noise ratio characteristics of impulse noise. Curve A = amplitude modulation receiver. Curve B = frequency modulation receiver with deviation ratio of unity. Curve C = frequency modulation receiver with deviation ratio of four.

The final measurements on impulse noise were made using the same procedure followed for the fluctuation noise measurements of Figure 9. Only peak voltage measurements were made on this type of noise. The curves are shown in Figure 15. It can be seen that the peak voltage characteristics of impulse noise are similar to those of fluctuation noise except for the location of the improvement threshold. For the receiver with a deviation ratio of four, the improvement threshold occurs at a carrier-noise ratio slightly above sixteen decibels as compared with slightly above eight decibels for fluctuation noise. The difference between the improvement thresholds for the two frequency modulation receivers is about fourteen decibels; the corresponding theoretical figure, which is equal to the ratio of the two deviation ratios, is twelve decibels. The theoretical difference between the strong-carrier fre-

quency modulation improvements for impulse and fluctuation noises, as indicated by the difference between the factors two and the square root of three respectively, is too small to be measurable with such variable quantities as these noise voltages.

Since the signal-noise ratios for the curves of Figure 15 were obtained by measuring the noise and signal in the absence of each other, the signal-depressing effect of the noise does not show up. However, in the case of impulse noise, these curves are more representative of the actual situation existing, because the noise depresses the signal for only a small percentage of the time. In the listening and oscilloscope observations conducted with carrier-noise ratios below the improve-

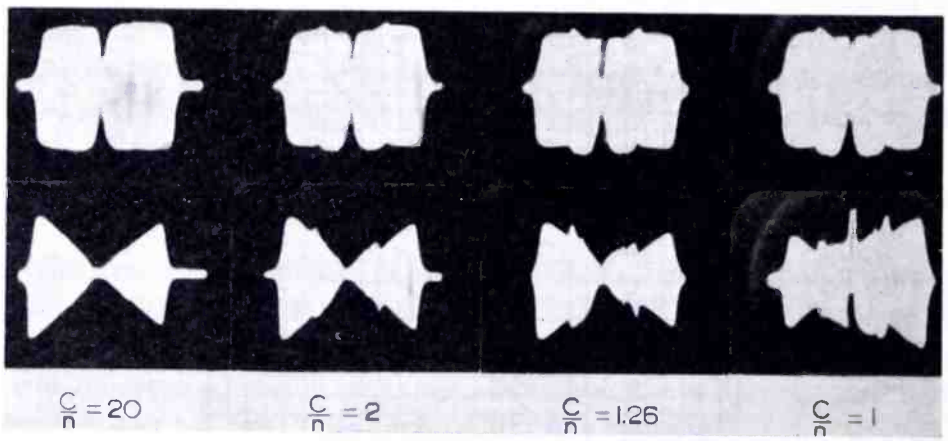


Fig. 16—Over-all transmission oscillograms of the frequency and amplitude modulation receivers. 1600-cycle low-pass filter in audio output. Top row = amplitude modulation, bottom row = frequency modulation. C/n = ratio between the carrier and the variable-frequency heterodyning voltages.

ment threshold, it was observed that at unity carrier-noise ratio the noise peaks depressed the amplitude of the signal to zero at the output of the detector. When the low-pass filter was inserted in the audio circuit, the impulse noise peaks depressed the signal about 2.5 decibels or reduced the amplitude from 100 per cent to 75 per cent. The effective signal-noise ratio is then increased from unity to $100/(100-75)=4$ or 12 decibels. This is equal to the theoretical reduction in peak voltage of impulse noise which would be effected by this four-to-one band width ratio. It is then evident that the reduction of the depth of the signal depression caused by the impulse noise is of the same magnitude as the reduction of the peak voltage of the noise alone.

Over-all Transmissions

The oscillograms of Figure 16 show the over-all transmissions of the amplitude and frequency modulation receivers at various carrier-noise ratios. These oscillograms were taken by tuning the receiver to a

carrier, and then, to simulate the noise, manually tuning a heterodyning signal across the intermediate-frequency channel. The audio beat output of the receiver was applied through the low-pass filter to the vertical plates of the oscilloscope. A bias proportional to the frequency change of the heterodyning voltage was applied to the other set of oscilloscope plates. Consequently the spectra obtained are those which would be produced by the combination of a single noise component of variable frequency and the carrier. At the higher carrier-noise ratios, the spectrum is rectangular for amplitude modulation and triangular for frequency modulation. The dip in the middle of the amplitude modulation spectrum is where the audio output is near zero beat. As the carrier-noise ratio is decreased, the frequency modulation spectrum deviates from its triangular shape and the wave form of the receiver output has increased harmonic content at the lower audio frequencies where the audio selectivity does not eliminate the harmonics.

The amplitude modulation spectra of Figure 16 also show the presence of added harmonic distortion on the lower modulation frequencies and lower carrier-noise ratios. However, the effect is so small that it is of little consequence.

The spectra of Figure 16 allow a better understanding of the situation which is theoretically portrayed by (7) of the theory.

Experimental Conclusions

It can be concluded that the experimental data in general confirm the theory and point out the following additional information:

The improvement threshold starts at a carrier-noise ratio about three or four decibels above equality of peak carrier and noise in the frequency modulation intermediate-frequency channel. Hence the full frequency modulation improvement may be obtained down to a peak carrier-noise ratio in the frequency modulation receiver of three or four decibels.

The root-mean-square fluctuation noise characteristics differ from the peak fluctuation noise characteristics for carrier-noise ratios below the improvement threshold. The improvement threshold starts at about the same peak carrier-noise ratio, but the improvement does not fall off as sharply as it does for peak signal-noise ratios. Thus, for carrier-noise ratios below the improvement threshold the energy content of the frequency modulation noise is reduced, but the peak characteristics are approximately the same as those of the amplitude modulation receiver. The characteristics are not exactly the same due to the frequency limiting which allows the noise peaks to depress the signal, but does not allow them to rise above the signal.

The crest factor of the fluctuation noise at the outputs of the frequency and amplitude modulation receivers is about thirteen decibels or 4.5 to one for the strong-carrier condition. The crest factor of amplitude modulation fluctuation noise remains fairly constant regardless of the carrier-noise ratio. At equality of peak carrier and peak noise in the frequency modulation intermediate-frequency channel, the crest factor of the noise in the output of the frequency modulation receiver rises to a value which counteracts the peak signal-noise ratio improvement over amplitude modulation; the improvement threshold manifests itself in this manner.

At the improvement threshold, the application of the audio selectivity reduces the signal depression due to a noise peak by the same ratio that it reduces the noise in the absence of the signal. Thus the depth of a noise depression in the signal is reduced by a ratio equal to the square root of the deviation ratio in the case of fluctuation noise, and equal to the deviation ratio in the case of impulse noise.

General Conclusions

The theory and experimental data point out the following conclusions:

A frequency modulation system offers a signal-noise ratio improvement over an equivalent amplitude modulation system when the carrier-noise ratio is high enough. For fluctuation noise this improvement is equal to the square root of three times the deviation ratio for both peak and root-mean-square values. For impulse noise the corresponding peak signal-noise ratio improvement is equal to twice the deviation ratio. When the carrier-noise ratio is about three or four decibels above equality of peak carrier and peak noise in the frequency modulation intermediate-frequency channel, the peak improvement for either type of noise starts to decrease and becomes zero at a carrier-noise ratio about equal to unity. Below this "improvement threshold," the peak characteristics of the frequency modulation receiver are approximately the same as those of the equivalent amplitude modulation receiver. The root-mean-square characteristics of the frequency modulation noise show a reduction of the energy content of the noise for carrier-noise ratios below the improvement threshold; this is evidenced by the fact that the improvement threshold is not as sharp for root-mean-square values as for peak values.

At the lower carrier-noise ratios, frequency modulation systems with lower deviation ratios have an advantage over systems with higher deviation ratios. Since the high deviation ratio system has a wider intermediate-frequency channel, more noise is accepted by that

channel so that the improvement threshold occurs at a higher carrier level in the high deviation ratio system than in the low. Hence the low deviation ratio systems retain their frequency modulation improvement down to lower carrier levels.

The peak voltage of fluctuation noise varies with band width in the same manner as the root-mean-square voltage, namely, as the square root of the band width. The peak voltage of impulse noise varies directly as the band width. In frequency modulation systems with a deviation ratio greater than unity, this difference in the variation with band width makes the improvement threshold occur at a higher carrier level with impulse noise than with fluctuation noise. Hence frequency modulation systems with higher deviation ratios are more susceptible to impulse noise interference.

Because of a phenomenon called "frequency limiting" the peak frequency deviations of the noise or the noise-plus-signal are limited so that the peak value cannot rise above the maximum peak value of the signal at the output of the detector. The application of audio selectivity reduces this maximum value of the noise so that fluctuation noise cannot rise to a value higher than the maximum value of the signal divided by the square root of the deviation ratio; the corresponding value of impulse noise cannot rise to a value higher than the maximum peak voltage of the signal divided by the deviation ratio. Inherent with this limiting effect is a signal-depressing effect which causes the fluctuation noise gradually to smother the signal as the carrier-noise ratio is lowered below the improvement threshold. However in the case of impulse noise, the signal depression is not as troublesome, and a noise-suppression effect is created which is similar to that effected in the recent circuits for suppressing impulse noise which is stronger than the carrier in an amplitude modulation system. When the deviation ratio is greater than unity, this frequency limiting is more effective than the corresponding amplitude modulation noise-suppression circuits; this is caused by the audio selectivity reducing the maximum peak value of the noise so that it is less than the peak value of the signal.

For carrier-noise ratios greater than unity, the application of frequency modulation to the carrier does not increase the noise above its value in the absence of applied frequency modulation.

At the transmitter, a four-to-one power gain is obtained by the use of class C radio-frequency amplification for frequency modulation instead of the customary class B amplification as is used for low level amplitude modulation. Therefore, for the same transmitter power input, a frequency modulation system will produce at its receiver a

carrier which is twice as strong as that produced at the receiver of an amplitude modulation system. This results in two effects: first, the frequency modulation improvement is doubled for carrier-noise ratios above the improvement threshold; second, when the improvement threshold occurs in the frequency modulation receiver, the carrier-noise ratio existing in the amplitude modulation receiver is one half of what it would have been without the transmitter power gain.

ACKNOWLEDGMENT

The author thanks Mr. H. H. Beverage and Mr. H. O. Peterson for the careful guidance and helpful suggestions received from them during the course of this work. The assistance of Mr. R. E. Schock in the experimental work is also appreciated.

THE SERVICE RANGE OF FREQUENCY MODULATION

By

Murray G. Crosby

Engineering Department, R.C.A. Communications, Inc.

Summary—*An empirical ultra-high-frequency propagation formula is correlated with experimentally confirmed frequency modulation improvement factors to develop a formula for the determination of the relation between signal-noise ratio and distance in a frequency-modulation system. Formulas for the distance between the transmitter and the point of occurrence of the threshold of frequency-modulation improvement, as well as formulas for the signal-noise ratio occurring at this distance, are also developed.*

Results of calculations and listening tests are described which evaluate the signal-noise ratio gain obtained by applying pre-emphasis to the higher modulation frequencies of both amplitude and frequency modulation systems.

Examples using a typical set of conditions are given and discussed.

IT HAS BEEN shown that the propagation characteristics of frequency modulation are such as to confine the use of this type of modulation insofar as telephony is concerned to the ultra-high frequencies which do not use the ionosphere as a transmission medium.[1] Consequently, for most purposes, calculations of service range of a frequency modulation transmitter may be based totally on the propagation characteristics of ultra-high-frequency waves and in this paper that base will be used. Formulas and empirical data now available make it possible to calculate the field strength at the receiver if transmitter power, antenna heights, and distance are known. Consequently, the signal-noise ratio for amplitude modulation may be determined if the field strength of the noise is known. From this known amplitude-modulation signal-noise ratio, the signal-noise ratio obtainable with a given frequency-modulation system may be determined by multiplying the corresponding amplitude-modulation signal-noise ratios by the frequency-modulation improvement factors as given in the author's previously published paper[2] on frequency-modulation noise characteristics. It is the purpose of this paper to perform this correlation of these propagation and frequency-modulation improvement formulas and develop formulas so that from the known constants of the frequency-modulation system, the signal-noise ratio at a given distance may be directly calculated.

Reprinted from *RCA Review*, Jan. 1940.

Amplitude-Modulation Signal-Noise Ratio vs. Distance

The following empirical formula has been given by H. H. Beverage[3] for calculating the field strength when the receiver is within the optical distance of the transmitter:

$$E \text{ (r-m-s volts per meter)} = \frac{88\sqrt{W}\,ah}{\lambda D^2} \qquad (1)$$

where W = effective watts radiated = power in antenna times antenna power gain over a one-half wave dipole,
a = the receiving antenna height in meters,
h = the transmitter antenna height in meters,
D = the distance in meters,
λ = the wavelength in meters.

This formula is to be used for calculating the field strength for distances within the horizon only. For distances beyond the horizon, Beverage used a graphical method of plotting the curve of field strength versus distance. In this graphical method the field strength versus distance curve according to equation (1) was plotted for distances out to the horizon and then the curve was continued for distances beyond the horizon, but with a slope of $1/D^n$ instead of $1/D^2$. The exponent "n" was determined empirically and varies with frequency in the manner shown in Figure 1 which is reproduced from Beverage's paper.

In place of the graphical construction of the curve for distances beyond the horizon, the formula given by (1) may be revised to be applicable to all distances, whether they be within or outside of the horizon, as follows:

$$E \text{ (r-m-s volts per meter)} = \frac{88\sqrt{W}\,ah\,D_h^{n-2}}{\lambda D^n} \qquad (2)$$

in which the exponent "n" is equal to two for distances within the horizon and is chosen from the curve of Figure 1 for distances beyond the horizon. D_h is the distance to the horizon in meters and is equal to $3550\sqrt{h} + 3550\sqrt{a}$ where a and h are the receiving and transmitting heights in meters. Thus, where $D < D_h$, $n = 2$, and where $D > D_h$, n is taken from Figure 1.

When the units of the formula given by (2) are converted to feet, microvolts, miles, and megacycles, the formula becomes:

$$E \text{ (r-m-s microvolts per meter)} = \frac{0.01052\sqrt{W}\,ahf\,D_h^{n-2}}{D^n} \qquad (3)$$

where W = effective watts radiated = power in antenna times antenna power gain over a one-half wave dipole,

a = receiving antenna height in feet,

h = transmitting antenna height in feet,

D = distance in miles,

D_h = distance to the horizon in miles = $1.22\sqrt{h} + 1.22\sqrt{a}$,

f = frequency in megacycles.

In comparing the calculated curves with the experimental data, Beverage states that "scattering and absorption, even in open country,

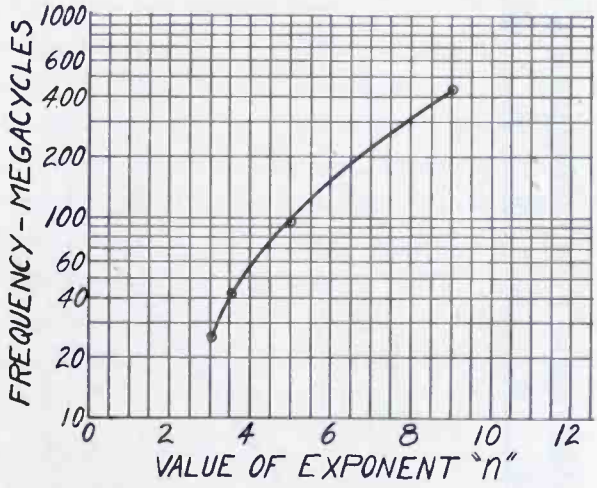


Fig. 1—Variation of exponent "n" in propagation formula when transmission is beyond the horizon.

tend to reduce the average intensity to something in the order of thirty to sixty per cent of the calculated value." In the following derivations an average experimental factor of forty-five per cent will be included to take into account this absorption and scattering.

It will be noted that this formula gives the average field intensity and does not take into account fading. More recent work by MacLean and Wickizer[4] shows the range of fading which may be expected for one set of transmission conditions. Thus, for transmission conditions reasonably close to the case treated by MacLean and Wickizer, the signal intensities at the fading minimums may be determined by applying a correction obtained from Figure 12 of the MacLean and Wickizer paper which is reproduced herewith as Figure 2. In the present paper, the formulas will be derived for the case of the average signal intensity and the fading correction will be applied to the examples given.

The peak carrier-noise ratio* obtained at a given distance may be determined by converting the voltage given by (3) into peak values (multiply by 1.414) and dividing by the peak noise voltage, N. N is the noise field strength as determined by means of a field-strength meter having a pass band characteristic equal to twice the band width of the audio spectrum which it is desired to receive. The experimental scattering and absorption factor may be taken into account by multi-

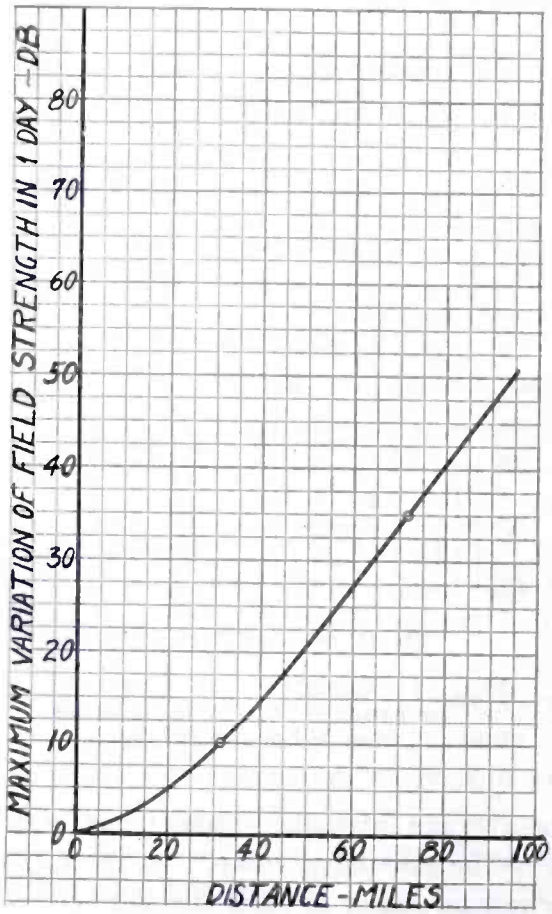


Fig. 2—Fading range of 50-megacycle transmission, transmitter antenna height = 1300 feet.

plying by 0.45. Thus, the amplitude-modulation peak carrier-noise ratio, which is approximately equal to the amplitude-modulation peak signal-noise ratio at the receiver output, for the case of one hundred per cent modulation, is given by:

* Throughout this paper carrier-noise ratio will refer to the ratio between the carrier and noise voltages as measured at the output of the intermediate-frequency channel of the receiver. Signal-noise ratio will refer to the ratio between the signal and noise voltages at the output terminals of the receiver. In an amplitude-modulation receiver the signal-noise ratio is usually substantially equal to the carrier-noise ratio, but in a frequency-modulation receiver the two quantities may differ greatly.

$$C/N = \frac{0.0067\sqrt{W}\,ahf\,D_h^{n-2}}{N\,D^n} = S_a/N_a \qquad (4)$$

where C = peak carrier-field strength,
N = peak noise-field strength as above defined,
S_a = peak audio-signal voltage, amplitude-modulation receiver,
N_a = peak audio-noise voltage, amplitude-modulation receiver.

Frequency-Modulation Signal-Noise Ratio vs. Distance

The author's previously cited paper[2] on frequency-modulation noise characteristics gives the improvement factors effected by a frequency-modulation system over an amplitude-modulation system. For the case of fluctuation noise the factor is $\sqrt{3}$ times the deviation ratio μ. (The deviation ratio, μ, is equal to F_d/F_a where F_d is the peak-frequency deviation due to modulation and F_a is the width of the audio channel of the system.) This factor holds for equal carriers fed to the two receivers and for the condition of a peak carrier at least 4 decibels above the peak noise in the frequency-modulation receiver intermediate-frequency channel. Thus, the signal-noise ratio at the output of the frequency-modulation receiver may be found by multiplying the signal-noise ratio given by (4) by the improvement factor or:

$$S_f/N_f = \sqrt{3}\,\mu\frac{0.0067\sqrt{W}\,ahf\,D_h^{n-2}}{N\,D^n}$$

$$= \frac{0.0116\,\mu\sqrt{W}\,ahf\,D_h^{n-2}}{N\,D^n} \text{ (fluctuation noise)} \qquad (5)$$

where S_f = peak audio-signal voltage, frequency-modulation receiver,
N_f = peak audio-noise voltage, frequency-modulation receiver.

When the received noise is impulse noise the improvement factor is equal to twice the deviation ratio or 2 μ. The corresponding signal-noise ratio is then:

$$S_f/N_f = \frac{0.0134\,\mu\sqrt{W}\,ahf\,D_h^{n-2}}{N\,D^n} \text{ (impulse noise)} \qquad (6)$$

It will be noted that in deriving the signal-noise ratio formulas of (5) and (6), the power gain normally effected by frequency modulation at the transmitter is automatically taken care of by the fact that W appears in both the amplitude- and frequency-modulation formulas. In the case of amplitude modulation the value of W used would normally be less than the corresponding value used in the frequency-

modulation formulas by a factor equal to the frequency-modulation power gain at the transmitter.

Effect of the Improvement Threshold

The presence of the phenomena called the "improvement threshold" places a rather definite maximum service range on a frequency-modulation system. The improvement threshold occurs at the point where the peak voltages of the noise and carrier in the intermediate-frequency channel of the frequency-modulation receiver are equal. The experimental work of the author's previously cited paper[2] shows that the full frequency-modulation improvement for fluctuation noise is not obtained until the carrier is at least four decibels above the improvement threshold or where the carrier-noise ratio is about four decibels. The nature of the improvement threshold is such that the signal-noise ratio drops rapidly as the carrier falls below the four decibel carrier-noise ratio. For the case of fluctuation noise, which is of a continuous nature, the noise smothers the signal in a manner which has been described in detail before[2]. Consequently the distance at which the improvement threshold occurs for fluctuation noise may be taken as the maximum service range of the frequency-modulation system. However, for impulse noise of the type which consists of sharp impulses having a low rate of recurrence, the situation is somewhat different. It has been pointed out to the author by V. D. Landon of RCA Manufacturing Company, that for the condition of no modulation present, if the receiver is carefully tuned so that the incoming carrier is approximately synchronized with the oscillation frequency of the impulse, the frequency-modulation improvement is maintained for impulse noise which is stronger than the carrier. This effect is shown in the experimentally determined curve of Figure 3 which also shows the effect of a slight detuning from the point of synchronism. For these curves, the peak carrier-noise ratio and the peak signal-noise ratios were measured by means of an oscilloscope coupled to the intermediate-frequency channel output and the audio output of the receiver, respectively. Curve B was taken with the receiver carefully tuned to the impulse noise minimum. Curve A was taken with the carrier detuned about 20 per cent of the maximum frequency deviation of the receiver used. It will be noted that the improvement threshold manifests itself sharply when the receiver is detuned. Hence, for all except the very low passages of modulation the threshold would be present since application of frequency modulation corresponds to a momentary detuning. However, the reduction of the noise during the idle periods of the modulation is undoubtedly very helpful.

The curves of Figure 3 also show the value of carrier-noise ratio required to obtain the full frequency-modulation improvement for impulse noise. It can be seen that for all practical purposes it can be assumed that the full improvement is obtained at a peak carrier-noise ratio of unity.

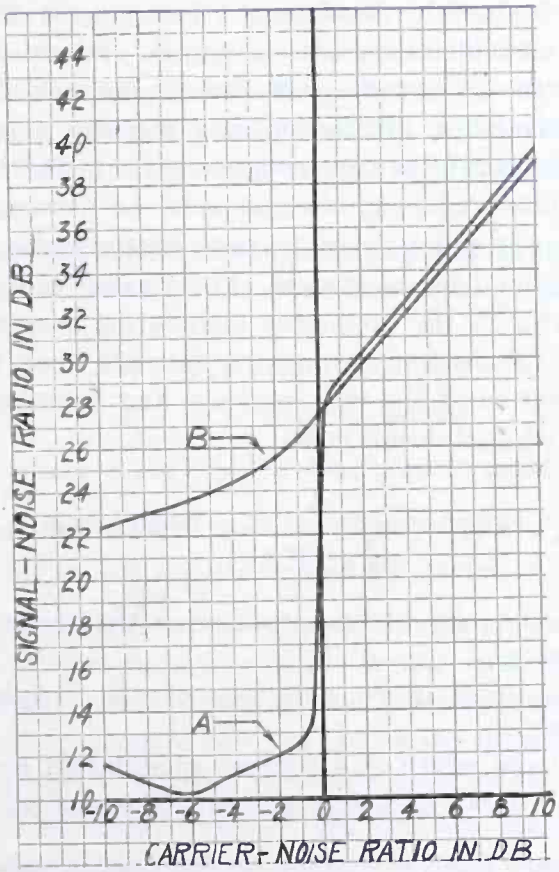


Fig. 3—Measured signal-noise ratio versus carrier-noise ratio characteristics with impulse noise. Receiver deviation ratio = 2.8. Curve A = receiver detuned by an amount equal to 20 per cent of the maximum frequency deviation. Curve B = receiver tuned for impulse noise minimum.

A further characteristic of the effect of impulse noise when it is stronger than the frequency-modulation carrier is the noise-silencing action which is present. This produces a sort of a minimum signal-noise ratio for carrier strengths below the improvement threshold when the carrier and pulse frequency are not synchronized. It has been shown[2] that this minimum signal-noise ratio is equal to the deviation ratio of the system. Thus, it may be seen that although the signal-noise ratio will become poorer below the improvement threshold, the signal

may still be serviceable. On the other hand, if the impulse noise is of a continuous nature like fluctuation noise, the noise discriminates against the signal below the improvement threshold and, therefore, smothers the signal in the same manner that fluctuation noise does.

The formula for the distance at which the improvement threshold occurs may be obtained by developing a formula for the distance from the transmitter at which carrier-noise ratios of four and zero decibels are obtained in the intermediate-frequency channel of the frequency-modulation receiver for fluctuation and impulse noises respectively. First—the equation for the carrier-noise ratio in the frequency-modulation intermediate-frequency channel must be determined. This may be derived from (4) which gives the peak carrier-noise ratio as defined for an amplitude-modulation system. For a given radiated carrier, the carrier-noise ratio in the frequency-modulation receiver is less than that defined by (4) by a factor which depends upon the ratio of the effective band widths of the two receivers and upon the type of noise.

For the case of fluctuation noise,

$$C_o/N_o = C/N \times \frac{1}{(F_{fi}/2F_a)^{1/2}} \quad \text{(fluctuation noise)} \qquad (7)$$

where C_o = carrier level in i-f amplifier of the frequency-modulation receiver,

N_o = noise level in i-f amplifier of the frequency-modulation receiver,

F_{fi} = i-f band width of frequency-modulation receiver,

F_a = band width of audio spectrum it is desired to receive.

For convenience let $Z_f = F_{fi}/2F_a$ for fluctuation noise. Substituting in (7), we obtain:

$$\frac{C_o}{N_o} = \frac{C}{N} \times \frac{1}{\sqrt{Z_f}} \quad \text{(fluctuation noise)} \qquad (8)$$

The equation corresponding to (8) for impulse noise, which varies directly with the band width instead of as the square-root of the band width, will be:

$$\frac{C_o}{N_o} = \frac{C}{N}\,\frac{1}{Z_i} \quad \text{(impulse noise)} \qquad (9)$$

where $Z_i = F_{fi}/2F_a$ for impulse noise.

The factors Z_f and Z_i, which are the ratios between one-half the intermediate-frequency channel width and the audio channel width of the frequency-modulation receiver for the two types of noise, are

noise-determining ratios and, therefore, must be expressed in terms of *equivalent* channel widths. The equivalent channel widths are different for the two types of noise; hence, a separate factor is used for each type of noise. For fluctuation noise, the equivalent channel width is determined by dividing the area under the energy response curve (the selectivity curve plotted with the ordinates squared) by the height of the curve at resonance. This gives the width of the rectangular channel which would be equivalent to the actual round-topped channel. For impulse noise, the equivalent channel width is determined by dividing the area under the amplitude response curve by the height of the curve at resonance.

A further simplification of the ratios, Z_f and Z_i, may be effected by expressing them in terms of the deviation ratio, μ. Thus,

$$Z_f = K_f \mu \text{ (fluctuation noise)} \tag{10}$$

$$Z_i = K_i \mu \text{ (impulse noise)} \tag{11}$$

$$\text{where } K_f = F_{fi}/2F_d \text{ (fluctuation noise)} \tag{10a}$$

$$\text{and } K_i = F_{fi}/2F_d \text{ (impulse noise)} \tag{11a}$$

in which F_d = maximum applied frequency deviation.

The factors K_f and K_i express the ratio between the equivalent band widths of the intermediate-frequency channel and the total plus and minus peak frequency deviation of the frequency-modulation system. In other words they specify how far out on the intermediate-frequency selectivity curve the frequency deviation may be carried.

In order to determine the formula for the carrier-noise ratio in the frequency-modulation intermediate-frequency channel for the case of fluctuation noise, (4) and (10) may be substituted in (8) which gives:

$$\frac{C_o}{N_o} = \frac{0.0067\sqrt{W}\,ahf\,D_h^{n-2}}{N\,D^n\sqrt{K_f\mu}} \text{ (fluctuation noise, peak values)} \tag{12}$$

Likewise, for impulse noise (4) and (11) may be substituted in (9) which gives:

$$\frac{C_o}{N_o} = \frac{0.0067\sqrt{W}\,ahf\,D_h^{n-2}}{K_i\,\mu\,D^n\,N} \text{ (impulse noise, peak values)} \tag{13}$$

(12) and (13) give the carrier-noise ratio in the intermediate-frequency channel of the frequency-modulation receiver for the two types of noise. Since it is known that the improvement threshold

occurs when these carrier-noise ratios are equal to four and zero decibels, respectively, the distance at which the improvement threshold will occur for a given set of transmission conditions may be determined by equating (12) to 1.585 (four decibels) and (13) to unity (zero decibels), and solving for the distance. Thus,

$$D_i = \left(\frac{0.0042 \sqrt{W}\, ahf\, D_h^{n-2}}{N \sqrt{K_f \mu}} \right)^{1/n} \quad \text{(fluctuation noise)} \qquad (14)$$

$$D_i = \left(\frac{0.0067 \sqrt{W}\, ahf\, D_h^{n-2}}{N K_i \mu} \right)^{1/n} \quad \text{(impulse noise)} \qquad (15)$$

in which D_i indicates the distance at which the improvement threshold occurs.

A study of (14) and (15) shows the effect of a variation of the transmission conditions. For both types of noise the distance to the improvement threshold is directly proportional to the $1/2n$ power of the watts radiated. Hence, for the higher radiation frequencies, where the exponent "n" is large, the improvement threshold distance increases more slowly with increase in power. For fluctuation noise, the improvement threshold distance is inversely proportional to the $1/2n$ power of the product of the factor K_f and the deviation ratio. Thus, as the deviation ratio is increased, the improvement threshold distance decreases and decreases at a less rapid rate when it is beyond the horizon and for the higher radiation frequencies where the exponent "n" is larger. The importance of making the factor K_f as small as possible, by carrying the frequency deviation out on the selectivity curve as far as possible, is also indicated. When the noise is impulse noise (formula 15) the improvement threshold distance is inversely proportional to the "n"th root of the product of the factor K_i and the deviation ratio. Consequently, with this type of noise, the improvement threshold distance decreases more rapidly as these factors are increased.

The distances given by (14) and (15) may be substituted in (5) and (6) to find the frequency-modulation signal-noise ratio existing at the improvement threshold as follows:

$$S_f/N_f \ (at\ D_i) = 1.585 \sqrt{3}\, \mu \sqrt{K_f \mu} = 2.74 \sqrt{K_f}\, \mu^{3/2} \ \text{(fluctuation noise)} \qquad (16)$$

$$S_f/N_f \ (at\ D_i) = 2 \mu K_i \mu = 2 K_i \mu^2 \ \text{(impulse noise)} \qquad (17)$$

Effect of Pre-Emphasizing the Higher Modulation Frequencies

As has been pointed out for the case of amplitude modulation[5], the use of pre-emphasis circuit at the transmitter and a de-emphasis circuit at the receiver produces an overall gain in signal-noise ratio. This gain depends upon the fact that the higher modulation frequencies of voice and program material are of such a small amplitude that their accentuation does not increase the peak voltage of the applied modulating wave as much as the restoring circuit at the receiver reduces the noise. Thus, when the pre-emphasis circuit is inserted at the transmitter, the peak voltage of the modulating wave may increase somewhat so that the modulation level must be lowered, but this loss at the transmitter is overshadowed by the gain at the receiver. Consequently there are two quantities which must be evaluated—the loss at the transmitter and the gain at the receiver.

In order to determine the loss introduced by the pre-emphasis circuit at the transmitter, the following experimental observations were made: The equivalent of a two-string oscillograph was arranged by using two electronically switched amplifiers to feed the vertical plates of an oscilloscope. The outputs of the amplifiers were common, with one of the separate inputs being fed by program material through a pre-emphasis circuit. The gain of the two amplifiers was equalized at a low modulation frequency where the pre-emphasis circuit was not effective. The two amplifiers were alternately switched on by means of a 60-cycle square-wave form and the oscilloscope sweep circuit synchronized with the 60 cycles. The resulting pattern on the oscilloscope screen consisted of two segments of sweep one of which was actuated directly from the program material and the other through the pre-emphasis circuit. Hence, an accurate method of comparing the peak voltages of the two waves could be provided by inserting an attenuator in the pre-emphasized circuit and setting for equal peak voltages indicated by the two traces on the oscilloscope.

The observations made with the electronic-switch oscillograph covered all types of program material and were made for two audio-band widths of 5 and 12 kilocycles. The results of the observations indicated that in general the insertion of the pre-emphasis circuit increases the peak voltage of the level about 2.5 decibels for both the 5 and 12-kilocycle band widths. On certain material such as guitar, harmonica, and piano solos, the level is raised about 4.5 decibels total, but the occasion of such rises is rather infrequent and their duration very short. Hence, a permanent attenuation of 2.5 decibels might be inserted and the volume-limiting equipment relied upon to take care of the occasional higher peaks.

The action of the de-emphasis circuit in reducing the noise at the receiver is somewhat greater for the case of frequency modulation than it is for amplitude modulation. The reason for this is shown in Figures 4 and 5 which show the noise spectrums obtained in the output of the amplitude and frequency-modulation receivers, respectively, with and without the use of a de-emphasis circuit such as would be used with a pre-emphasizing circuit in accordance with R.M.A. Television Transmission Standard M9-218 at the transmitter. In the case of amplitude modulation, the normally evenly distributed noise is concentrated at the lower modulation frequencies. In the case of frequency modulation, the triangular noise spectrum is changed to a spectrum

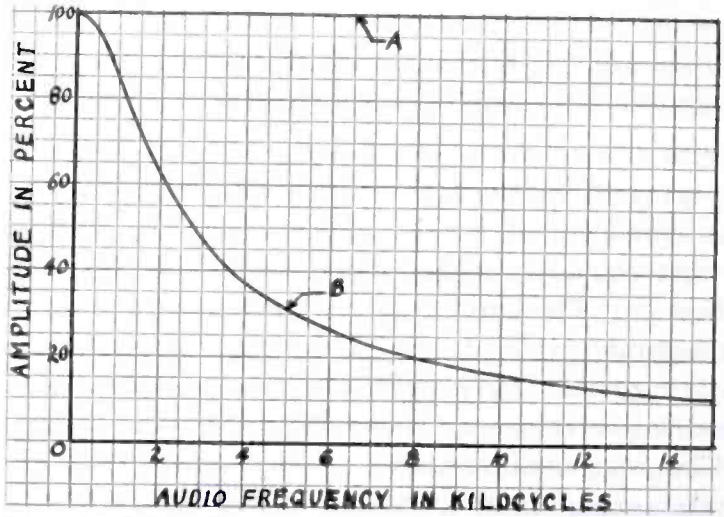


Fig. 4—Amplitude-modulation audio noise spectrums.
Curve A = no de-emphasis used.
Curve B = de-emphasis in accordance with R.M.A. Television Transmission Standard M9-218.

which is practically flat except for the falling off at the lower modulation frequencies.

In determining the relative figures of merit for the noise spectrums of Figures 4 and 5, or, in other words, the noise gains produced by the de-emphasis circuit, there are two determinations which are of importance. The first is the objective comparison which has to do with the relative strengths of the noise as would be measured on a meter. The second is the subjective comparison which takes into consideration the manner of utilizing the signal in the presence of the noise. For program or voice reception, the subjective comparison would be determined by a listening test.

The objective comparison of the noise spectrums of Figures 4 and 5 may be calculated for fluctuation noise by comparing the squared-ordinates areas of the spectrums. Such a comparison gives the ratio

of the energies passed by the two spectrums; the root-mean-square voltage ratio is the square-root of this area ratio. The corresponding comparison for impulse noise may be calculated by comparing the areas of the spectrums. The ratio of the areas of the spectrums gives the peak voltage ratio directly for impulse noise. These areas may be obtained by integration or they may be plotted and a planimeter used. The curves of Figure 6 show the results of such a determination by the use of integration. These curves give the signal-noise ratio gain effected by the de-emphasis circuit alone. To obtain the overall gain due to pre-emphasis, the transmitter loss of 2.5 decibels must be subtracted from the values indicated by the curves.

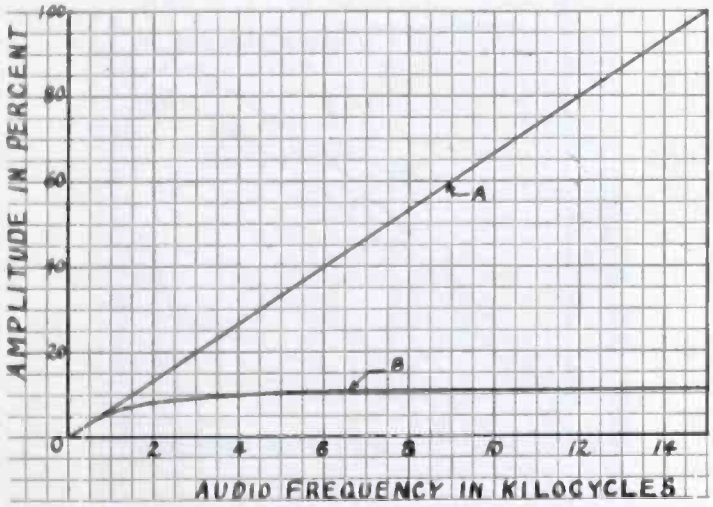


Fig. 5—Frequency-modulation audio noise spectrums.
Curve A = no de-emphasis used.
Curve B = de-emphasis in accordance with R.M.A. Television Transmission Standard M9-218.

In order to obtain the subjective gain effected by the de-emphasis circuit, listening tests were conducted in which the de-emphasis circuit was switched in and out and the annoyance effect of different types of noise compared. Both fluctuation noise and impulse noise were obtained from amplitude- and frequency-modulation receivers to mix with program to produce a signal-noise ratio. The de-emphasis circuit was then switched in and out of the noise while an attenuator was varied to balance the annoyance effect. Audio band widths of 5 and 12 kilocycles were used. The averaged observations of two observers indicated the rather unexpected result that for program and music reception, the subjective effect is practically the same as the objective effect. That is to say that the objective gains as portrayed by Figure 6 may also be taken as the subjective gains that would be realized in a

listening test and the overall gain is obtained by subtracting the transmitter loss of 2.5 decibels from the values obtained from Figure 6.

For speech reception, where the noises were balanced for equal intelligibility, it was found that there was little or no gain effected by the de-emphasis circuit. For the case of amplitude-modulation noise, the use of pre-emphasis would apparently entail a slight loss in intelligibility when the 2.5 decibel transmitter loss was subtracted. With frequency-modulation noise, an intelligibility gain of a few decibels would be realized.

In connection with the listening tests another interesting observa-

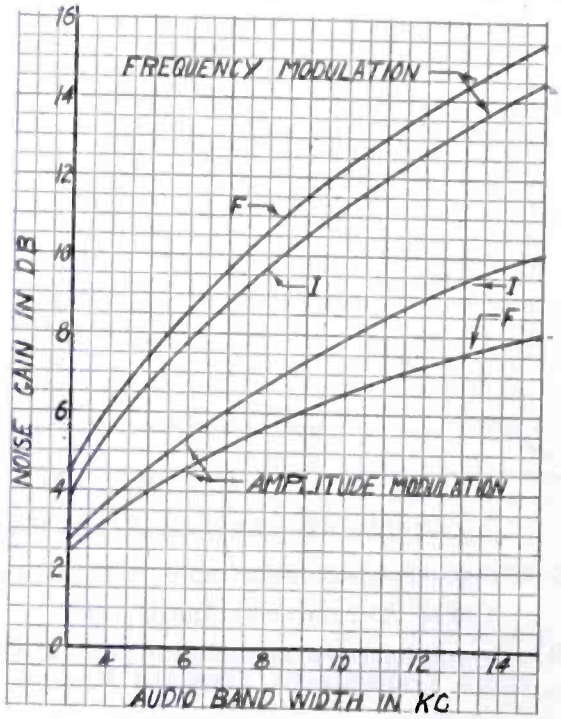


Fig. 6—Calculated noise gains effected by the de-emphasis circuit.
Curves F = fluctuation noise.
Curves I = impulse noise.

tion was made regarding the relative annoying effects of the triangular frequency-modulation noise spectrum and the rectangular amplitude-modulation noise spectrum using tube hiss as a noise source. In this test the noises were first balanced with a meter and then they were mixed with speech and the relative levels readjusted until they impaired the intelligibility of the speech the same amount. The averaged observations of three observers showed that about 8 decibels more noise could be tolerated with the triangular frequency-modulation noise than with the rectangular amplitude-modulation noise.

Example

The curves of Figures 7, 8, 9, and 10 have been calculated for the following assumed transmission conditions for a high-fidelity broadcast system:

Transmitting antenna height	= 800 feet.
Receiving antenna height	= 30 feet.
Audio channel	= 15 kilocycles.
Frequency	= 42 megacycles.
Maximum frequency deviations	= 20 and 75 kilocycles.

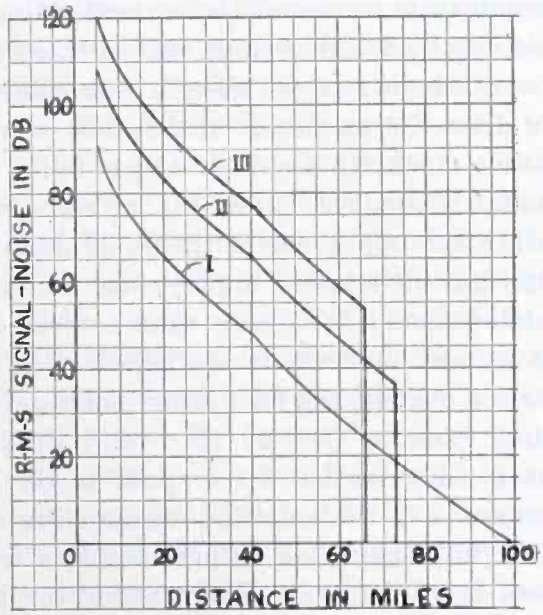


Fig. 7—R-M-S signal-noise ratio versus distance. Fluctuation noise = 1 peak microvolt per meter.
Curve I = amplitude modulation, 500 watts radiated.
Curve II = frequency modulation, 1000 watts radiated, maximum frequency deviation = 20 kilocycles.
Curve III = frequency modulation, 1000 watts radiated, maximum frequency deviation = 75 kilocycles.

Powers of one and 100 kilowatts radiated were used for frequency modulation and these powers were halved for the corresponding amplitude-modulation calculations. This two-to-one power gain effected at the frequency-modulation transmitter was taken instead of the usual four-to-one factor since it represents the gain that would be effected by a frequency-modulation system over the most efficient amplitude-modulation system which is the high-level type of modulation system. If the modulator tubes were paralleled with the final amplifier tubes

in such a system, the increase in power would be two-to-one. The usual factor of four-to-one assumes the more inefficient types of low-level modulation.

The values of signal-noise ratio obtained for all of the curves were corrected by subtracting one-half the maximum fading range as portrayed by Figure 2. Formula (4) was used to calculate the average signal-noise ratio for amplitude modulation and the fading correction was applied to obtain the minimum signal-noise ratio. It was assumed that the transmission conditions of the fading correction curve were near enough to the conditions of this example to allow this direct correction without interpolation.

Where the propagation curves of the corresponding amplitude-modulation system were to be determined as well as those of the frequency-modulation systems, which was the case in this example, the simplest procedure was to draw the amplitude-modulation curves and then construct the frequency-modulation curves which follow lines parallel to the amplitude-modulation curve, but at higher signal-noise ratio levels for distances within the improvement threshold distances. To do this the average amplitude-modulation signal-noise ratios for fluctuation noise were calculated from (4). These were corrected for fading and the pre-emphasis gain of 5.6 decibels was added. For the low-deviation frequency-modulation system, the fluctuation noise gain is equal to 1.73 times the deviation ratio, μ ($\mu = 20/15 = 1.33$) or 7.2 decibels. The use of pre-emphasis adds another 7.4 decibels to the frequency-modulation gain as compared to the amplitude-modulation system with pre-emphasis since the total gain due to pre-emphasis on the frequency-modulation system is 13 decibels. The transmitter power gain adds another 3 decibels. Hence the curve for frequency modulation with a deviation of 20 kilocycles is $7.2 + 7.4 + 3 = 17.6$ decibels higher than the corresponding amplitude-modulation curve for the region within the threshold of improvement distance of the frequency-modulation system. The frequency-modulation system with a 75-kilocycle deviation has a larger deviation ratio ($\mu = 75/15 = 5$) so that its total gain over the amplitude-modulation system with pre-emphasis is $18.7 + 7.4 + 3 = 29.1$ decibels.

For the case of impulse noise, the frequency-modulation gains are equal to twice the deviation ratio and the pre-emphasis gains are also different. For amplitude modulation, the pre-emphasis gain is 7.5 decibels and for frequency modulation it is 4.5 decibels more or a total of 12 decibels. Hence, the gain of the frequency-modulation system using a 20-kilocycle deviation is $8.5 + 4.5 + 3 = 16$ decibels. The corresponding gain for the system using a deviation of 75 kilocycles is $20 + 4.5 + 3 = 27.5$ decibels.

In order to determine the signal-noise ratios at the improvement threshold distances, the gains due to pre-emphasis must be added to the signal-noise ratios calculated from (16) and (17). The pre-emphasis gains to be added in this case are the total gains of 13 and 12 decibels for fluctuation and impulse noise respectively. Before formulas (16) and (17) can be applied, the factors K_f and K_i must be evaluated. These factors are the ratios between the equivalent band widths of the frequency-modulation receiver intermediate-frequency channel and the total plus and minus frequency deviation. It is obvious that, in order to obtain the greatest distance to the improvement threshold, the equivalent band width must be made as small as the frequency deviation of the system will allow. The limitations encountered are the introduction of harmonic distortion and the introduction of amplitude modulation due to the frequency variation exceeding the flat-topped portion of the selectivity characteristic. In practice the limiter tends to take care of this departure from a flat-topped selectivity characteristic. Some rather preliminary measurements have indicated that the deviation may extend out to about 2.5 decibels down (down to 75 per cent) on the sides of the selectivity curve without producing harmonics which are too high for high-fidelity reception.

A study of typical intermediate-frequency equivalent band widths has shown that the band width 2.5 decibels down is about 90 per cent of the equivalent band width for fluctuation noise and about 78 per cent of that for impulse noise. The factors K_f and K_i are equal to the reciprocals of these percentages or 1.1 and 1.3, respectively.

The above evaluations of K_f and K_i are based on somewhat incomplete distortion measurements and assume no guard band to take care of tuning drift in the case of the frequency-modulation systems. Further work is undoubtedly necessary in these respects, but it is believed that this evaluation will serve the purposes of this paper.

With these evaluations of K_f and K_i, formulas (16) and (17) simplify to:

$$S_f/N_f \text{ (at } D_i) = 2.9\,\mu^{3/2} \text{ (fluctuation)} \qquad (16a)$$

$$S_f/N_f \text{ (at } D_i) = 2.6\,\mu^2 \quad \text{(impulse)} \qquad (17a)$$

Applying (16a) to the system with a deviation of 20 kilocycles and adding the gain due to pre-emphasis of 13 decibels gives 25.9 decibels for the peak signal-noise ratio at the improvement threshold distance. The same calculation for the case of a 75-kilocycle deviation gives 43.2 decibels. For impulse noise (17a) is used and a pre-emphasis gain of 12 decibels is added to give ratios of 25.2 and 48.3 decibels for the two values of frequency deviation.

The curves of Figures 7 and 8 for fluctuation noise assume a thermionic agitation and tube-hiss noise level equivalent to one peak microvolt per meter (0.224 r-m-s microvolts per meter). This corresponds to an r-m-s noise voltage of 0.52 microvolts in series with a dummy antenna at the input terminals of the amplitude-modulation receiver. Such a noise voltage is about that which would be obtained with good design using an 1852 radio-frequency amplifier tube. All of the ratios for these curves were converted from peak to r-m-s ratios to correspond to the general practice in considering this type of noise. This was done by adding 10 decibels to the ratios. The figure of 10

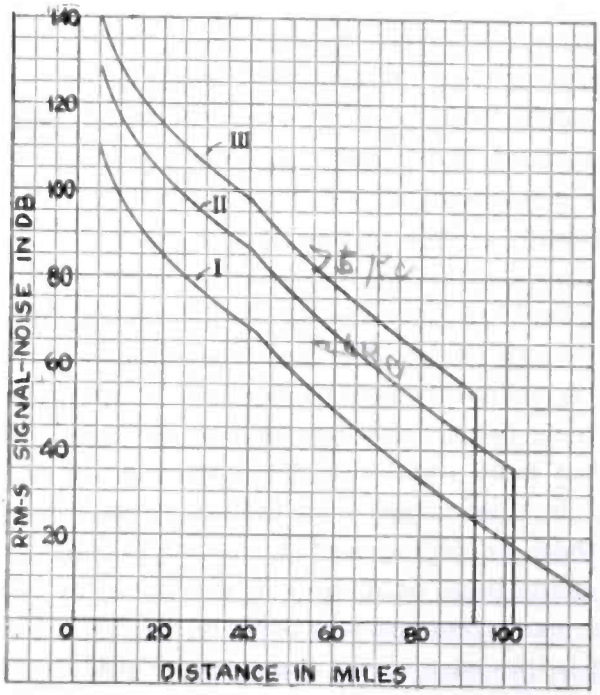


Fig. 8—Same as Fig. 7, but with amplitude modulation radiated power = 50 kilowatts, frequency-modulation radiated power = 100 kilowatts.

decibels is obtained from the data of the author's previous paper[2]. In that paper the crest factor of fluctuation noise is evaluated at 13 decibels. Subtracting the 3 decibel crest factor of the signal gives what might be termed the crest factor of the signal-noise ratio as 10 decibels.

The curves of Figures 9 and 10 for impulse noise assume a noise field intensity of 100 peak microvolts per meter for an effective band width of 30 kilocycles (audio band width = 15 kilocycles). This intensity is about that which would be received with horizontal polarization from the ignition system of the average automobile at a location about 125 feet from the road over which the automobile travels. The curves for this type of noise have been plotted with peak signal-noise ratios

since root-mean-square values have little significance due to the very high and variable crest factor.

At distances beyond the improvement threshold for impulse noise, the curves are plotted for the condition of full modulation in which the improvement threshold manifests itself. Under this condition, the signal-noise ratio is limited to a value which is equal to the deviation ratio when pre-emphasis is not being used. It will be remembered that this ratio is what might be termed the "silencing" signal-noise ratio since the noise tends to punch holes in the signal, but when the noise pulses have a relatively slow rate of recurrence, this condition is quite

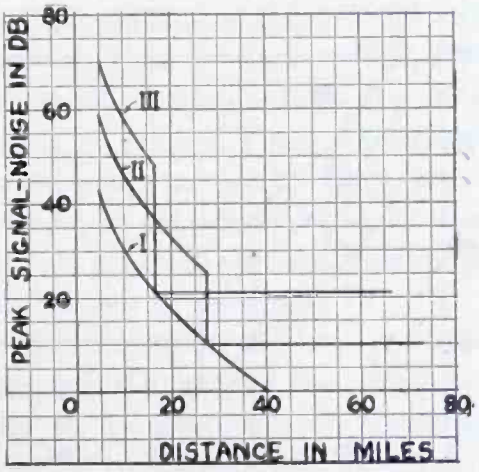


Fig. 9—Peak signal-noise ratio versus distance. Impulse noise = 100 peak microvolts per meter.
Curve I = amplitude modulation, 500 watts radiated.
Curve II = frequency modulation, 1000 watts radiated, maximum frequency deviation = 20 kilocycles.
Curve III = frequency modulation, 1000 watts radiated, maximum frequency deviation = 75 kilocycles.

tolerable. On the other hand, if the pulses have a high rate of recurrence, the noise tends to smother the signal and service is limited to the improvement threshold distance as in the case of fluctuation noise. When pre-emphasis is being used, a gain is added to this limited signal-noise ratio which is equal to the pre-emphasis gain of 7.5 decibels for amplitude-modulation noise. The pre-emphasis gain for amplitude modulation noise is used in this case since the noise in the silencing condition has lost its triangular spectrum characteristic of the frequency-modulation noise received above the improvement threshold. Thus, for the system with a 20-kilocycle deviation, the signal-noise ratio in the silencing condition beyond the improvement threshold dis-

tance is $2.5 + 7.5 = 10$ decibels. For the system with a 75-kilocycle deviation it is $14 + 7.5 = 21.5$ decibels.

At distances beyond the improvement threshold for impulse noise when the modulation is in the idle condition, the signal-noise ratios are somewhat higher than those shown by the curves due to the synchronization effect pointed out by V. D. Landon. This results in a considerable reduction of the annoyance effect of the noise. Hence, the curves may be taken as somewhat pessimistic, but satisfactory for comparison purposes.

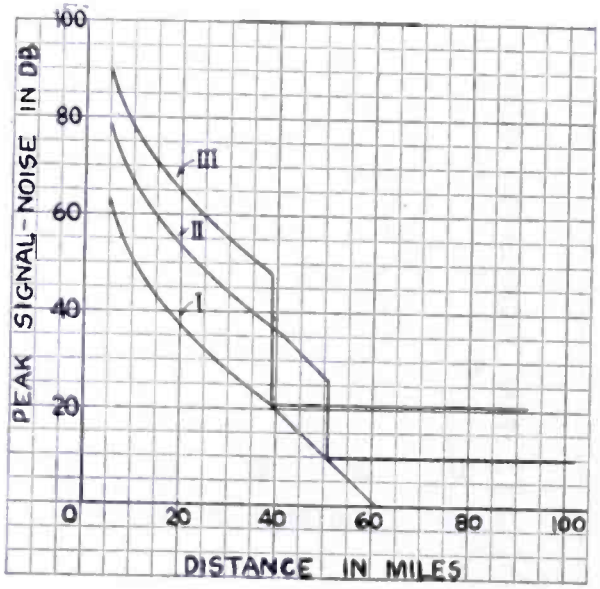


Fig. 10—Same as Figure 9, but with amplitude-modulation radiated power = 50 kilowatts, frequency-modulation radiated power = 100 kilowatts.

Discussion

It is apparent from the curves that the maximum distance is served if the maximum frequency deviation is such that the minimum tolerable signal-noise ratio exists at the improvement threshold distance. Hence, the choice of the optimum deviation hinges on the definition of the minimum tolerable signal-noise ratio. If the figure of 30 decibels root-mean-square for fluctuation noise is taken as acceptable (this figure has appeared in the literature[6,7] as a commercially satisfactory signal-noise ratio), it is seen from Figures 7 and 8 that a maximum deviation of 20 kilocycles is more than adequate for all distances out to the improvement threshold distance since the lowest signal-noise ratio, which occurs at the improvement threshold distance for this type of noise, is 36 decibels. A deviation of less than

20 kilocycles is probably inadvisable since this would prevent the possible future transmission of an audio fidelity of zero to 20 kilocycles.

It is seen from Figures 7 and 8 for fluctuation noise that at the distance corresponding to the improvement threshold for a deviation of 75 kilocycles, a broadcasting service using a deviation of 20 kilocycles yields a signal-noise ratio of 42 decibels. For the same distance, a service using a 75-kilocycle deviation yields a signal-noise ratio of 53 decibels. At this point it is apparent that a 20-kilocycle deviation produces a signal-noise ratio which is comfortably above what would be considered commercially satisfactory, with an attendant conservation of band width in the available portion of the frequency spectrum. However, if it were assumed that a signal-noise ratio of 53 decibels is necessary, this ratio may be obtained with a deviation of 20 kilocycles at a somewhat shorter distance from the transmitter. For the conditions of the curves of Figures 7 and 8 this distance would be 20 per cent shorter. On the other hand, a 20-kilocycle deviation is capable of furnishing what would be considered better than acceptable service out to a distance which is about 10 per cent greater than the maximum service range for a 75-kilocycle deviation.

When the noise is impulse noise as portrayed by Figures 9 and 10, it can be seen that the general shape of the curves are similar to those for fluctuation noise. If the frequency of recurrence happens to be high so that the noise is continuous, the silencing properties of the frequency modulation cannot be taken advantage of. Instead, the noise, if stronger than the carrie , will depress the signal so as to smother it. For this type of noise a low-deviation service will have a greater range than a high-deviation service by an amount that is considerably larger than the corresponding case for fluctuation noise.

If the impulse noise happens to be automobile ignition where the rate of recurrence of the pulses is rather infrequent and the time duration of the impulses short, the noise-silencing properties of the frequency-modulation system may be taken advantage of when the noise is stronger than the carrier. With this type of noise, the effectiveness of the silencing action increases as the deviation is increased. However, for this type of noise, even a system with a low deviation produces a noise output which has a rather low annoyance value.

It is apparent that the service obtainable where fluctuation noise predominates, may be predicted with a fair degree of accuracy. However, owing to the highly variable character and distribution of impulse noise, predictions are difficult regarding this type of noise. Hence, it is felt that it is highly desirable to conduct field tests to further

compare the results of different deviations under actual service conditions.

Conclusions

From the above considerations, it seems apparent that if fluctuation noise is the primary limitation, a frequency-modulation broadcasting system using a deviation of 20 kilocycles will produce a signal-noise ratio greater than the values normally considered acceptable, for all distances out to the distance at which the improvement threshold occurs. If the limitation is impulse noise a similar relationship exists, but it is felt that further field work is desirable to study the results under actual service conditions.

The effect of increasing the deviation of the system is to reduce the distance at which the threshold effect is realized, but at the same time the signal-noise ratio at distances equal to or less than that distance is improved.

The interests of band width conservation are, of course, best served by a choice of the lowest value of deviation that will yield an acceptable signal-noise ratio out to the threshold distance. It would also seem that the same choice will yield the desired service at the lowest cost.

In order to evaluate carefully these effects under a large number of actual service conditions, a series of field tests are being undertaken, using receivers designed for optimum performance at each of the different values of deviation under consideration.

Acknowledgment

The guidance and helpful suggestions of Mr. H. O. Peterson, and the assistance of Mr. R. E. Schock in the experimental work of this paper, are gratefully acknowledged.

References

[1] Murray G. Crosby, Frequency-Modulation Propagation Characteristics, *Proc. I.R.E.*, Vol. 24, No. 6; June (1936). (See Page 143.)

[2] Murray G. Crosby, Frequency-Modulation Noise Characteristics, *Proc. I.R.E.*, Vol. 25, No. 24; April (1937). (See Page 160.)

[3] H. H. Beverage, Some Notes on Ultra-High-Frequency Propagation, *RCA Review*, Vol. 1, No. 3, January (1937).

[4] K. G. MacLean and G. S. Wickizer, Notes on the Random Fading of 50-Megacycle Signals Over Nonoptical Paths, *Proc. I.R.E.*, Vol. 27, No. 8; August (1939).

[5] In an unpublished engineering report written in July (1933) and entitled: "Tests of a System for Reducing Interfering Noises in Radio Transmission", Messrs. I. Wolff, G. L. Beers, and L. F. Jones of RCA Manufacturing Company describe tests which indicated improvements of from 8 to 10 decibels in an amplitude-modulation program system.

[6] I. R. Weir, Comparative Field Tests of Frequency-Modulation and Amplitude-Modulation Transmitters, *Proc. The Radio Club of America*, Vol. 16, No. 2; July (1939).

[7] C. V. Aggers, Dudley E. Foster, and C. S. Young, Instruments and Methods of Measuring Radio Noise, *A.I.E.E. Technical Paper 39-144*, Preprinted for A.I.E.E. Pacific Coast Convention, San Francisco, Calif., June 26-30, 1939.

PRACTICAL APPLICATION OF AN ULTRA-HIGH-FREQUENCY RADIO-RELAY CIRCUIT*

By

J. Ernest Smith, Fred H. Kroger, and R. W. George

R.C.A. Communications, Inc., New York, N. Y.

Summary—The utilization of an ultra-high-frequency radio circuit for the transmission of telegraph, teletype printer, and facsimile signals is described. The operating procedure is stressed throughout, particularly with regard to equipment maintenance tests and the methods employed to determine the conditions of the circuit such as degree of modulation, signal-to-noise ratio, etc. Considerations are presented with respect to the most efficient division of the total modulation band into the communication channels, as well as the signal-to-noise ratios required for the different types of service. It is found that fading, static, and weather conditions at these frequencies are of little importance as to their effect on the economic use of the circuit. However, diathermy machines and similar sources of disturbance are troublesome and their effect must be minimized. Experience during the past year and a half indicates that the dependability of the ultra-high-frequency radio circuit is of a high order.

Introduction

While the theory of ultra-high-frequency propagation has been studied for several years, there is little information available on the economic utilization of these frequencies for commercial purposes. To obtain such data, the engineers of R.C.A. Communications, Inc., have designed and installed an experimental three-link two-way ultra-high-frequency radio relay between New York and Philadelphia. For the past year and a half this circuit has been utilized by the traffic department as a regular communication facility. This paper will describe the circuit equipment and the means employed to insure dependable operation.

Transmitter

In general, the transmitter construction followed conventional design. Some features however are unique and worthy of detailed description. A close-up view of the transmitter is shown in Figure 1. The rack on the left houses the high-voltage direct-current supply. The center unit contains the modulator and five test meters. The tank-like structure at the right houses the resonant line control, master oscillator, and power amplifier.

* Decimal classification: R480. Presented before I.R.E. New York Meeting, April 7, 1937.

Reprinted from *Proc. I.R.E.*, November, 1938.

The five meters together with their proper cables can be plugged into any of the nineteen jacks visible in the picture. These measurements form a part of the daily transmitter test schedule. Each jack is marked with its normal reading so that incorrect operation is easily determined. The racks are built on rollers and may be pulled forward on special skids thus facilitating test and repair work.

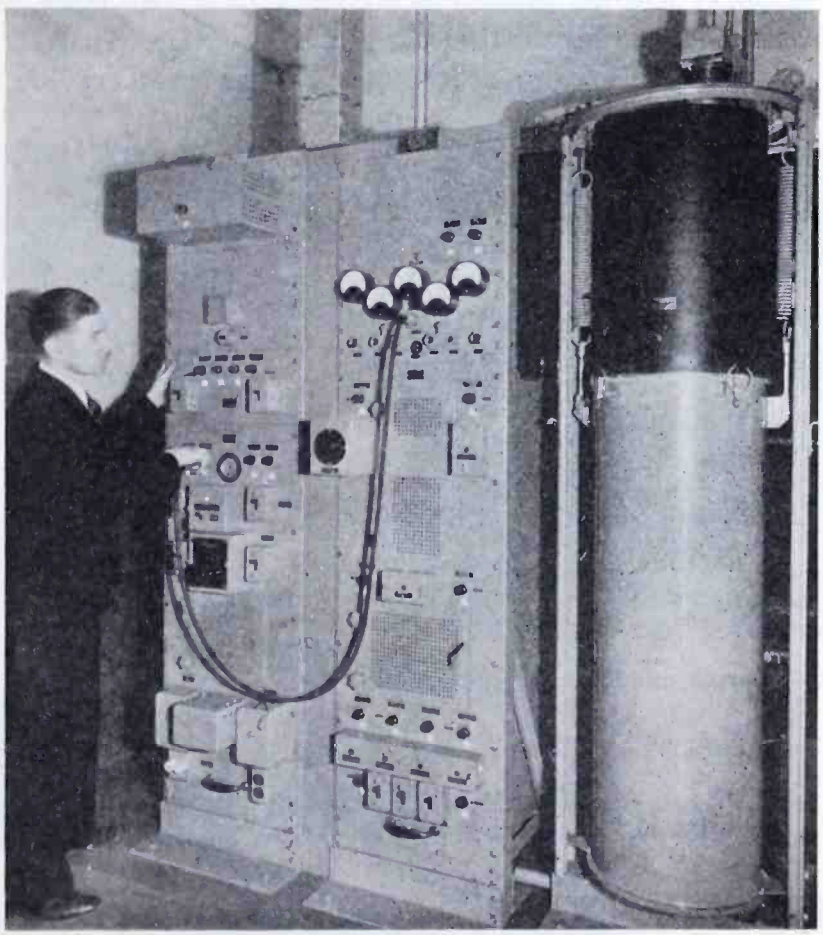


Fig. 1—Front view of an ultra-high-frequency transmitter (installed).

The radio-frequency unit in the tank structure is suspended by springs to reduce vibration. Holes through the black hood allow adjustment of the radio-frequency circuits by means of tuning rods. The small tubular projections fitted with caps and chains permit the insertion of the special thermojunction microammeter shown in Figure 2. This meter, when held against the glass envelope of a vacuum tube, can be used to predict the life of the tube since experience has shown that subnormal temperature is a sign of low filament emission. The upper portion of the radio-frequency unit is painted black since it was found that the difference in heat radiation properties of black and

green paint was sufficient to change the original calibration determined for the test model. The locations of the radio-frequency tubes and their associated circuits are illustrated in Figure 3.

The master oscillator has a tuned grid circuit consisting of a concentric line having a Q in the order of 15,000. A buffer stage was omitted since it was found that without a buffer stage between the power amplifier and the master oscillator there is frequency modulation of only 0.005 per cent at 100 per cent modulation. There is also a frequency shift of the same order when the transmission line is dis-

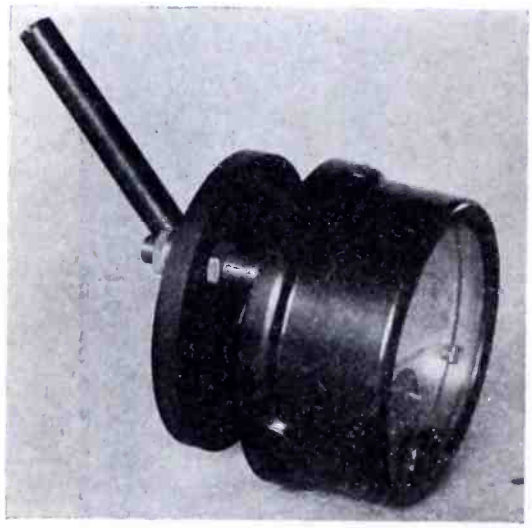


Fig. 2—Thermojunction microammeter.

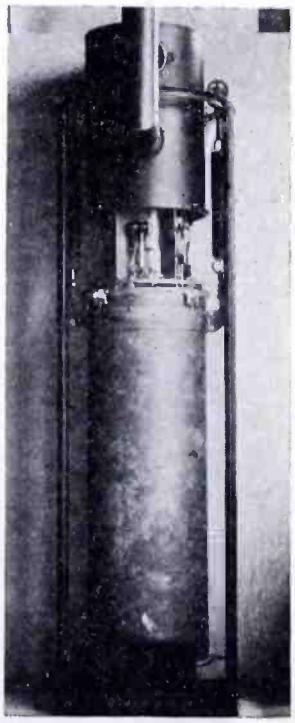


Fig. 3—Radio-frequency unit of an ultra-high-frequency transmitter.

connected or badly matched to the antenna. A set of daily measurements covering a space of two weeks indicated that all transmitters held to a constant frequency within plus or minus 0.005 per cent.

These transmitters are capable of 200 watts output with a carrier frequency of 100 megacycles and a 20-kilocycle modulation band. Normally only half power is used except in extreme conditions of low signal level such as might occur with heavy ice formation on the transmitting antenna.

Receivers

The receivers throughout the system are of the superheterodyne type and utilize loaded-line-section resonators for high-frequency selectivity and oscillator-frequency stabilization. The rear view of a typical

receiver unit is shown in Figure 4. Some features of the receiver may be of interest.

The high-frequency converter takes the signal from a balanced concentric transmission-line system, through a concentric preselector circuit and a similar concentric resonant circuit to the grid of an RCA 955 tube detector. A stable concentric resonant circuit controls the 955

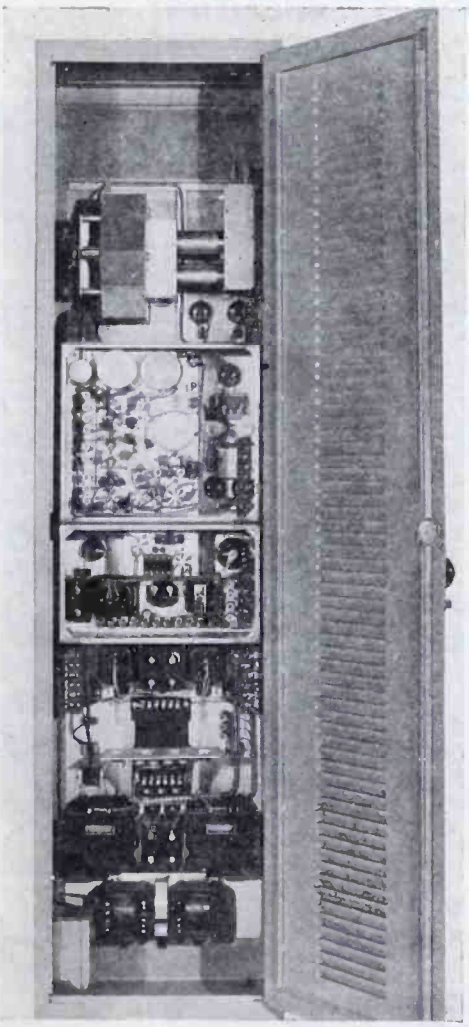


Fig. 4—Rear view of an ultra-high-frequency receiver.

heterodyne oscillator. The first detector output of 30 megacycles is amplified by two stages using RCA 954 tetrodes, then converted to a lower intermediate frequency of 5.5 megacycles which is amplified before final detection.

Unusually flat automatic gain control is employed to maintain the correct modulation level over the relay circuit. A special resonant circuit and indicator provides means of tuning a signal to the center of the intermediate-frequency amplifier band without interfering with the

receiver operation. The operating conditions of the tubes can be determined readily by a group of test meters and switches. All tubes are operated conservatively which materially increases their life.

The power supply is of rugged design and incorporates regulation

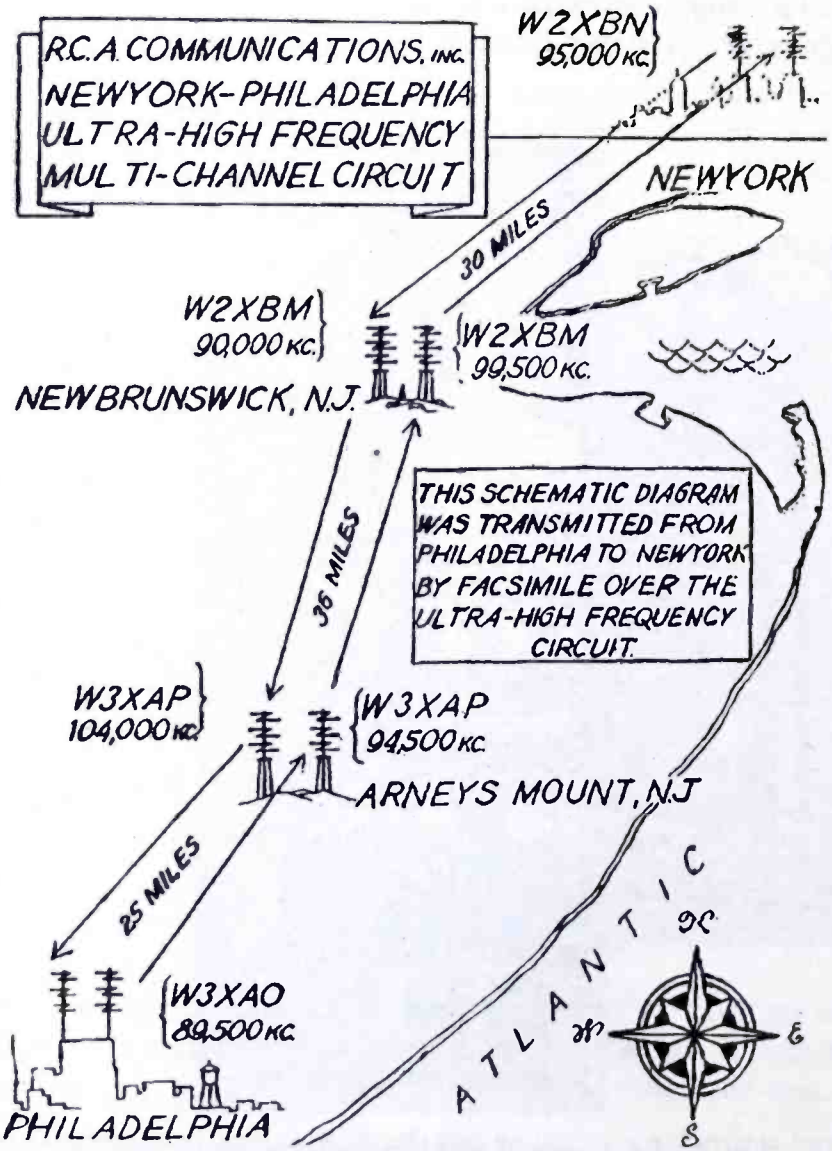


Fig. 5—Schematic diagram of the New York-Philadelphia ultra-high-frequency relay system.

of the 110-volt alternating-current supply as well as regulation of voltage for the high-frequency tubes and gain-control circuits. The receiver is thoroughly shielded and the construction is such as to permit easy servicing of all parts which are potential sources of trouble. The audio-frequency response of the receiver at present is from 100 cycles to 20 kilocycles with a maximum variation of plus or minus three tenths of

a decibel, and with harmonic distortion less than two per cent for the required output level.

Over-All Circuit

A schematic diagram of the over-all circuit is shown in Figure 5. Two intermediate relay points are used to insure dependable operation. An interior view of one of the relay stations is shown in Figure 6. The heights of the transmitting and receiving antennas are such that

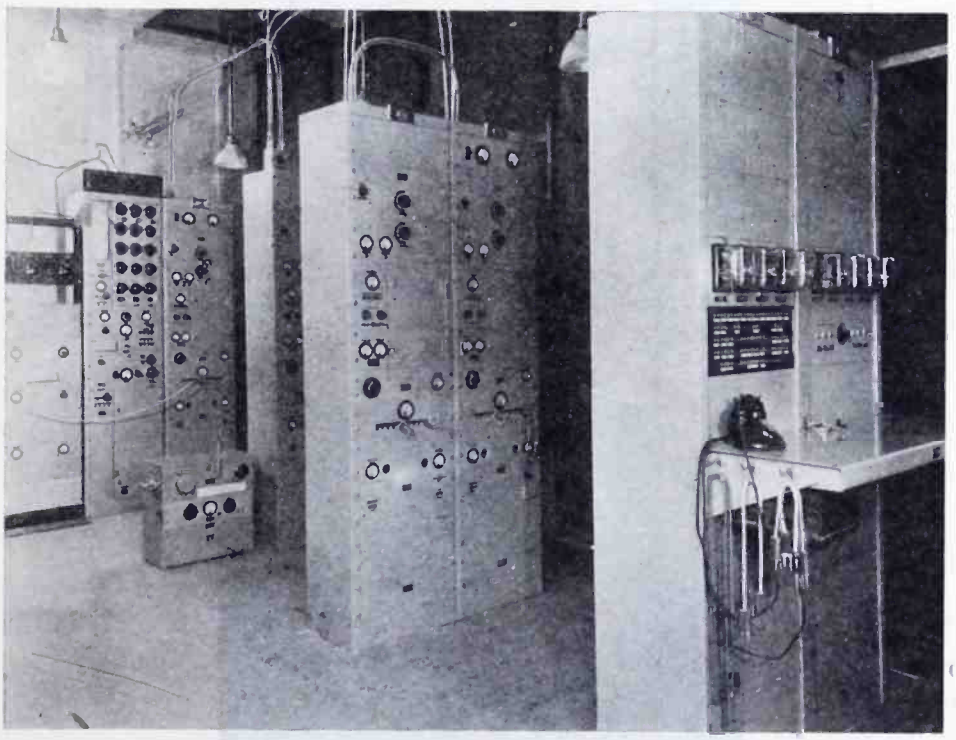


Fig. 6—Interior view of the relay station at New Brunswick.

the two end relay links are within the optical range and short enough to lay down a strong signal well above the high noise level of the urban terminals. The New Brunswick-Arney's Mount link is just below grazing incidence due to an intervening hill near Arney's Mount. The entire circuit is over land thus permitting the use of horizontal polarization.[1] All antennas are directive; the pine-tree array of horizontal dipoles as shown in Figure 7 is generally employed for both transmitting and receiving although at some points diamond-type receiving antennas are preferred.

---

[1] Bertram Trevor and P. S. Carter, "Notes on Propagation of Waves Below Ten Meters in Length," *Proc. I.R.E.*, vol. 21, pp. 387-426; March, (1933).

The carrier frequencies are staggered according to an allocation plan described in a previous article.[2] This arrangement permits a maximum utilization of the available frequency spectrum and minimizes the probability of overlapping or cross talk between the north-

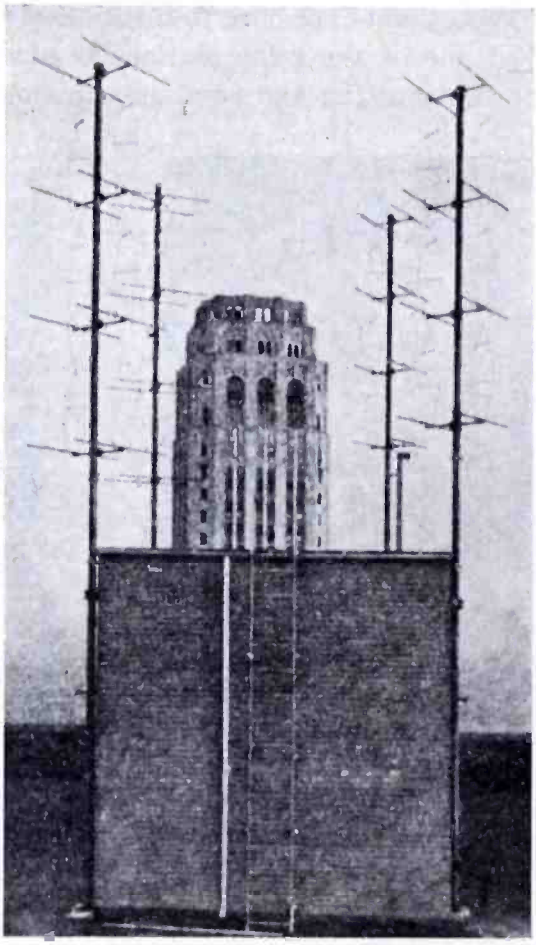


Fig. 7—The transmitting antennas at New York.

and south-bound circuits or between links of the same circuit. The Arney's Mount relay is unattended except for weekly inspections and its transmitter frequencies are monitored at New Brunswick to comply with the requirements of the Federal Communications Commission.

This radio circuit has been in operation for approximately 5,000 hours. During that time several circuit failures have occurred. The total percentage outage time of the circuit together with its cause is given in the following tabulation.

[2] H. H. Beverage, "The New York-Philadelphia Ultra-High-Frequency Facsimile Relay System," *RCA Rev.*, vol. 1, pp. 15-31; July, (1936).

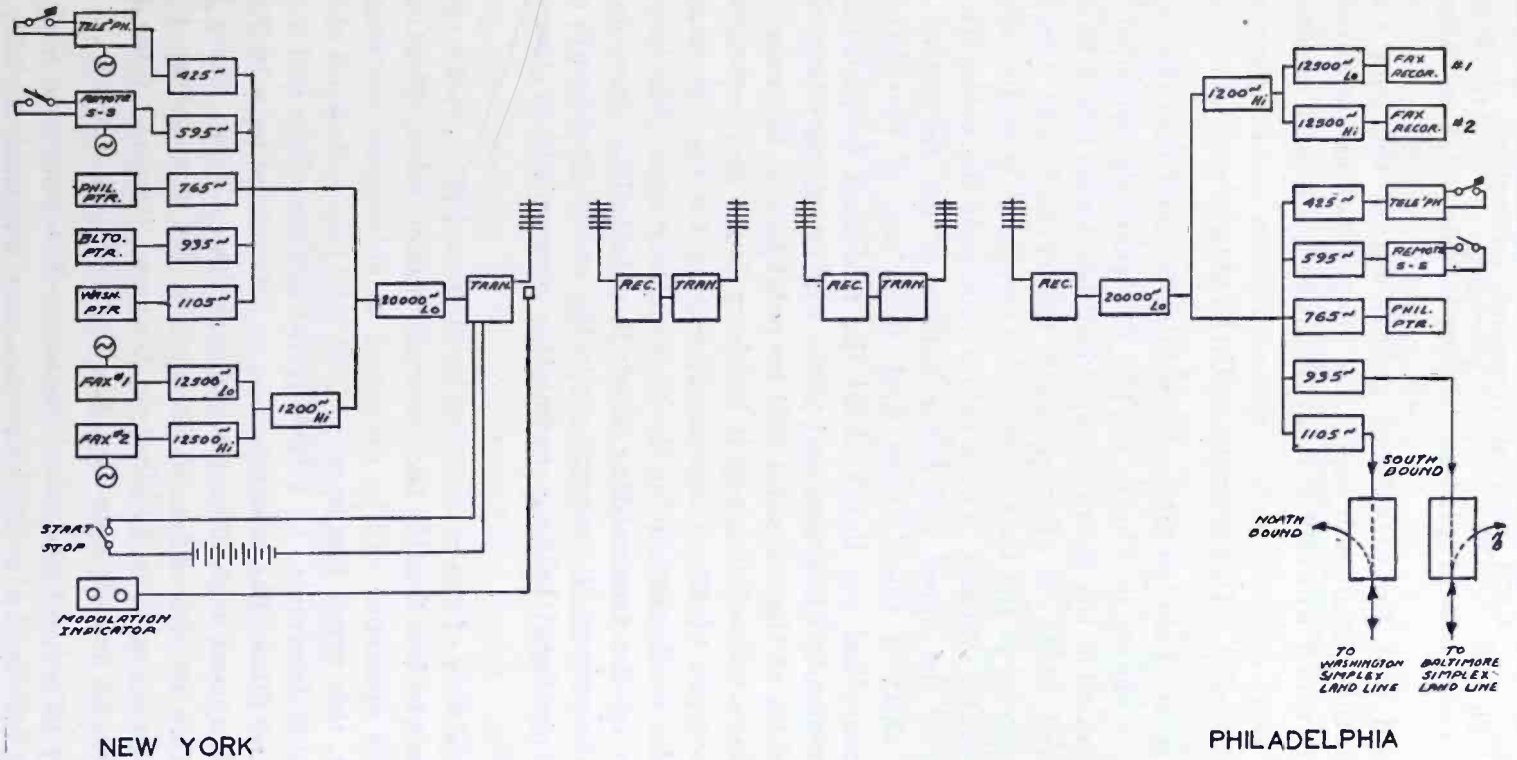


Fig. 8—Schematic diagram of terminal equipment for one direction. Duplicate equipment is used for opposite direction.

Transmitter tubes ..	0.0067 per cent	Receiver equipment.	0.015 per cent
Receiver tubes	0.074 per cent	Incorrect operating procedure	0.118 per cent
Transmitter equipment	0.018 per cent	Power failure	0.0018 per cent

It is believed that the experience gained during the past eighteen months with this circuit will enable the adoption of an improved daily test schedule which will still further decrease the loss of service time.

Description of Channels

The division of the 20-kilocycle modulation band into the channels now in use is shown in Figure 8. The five channels indicated in the upper left section of the figure are standard band-pass filters having an effective band width of approximately 100 cycles.[3] The frequencies shown correspond to the tone frequencies employed in audio-frequency carrier-telegraph systems. These values represent the geometric mean frequencies of the upper and lower cutoffs of the respective filters. The phase shift of these filters, and, therefore, of the channels is neither symmetrical nor linear about the mid-band frequencies. The resulting frequency distortion and phase distortion have been studied and the results of these studies will be published in the near future. For these narrow-band channels it has been found more economical to employ two types of filters, one configuration for the receiving type and a simpler configuration for the transmitting type. The lesser discrimination of the transmitting filter is permissible since its main function is to prevent the appearance in the receiving channels of beat frequencies produced between modulation components of the several keyed tones.

The remainder of the modulation band from 1200 cycles to 20 kilocycles is more or less flexible for facsimile transmission. Possible divisions of this spectrum will be discussed in subsequent sections. For the present, this range has been divided into two wide-band channels. One channel is determined by a 1200-cycle high-pass filter and a 12,500-cycle low-pass filter. The other channel is determined by a 12,500-cycle high-pass filter and a 20,000-cycle low-pass filter. These filters have a discrimination on the order of 60 decibels for frequencies in their attenuating ranges. The 20-kilocycle low-pass "roofing" filter is provided to insure a definite upper limit to the modulation band.

In order to prevent undesired frequency distortion, the modulation band from 300 cycles to 20 kilocycles was designed to have a flat

---

[3] The exact band width may be expressed only for idealized cases. See K. Kupfmueller, "On Relations Between Frequency Characteristics and Transients in Linear Systems," *Elec. Nach. Tech.*, January, (1928).

attenuation-versus-frequency response plus or minus one decibel for the over-all circuit. The over-all characteristic was compensated by a two-section lattice-type equalizer installed at the incoming terminals of the north- and south-bound circuits.

Channeling Considerations

In multichannel communication systems of this type, the division of the modulation band is of primary importance. Where both narrow- and wide-band channels are to be provided, it is more economical from the standpoint of filter design to place the narrow-band channels at the low-frequency edge of the available band. This is due to the fact that, for a particular filter configuration, the steepness of cutoff, i.e., the discrimination between attenuating and pass-bands, varies inversely as the cutoff frequency. Simpler and, consequently, less expensive band-pass filter configurations can be employed at low frequencies for the same discrimination. This arrangement has the further advantage of less frequency wastage between channels.

Wherever phase or frequency compensation is required for a particular channel, it is generally more economical to design compensating networks for that channel than to redesign the channel filter to have improved characteristics. These improved characteristics require special filter elements such as very high-Q inductances. High-Q elements are more susceptible to mechanical shock and to temperature variations. It is difficult to design inductances with alloy cores having modulation products below minus fifty decibels. High-quality channels of very low noise level should therefore utilize air-core inductances and a more efficient division of the band may be made with these channels at the higher frequencies where higher values of Q are automatically obtained.

Tone-Level-Setting Procedure

In a central control terminal where both land-line and radio facilities transmit the outgoing signals, it is desirable that the characteristic impedance and required tone levels be identical at the master control board. This arrangement makes for flexibility since no manual adjustments are required by the operator. Moreover, a simple procedure should be sufficient to check a particular type of terminal equipment. Also, in a multichannel setup, the design should be such that maximum energy will be obtained in an individual channel for a fixed peak value of combined tone amplitude. For the ultra-high-frequency radio circuit this peak value should correspond to 100 per cent modulation. Obvi-

ously, the combined peak value is a function of (a) the number of tones, (b) their phase relationships, and (c) the duration of mark on each channel.

The terminal equipment for the ultra-high-frequency control is shown in Figure 9. All of the keyer units, whether they be for telegraph, Teletype printer, or facsimile transmission, are adjusted to the same output level. This level is standard throughout the central office so that any service may be transmitted over any facility. In order that this tone level will have its maximum permissible amplitude at the line (or transmitter) input, the booster-amplifier—master-pad combination is adjusted for the number of channels provided by the filter

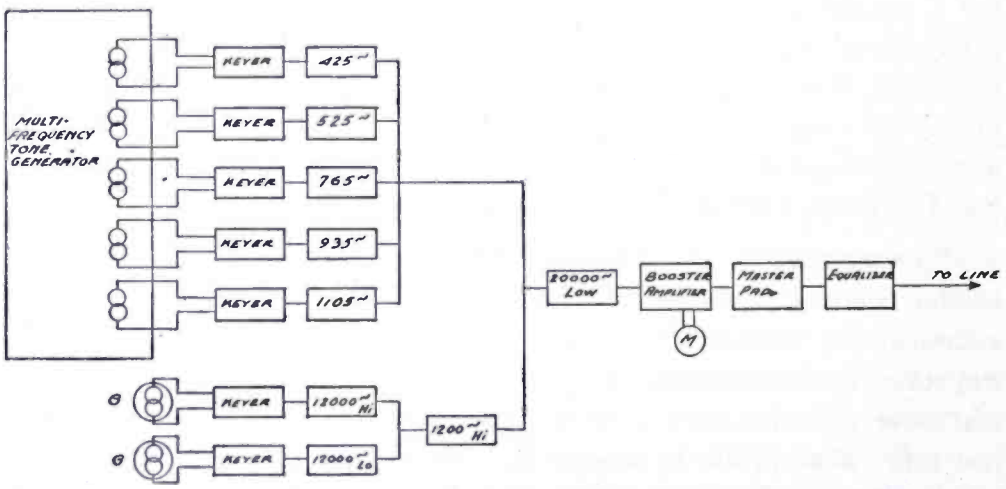


Fig. 9—Schematic diagram to illustrate tone-level procedure.

group. When less than the total number of channels are used the master pad may be calibrated against the number of channels. For routine checking of the terminal equipment the control operator applies a single steady tone from one keyer unit. The booster-amplifier output as indicated on a rectox-type meter should then read the value specified for that terminal. By inserting each keyer unit in turn he can check the complete terminal in a few seconds' time.

The calibration of the master pad depends upon the maximum peak amplitude of the combined tone output which in turn depends upon the three conditions stated above. Neglecting condition (c) for the moment, if n tones of equal amplitudes and random-phase relationships were employed, i.e., n nonsynchronized tone sources, they would all add in phase at some instant thus producing an instantaneous peak amplitude of n times a single-tone value. The master pad should then have a loss corresponding to a voltage reduction of $1/n$.

In practice, the in-phase peaks of short duration may be threshold limited to a certain value depending upon the amount of distortion permissible in the various channels.

If, however, the phase relationships of the various tones are fixed as would be obtained for synchronized oscillators or multitone generators the conditions are quite different. In order to present this analysis clearly, a practical example is worked out in detail.

Suppose a multitone generator produces twelve frequencies 450, 630, 810, etc., when rotating at 1800 revolutions per minute or 30 revolutions per second. We observe immediately that the armature windings are in the ratio 15, 21, 27, etc., so that we need to calculate

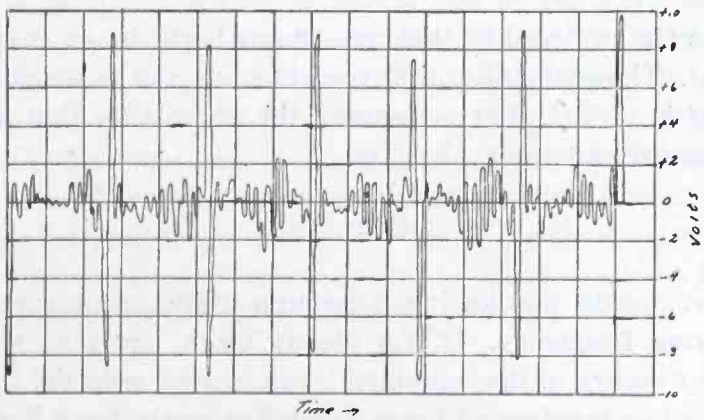


Fig. 10—Resultant of 12 combined tones of a multitone generator.

only the vector sum of the 12 tones for one revolution of the armature. Figure 10 shows the value of the twelve combined tones during one revolution on the basis that the tones are arranged to produce zero amplitude when all are in phase. The ordinates are in multiples of a single-tone amplitude. We see that the master pad would need to reduce the combined tones 1/9.5 or 19.6 decibels, whereas for random tone sources the reduction would be 1/12 or 21.6 decibels. This means the multitone generator could furnish a 2-decibel greater level in each channel than would be permissible with random-phase oscillators.

The above analysis was made on the basis that continuous tones were supplied to all channels. Actually, the channels will be keyed on and off according to the type of communication transmitted. It is apparent that the number of combined peak maxima occurring during an interval of time will decrease as the continuous tones are replaced with keyed tones.

To obtain an idea of the effect of keying on the peak-signal amplitude, we shall continue the analysis of the 12 tones considered above.

Let us assume these 12 tones are to be used for 12 telegraph channels and that they are all keyed at 50 words per minute or its equivalent, 20 cycles per second. An analysis of a typical message indicates the signal "on" time to be very nearly equal to the signal "off" time for the international Morse code. The probability that channel No. 1 will be "marking" at a particular instant is then equal to ½. The probability that this channel will be marking over a finite interval $\triangle t$ is somewhat less than ½, but for our purpose $\triangle t$ will always be of small magnitude compared to the marking interval.

We may then write

$$P \Rightarrow \tfrac{1}{2}$$

where P is the probability that one channel will be on mark over the interval $\triangle t$. The probability that two channels will be marking over the interval $\triangle t$ is $(1/2)^2$. For n channels the probabiilty that they will all be marking over an interval $\triangle t$ is

$$P_n \Rightarrow (\tfrac{1}{2})^n$$

or, in other words, one keying maximum would occur every 2^n cycles of the keying frequency. If, for steady tones, one tone maximum of duration $\triangle t$ occurs in the combined tones every t seconds, the probability that the two maxima will occur simultaneously for n keyed tones is

$$P_k \Rightarrow \frac{\triangle t}{t2^n}$$

or one maximum occurs for every $2^n t/\triangle t$ cycles of the keying frequency.

We may then write

$$T = \frac{2^n t t_1 \text{ hours}}{3600 \triangle t}$$

where,

T = time required for 1 maximum
n = number of telegraph channels
t = time for 1 maximum with n steady tones
t_1 = time of 1 keying cycle
$\triangle t$ = duration of peak maximum.

For our example above, all channels are keyed at 20 cycles per second and the keying relationships between channels are random. We have $n = 12$ and $t_1 = 0.05$ second. From Figure 10, $t \doteq 12.5 \triangle t$.

Substituting

$$T = \frac{2^{12}(12.5)\,(0.05)}{3600} = 1.06 \text{ hours}$$

we observe that for the example given when all 12 tones are keyed, a peak maximum occurs on the average of once every hour. It is clear that partial elimination of these occasional peaks by limiting would produce very little distortion in the received channels, and, at the same time, would allow a greater energy level to be transmitted in the individual channels. The lowest limiting threshold value that can be permitted depends upon the rate of recurrence of the peak maxima, the duration of the peak impulse, the frequency spacing of the channels, and the distortion tolerance of the receiving equipment. It is therefore best determined by experiment. Moreover, in practice, long marking intervals will occur on some channels so that the actual peak amplitude will occur more frequently than is indicated by the idealized example above. It is clear, however, that higher audio-frequency signal levels or a greater number of useful channels can be made available for a given transmitter power by use of the limiting principle.

Signal-to-Noise Measurements

An important figure of merit of a transmission channel is its signal-to-noise ratio. A particular type of recording device has a minimum signal-to-noise value below which failures occur. In designing a channel for a certain facility it is necessary to compromise between the band width desired for fidelity of the received signal and that required to maintain a usable signal-to-noise ratio. Recent studies indicate that impulse noise varies directly with the band width and fluctuation or smooth noise varies directly with the square root of the band width. Preliminary tests showed the audible noise over the ultra-high-frequency circuit to be generally of the fluctuation type except during occasional intervals of diathermy or motor-ignition interference, at which time the impulse-type of noise predominated.

In order to obtain quantitative noise data over the radio circuit a set of measurements was made. Tones were applied at Philadelphia in various combinations and the outputs of the different channels at New York were compared for signal levels, cross talk, and noise values. Some of these results are shown in Tables I and II.

TABLE I

Signal-to-Noise Measurement

Channel (cycles, narrow band)	Channel Output at New York in Decibels	
	(a) No Signal at Philadelphia	(b) Single 595-cycle tone (100 per cent modulation)
425	—59	—43
595	—58	0
765	—60	—42
935	—56	—58
1105	—58	—54
1275	—59	—52
1445	—59	—58
(kilocycles, wide band)		
3-17	—35	—35

TABLE II

Observation for Effect of Cross Talk into 595-Cycle Channel from Other Channels

Inserted tones	Output of 595-cycle channel Decibels
425-cycle tone, 14 decibels below 100% modulation, steady mark	—42
595-cycle tone, 14 decibels below 100% modulation, steady mark	0
935-cycle tone, 14 decibels below 100% modulation, steady mark	—42
8-kilocycle tone, 14 decibels below 100% modulation, steady mark	—43
425-cycle + 935-cycle tone, 14 decibels below 100% modulation, steady mark	—42
425-cycle + 8-kilocycle tone, 14 decibels below 100% modulation, steady mark	—42
935-cycle + 8-kilocycle tone, 14 decibels below 100% modulation, steady mark	—43
425-cycle tone keyed by hand	—41
935-cycle tone keyed with RY's on printer	—42
8-kilocycle tone keyed by facsimile	—42
425-cycle + 935-cycle tone keyed by hand and printer	—40
425-cycle + 8-kilocycle tone keyed by hand and facsimile	—41
935-cycle + 8-kilocycle tone keyed by printer and facsimile	—42

It was desired to use one channel of the system as a daily indicator of the noise condition on the circuit. Table I, the second column gives the noise levels in the narrow- and wide-band channels when no signal is applied at Philadelphia. Table I, the third column shows the corresponding measurements with a 595-cycle tone modulating the Philadelphia transmitter 100 per cent. It is seen that the discrimination of the adjacent-channel filters is not sufficient to reduce the 595-cycle signal level below their normal noise levels. This cross-talk value would be still further increased if the 595-cycle tone were keyed. It is therefore apparent that none of the narrow-band channels can serve as a criterion of the circuit noise. The wide-band channel on the other hand has an inherent noise level greater than the cross talk from the 595-cycle tone. Subsequent measurements indicated the wide-band channel to be satisfactory for this purpose.

Table II shows the cross talk obtained in the 595-cycle narrow-band channel for various keyed and steady tones. For these filters it is seen that keying produces about one decibel greater cross talk in the adja-

cent channels than steady tones. The other channels are generally unaffected due to their greater "frequency distance."

As stated above, the general noise condition of the ultra-high-frequency circuit is believed to be determined by smooth-type noise. To check this we may use the relation that the noise varies as the square root of the band width. In Table I the wide-band channel is 14 kilocycles and the narrow-band channels are approximately 100 cycles. We should accordingly expect a noise increase in the wide-band channel of 11.83:1 (21.5 decibels) which agrees closely with the measured values.

During the occasional intervals when impulse noise predominates, which occurs when diathermy or motor ignition is picked up at any of the transmitters, the wide-band channel may be rendered useless for facsimile communication. The interference appears as wavy diagonal lines in the received facsimile copy. This type of interference does not disturb the normal operation of the narrow-band channels.

Normal atmospherics are of little importance at these frequencies. The automatic volume control incorporated in the receivers compensates for all signal variations due to weather conditions except for exceedingly heavy ice formation on the antennas.

The effects of fading may be considered negligible. During a year's operation of the circuit, three short fading periods occurred, each of about two minutes duration. They are known to be in the New Brunswick-Arney's Mount radio link, and, apparently, have no correlation with the fading periods of normal short-wave frequencies. One theory explains this phenomenon as being due to cancellation of the refraction field by the diffraction field. However, it is not felt that sufficient data are yet available to furnish a complete explanation.

On the whole, the ultra-high-frequency radio circuit has been found quite satisfactory for communication purposes and it is expected that wider applications will demand an increasing utilization of this frequency range.

U-H-F EQUIPMENT FOR RELAY BROADCASTING

By

W. A. R. Brown

Assistant Development Engineer, National Broadcasting Company, Inc.

Summary—Several types of portable transmitting and receiving equipment, which have recently been developed to meet the increasing demands imposed upon operating facilities by the constantly expanding scope of ultra-high-frequency relay broadcasting are described, and their functions in relay-broadcast operations outlined in this paper.

RELAY broadcasting is a recent designation for a form of program transmission of some years' standing which has steadily increased until it is now an integral part of radio broadcasting. This interesting and colorful service furnished to the listener is another step in the expansion of broadcasting brought about by the desire of the broadcaster to better serve the public.

A number of intermediate and ultra-high frequencies have been allocated to this service. The ultra-high-frequency allocations in the 30-42 Mc band are particularly suitable for many phases of relay-broadcast operation because they permit the use of radiating systems of comparatively high efficiency, but of small physical dimensions. Readily portable equipment can be used because the efficiencies of many standard tubes are still high at these frequencies. Propagation characteristics are also suitable as attenuation is not excessive and wave interference is not too pronounced.

In relay-broadcast operations the radio link permits the presentation of programs from points where wire facilities are not available or their installation is impractical. These program-origination points may be almost anywhere: aeroplanes, ships, automobiles, golf courses, airports.

Before this service reached its present stage of efficiency a long period of experimental development was necessary. Work in this field up to 1936, and particularly the work at 300 Mc, has been described previously.[1] This paper will describe briefly the u-h-f relay-broadcast

[1] "Micro-waves in NBC Remote Pick-ups," R. M. Morris, *RCA Review*, July, 1936.

Reprinted from *RCA Review*, October, 1938.

equipment which NBC has since developed and is now standard in all its divisions throughout the country.

Equipment for this service must necessarily be portable, rugged, efficient, easily operated, and reliable in operation. It must also operate under a wide variety of conditions.

The present u-h-f equipment consists of five major units: the 25-watt general-utility transmitter, the 2-watt pack transmitter, the 0.2-watt miniature transmitter, the program receiver, and the cue receiver.

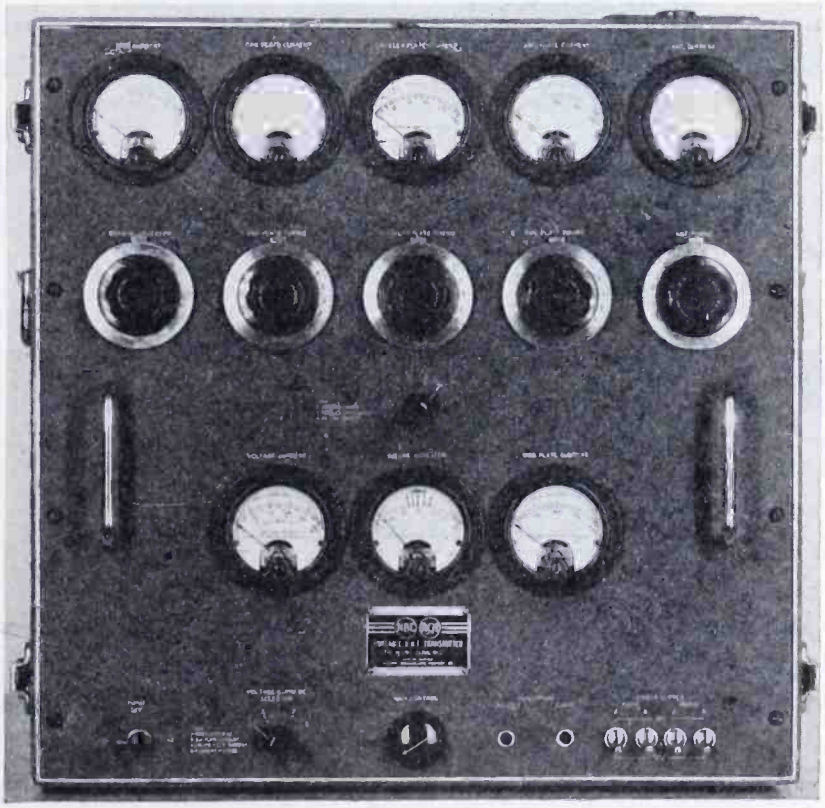


Fig. 1—Utility transmitter.

The utility transmitter, with 25 watts output, is the highest powered portable transmitter of our u-h-f relay-broadcast line. Its power renders it useful for a wide variety of applications, including (a) mobile transmission from automobiles, trains, aeroplanes; (b) as a base transmitter at wire-line terminals for cue transmission and communication to program-origination points in the field; (c) as an automatic relay-transmission point when low-power program-origination transmitters cannot reach the wire-terminal location directly, etc.

The 25 watts power output represents a compromise between tube efficiencies and the necessity for keeping the equipment, including the

associated a-c power supply, sufficiently small in size and low in weight to be readily portable. The transmitter, without the power supply, is shown in Figure 1. Its size is 19″ × 19¼″ × 10½″ and its weight is 73 pounds. For operation from 110-volt d-c sources a small auxiliary motor-generator is employed with the power supply. Storage battery operation of the transmitter is accomplished by the use of a dynamotor. This power-supply unit may also be used to provide power for the intermediate-frequency relay-broadcast transmitter which is identical in size and output.

Conventional circuits are utilized in the transmitter. The schematic diagram is shown in Figure 2. Harmonic V-cut crystals with low-temperature coefficients are employed to control the transmitter fre-

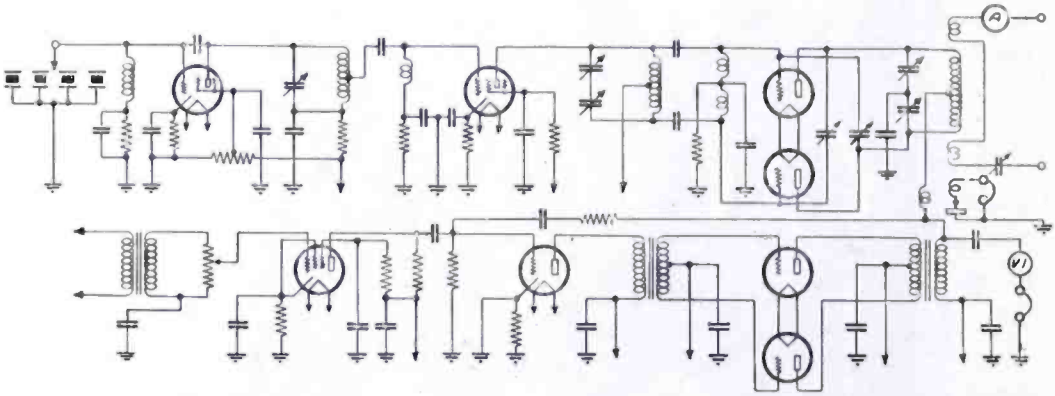


Fig. 2—Schematic of utility transmitter.

quency and there is provision for switching to any of four crystals. Plate modulation is employed and 100 per cent modulation may be obtained with less than 4 per cent harmonic distortion (arithmetical sum).

The r-f section is contained in the upper half of the assembly and consists of three stages: a crystal oscillator, a doubler, and a push-pull power amplifier. The audio-frequency section is contained in the lower half of the assembly and also consists of three stages: a speech amplifier, a driver, and a push-pull Class AB modulator. About 8 db of negative feedback is utilized to reduce distortion. The frequency characteristic is peaked at 8,000 cycles, to improve the signal-to-noise ratio at the higher audio frequencies, and is down 3 db at 50 cycles.

The antenna-coupling system permits the use of either a grounded quarter-wave radiator or a remotely located radiator fed by a two-wire or concentric transmission line.

Both aural and visual monitoring are obtained from the secondary of the modulation transformer, the former with headphones and the

latter with a level indicator. Radio headphone monitoring is obtained from an enclosed crystal detector and pick-up coil. Comprehensive metering facilities are provided, for plate currents of all tubes, grid currents of all r-f tubes, modulator-bias voltage, and filament voltage. Some of these are indicated directly and others by a universal meter and selector switch. Careful consideration was given in the

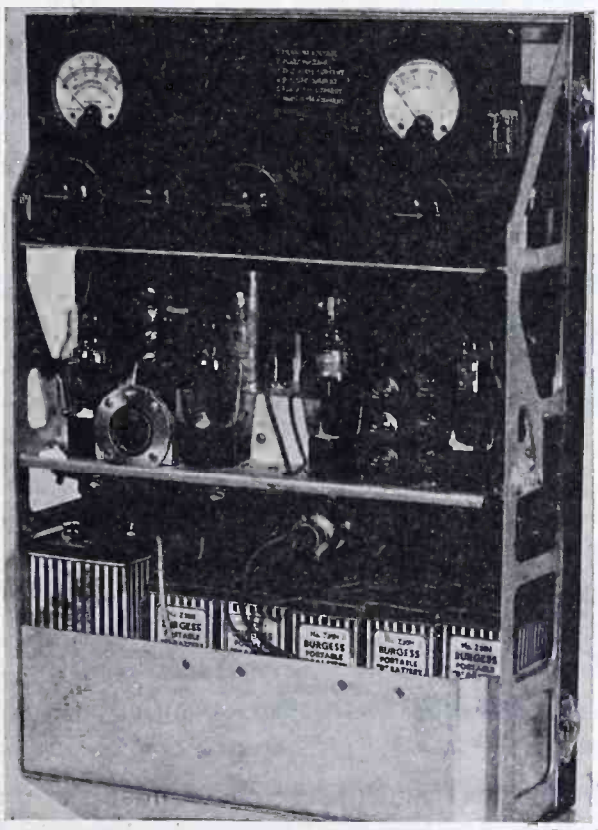


Fig. 3—Pack transmitter.

design to accessibility of parts to facilitate repairs if necessary while in the field.

The 2-watt pack transmitter, shown in Figure 3, is a highly portable, medium-powered, self-contained unit designed primarily to provide relay-broadcast transmission facilities while carried as a pack. It is normally carried in a leather case securely strapped to the announcer's back and the only piece of external equipment is the microphone, which is carried by hand. Thus the announcer is free to move around wherever necessary. With an output power of 2 watts it is used chiefly for transmission over distances of a few hundred yards to a mile, as for example, from locations on a golf course to the line terminal in the club house. Under more favorable conditions, the

range is considerably greater. For instance, when receiving on top of the RCA Building in New York City, at an elevation of 850 feet, this transmitter is frequently employed for program transmission from vessels in the harbor, some four or five miles distant. Under highly favorable transmission conditions, such as from aeroplanes, program transmission ranges of fifteen miles or more are obtained. It is also employed sometimes as a cue transmitter when the relay-broadcast circuits are short.

The design of this pack transmitter is based upon several years experience in this phase of broadcasting. Although the basic circuit is conventional, several features have been incorporated which con-

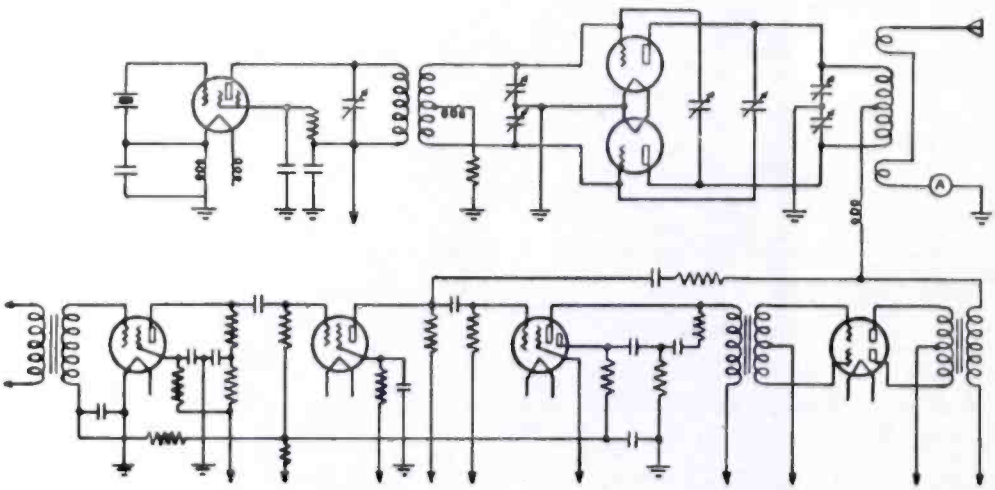


Fig. 4—Schematic of pack transmitter.

tribute to high efficiency and utility and render the unit rather outstanding in its field. A schematic diagram is shown in Figure 4. The power output of 2 watts represents, of course, the inevitable compromise between size, weight, tube capabilities, power supply, etc. The total weight, including the self-contained power supply, is 33 pounds and the dimensions of the transmitter are 17½" high, 12⅝" wide, and 4½" deep. Figure 3 shows the three sections into which the unit is divided mechanically. The upper section contains the r-f equipment, the middle section the a-f equipment, and the lower section the power supply. Any or all of these sections may be readily removed from the chassis. This greatly facilitates maintenance.

For obvious reasons the transmitter is crystal controlled. Plate modulation is employed and 100 per cent modulation is obtained with low harmonic distortion using negative feed-back. One of the distinctive features of this transmitter is automatic audio-gain control. This is a very valuable operating feature as it reduces the extreme

vocal dynamic ranges which frequently occur in this type of broadcasting and assures high modulation levels without over or under modulation for all except extremely abnormal voice levels.

The r-f section consists of two stages: a crystal-controlled electron-coupled oscillator-doubler and a push-pull power amplifier. To obtain 2 watts output it is necessary to operate all tubes somewhat above normal filament potentials. Tube life is shortened under these conditions and tubes are discarded after 50 hours operation to provide a wide margin of safety.

The a-f section contains four stages: a two-stage speech amplifier, a driver, and a modulator. These possess sufficient gain, after utilizing degeneration, to permit full modulation with an RCA 50-A inductor microphone. Correct polarization of the audio circuits to obtain maximum undistorted power is another operating feature which contributes to the efficiency of this transmitter. By taking advantage of the phenomenon that the peak amplitudes of most voice waves are highest in one direction, it is possible to secure greater than 100 per cent modulation upward while not exceeding 100 per cent downward. Automatic audio-gain control is obtained by feeding a portion of the rectified-driver output to the input of the speech amplifier through a resistor-condenser combination of suitable delay characteristics. The frequency characteristic is purposely peaked about 3 db from 5,000 to 10,000 cycles and is down 3 db at 60 cycles. This departure from a flat characteristic is an attempt to compensate for difficulties encountered in operation. Our experience has indicated that general crowd noise and wind noise is often most pronounced at frequencies below 80 cycles while the hiss type interference experienced at the receiving point from ignition and diathermy frequently appears most pronounced at the higher audio frequencies.

The antenna construction and installation is particularly convenient for operation. It is a continuously-adjustable four-section semi-flexible telescopic assembly which is an integral part of the transmitter and telescopes into the case when not in use. The system is designed to operate as close to resonance as is necessary to obtain correct loading of the power amplifier.

Two meters are contained in the transmitter assembly. One indicates antenna current and the other, by selective switching, indicates filament and plate voltages, oscillator, power-amplifier, and modulator-plate currents, and power-amplifier grid current.

Standard dry batteries are employed for the power supply which has sufficient capacity for nine hours continuous operation. To obtain this length of service a reserve 45-volt battery is included in the power

supply and is switched into service after about six hours transmission. This is an operating convenience which permits considerable test transmission to be conducted prior to a broadcast and still allows a comfortable margin of available power for a lengthy period of program operation. For continuous operation at a fixed location, as for cue-channel transmission, provision has been made for plugging in an external heavy-duty battery supply.

The miniature transmitter, shown in Figure 5, is the smallest and most recent of our u-h-f transmitters. It is a small, entirely self-contained, extremely portable, low-powered transmitter for short-

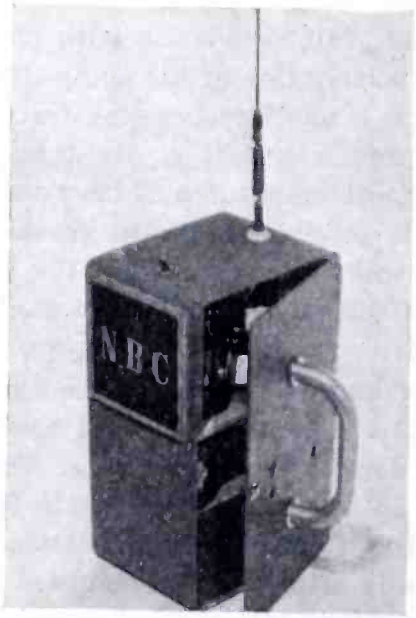


Fig. 5—Miniature transmitter.

range transmission. Although its small size is productive of considerable novelty appeal the unit was actually developed to fill a definite operating need and has already introduced a new technique in program presentation. It is particularly useful for program origination in crowded areas, or on locations where the size of the equipment must be kept to an absolute minimum or where transmission distances are short. One of its most distinguishing features, and one which at times is a valuable operating convenience, is that it can be readily passed from hand to hand. The working range is dependent upon propagation conditions and the signal-to-noise ratio at the receiver and may vary from a few hundred feet to half a mile.

The transmitter, with built-in microphone and power supply, is housed in an aluminum case 10″ × 4½″ × 5¾″ and weighs 7½ pounds. Despite its small size an output of 0.2 watt is obtained. Crystal control

assures frequency stability and automatic audio-gain control maintains proper modulation levels.

The circuits employed are quite conventional and the efficiency of the unit is the result of good detail design. A schematic is shown in Figure 6. The r-f section consists of a crystal oscillator-doubler stage and a power amplifier while the audio-frequency section consists of a single-stage speech amplifier and a modulator. Heising modulation is employed and modulation values of approximately 90 per cent are obtained by the expedient of dropping the power-amplifier plate voltage.

Automatic audio-gain control is particularly effective in the maintenance of proper modulation levels under the conditions frequently encountered in operation of this transmitter. With 12-db peak gain

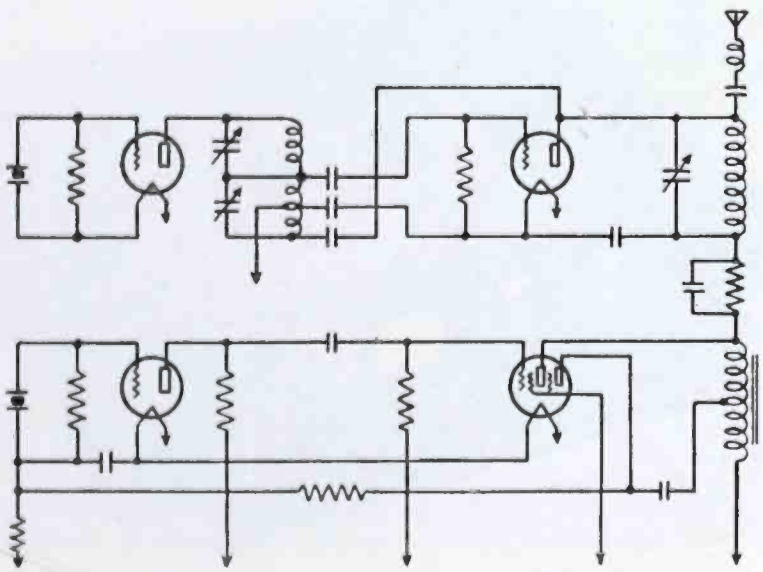


Fig. 6—Schematic of miniature transmitter.

reduction it effectively eliminates overloading of the transmitter except on very pronounced peaks.

A crystal microphone is employed in the interest of weight reduction and is concealed behind a metal grille on the front of the unit. The overall audio-frequency characteristic of the transmitter drops off rapidly above 5500 cycles and is therefore not as good as the larger transmitters. It is quite adequate, however, for voice transmission which is the only type normally employed with this unit.

The antenna is fed through an insulator in the top of the housing. It consists of a short dural rod with a spring coil at the lower end which provides the necessary mechanical flexibility for the antenna and also some electrical loading.

The power supply is a special battery block in which is combined the filament, plate, and bias supplies. This is sufficient for nine hours

continuous operation and is so designed that the *A* and *B* capacities are exhausted practically simultaneously.

Jacks and an external meter permit determination of oscillator, power-amplifier, and modulator-plate currents. Tuning is also accomplished externally by the use of a wand which is inserted in small openings in the housing to vary the tuning condensers.

The program receiver shown in Figure 7 is a portable, high-quality, superheterodyne receiver designed primarily for reception of crystal-controlled relay-broadcast transmitters. This type of receiver possesses a number of advantages over the super-regenerative receivers previously used for this service, one of the most important being the

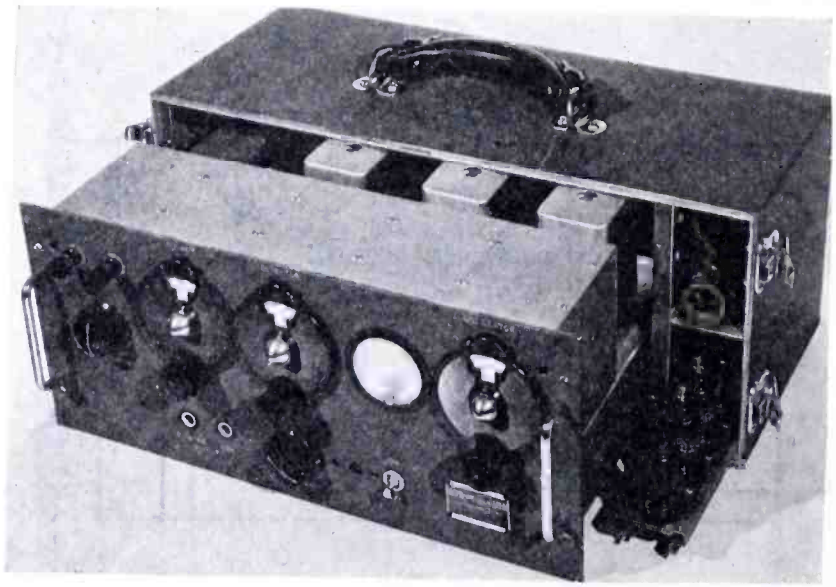


Fig. 7—Program receiver.

absence of re-radiation. The receiver dimensions are 19¼" × 9¾" × 8½", and its weight, exclusive of the external battery-power supply, is 26 pounds.

Conventional superheterodyne circuits, with avc, are employed, as follows: one stage r-f amplifier, first detector, oscillator, two-stage i-f amplifier, and second detector-audio amplifier. Figure 8 shows the schematic diagram. The oscillator operates 5.9 Mc above the signal frequency. This choice of intermediate frequency was a matter of design convenience to obtain the desired band-width characteristic.

The design and operating characteristics of the receiver were based largely upon the conditions under which the unit would operate in the field. At that time there were still some self-excited transmitters in use and provision was therefore made for variation of the i-f amplifier band width, up to a maximum of 100 kc, by the inclusion of variable

coupling in each i-f transformer. With a 60-kc band pass the selectivity is sufficient to reduce an undesired signal 75 kc off resonance by approximately 44 db. The selectivity can be improved considerably when necessary. Image response under above conditions is minus 72 db.

High gain is not essential in this receiver since in normal broadcast operation it would feed into a field-program amplifier with 100 db gain. There is sufficient amplification, however, to give a plus 4-db level at the receiver output with 100 microvolts signal input. With this gain the audio characteristic is practically flat from 30 to 10,000 cycles. It may be of interest to note that a voltage gain of 15 is obtained in the r-f amplifier stage with an RCA acorn tube.

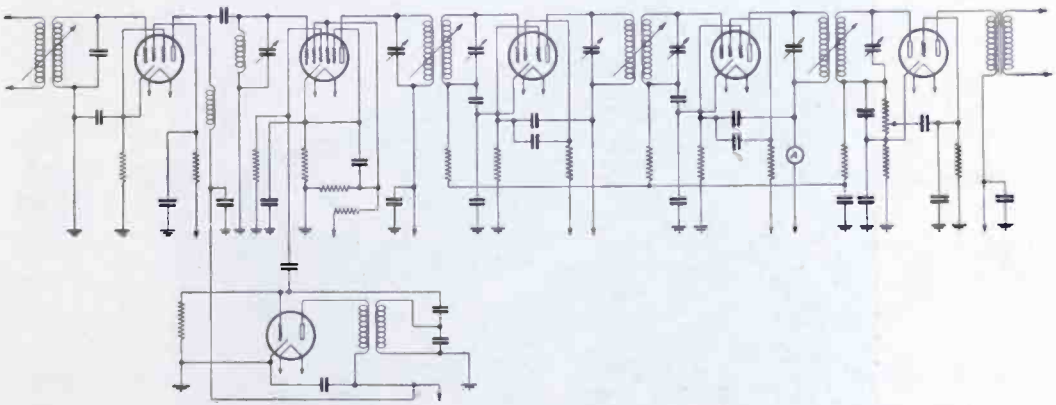


Fig. 8—Schematic of program receiver.

A half-wave antenna is usually employed for reception and the input circuit is designed to operate from a low-impedance two-wire transmission line. Although the minimum useful signal input is about 10 microvolts, in the absence of external noise, the utility of the receiver is dependent upon the signal-to-noise ratio and for satisfactory program use a ratio of about 25 db or more is desirable. Receiving conditions on the roof of the RCA Building in New York City may be cited as a typical example. Here, when the noise input averages about 30 microvolts, signal inputs of about 600 microvolts are considered satisfactory for program use. This corresponds to a signal-to-noise ratio of 26 db.

There are three tuning controls: r-f amplifier, detector, and oscillator. Separate controls eliminate any possible difficulties of gang tuning and simplified design. No operating difficulties are introduced since relay-broadcast frequencies are limited in number and the tuning dials are marked for these frequencies. However, a complete oscillator-dial calibration is provided to assist in tuning other frequencies, if necessary.

A very valuable operating feature is a tuning meter which indicates the i-f amplifier plate current and is calibrated against signal input. This permits easy determination of relative signal strengths at various locations, good and poor transmission areas and also signal-to-noise ratios.

Headphone monitoring facilities are provided and a fixed 20-db pad is incorporated in the receiver for matching impedances when feeding a field amplifier.

Fig. 9—Cue receiver.

The cue receiver shown in Figure 9 is a small, light-weight, entirely self-contained superheterodyne receiver designed primarily for the reception of cues and instructions at relay-broadcast program-origination points. Under these conditions extreme portability and a high degree of sensitivity are frequently essential, but the quality need not be more than sufficient for good speech intelligibility. When used in conjunction with the pack transmitter the engineer who accompanies the announcer carries the receiver in a leather case slung across the chest. All controls are within easy reach and it can be carried for hours without discomfort. The total weight of the complete unit and carrying case is only 13 pounds and its size is but $9\frac{3}{4}'' \times 4\frac{3}{4}'' \times 10\frac{1}{4}''$. The power supply is the same battery block that is used in the minia-

ture transmitter and will provide eight hours continuous service. Battery replacement is convenient and rapid.

Conventional superheterodyne circuits, with avc, but no pre-selection, are employed in the following circuit arrangement: first detector, oscillator, two-stage i-f amplifier, second detector-audio amplifier. The schematic is shown in Figure 10. The elimination of an r-f amplifier stage was dictated by the necessity for a minimum of weight, although it introduced certain disadvantages.

One interesting design feature is the use of a triode tube in parallel with the triode section of a pentagrid converter to increase

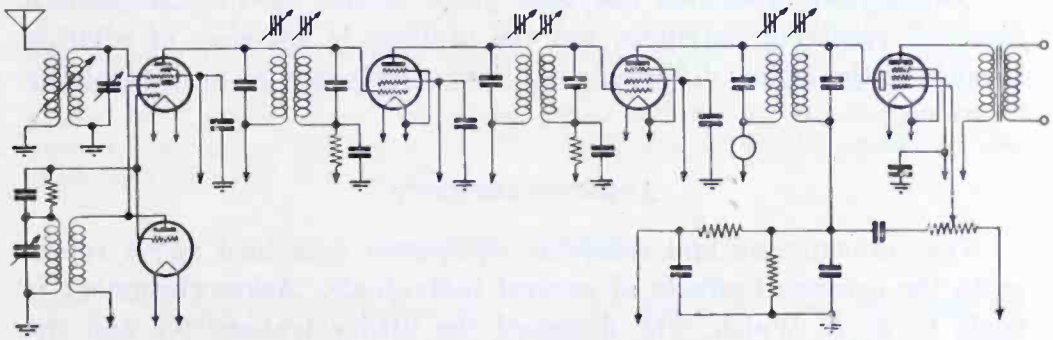


Fig. 10—Schematic of cue receiver.

the mutual conductance of the latter and thereby improve its oscillation performance at the ultra-high frequencies. The oscillator operates 4.1 Mc below the signal frequency. Transformers with adjustable iron cores to facilitate circuit alignment are employed in the i-f amplifier. Although the selectivity is limited somewhat by lack of pre-selection it is sufficient for relay-broadcast operation. A signal 75 kc off resonance is down 23 db. Image response is approximately minus 35 db. The antenna is of the telescoping, flexible, whip type, similar to that used in the pack transmitter. It is an integral part of the unit and telescopes into the receiver when not in use.

Many of the operating features are similar to those of the program receiver and for the same reasons. Two tuning dials are employed: first detector and oscillator. Dial settings for the usual relay-broadcast frequencies and oscillator-dial calibration for spotting other frequencies render operation simple. A tuning meter for indicating relative signal intensities has also been incorporated in this receiver and operates in the same manner as in the program receiver.

The output transformer is specifically designed to work into high-impedance headphones and therefore this receiver is not suitable for feeding an amplifier or line. A manually operated gain control is also provided to keep the phone levels below values where acoustical feed-back to the associated pack transmitter may occur.

At these operating frequencies there are several advantages in favor of horizontally polarized transmission, such as somewhat lower noise levels, and higher average signal strengths. However, in many phases of relay-broadcast operation these advantages cannot be realized because physical limitations permit only vertical transmission. This is particularly true with the pack and miniature transmitters.

With the low power of these transmitters the operating personnel must therefore exercise a high degree of engineering ability to cope successfully with the problems which the varied conditions of operation produce.

Considerable attention has been given to the use of standardized directive receiving antennas, but the problem is not easy of solution because of the widely different conditions which exist at most receiving locations.

Acknowledgment

The transmitting and receiving equipment described above represents the combined efforts of several individuals. Acknowledgment is made to A. A. Walsh, who designed the utility transmitter and the pack transmitter; to J. L. Hathaway, who designed the miniature transmitter; to W. C. Resides, who designed the program receiver and the cue receiver; and to R. M. Morris, Development Engineer, under whose direction the development was conducted.

WIDE-BAND VARIABLE-FREQUENCY TESTING TRANSMITTERS

By

G. L. Usselman

Engineering Department, R.C.A. Communications, Inc.

Summary—*A wide-band variable-frequency transmitter, designed primarily for testing purposes is described. The transmitter consists of a single-tube oscillator stage with a high degree of frequency control. Two models have been built which operate in the frequency range of 40 to 100 megacycles, with 500 to 1000 watts output. Means are also provided for telegraphically keying the transmitters. This type of transmitter supplies a need for a source of high-frequency power which is continuously variable over a wide range of frequency, but which will accurately maintain any frequency to which it is adjusted.*

Introduction

DURING the early part of the year 1937 the need was felt for more information on the propagation characteristics and multi-path phenomena of radiated electric waves over wide-frequency bands in order that we might attack more intelligently the problem of television transmission. As a part of the program to supply this information the Transmitter Research and Development Laboratory at Rocky Point, Long Island constructed and installed in the top of the Empire State Building one each of two new types of variable frequency transmitters including their antennas. These transmitters were used for survey tests in collaboration with the Receiver Research and Development Laboratory at Riverhead, Long Island.[1] The transmitters have rather accurate frequency control and are designed to operate at approximately 150 and 90 megacycles with a power output of 500 to 1000 watts. Comparatively large power output was required in order to simplify and to increase the reliability and accuracy of the field measurements.

It might also be stated that in the past the testing of any apparatus such as antennas, to obtain characteristic data at different frequencies, required the construction of a number of models, or the adjustment of a more or less fixed-frequency test transmitter in several steps of frequency. These methods are both slow and expensive. Consequently, for a long time we have felt the need of means to reduce the time and expense necessary for taking these measurements by making available sources of high-frequency power which are continuously variable over

Reprinted from *RCA Review*, April, 1939.

a wide range of frequency, but which will accurately maintain any frequency to which they are adjusted. This need has been supplied to a great extent in our case by one of these new types of variable-frequency transmitters.

Description

Because of its simplicity of design, accuracy of frequency control, and because it fills a general need for testing purposes it is thought that a description of the 90-Mc survey transmitter would be of general interest.

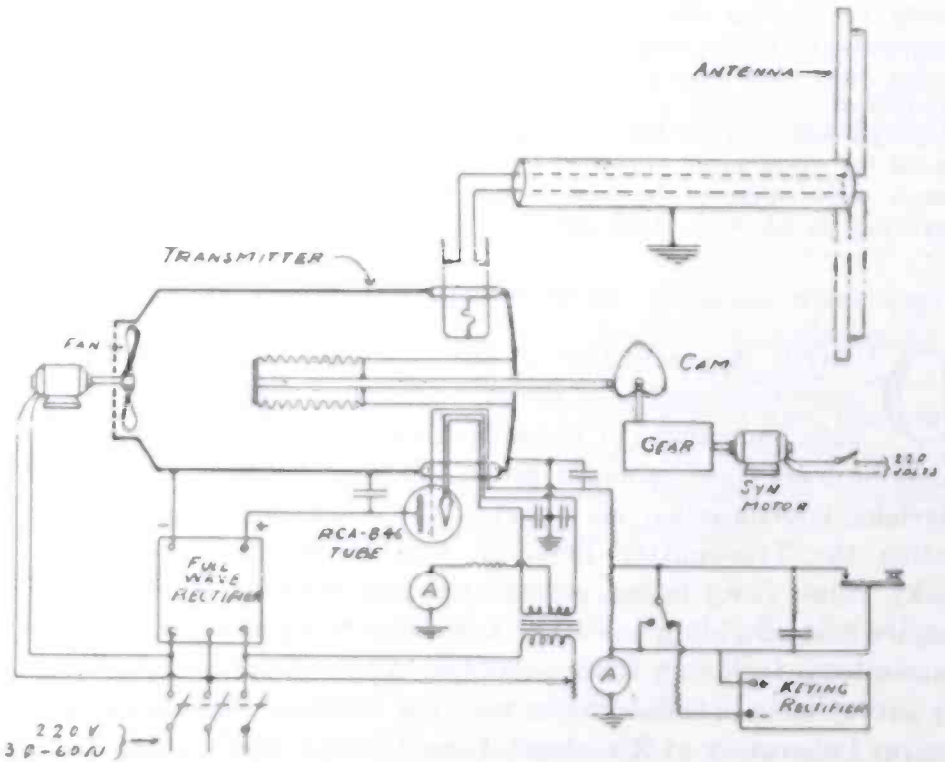


Fig. 1.—Schematic diagram of the variable frequency transmitter.

The transmitter consists of a single-tube oscillator stage with a high degree of frequency control. The principle by which this type of transmitter operates may best be understood by referring to Figure 1. The transmitter is a "grounded anode" type of oscillator in which the radio-frequency voltage of the oscillator tube grid is in phase with that of the filament, but the grid is driven to a greater amplitude. The oscillator driving power is applied through the tube filament. The only oscillating tank, or fly-wheel, circuit in the transmitter consists of a quarter-wave section of two concentric tube conductors which act as a frequency-control line. This line is constructed of copper tubing. One end of the two concentric conductors is shorted together and supported by a heavy copper disk. The grid and filament of the transmitter

tube are coupled to the high-current, or short circuited, end of the frequency-control line through loop conductors which are located near the inner concentric conductor. The outer ends of these loops are grounded for radio-frequency currents to the outer concentric conductor through by-pass condensers before being connected to their respective power sources. In general the length of the frequency-control line determines the operating frequency. The outer line conductor

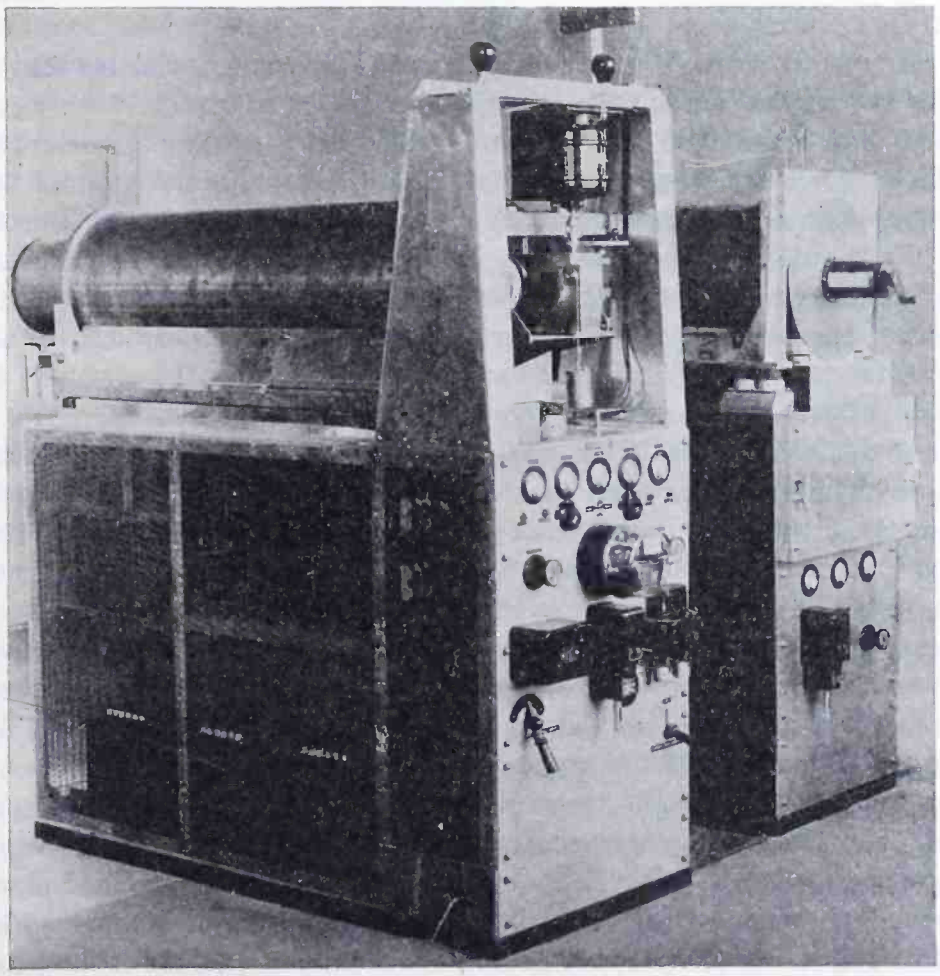


Fig. 2.—Front view of variable frequency transmitters.

is made considerably longer than the inner conductor to reduce influence of end effect when the frequency is varied through the band and to make the wave length correspond closely to the physical length of the inner conductor. It is also to be noted that the inner conductor is generally slightly shorter than a quarter-wave length due to its end capacity. In this case the line lengths are 93 to 96 per cent of quarter-wave length depending upon the transmitter frequency. The end capacity is a larger percentage factor affecting the line length, at the

higher frequencies. The ratio of diameters of outer to inner conductor was made 3.6 because this results in optimum power factor for the line as an oscillating circuit.[2] Also, since the power factor is inversely proportional to the diameters of the conductors, relatively large diameters were used. The outer conductor is 16 inches diameter, which was the largest size seamless copper tubing readily obtainable. There have been published a number of papers pertaining to line control for transmitters and line calculations, to which the reader is referred for detail information.[3]

A general idea of the appearance of the survey transmitter may be obtained by referring to the left-hand unit in Figure 2. The right-hand unit in Figure 2 is a second transmitter of this general type which was built later. This illustration shows the transmitter frequency-control line mounted with rubber shock absorbers on the top of the main plate rectifier. On the front panel of the rectifier are mounted the meters, relays and various controls, while in front of the frequency-control line are mounted the frequency-sweep motor and cam arrangement. The shorted end of the line faces the reader. The transmitter oscillator tube is an air-cooled 846 which is mounted on the under side of the frequency-control line just behind the rectifier panel. The glass envelope of the tube projects up through a large opening in the outer concentric conductor of the frequency control line. The grid and filament of the tube are inductively coupled to the inner conductor of the line by means of loop-shaped conductors. In order to secure proper oscillations the grid loop must be more closely coupled to the frequency-control line than the filament is coupled. The amount of coupling required was determined by experiment. These loops are assembled together in a compact unit, and the outer ends of the loops are by-passed for radio-frequency currents to the outer concentric conductor of the frequency-control line with specially constructed mica condensers. The coupling loops also serve as grid and filament power leads. The anode of the 846 tube is fitted by means of a taper socket into a cylindrical shaped unit having a large number of air-cooling fins. This cooling unit is constructed entirely of copper for maximum heat-conducting efficiency. The clamp which supports the anode cooling unit, together with a sheet of dielectric material, fits against the outer concentric conductor of the frequency-control line in such a way as to form the anode radio-frequency by-pass condenser to ground. Means are provided for unclamping and lowering the entire tube mounting assembly consisting of the anode cooling unit, the tube, the coupling loops and the by-pass condensers. These parts can then be inspected and the tube can be changed through a door in the transmitter shielding. The other end of the frequency-control line contains a ventilating

fan, mounted on rubber shock absorbers, for cooling the tube anode. This end of the line is also fitted with a screen cover to protect the fan and for electrical shielding. An air-flow interlock is mounted under the anode cooling fins to trip off the anode power supply should the air flow become insufficient.

A large portion of the inner conductor of the frequency-control line, at the open-circuit end, is constructed of copper and silver-plated metal bellows. A metal rod extends from the end bellows back through the inner conductor to a cam arrangement. This provides the means for expanding or contracting the bellows which changes the length of the inner conductor, thereby lowering or raising the transmitter frequency in response to the shape of the cam. For the survey tests, this transmitter was designed with a cam arrangement to give it a linear frequency sweep between 81 and 86 megacycles at either one or six sweep cycles per minute as desired.

The transmitter control circuits are designed for switching to either remote or local control for stopping, starting, and telegraph keying the transmitter. Keying of the transmitter is accomplished by applying a high negative grid bias to the oscillator tube from a small rectifier provided for the purpose. The normal working grid bias is obtained by passing the d-c grid current through a resistor. All access doors for tube replacement, etc., are equipped with interlocks to remove dangerous voltages should a door be opened with the power on. The main rectifier, which uses six 872A tubes in a three-phase full-wave circuit, is provided with a hand-operated tap switch for the plate transformers. This enables the rectifier to supply d-c anode current to the 846 transmitter tube in several steps of potential ranging from 2700 to 7700 volts. Two filament rheostats and a voltmeter are provided for proper adjustment of filament heating power. Time delay relays protect the filaments against improper starting. An a-c overload relay and status lights are provided for the main rectifier. An over-current relay, an ammeter and a voltmeter are provided in the anode circuit of the transmitter tube while a milliammeter is provided in the grid circuit. The whole transmitter unit is well shielded.

The transmitter output leads are located at the top and front of the unit where an insulated loop conductor extends down into the concentric frequency-control line. This loop can be moved up or down to adjust its distance from the inner conductor of the line for load coupling variation. A sliding ground connection is provided on the coupling loop for balancing the output voltages.

Figure 3 shows the antenna which was used with this transmitter at the top of the Empire State Building. The antenna was one of the folded dipole type. Handles were provided by means of which the

antenna could be turned to a vertical or horizontal position in order to change its polarity during the survey tests.

The right-hand unit in Figure 2 illustrates the second variable frequency transmitter which has been built of this particular type. This transmitter was built primarily for laboratory use. Although it does not have as many refinements as the first transmitter, it has certain other advantages. This transmitter is not equipped with a plate power rectifier and the control circuits are not so elaborate, the transmitter control being entirely local. However, the outer conductor of

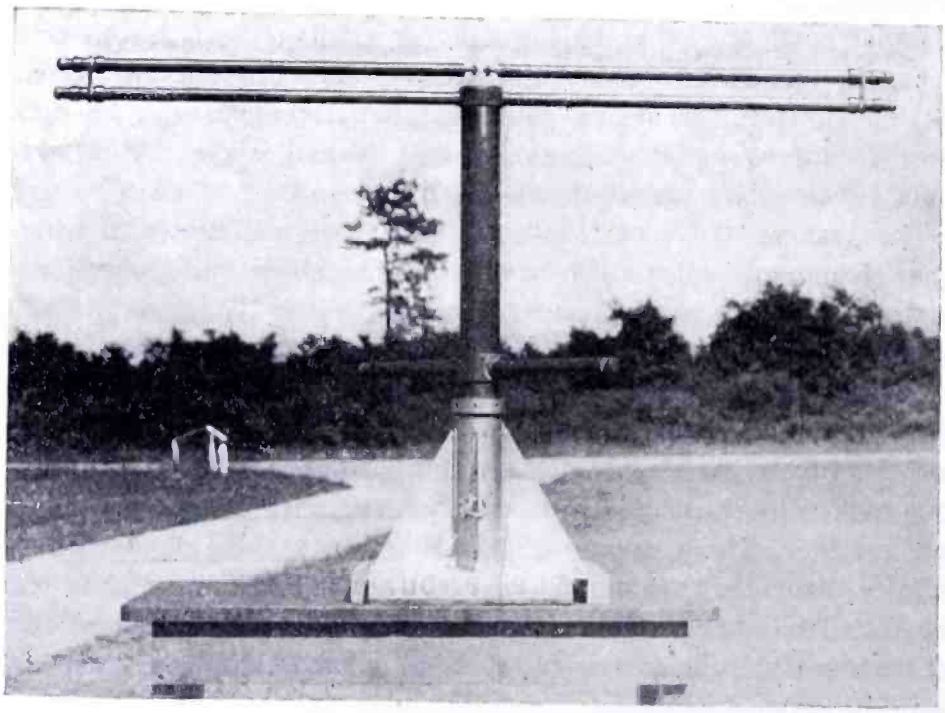


Fig. 3.—Folded dipole type of antenna.

the frequency-control line is longer which gives the transmitter a greater possible frequency range and the metal bellows section of the inner conductor is longer, which provides a greater range of continuous frequency adjustment. Instead of the cam arrangement, the second transmitter is provided with a hand-operated crank with a vernier scale, which turns a threaded rod for changing the length of the inner conductor of the frequency-control line. This permits very accurate frequency adjustment. There are also other improvements in the second transmitter such as larger radio-frequency by-pass condensers for the grid and filament coupling loops and the greater accessibility of the transmitter tubes by the easy removal of the upper section of the front panel. These various features of the second transmitter make it particularly useful for general laboratory work.

PERFORMANCE

The power output from each of these transmitters is 500 to 1000 watts depending upon the frequency and the transmitter adjustments. The transmitter plate efficiency varies from 40 to 50 per cent. This also depends upon the frequency and the transmitter adjustments. The first transmitter has a maximum variable frequency range of 8 per cent at about 85 megacycles. The second transmitter has a maximum frequency operating range of 40 to 100 megacycles by substituting different lengths of inner concentric line conductor. The continuously variable frequency range of the second transmitter is 7 per cent at 40 megacycles and 12 per cent at 100 megacycles. The frequency of these transmitters may be quickly and easily set to any value in the adjustable range with less than one-tenth of one per cent error from the calibrated value. For optimum results, the size of the grid- and filament-coupling loops should be changed for any considerable change in the length of inner conductor of the frequency-control line. In general, larger loops for coupling to the frequency-control line are required for the grid and filament of the transmitter tube at the lower frequencies. With proper design these transmitters are not troubled with parasitic oscillations. However, if the coupling loops have too much inductance and are too loosely coupled to the frequency-control line, parasitic oscillations may occur. These parasitic oscillations have been known to occur at a frequency which caused the frequency-control line to oscillate on a harmonic as well as the fundamental frequency. The remedy for this parasitic trouble is to use a coupling loop assembly having lower self-inductance and to couple it more closely to the frequency-control line.

REFERENCES

[1] R. W. George, "A Study of Ultra-High-Frequency Wide-Band Propagation Characteristics". *Proc. I. R. E.*, Vol. 27, January, 1939. (See Page 90). R. W. George, "Field-Strength Measuring Equipment for Wide-Band U-H-F Transmission". *RCA Review*, April, 1939. (See Page 262).

[2] C. W. Franklin British Patent No. 284,005 and corresponding U.S. Patent No. 1,937,559.

[3] Sterba and Feldman, "Transmission Lines for Short-Wave Radio Systems", *Proc. I. R. E.*, July, 1932. F. E. Terman, "Resonant Lines in Radio Circuits", *Electrical Engineering*, July, 1934. Conklin, Finch and Hansell, "New Methods of Frequency Control Employing Long Lines", *Proc. I. R. E.*, November, 1931. C. W. Hansell and P. S. Carter, "Frequency Control by Low-Power Factor Line Circuits", *Proc. I. R. E.*, April, 1936. (See Page 53). P. S. Carter, "Charts for Transmission Line Measurements and Computations". *RCA Review*, January, 1939.

FIELD STRENGTH MEASURING EQUIPMENT FOR WIDE-BAND U-H-F TRANSMISSION

By

R. W. George

R.C.A. Communications, Inc., Riverhead, N. Y.

Summary—Mainly in the interest of television service, some quantitative data were required on the characteristics of ultra-high-frequency propagation paths. Portable measuring equipment and methods devised to obtain these data are described.

Transmitter systems giving constant output over the ranges of 81 to 86 Mc and 140 to 145 Mc were used, the frequency of each transmitter varied at the rate of 166 kc per second from one extreme of its range to the other. Field strength was measured at each 70-kc increment of the frequency range, a measurement being recorded automatically as the signal frequency passed through the receiver pass-band. The circuits were arranged to hold the recorder indication until changed by the following measurement. About 70 measurements were recorded in one-half minute by means of a novel automatic tuning arrangement.

The signal generator and calibration methods used are also described.

FIELD strength versus frequency measurements afford a means of studying the characteristics of propagation paths, especially where the difference between direct and indirect path lengths is relatively small. It is the object of this paper to describe some equipment with which such measurements were made during a recent study of ultra-high-frequency transmissions from the Empire State Building.[1] In general, the problem was resolved to one of measuring and recording received field strength as the transmitter frequency varied at a constant rate of change over a 5-megacycle range. Measurements over this frequency range are sufficient to indicate the existence, and some important characteristics, of indirect paths which have time delays long enough to cause distortion in television reception.

In the past, one method of making such measurements has been to adjust manually the transmitter and receiver frequency step by step for each measurement to be made.[2] This method is obviously slow and requires perfect coordination between the transmitter and

[1] "A study of U-H-F Wide-Band Propagation Characteristics" by R. W. George, *Proc. I. R. E.*, Vol. 27, January, 1939. (See Page 90).

[2] P. S. Carter and G. S. Wickizer, "Ultra-High-Frequency Transmission Between the RCA Building and the Empire State Building in New York City", *Proc. I.R.E.*, Vol. 24, August, 1936.

Reprinted from *RCA Review*, April, 1939.

receiver operators. For the purpose of making a rather extensive survey, it was considered desirable to speed up the operation to permit a greater number of measurements to be taken, and to record automatically the measurements in a form readily utilized.

One relatively simple measuring system to accomplish this result might consist of a wide-band receiver associated with a suitable recorder. Such a receiver, having constant response over a 5-megacycle range has the disadvantage that its inherently high noise equivalent would not permit satisfactory recording of weak signals. Otherwise, this method would be very convenient as no tuning of the receiver to the changing transmitter frequency would be necessary. The measuring equipment to be described was used to record a large number of measurements over the frequency range with a result similar to that which would be had if continuous measurements with frequency were made. It also embodied a conventional type of receiver which permitted measurements at low field intensities. By making the transmitting and receiving equipment conform to the following conditions, the recorded data required no corrections in order to determine the field strength at a given frequency.

The transmitting systems were so designed that the radiated power was constant as the frequency varied from 81 to 86 megacycles or 140 to 145 megacycles[3]. The transmitted frequency varied at the rate of 166-kc change per second and required one-half minute to go from one extreme of the frequency range to the other. By the use of half-wave doublet receiving antennas having substantially constant response over the desired 5-megacycle range, connected to a 75- or 100-ohm transmission line terminated with resistance at the receiver, constant voltage was delivered to the receiver for constant field strength at the antenna.

An ultra-high-frequency, triple-detection receiver having a 100-kilocycle i-f pass-band was adapted by ganging the ultra-high-frequency detector and heterodyne-oscillator controls and adjusting the circuits to extend the 5-megacycle signal range over the major portion of the control dial. The response of the receiver was adjusted to be constant over the required tuning range. It was not considered feasible to incorporate conventional automatic-frequency control of the ultra-high-frequency circuits because of the required flexibility of the measuring system. An automatic step-by-step frequency control, coordinated with the measuring system, was used which will be understood more clearly after the following description of the functions of this system.

[3] G. L. Usselman, "Wide-Band Variable-Frequency Testing Transmitters," *RCA Review*, April, 1939. (See Page 255).

The most suitable time to make a measurement would be when the signal, changing in frequency, was at the mid-band of the intermediate-frequency amplifier. Thus, the signal is in the receiver about three-tenths of a second before it reaches the middle of the pass-band which gives the diode time to reach an output corresponding to the input signal. The operation of the measuring functions is indicated in Figure 1. An audio tone which was utilized to cause a measurement to be made, was obtained by combining, in a separate detector, the intermediate-signal frequency and a beat oscillator tuned to the mid-band frequency of the second intermediate-frequency amplifier. This tone, fed through an audio amplifier and associated rectifier, caused the relay R_1 to connect the diode to the grid of the first d-c amplifier tube for the duration of the tone which was about 0.12 second. The small condenser C_1 was charged to the corresponding diode voltage and maintained this voltage on the grid of the first d-c amplifier tube after the signal had passed out of the audio-frequency pass-band and the diode was disconnected. To accomplish this, it is important that the grid current of the first d-c amplifier be extremely low, a condition that was easily obtained by the use of a G.E., FP-54 type tube. It was necessary for the capacity of C_1 to be rather small, on the order of 0.01 microfarad, for it to be fully charged in the short charging time allowed. With this arrangement, over an hour was required for the grid voltage to decay to 37 per cent of its original value and, of course, no appreciable change occurred between measurements. The d-c amplifier had a linear input-output characteristic and operated a conventional recorder. This method of holding the recorder at its last measurement until a subsequent measurement is made, was advantageous in that a sequence of measurements gave an approximately smooth curve. It also gave the recorder-movement time to indicate the true amplitude of each measurement.

With the system as described up to this point, it will be apparent that if the receiver is tuned to say the low-frequency end of the signal-frequency range, a measurement will be made as the transmitter frequency, increasing from its minimum value, passes through the pass-band of the receiver. After this measurement, if the receiver is quickly tuned to a higher frequency, the increasing signal frequency will again come into the audio-frequency pass-band and cause another measurement to be made. By carrying on this procedure a complete series of measurements can be made. In advancing the receiver tuning, the receiver is tuned through the signal frequency, but this can be done so quickly that the signal is not in the pass-band long enough to cause a false measurement to be made. Tuning by hand in this man-

ner permitted about fifty measurements to be made over the 5-megacycle range in one-half minute. These measurements were spaced with fairly equal frequency increments according to the skill of the operator. The automatic receiver-tuning means permitted about 70 measurements to be made with substantially equal frequency increments between measurements, and when the relay R_1 returned to normal signifying that a measurement was completed, it caused another relay, R_2, to be operated which in turn caused the receiver tuning to be changed by means of a modified automatic-telephone circuit-selector mechanism. This tuning device simply notched up the tuning controls to a higher frequency by approximately 70-kilocycle incre-

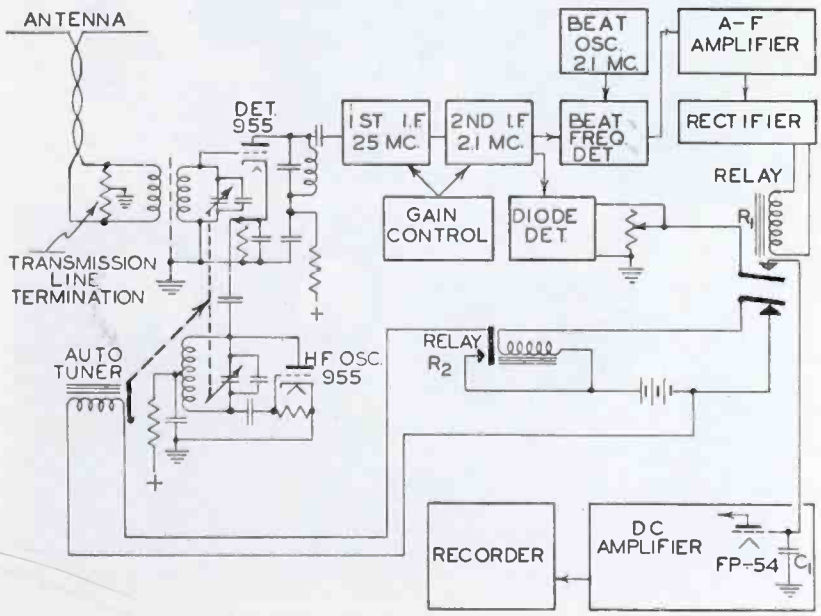


Fig. 1—Diagram showing elements and functions of the measuring system, except the signal generator.

ments. At the instant an incoming signal caused relay R_1 to operate, the holding current in relay R_2 was removed which in turn caused current to be supplied to the auto tuner. The auto tuner was so arranged that the application of current in its coil caused a pawl to engage an advanced tooth on an associated ratchet which was geared with the tuning controls. As the signal frequency increased, the audio-beat note passed out of the response band of the a-f amplifier and thus restored R_1 to normal as shown in Figure 1. It will be noted that this last operation first disconnected the diode from the d-c amplifier and then applied current to the relay R_2. This removed the current in the auto tuner, the armature of which in being restored to normal by a spring caused the receiver tuning to be changed by means of the pawl and ratchet.

As a matter of uniform procedure, the system was designed so that measurements were taken with increasing frequency. Thus, a few minutes of observation indicated the proper gain of the receiver to use for the range of signal levels available and the measurement was started by disengaging the automatic-tuning drive and tuning the receiver to the lowest frequency. The automatic-tuning drive was

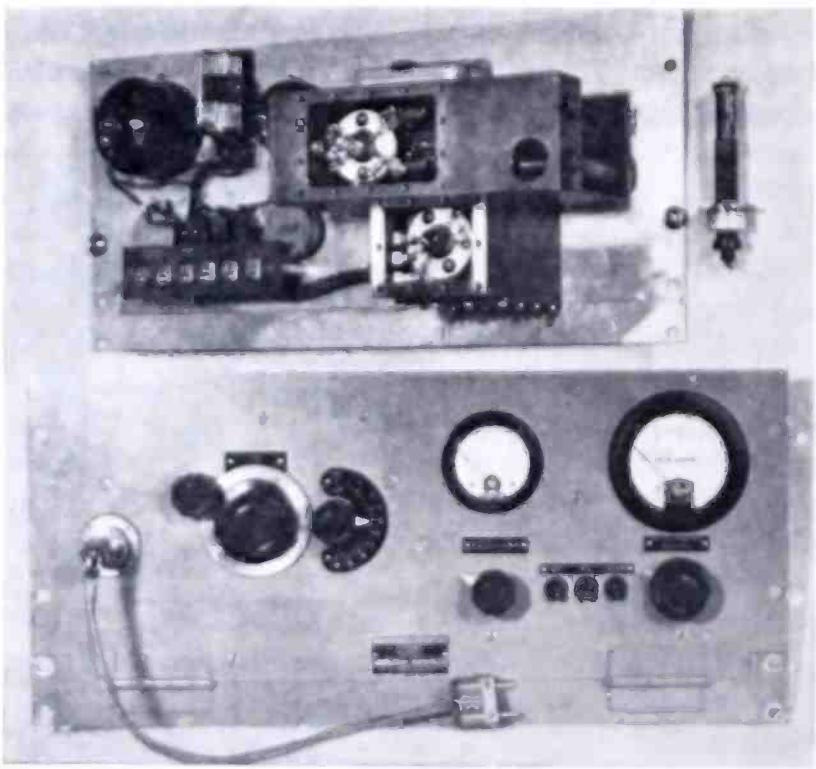


Fig. 2—Standard u-h-f signal generator. In the top view (rear) the shield covers are removed exposing the u-h-f oscillator in the upper compartment, and the voltmeter in the lower compartment. Through the circular opening can be seen part of the inductance, L_1. At the extreme right is shown the movable element of the attenuator.

then re-engaged and the series of measurements started when the signal began increasing in frequency after reaching its lowest frequency.

With this system, it was only necessary to keep the transmitter-frequency sweep in continuous operation and to observe a predetermined schedule for polarization of the transmission. Maximum stability of the measuring system was obtained by the use of voltage regulated-power supplies. Normally the equipment was operated from 110-volt a-c sources.

Calibration of Equipment

For all such measurements, some means of comparing data with a common reference is necessary. The signal generator used as a reference also provided a fairly accurate means of measuring field strengths. This means was embodied in the 74-ohm output circuit of the signal generator which is known to be substantially equivalent to the impedance of a half-wave doublet antenna. Thus, by substituting the signal generator for the antenna, the equivalent voltage in the antenna was known, which with its known effective height provided the corresponding field strength indicated by the receiver output. As it was somewhat cumbersome to check the overall gain of the measuring system frequently by this method, it was found practical to measure the transmission-line loss and calibrate the measuring system by connecting the signal generator to the receiver-input terminals, including the transmission-line termination, whenever a calibration was necessary.

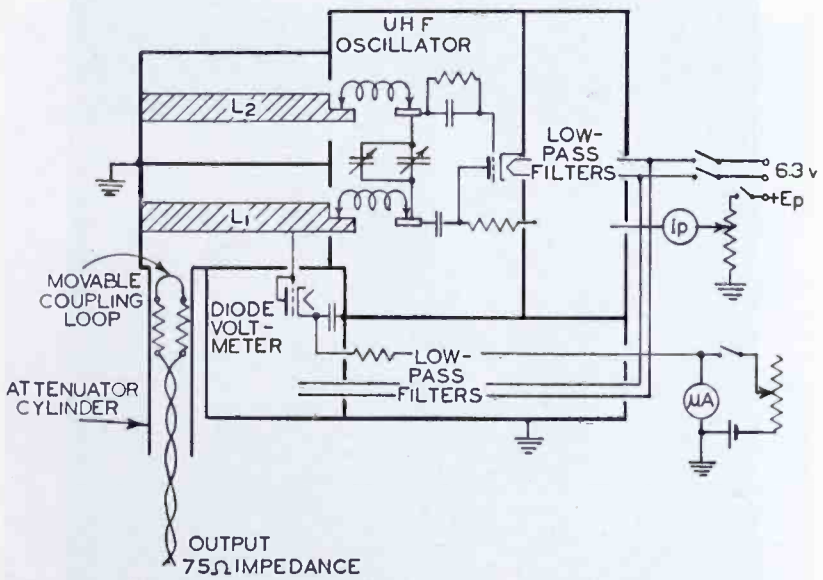


Fig. 3—Diagram of u-h-f signal generator.

The signal generator also provided means of checking the input-output response and the flatness of response over the frequency range of the measuring system. Final evidence that the entire system had constant response over both frequency ranges was obtained by making measurements in an open, unobstructed field where no serious indirect-path waves, with the exception of the wave reflected from the ground, were expected. The path length of the wave reflected from the ground was so nearly equal to the direct wave-path length that it caused no appreciable field strength change over the 5-Mc range with the result that the recorder indicated substantially constant field strength.

Details of the ultra-high-frequency signal generator, Figure 2, are given in the diagram, Figure 3. It will be seen that two sections of concentric line, L_1 and L_2, comprise part of the inductance in the ultra-high-frequency oscillator circuit. The voltage across a portion of the inductance, L_1, can be set to calibrated values by means of the vacuum-tube voltmeter and plate-voltage control of the oscillator. Output from the oscillator is taken by inductively coupling a small loop with the maximum-current end of the inductor L_1. The coupling

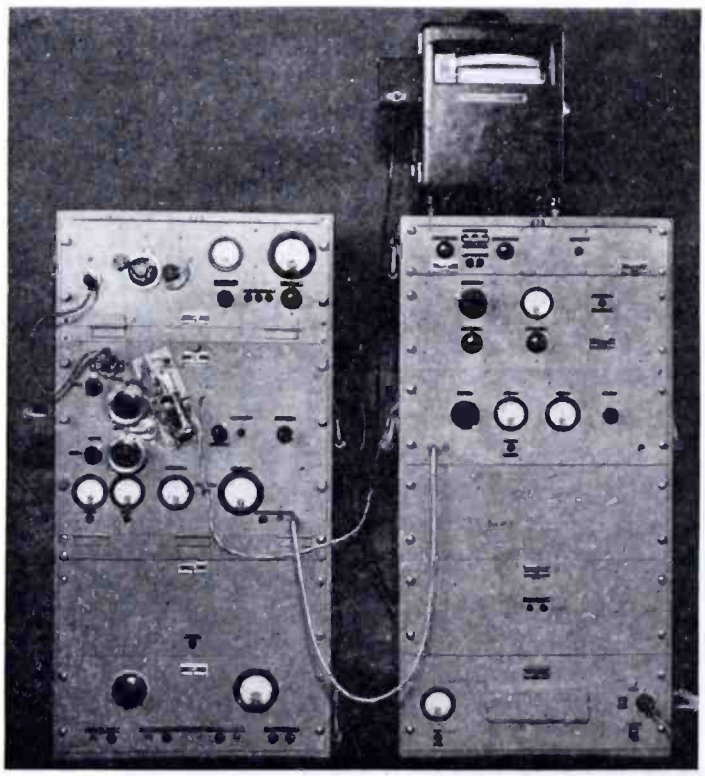


Fig. 4—Measuring equipment in portable racks. The units are, beginning at the top of the left-hand rack, standard signal generator, u-h-f receiver with the automatic-tuning device attached, power-supply unit, and spare audio amplifier. In the right-hand rack are, the recorder (on top), final d-c amplifier, FP-54 d-c amplifier, tone rectifier, regulated power supply, and a-c voltage regulator.

loop is connected to a short 74-ohm transmission line with a 37-ohm resistor on either side thus providing a 74-ohm output circuit. The coupling of the loop, and thus the output, is controlled by a calibrated threaded movement which moves the coupling loop inside a metal sleeve. This attenuator gives a range of 60 db with about 1¼ inches displacement, the calibration of which is substantially independent of frequency. The signal generator was calibrated by comparison with other standard-signal generators at 30 and 40 megacycles. Some satisfactory calibration checks have also been obtained with field-strength

measurements at frequencies between 80 and 100 megacycles compared with field-strength measurements made with half-wave antennas incorporating a current- or voltage-measuring means at the center. A frequency range of from 30 to 200 megacycles is obtained by the use of two sets of plug-in inductance elements.

Miscellaneous

Several items were found useful in facilitating measurements in the field. Simple attenuator pads made up with quarter-watt carbon resistors, were inserted in the transmission line at the receiver to obtain a loss of as much as 24 db. These were necessary in some cases

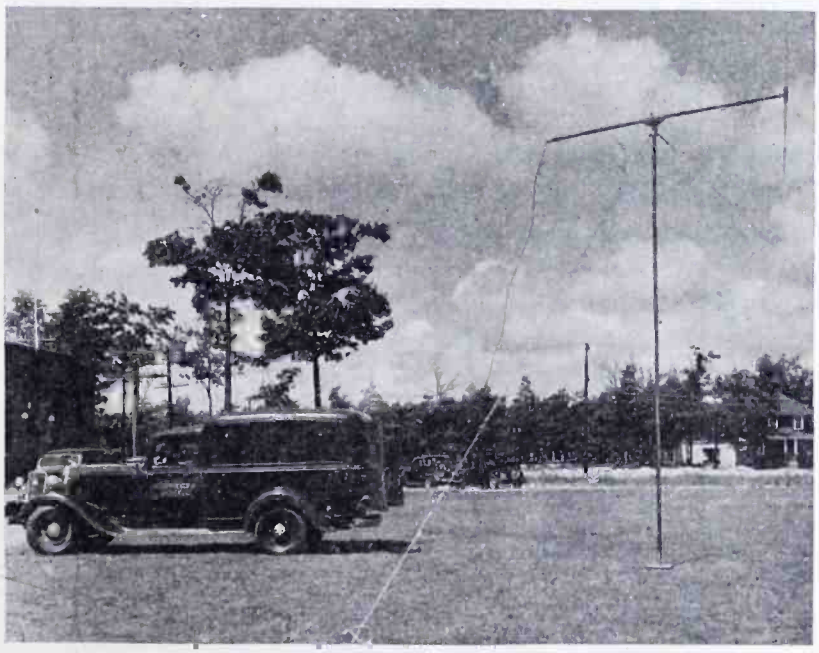


Fig. 5—Portable mast set-up, showing arrangement of antenna and transmission line.

where the received signal was several hundred millivolts. The frequency-response characteristic of these pads was found to be quite satisfactory over the required frequency range and their loss was conveniently measured by means of the signal generator.

The measuring equipment, Figure 4, excepting the recorder, was mounted in two portable racks which were about 40 inches high and accommodated standard 19-inch panel units. Suitable clamps were provided to fasten the racks in place in a ¾-ton panel truck. Measurements were made without removing the equipment from the truck whenever feasible.

The antenna-mounting means were extremely flexible in order to support the antenna satisfactorily at a large variety of receiving loca-

tions. A sectionalized mast, Figure 5, was provided which could be set up at various heights up to 22 feet. The top of the mast supported the center of a 10-foot cross-arm along which the transmission line was draped. The antenna made of two ¼-wavelength aluminum tubes was mounted by a bakelite clamp which in turn was supported on the end of a pole or at one end of the cross-arm on the mast. The cross-arm was mounted so it could be turned through 90 degrees by means of ropes, enabling easy control of the antenna polarization. The position of the antenna could also be changed over a circle of 5-foot radius by simply turning the mast, the three guy ropes being fastened to a movable collar at the top.

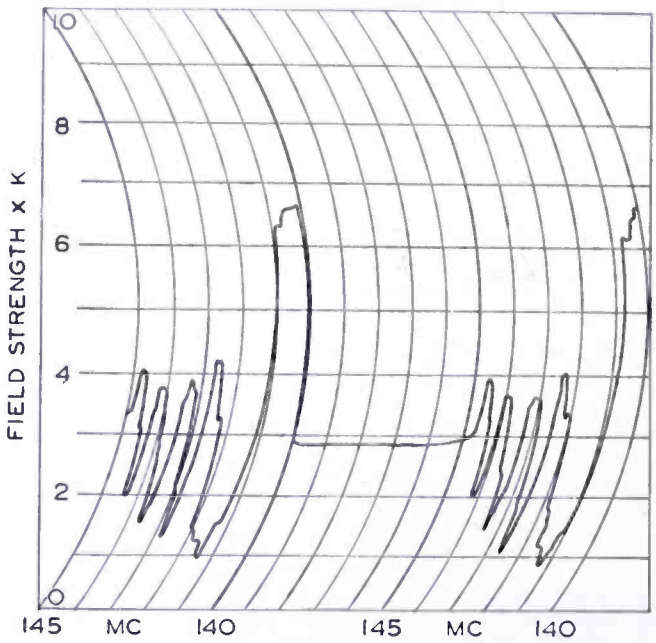


Fig. 6—Two frequency versus amplitude curves recorded under unchanged conditions.

The transmission lines used were of the twisted pair, Latox rubber-insulated type. For required lengths under 60 feet, a 74-ohm (characteristic impedance) line was available which had a loss of 10½ db per 100 feet at 83 megacycles and about 15 db per 100 feet at 145 megacycles. Longer lines sometimes used, having a characteristic impedance of 100 ohms, had a loss of 6½ db per 100 feet at 83 megacycles and 9½ db per 100 feet at 145 megacycles.

The recorder was critically damped and followed the changing signal levels quite accurately except in one case where extremely large and fast variations were encountered. To insure a uniform time or frequency scale, the chart was driven at the rate of three inches per minute by a small synchronous motor.

In general, the measuring equipment was quite satisfactory and gave consistent repeat measurements. This performance is illustrated by the two samples of recorded data shown in Figure 6 which were obtained in sequence with unchanged receiving conditions. The time interval between curves is one-half minute which was required for the transmitter frequency to return from 145 Mc to the starting frequency of 140 Mc. During this interval, the receiver tuning was changed to 140 Mc and the automatic-tuning device re-engaged so that measurements were repeated when the transmitter frequency began increasing from 140 Mc. Since only one-half minute was required in which to make a complete series of measurements, a large number of measurements could be made at each location with a variety of antenna positions.

ACKNOWLEDGMENT

The measuring methods embodied in the described equipment were developed under the guidance of Mr. H. O. Peterson. Mr. K. G. MacLean gave valuable assistance in adjusting the equipment to give the desired performance.

A NEW METHOD FOR MEASUREMENT OF ULTRA-HIGH FREQUENCY IMPEDANCE

By

Stuart W. Seeley and William S. Barden

License Laboratory, Radio Corporation of America

1. Foreword

SINCE the earliest days of the radio art, progress has been assisted by continuous improvement of measurement technique, whereby the reduction to practice of important discoveries and the refinement of design have been expedited, and maximum performance achieved.

The contemporary development of a precise measurement technique is relatively slow: partly because it is less interesting and at times less important than the experimental demonstration of new ideas, and partly because precision of measurement is often difficult when simple or convenient methods are needed. As the improvised methods of rough measurement give way to suitably refined and precise techniques, the advance of a new art can be engineered straightforwardly, and more economically.

In several important respects, the present status of ultra-high-frequency practice marks it as a new art. Meritorious results have been achieved with short and ultra-short waves. Nevertheless, from the standpoint of good engineering practice, the measurement of ultra-high-frequency circuit properties has remained essentially undeveloped.

This paper is concerned chiefly with a new and simple method for accurate measurement of resistance and reactance at high and ultra-high frequency. The method is described in Section 3, and treated analytically in Section 4. Section 2 is a brief discussion of preliminary considerations.

It is believed that the method described is not only extremely simple and convenient to use, but that it provides a degree of accuracy much higher than that provided by previous and more complicated systems.

2. Factors Involved in Impedance Measurements at Ultra-High Frequency

As regards measurement at high frequency, the teachings of experience emphasize the importance of increased care and precaution as the

frequency is increased. There are various reasons for the difficulties of precise measurement at high and "ultra-high" frequency. Broadly speaking, there is one basic reason, namely, that by virtue of slight imperfections arising from physical limitations of construction, circuit elements (including certain meters) are not strictly circuit constants when the frequency is variable. Not only are inductance, capacitance, and resistance standards dependent upon frequency, but the inductance and capacitance of circuit wiring acquires significance at high frequency. Hence, methods of measurement which can be used safely at low frequency are often found to be definitely unsatisfactory at high (or very high) frequency.

Of course it is not impossible to proceed very carefully and to develop a frequency calibration of standards, such that their use at very high frequency becomes feasible. However, that is not a straightforward procedure from some important standpoints, and proof of the accuracy of the standards is likely to be quite involved. Experience with the accurate measurement of resistance and reactance at high frequency often leads to the desire for a method which minimizes the required amount of absolute information to the point of feasibility.

Suppose it be required to measure the 50-megacycle resistance of a 100-ohm carbon resistor. A "Q" meter is likely to be readily available, but 100 ohms is much too high to use in series with the tuned circuit, and much too low to use in shunt. If the 100-ohm resistor be connected across an appropriately coupled secondary, thus adapting it to the "Q" meter, a knowledge of the degree of coupling at once becomes involved. Eventually, as one thing leads to another, a complicated procedure ensues and leads to a questionable result. Of course other methods may be considered, one being the volt-ampere scheme. It works well on paper, but a thermal milliameter is not readily adaptable to high-frequency practice. Most radio laboratories are not equipped to measure high-frequency current accurately, although recent vacuum-tube voltmeter developments render the reasonably accurate measurement of ultra-high-frequency voltage not particularly difficult.

However, it is still well to avoid dependence upon absolute information of voltage when measuring resistance or reactance at ultra-high frequency. For example, even with a truly modern vacuum-tube voltmeter, it is safer to assume that the deflection vs. voltage characteristic is the same *on a percentage basis* at ultra-high frequency as it is at low frequency, than to assume that the voltage measurement of the vacuum-tube voltmeter is independent of frequency up to (say) 100 megacycles. At ultra-high frequency, certain vacuum-tube voltmeters can be relied upon to obey their low-frequency deflection law on a

percentage basis. Accordingly, a preferred method of resistance or reactance measurement at ultra-high frequency should involve a *ratio* of voltages, *determined as a ratio directly*, rather than involve absolute measurements of voltage.

At ultra-high frequency, a voltage ratio may be determined by using a vacuum-tube voltmeter without self bias on any electrode, and employing only the limited voltage range wherein the deflection is proportional to the square of the voltage. It can be shown analytically that for various tubes this range of deflection is from zero to 50 or 100 microamperes. Accurate low-frequency tests for this square law often reveal it so well that any discrepancy is likely to be due to an error in measuring the voltage (even at 60 cycles). With the range of square law operation known, it is quite proper to accept the assumption that the vacuum-tube voltmeter obeys a square law at ultra-high frequency. Hence it becomes a simple matter to learn accurately a voltage ratio at ultra-high frequency: e.g., a deflection ratio of 2 corresponds to a voltage ratio of $\sqrt{2}$.

Without knowledge of current or voltage magnitude, it is axiomatic that the measurement of any impedance (resistance or reactance) requires that at least one circuit element be a known quantity on which precision of measurement is dependent. From this standpoint, both thought and experience lead to the choice of a physically small air condenser as the circuit element whose assumed properties at ultra-high frequency are based on low-frequency measurement. Moreover, it is preferable to require knowledge only of the incremental capacitance due to rotation of plates from any predetermined position, rather than to involve the entire capacitance (which includes strays). This preferred character of absolute information, and dependence upon a directly determined voltage ratio instead of upon absolute magnitudes of voltage, plus the normally required degree of careful circuit management at very high frequency, are employed in the method described in Section 3. This method is an adaptation of a principle which first may be considered briefly, as follows:

Consider a simple series circuit comprising an inductance L, a capacitance C, a resistance R (due chiefly to the coil), and an impressed voltage E at the frequency of resonance. Q having its standard definition of $\frac{\omega L}{R}$, the condenser voltage at resonance is QE. Resonance is established by adjusting C for maximum deflection of a vacuum-tube voltmeter across C. By changing C from its value at resonance, detune the system until the vacuum-tube voltmeter deflection corre-

sponds to $\frac{1}{\sqrt{2}} \times 100$ per cent of the condenser voltage at resonance. The system is detuned by means of an incrementally calibrated vernier condenser. Note the change in capacitance, and designate it by $\triangle C$. $\frac{1}{\omega(\triangle C)}$ is numerically equal to the parallel resonance-tuned impedance which is developed across the condenser terminals: i.e.,

$$\frac{1}{\omega(\triangle C)} = Q\omega L = \frac{Q}{\omega C} = \frac{L}{CR}$$

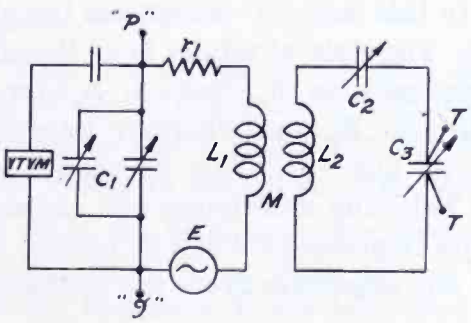


Fig. 1

This principle, as here outlined on the basis of its preferred character of data, underlies the method of measurement which is described in Section 3.

The foregoing discussion of a measurement principle is trivially in error. The error is definitely unimportant when Q is normally high. Meticulous precision, aside from being unnecessary and usually absurd, almost invariably leads to much complexity of both method and explanation. Also in the following descriptions, commission of negligible error is preferred to useless confusion.

Suitably high-Q circuits in the measuring device can be realized readily, thus rendering the result as nearly true as is necessary or experimentally demonstrable. When the error of analysis is of the order of only one- or two-tenths of a per cent, of course the ultimate error in measurement is due mainly to other causes.

3. Description of Method

Note Figure 1. Circuit 1 comprises the series connection of a fixed inductor L_1, a variable condenser C_1, and a voltage E at the frequency involved. E may be obtained by a magnetic coupling of Circuit 1 to a

driver, provided that the incidental coupling of the driver to Circuit 2 is negligible, and provided that the coupling of Circuit 1 to the driver is so loose (requiring sufficient power, of course) that a detuning of Circuit 1 does not result in an appreciable change in the induced voltage E. A good method of obtaining E is by means of resistance coupling to a driver (not shown). In this case, the resistance placed in the position of E should be very small in comparison with r_1 on Figure 1. r_1 is the effective series resistance of Circuit 1, and is due mainly to the low resistance of the high Q inductor winding.

Another good (and in some cases a more convenient) means of driving Circuit 1 is to connect the points "p" and "g" (on Figure 1) into the plate circuit of an r-f pentode whose grid is driven by a signal generator. In this case, "p" designates the plate connection, and "g" the r-f ground. The state of affairs is as though the series voltage E were actually impressed as in Figure 1. A current $E_g G_m$ flows into a shunt combination of R_p and whatever impedance presents itself across the points "p" and "g". This impedance and R_p need not be known. All of the following description and explanation refer to Figure 1 with the series impressed voltage E, but the use of a driver tube need not result in any departure from the method of measurement as outlined.

Circuit 2 comprises the series connection of a fixed inductor L_2, a variable air condenser C_2, and another variable air condenser C_3. Circuit 2 is magnetically coupled (adjustable) to Circuit 1. Conveniently accessible terminals T are placed across C_3. Circuit 2 is shown without any resistance. A normally high Q_2 renders this resistance unimportant.

C_2 is used to tune L_2 when C_3 is shorted by means of a small strip of crimped copper which is gently inserted between plates at (or near) the middle of the condenser. Of course this means of shorting C_3 is preferable to connecting a wire between the terminals T, because L_2 includes circuit wiring effects. It is important that the adjustment of C_2 to tune L_2 be made when C_3 is shorted without changing the effective value of L_2. The adjustment of C_2 is not so critical that perfection is demanded in the matter of maintaining a constant effective value of L_2 with and without C_3 shorted, but the diminution of slight errors naturally avoids a large error due to their possible sum.

Of course there is likely to be an appreciable degree of capacitive coupling of Circuits 1 and 2, since L_1 and L_2 are occasionally in fairly close proximity in order to obtain sufficient magnetic coupling. However, the capacitive coupling as well as the magnetic coupling need not be known. An effective value of M results from the combined coup-

lings, but its actual value is not involved in the evaluation of resistance or reactance from the data obtained.

C_1 and C_2 need not be known. Their adjustments are made by observing the vacuum-tube voltmeter across C_1. The value of C_3, however, must be known on an incremental basis. The dial on C_3 indicates zero capacitance when the plates are in any predetermined position, such as the position of minimum capacitance. $\triangle C_3$ is the required information, and a physically small variable air condenser may be relied upon at very high frequencies where other devices are not to be trusted.

Suitable values of L_1, C_1, L_2, and C_2 depend upon the frequency (or range of frequencies) involved. In general, convenience may be favored in this respect.

Normally good technique should be employed in the physical disposition and wiring of circuit elements. For example, the leads from the inductor L_2 should be fairly short, and separated not less than (say) ½ inch. This results in negligible stray capacitance across L_2.

The leads from C_3 to the terminals (so-called) T should be short and essentially non-inductive: e.g., strip type conductors. To this point, reference to the terminals T has been made in an essentially diagrammatic sense. It is altogether desirable to construct the strip type conductors to serve as terminals by bending a short strip back on itself, thus forming a terminal of the spring jaw type. At ultra-high frequency, it is important that these conductors (or terminals) lead to the condenser C_3 *directly*, instead of to other points on circuit wiring which are merely diagrammatically across C_3. This precaution is not on account of circuit wiring capacitance across C_3 (since only $\triangle C_3$ is involved as regards a known quantity), but because of the lead inductance of condenser C_3 "itself". C_3 (or rather $\triangle C_3$) is calibrated at low frequency. In order that this calibration shall hold at very high frequencies, the lead inductance of C_3 must be duly minimized. This lead inductance, as such, is only whatever inductance exists between the electrical center of C_3 and the points whereat the strip type conductors (terminals) are connected to C_3. Lead inductance from C_3 to C_2 or to L_2 is not involved in this respect, because that lead inductance is effectively part of L_2 when C_2 is adjusted for resonance in Circuit 2 as described subsequently. The physical arrangement of parts should be such that the short terminal-conductors across C_3 are directly exposed in order to permit an essentially leadless connection to an external impedance which is to be measured. In many cases these terminal conductors may be dispensed with entirely, as by soldering the external device directly across C_3.

The vacuum-tube voltmeter is used as an indicator, but its law of indication must be known qualitatively, and it should be of ultra-high frequency design so that at 50 or 100 megacycles it may be relied upon to obey whatever law of indication it is found (or made) to have at low frequency. As noted in Section 2, square law indication is natural, provided that the applied signal voltage is not excessive.

A device whose resistance and reactance are to be measured is connected across the terminal-conductors T. The process of measurement is as follows.

(a) With C_3 at its minimum value (i.e., $\triangle C_3$ equal to zero), and without any connections to terminals T, impress E at the frequency involved, and adjust C_1 for resonance of Circuit 1.

(b) Short C_3 by the means noted above. Then adjust C_2 for resonance of Circuit 2, by observing minimum deflection of the vacuum-tube voltmeter. Remove the shorting device from C_3.

(c) Ensure that Circuit 1 is exactly resonant, by means of a slight readjustment of C_1 or by a slight change in frequency. Then obtain 70.7 per cent of that voltage across C_1 by either: (1) decreasing the frequency of E, or (2) by decreasing C_1.

(d) Retune the system to exact resonance by adjusting C_3. This is done most accurately by increasing C_3 so that the vacuum-tube voltmeter deflection passes through its maximum value and returns to the value obtained in (c). Note that increase of C_3. Half of that value is $\triangle C_3$, and upon setting C_3 at $\triangle C_3$ the system is tuned to resonance as precisely as necessary (or possible).

(e) Note $\triangle C_3$ and the vacuum-tube voltmeter deflection. Then connect across T the impedance to be measured. This results in a decreased deflection of the vacuum-tube voltmeter. By a further adjustment of C_3, retune (if necessary) to maximize the second deflection. Note the necessary change of C_3, and designate it by $\triangle' C_3$. Also note whether it is necessary to increase or to decrease C_3 in order to retune the system after connecting across T the impedance to be measured.

(f) Knowing the deflection law of the vacuum-tube voltmeter, the first and second deflections in (e) may readily be translated as the ratio of the first voltage to the second voltage: i.e., $\frac{E_1}{E_2}$. Let $P = \frac{E_1}{E_2} - 1$.

(g) To this point, three quantities have been learned: $\triangle C_3$, $\pm\triangle' C_3$,

and P. These three known quantities are used in the simple evaluation of R and X, as follows:

$$R = \frac{1}{P\omega(\triangle C_3)} \qquad X = \frac{1}{\omega(\triangle' C_3)}$$

R and X are a resistance and a reactance in *shunt*, which, at ω, are the equivalent of whatever impedance is connected across T for measurement. When that impedance is a single arm comprising series elements, R and X may readily be converted to their equivalent series combination. The conversion yields the values of the series elements which were unknown. However, it may not be necessary to make the conversion, since in some cases the shunt equivalent of a series circuit may be used in analysis as appropriately as the actual series combination.

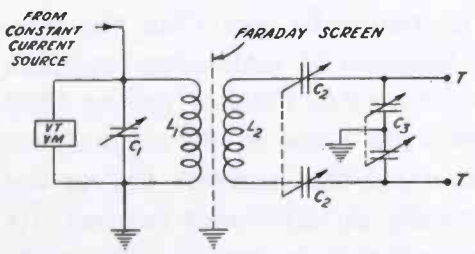


Fig. 2—A new method for measurement of ultra-high-frequency impedance.

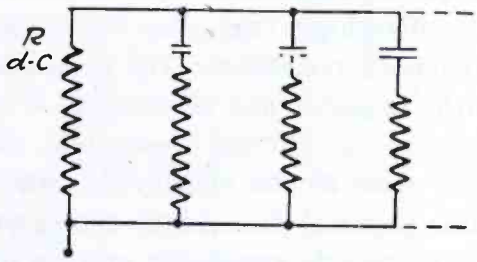


Fig. 3
(Probable Circuit Equivalent of Typical Carbon Resistor)

The interpretation of $\pm\triangle' C_3$, leading to $\pm X$, is as follows. In step (e) when C_3 must be *de*creased in order to retune the system, X is a *capacitive* reactance. The system was in tune before the unknown impedance was connected across T. To restore resonance, of course it is necessary to *remove* the same amount of capacitance from C_3 as is *added* across C_3 by connecting an unknown impedance across C_3. Similarly, a necessary *in*crease of C_3 to restore resonance means that X is an *inductive* reactance.

The foregoing process of measurement is explained analytically in Section 4 after next noting a few results of measurement. Since it is likely that the high or ultra-high frequency performance (or particular properties) of transmission lines may be an important object of measurement, a compendium on distributed LCR systems is included at the end of this paper.

The circuit used for the measurement of impedances of balanced devices, such as balanced transmission lines, is shown in Figure 2. It differs from the basic circuit of Figure 1 by virtue of the use of split

condensers C_2 and C_3, which results in a symmetrical disposition of all elements with respect to ground. The condensers labelled C_2 in Figure 2 are separate components, which are mechanically ganged. C_3 is a split stator, split rotor condenser, with stators connected to the terminals *T T*.

It may be necessary to provide for the insertion of a Faraday screen between L_1 and L_2, to minimize the capacitive coupling unbalance which may occur as the result of driving the primary coil from an unbalanced voltage source.

A few 0.5-watt carbon resistors having from 100 to 5,000 ohms were measured at ultra-high frequency up to 100 megacycles. Resistors of this type show *less* resistance at ultra-high frequency than at low frequency or direct current, but this effect is negligible for low values of resistance, such as 100 ohms. At higher values of low-frequency resistance, such as several hundred or a few thousand ohms, the ultra-high frequency resistance is materially less than the low-frequency resistance. The percentage decrease of resistance increases with frequency and resistance. A 0.5-watt carbon resistor having 5300 ohms direct current resistance, shows 4100 ohms at 50 megacycles, 3600 ohms at 100 megacycles, whereas a 100-ohm resistor having the same physical dimensions shows practically no difference between its 100-megacycle resistance and its direct-current resistance. The state of affairs is as though the physical structure of these carbon resistors were such that the electrical circuit equivalent is probably as indicated on Figure 3, where the parallel arms have unlike values of capacitance and resistance such as to cause the a-c resistance to decrease with increasing frequency. However, typical carbon resistors, as measured, show a shunt capacitance merely of the order of 0.5 $\mu\mu$fd, and some of this is due to resistor leads.

Note on Skin Effect: It is not illogical to assume that skin effect may be involved in ultra-high-frequency resistance of these typical carbon resistors, but theory does not predict skin effect to any appreciable degree. The following expression (1) is a rigorous theoretical formula for the calculation of skin effect,

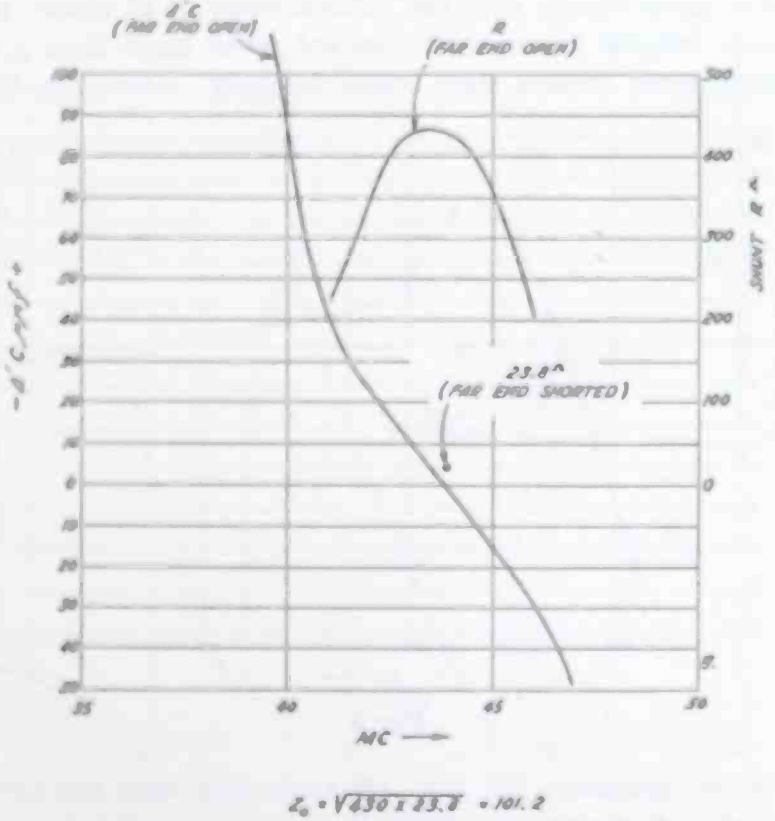
$$r = R_o \left[1 + \left(\frac{1}{12}\,\frac{2\pi FS\mu}{R_o 10^9}\right)^2 - \left(\frac{1}{180}\,\frac{2\pi FS\mu}{R_o 10^9}\right)^4 + \cdots \right. \qquad (1)$$

where r and R_o are respectively the a-c and d-c resistances of the entire length of a cylindrical conductor, S is the length (centimeters) of the conductor, μ is the permeability, and F is the frequency (cy/sec).

Consider a typical carbon resistor which is 2 centimeters long, having 100 ohms d-c resistance, μ taken as unity, and the frequency F taken as 50,000,000. In this case, the theoretical formula (1) results as follows.

$$r = 100\,[1 + .000131 - .000000013 + \cdots \qquad (2)$$

(1) shows only the first few terms of an infinite series. From (2) it is evident that in the case of this 100-ohm carbon resistor



$Z_0 = \sqrt{430 \times 23.8} = 101.2$

Fig. 4

(1) contains enough terms as it stands. Incidentally, this does not mean that the calculation of skin effect is ordinarily quite easy. When dealing with copper or iron at very high frequencies, (1) must be extended to include 10 or 20 terms. In such cases, other formulae for approximation are preferable. When an evaluation, as in (2), is carried out for higher values of d-c resistance, such as several hundred or a few thousand ohms, S being 2 or 3 centimeters, and the frequency being increased to 100 megacycles, it is likewise found that skin effect is entirely negligible in the case of typical carbon resistors.

Another example of measurement is shown by Figure 4, concerning $\triangle'C_3$ and R vs. frequency, at and near full-wave resonance of a certain transmission line having its far end open. Note that the maximum value of R obtains at nearly, but not quite the frequency which results in $\triangle'C_3$ equal to zero.

Figure 5 shows $\triangle'C_3$ as seen by a generator at the center of a half-wave doublet having resonance at 36.3 megacycles. Of course X may be evaluated directly from these data, as per step No. (g). The resistance at resonance was found to be 64 ohms. This value was not expected to agree closely with the natural value of 72 ohms (radiation resistance), because the doublet was not much more than a quarter-wave length away from a steel frame building. Above the frequency

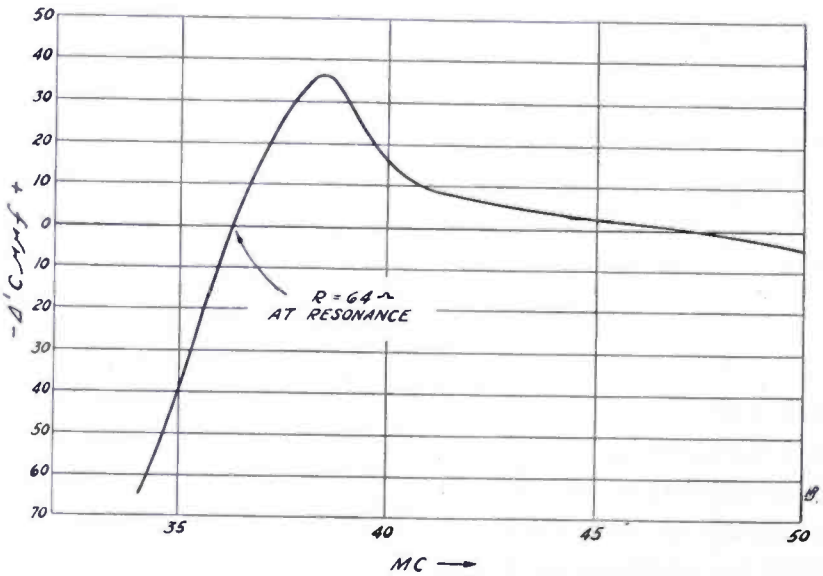


Fig. 5

of half-wave resonance, the $\triangle'C_3$ characteristic does not agree with the idealistic characteristic which would be approximated were the steel frame building not involved.

Figure 6 shows data taken when looking into a transmission line which was terminated by a doublet having approximately 20 per cent less length than the doublet of Figure 5.

These data are given to show that ultra-high-frequency impedances can be measured without making any uncertain assumptions.

4. Analytical Explanation of Method

Note Figure 7, where the shunt impedances Z_x and Z_r are across C_3. Let $\omega L_2 = \dfrac{1}{\omega C_2}$. Disregard the negligible quantity r_2. Let Z_2 be

the impedance seen by a series generator in Circuit 2, when Circuit 1 is isolated therefrom.

$$Z_2 = \frac{Z_x Z_r}{Z_x + Z_r}$$

Upon coupling Circuits 1 and 2, the impedance appearing in Circuit 1 due to Circuit 2 is $\frac{(\omega M)^2}{Z_2}$ or, $\frac{(\omega M)^2}{Z_x} + \frac{(\omega M)^2}{Z_r}$ (3).

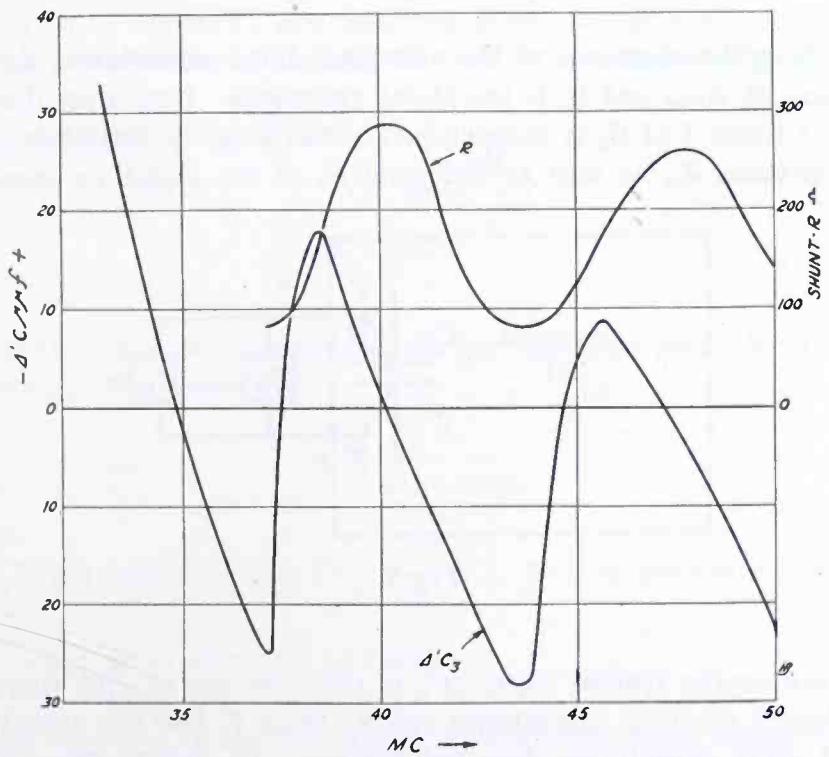


Fig. 6

Note that each shunt impedance has a strictly independent influence on Circuit 1: i.e., change or remove either Z_x or Z_r, and the other still has its former effect on Circuit 1.

Z_x and Z_r may be any types of impedance. Let Z_x be due to the condenser C_3 in Figure 1, and let Z_r be a pure resistance. Then the first term in (3) is a change of reactance seen in Circuit 1, and the second term is a change of resistance seen in Circuit 1. Note particularly that the change in reactance is independent of the change of resistance. Hence a net reactance of zero in Circuit 1 may be obtained by adjusting either C_1 or C_3, without affecting the total resistance in

Circuit 1. Also note that since (3) is general, Z_s and Z_r may be the separate reactances of each of two condensers in shunt. Or, the first term of (3) may be written for two shunt condensers, retaining the second term on account of a shunt resistance: i.e., across T let there be (a) some stray capacitance including the unknown minimum capacitance of C_3, (b) another capacitance $\triangle C_3$, and (c) a resistance. Then the impedance appearing in Circuit 1 due to Circuit 2 is:

$$\frac{(\omega M)^2}{Z_s} + \frac{(\omega M)^2}{Z_c} + \frac{(\omega M)^2}{Z_r} \tag{4}$$

where Z_s is the reactance of the unknown shunt capacitance, Z_c is the reactance of $\triangle C_3$, and Z_r is the shunt resistance. Note that the influence in Circuit 1 of C_3 is independent of the stray or unknown capacitive reactance Z_s, as well as independent of the shunt resistance Z_r.

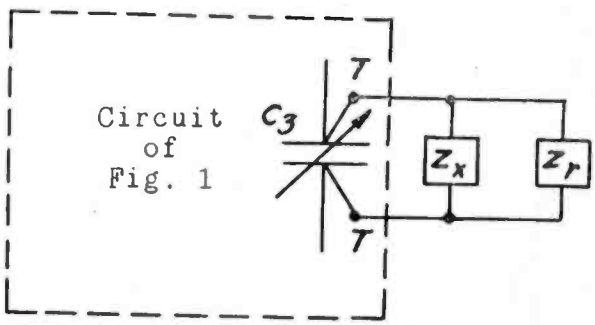


Fig. 7

After tuning the system when $\triangle C_3$ is zero, the use of $\triangle C_3$ therefrom is as though no stray capacitance exists across T, and the actual presence of stray capacitance does not prevent $\triangle C_3$ and Z_r from having distinctly independent effects in Circuit 1.

Tune Circuit 1 when $\triangle C_3$ is zero. Then slightly decrease the frequency of the impressed voltage E (note Figure 1), or decrease C_1, such that the vacuum-tube voltmeter gives a deflection corresponding to 70.7 per cent of the former (resonant) voltage across C_1. Since 70.7 is $1/\sqrt{2} \times 100$, the adjustment to 70.7 per cent of the resonant response means that the resistance r_1 (of Circuit 1) is equal to the net reactance of Circuit 1. Next restore resonance in Circuit 1 by increasing C_3 by an amount $\triangle C_3$. This means that ωM presents $\triangle C_3$ to Circuit 1 as a reactance which is numerically equal to the resistance of Circuit 1, because, before $\triangle C_3$ was introduced, Circuit 1 was detuned such that its net reactance was equal to its resistance. Hence,

$$\frac{(\omega M)^2}{\frac{1}{\omega \triangle C_3}} = r_1 \text{ or, } (\omega M)^2 = \frac{r_1}{\omega \triangle C_3} \tag{5}$$

Since series resonance is established in Circuit 1 both before and after the connection of an unknown impedance across T, and since this unknown impedance, whatever be its physical embodiment, is equivalent to a shunt combination of R and X, and remembering that the adjustment of C_3 to establish resonance is independent of R, it necessarily follows that the first and second voltages (E_1 and E_2 as per Section 3) across C_1 are inversely proportional to the corresponding total resistances in Circuit 1: i.e., r_1 determines E_1, and $r_1 + \frac{(\omega M)^2}{R}$ determines E_2. Hence,

$$\frac{E_1}{E_2} = \frac{r_1 + \frac{(\omega M)^2}{R}}{r_1} \tag{6}$$

In (6), for $(\omega M)^2$ substitute its equivalent as given by (5), and solve for R. The result is,

$$R = \frac{1}{P\,\omega(\triangle C_3)}$$

where P (as defined in step f) is $\frac{E_1}{E_2} - 1$. This is the expression for R which is stated in step (g) of Section 3.

That $\frac{1}{\omega(\triangle' C_3)}$ is the numerical value of the shunt reactance X is self-evident from the derivation and discussion of (3) and (4). When maximizing E_2, the adjustment of C_3 yields zero reactance in Circuit 1 as though R were not involved at all. As stated in Section 3, $\triangle' C_3$ is the change in C_3 from the $\triangle C_3$ setting. Though Section 3 describes the measurement of R and X jointly, it is self-evident that either R or X may be measured without measuring both, upon desire.

5. Note on Transmission Line Measurements

Transmission line considerations are of particular interest in ultra-high-frequency practice. On account of the very high frequency, only 50 feet of cable is effectively a long transmission line. With home or apartment television installations in view, a transmission line from

the out-door doublet to the receiver may be several wave-lengths long. Such a line is electrically as long as several hundred or a few thousand feet of r-f transmission line in the standard broadcast band.

The proper use of a transmission line generally requires that its surge impedance be known, and it is usually desirable to know that the attenuation is acceptably low. Also, it may be desirable to take measurements involving the voltage rise of a resonant length of transmission line. It is thought that the various likely objectives concerning transmission lines at ultra-high frequency, particularly from a measurement standpoint but also as regards performance, warrant this inclusion of certain formulae which are directly applicable to systems having uniformly distributed inductance, capacitance, and resistance.

Whether the transmission line be a pair of uniformly twisted wires, a pair of parallel wires, or a concentric cable, the inductance, capacitance, and resistance are for all practical purposes uniformly distributed. The theory of such a system is found in numerous texts, notably "Calculation of Alternating Current Problems" by Louis Cohen (McGraw-Hill), which is particularly informative.

In standard form, $V = \alpha + j\beta$, where V is the propagation constant of the line. α is the attenuation constant, and β is the velocity constant. "α" is a measure of the transmission loss along the line, due to both series and shunt resistance per unit length of line. The velocity of transmission is $\frac{\omega}{\beta}$, and may be taken as the velocity of light in the case of parallel wires with air dielectric, but may be found as much as 50 per cent less than the velocity of light in the case of other dielectrics (such as rubber, etc.). The general expressions for α and β are seldom (if ever) used in practice. Their simplification first depends upon the nature of the line loss, involving predominance of series resistance, or shunt resistance, etc., but both the measurement and the use of "α" do not depend upon the details of its constituency, and α has only an insignificant effect on β. Hence the altogether acceptable approximation: $\beta = \omega\sqrt{LC}$, where L and C are the inductance and capacitance per unit length of line. There need be no concern over the unit of length, finite or differential, because β need not be involved alone, but rather βS, where S is the entire length of the transmission line. Ultimate formulae need involve simply the total inductance and the total capacitance of the entire length of the line. Likewise, α need not be involved alone, the quantity αS being a measure of the loss for the entire length of the transmission line. The line loss is 8.686 (αS) decibels.

Consider a length S of transmission line, open at the far end (excepting the negligible capacitance of a modern vacuum-tube voltmeter there connected). The voltage at the far end of the line is,

$$\frac{2E_g}{\sqrt{2\cosh(2\alpha S)+2\cos(2\beta S)}} \tag{7}$$

where E_g is a voltage (at ω) impressed at the transmitting end of the line. Suppose that ω be varied. Then the term $2\cos(2\beta S)$ becomes periodic, and ranges between $+2$ and -2. This causes (7) to show that the voltage at the far end of the line is a maximum at a fundamental frequency and at all odd harmonics thereof. The fundamental, or lowest frequency to yield a maximum voltage at the far end of the line, is found by setting βS equal to $\pi/2$ radians. At the next higher frequency yielding a maximum voltage, βS is equal to $3\pi/2$ radians, etc., etc. Were the line loss equal to zero, i.e., no series or shunt resistance along the transmission line, α would be zero, and since cosh (zero) is unity, the term $2\cosh(2\alpha S)$ would equal 2. This would cause the denominator of (7) to be zero at values of ω which result in the expression:

$$2\cos(2\beta S)=-2$$

Thus would the voltage rise of the line be infinity at the fundamental frequency and at all of its odd harmonics. Of course the line loss is never zero, hence the term $2\cosh(2\alpha S)$ is bound to be greater than 2. Ordinarily, it exceeds 2 by a relatively small quantity such as 0.05, or 0.1, or 0.2, such that the resonant rise of the line is usually between 10 and 100.

(7) suggests a method of learning αS. Choose a frequency (or length of line) resulting in a maximum voltage at the far end of the line. Read it, and also read E_g. Then the term $2\cosh(2\alpha S)$ is the only unknown quantity in (7). $2\cosh(2\alpha S)$ may be evaluated directly, and αS then found by reference to a table of hyperbolic functions. In case the transmission line is very efficient, leading to voltage measurement difficulties on account of a high ratio of output to input voltages, a long length of line may be used in order to increase αS to a readily measurable value. In this manner, αS per foot, or per meter of line may be learned.

Seldom is it necessary to deal with (7) in its hyperbolic form. (7) is a strictly rigorous expression. It may be greatly simplified at resonant values of ω by accepting a slight error. The hyperbolic term may be eliminated as follows: Identity:

$$\cosh(2aS) = 1 + \frac{(2aS)^2}{2!} + \frac{(2aS)^4}{4!} + \text{..etc.} \qquad (8)$$

In (8), $1 + \frac{(2aS)^2}{2!}$ is all that need be retained when aS is not greater than 0.35, provided that an error of 1 per cent is tolerable. The error decreases rapidly as aS decreases. Upon substituting the first two terms of (8) into (7), it is found directly that aS is simply the reciprocal of the voltage rise at resonance. This simple method of close approximation seems justified by the likely errors of observation at ultra-high frequency.

When employing the foregoing method of determining aS, a signal generator should not be relied upon in the matter of impressed voltage magnitude at ultra-high frequency. This precaution is not entirely because the signal generator may be in error. It is notably because of the output leads in the signal generator. At very high and ultra-high frequency, many signal generators should be regarded as having a short transmission line within. Therefore, the resonant rise method of learning aS should disregard the signal-generator circuit properties by placing a vacuum-tube voltmeter at the actual input end of the transmission line under test, and at the far end of the line. Then vary the frequency in the desired region, until the maximum ratio of output to input voltage is found. aS is the reciprocal of that ratio, at the frequency involved, and signal-generator output wiring effects are thus eliminated.

Ordinarily, it is desirable to know or to consider the impedance seen when looking into a transmission line with its far end terminated in a prescribed manner. Designating this input impedance by z,

$$z = \frac{Z_o \sinh(VS) + Z\cosh(VS)}{\cosh(VS) + \frac{Z}{Z_o}\sinh(VS)} \qquad (9)$$

where Z is the impedance (which may be complex or not) looked into by the far end of the line, and Z_o is the surge impedance of the line itself. Z_o is ordinarily defined and used as $\sqrt{\frac{L}{C}}$, where L and C are the inductance and capacitance per unit length of line, or of the entire length of line. Rigorously, Z_o is much less simple than that, and also it is a complex quantity. Invariably, however, this simple definition of Z_o is easily a sufficiently close approximation.

Since (9) is general, let Z (the impedance connected to the far end of the line) be equal to the surge impedance Z_0. Then z, the impedance looking into the line, becomes simply Z_0, all other terms vanishing. Note that this is true not merely for resonant frequencies of the line, but that it holds for any frequency when Z is equal to Z_0, and the transmission line is said to be critically damped. In this case, the transmission line has the effect of moving a physically remote impedance to the actual position of the input terminals of the line. Of course this is a steady-state condition. Since the input impedance is equal to the terminating impedance, the transmission line may be said to act like a 1:1 transformer as regards impedance ratio. (This statement has its limitations: e.g., it does not deal with phase relations. The voltage across the terminating impedance Z lags the voltage impressed on the line, due to the finite velocity of transmission.)

From (9), it may be seen directly that a transmission line may be made to serve as an impedance-matching transformer, although in this case the system does not act like a transformer at all frequencies, but only at the frequency of quarter-wave resonance, three-quarter wave resonance, etc., and a low-loss line should be used. This impedance-transformation property is noted from (9) without confusion by letting "α" be zero, hence V is simply $j\beta S$, and then let βS be $\pi/2$ (or $3\pi/2$, etc.) radians. The direct result is,

$$z = \frac{Z_0^2}{Z}$$

wherein Z_0^2 is analogous to (turns ratio)^2, or $(\omega M)^2$ in transformer terminology.

It may be desirable to learn the value of Z_0 at high or ultra-high frequency. Z_0 is the geometric mean of the input impedances with the far end of the line open and with it shorted, as seen directly from the general expression (9) for z: i.e., divide numerator and denominator by Z_0, and then let Z be zero. Thus, with the far end shorted, z has the value $Z_0 \tanh (VS)$. Returning to (9), divide numerator and denominator by Z, then let Z be infinity. Thus, with the far end of the line open, z has the value $Z_0 \coth (VS)$. Then upon taking the square root of the product of these two values of z, the result is simply Z_0. In order to keep free of complex quantities, these two measurements of the input impedance z may be taken at a frequency of resonance. A suitably long test line should be used in order to prevent the high value of z from being too high, or the low value of z from being too low for accurate measurement. Z_0 measured at very high or ultra-high frequency is not likely to differ greatly, but may differ somewhat, from the low (audio) frequency determination of Z_0.

When measuring Z_o at ultra-high frequency, the method described in Section 3 leads to data as exemplified by Figures 5, 6, or by Figure 4 in particular where the $\triangle' C_3$ characteristic reveals resonance at 43.9 megacycles. At this frequency, the input impedance (a resistance) is 430 ohms when the far end of the line is open, and 23.8 ohms when the far end is shorted. Z_o, the geometric mean of these quantities, is 101 ohms. A longer test line would have been used in case the measured value of Z_o were for important use. Precision of measurement is enhanced by employing a length of test line to result in not less than (say) 50 ohms as the lower value of input impedance.

Z_o, aS, and z are involved in simple relationship at quarter-wave resonance, three-quarter wave resonance, etc., the far end of the line being open or shorted as prescribed. This relationship is here derived with enough detail to exemplify the manner in which many transmission formulae may be dealt with.

Again taking the general expression (9) as a basis, divide numerator and denominator by Z_o. Then let Z be zero, meaning that the far end of the line is shorted. The immediate result is (10).

$$z = Z_o \frac{\sinh (VS)}{\cosh (VS)} \tag{10}$$

Since $V = a + j\beta$, $VS = aS + j\beta S$.

Identities:

$$\left.\begin{aligned} \sinh (aS + j\beta S) &= \sinh (aS) \cdot \cos (\beta S) \\ &\quad + j\cosh (aS) \cdot \sin (\beta S) \\ \cosh (aS + j\beta S) &= \cosh (aS) \cdot \cos (\beta S) \\ &\quad + j\sinh (aS) \cdot \sin (\beta S) \end{aligned}\right\} \tag{11}$$

Substitute these identities in (10), and in so doing substitute zero for $\cos (\beta S)$ and ± 1 for $\sin (\beta S)$, because (βS) is $\pi/2$, or $3\pi/2$, etc., i.e., it has been prescribed that the line is quarter-wave resonant, or three-quarter wave resonant, etc. Directly after these substitutions in (9), (12) results.

$$z = Z_o \frac{\cosh (AS)}{\sinh (AS)}, \text{ or, } Z_o \coth (aS) \tag{12}$$

hence,

$$\tanh (aS) = \frac{Z_o}{z} \tag{13}$$

Were the far end of the line open instead of closed, the numerator and denominator of (9) would have been divided by Z instead of by Z_o,

then Z would have been set equal to infinity instead of equal to zero, and the immediate result would have been $Z_0 \dfrac{\cosh (VS)}{\sinh (VS)}$ instead of (10). Therefrom, the same substitutions would have been made, and, instead of (12), the result would have been $Z_0 \tanh (aS)$ or

$$\tanh (aS) = \frac{z}{Z_0} \tag{14}$$

At resonance as prescribed, z in (13) and (14) is a pure resistance. (13) and (14) relate (aS), the surge impedance Z_0, and the input resistance at resonance. Choice between (13) and (14) depends upon ease of accuracy of measurement.

As done in the case of (7), (13) and (14) may be freed of their hyperbolic form without involving more than a 1 per cent error when aS is not greater than 0.15. The simplification is effected by merely noting that the hyperbolic tangent of a quantity is nearly equal to that quantity, when the quantity is sufficiently small. The discrepancy is 1 per cent when the quantity is 0.15. (A greater value of aS, .35, is allowable in the similar simplification of (7) when 1 per cent error is tolerable.) Hence, for small values of aS, (15) obtains,

$$\begin{aligned} &\text{Far end of line shorted} \quad (aS) = \frac{Z_0}{z} \\ &\text{Far end of line open} \quad\quad (aS) = \frac{z}{Z_0} \end{aligned} \tag{15}$$

Since the greater value of z obtains when the far end of the line is shorted, this condition is preferable when using the method of measurement described in Section 3.

The expressions (13), (14), and (15) are quite simple, but peculiar circumstances sometimes lead to a choice of less intermediary simplicity in favor of over-all convenience. By a procedure similar to the derivation of (13) and (14), alternative (though less simple) expressions are straightforwardly formed: e.g., with the far end of the line shorted, and at quarter-wave resonance or at any odd multiple thereof

$$(aS) = \frac{\cosh^{-1} \dfrac{A+1}{A-1}}{2} \quad \text{where } A = \left(\frac{z}{Z_0} \right)^2$$

This expression is rigorous for all values of (aS).

6. Regarding C_3

Referring to Figure 1 and to Sections 3 and 4, the variable air condenser C_3 is the only circuit element which is used as a standard. It should be a physically small condenser of the very low-loss type, and also should embody mechanical excellence. $\triangle C_3$ is the quantity involved in the evaluations of R and X, hence C_3 is calibrated on an incremental basis. As regards the maximum capacitance of C_3, at ultra-high frequency $\triangle C_3$ may be only 1 or a few $\mu\mu$fd when measuring a high resistance such as 5 or 10 thousand ohms, whereas for lower measurable values of resistance $\triangle C_3$ may be from 15 to 30 $\mu\mu$fd. Other matters affect the choice of C_3 as regards its maximum capacitance. As described in Section 3, a precise adjustment of C_3 requires a setting on each side of resonance in order to accurately determine $\triangle C_3$, thus increasing the necessary maximum capacitance of C_3. Also, the measurement of reactance may, on occasion, require that $\triangle' C_3$ be 30 or 40 $\mu\mu$fd. Due to the likely observational error when reading only a few micromicrofarads on a scale which extends to 40 or 50 $\mu\mu$fd, two conveniently interchangeable condensers may be used: e.g., a 15 $\mu\mu$fd condenser and a 50 $\mu\mu$fd condenser; or, a single condenser having a maximum capacitance of 40 or 50 $\mu\mu$fds may be used when it has shaped plates in order to expand the scale in the region of a few micromicrofarads. Since $\triangle C_3$ is the quantity involved, a properly fitted reduction gear drive may be used advantageously by placing a graduated dial on the shaft of the smaller gear.

C_3 is calibrated at audio frequency. A suitable bridge is employed. C_3 is connected directly in shunt to a precision condenser, hence the calibration of C_3 is a matter of direct substitution. A preferred type of precision condenser for this use as a comparator is described below.

The capacitance formed by two concentric cylindrical surfaces with air dielectric is

$$\frac{.6128}{\log_{10}\left(\frac{D_1}{D_2}\right)} \mu\mu\text{fd}$$

per inch of axial length, where D_1 and D_2 are respectively the larger and the smaller diameters of the concentric cylinders. Accordingly, a diameter ratio of 2.035 results in 2 $\mu\mu$fds per inch of length. This affords a convenient method of making an accurate variable condenser whose capacitance need not be more than 20 $\mu\mu$fds, in which case the working length of the cylinders would not be more than 10 inches. As regards actual diameters, the larger circle may be the accurately

measured inside diameter of a brass tube having an inside diameter of approximately 1 inch, and the smaller circle may be the diameter of a brass rod which is placed in a lathe and machined to a tolerance of 0.0005". By means of bakelite bearings, this rod slides co-axially through the outer tube, causing a capacitance change of 2 $\mu\mu$fds per inch of motion when the diameter ratio is chosen for that result, and the rod is graduated simply in half inches and subdivisions. Of course end effects must be considered in a precision condenser of this type. The variation of stray capacitance due to end effects is minimized to a definitely negligible quantity when the diameters are as noted above, by providing an end overlap of not less than 10 times the smaller diameter: i.e., the rod is graduated only for a range of movement wherein neither end of the rod comes within 5 inches of alignment with either end of the stationary tube.

A bakelite grip is placed on one end of the movable rod, and a cylindrical copper shield is placed outside of the stationary tube and connected to the brass rod. A precision condenser of this type is immediately useful in calibrating C_3 at audio frequency. A tubular condenser, as described, is too long to be regarded as a capacitance which is independent of frequency at ultra-high frequency, and hence should not be used as C_3 in the measurement method outlined in this paper.

A SURVEY OF ULTRA-HIGH-FREQUENCY MEASUREMENTS

By

L. S. Nergaard

Research and Engineering Department, RCA Manufacturing Company, Harrison, N. J.

Summary—*A simple magnetron signal generator is described. The more useful transmission-line and skin-effect formulas are listed. In connection with transmission lines, it is pointed out that even at very high frequencies the quadrature component of the characteristic impedance cannot be neglected. Methods which have been used for the measurement of the following quantities are described:*

(1) Wavelength—Wavelength has been measured by reflection of waves in free-space and transmission-line wavemeters. A method of determining the end correction of a transmission-line wavemeter is described.

(2) Power—Thermocouples for the measurement of small powers are described. The use of incandescent lamps for the measurement of large powers is discussed.

(3) Voltage—Diode voltmeters and thermocouples have been used for the measurement of voltage. The errors of these voltmeters are discussed.

(4) Reactance—Reactance has been measured by tuning the unknown reactance to resonance with a transmission line of known characteristics. The method is illustrated by two examples: the determination of the resonant wavelength, the inductance and capacitance of a diode, and the calibration of a variable condenser.

(5) Resistance—Resistance has been measured by the substitution method, the resistance-variation method, and the reactance-variation method. The reactance-variation method has been used in all of its forms, namely, capacitance variation, line-length variation, and frequency variation. The pertinent formulas and examples of the methods are given. A transmission line and condenser for the measurement of resistance by the capacitance-variation method are described.

(6) Current—The measurement of current with thermocouples is discussed.

I. Introduction

THE last five years have brought ever-increasing activity in the ultra-short-wave field and a gradual shift of the work on the more conventional tubes and circuits from a research to an engineering basis. This shift is, in a large measure, due to the development of a measurement technique for the ultra-high-frequency region. This development is by no means complete. Much remains to

Reprinted from *RCA Review*, October, 1938.

be done, both in improving the precision of the measurements and in extending the frequency range over which reliable measurements can be made. However, it is sometimes advantageous to pause and take stock of the methods available. It was with this thought in mind that the present paper was undertaken.

This paper is a survey of those methods of measurements which have proved most useful in this laboratory. The writer is well aware that, in limiting himself to a report of the methods used in a single laboratory, he is excluding the very valuable work done elsewhere. His apology for this exclusion is that it permits him to describe only those methods of which he has first-hand knowledge, i.e., the methods whose value he is best able to appraise.

II. Signal Generators

A good source of radio-frequency power is essential to radio-frequency measurements. The requirements of a signal generator are much the same at any frequency and include (1) adequate output, (2) a wide frequency range, (3) freedom from harmonics, (4) frequency and voltage stability, and (5) good shielding. It is also desirable that the generator be flexible so that the frequency and output can be varied widely with ease.

Two types of oscillators are commonly used, namely: (1) the negative-grid triode and (2) the magnetron. Triode oscillators can be made to operate down to a wavelength of 20 cm with an output of 2 watts.[1] For wavelengths above 80 cm, the RCA-955 and the RCA-834 have been used in this laboratory. Below 80-cm wavelength, the feedback becomes critical of adjustment, and it is usually necessary to tune the filament circuit. Therefore, a triode oscillator for wavelengths less than 80 cm requires a minimum of two circuit adjustments to change wavelength and becomes less flexible than is desirable.

Magnetrons operating in the negative-resistance mode will deliver as much as 80 watts at a wavelength of 19 cm.[2] This output was obtained from a water-cooled internal-circuit tube. Because of the fixed internal circuit, such a tube lacks the flexibility desirable in a signal generator. Messrs. G. R. Kilgore and T. H. Clark of this laboratory have designed a small split-anode magnetron having an anode diameter of 2.5 mm and very short anode leads.* This tube delivers 1 watt at a wavelength of 40 cm with 300 volts on the anodes and a field

* With the exception of the RCA-955 and the RCA-834, all of the tubes and thermocouples described in this paper have been constructed for measurements incidental to our short-wave research.

strength of 1500 gausses. The wavelength can be varied from 40 cm to 200 cm by merely varying the length of the anode transmission line. The output is readily controlled by varying the filament current. Three signal generators employing these tubes are now in use in this laboratory and are proving very usable and useful. The completely shielded signal generator described in the writer's previous measurement paper[11] now uses this type of tube. For many rough measurements the oscillator need not be shielded. Figure 1 shows a simple signal generator of the unshielded type.

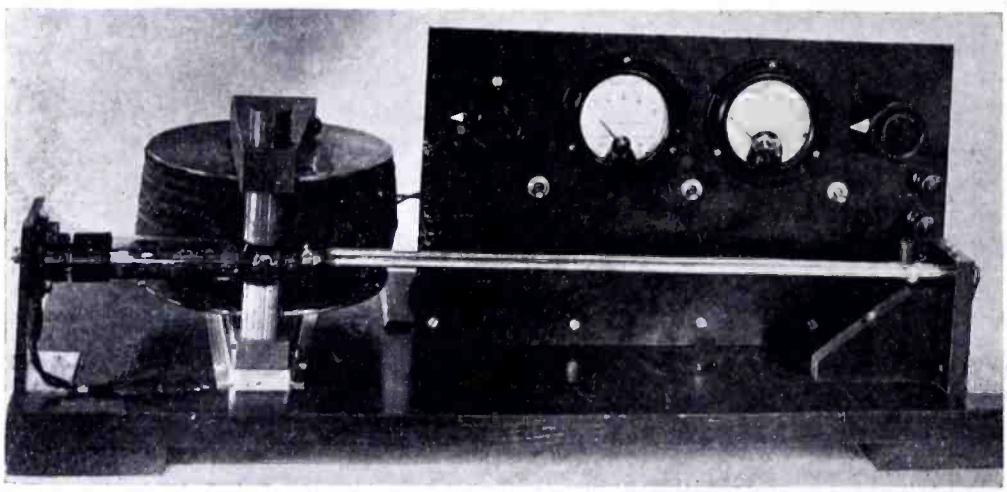


Fig. 1—An unshielded magnetron signal generator.

III. Circuits

At short wavelengths, transmission lines have almost entirely supplanted the lumped circuit elements used at long wavelengths. This change has taken place for two reasons; namely, lumped circuit elements for short wavelengths are difficult to construct because of their small physical size, and are also difficult to adjust once they have been constructed.

The following example illustrates the difference in physical size of a lumped circuit and a transmission line, both to tune to a wavelength of 50 cm. Suppose the condenser of the lumped circuit consists of coaxial circular plates 2 cm in diameter and spaced 0.1 cm. apart. The capacitance of the condenser is approximately 2.8 $\mu\mu f$ and the required inductance is approximately $1/\omega^2 C$, or 2.5×10^{-8} henry. A circular loop of No. 14 wire, 2 cm in diameter, has approximately the required inductance. For a transmission line open at one end and short-circuited at the other end, the length of the line for resonance

is 12.5 cm. It is interesting to note that the current in the lumped inductance is uniform within 10 per cent, whereas the current in the transmission line is zero at the open end and a maximum at the short-circuited end.

Because of the great utility of transmission lines, a few of the more important formulas† will be listed in the forms in which they have been found most useful. Let

l = length of line in cm
r_o = total resistance per unit length of line in ohm · cm^{-1}
L_o = inductance per unit length of line in henrys · cm^{-1}
C_o = capacitance per unit length of line in farads · cm^{-1}
G_o = leakage conductance per unit length of line in mhos · cm^{-1}
α = attenuation constant in cm^{-1}
β = phase constant in cm^{-1}
m = propagation constant in $cm^{-1} = \alpha + j\beta$
γ = resistive component of the characteristic impedance in ohms
δ = reactive component of the characteristic impedance in ohms
Z_o = characteristic impedance of the line in ohms $= \gamma - j\delta$
$j = \sqrt{-1}$
$\omega = 2\pi f$
f = frequency in cycles per sec
λ = wavelength in cm
E_s = voltage across sending end of line in volts
E_r = voltage across receiving end of line in volts
I_s = current at sending end of line in amperes
I_r = current at receiving end of line in amperes
Z_r = impedance across the receiving end in ohms
Z_s = impedance looking into sending end of line in ohms

The properties of a transmission line are most conveniently expressed in terms of the characteristic impedance, the propagation constant, and the length of the line. Because the expressions relating the characteristic impedance and the propagation constant to the resistance, inductance, conductance, and capacitance per unit length of line are unwieldy, it is advantageous to have approximate expressions which are simple, yet adequately accurate. The exact expressions for the real and imaginary parts of the propagation constant and characteristic impedance are:

$$\alpha = \omega\sqrt{L_oC_o}\cdot\left\{\frac{1}{2}\left[1+\left(\frac{r_o}{\omega L_o}\right)^2\right]^{\frac{1}{2}}\cdot\left[1+\left(\frac{G_o}{\omega C_o}\right)^2\right]^{\frac{1}{2}}\right.$$

(Formula continued on next page)

† A consistent notation is used throughout the paper. For convenience, a table of the symbols used is given in Appendix A.

$$-\frac{1}{2}\left[1-\frac{G_o r_o}{\omega^2 L_o C_o}\right]\Bigg\}^{\frac{1}{2}}$$

$$\beta=\omega\sqrt{L_oC_o}\cdot\left\{\frac{1}{2}\left[1+\left(\frac{r_o}{\omega L_o}\right)^2\right]^{\frac{1}{2}}\cdot\left[1+\left(\frac{G_o}{\omega C_o}\right)^2\right]^{\frac{1}{2}}+\frac{1}{2}\left[1-\frac{G_o r_o}{\omega^2 L_o C_o}\right]\right\}^{\frac{1}{2}}$$

$$\gamma=\sqrt{\frac{L_o}{C_o}}\cdot\left[1+\left(\frac{G_o}{\omega C_o}\right)^2\right]^{-\frac{1}{2}}\cdot\left\{\frac{1}{2}\left[1+\left(\frac{r_o}{\omega L_o}\right)^2\right]^{\frac{1}{2}}\cdot\left[1+\left(\frac{G_o}{\omega C_o}\right)^2\right]^{\frac{1}{2}}+\frac{1}{2}\left[1+\frac{G_o r_o}{\omega^2 L_o C_o}\right]\right\}^{\frac{1}{2}}$$

$$\delta=\sqrt{\frac{L_o}{C_o}}\cdot\left[1+\left(\frac{G_o}{\omega C_o}\right)^2\right]^{-\frac{1}{2}}\cdot\left\{\frac{1}{2}\left[1+\left(\frac{r_o}{\omega L_o}\right)^2\right]^{\frac{1}{2}}\cdot\left[1+\left(\frac{G_o}{\omega C_o}\right)^2\right]^{\frac{1}{2}}-\frac{1}{2}\left[1+\frac{G_o r_o}{\omega^2 L_o C_o}\right]\right\}^{\frac{1}{2}}$$

In practical short-wave transmission lines, $G_o=0$ and $r_o/\omega L_o << 1$. Therefore, the above relations can be written in the following simple forms which are sufficiently accurate for all practical short-wave transmission lines.

$$\alpha=\frac{r_o}{2}\sqrt{\frac{C_o}{L_o}}=\frac{r_o}{2\omega L_o}\beta,\quad \beta=\omega\sqrt{L_oC_o}=\frac{2\pi}{\lambda},\quad \gamma=\sqrt{\frac{L_o}{C_o}}$$

and

$$\delta=\frac{r_o}{2\omega L_o}\sqrt{\frac{L_o}{C_o}}=\frac{r_o}{2\omega L_o}\gamma \qquad \text{(III-1)}$$

Hence,

$$m=\left(\frac{r_o}{2\omega L_o}+j\right)\cdot\frac{2\pi}{\lambda}$$

$$Z_o = \left(1 - j\frac{r_o}{2\omega L_o}\right)\sqrt{\frac{L_o}{C_o}} \qquad \text{(III-2)}$$

It apparently has been customary to neglect the quadrature component of the characteristic impedance in short-wave transmission-line calculations. This procedure may lead to 50 per cent error in calculating the input impedance of a transmission line with the receiving end short-circuited. That this statement is true is readily verified for the case when the length approaches zero.

The following relations have been found useful:

$$Z_s = Z_o \frac{Z_r \cosh ml + Z_o \sinh ml}{Z_r \sinh ml + Z_o \cosh ml} \qquad \text{(III-3)}$$

$$\frac{I_r}{I_s} = \frac{Z_o}{Z_r \sinh ml + Z_o \cosh ml} \qquad \text{(III-4)}$$

$$\frac{E_r}{E_s} = \frac{I_r Z_r}{E_s} = \frac{Z_o}{Z_r \cosh ml + Z_o \sinh ml} \qquad \text{(III-5)}$$

Two special cases are of particular interest:

a) If the receiving end of the line is short-circuited ($Z_r = 0$), the sending-end impedance is $Z_o \tanh ml$

or

$$Z_s = \frac{\frac{r_o l}{2}\left[1 + \frac{\sin\frac{4\pi l}{\lambda}}{\frac{4\pi l}{\lambda}}\right] + j\sqrt{\frac{L_o}{C_o}}\sin\frac{2\pi l}{\lambda}\cdot\cos\frac{2\pi l}{\lambda}}{\cos^2\frac{2\pi l}{\lambda} + \left(\frac{r_o l}{2}\sqrt{\frac{C_o}{L_o}}\right)^2 \sin^2\frac{2\pi l}{\lambda}} \qquad \text{(III-6)}$$

When $l = \frac{\lambda}{2}$, then $Z_s = \frac{r_o \lambda}{4}$ (III-7)

and the circuit acts as a resonant circuit.

When $l = \frac{\lambda}{4}$, then $Z_s = \frac{8L_o}{r_o \lambda C_o}$ (III-8)

and the circuit acts as an anti-resonant circuit.

When

$$\cos\frac{2\pi l}{\lambda} >> \frac{r_o l}{2}\sqrt{\frac{C_o}{L_o}}\sin\frac{2\pi l}{\lambda}, \quad \text{then}$$

$$Z_s = \frac{r_o l}{2}\left[1 + \frac{\sin\dfrac{4\pi l}{\lambda}}{\dfrac{4\pi l}{\lambda}}\right]\sec^2\frac{2\pi l}{\lambda} + j\sqrt{\frac{L_o}{C_o}}\tan\frac{2\pi l}{\lambda} \qquad \text{(III-9)}$$

The imaginary term here is much greater than the real, so that the circuit acts as an inductance with a small loss. In the limiting case when $l \longrightarrow 0$

$$Z_s \longrightarrow (r_o + j\,\omega L_o)\,l \qquad \text{(III-10)}$$

b) If the receiving end of the line is open ($Z_r = \infty$), the sending-end impedance is $Z_o \coth ml$

or

$$Z_s = \frac{\dfrac{r_o l}{2}\left[1 - \dfrac{\sin\dfrac{4\pi l}{\lambda}}{\dfrac{4\pi l}{\lambda}}\right] - j\sqrt{\dfrac{L_o}{C_o}}\sin\dfrac{2\pi l}{\lambda}\cdot\cos\dfrac{2\pi l}{\lambda}}{\sin^2\dfrac{2\pi l}{\lambda} + \left(\dfrac{r_o l}{2}\sqrt{\dfrac{C_o}{L_o}}\right)^2\cos^2\dfrac{2\pi l}{\lambda}} \qquad \text{(III-11)}$$

When $l = \dfrac{\lambda}{2}$, then $Z_s = \dfrac{4L_o}{r_o\lambda C_o}$ (III-12)

and the circuit acts as an anti-resonant circuit.

When $l = \dfrac{\lambda}{4}$, then $Z_s = \dfrac{r_o\lambda}{8}$ (III-13)

and the circuit acts as a resonant circuit.

When $\sin\dfrac{2\pi l}{\lambda} >> \dfrac{r_o l}{2}\sqrt{\dfrac{C_o}{L_o}}\cos\dfrac{2\pi l}{\lambda}$ then

$$Z_s = \frac{r_o l}{2}\left[1 - \frac{\sin\frac{4\pi l}{\lambda}}{\frac{4\pi l}{\lambda}}\right]\csc^2\frac{2\pi l}{\lambda} - j\sqrt{\frac{L_o}{C_o}}\cot\frac{2\pi l}{\lambda} \qquad \text{(III-14)}$$

and the circuit acts as an imperfect condenser. In the limiting case

when $l \longrightarrow 0$, then $Z_s \longrightarrow \frac{1}{j\omega C_o l}$ (III-15)

Therefore, transmission lines can be used as resonant and anti-resonant circuits, and as inductances and capacitances.

In the lumped circuits used at long wavelengths, the circuit losses are chiefly in the nature of ohmic losses. At short wavelengths where transmission lines are used as circuit elements and the spacings between conductors of the lines are an appreciable fraction of a wavelength, radiation resistance becomes appreciable. In some short-wave transmission lines, the radiation resistance amounts to half of the total resistance. Furthermore, skin-effect, which results in minor corrections to the ohmic resistance at long wavelengths, becomes an important factor in resistances at short wavelengths.

The resistance per unit length of a solid cylindrical conductor is[4, 5, 6]

$$R_o = R_e\frac{P\rho}{2\pi a} \times 10^{-6}\frac{J_o(Pa)}{J_1(Pa)} \qquad \text{(III-16)}$$

in which

R_o = resistance in ohms per cm length
R_e = "real part of"
J_n = Bessel function of order n
a = radius of conductor in cm
ρ = resistivity of the conductor in microhm · cm

$$P = 2\pi\sqrt{\frac{\mu f}{\rho \times 10^3}}(1 - j)$$

μ = magnetic permeability
f = frequency in cycles per second
$j = \sqrt{-1}$

Conductors at short wavelengths generally fall into one of two classes:

(1) Those for which the skin effect is small. These generally have very high resistivities or very small diameters. Thermocouple heaters are of this class. In this case, Formula (III-16) reduces to

$$R_o = \frac{\rho \times 10^{-6}}{\pi a^2} \left\{ 1 + \frac{1}{3} \left[\pi a \sqrt{\frac{\mu f}{\rho \times 10^3}} \right]^4 \right\} \qquad \text{(III-17)}$$

which is valid when

$$a \sqrt{\frac{\mu f}{\rho}} < 10$$

(2) Those for which the skin effect is very large. This class includes all of the good conductors normally used in transmission lines. In this case, Formula (III-16) becomes

$$R_o = \frac{3.16 \times 10^{-8}}{a} \sqrt{\rho \mu f} \qquad \text{(III-18)}$$

which is valid when

$$a \sqrt{\frac{\mu f}{\rho}} \geqslant 180$$

Because of the skin effect at short wavelengths, the current in a good conductor flows almost entirely on the surface of the conductor. Therefore, the internal inductance becomes negligible compared with the external inductance. This leads to a very simple and useful formula for the real part of the characteristic impedance of a transmission line. When the internal inductance is negligible and $r_o/\omega L_o << 1$, the phase velocity

$$v = \frac{1}{\sqrt{L_o C_o}} \qquad \text{(III-19)}$$

is equal to the velocity of propagation of light in free space for all practical purposes. Therefore,

$$\sqrt{\frac{L_o}{C_o}} = \frac{1}{vC_o} = \frac{10^{-10}}{3C_o} \qquad \text{(III-20)}$$

Hence, the real part of the characteristic impedance can be calculated from the capacitance per unit length of line.

The two types of transmission lines commonly used are the parallel-wire line and the concentric transmission line. The concentric line has the lower attenuation of the two and, therefore, gives the higher impedance when used as an anti-resonant circuit. Its lower attenuation is due to its negligible radiation resistance and to the relatively low resistance of the outer conductor. Whether or not a concentric line has any radiation resistance at all except that due to an open end is a controversial point.[7,8,9] However, the concentric lines used in this laboratory have always given higher anti-resonant impedances than equivalent parallel-wire lines, and no radiation from a concentric line with closed ends has ever been detected. In some cases it is necessary to use parallel-wire lines. For example, the split-anode magnetron requires such a line. Push-pull devices in general fit naturally into lines which are balanced with respect to ground. When it has been desirable to shield a parallel line, the line has been enclosed in a metallic cylinder. The cylinder reduces the radiation resistance, but increases the ohmic resistance because of circulating currents in the cylinder.

IV. Measurement of Wavelength

Wavelength is usually measured by reflection of the waves in free space or by a transmission line. To measure wavelength by reflection in free space an indicator such as a crystal detector and microammeter in series is set up in the vicinity of the source. Then as a plane metallic reflector is moved in a direction normal to its surface and roughly in the direction of the indicator, the positions of the reflector for maximum readings of the microammeter are noted. The distance between positions for successive maxima is a half wavelength.

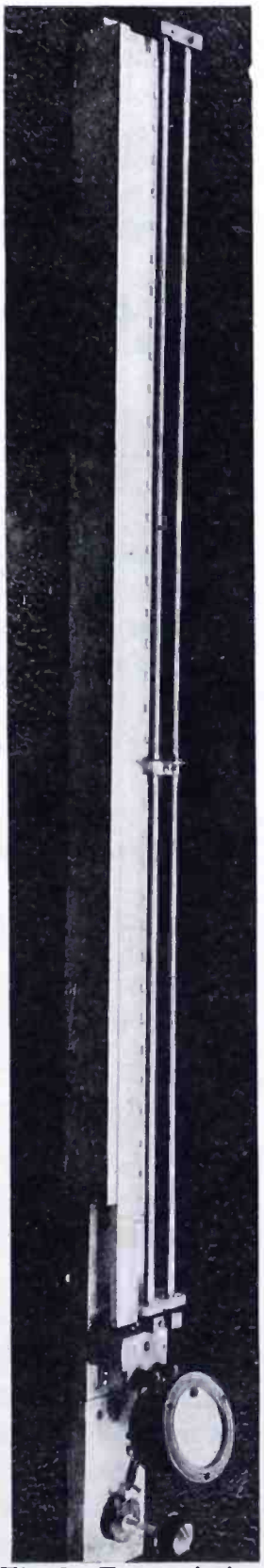


Fig. 2—Transmission-line wavemeter.

The measurement of wavelength with a transmission line consists in measuring the distance between the positions of a movable short-circuit on the line for successive maximum readings of an indicator loosely coupled to the line. The wavelength is then[10]

$$\lambda = 2\triangle l\left[1+\left(\frac{r_o}{2\omega L_o}\right)^2\right] \qquad \text{(IV-1)}$$

where $\triangle l$ = distance between positions for successive maxima

$$\text{Usually } \left(\frac{r_o}{2\omega L_o}\right) << 1$$

so that

$$\lambda = 2\triangle l \qquad \text{(IV-2)}$$

Frequently, it is desirable to keep the length of the line to be used as a wavemeter as short as possible. Then, ideally, the line length is just a quarter wavelength. In practice. the short-circuit on the line has a small inductance L, so that the condition for resonance is

$$\omega L - \sqrt{\frac{L_o}{C_o}}\cot\frac{2\pi l_o}{\lambda} = 0 \qquad \text{(IV-3)}$$

Because L is small, it is permissible to write

$$\omega L\sqrt{\frac{C_o}{L_o}} = \tan\omega L\sqrt{\frac{C_o}{L_o}}$$

in which case (IV-3) becomes

$$\cot\left(\frac{2\pi l_o}{\lambda}+\omega L\sqrt{\frac{C_o}{L_o}}\right) = \cot\frac{2\pi}{\lambda}\left(l_o+\frac{L}{L_o}\right) = 0 \qquad \text{(IV-4)}$$

Then the wavelength is

$$\lambda = 4\left(l_o+\frac{L}{L_o}\right) \qquad \text{(IV-5)}$$

Figure 2 is a photograph of a wavemeter using this principle. A crystal detector and microammeter in series, capacitively coupled to the open end of the line, are used as an indicator. Figure 3 is the calibration curve of this wavemeter. Notice the correction resulting from the inductance of the short-circuit. The inductance in this case is $L = 8.8\times10^{-9}$ henry.

V. Measurement of Power

In general, the requirements of a device for measuring power at short wavelengths are:

(1) It should have adequate sensitivity.

(2) Its resistance should be high so that substantially all the power is dissipated in the device and a negligible amount in other circuit resistances.

(3) It should be free from stray capacitances and inductances which make it difficult to match the impedance of the device to the impedance of the source.

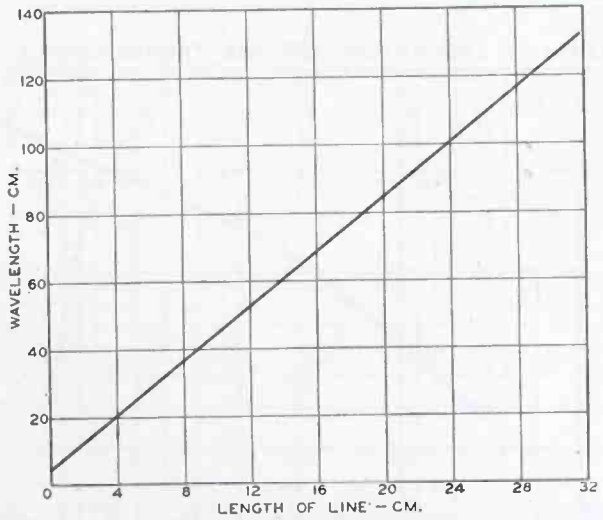


Fig. 3—Calibration of the transmission-line wavemeter.

(4) The calibration should be independent of frequency so that calibration at a low frequency, say 60 cycles per second, is possible.

For powers less than one watt the vacuum thermocouple has been found to be the most satisfactory power-measuring device. Thermocouples for this purpose have been built.[11] The salient features of these thermocouples are:

(1) The heaters are short and straight so that the current is uniform along the length of the heater within 1 per cent for wavelengths above $\lambda = 18$ cm.

(2) The heater resistances are large compared to the circuit resistances likely to be encountered in practical circuits. The resistances range between 5 and 1000 ohms depending on the type of heater used.

(3) The couple and its leads are arranged in an acute V perpendicular to the axis of the heater to minimize coupling between the heater and couple circuits.

(4) The couple and heater are supported by the leads through the radial seal of the acorn-type bulb. No glass beads or spacers are used to support the elements. The most troublesome shunting capacitance is thereby eliminated. The thermocouples having a heater resistance of 1000 ohms described in an earlier paper[11] had glass beads to support the elements. These couples have now been made without beads.

(5) These thermocouples have been made with sufficient sensitivity to measure 0.1 milliwatt.

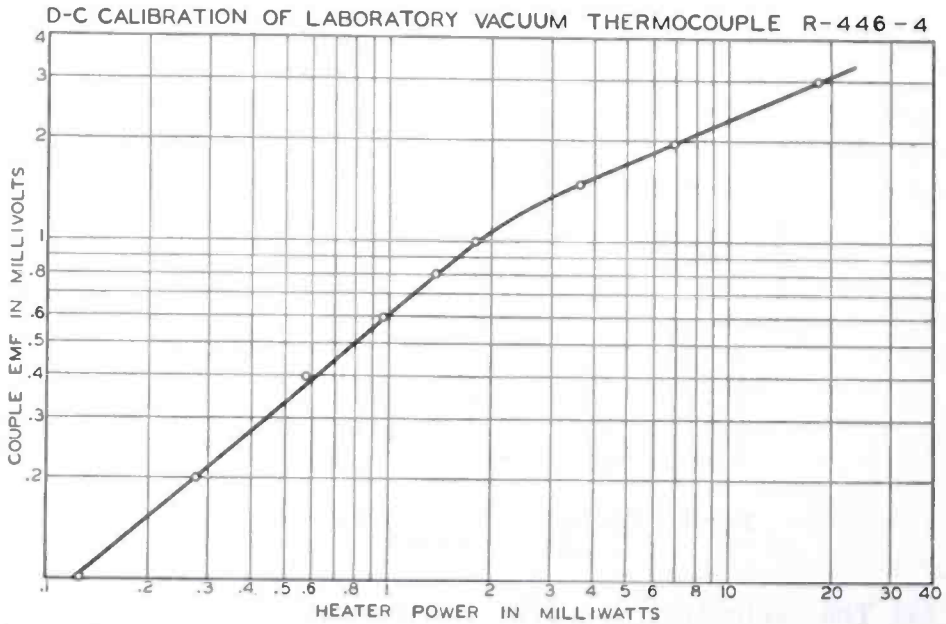


Fig. 4—D-C calibration of laboratory vacuum thermocouple R-446-4. The couple e.m.f. was read on a meter of 10 ohms resistance.

Figure 4 is the calibration curve of one of the more sensitive thermocouples. This thermocouple has a carbon heater of 0.0003-inch diameter and an iron and gold-palladium couple, both the iron and gold-palladium wires having a diameter of 0.0005 inch. The heater resistance is 1000 ohms.

Vacuum thermocouples of the type just described cannot be constructed to dissipate powers in excess of one watt without sacrificing their desirable characteristics. To dissipate considerable power without burning out the heater and couple, the area of the heater must be made large. This can be accomplished by using an undesirably low heater resistance, or by making the heater long and of small diameter, in which case the current is no longer uniform throughout the length

of the heater and the use of a low-frequency calibration is impossible. To overcome these difficulties, indirectly heated thermocouples have been built. These "thermocouples" have a long heater of fine wire enclosed in a metal box of high thermal conductivity. The box then absorbs all the heat dissipated by the heater, whatever the current distribution on the heater, and assumes a temperature depending on the amount of heat dissipated. A thermocouple attached to the outside of the box measures its temperature. These "thermocouples" require about 15 minutes to reach a steady-state temperature so they are not particularly useful when time is a consideration. Figure 5 shows the calibration curve of such a thermocouple.

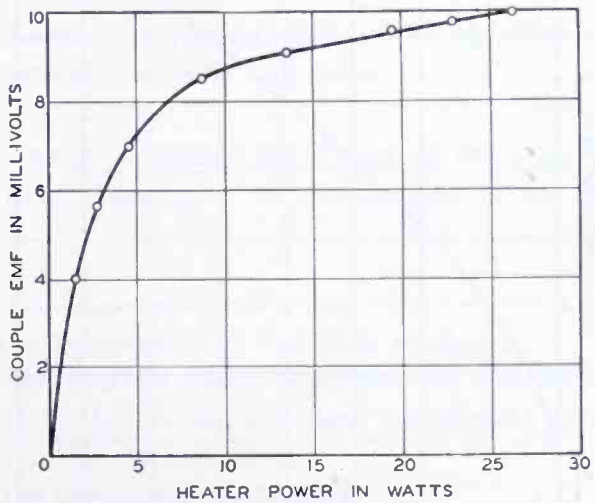


Fig. 5—D-C calibration of laboratory thermocouple R-356. The couple e.m.f. was read on a meter of 10 ohms resistance.

For the measurement of powers of the order of one watt, small diodes can also be used. In these the power is dissipated in the filament and the temperature-limited emission serves as a measure of the power dissipated in the filament. The filament can be made short to insure a uniform current distribution, and the resistance can be made of the order of 50 ohms. The anode voltage must be high enough to insure temperature limitation of the current at all times. The anode dissipation is, therefore, a limitation which considerably restricts the range of power which can be measured with such a diode.

For the measurement of powers between 2 watts and 100 watts, incandescent lamps are quite commonly used. The power to be measured is dissipated in a lamp and a second lamp run from a 60-cycle source is adjusted until the two have the same brilliancy. The adjustment is made visually and with considerable accuracy. To check the accuracy of the method, measurements were made independently by

three observers and were found to agree within 5 per cent. Considerable inaccuracy doubtless results from non-uniform heating of the lamp filament at short wavelengths. However, there is something very convincing about seeing a lamp brilliantly lighted. A refinement of this method consists in measuring the light output of a lamp with a phototube. The lamp and cell as a unit are calibrated at 60 cycles per second. Figure 6 shows the calibration curve of such a lamp and phototube.

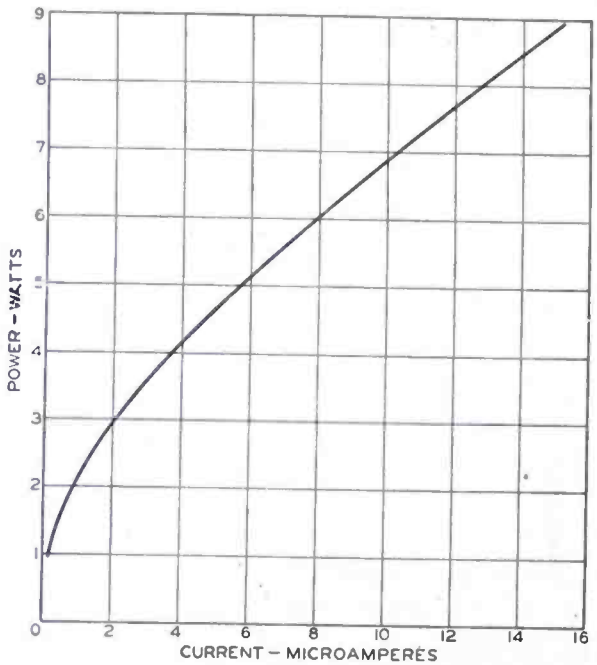


Fig. 6—Low-frequency calibration of an incandescent lamp and phototube used for power measurements.

VI. The Measurement of Voltage

The requirements of a good voltmeter are:

(1) It should have adequate sensitivity.

(2) Its input impedance should be high so that it does not perturb the circuit to which it is connected.

(3) Its calibration should be independent of frequency.

Triode voltmeters have these attributes at long wavelengths so it was natural to try to use them at ultra-short wavelengths. It was soon found that they loaded tuned circuits very considerably. As a matter of fact, their input resistance appeared to be less than 10^4 ohms at a wavelength of 100 cm. Because the calibrations of the existing volt-

meters were not known, only estimates of impedances could be made at the time. Diode voltmeters of the type shown in the inset of Figure 7 were then tried. The diodes used had an anode diameter of 0.012 inch and a cathode diameter of 0.0026 inch. These voltmeters were found to have an input impedance of the order of 10^5 ohms.

In comparing diode voltmeters at various wavelengths, two sources of error were found:

(1) A voltage step-up between the external terminals and the diode electrodes due to series resonance between the lead inductances and the interelectrode capacitance of the tube. This error is usually referred to as the resonance error.

(2) A non-linearity in calibration due to the transit time of electrons between cathode and anode.

The resonance error is usually the larger of the two. Fortunately, it is easier to correct because it is independent of the applied voltage.

Let

L_d = the inductance of the leads of the diode
C_d = the interelectrode capacitance of the diode
$\lambda_r = 2\pi v \sqrt{L_d C_d}$ = the resonant wavelength of the diode
$v = 3 \times 10^{10}$ cm per second
λ = the operating wavelength
E_2 = the voltage across the terminals of the diode
E_1 = the voltage across the electrodes of the diode

Now L and C are in series, so

$$\frac{E_1}{E_2} = \frac{1}{1 - \omega^2 L_d C_d} = \frac{1}{1 - \left(\frac{\lambda_r}{\lambda}\right)^2} \qquad \text{(VI-1)}$$

This equation assumes, of course, that λ is sufficiently remote from λ_r so that the resistance of the leads may be neglected. Once the resonant wavelength λ_r has been measured, the reading of the voltmeter is easily corrected for this error by multiplying the indicated voltage by $1 - (\lambda_r/\lambda)^2$. The determination of the resonant wavelength will be discussed in the section devoted to the measurement of reactance.

The nature of the transit-time error can be seen from the following argument: Consider a diode and condenser connected across a source

of constant high-frequency voltage. When the voltage is applied, the diode will pass current on the positive peaks of voltage and charge the condenser. As the charge on the condenser increases, the peak positive swing across the diode electrodes decreases. If the time of transit of the electrons were zero, the condenser would charge up to the peak value of the alternating voltage, i.e., until the electrons at the cathode felt no accelerating force at any part of the cycle. Actually, the transit time is not zero so the condenser will charge only to such a voltage that electrons just fail to reach the anode before their velocity becomes zero because of the retarding field on the inverse half-cycle. Since the condenser does not charge to the peak of the alternating voltage when the transit time is appreciable, the voltmeter reads low. Actual voltmeters have a resistance across the condenser. However, when the time constant of the resistor-capacitor combination is large compared to the period of the applied voltage, the resistance changes the picture inappreciably. This point has been checked by using an electrometer in place of the resistor and microammeter. Because the finite transit-time makes the diode cut off at a lower voltage at short wavelengths than at long wavelengths, this effect is sometimes called *premature cut-off*.

The transit-time error was first studied theoretically.[11,12,13] The analysis led to the following approximate formula for the error in cylindrical diodes.

$$\frac{\triangle E}{E} = -\frac{\Gamma K d}{\lambda \sqrt{E}} \qquad \text{(VI-2)}$$

in which

$\triangle E$ = the error in peak volts
E = the applied peak voltage
d = the spacing between cathode and anode in cm
λ = the wavelength in cm
Γ = a constant to be determined experimentally
K = a function of the ratio of anode diameter to cathode diameter which, in effect, reduces a cylindrical diode to an equivalent parallel-plane diode[14]

Because all voltmeters suffer transit-time errors, no standard voltage could be established. Therefore, the formula had to be verified and the constant Γ determined by comparing voltmeters of different spacings. To accomplish this a series of diodes of different spacings was built. To verify the dependence of error with wavelength and time constant of the resistance-condenser combination is large com-

voltage, the discrepancy in readings of a pair of voltmeters employing two of the diodes was measured as a function of wavelength and voltage. To verify the variation of error with spacing and to determine the constant Γ, the discrepancy in readings between each diode of the series and a "standard" diode was measured. This procedure made it possible to plot Γ in Equation (VI-2) as a function of Kd. A straight line resulted. The slope Γ had the value of 562 volts$^{1/2}$.

In Figure 7 are shown the calibration curves at two wavelengths of a diode voltmeter employing a diode having an anode diameter of

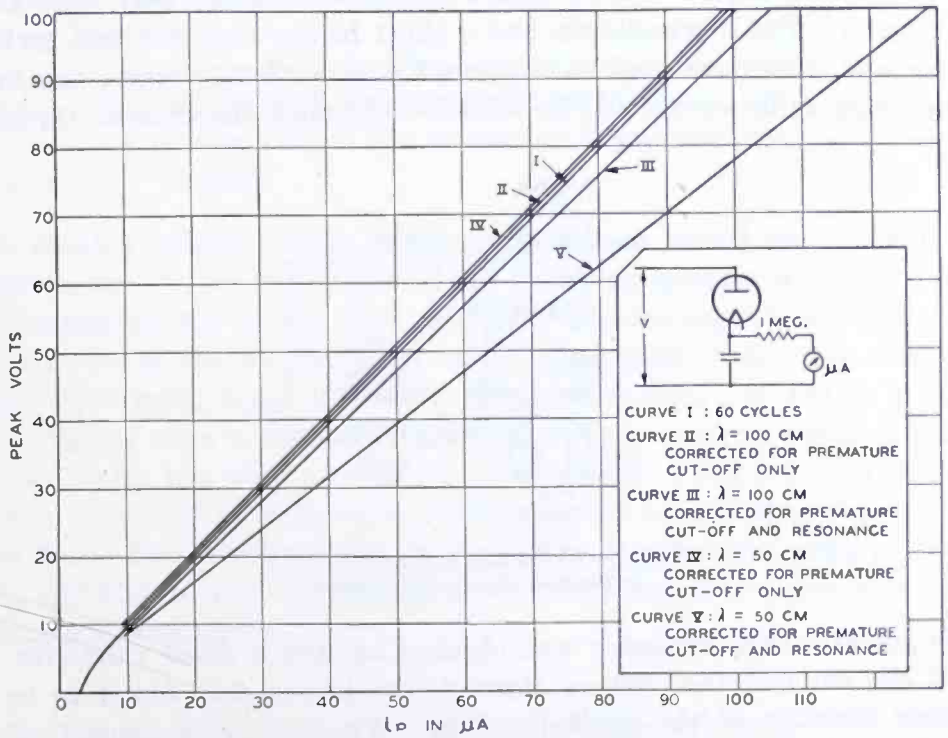


Fig. 7—Calibration of laboratory diode R-353 as a peak voltmeter.

0.009 inch and cathode diameter of 0.0026 inch. Note that the resonance error is by far the larger of the two errors. For this diode the resonance correction is $1-(25/\lambda)^2$ and the premature cut-off correction is $\Delta E = -7.5\sqrt{E}/\lambda$

In making voltage measurements on parallel-wire transmission lines, it is advantageous to have a balanced voltmeter. Figure 8 shows the schematic circuit diagram of such a voltmeter. It will be noticed that each cathode return has its own resistor. If the diodes were perfectly matched and the line perfectly balanced, the cathodes could be returned to ground through a common resistor. However, if these conditions do not obtain, one of the diodes will cut off before the other

and the indicated voltage will be merely the voltage from one side of the line to ground. When separate resistors are used, this difficulty is overcome and the indicated voltage is the total voltage between the lines.

The RCA-955 "acorn" triode with grid and anode tied together is frequently used as a diode voltmeter. Its resonance and premature cut-off corrections, respectively, are approximately $1-(40/\lambda)^2$ and $\Delta E = -30\sqrt{E}/\lambda$. While the RCA-955 is not quite as satisfactory as a diode built for the purpose of measuring voltage, its ruggedness and ready availability make it very useful for this purpose.

A thermocouple in series with a resistance has also been used as a voltmeter. The thermocouple had a short heater of 0.002-inch carbon wire and, therefore, had no appreciable skin-effect. Hence, the low-frequency calibration could be used to determine the current through

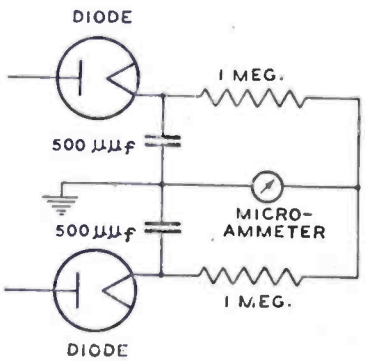


Fig. 8—Schematic diagram of a balanced diode voltmeter.

the heater. This voltmeter was checked against a diode voltmeter at $\lambda = 184$ cm and the current through the heater was found to be a linear function of the applied voltage. Whereas the measured value of the resistance in series with the thermocouple was 40,000 ohms for direct current, the apparent resistance of the voltmeter obtained by dividing the applied voltage by the current through the thermocouple was only 14,300 ohms. This decrease in resistance is due to the stray capacitances of the resistor. The writer would hesitate to use such a voltmeter without a calibration made with the resistance and thermocouple in the exact positions in which they are to be used. With such a voltmeter extreme care must be taken to prevent stray fields from reaching the couple circuit. Because the couple circuit has a very low resistance compared to the heater circuit, a stray field which does not affect the heater circuit appreciably can cause large circulating currents in the couple circuit. Such currents greatly disturb the calibration of the instrument and may have dire effects on the thermocouple.

The writer recently burned out three thermocouples in ten minutes before he obtained shielding enough to make operation safe for the thermocouples.

It is sometimes necessary to obtain known voltages whose magnitudes are smaller than those which can be measured with existing voltmeters. For instance, it may be necessary to calibrate a crystal detector to be used in field-strength measurements. To provide such voltages, a capacitance attenuator has been built. This attenuator consists of two parallel, circular condenser plates enclosed in a concentric cylinder. The capacitance between plates in such a condenser varies as[15] $\exp(-K_1 z/a)$ when $z/a > 1$

where

z = spacing between plates in cm
a = radius of enclosing cylinder in cm
K_1 = first root of the zero-order Bessel function
= 2.405

The spacing between plates in the attenuator is varied by means of a micrometer screw. The radius of the enclosing cylinder and the pitch of the micrometer thread were so chosen that the capacitance changes by a factor of ten for ten turns of the micrometer head. The dimensions were made small so that an attenuation range of 100 to 1 can be obtained at a wavelength of 50 cm with a spacing between condenser plates less than a tenth of a wavelength. This prevents errors due to the radiation field which attenuates less rapidly than the induction field. The attenuator has been calibrated at wavelengths of 151, 129, and 112 cm and the attenuation has been found to agree with the calculated attenuation.

VII. The Measurement of Reactance

Reactance is usually measured by connecting the unknown reactance across one end of a transmission line, coupling the line loosely to a generator, and tuning the circuit to resonance by varying the length of the transmission line. If the unknown reactance X is tuned to resonance with a short-circuited line, the reactance is

$$X = -\sqrt{\frac{L_o}{C_o}} \tan \frac{2\pi l_o}{\lambda} \qquad \text{(VII-1)}$$

It may be necessary to take into account the inductance of the "short circuit". This can be done as described in the section on the measurement of wavelength. If an open-ended line is used, the reactance is

$$X = \sqrt{\frac{L_o}{C_o}} \cot \frac{2\pi l_o}{\lambda} \qquad \text{(VII-1-a)}$$

This method is usually simple and straightforward and is perhaps best demonstrated by example.

As a first example, consider the measurement of the resonant wavelength of a diode of the type described in the discussion of voltmeters. Before apparatus is set up, it is usually advantageous to theorize briefly in order to determine what is required of the apparatus. The reactance of a diode is

$$X = \omega L_d - \frac{1}{\omega C_d}$$

Therefore, the condition for resonance when a short-circuited line is used is

$$\omega L_d - \frac{1}{\omega C_d} = -\sqrt{\frac{L_o}{C_o}} \tan \frac{2\pi}{\lambda}\left(l_o + \frac{L}{L_o}\right) \qquad \text{(VII-2)}$$

The correction for the inductance of the "short-circuit" has been included.

Now we can write

$$\omega L_d - \frac{1}{\omega C_d} = \omega L_d \left[1 - \frac{1}{\omega^2 L_d C_d}\right] = \frac{2\pi v L_d}{\lambda}\left[1 - \left(\frac{\lambda}{\lambda_r}\right)^2\right]$$

On substituting this relation in Equation (VII-2), we obtain

$$\lambda \tan \frac{2\pi}{\lambda}\left(l_o + \frac{L}{L_o}\right) = -2\pi v L_d \cdot \sqrt{\frac{C_o}{L_o}}\left[1 - \left(\frac{\lambda}{\lambda_r}\right)^2\right]$$

$$= +2\pi \frac{L_d}{L_o}\left[\left(\frac{\lambda}{\lambda_r}\right)^2 - 1\right] \qquad \text{(VII-3)}$$

Therefore, if we plot $(\lambda \tan 2\pi/\lambda)(l_o + L/L_o)$ versus λ^2, we should obtain a straight line.

The intercept on the λ^2 axis is λ_r^2, and the slope is

$$S = \frac{2\pi}{\lambda_r^2}\frac{L_d}{L_o}$$

Hence, the series inductance of the diode is

$$L_d = \frac{\lambda_r^2}{2\pi} SL_o \qquad \text{(VII-4)}$$

Finally, we can obtain the interelectrode capacitance of the diode from the relation

$$C_d = \frac{\lambda_r^2}{(2\pi v)^2 L_d} \qquad \text{(VII-5)}$$

To determine λ_r, L_d, and C_d, we require the following data:

(1) The inductance per unit length L_o of the line

(2) The inductance L of the "short circuit"

(3) The length of line l_o for resonance for a series of values of the wavelength λ

To obtain these data, a transmission line consisting of two ⅜-inch copper tubes with a center-to-center spacing of ¾ inch was used. For this line

$$C_o = 2.11 \times 10^{-13} \text{ farad per cm}$$

$$L_o = 5.26 \times 10^{-9} \text{ henry per cm}$$

$$\sqrt{L_o C_o} = 158 \text{ ohms}$$

The length of the line could be varied by moving a movable short-circuit. The inductance of the short circuit was determined in the manner described in connection with wavemeters, and was found to be $L = 5.89 \times 10^{-8}$ henry. Then, the correction in length for this inductance was $L/L_o = 1.12$ cm. The diode was connected to the open end of the line. In this case it was convenient to use the diode as an indicator of resonance. In order to avoid extraneous effects due to the filament leads of the diode, these leads were brought through one of the copper conductors of the transmission line to the inactive part of the line which was essentially at ground potential. The length of the line for resonance as a function of wavelength was then measured, and $(\lambda \tan 2\pi/\lambda)(l_o + L/L_o)$ plotted against λ^2. Figure 9 shows this plot. The intercept on the λ^2 axis is 550. Therefore, the resonant wavelength is $\lambda_r = \sqrt{550} = 23.4$ cm. The slope of the line is $S = 4.67 \times 10^{-2}$ cm^{-1}. Hence, from Equation (VII-4), the diode inductance is

$$L_d = \frac{(23.4)^2}{2\pi} \times 4.67 \times 10^{-2} \times 5.26 \times 10^{-9} = 2.15 \times 10^{-8} \text{ henry}$$

and the capacitance, from Equation (VII-5) is

$$C_d = \frac{(23.4)^2}{(2\pi \times 3 \times 10^{10})^2 \times 2.15 \times 10^{-8}} = 0.719 \times 10^{-12} \text{ farad}$$

Summarizing these results, we have

$\lambda_r = 23.4$ cm = resonant wavelength of diode
$L_d = 2.15 \times 10^{-8}$ henry = inductance of diode leads
$C_d = 0.719 \times 10^{-12}$ farad = interelectrode capacitance

As a second example, we shall consider the calibration of a symmetrical variable condenser used in resistance measurements.

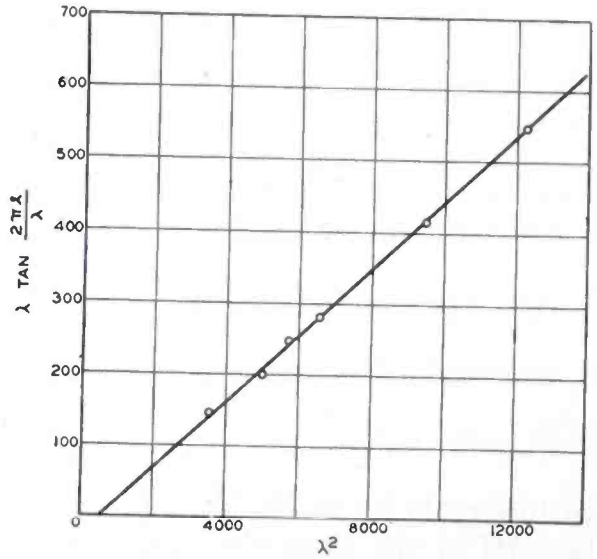


Fig. 9—Determination of the resonant wavelength of a diode.

The condenser is essentially a concentric cylinder condenser with the outer conductor split symmetrically into two electrodes. The inner cylinder has a diameter of 5/16 inch; and the outer conductor has a diameter of 3/8 inch. The angular aperture of the slits dividing the outer conductor into two segments is 60 degrees. The capacitance between the two segments is varied by moving the central cylinder axially with a micrometer screw. When the segments of the outer cylinder are connected across a symmetrical circuit, the central conductor, which is not connected to either segment, assumes ground potential and may be connected to ground if advantageous.

The purpose of the measurement to be described was to determine the change in capacitance between electrodes per division on the micrometer head. Figure 10 is a photograph of this condenser. The condenser was mounted at the end of the transmission line used in

the previous example. The voltmeter was left in place to indicate resonance. The capacitance of the condenser can be written

$$C_1 + t \cdot C_2 \qquad \text{(VII-6)}$$

in which

C_1 = the zero capacitance
C_2 = the capacitance per division of the micrometer head of the condenser
t = reading of the micrometer

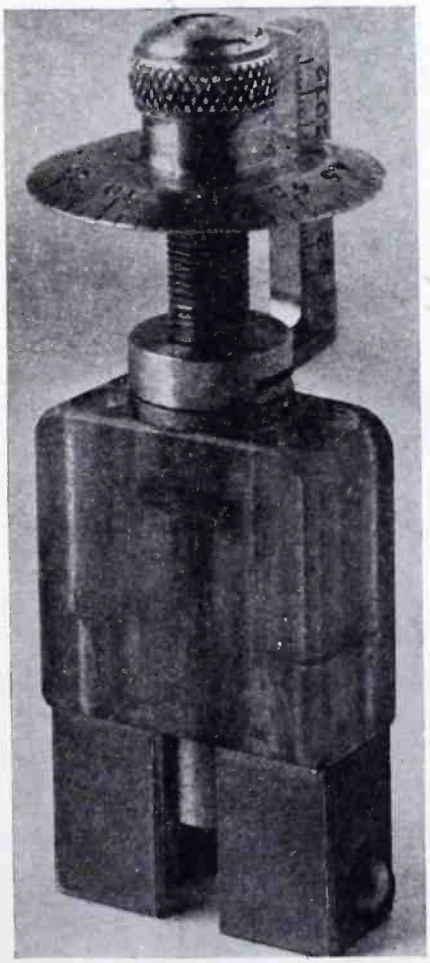


Fig. 10—Balanced condenser for the measurement of resistance by the reactance-variation method.

The diode voltmeter was in parallel with the condenser, so the reactance of the combination was

$$X = \frac{1}{-\omega(C_1 + t \cdot C_2) + \dfrac{1}{\omega L_d - \dfrac{1}{\omega C_d}}} \qquad \text{(VII-7)}$$

During the calibration at a given wavelength, all of the quantities except t in this expression are constant. Hence, for convenience, Equation (VII-7) can be written

$$X = \frac{1}{-\omega C_2 (t + A)} \tag{VII-8}$$

Then the equation (VII-1) for resonance becomes

$$\frac{1}{\omega C_2 (t + A)} = \sqrt{\frac{L_o}{C_o}} \tan \frac{2\pi}{\lambda} \left(l_o + \frac{L}{L_o} \right)$$

or

$$\cot \frac{2\pi}{\lambda} \left(l_o + \frac{L}{L_o} \right) = \omega C_2 \sqrt{\frac{L_o}{C_o}} (t + A) = \frac{2\pi}{\lambda} \frac{C_2}{C_o} (t + A) \tag{VII-9}$$

Therefore, when $\cot 2\pi (l_o + L/L_o)/\lambda$ is plotted against t, a straight line results, the slope of which is

$$S = \frac{2\pi}{\lambda} \frac{C_2}{C_o}$$

Then the required capacitance is

$$C_2 = \frac{\lambda S}{2\pi} C_o \tag{VII-10}$$

The measurement consists in measuring l as a function of t. Figure 11 is a plot of $\cot 2\pi (l_o + L/L_o)/\lambda$ versus t at a wavelength of 98.4 cm. The slope of the line is

$$S = 1.46 \times 10^{-4}$$

Therefore, the capacitance per division on the micrometer head is

$$C_2 = \frac{98.4 \times 1.46 \times 10^{-4}}{2\pi} \times 2.11 \times 10^{-13} = 4.82 \times 10^{-16} \text{ farads}$$

A calibration was made at a wavelength of 308 cm; the result checked the value found above within one per cent, which is within the experimental accuracy.

VIII. The Measurement of Resistance

Resistance has been measured by three methods:

(1) Substitution method
(2) Resistance-variation method
(3) Reactance-variation method

At short wavelengths all of these methods employ a tuned circuit loosely coupled to a generator. This arrangement is equivalent to connecting the tuned circuit in series with a very high impedance across

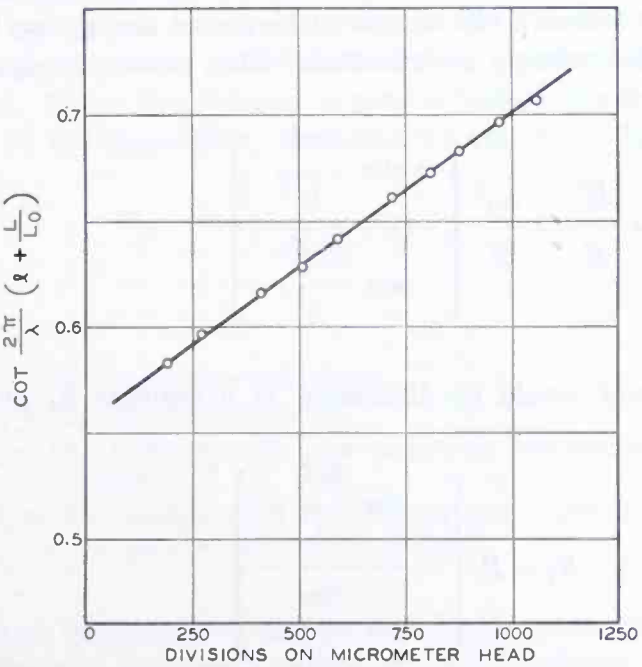


Fig. 11—Determination of the calibration of the balanced condenser.

a constant-voltage generator. Then the voltage developed across the tuned circuit is proportional to its impedance.

J. M. Miller and B. Salzberg of this laboratory have used the substitution method to study commercial resistors at frequencies up to 250 megacycles and to measure dielectric losses.[16] In this case, the tuned circuit consists of a transmission line of low characteristic impedance, terminated by a short circuit at one end and by a capacitance at the other end. The voltage distribution along the line is then

$$E = E_1 \frac{\sin \dfrac{2\pi l}{\lambda}}{\sin \dfrac{2\pi l_1}{\lambda}} \qquad \text{(VIII-1)}$$

where

l_1 = the total length of line
E_1 = the voltage at l_1
l = the distance from the short circuit to any point on the line
E = the voltage at l

If a resistance R is placed across the line at l and the current through the resistor is inappreciable compared to the circulating current in the line at that point, the voltage distribution will not be altered. When the circuit is resonant, the circulating current is very high so that resistances as low as 8,000 ohms can be placed across the line without disturbing the voltage distribution. The power dissipated in the resistor is

$$W = \frac{E^2}{R} = \frac{E_1^2}{R}\left[\frac{\sin\frac{2\pi l}{\lambda}}{\sin\frac{2\pi l_1}{\lambda}}\right]^2$$

The same power would be dissipated in a resistor R_1 at the end of the line if

$$R_1 = R\left[\frac{\sin\frac{2\pi l_1}{\lambda}}{\sin\frac{2\pi l}{\lambda}}\right]^2 \qquad \text{(VIII-2)}$$

Therefore, placing a resistor R at l is equivalent to placing a resistor R_1 at the end of the line. This assumes that the equivalent resistance R_1 is small compared to the resonant impedance of the line alone. In the line used by Miller and Salzberg, $l_1 << \lambda/4$ so that

$$R_1 = R\left(\frac{l_1}{l}\right)^2 \qquad \text{(VIII-3)}$$

A series of resistors of various values were placed across the line at such positions that R_1 remained constant, as evidenced by the constant reading of a vacuum-tube voltmeter across the line. For low values of the particular type of resistor used, they found the resistance measured at low frequencies in agreement with Equation (VIII-3). Since it is quite unlikely that these resistors should all have the same percentage error, they assumed that the high-frequency values

were the same as the low frequency. Having thus established a number of standard resistances, other resistors were measured by comparing them with the standards through relation (VIII-3). It was possible by this means to measure very high values of resistance. It is of interest to note that certain ½-watt resistors of the ceramic type, in the resistance range from 10,000 to 100,000 ohms, show less than 10 per cent deviation from their low-frequency values up to 250 megacycles.

The resistance-variation method consists in placing the unknown resistance and a voltmeter across a resonant circuit which is loosely coupled to a generator. The voltage across the tuned circuit is then read. A known resistance is added across the circuit and the voltage is again read. Since the voltage across a loosely coupled circuit is proportional to the impedance, the unknown resistance has a value

$$R_x = R_s\left(\frac{E_x}{E_s} - 1\right) \qquad \text{(VIII-4)}$$

in which

R_x = the unknown resistance
R_s = the known resistance
E_x = the voltage with the unknown resistance across the circuit
E_s = the voltage with the known and unknown resistances in parallel across the circuit

If the resonant impedance of the tuned circuit is comparable to the unknown resistance, the circuit impedance must be taken into account. The resonant impedance can be measured by the same method. This method has been useful for wavelengths greater than 120 cm ($f < 250$ Mc) where the work of Miller and Salzberg has provided a series of known resistances.

The reactance-variation method takes three forms at short wavelengths:

(1) The capacitance-variation method
(2) The line-length-variation method
(3) The wavelength-variation method

In all these forms, the unknown resistance is placed across a tuned circuit, usually a transmission line terminated at one end by a short circuit and at the other by a capacitance. The tuned circuit is loosely coupled to a generator and the voltage across the line at some point is read.

The capacitance-variation method is perhaps the simplest method to use. A variable condenser of known calibration is connected across the circuit. Then the admittance of the circuit is

$$Y = G + j(B + \omega C) \qquad \text{(VIII-5)}$$

in which

G = the parallel conductance of the line and unknown impedance
B = the parallel susceptance of the line and unknown impedance
ωC = the susceptance of the condenser

Let $$\omega C = \omega C' + \omega \triangle C$$

where C' = the capacitance which makes the circuit resonant.

Then $$Y = G + j\,[B + \omega C' + \omega \triangle C] = G + j\omega \triangle C \qquad \text{(VIII-6)}$$

Now the voltmeter reading is proportional to the magnitude of the impedance. Therefore

$$\frac{1}{E^2} = k\,[G^2 + (\omega \triangle C)^2] \qquad \text{(VIII-7)}$$

where k is a proportionality factor. At resonance, $\triangle C = 0$ and the voltage E_o is given by

$$\frac{1}{E_o^2} = kG^2$$

Then

$$\frac{E_o^2}{E^2} = 1 + \left(\frac{\omega \triangle C}{G}\right)^2 \qquad \text{(VIII-8)}$$

Therefore, when the circuit is detuned until the voltage drops to 0.707 times the voltage at resonance, the resistance of the circuit is equal to the change in capacitance reactance:

$$R = \frac{1}{G} = \frac{1}{\omega \triangle C} \qquad \text{(VIII-9)}$$

The condenser described in Section VII as built for measuring resistances by this method. Because of its symmetrical form, it is particularly useful for measuring the resonant impedances of parallel-wire transmission lines.

A special transmission line and condenser system has been built for measuring resistance. This is shown schematically in Figure 12. Figure 13 is a photograph of the instrument. The system consists of a concentric transmission line with one end terminated in a short-circuit and the other terminated in a concentric-cylinder condenser. The short-circuit is mounted on a brass cylinder which is driven by a lead screw to make adjustments of line length. A second short-circuit

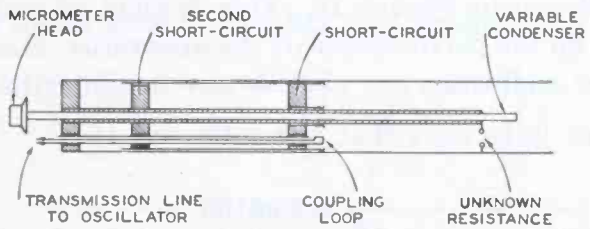


Fig. 12—Schematic diagram of a transmission line and condenser for the measurement of resistance by the reactance-variation method.

is provided to suppress standing waves on the inactive part of the line. The inner conductor of the variable condenser is brought back to the inactive part of the transmission line through the central conductor of the line. A micrometer screw controls the variable condenser. The change in capacitance of the condenser per turn of the micrometer head is $C_T = 4 \times 10^{-14}$ farad when the length of the inner conductor is very small compared to a wavelength. Since this is usually not true, the voltage distribution on the condenser must be taken into account. When this is done the change in capacitance per turn is $C_T = 4 \sec^2 \frac{2\pi l}{\lambda} \times 10^{-14}$ where l is the length of the inner conductor of the condenser. The line is excited by a small coupling loop near the short-circuit.

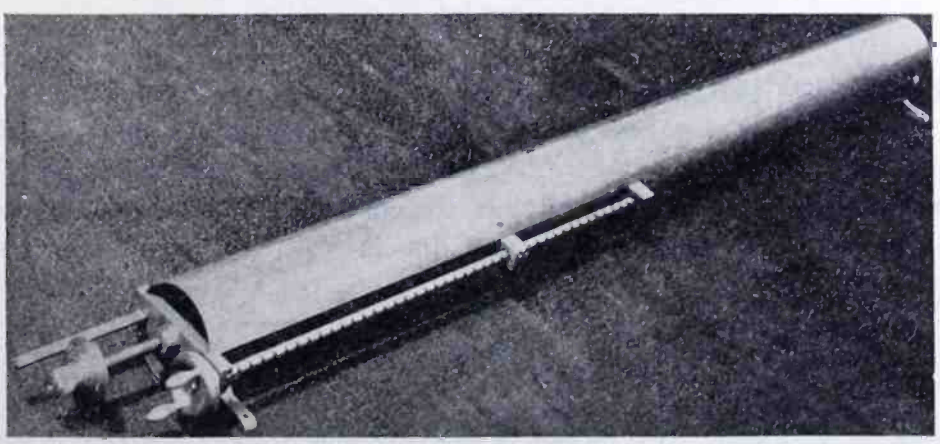


Fig. 13—The transmission line and condenser for the measurement of resistance.

Figure 14 shows the resonance curve of a diode voltmeter at a wavelength of 184 cm measured with this apparatus. R in the figure is the reading of the micrometer in turns. This curve appears symmetrical and of the normal form. However, to check the constancy of the voltage supply and coupling, the reciprocal of the voltage squared was plotted against the square of the increment of R from the value at resonance. This is essentially a plot of Equation (VIII-7). The curve is shown in Figure 15. This method of plotting gives an excellent check on the performance of the apparatus. Since the squares of the pertinent quantities are plotted, any irregularities are doubled. From the curve it is seen that the intercept is

$$\frac{1}{E^2} = 0.00135$$

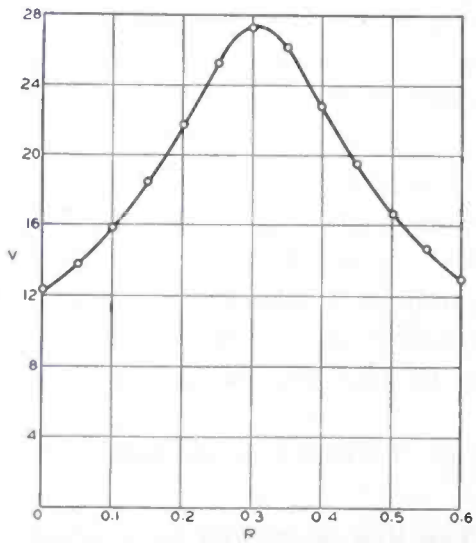


Fig. 14—Resonance curve of a diode voltmeter across the transmission line for the measurement of resistance.

When $\frac{1}{E^2}$ has twice this value

$$R = \sqrt{0.0225} = 0.150$$

The length of the inner conductor of the condenser was 5.5 cm. Hence, the resistance of the voltmeter and line is

$$R = \frac{\cos^2 \frac{2\pi l}{\lambda}}{\omega C_T \cdot \triangle R} = \frac{\cos^2 \frac{2\pi \times 5.5}{184}}{\frac{2\pi}{184} \times 3 \times 10^{10} \times 4 \times 10^{-14} \times 0.150}$$

$$= 157{,}000 \text{ ohms}$$

The length-of-line variation method is very similar to the capacitance-variation method except that the length of a transmission line instead of the capacitance of a condenser is varied. Let

G = the parallel conductance of the unknown resistor
G_L = the parallel conductance of the line
B = the parallel susceptance of the unknown resistor

Then the total admittance of the line and resistor is

$$Y = G + G_L + j\left[B - \sqrt{\frac{C_o}{L_o}}\cot\frac{2\pi l}{\lambda}\right] \qquad \text{(VIII-10)}$$

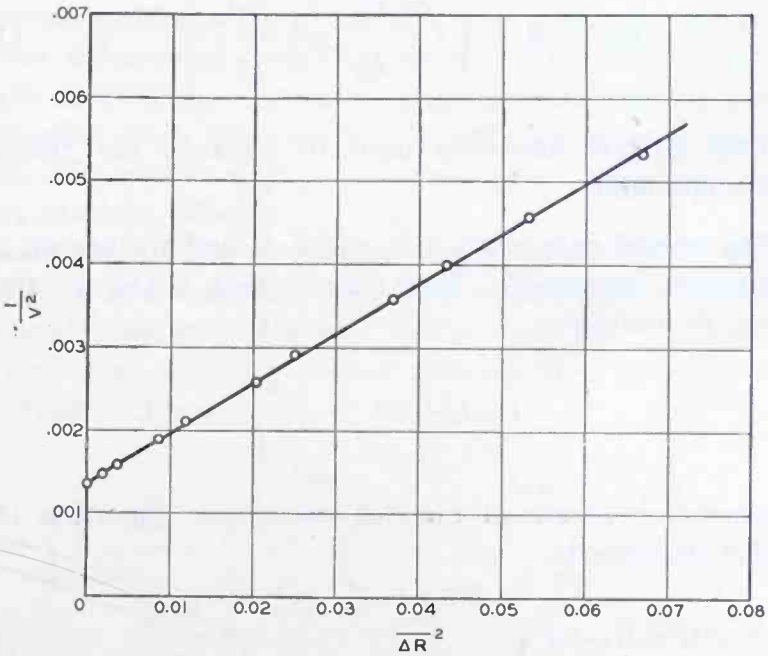


Fig. 15—Determination of the input resistance of a diode voltmeter.

G_L is a function of l/λ. However, since G_L is small, less than 5×10^{-5} mho for practical transmission lines, its variation with l/λ causes no difficulty in measurements. The reasons will be detailed in connection with particular applications of the formula. The condition for resonance is

$$B = \sqrt{\frac{C_o}{L_o}}\cot\frac{2\pi l_o}{\lambda} \qquad \text{(VIII-11)}$$

where l_o = the length, which satisfies (VIII-11). Hence, B is determined by measuring the length for a maximum voltage reading.

The measurement of G separates conveniently into two cases:

(a) The first case is that in which $G >> G_L$. Then G_L can be neglected and Equation (VIII-10) becomes:

$$Y = G + j\left[B - \sqrt{\frac{C_o}{L_o}} \cot \frac{2\pi l}{\lambda} \right] \qquad \text{(VIII-12)}$$

B in this relation is first measured as outlined above. Then the line length is changed until the voltage drops to 0.707 times its resonant value, and the line length l_1 for this condition is measured. The conductance is then

$$G = \left\{ B - \sqrt{\frac{C_o}{L_o}} \cot \frac{2\pi l_1}{\lambda} \right\} \qquad \text{(VIII-13)}$$

This method has been used to measure the resistance of thermocouples.

(b) The second case is that in which G and G_L are of the same order of magnitude. Then because both G and G_L are small, l can be written

$$l = l_o + \Delta l, \qquad \frac{\Delta l}{l_o} << 1$$

where l_o = length of line for resonance. Equation (VIII-10) then reduces to

$$Y = G + G_L + j \sqrt{\frac{C_o}{L_o}} \frac{2\pi \Delta l}{\lambda} \csc^2 \frac{2\pi l_o}{\lambda} \qquad \text{(VIII-14)}$$

The conductance of the line and resistor are determined by changing the length such an amount Δl_o that the voltage drops to 0.707 times its resonant value. Then

$$G + G_L = \sqrt{\frac{C_o}{L_o}} \cdot \frac{2\pi \Delta l_o}{\lambda} \csc^2 \frac{2\pi l_o}{\lambda} \qquad \text{(VIII-15)}$$

During the detuning G_L also changes. However, the real and imaginary parts of Equation (VIII-14) are of the same order of magnitude. Therefore, the change in G_L will be of the order of G_L^2, which may be neglected in view of the small magnitude of G_L.

The latter method is used extensively to determine the impedance of quarter-wave transmission lines. For such lines Equation (VIII-15) becomes

$$G = \frac{\pi}{2}\sqrt{\frac{C_o}{L_o}}\,\frac{\Delta l_o}{l_o} \qquad \text{(VIII-16)}$$

P. D. Zottu of this laboratory has used this method to determine the resistance of circular conductors.[17] In essence, his method is the following. The inner conductor of a quarter-wave concentric transmission line was made of the material whose resistance was to be measured. The impedance of the line was determined from the appropriate measurements and Equation (VIII-16). The resistance per unit length of line was determined from Equation (III-8). The resistance per unit length of the outer conductor of the line, previously determined, was subtracted from this result. The remainder was the required resistance. His results are in good agreement with the asymptotic skin-effect formula (III-18).

The wavelength-variation method is so similar to the other methods that it will not be described in detail. It is sufficient to say that when the wavelength has been shifted from the resonant wavelength by an amount such that the line voltage has dropped to 0.707 times its resonant value

$$G + G_L = \sqrt{\frac{C_o}{L_o}}\cdot\frac{2\pi l}{\lambda_o}\cdot\frac{\Delta\lambda_o}{\lambda_o}\left[1 \pm \frac{\sin\dfrac{4\pi l}{\lambda_o}}{\dfrac{4\pi l}{\lambda_o}}\right]\csc^2\frac{2\pi l}{\lambda_o} \qquad \text{(VIII-17)}[18]$$

in which

G = the conductance of the unknown
G_L = the conductance of the line
λ_o = the resonant wavelength
$\Delta\lambda_o$ = the change in wavelength from λ_o to reduce the voltage to 0.707 times its value at resonance.

This formula assumes that G and G_L are small, i.e., less than 10^{-3} mho.

Mr. G. R. Kilgore of this laboratory has used this method to measure the resonant impedance of the anode circuit of a magnetron. An oscillator was loosely coupled to the anode circuit. A second fixed-frequency oscillator was set up sufficiently remote from the circuit

to induce no perceptible voltage across the circuit. To measure the difference in frequency between the two oscillators, a crystal detector was connected between the ground and antenna terminals of a broadcast receiver and the difference in frequency read on the calibrated tuning dial. The frequency of the oscillator coupled to the magnetron circuit was varied and the frequency shift from resonance necessary to reduce the resonant voltage across the line to the 0.707 point was measured. In this work, Mr. Kilgore found that a slight amount of unbalance in a parallel-wire line greatly reduces the resonant impedance, as much as 50 per cent in some cases. This reduction in impedance is apparently due to the increase in the radiation resistance of the line with unbalance.

In the course of his experimental work on the measurement of impedance, the writer encountered an anomoly which may considerably trouble others. He found that resonance curves in some instances were asymmetrical near the resonance point. The theory was carefully checked and no trace of such an asymmetry was disclosed. The difficulty finally was traced to the circuit coupling the generator to the measuring circuit. The coupling circuit had a very high resonant impedance. Therefore, when it was tuned slightly off-resonance where the slope of the curve was very steep, any slight reaction of the measuring line on the circuit detuned it sufficiently to reduce its circulating current appreciably. When the coupling circuit was loaded with resistance and tuned to resonance, the difficulty was completely eliminated.

IX. The Measurement of Current

The writer and his associates have had little occasion to measure current at ultra-short wavelengths. In the past, measurements of voltage, reactance, and resistance have given all the required information. Therefore, the measurement of current lacks the background of experience which the other measurements have. However, it is apparent that instruments and a technique for measuring current will be convenient if not necessary in the future. With this in mind, some work on the measurement of current has been done.

The attributes of a satisfactory device for the measurement of current are:

(1) It should have adequate sensitivity

(2) It should have a low resistance

(3) The calibration should be independent of frequency so that the device may be calibrated at 60 cycles per second

The vacuum thermocouple has these attributes in the low-current range. The most sensitive short-wave thermocouple described in Section V will measure a current of 1 milliampere. The resistance of this thermocouple is 1000 ohms, which is not low. However, if a similar thermocouple were built for current measurement, the resistance could be reduced about 50 per cent without seriously reducing the sensitivity. For higher-current ranges, the sensitivity and low-resistance requirements are easily met.

The factors which make the calibration of a thermocouple dependent upon the wavelength of operation are:

(1) The current distribution along the heater

(2) The change in heater resistance due to skin effect

The effect of current distribution along the heater can be taken care of in the same manner in which it was taken care of in the case of power-measuring thermocouples, namely, by making the heater short. Since it is desirable to have a low heater resistance in the case of current-measuring thermocouples, the problem is simpler than in the case of power-measuring thermocouples in which a high heater resistance is desirable. Past experience indicates that a thermocouple with 1-ma sensitivity and a heater length of 4 mm can be built. Such a thermocouple would have less than one per cent error due to current distribution for $\lambda > 9$ cm.

For the measurement of current it is particularly desirable that the thermocouple heater have no skin-effect. A thermocouple depends on the temperature rise of the junction for its operation. Therefore, the resistivity of the heater varies widely over the operating range of the device. The resistance of a conductor having appreciable skin-effect does not vary linearly with the resistivity. Hence, the skin-effect correction varies not only with frequency, but also with the amount of power being dissipated in the conductor. Consequently, a thermocouple to be used for current measurement which has appreciable skin-effect must have its resistance measured over the entire operating range at each frequency at which it is to be used. Such an instrument would be quite inconvenient to calibrate.

From Equation (III-17) it can be readily shown that in order that skin-effect in a solid cylindrical conductor give rise to less than 1 per cent increase in resistance over the d-c value, the following relation must obtain

$$a\sqrt{\frac{\mu f}{\rho}} < 4.27 \qquad \text{(IX-1)}$$

For the materials used as heaters in thermocouples

$$\mu = 1$$

so Equation (IX-1) can be written

$$a^2 \frac{f}{\rho} < 18.2 \qquad \text{(IX-2)}$$

Now the resistance per unit length of a cylindrical conductor is

$$R_o = \frac{\rho}{\pi a^2} \times 10^{-6} \text{ ohms cm}^{-1} \qquad \text{(IX-3)}$$

The substitution of this relation in Equation (IX-2) gives

$$R_o > 1.75\, f \times 10^{-8} \text{ ohm cm}^{-1} \qquad \text{(IX-4)}$$

as the condition that skin-effect be negligible. Therefore, the condition that the resistance of a cylindrical conductor differ by less than one per cent from the d-c value at any wavelength above 1 cm is that

$$R_o > 525 \text{ ohm cm}^{-1} \qquad \text{(IX-5)}$$

i.e., that the resistance per unit length exceed 525 ohms per cm.[19] This condition is satisfied by most carbon conductors of less than 0.002 inch diameter. Therefore, skin-effect is a minor problem in low-current thermocouples.

As noted in Section VIII, the resistance of a few thermocouples at short wavelengths has been measured. Some of these thermocouples had carbon heaters. For these thermocouples no deviation from the d-c resistance was found. Perhaps the results should be construed as a check on the method of measurement.

For currents exceeding 100 ma, the skin-effect in solid conductors becomes troublesome. Circuits carrying high currents are usually low-resistance circuits. Therefore, the resistance of a thermocouple to be placed in such a circuit must be very low if it is not to perturb the circuit. It is readily seen from relation (IX-4) that a solid conductor must have a high resistance per unit length if skin-effect is to be inappreciable. Hence, the solid conductor is inherently unsuited for low values of resistance without skin-effect.

The best solution to the problem of measuring currents exceeding 100 ma appears to be the measurement of the voltage drop across a low resistance placed in the circuit. Thermocouples having heater resistances of the order of 100 ohms with no appreciable skin-effect will readily measure voltages of the order of 0.1 volt. It is therefore feasible to measure currents as low as 100 ma by this method with an added series resistance of only one ohm.

X. Conclusion

Methods for the measurement of wavelength, voltage, current, power, resistance, and reactance have been described. All of these methods with the exception of that for the measurement of current have been in daily use for several years. Therefore, they represent a large background of practical experience. No doubt the future will bring many improvements in both methods and technique. However, the writer feels that the methods of the future will be fundamentally the same as those of the present. After all, they are merely adaptations of the methods described in that familiar measurement classic, Bulletin 74 of the Bureau of Standards.

The writer gratefully acknowledges that he has drawn freely on the ideas and experience of his colleagues, Dr. J. M. Miller, Dr. A. V. Haeff, Mr. G. R. Kilgore, and Mr. P. D. Zottu. The specific references to their work in no way define the extent of their contributions to the work reported in this paper.

Appendix A

List of Symbols

A = a constant
a = radius of a conductor in cm
B = susceptance in mhos
C_d = interelectrode capacitance of a diode
C_o = capacitance per unit length of transmission line in farads · cm^{-1}
C_T = capacitance per turn of a micrometer condenser
d = interelectrode spacing of a diode in cm
E_r = voltage across the receiving end of a transmission line in volts
E_s = voltage across the sending end of a transmission line in volts
G = conductance in mhos
G_L = sending-end conductance of a transmission line in mhos
G_o = leakage conductance per unit length of transmission line in mhos cm^{-1}
I_r = current at the receiving end of a transmission line in amperes
I_s = current at the sending end of a transmission line in amperes
$J_n(x)$ = Bessel function of order n
K = a function of the ratio of anode diameter to cathode diameter which, in effect, reduces a cylindrical diode to a parallel-plane diode
K_1 = first root of zero-order Bessel function = 2.405
L_d = inductance of the leads of a diode in henrys
L_o = inductance per unit length of transmission line in henrys · cm^{-1}
l = length of a transmission line in cm

l_o = that length of transmission line which makes a circuit resonant

m = propagation constant of a transmission line in $cm^{-1} = \alpha + j\beta$

$$P = 2\pi \sqrt{\frac{\mu f}{\rho \times 10^3}} (1 - j) \text{ cm}^{-1}$$

R_e = "real part of"

R_o = resistance per unit length of a conductor in ohm cm^{-1}

r_o = total resistance per unit length of transmission line (per loop cm) in ohms $\cdot$ cm^{-1}

S = the slope of a curve

v = phase velocity of a wave on a transmission line in cm $\cdot$ sec^{-1}

Z_o = characteristic impedance of a transmission line in ohms $= \gamma - j\delta$

Z_r = impedance across the receiving end of a transmission line in ohms

Z_s = impedance looking into the sending end of a transmission line in ohms

z = spacing between condenser plates of attenuator (Section VI) in cm

α = attenuation constant of a transmission line in cm^{-1}

β = phase constant of a transmission line in cm^{-1}

Γ = diode voltmeter transit-time error constant (Section VI) in $volts^{\frac{1}{2}}$

γ = real component of the characteristic impedance of a transmission line in ohms

Δ = an increment, for example, Δl = increment in length

δ = reactive component of the characteristic impedance of a transmission line in ohms

ε = base of Napierian logarithms

λ = wavelength in cm

λ_r = resonant wavelength of a circuit in cm

μ = magnetic permeability; when used as prefix, denotes micro-

ρ = resistivity in microhm cm

$\omega = 2\pi f$

Bibliography

[1] A. L. Samuel, "A Negative Grid Triode Oscillator and Amplifier for Ultra-High Frequencies", *Proc. I.R.E.*, Vol. 25, No. 10, p. 1243, October (1937).

[2] O. Pfetscher and W. Puhlmann, "On Habann Tubes of Large Power for Ultra-High Frequencies", *Hochfrequenz und Elektroakus*, Vol. 47, pp. 105-115, (1936).

[3] E. J. Sterba and C. B. Feldman, "Transmission Lines for Short-Wave Radio Systems", *Proc. I.R.E.*, Vol. 20, No. 7, p. 1163, July (1932).

[4] Russell, "Skin Effect in Concentric Mains", *Phil. Mag.*, Vol. 17, p. 524, (1909).

[5] "Radio Measurements and Instruments", *Bulletin 74, Bureau of Standards.*

[6] "Funktionentafeln", Jahnke und Emde, *Teubner*, p. 142.

[7] Lester E. Reukema, "Transmission Lines at Very High Frequencies", *Elect. Eng.*, Vol. 56, No. 3, p. 1002, August (1937).

[8] Discussion on "Transmission Lines at Very High Frequencies", *Elect. Eng.*, Vol. 57, No. 2, p. 104, February (1937).

[9] S. A. Schelkunoff, "Some Equivalence Theorems of Electromagnetics and Their Application to Radiation Problems", *BSTJ*, Vol. 15, No. 1, p. 92, January (1936).

[10] August Hund, "Theory of the Determination of Ultra-Radio Frequencies by Standing Waves on Wires", *Scientific Papers of the Bureau of Standards*, No. 491.

[11] L. S. Nergaard, "Electrical Measurements at Wavelengths Less than Two Meters", *Proc. I.R.E.*, Vol. 24, No. 9, p. 1207, September (1936).

[12] C. L. Fortescue, "Thermionic Voltmeters for Use at Very High Frequencies", *Jour. I.E.E.*, Vol. 77, No. 456, p. 429, September (1935).

[13] E. C. S. Megaw, "Voltage Measurement at Very High Frequencies", *Wireless Engineer*,

Part I, Vol. 13, No. 149, p. 65, February (1936).
Part II, Vol. 13, No. 150, p. 135, March (1936).
Part II, Vol. 13, No. 151, p. 201, April (1936).

[14] W. R. Ferris, "Input Resistance of Vacuum Tubes as Ultra-High-Frequency Amplifiers", *Proc. I.R.E.*, Vol. 24, No. 1, p. 82, January (1936).

[15] The potential distribution between the inside of a cylinder and a disc electrode within the cylinder was studied by W. Schottky and J. v. Issendorf, *Zeit. f. Physik*, Vol. 31, pp. 163-201, (1925). They obtained the relation

$$V = V_0 J_0 \left(K_1 \frac{r}{a} \right) e^{-K_1 \frac{Z}{a}}$$

for the potential V at any point (r, Z) in terms of the potential V_0 at point (0, 0). The particular type of attenuator described in the present paper was described to me by J. F. Dreyer, Jr., formerly of this Company, in 1934. He had used it some years previously in measurement work. Similar attenuators are described in a paper entitled, "The Design and Testing of Multi-range Receivers", by D. E. Harnett and N. P. Case, *Proc. I.R.E.*, Vol. 23, No. 6, p. 578, June (1935).

[16] J. M. Miller and B. Salzberg, "Measurements of Admittances at Ultra-High Frequencies", *RCA-Review*, April, 1939.

[17] P. D. Zottu, "Resistance and Permeability Measurements at Ultra-High Frequencies", presented at *U.R.S.I.* meeting, April, 1938.

[18] Use + sign when $B = \omega C$ and − sign when $B = -1/\omega L$.

[19] These considerations were discussed in 1934 by P. D. Zottu in an unpublished paper entitled "Memorandum on High-Frequency Resistance with Negligible Skin-Effect".

VACUUM TUBES OF SMALL DIMENSIONS FOR USE AT EXTREMELY HIGH FREQUENCIES*

By

B. J. Thompson and G. M. Rose, Jr.

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary—*This paper describes the construction and operation of very small triodes and screen-grid tubes intended for reception at wavelengths down to 60 centimeters with conventional circuits.*

The tubes represent nearly a tenfold reduction in dimensions as compared with conventional receiving tubes, but compare favorably with them in transconductance and amplification factor. The interelectrode capacitances are only a fraction of those obtained in the larger tubes.

The triodes have been operated in a conventional feed-back oscillator circuit at a wavelength of 30 centimeters with a plate voltage of 115 volts and a plate current of 3 milliamperes.

Receivers have been constructed using the screen-grid tubes which afford tuned radio-frequency amplification at 100 centimeters and 75 centimeters, a gain of approximately four per stage being obtained at the longer wavelength.

Introduction

In recent years interest in the possibilities of radio communication at wavelengths of less than three meters has been greatly increased, because of the imminent saturation of the spectrum of greater wavelengths and of the peculiar properties expected of these short waves. As a study of radio transmission requires transmitting and receiving apparatus, much effort has been devoted to the development of equipment suitable for such wavelengths, the types of tubes and circuits used at longer wavelengths having proved unsatisfactory. It is the purpose of this paper to describe a study of the possibilities of vacuum tubes of very small physical dimensions for use in radio reception at wavelengths as short as 60 centimeters.

As the minimum wavelength of commercial radio communication has been reduced, refinements of the previously existing types of receiving apparatus have been made, until it is now possible to use either tuned radio-frequency amplification or superheterodyne amplification followed by a triode type detector at wavelengths as short as five meters. It had been found that the vacuum tubes constituted the limit-

* Decimal classification: R331. Presented before I.R.E. Eighth Annual Convention, Chicago, Illinois, June 26, 1933. Presented before New York Meeting, October 4, 1933.

Reprinted from *Proc. I.R.E.*, December, 1933.

ing factors at about ten meters, and subsequent reduction in wavelength has been made possible by improvements in tube design. In all of this work the apparatus in use differs from the conventional long-wave apparatus only in refinement. The limit of the improvements by this method seems definitely to have been reached at about three to five meters wavelength, due to various characteristics of the tubes.

As what appeared to be a wall was reached in this refinement of conventional long-wave apparatus, it was natural that investigators should seek other and radically different means for reception. By far the greatest amount of attention has been devoted to the study of Barkhausen-Kurz[1] or Gill-Morrell[2] oscillations, which may readily be obtained at wavelengths as short as thirty centimeters. Many schemes for the use of these oscillations in reception have been described, some of the circuits resembling the well-known superregenerative detector, others being used as heterodyne detectors, but the majority being more obscure in their principles of operation.[3] In general, all use only one tube—or one stage—at the ultra-high frequency, the amplification being carried out at an intermediate or low frequency. Other schemes have been proposed, using oscillating magnetrons, for example.

It is the authors' experience that these methods all suffer from one or more serious faults when considered from the standpoint of practical use. Nearly all of the methods are wasteful of plate power. Many are insensitive in the extreme. The more sensitive are unstable, in general. Tuning is broad. The most serious faults shared by all are limitation of sensitivity, due to the fact that no amplification may be had ahead of the detector, and radiation from the oscillator which is coupled directly to the antenna.

The purpose of the work described in this paper was to reduce the lower wavelength limit of the conventional types of tubes and circuits in order to obtain their advantages of simplicity at wavelengths below one meter.

Theoretical Considerations

The limitations imposed on tuned radio-frequency reception at the present lower wavelength limit are due to a number of factors. These are:

---

[1] H. Barkhausen and K. Kurz, *Phys. Zeit.*, vol. 21, no. 1; January, (1920).

[2] E. W. B. Gill and J. H. Morrell, *Phil. Mag.*, vol. 44, no. 161; July, (1922).

[3] A considerable bibliography on this subject is given by W. H. Wenstrom, *Proc. I.R.E.*, vol. 20, no. 1, pp. 95-112; January, (1932). The papers by Hollmann, Uda, Okabe, and Beauvais, in particular, discuss receiving circuits.

(1) The interelectrode capacitances of the tube are so great that, with the addition of the tuning capacitance, the L/C ratio is too low for a value of impedance sufficient to afford appreciable amplification.

(2) The lead inductances of the tube are so great that much of the output voltage of the tube appears inside the bulb, where it is unavailable.

(3) The interelectrode capacitances and lead inductances form a tuned circuit at a wavelength well above the limit desired.

(4) The time of transit of the electrons across the space between electrodes becomes an appreciable part of a period which results in a reduction in the effective amplification of the tube.

(5) As the wavelength is reduced the radio-frequency resistance of the circuit is increased with a consequent reduction in resonant impedance.

It will be seen that most of these limitations are associated with too large a ratio of some fixed characteristic of the tube to those characteristics of the circuit which vary with frequency. These characteristics of the tube are fixed only for a given design; however, any change in design which results in lower transconductance, as would be the case with increased interelectrode spacing to reduce capacitance, cannot be considered a genuine improvement.

In a vacuum tube, if all linear physical dimensions are kept in a fixed ratio to each other there will be no change in transconductance, plate current, or amplification factor at fixed operating voltages, no matter what changes are made in the actual magnitude of the linear dimensions. On the other hand, the values of interelectrode capacitance, lead inductance, and time of electron transit are in direct proportion to the magnitude of the linear dimensions.[4]

These considerations lead directly to the principle on which this work is based: for optimum design at any wavelength, all tube and circuit linear dimensions should be in proportion to the wavelength. This principle is modified in practice since, at longer wavelengths, there is no advantage in making the tube of large size and it becomes inconvenient to make the tuned circuit of optimum dimensions.

Unfortunately, it is not to be expected that this proportionality of dimension will result in constant amplification, since the resonant impedance, L/RC, is reduced as the wavelength becomes shorter. However, on the basis of Butterworth's formulas for high-frequency resis-

[4] A statement and proof of part of this may be found in that remarkable paper by Langmuir and Compton, "Electrical Discharges in Gases," Part II, *Rev. Mod. Phys.*, vol. 3, no. 2, p. 252; April, (1931).

tance[5] and Coffin's formula for inductance, with Rosa's correction,[6] a coil having the following dimensions:

$$\begin{aligned} \text{length} &= 1/8 \text{ in.} \\ \text{diameter} &= 1/8 \text{ in.} \\ \text{turns} &= 5 \\ \text{wire diameter} &= 0.020 \text{ in.} \end{aligned}$$

should have a resistance of 0.37 ohm and an inductance of 5.87×10^{-8} henry at 50 centimeters wavelength. A capacitance of 1.2×10^{-12} farad would be required for resonance, giving an impedance, L/RC of 132,000 ohms. This high value in comparison to those obtained at longer wavelengths may be accounted for in part by the fact that the

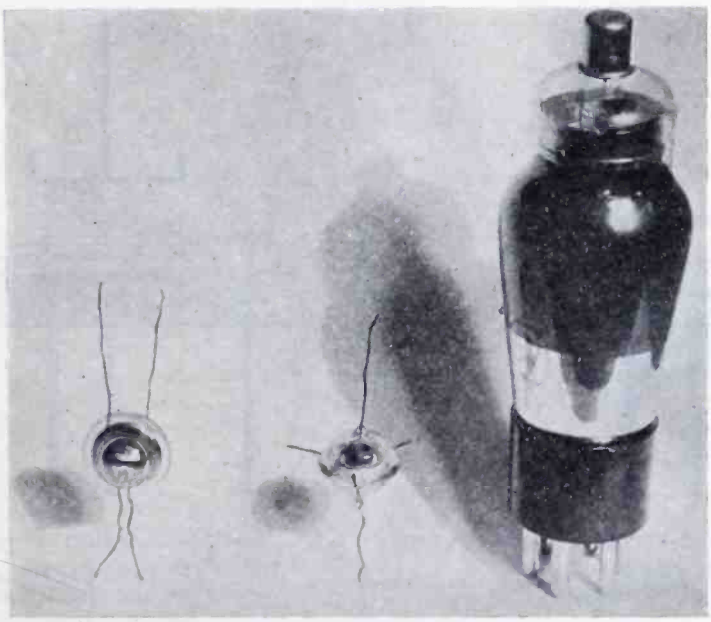


Fig. 1—Small triode and screen-grid tube compared with a conventional size 57 type tube. The triode is shown at the left.

coil is of much more nearly optimum design than those used at the longer wavelengths.

Since tubes of conventional size have been found to have a lower wavelength limit of about five meters, the principle of proportionality requires a tenfold reduction of linear dimension to produce a tube capable of amplification at a wavelength of 50 centimeters.

Both screen-grid tubes and triodes representing such reductions have been constructed and studied in operation.

---

[5] S. Butterworth, *Exp. W. & W. Eng.*, vol. 3, nos. 31, 32, and 34, pp. 203-210; 309-316; 417-424; April, May, and July, (1926).

[6] E. B. Rosa and F. W. Grover, Scientific Paper of the Bureau of Standards No. 169, third edition, pp. 117-122, December, (1916).

Tube Structure

A photograph of these tubes is shown in Figure 1. A conventional size type 57 tube serves as a standard of comparison. The largest dimension of either of these small tubes is less than three quarters of an inch, and the elements themselves are correspondingly small.

Both types of tubes are of parallel plane construction and have indirectly heated cathodes.

In the triode the parts are sufficiently light in weight to permit supporting them on their lead wires alone. This has resulted in the elimination of capacitances which would otherwise be present between the various elements and the support structure. Both plate and cathode

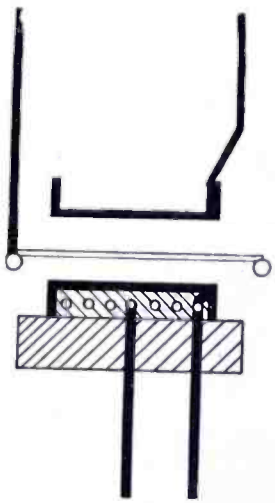


Fig. 2—Cross-section view showing the structure of the small triode.

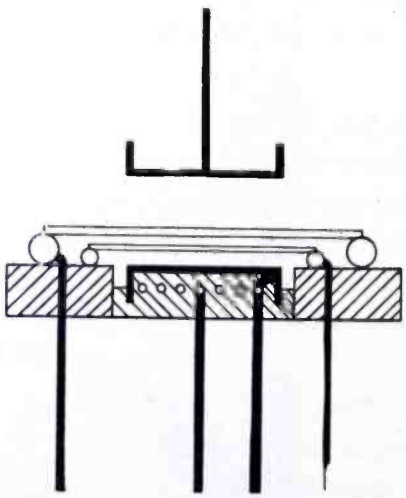


Fig. 3—Cross-section view showing the structure of the small screen-grid tube.

are of the same shape, consisting of two small metal cups placed back to back with the grid interposed between them. The cathode cup has within it a small heater; its outer surface is coated with the emitting material. The grid is of mesh fastened on a support ring. The inter-electrode spacings are only a few thousandths of an inch. Figure 2 shows the general construction.

The assembly scheme of the triode cannot be satisfactorily applied to the screen-grid tube because of mechanical complications arising from the presence of the second grid. A different method is used which is productive of a stronger and more rigid assembly. The tetrode parts, however, are of the same size and shape as those of the triode, with the addition of the screen grid which is similar to the control grid though somewhat larger.

A small ceramic disk serves as a foundation upon which the tube parts, with the exception of the plate, are assembled. It, therefore, acts as a common supporting insulator. The correct spacings between the grids and the cathode are obtained by adjusting their individual distances from this insulator. As the distance from the plate to the screen grid is not so critical, the plate is supported by its lead wire from the glass bulb. The spacings between the other parts is again only a few thousandths of an inch. The general construction is shown in Figure 3.

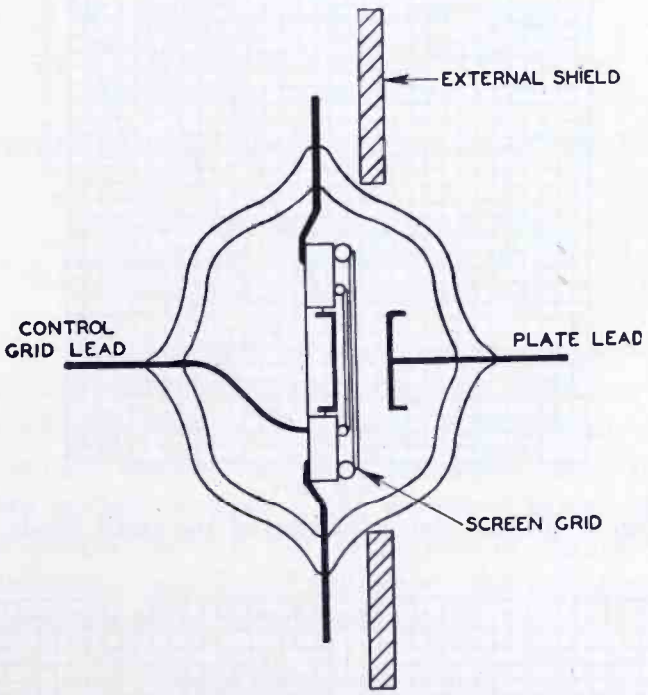


Fig. 4—View showing the screening arrangement used with the small screen-grid tube.

The bulbs used to enclose both types of tubes are in two parts which are more or less hemispherical in shape. These two parts are placed together with the mount inside, and a seal is made between them. All of the triode leads pass through this seal, thereby eliminating the need for a stem as ordinarily used. In the tetrode separate seals are made at opposite ends of the bulb for the plate and control-grid leads while all of the remaining leads come out through the main seal.

In the case of the tetrode this general arrangement has advantages from the standpoint of screening. The mount is so placed in the bulb that its screen grid lies just above the plane of the seal and extends almost to the glass. It is readily seen from Figure 4 that when the external shield is placed as indicated the plate is effectively isolated

from the control grid. The screen-grid lead is quite short and comes out adjacent to the external shield where it can be readily grounded,

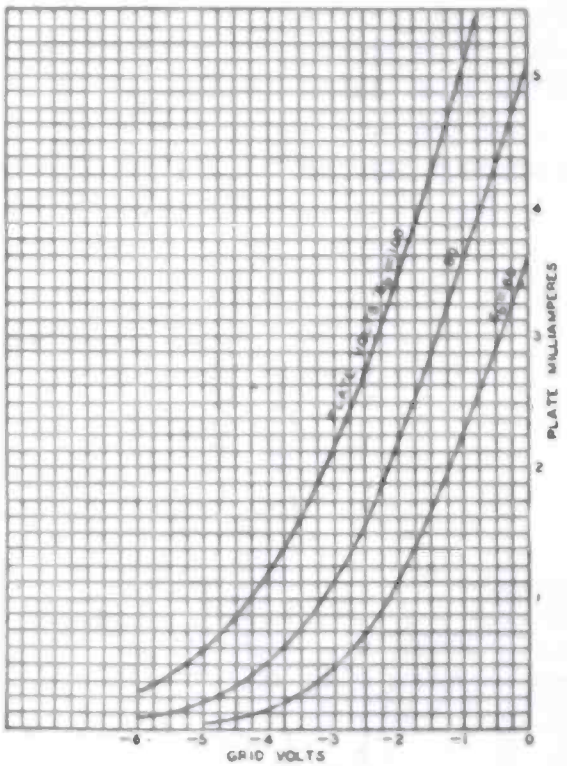


Fig. 5—Mutual characteristics of the small triode.

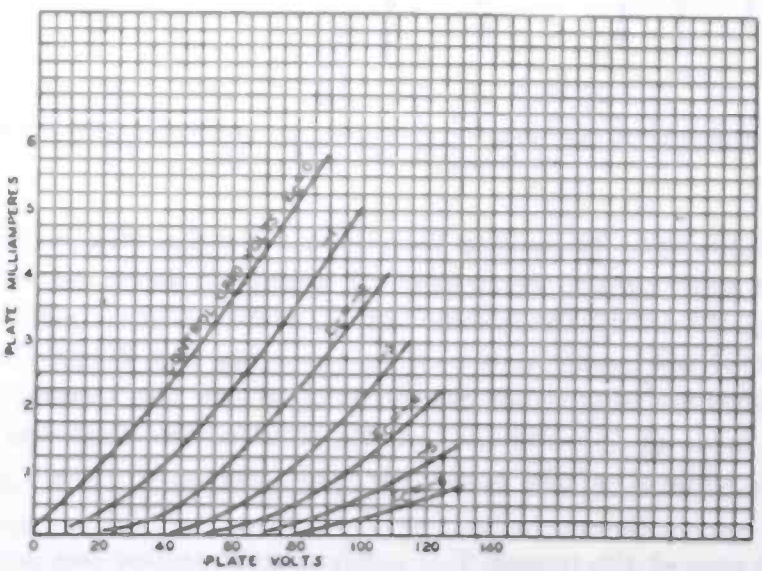


Fig. 6—Plate characteristics of the small triode.

thus minimizing screen-lead impedance. This holds true for the heater and cathode leads likewise.

Electrical Characteristics

From an examination of the static characteristics of the triode, which are shown in Figures 5 and 6, it is readily seen that these characteristics are directly comparable both as to magnitude and shape with those of an ordinary triode. Under the operating conditions, plate voltage = 67.5 volts and grid voltage = −2 volts, the values of the various parameters are as follows:

Plate current = 4 ma
Plate resistance = 9,500 ohms
Transconductance = 1550 μa/v
Amplification factor = 14.7.

The interelectrode capacitances for these tubes have been measured as follows:

Grid-cathode capacitance = 0.7 $\mu\mu$f
Plate-cathode capacitance = 0.07 $\mu\mu$f
Plate-grid capacitance = 0.8 $\mu\mu$f.

As might be predicted from the results of the measurements on the triode, the tetrode characteristics are likewise similar to those of the larger tubes of this sort. A family of plate-current—plate-voltage curves is shown in Figure 7. Points were not taken for the lower values of plate voltage because of the excessive values of screen-grid current. The mutual family of curves is given in Figure 8. Under the operating conditions, control-grid voltage = −0.5 volt, screen-grid voltage = 67.5 volts, and plate voltage = 135 volts,

Plate current = 4.0 ma
Transconductance = 1100 μa/v
Plate resistance = 360,000 ohms
Amplification factor = 400.

The values of the interelectrode capacitances are:

Input capacitance = 2.5 $\mu\mu$f
Output capacitance = 0.5 $\mu\mu$f
Plate-grid capacitance = 0.015 $\mu\mu$f.

Operation

Tests have been made upon both the triodes and the screen-grid tubes to determine how well they will perform in conventional circuits at wavelengths much lower than the minimum at which ordinary tubes will function.

The minimum wavelength at which a triode will generate oscillations offers a means for comparing it with ordinary tubes. The value

of this minimum wavelength of oscillation is of particular interest here because it shows how closely a normal feed-back oscillator can

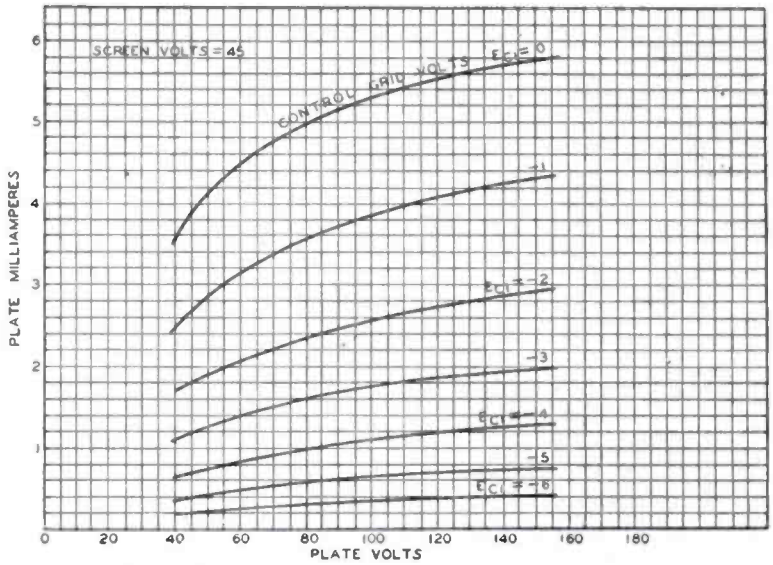


Fig. 7—Plate characteristics of the small screen-grid tube.

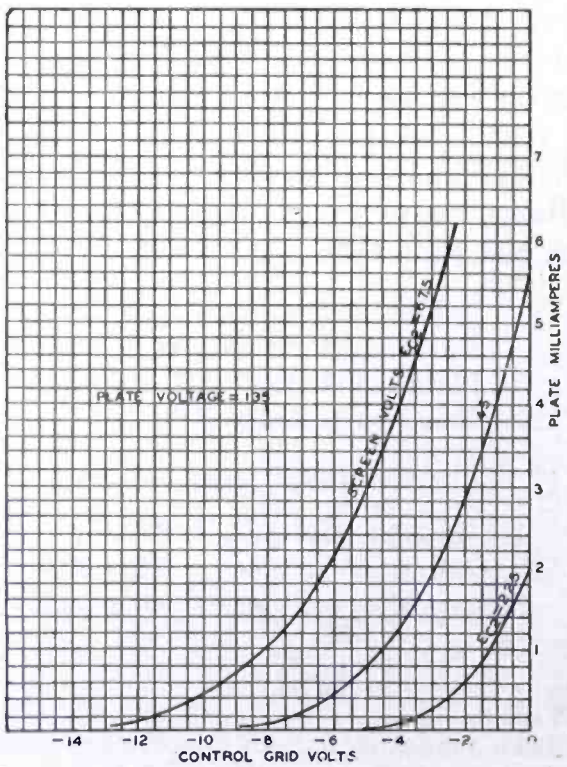


Fig. 8—Mutual characteristics of the small screen-grid tube.

approach those wavelengths generated almost solely by Barkhausen tubes and circuits.

An inductive feed-back oscillator was set up whose inductance consisted of several turns of small copper wire wound in a solenoid about

one-eighth of an inch in diameter tuned only by the tube interelectrode capacitances. The circuit is given in Figure 9, while a photograph of

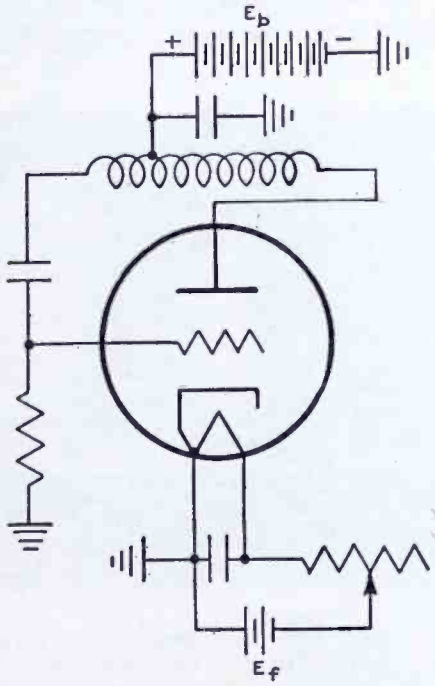


Fig. 9—Circuit diagram of the ultra-high-frequency oscillator using the small triode.

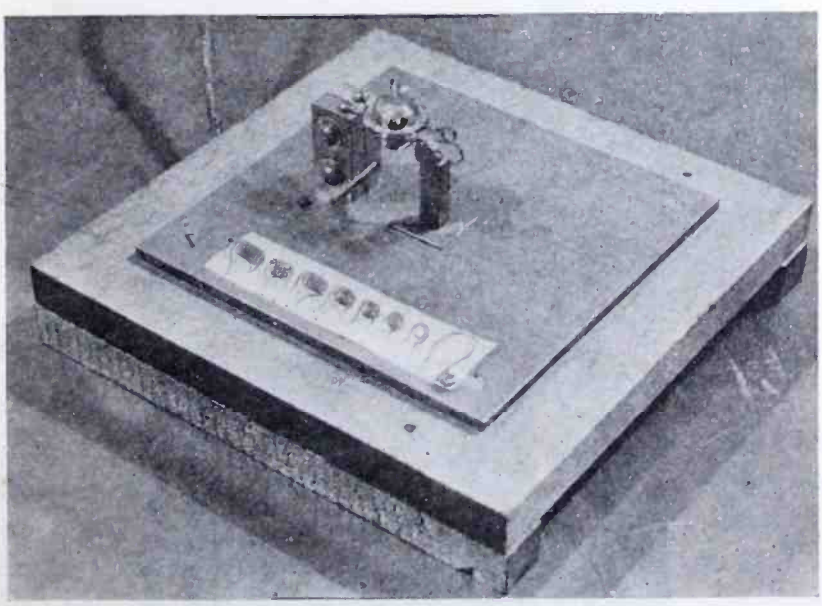


Fig. 10—Photograph of the ultra-high-frequency oscillator. Several tuning coils for different wavelengths are shown.

the oscillator is shown in Figure 10. With a coil of six turns very stable 65-centimeter oscillations were produced with as low as 45 volts on the plate of the tube. Smaller coils gave shorter wavelengths with

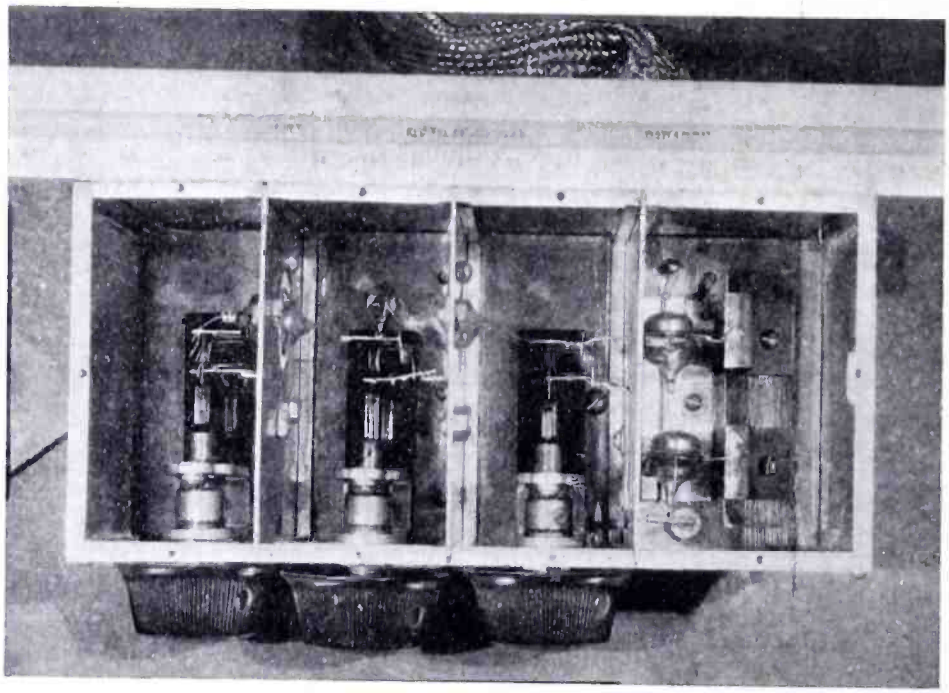


Fig. 11—Photograph of a tuned radio-frequency receiver for a wavelength of 100 centimeters.

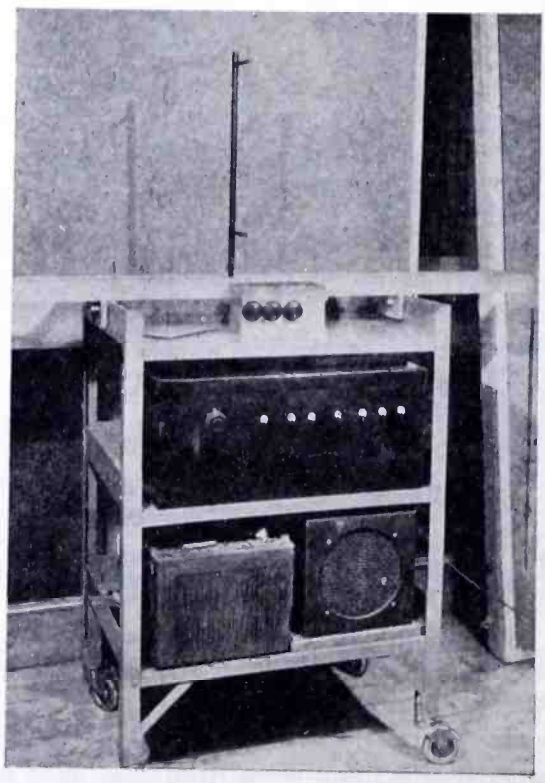


Fig. 12—Photograph of the complete 100-centimeter receiver arrangement.

continued stability until a minimum wavelength of slightly below 30 centimeters was reached with a coil of only one turn. Oscillations at this wavelength could be sustained with as low as 115 volts on the plate of the tube and with a plate current of approximately 3 milliamperes.

Owing to the difficulty of making quantitative measurements of radio-frequency amplification at wavelengths of one meter and less, the gain realizable by the use of the screen-grid tubes was determined by their operation in actual receiving sets. The first set consisted of two stages of tuned radio-frequency amplification, a detector, and one stage of audio-frequency amplification. The screen-grid tubes were

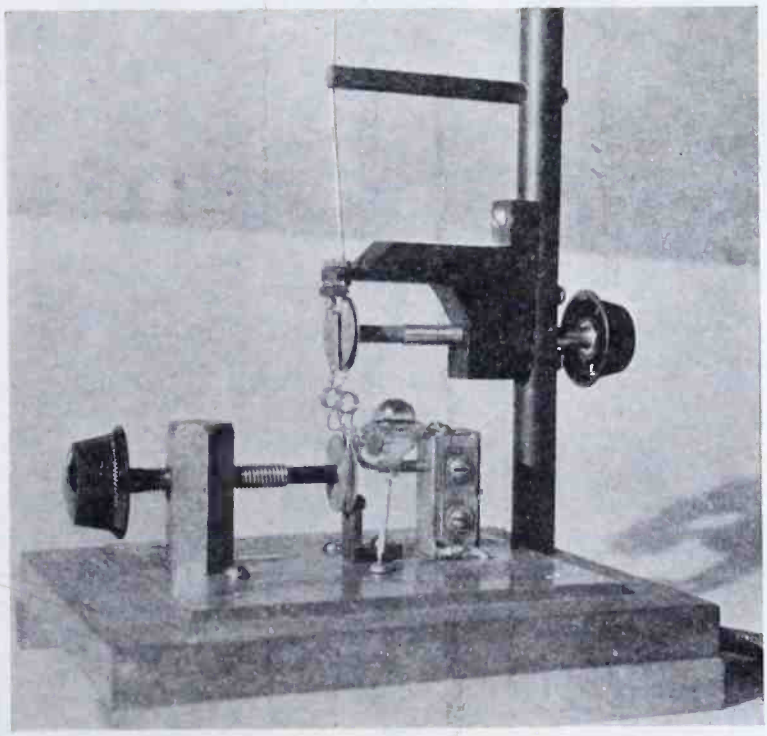


Fig. 13—Photograph of the 100-centimeter oscillator.

used in the radio-frequency amplifier stages and the small triodes as the detector and audio amplifier. The whole set, of which photographs are shown in Figures 11 and 12, was enclosed in a brass box seven inches long, three inches high, and three inches wide. Small coils such as those used in the oscillator tuned by almost equally small variable condensers constituted the tuned circuits. In order to prevent any signal pick-up except through the antenna, the batteries and all external leads were enclosed in metal shielding. With the set so shielded there was no trace of oscillation in any of the circuits. The tuning range of the receiver was from about 95 to about 110 centimeters.

An oscillator operating at a wavelength of 100 centimeters, consisting of one of the small triodes modulated by a broadcast receiver and loosely coupled to a half-wave radiator, was set up in an open area. The total plate power supplied to the oscillator was 68 milliwatts. Photographs of this transmitter are shown in Figures 13 and 14.

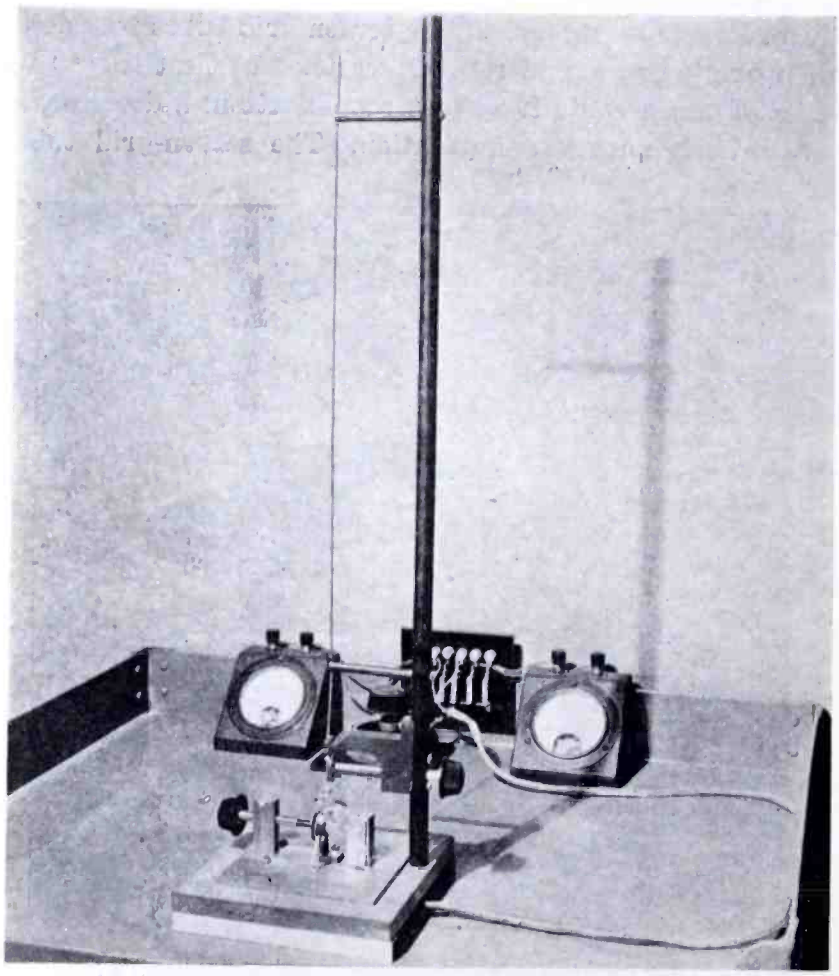


Fig. 14—Photograph of the 100-centimeter transmitter.

With the receiver located about 200 feet from the transmitter, signals of good strength were received with the half-wave receiving antenna coupled to the input of the first radio-frequency stage but none could be heard with the antenna coupled directly to the detector. From other listening tests it was estimated that the gain per stage was of the order of four.

The second receiver was constructed to operate at 75 centimeters or thereabouts. This set was not so elaborate as the one previously described, but more care was taken in placing the tubes so that all

circuit connections would be shorter than before. It consisted of one stage of radio-frequency amplification and a grid-leak detector. As before, the set was enclosed in a small brass box and completely shielded. Figure 15 shows a photograph of this receiver.

Inasmuch as it was desired to obtain as high a value of input and coupling circuit impedances as possible, tuning condensers were eliminated and use was made of the tube interelectrode capacitances only. This made it necessary to fix the tuning of the set. The initial tuning necessary to line up the amplifier and detector circuits at approxi-

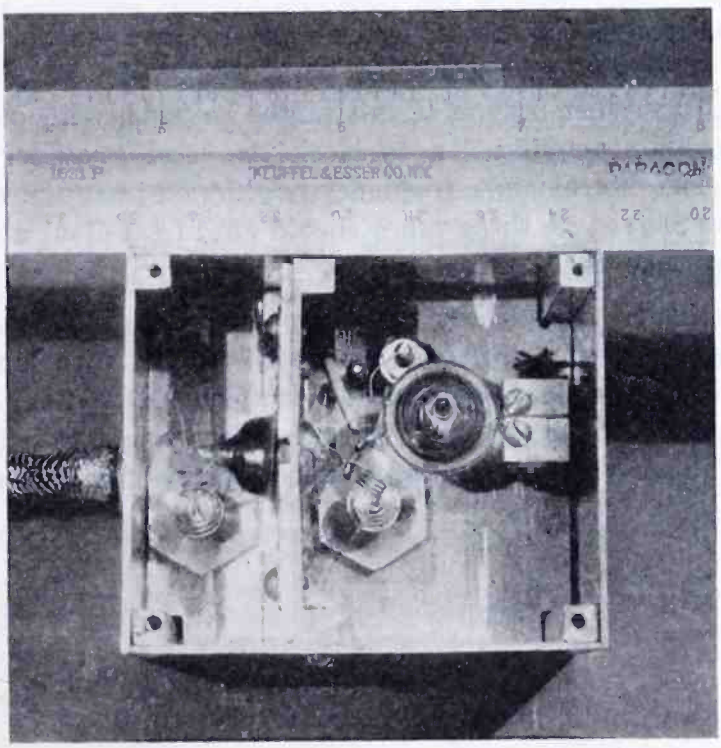


Fig. 15—Photograph of the tuned radio-frequency receiver for a wavelength of 75 centimeters. The scale is marked in inches.

mately 75 centimeters was accomplished by changing the turn spacing of the tuning coils, thereby varying their inductances. The frequency of the transmitter was adjusted by means of a variable condenser to bring it into tune with the receiver.

Following much the same procedure as before, except that the distance from oscillator to receiver was less and the tests were carried out in a large shielded room, the receiver output when the antenna was coupled to the radio-frequency stage was compared to its output when the antenna was coupled directly to the detector. Again these comparisons were qualitative rather than quantitative. While the contribution of the radio-frequency stage was found to be small, it did

furnish some gain as evidenced by the increase in output when it was operating.

Conclusion

While no claim is made to optimum design of either tubes or circuits, it has been demonstrated that it is possible to produce tubes of small physical dimensions with characteristics which permit radio-frequency amplifiers, oscillators, and detectors to be used at wavelengths well below one meter in the conventional manner. It may be of interest to consider the significance of these results.

A sensitive, compact, and economical receiver should be only a problem of design in accordance with well-known principles. For example, a superheterodyne circuit might be used, with one stage of radio-frequency amplification to block the local oscillator from the antenna. The intermediate frequency might well be in the range from two to five meters, as these tubes should afford excellent amplification at such wavelengths.

In conclusion the authors wish to point out that the tubes which have been described were made in the laboratory with the object of demonstrating certain fundamental principles, rather than of producing a commercial tube design. However, it is hoped that these principles will be of value in the future development of special short-wave tubes.

SIMPLE ANTENNAS AND RECEIVER INPUT CIRCUITS FOR ULTRA-HIGH FREQUENCIES

By

R. S. Holmes and A. H. Turner

RCA Manufacturing Company, Inc., Camden, N. J.

Summary—This paper describes some of the factors affecting the performance of antennas, transmission lines and receiver input circuits. The effect of antenna selectivity, directivity and gain is discussed, and selectivity curves for some simple antenna arrangements are given. Some of the factors governing selectivity, gain and signal-to-receiver hiss ratio are discussed. The discussion covers both narrow and wide band circuits.

Simple Ultra-High-Frequency Receiving Antennas

ANTENNAS for ultra-high frequencies are usually complete within themselves so that the earth does not form part of the antenna circuit. However, the strengths of the received signals are still influenced by the presence of the earth insofar as it obstructs direct propagation or provides additional indirect paths by reflection. For best reception it is usually desirable to support the antenna high above the earth or surrounding obstructions.

The three general types of antennas most frequently used for u.h.f. are:

1. Half-wave dipoles and dipole arrays.
2. Tilted wires and tilted wire arrays.
3. Open and closed circuit loops.

The half-wave dipole is more popular than any other type of u-h-f antenna because it is simple to construct and install, and because it receives more energy than other types of similar dimensions. It is used as the standard of comparison for all other antenna types. In its simplest form the dipole consists of a straight wire or rod slightly shorter than one-half of the wavelength of the desired signal. The transmission line is usually connected in series with the dipole at a center opening. However, a shunt connection may be made by connecting the transmission line across a part of the continuous conductor. This shunt connection has the disadvantage that the triangular-shaped loop formed by the fanned line and the included part of the dipole distorts the normally symmetrical dipole directivity pattern. The directivity pattern of a dipole antenna follows the cosine law, with minimum pickup in the direction along the antenna axis.

To obtain more signal voltage and at the same time greater directivity, another dipole slightly longer than the antenna is often placed behind the antenna to serve as a "reflector" of the received energy. Additional signal voltage and directivity may be obtained by placing still another dipole slightly shorter than the antenna out in front to act as a "director" of the received energy.

The selectivity characteristic of a simple dipole loaded with a matched transmission line is shown in Figure 1. A reflector increases the slope of the selectivity characteristic on the low-frequency side of resonance and a director similarly steepens the high-frequency side.

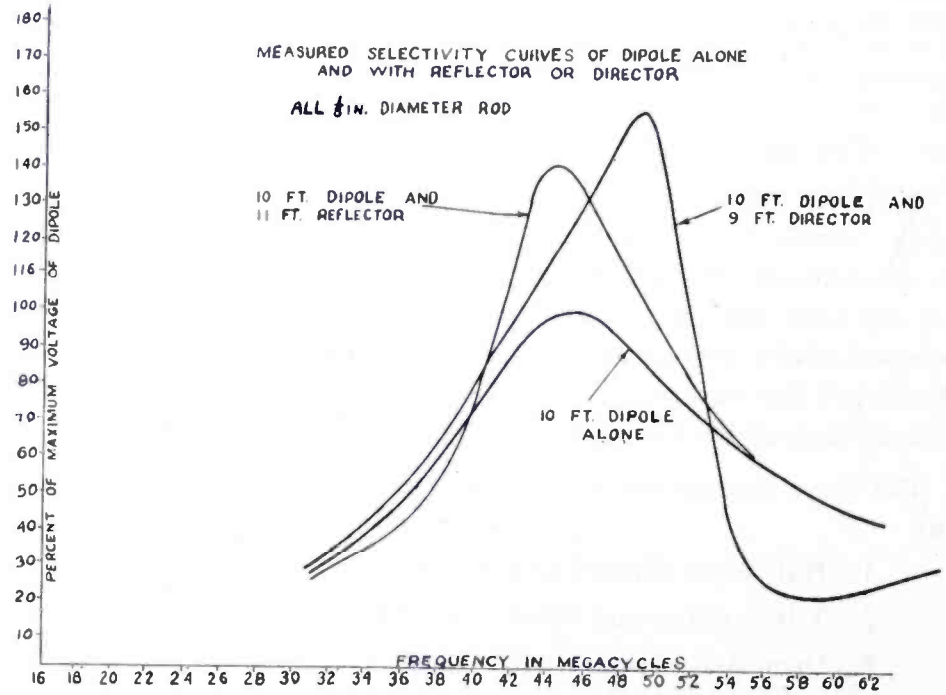


Fig. 1.

There are some applications, such as in television services, where antenna selectivity is undesirable. Therefore, antennas for these services should be designed to have a minimum of reactance and a maximum of radiation resistance. Increasing the conductor size of dipole antennas reduces their reactance appreciably without much loss of radiation resistance. Inasmuch as the decrease in reactance results from a decoupling of the current filaments as these filaments are spread over a larger surface, the same effect can be approximated by replacing the continuous conductor surface with small parallel-spaced wires. This open-grid construction of low-reactance dipoles both reduces the wind resistance and saves on conductor material.

The two RCA commercial television-receiver antennas illustrated in Figure 2 are single dipoles each consisting of two spaced conductors lying in the same vertical plane and connected in parallel. Antenna A is shown fitted with reflector dipoles to be used where greater signal voltage and directivity are required. The two lower curves of Figure 3 indicate the wider frequency response obtained by fanning the simple dipole as is done in antenna B of Figure 2. Spreading the conductors as in antenna A makes the response even wider.

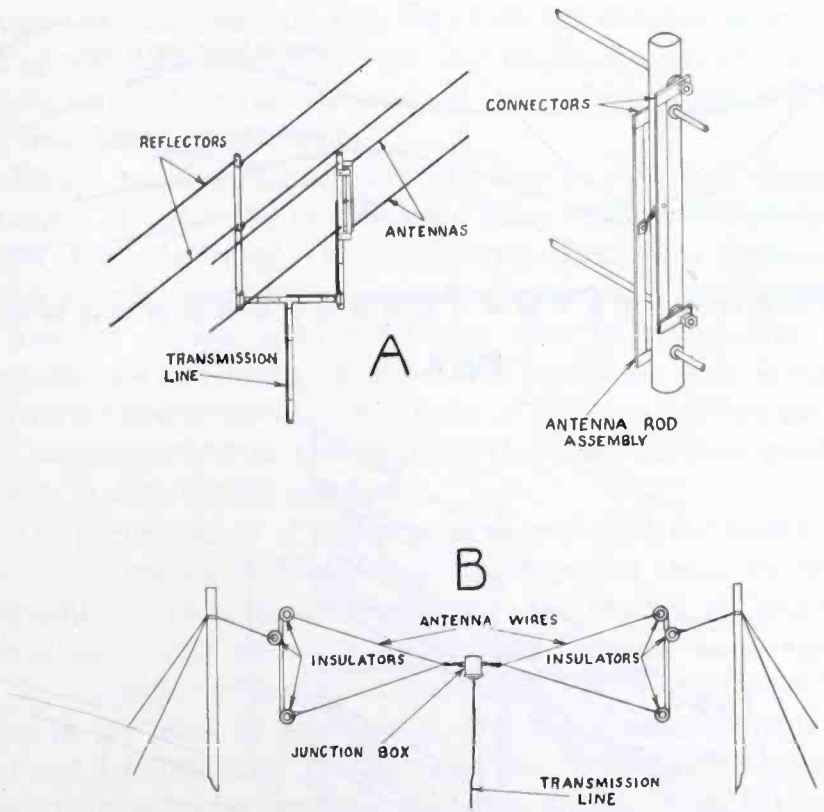


Fig. 2.

At television receiver locations where the single dipole and reflector do not provide sufficient signal, four low-reactance elements, each one-half wave long, may be combined in series-parallel. As shown in Figure 3, this combination of four elements increases the received energy to approximately 2½ times that of a single element without increasing the selectivity. A reflector grid behind the array approximately doubles the received energy, the same as in the case of a single dipole.

An array of this type consisting of four half-wave fans and reflector was installed atop the 250-foot RCA antenna tower at the New York World's Fair. Figure 4 shows the antenna conductor arrangement without the complications of supporting members.

The vertical center-to-center spacing of the fans is usually made equal to one-half of the wavelength of the most desired signal for several reasons:

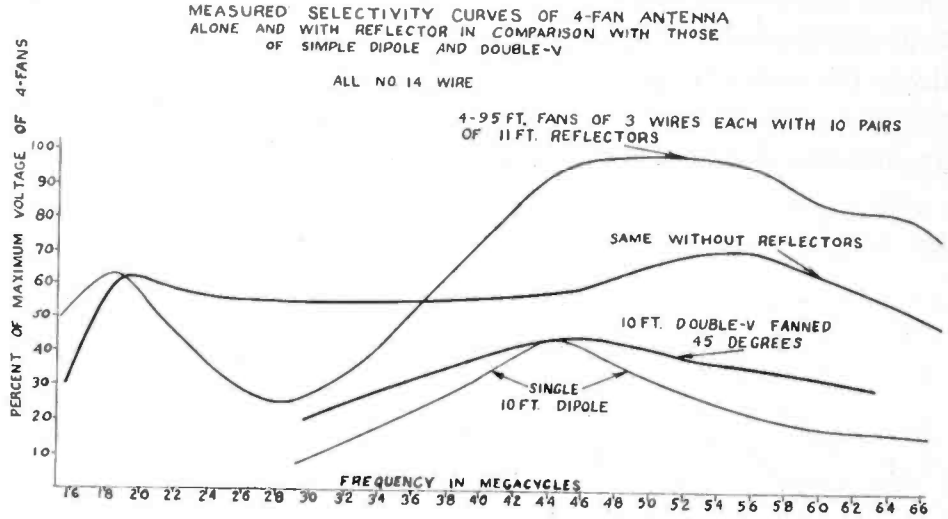


Fig. 3.

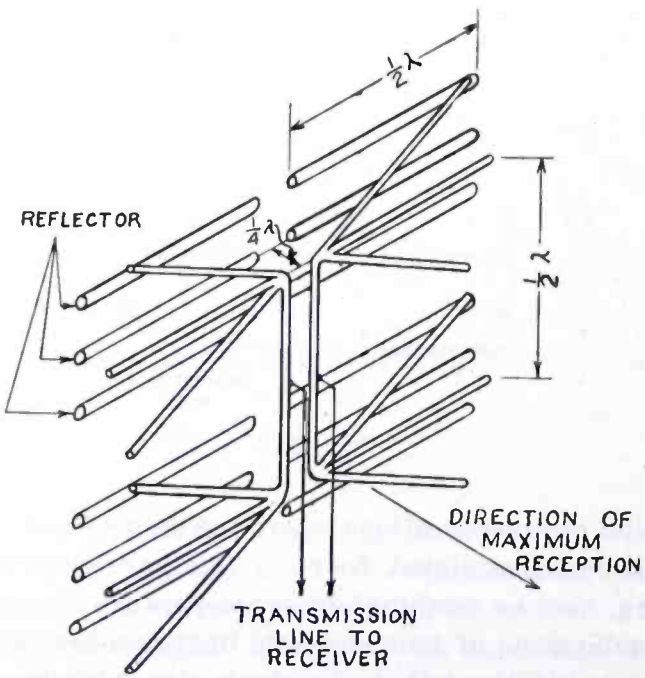


Fig. 4.

1. Interference arriving from directly below or directly above is cancelled in the antenna.
2. The two vertical one-quarter wavelength branch lines connecting the main transmission line to each pair of fans can serve as impedance transformers when necessary.

3. The horizontal and vertical dimensions in the plane of the antenna are nearly equal and consequently the horizontal and vertical directivities are also nearly equal.

The apex-to-apex impedance of one pair of fans with reflector is approximately 750 ohms at resonance. There are two such impedances in parallel across the main transmission line. Therefore, no impedance transformation is required by the branch lines if the main transmission line has an impedance of 375 ohms. It is interesting to observe that fanning the elements not only broadens the antenna selectivity curve, but at the same time brings the end impedance down to a value that can be matched by the conventional open-wire transmission line without impedance transformers.

Where more antenna selectivity can be tolerated, dipoles can be arranged in series and in parallel in many other combinations for high signal gain. However, these combinations require phasing elements which greatly restrict the useful frequency range of a given array.

Leaving dipoles and considering now the tilted-wire family of antennas, we find that in general they have high gain and directivity and very little selectivity. But to be of much value they must be several wavelengths long. Consequently they have not been generally used except in commercial services.

The correct angle of tilt or slope is such that the length of wire is one-half wavelength greater than its projection along the direction of propagation. Thus the longer the wire the smaller the angle the wire should make with the direction of wave propagation in space.

Two sloping wires may be combined to form a balanced V-antenna lying in the plane of polarization, and two V-antennas may be joined to form a diamond or rhombic pattern. With the far ends of these tilted-wire antennas open, they receive signals from the back nearly as well as from the front because the energy from one direction travels out to the open ends and is reflected back to the receiver. Resistive termination may be applied to the open ends if reception from the back direction is undesired.

In England the phase velocity along tilted wires has been increased by inserting many capacitors at regular intervals along the wire. This permits a greater angle of tilt and consequently exposes the wire to a longer wave-front.

Loop antennas have not been generally used for u-h-f reception because they offer no advantages over dipoles. For a given resonance frequency, the loop effective height (and induced voltage) decreases as the number of turns increases. Thus, the convenience of small compact loop antennas can not be had without a considerable loss of

signal voltage. Furthermore, locating a small loop antenna at the receiver will usually be found to provide much less signal voltage than placing the same antenna above the roof.

U-H-F Transmission Lines

Several types of transmission lines are available to transfer the signal energy from the antenna to the receiver. Of these the rubber-insulated twisted-pair lines are the most generally used because they are inexpensive and easy to install. At 50 megacycles the losses in these lines vary from about 4 to 10 decibels per hundred feet, the most expensive having the best rubber and the least loss. Losses of a typical rubber-insulated twisted pair intended for use with all-wave receivers are 2.1 and 9.5 decibels per hundred feet at 10 and 66 megacycles, respectively.

For the rubber-insulated lines the dielectric losses and conductor losses are of the same order of magnitude at ultra-high frequencies. The conductor loss is inversely proportional to the surge impedance of the line. All solid dielectrics have more loss than air. Therefore, for least loss the lines should have little solid dielectric and wide spacing between conductors (high impedance), with a limit at about 5 per cent of the wavelength to keep line radiation negligible. This, of course, means that impedance transformers will be required if the terminating impedances are appreciably lower than the line. Quarter-wave sections of line are frequently used as impedance transformers, but have the disadvantage that they introduce frequency selection. A less selective arrangement is to distribute the impedance transformation over the entire length of the line by varying the line spacing exponentially, although the type of taper is not very important as long as there are no abrupt discontinuities along the line.

Coaxial transmission lines with air or solid low-loss dielectric are used for ultra-high frequencies in special applications, but are usually more expensive than open-spaced wires having the same loss.

U-H-F Receiver Input Circuits

The major function of the receiver input circuit is to step up the voltage of the incoming signal from the antenna transmission line to the grid of the first tube and to provide selectivity against unwanted responses and overload from unwanted signals.

For the relatively low modulation frequencies such as are used for voice and facsimile transmission, single-tuned input circuits may be used without added circuit loading. This applies to reception of

any signal where the total frequency swing or band is not greater than two or three hundred kilocycles. But for television reception, it is often desirable to pass both audio and video signals of each channel through a single station selector. This may require a pass-band of more than 4 megacycles, which may be obtained by using double-tuned closely-coupled circuits, properly loaded.

The importance of obtaining maximum voltage gain to the grid of the first tube in ultra-high-frequency receivers should be especially stressed because it has a direct bearing on the ratio of signal to internal receiver noise at low signal levels. At lower frequencies the impedance in the grid circuit of the first tube is usually so high that thermal-agitation hiss, amplified through the first tube, exceeds the random hiss of the space current in its plate circuit. But in most u-h-f receivers, particularly for television, the first-tube grid-circuit impedance is so low that the space-current hiss predominates. This means that all gain in signal voltage up to the plate circuit of the first tube gives a nearly proportionate improvement in signal-to-hiss ratio. The first-tube grid circuit should, therefore, have an impedance limited only by tube-input and line-termination losses and should have a high L/C ratio. This means that the conventional variable-capacitor tuners used at lower frequencies may advantageously be replaced at ultra-high frequencies by inductance tuners having much higher reactance.

Inductance tuners used to obtain maximum grid-circuit impedance may be placed in three groups:

1. Continuously variable types, tunable by
   - (a) Changing the effective length of wire in the circuit with sliding contacts.
   - (b) Changing the permeability of the core.
   - (c) Coupling a shorted turn to the coil.
2. Switch types
   - (a) Switching coil taps.
   - (b) Switching in different coils.
3. Continuous coverage, but not continuous in operation
   - (a) Combinations of No. 1 and No. 2.

The continuously variable sliding-contact type has mechanical and electrical disadvantages. The variable magnetic-core tuners operate smoothly as far as they go, but the range is restricted by the variation in permeability obtainable. This also applies to the shorted-turn

tuners which are somewhat inferior electrically to the magnetic-core tuners because they require more wire for a given maximum inductance.

The switch type is simple and reliable, but is not continuously variable. However, for the many applications where adjacent or widely separated bands of frequencies are to be received, as in television, the switch tuner is admirably suited.

Two typical television tuners of the latter type are shown in Figure 5. The model A switches two sets of coupled-circuit coils and a set of oscillator coils to five television frequency channels. This arrangement requires more space than would three coils with five taps, but permits individual adjustment at each frequency by means of movable powdered-iron cores.

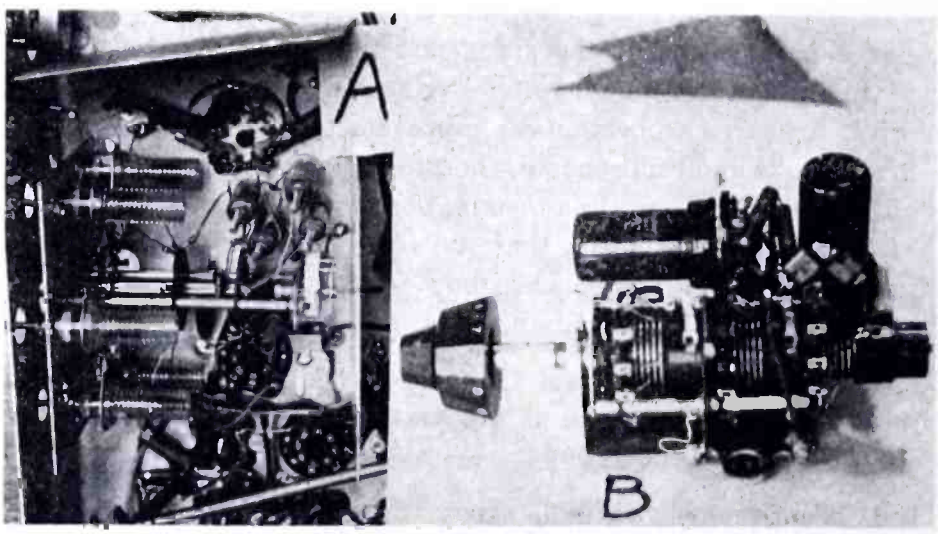


Fig. 5.

The model B in Figure 5 is of the tapped-coil variety with the three coils mounted coaxially and surrounding the switch control shaft and adjacent to the switch wafers. This model provides for tuning to twelve selected channels, each 6 megacycles wide, between the frequency limits of 38 and 108 megacycles. The oscillator coil is on the smaller form at the rear and contains a movable powdered-iron core operated by an inner control shaft with coaxial knob at the front. The iron core varies the oscillator frequency 8 to 10 megacycles at each channel, thus giving the station selector overlapping coverage for the assigned television stations and random sound stations over the range 38 to 108 megacycles.

There is another important difference between the models of Figure 5. Model B has a single untuned low-impedance primary tightly coupled to the first tuned circuit, the turns of which are shorted by

the tap switch. The primary has been placed in such a position that its impedance remains practically constant for all channels. Model A of Figure 5 contains a separate tuned primary for each channel. This latter arrangement is more flexible, but requires more space.

Oscillator frequency stability is an important problem in u-h-f receivers, particularly if the intermediate-frequency amplifier must have a very narrow pass-band. Oscillator circuit Q is lower and tube space-charge effects are more detrimental than at lower frequencies. The two typical television station selectors of Figure 5 have been used to feed sound intermediate frequencies into 9 megacycle amplifiers having pass-bands 50 to 100 kilocycles wide. Model B has a plate-tuned oscillator circuit with considerable impedance step-down to the

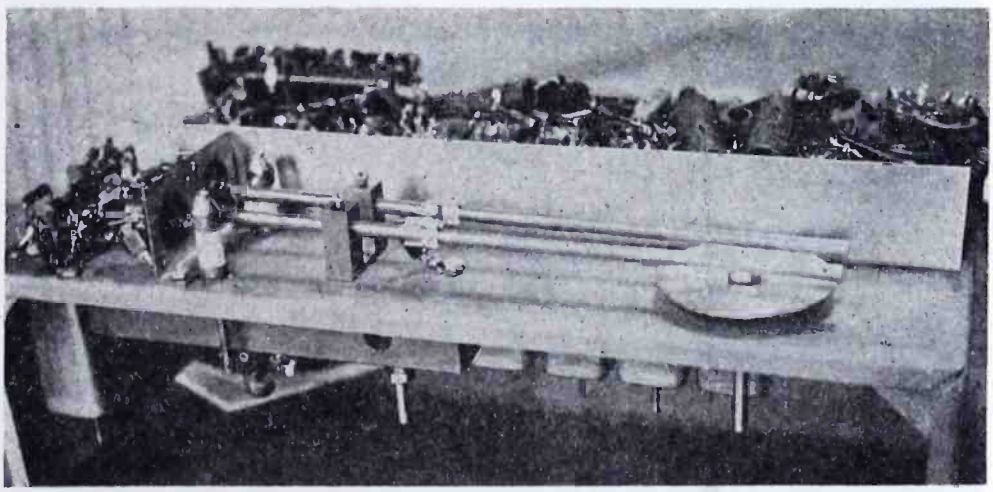


Fig. 6.

grid. The frequency change during the first few minutes of warm-up is appreciable and occurs because the tube-plate capacitance is too large a part of the total circuit capacitance. Model A has better frequency stability at the expense of oscillator output voltage because an 80-micromicrofarad capacitor is permanently connected across the oscillator circuit.

The RCA type 6J5 triode is used as the local oscillator in both of these station selectors and easily reaches 120 megacycles in Model B. For higher frequencies the RCA type 955 acorn-size tube is more suitable.

Station selection and oscillator circuits for frequencies above 200 megacycles are usually made with distributed instead of lumped constants, partly to obtain lower resistance and partly to obtain circuits with larger physical dimensions for easier adjustment. The input circuits of a receiver for 342 megacycles are shown in Figure 6 and

the oscillator circuit in Figure 7. The parallel rods form two hairpin circuits coupled by the reactance of the supporting posts which are common to both circuits. These circuits are essentially fixed tuned for one channel 12 megacycles wide.

The oscillator circuit consists of a seven-inch length of one-inch square tubing with a slightly shorter length of ¼-inch diameter tubing supported inside by a block at one end. The oscillator plate lead enters the inner tubing near its free end and comes out at the low-potential end. The cathode and one heater lead connect to the inner tubing approximately one-fourth of its length from the low-potential end. The other heater lead enters the inner tubing at the same point. The

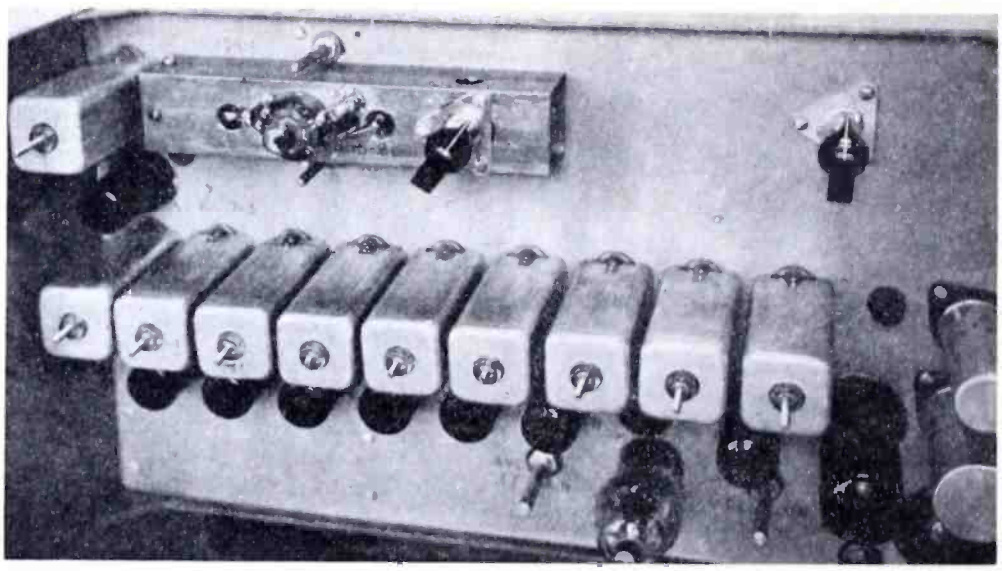


Fig. 7.

oscillator grid connects to ground (square tubing) through grid leak and capacitor.

Oscillator voltage is applied to the push-pull converter tubes (RCA type 954's) by connecting their paralleled cathodes to a low-potential point on the oscillator inner conductor. Accordingly, if the tube capacitances are well balanced very little oscillator energy gets to the receiving antenna by way of the push-pull converter grid circuit.

Perhaps just as important as high grid-circuit impedance for maximum signal-to-hiss ratio are the characteristics of the first tube in the receiver. The best first tube is not necessarily the tube capable of giving the greatest voltage gain, but rather the tube with the best gain per milliampere of space current.

The signal-to-hiss ratio for any type of tube is usually better when the tube is operated as an amplifier rather than as a frequency con-

verter. Also, a radio-frequency amplifier stage reduces oscillator radiation considerably by keeping the oscillator energy from being fed back into the receiving antenna.

Oscillator radiation from u-h-f receivers without r-f stages is more serious than that from conventional low-frequency receivers because of the converter arrangement used to obtain the best possible signal-to-hiss ratio. At lower frequencies it is customary to apply oscillator and signal voltages to different converter grids which are usually shielded from each other so that oscillator energy cannot get to the antenna by way of the signal grid circuit. But for u-h-f receivers where converter efficiency is at a premium, better performance can be obtained by applying both oscillator and signal voltages to the same grid.

PART II

MAGNETRON OSCILLATORS FOR THE GENERATION OF FREQUENCIES BETWEEN 300 AND 600 MEGACYCLES*

By

G. R. KILGORE

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary—*The need for vacuum tube generators capable of delivering appreciable power at frequencies from 300 to 600 megacycles is pointed out and the negative resistance magnetron is suggested as one of the more promising generators for this purpose.*

An explanation of the negative resistance characteristic in a split-anode magnetron is given by means of a special tube which makes possible the visual study of electron paths. In this manner it is demonstrated how most of the electrons starting toward the higher potential plate reach the lower potential plate.

From the static characteristics it is shown how the output, efficiency, and load resistance can be calculated, and from this analysis it is concluded that the negative resistance magnetron is essentially a high efficiency device at low frequencies.

Measurements of efficiency at ultra-high frequencies are given for several magnetrons under various operating conditions. It is concluded from these measurements that the decrease of efficiency at very high frequencies is mainly due to electron-transit-time effects. A general curve is given showing efficiency as a function of the "transit-time ratio." This curve indicates that for a transit time of one-fifteenth of a period, approximately fifty per cent efficiency is possible; for one-tenth of a period, thirty per cent; and for one-fifth of a period, the efficiency is essentially zero.

Two methods are described for increasing the plate-dissipation limit. One method is that of increasing the effective heat-dissipating area by the use of an internal circuit of heavy conductors. The other method is that of a special water-cooling arrangement which also makes use of the internal circuit construction.

Examples of laboratory tubes are illustrated, including a radiation-cooled tube which will deliver fifty watts at 550 megacycles and a water-cooled tube which will deliver 100 watts at 600 megacycles.

I. INTRODUCTION

THE demand for more ultra-high-frequency channels has necessitated the development of generators for frequencies of higher and higher order. This development has progressed in two directions; the extention of the upper frequency limit of conventional oscillators and amplifiers, and the investigation of other types of generators

* Decimal classification: R355.9. Presented before I.R.E. Tenth Annual Convention, Detroit, Michigan, July 1, 1935.

Reprinted from *Proc. I.R.E.*, August, 1936.

especially adapted to ultra-high frequencies, such as Barkhausen-Kurz oscillators,[1] and magnetron oscillators.

A review of the work done shows a tendency to concentrate on the second line of development with an emphasis on obtaining the very highest frequencies rather than on obtaining appreciable power at frequencies just above the limit of conventional tubes. As a result little work has been done until very recently towards the generation of power at frequencies between 300 and 600 megacycles.

At the same time the advancement of the receiving tube art with the introduction of the "acorn" type tube[2,3] has made it possible to build practical receivers for frequencies somewhat above 300 megacycles. This fact brings nearer the practical utilization of these frequencies and makes it more important to obtain satisfactory generators.

While considerable progress has been made in extending the usefulness of the feed-back oscillator above 300 megacycles, the possibilities of other means of generation cannot be disregarded. One of the less conventional means which shows promise from the standpoint of output and efficiency is the magnetron oscillator.

Magnetron oscillators for generation of ultra-high frequencies can be classed as "electronic oscillators"[4,5,6,7,8,9] and "negative resistance" oscillators,[10] the former being of little importance in the frequency range under consideration. However, for the sake of clearness both types will be defined.

---

[1] H. Barkhausen and K. Kurz, "Shortest Waves Obtainable With Valve Generators," *Phys. Zeit.*, Vol. 21, pp. 1-6; January (1920).

[2] B. J. Thompson and G. M. Rose, Jr., "Vacuum Tubes of Small Dimensions For Use at Extremely High Frequencies," *Proc. I.R.E.*, Vol. 21, pp. 1707-1721; December (1933). (See Page 334).

[3] B. Salzberg and D. G. Burnside, "Recent Developments in Miniature Tubes," *Proc. I.R.E.*, Vol. 23, pp. 1142-1157; October (1935).

[4] This type of magnetron oscillator was first described in the literature by Zacek[5] in 1924, and was later discussed in papers by Okabe,[6] Yagi,[7] Kilgore,[8] Megaw,[9] and others. It is sometimes referred to as a "Magnetostatic oscillator."[8]

[5] A. Zacek, "A Method of Generating Short Electromagnetic Waves," *Casopis pro Pestovani Mathematiky a Fysiky* (Prague), Vol. 53, p. 378; June (1924); (summary in *Zeit. für Hochfrequenz.*, Vol. 32, p. 172, (1928).

[6] K. Okabe, "Ultra-Short Waves from Magnetrons," *Jour. I.E.E.* (Japan), p. 575, June (1927).

[7] H. Yagi, "Beam Transmission of Ultra-Short Waves," *Proc. I.R.E.*, Vol. 16, pp. 715-740; June (1928).

[8] G. R. Kilgore, "Magnetostatic Oscillators for Generation of Ultra-Short Waves," *Proc. I.R.E.*, Vol. 20, pp. 1741-1751; November (1932).

[9] E. C. S. Megaw, "An Investigation of the Magnetron Short-Wave Oscillator," *Jour. I.E.E.* (London), Vol. 72, pp. 326-348; April (1933).

[10] Otherwise referred to as a "dynatron magnetron"[9] and as a "Habann generator."

An *electronic magnetron* oscillator can be defined as one which operates by reason of electron-transit-time phenomena and in which the frequency is essentially determined by the electron-transit time. Although this type of oscillator is capable of generating the very highest frequencies obtainable with vacuum tubes, it has an inherently low efficiency (approximately ten per cent) and a very limited output. At frequencies between 300 and 600 megacycles, the possible output is much smaller than that obtainable from the negative resistance magnetron.

A *negative resistance* magnetron oscillator is defined as one which operates by reason of a static negative resistance between its electrodes and in which the frequency is equal to the natural frequency of the circuit. In its usual form it consists of a cylindrical plate and coaxial filament, the plate being split into two or more segments. Both the two-segment and the four-segment[11,12,13,14] form are being used with success, but in this paper the discussion will be limited to the two-segment type.

The basic idea of the negative resistance magnetron was disclosed by Habann[15] in 1924. Since that time a number of papers on the subject have appeared; notably those of Spitzer and McArthur,[16] Megaw,[9] and Slutzkin[17] and his associates. Although the present paper necessarily covers some of the same ground as the previous papers, it represents an independent investigation of the subject by the writer in the past few years.

It is the object of this paper to discuss the two-segment negative resistance magnetron with regard to mechanism of oscillation, limitations in efficiency and power output at ultra-high frequencies, and its application to generation of large power output at frequencies from 300 to 600 megacycles.

---

[11] The four-segment construction appears to have been first mentioned in the literature by Yagi[7] and later discussed by Posthumous,[12] Runge,[13] and others. Recently there has been considerable discussion as to whether the four-segment tube can be classed as a negative resistance oscillator.[14] The present writer feels that there is enough difference between the two-segment and four-segment tubes at ultra-high frequencies to warrant a separate treatment of the two types.

[12] K. Posthumous, "Oscillations in a Split-Anode Magnetron," *Wireless Engineer*, Vol. 12, pp. 126-132; March (1935).

[13] W. Runge, "Four-Segment Magnetron," *Telefunken Zeitung*, Vol. 15, p. 69; December (1934).

[14] E. C. S. Megaw and K. Posthumous, "Magnetron Oscillators," *Nature*, Vol. 135, p. 914; June 1 (1935).

[15] E. Habann, "A New Vacuum Tube Generator," *Zeit. für Hochfrequenz.*, Vol. 24, pp. 115-120; 135-141, (1924).

[16] E. D. McArthur and E. E. Spitzer, "Vacuum Tubes as Ultra-High Frequency Generators," *Proc. I.R.E.*, Vol. 19, pp. 1971-1982; November (1931).

[17] A. A. Slutzkin, "Theory of Split-Anode Magnetrons," *Phys. Zeit. der Sowietunion*, Vol. 6, pp. 280-292, (1934).

II. Theory of Negative Resistance Magnetron Oscillators

Before considering the negative resistance magnetron at ultra-high frequency it is well to study the fundamental principles underlying its operation.

The usual circuit of a split-anode magnetron oscillator is shown in Figure 1. Oscillations can be started by applying a magnetic field of proper magnitude parallel to the filament. The value of magnetic field required is somewhat beyond the "critical" value, which is defined as the field required to cause all of the electrons to miss the plate when both plate halves are at the same potential. The expression for the "critical field"[18] is

$$H_c = \frac{6.72}{R_a} \sqrt{E_0} \qquad (1)$$

where,

R_a = anode radius in centimeters
H_c = critical field in gausses
E_0 = average plate potential in volts.

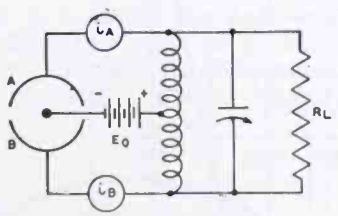


Fig. 1

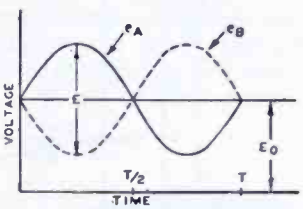


Fig. 2

Fig. 1—Two-segment magnetron oscillator circuit.

Fig. 2—Instantaneous potentials on the plate halves of a two-segment magnetron oscillator.

During the oscillation cycle, the instantaneous potentials on the plate halves can be represented as shown in Figure 2.

It is possible to demonstrate the reason for oscillation by referring to the volt-ampere characteristics, which can be shown in a number of ways. Probably the best method of representing these characteristics is illustrated in Figure 3. For this example, a tube having a 0.5-centimeter diameter plate was used and the curves were taken for the condition of 500 volts average plate potential and a magnetic field equal to approximately 1.5 times the "critical field."

The method of taking these characteristics was to increase the potential of plate *A* by increments and, at the same time, decrease the potential of plate *B* by the same increments, so as to simulate condi-

[18] A. W. Hull, "Effect of a Uniform Magnetic Field on the Motion of Electrons Between Coaxial Cylinders," *Phys. Rev.*, Vol. 18, pp. 31-57; September (1921).

tions during oscillation. When the currents to the plate segments are measured, it is found that more current flows to plate B than to plate A even though plate B is at the lower potential. Furthermore, as the potential difference $(E_A - E_B)$ is increased, up to a certain point, the excess of current to plate B increases. The current $(I_A - I_B)$ plotted against the potential $(E_A - E_B)$ gives the curve OPB of Figure 3, the portion OP of which represents a negative resistance across the circuit. This negative resistance is sufficient to account for self-sustained oscillations.

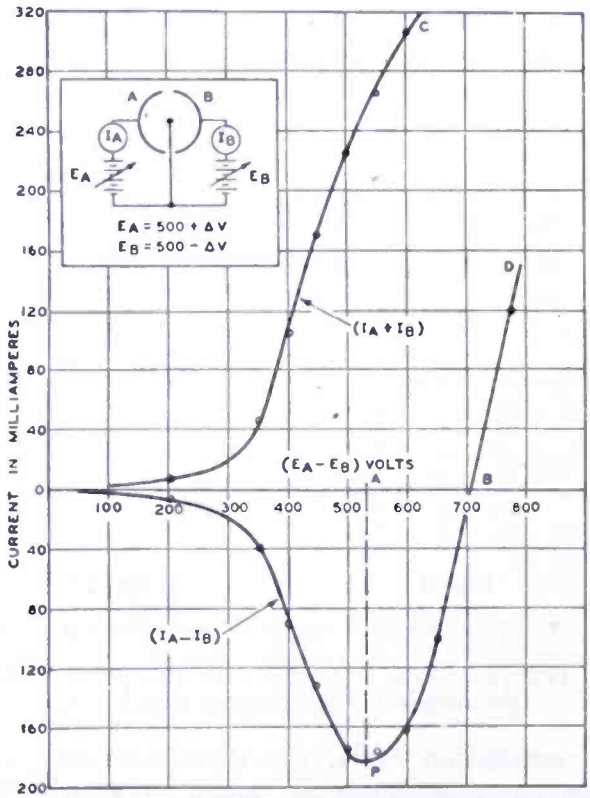


Fig. 3—Static characteristics of a two-segment magnetron.

To understand why such a characteristic should exist, it is necessary to study the electron paths under various potential conditions. With equal potentials on the plate halves, and with magnetic field beyond the "critical value," the electron paths are symmetrical curves of the type shown in Figure 4. However, when the plate halves are at different potentials, say $E_A = +150$ and $E_B = +50$, the paths are more complicated. An approximate idea of what an electron will do in this case can be had by studying an electrostatic-flux plot as shown in Figure 5. Consider first the case of an electron starting toward the high potential plate. The electron after passing the slot plane will enter a low potential region which *decreases the radius of curvature* and causes the electron to curve back somewhat short of the filament. This results

in the electron describing one or more loops, finally landing on the lower potential plate in most cases as shown in Figure 5.

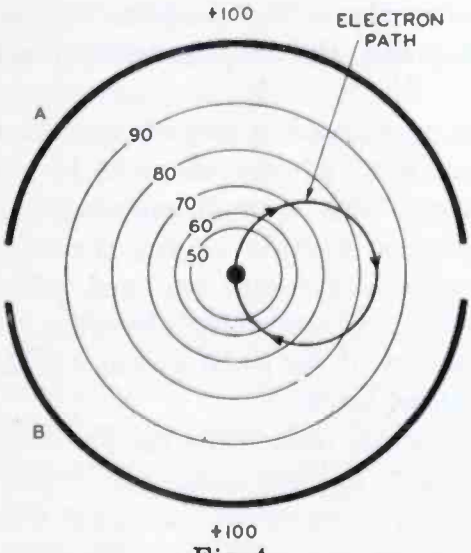


Fig. 4

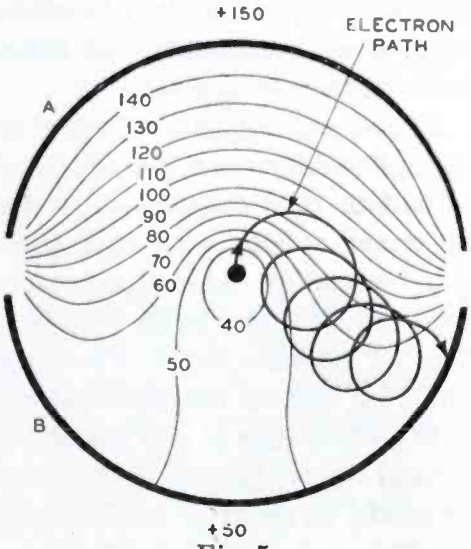


Fig. 5

Fig. 4—Electron path in a two-segment magnetron when the plate halves are at the same potential and the magnetic field is 1.5 times the critical value. (In Figs. 4, 5, and 6 the lightweight lines represent equipotential surfaces.)
Fig. 5—Electron path in a two-segment magnetron when the plate halves are at different potentials and the electron starts toward the higher potential plate. Magnetic field 1.5 times critical value.

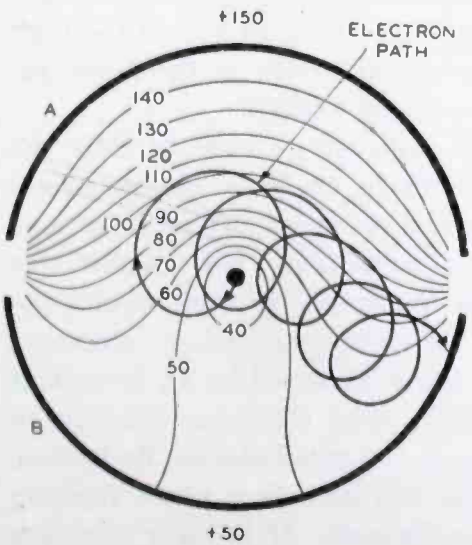


Fig. 6

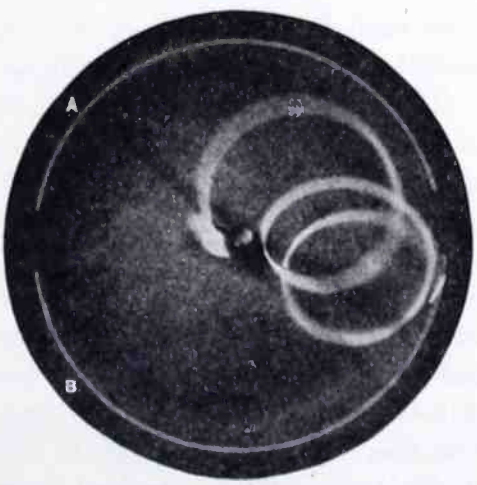


Fig. 7

Fig. 6—Path of an electron starting toward the lower potential plate. Magnetic field 1.5 times critical value.
Fig. 7—Photograph of ionized path of an electron stream starting toward the higher potential plate. Magnetic field 1.25 times critical. $E_A = +300$ volts, $E_B = +250$ volts.

On the other hand, the electrons which start toward the lower potential plate will pass the slot plane into a higher potential region

with a resulting *increase in radius of curvature*, and a consequent encircling of the filament as shown in Figure 6. In this case, it is more difficult to say what the ultimate destination of the electrons will be, but it appears probable that these electrons will also eventually reach the lower potential plate.

For an experimental check of these predictions, a special magnetron was built which made possible a visual study of electron paths by gas ionization. The cathode was constructed with a small emitting spot and made rotatable so that the electrons starting in any direction could be studied. Argon gas of a few microns pressure was used, which made the electrons beam just visible without essentially changing the shape of the beam trace. The terminal spot of the beam was also made visible by coating the plate halves with willemite.

With this tube it was possible to illustrate beautifully the predicted paths of the type shown in Figures 5 and 6. In some cases of high magnetic field, as many as ten or more loops were observed. A typical photograph of an electron-beam trace is shown in Figure 7 for the condition of $E_A = +300$, $E_B = +250$, and magnetic field equal to about 1.25 times the critical value. A systematic study was made of the electron paths by varying the direction of emission, the ratio of plate potentials, and the strength of magnetic field. The conclusion drawn from these observations is that, for sufficiently high magnetic field (one and one-half to two times critical) and with the ratio of E_A to E_B not too high (less than four to one), most of the electrons arrive at the lower potential plate no matter in what direction they started.

However, this fact in itself is not sufficient to explain fully the negative resistance characteristic. In addition, the space-charge effects must be considered. A complete analysis of the space-charge conditions in a magnetron is too involved to be attempted here, but a qualitative picture can be given as follows: With the magnetic field beyond the critical value and the plate halves at the same potential, no electrons will reach either plate; but as soon as E_A is increased by an increment and E_B decreased by the same increment, some electrons will flow to plate B as illustrated above. Because of the space-charge limitation, however, the number of these electrons will be only a small fraction of the total number emitted from the filament. It is clear that the space charge for a given current will be much higher than without magnetic field because the electrons describe several orbits before reaching the plate, thus contributing more to the space charge. Now a study of the electron paths shows that an increase in $(E_A - E_B)$ causes the electrons to describe fewer orbits. Therefore, an increase in $(E_A - E_B)$ will result in a smaller space charge and a, consequently, greater current to B.

This analysis is sufficient to explain the negative resistance characteristic such as is represented by the portion OP of the $(I_A - I_B)$ curve in Figure 3. The other part of the curve, PBD, can be explained by the fact that, as the ratio of E_A to E_B is increased, electrons eventually begin to arrive at plate A until, ultimately, more electrons are arriving at A than at B.

III. Calculation of Performance from the Volt-Ampere Characteristics

Having explained the reason for the typical volt-ampere characteristics in a split-anode magnetron it is interesting next to use these characteristics to calculate the oscillator performance. In this analysis, it is assumed that the transit time is very small compared to a period. Referring to the notation of Figures 1 and 2, and letting $e_A = E_0 + E/2 \sin \omega t$ and $e_B = E_0 - E/2 \sin \omega t$, it is easy to derive the following expressions:

$$\text{Plate Loss} = \frac{1}{T}\int_0^T e_A i_A dt + \int_0^T e_B i_B dt$$

$$= \frac{E_0}{T}\int_0^T (i_A + i_B)\,dt + \frac{E}{2T}\int_0^T (i_A - i_B)\sin \omega t\, dt. \tag{2}$$

$$\text{Power Input} = \frac{E_0}{T}\int_0^T (i_A + i_B)\,dt. \tag{3}$$

$$\text{Power Output } (P_0) = -\frac{E}{2T}\int_0^T (i_A - i_B)\sin \omega t\, dt. \tag{4}$$

$$\text{Efficiency} = \frac{E}{2E_0}\,\frac{-\int_0^T (i_A - i_B)\sin \omega t\, dt}{\int_0^T (i_A + i_B)\,dt}. \tag{5}$$

$$\text{Load Resistance} = \frac{E^2}{2P_0}. \tag{6}$$

It is clear that curves OPB and OC of Figure 3 give all the information necessary to calculate the above quantities. If an amplitude E is assumed, the instantaneous values of $(i_A - i_B)$ and $(i_A + i_B)$ can be read from Figure 3 and then, by numerical integration, power input and power output can be evaluated from (3) and (4).

The conditions for maximum output and maximum efficiency can be determined by taking several values of amplitude. This analysis was

carried out for the example of Figure 3; the results are shown in Figure 8. As could be expected, the peak output occurs at an amplitude nearly equal to *OP*, and the maximum efficiency occurs at somewhat lower amplitude.

In general, it is found that as the magnetic field is increased the crossing point *B*, of curve *OPB*, moves further out. The result is an increase in maximum output and efficiency.

In the above example, the maximum efficiency is about thirty-four per cent which by no means represents the best efficiency obtainable in a tube of this type. Unfortunately in the example given, the static curves could not be taken with higher magnetic fields because of

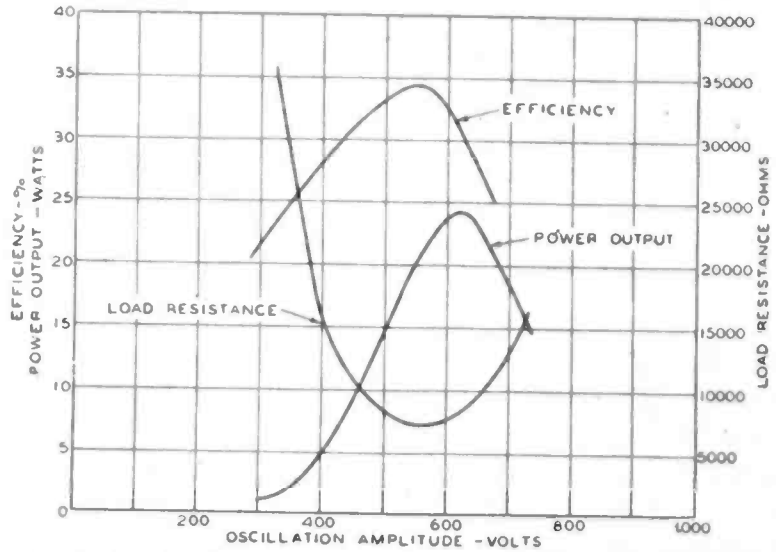


Fig. 8—Performance curves of a two-segment magnetron, calculated from the static characteristics.

problems of oscillation and electron bombardment of the leads. With higher magnetic fields, higher efficiencies may be expected. Megaw,[9] in a similar analysis, has calculated efficiencies as high as forty-five per cent for a tube operating under somewhat more favorable conditions than the above example.

IV. Limitations of Efficiency at Ultra-High Frequencies

Measured efficiencies of magnetrons at low and medium frequencies agree well with those predicted from the volt-ampere characteristics, but at very high frequencies it is found that the efficiency is considerably reduced. An experimental study of the efficiency of several magnetrons at very high frequencies was made to determine the main causes of decreased efficiency.

The problem of measuring efficiency at frequencies above 300 megacycles is very difficult because of the lack of a means of measuring

power output, which is accurate and at the same time flexible enough to be used under a wide variety of conditions. The method finally adopted was that of absorbing power in a lamp previously calibrated photometrically on direct current. The obvious error in this method is the nonuniform heating of the lamp filament at high frequencies. However, it is estimated that, by the use of specially designed lamps, the error in the measurements was held to within plus or minus twenty per cent.

Figure 9 shows the measured efficiency as a function of frequency for plate diameters of 0.5 and 1.0 centimeter. When these curves were taken, the plate potential was held at 500 volts and the magnetic field

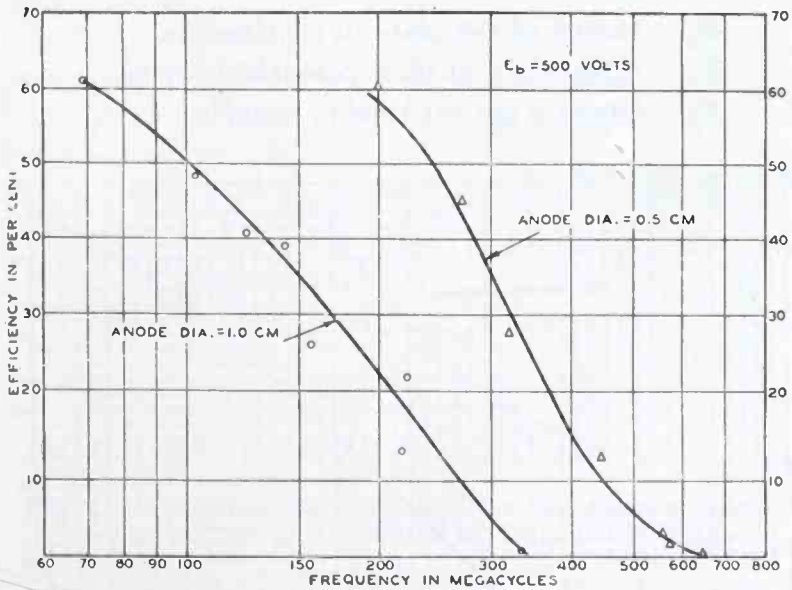


Fig. 9—Efficiency versus frequency for two sizes of magnetrons operating at the same plate potential.

was adjusted to give maximum efficiency at each point. The general shape of the curves for the two tubes is approximately the same; but, at a given efficiency, the frequency for the smaller plate diameter is roughly twice that for the larger. This seems to indicate that the higher efficiency for the smaller diameter tube is due to the shorter electron-transit time, and suggests that the decrease in efficiency at higher frequencies is because of the appreciable transit time.

If the efficiency is mainly a function of the ratio of transit time to period, then it might be expected that, at a given frequency, the efficiency will increase with plate voltage. This is borne out by the curve shown in Figure 10 where efficiency is plotted as a function of plate voltage for a tube operating at 440 megacycles. The tube used in this test had a plate diameter of 0.5 centimeter and was of the internal

circuit construction described in Part V. The magnetic field for each reading was adjusted for the best efficiency.

To illustrate further the relation between efficiency and transit time, the efficiency data of Figures 9 and 10 with some additional data are plotted as a function of the ratio of transit time to period as shown in Figure 11. The value of transit time used is an effective direct-current transit time[19] given by the expression,

$$T_0 = 2.65 \times 10^{-8} \frac{R_a}{\sqrt{E_0}} \tag{7}$$

where,

R_a = radius of the plate in centimeters,
E_0 = direct-current plate potential in volts,
T_0 = effective transit time in seconds.

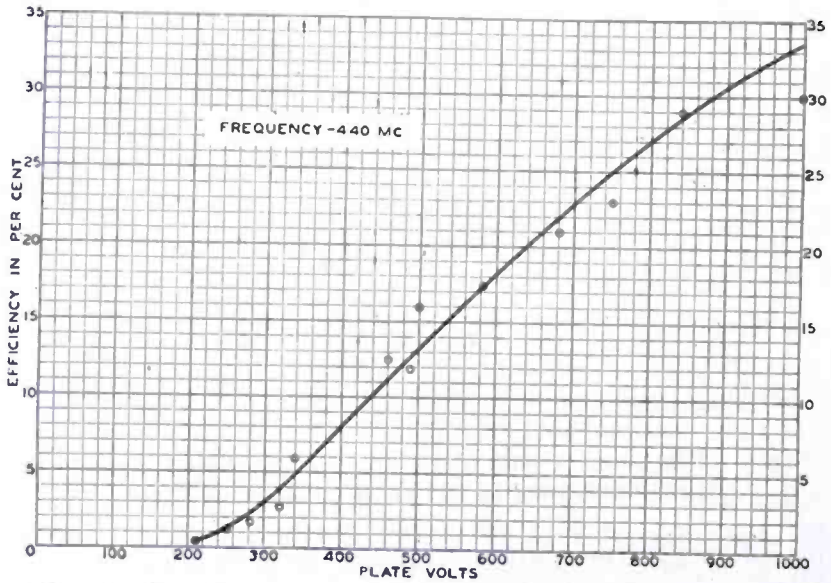


Fig. 10—Efficiency versus plate voltage for a magnetron operating at a frequency near the high-frequency limit.

The fact that all points from the several sources lie fairly close to a smooth curve is good evidence that the decrease in efficiency is, for the most part, due to transit-time effects. Examination of the curve shows that, if the transit-time ratio is below one-fifteenth of a period, efficiencies as high as fifty per cent can be expected, and that, even at

[19] The calculation of this transit time assumes a uniform velocity ($v_0 = 5.95 \times 10^7 \sqrt{E_0}$) and a semicircular path of a diameter equal to R_a. This transit time is equal to one half the orbital time of an electron traveling with a velocity v_0 in a magnetic field $H_c = 6.72/R_a \sqrt{E_0}$.

one-tenth of a period, thirty per cent efficiency is possible, but that, at one-fifth of a period, the tube will almost fail to oscillate.[20]

As a practical example, an efficiency of thirty per cent can be obtained at 600 megacycles with a tube having a plate diameter of 0.5 centimeter and a plate potential of 1500 volts.

Another factor to be considered in connection with attaining high efficiency at ultra-high frequencies is the value of magnetic field required. It is possible to demonstrate that, for a given efficiency and frequency, the value of the magnetic field is definitely determined,

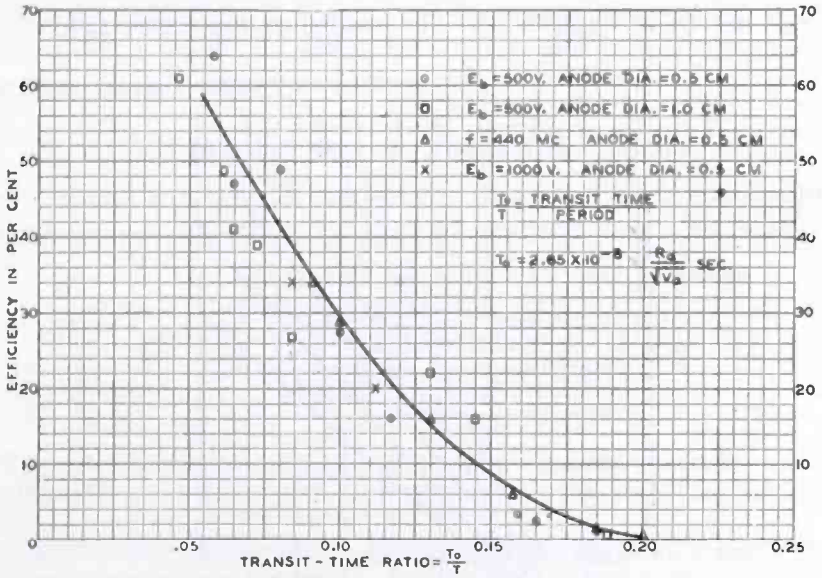


Fig. 11—Combined efficiency data for two-segment magnetrons, plotted as a function of transit-time ratio.

regardless of what plate voltage or plate diameter is used. This can be shown by expressing the transit time as a function of the magnetic field alone. This is possible, since E_0 and R_a are connected through the relation

$$H = k \frac{6.72}{R_a} \sqrt{E_0} \tag{8}$$

where k in practice lies between 1.5 and 2.0. Substituting (8) and (7) gives,

$$\frac{T_0}{T} = 1780 \frac{kf}{H} \times 10^{-10} \tag{9}$$

[20] When this general relation is compared with the data given by Megaw[9] a fairly good agreement is found. A similar comparison with the work of Slutzkin shows much poorer agreement, the efficiencies given by Slutzkin being generally higher.

where,

f = frequency in cycles per second

H = magnetic field in gausses

$\frac{T_0}{T}$ = ratio of transit time to period.

This expression for transit-time ratio can now be combined with the efficiency curve of Figure 11 to give the magnetic field for any frequency and efficiency. This relation can best be illustrated by a chart of the type shown in Figure 12. The values of magnetic field obtained

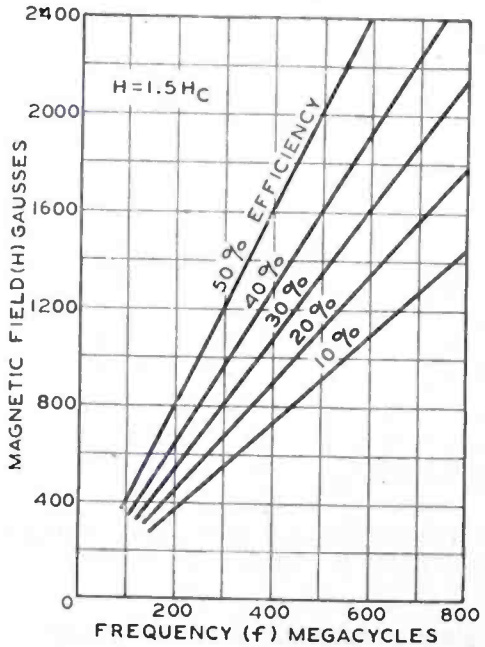


Fig. 12—Constant efficiency curves of a negative resistance magnetron showing the magnetic field strength required for any given efficiency and frequency.

from this diagram are only approximate because the value of k may vary considerably in practice and because the efficiency measurements are subject to a fairly large error.

It is interesting to compare the magnetic field required by a negative resistance oscillator to that required by an electronic oscillator operating at the same frequency. If the comparison is made on the basis of ten per cent efficiency (limit of electronic oscillator) it is found that the negative resistance magnetron requires approximately four times the field strength.[21]

[21] This follows from the approximate relation for electronic oscillators that

$$H = 12{,}000/\lambda \text{ cm.}$$

In the discussion so far it has been assumed that circuit loss plays an unimportant role. This is contrary to the statement often made that circuit loss is the limiting factor in magnetron oscillators at ultra-high frequencies. However, it has been the experience of the author that, with proper care in design, circuits can be built which have negligible loss even at 600 megacycles. This is accomplished by using close-spaced leads to reduce radiation and by making the surface area of the leads sufficient to give small high-frequency resistance. A very low loss circuit has been obtained by the use of an internal copper circuit which will be described in Part V. The conclusion can be drawn that magnetron circuits can be designed so that the real limiting factor is the electron transit time.

V. Limitations in Plate Dissipation

The preceding analysis has shown that, for high efficiency in the 300- to 600-megacycle range, a small anode diameter is required (about

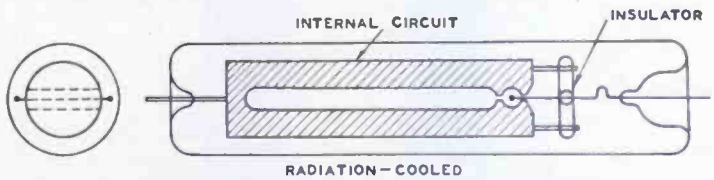


Fig. 13—Sectional view of an internal circuit radiation-cooled magnetron for obtaining high power at ultra-high frequencies.

0.5 centimeter for 1500 volts). It has also shown that a high magnetic field is required, a fact which limits the length of plate to a few centimeters for a magnet of reasonable dimensions. Both of these facts definitely limit the plate size and, consequently, impose a serious limitation on the possible plate dissipation. At first thought, it appears that the maximum dissipation of such a tube would be of the order of ten watts, but further study shows that this value can be increased by a large factor.

Although the dimensions of the plate cylinder are small, the radiating surface can be increased considerably by using a heavy walled plate to increase the outside area. The surface can be still further increased by placing the oscillating circuit within the bulb, as illustrated in Figure 13. When conductors of large cross section and good thermal conductivity are used the whole circuit is essentially at the same temperature, and its entire surface is effective in radiating heat. In this manner the radiating surface can be increased by a factor of the order of twenty to one.[22] Figure 14 illustrates a tube of this

[22] Shortly prior to the time at which the author constructed the first tube of this type, Mr. P. D. Zottu of this laboratory designed an internal circuit magnetron embodying the principal features described here.

construction which has a safe plate dissipation of 200 watts and will deliver an output of about fifty watts at a frequency of 550 megacycles with an efficiency of about thirty per cent. In this instance, the circuit was made of copper, which not only gives good thermal conductivity but also results in a very low loss circuit. To increase the emissivity the outer surface was carbonized. Incidentally, the carbonization may be expected to cause but little increase in the high-fre-

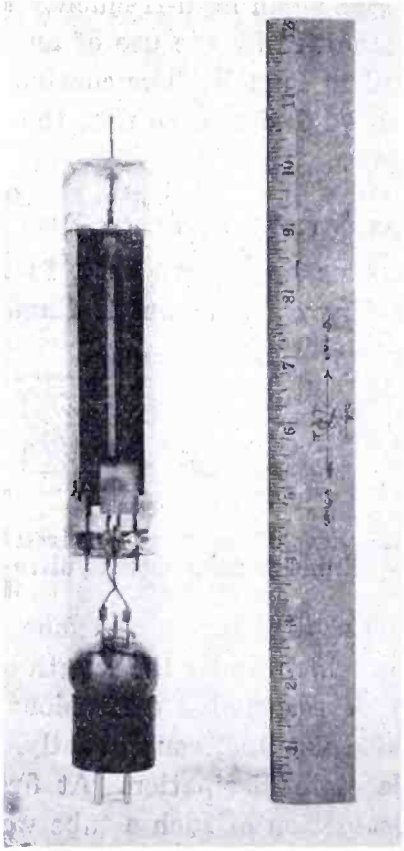


Fig. 14—Photograph of an internal circuit radiation-cooled magnetron oscillator for 550 megacycles.

quency resistance because most of the current in such a structure flows on the inner surface. The method of coupling this tube to the load was to use a parallel-wire transmission line, the closed end of which was inductively coupled to the internal circuit of the tube.

It is obvious that the internal circuit construction limits a given tube to operation over a relatively narrow frequency band. However, such a limitation may not be so serious in an ultra-high-frequency tube as it would be in a tube intended for lower frequency applications. Moreover, it may be pointed out that, aside from the advantage of high dissipation, the internal circuit construction becomes a necessity at frequencies around 400 megacycles, because of the impossibility of

building external circuit tubes which will tune to such high frequencies.

To obtain still greater plate dissipation, it is necessary to resort to the use of some cooling liquid such as water. The problem of water-cooling in a split-anode magnetron for very high frequencies is not so simple as in the conventional three-electrode tube. In a magnetron, the power is dissipated over a comparatively small area. This makes it necessary to conduct the heat away to a larger surface which can be effectively water-cooled. There are a number of ways in which this can be accomplished; one example is illustrated by Figure 15. Here, also, an internal circuit of heavy copper conductors is used, but in this case the circuit is conductively coupled to the load by two leads brought out in a plane at right angles to the plane of the filament leads. Figure 16 shows a photograph of a laboratory tube of this construction with an internal circuit and plate of approximately the same

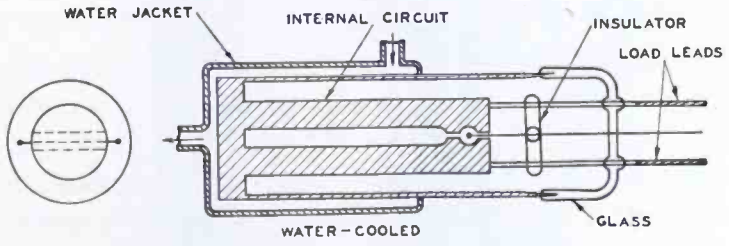


Fig. 15—Sectional view of one type of water-cooled magnetron for high power at ultra-high frequencies.

dimensions as the radiation-cooled tube of Figure 14. This particular tube will dissipate more than 500 watts and will deliver an output of approximately 100 watts at a frequency of 600 megacycles with an efficiency of about twenty-five per cent.

VI. Miscellaneous Limitations

Besides the factors discussed in the previous sections, there are two other factors that limit to some degree the output obtainable from a magnetron oscillator. One of these, existing in radiation-cooled tubes, is the electron bombardment of the glass walls opposite the ends of the plate due to the focusing effect of the magnetic field. A solution to this problem has been found by adding shielding electrodes at the plate ends. Shielding electrodes of this type can be seen in the illustration of Figure 14.

The other factor which is somewhat more serious is a phenomenon termed "filament-bombardment effect." This effect, observed by the author several years ago, has been mentioned by a number of writers

on magnetrons.[23] The effect manifests itself as an increase of filament temperature under certain conditions of high magnetic field and high plate voltage, and sometimes results in unstable operation of the tube.

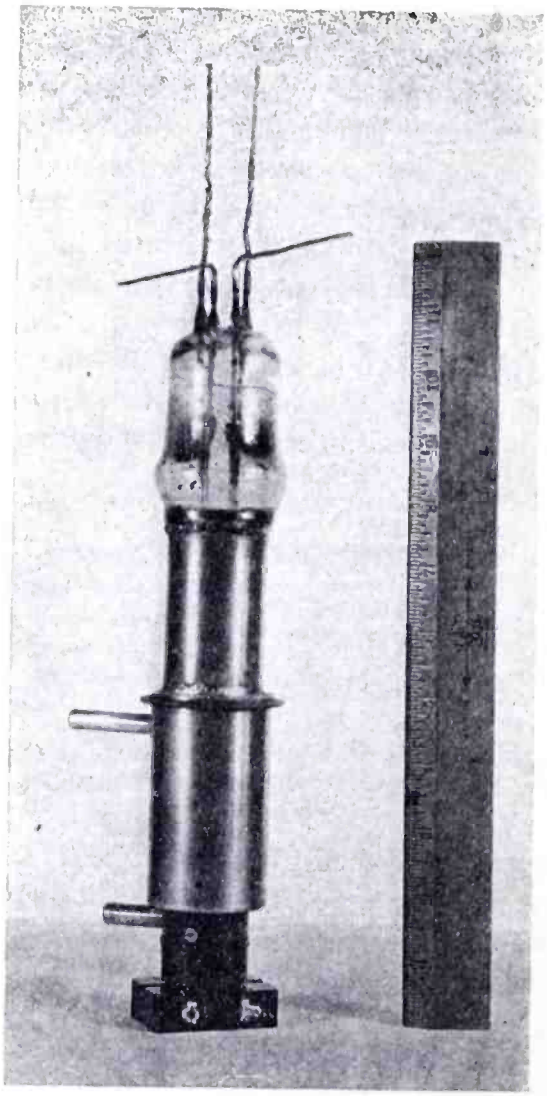


Fig. 16—Photograph of a water-cooled magnetron oscillator for 600 megacycles.

[23] Megaw,[24] Slutzkin[25, 26] and others have described this effect and Langmuir and Found[27] have observed a related phenomenon in connection with electron scattering.

[24] E. C. S. Megaw, "A New Effect in Thermionic Valves at Very Short Wave Lengths," *Nature*, Vol. 132, p. 854; December 2 (1933).

[25] A. A. Slutzkin, S. J. Brande, and I. M. Wigdortschik, "Generation of ion currents in high vacuum by the help of magnetic fields," *Phys. Zeit. der Sowietunion*, Vol. 6, pp. 268-279 (1934).

[26] A. A. Slutzkin, et al., "Production of Electromagnetic Waves Below Fifty Centimeters," *Phys. Zeit. der Sowietunion*, Vol. 6, pp. 150-158 (1934).

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

In extreme cases, the filament can receive sufficient energy from the plate circuit to permit operation of the tube with the usual filament supply disconnected. The cause of this phenomenon has not been fully explained, but it appears to be due to a bombardment of the filament by electrons.

Although the filament-bombardment effect is sometimes troublesome, it can generally be avoided by using heavy filaments and by operating the tube at somewhat reduced plate voltage and magnetic field strength.

VII. Conclusion

It has been demonstrated by theory and experiment that the negative resistance magnetron is essentially a high efficiency device at low frequencies, and that the decrease of efficiency at high frequencies is mainly due to transit-time effects. As applied to frequencies between 300 and 600 megacycles, it is shown that this type of oscillator can be expected to give efficiencies of the order of fifty to thirty per cent.

Methods are described by which the inherently small plate-dissipation limit can be extended by twenty to fifty times, and by which it is possible to realize power outputs of the order of fifty to 100 watts in the 300- to 600-megacycle range. The output and efficiencies obtained compare favorably with those of conventional tubes at much lower frequencies, but it is not to be inferred that magnetrons will necessarily supplant other types of generators. Problems of modulation and frequency stability are still to be met and in some applications the supplying of a high magnetic field may be inconvenient.

In conclusion, the author wishes to point out that the specific tubes described are not to be regarded as commercial designs. They are, rather, laboratory tubes built to demonstrate certain principles which, it is hoped, will prove useful in future designs of ultra-high-frequency generators.

Acknowledgment

The writer desires to express his indebtedness to Mr. B. J. Thompson and Mr. H. C. Thompson for many helpful suggestions, and to Mr. T. H. Clark for his assistance in the experimental work.

AN ULTRA-HIGH-FREQUENCY POWER AMPLIFIER OF NOVEL DESIGN

By

ANDREW V. HAEFF

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

FOR many high-frequency applications a non-regenerative power amplifier is of primary importance. It is the purpose of this paper to describe an amplifier of a novel type with which successful operation is obtained at frequencies much higher than those which can be handled by conventional devices of comparable power capabilities.

Some of the principal difficulties encountered in the operation of amplifiers at very high frequencies can be ascribed to the following causes.

1. The finite electron transit time resulting in excessive input loading and a loss in transconductance.

2. Abnormally high coupling between the output electrode and input electrode causing either regeneration or excessive loading of the output circuit by the reflected input losses with a consequent loss of power and efficiency.

3. High circulating currents and increased resistance at high frequencies lowering circuit efficiency.

In the new amplifier the electron-transit-time effects are minimized by utilizing electrons of high velocity. This is accomplished without increased dissipation and loss in efficiency by separating the functions of the output electrode and of the current-collecting electrode and by making use of electron focusing. The output-input coupling is reduced to a negligible value by screening and separation of the respective electrodes and circuits. The high-frequency losses due to circulating currents are minimized by using current-carrying electrodes of large periphery.

The principle used in exciting the output tank circuit of the new amplifier differs from that of conventional tubes. Figure 1 represents a quarter-wave concentric tank circuit with a hollow inner conductor. Let us suppose that a negatively charged body is passed through the inner conductor from left to right. In Figure 1a we see the conditions of charge distribution on the circuit as the body is introduced into

Reprinted from *Electronics*, February, 1939.

the inner conductor. There is a positive charge, equal to the negative charge, induced on the inside of the inner conductor near the body. However, no charge appears on the outer surface of the inner conductor. The induced charge moves with the charged body along the inner surface until the end of the inner conductor is reached. During the passage of the charged body across the gap *a-b* the charge is

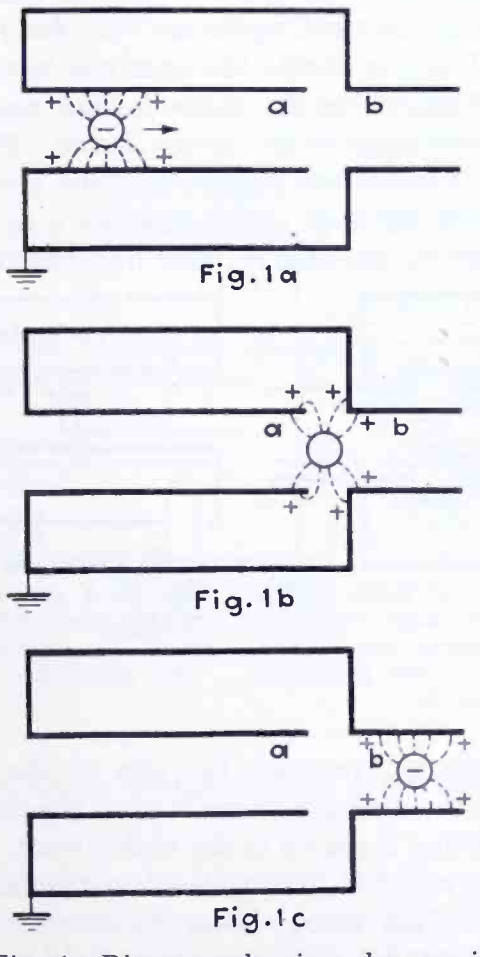


Fig. 1—Diagram showing changes in distribution of image charge as a charged particle moves inside the quarter-wave tank circuit.

partially imaged on the end of the inner conductor and partially on the outer conductor as shown in Figure 1b. The passage of the charged body beyond the gap *a-b* into the aperture in the outer conductor causes all of the induced charge to appear on the inner surface of the extension of the aperture (Figure 1c). The induced charge, in transferring from the end of the inner conductor to the aperture extension, flows back over the outer surface of the inner conductor, thus constituting a current flow in the tank circuit.

Figure 2 illustrates the configuration of the electric and magnetic fields within the resonating space of the tank circuit when the latter is excited. The solid lines represent the electric field distribution and the circles represent the magnetic lines of force. The dashed lines represent the equipotential surfaces in the gap. Along the major part of the length of the tank the direction of the electric field is substantially radial. However, at the gap *a-b* the electric field has an axial component. The field does not penetrate very far inside the open end of the inner conductor or inside the aperture in the outer conductor but is confined effectively to the space defined approximately by the limiting equipotential lines shown in the figure. The space inside the inner conductor and inside the extension of the aperture is essentially field free. Therefore, no work will be done on a charge moving inside the inner conductor by the electric field until the charge reaches the

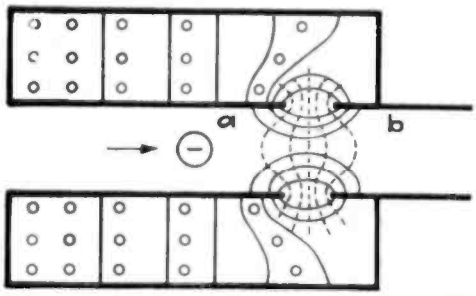


Fig. 2—Distribution of fields inside resonating space of tank circuits; fields do not penetrate inside the inner conductor, *a*, or the aperture extension, *b*.

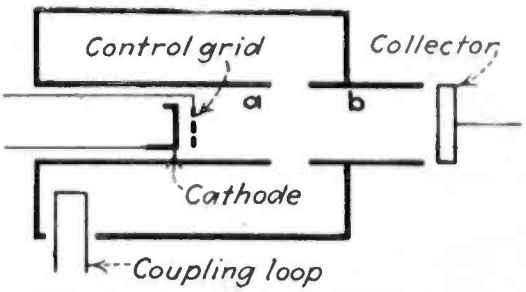


Fig. 3—A schematic diagram illustrating the arrangement of electrodes for exciting the tank circuit by modulated electron stream.

gap *a-b*. If the charge traverses the gap at the instant when the electric force is in the direction from *b* to *a*, the charge will be decelerated, its energy being given up to the tank circuit. A charge crossing the gap during the opposite half cycle when the field is reversed will be accelerated, absorbing energy from the circuit. If the number of charges traversing the gap during the first half cycle is greater than during the second, the net effect will be that energy is supplied to the tank circuit.

We conclude, then, that the tank circuit may be excited by passing groups of electrons at the proper frequency across the gap between the inner conductor and the outer conductor. The motion of electrons through the interior of the inner conductor has no effect on the current in the tank circuit. Also, high-frequency electromagnetic fields which will be generated within the resonating space of the tank circuit penetrate but a short distance inside the screening electrode *b*, so that the electrons will be influenced by these fields only during their passage across the gap.

Figure 3 shows schematically how a tube may be combined with the tank circuit to operate on this principle. A conventional grid-cathode structure may be used to obtain pulses of electrons. A collector electrode may be placed beyond the screening electrode. If a high potential is applied between the cathode and the electrodes *a*, *b*, and also between the collector and the cathode, a stream of electrons from the cathode will flow towards the collector. If a high-frequency voltage is applied between the control grid and the cathode, the electron stream will be periodically modulated in intensity. Pulses of electrons traversing the gap *a-b* will induce high-frequency currents between electrodes *a* and *b*. If the excitation frequency is adjusted to the

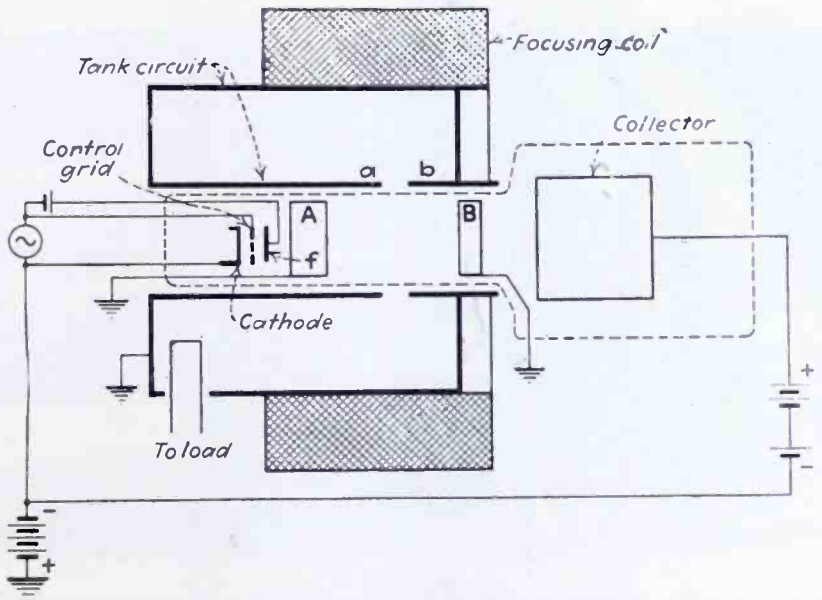


Fig. 4—Complete schematic wiring diagram of the amplifier and new tube with all essential parts indicated.

resonant frequency of the tank circuit, a high impedance will exist across the gap *a-b* at this frequency. Consequently, the induced currents will produce a high radio-frequency voltage across the gap. The phase of this voltage at or near resonance will be such as to decelerate electrons traversing the gap during the half period of maximum intensity of electron current in the stream.

The energy lost by the electrons is transformed by the tank circuit into the energy of the electromagnetic field within the space between the inner and outer conductors. This energy is then transferred to the useful load by means of a coupling loop, as shown in Figure 3.

The high-frequency electromagnetic field existing in the resonant space of the tank circuit penetrates only a short distance inside the cylinder *a* and inside the screening electrode *b*. Therefore, by positioning the control electrode and the collector at suitable distances from

the gap *a-b*, the coupling between the output circuit and the respective electrodes can be reduced to a negligible value.

To minimize the transit-time effects the electrodes *a* and *b* can be operated at suitably high potentials with respect to the cathode. The adjustment of these potentials is not at all critical because the functioning of the tube does not depend critically upon the electron transit

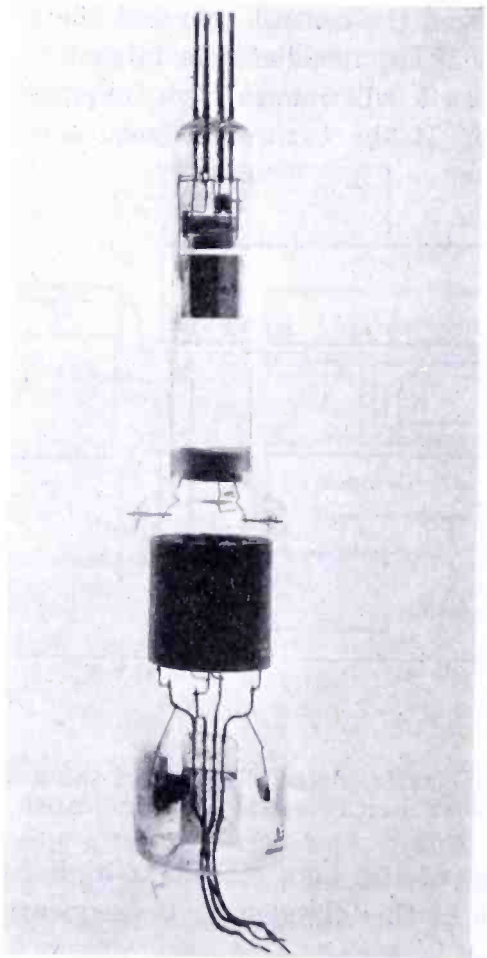


Fig. 5—Photograph of new type of tube in developmental stage.

time. The current-collecting electrode can be operated at a much lower potential and is usually operated at a potential just sufficient to collect all decelerated electrons in order to obtain high efficiency. By the use of electrostatic or magnetic focusing the electron stream can be prevented from impinging on the high-potential electrodes *a* and *b*. Therefore, these electrodes do not dissipate energy. It follows that all of the power developed by the tube is taken from the low-voltage supply used for the collector.

Figure 4 shows schematically one complete arrangement. A focusing electrode *f* and accelerating electrodes *A* and *B* are mounted inside the glass envelope in addition to the cathode, the control grid, and the collector of the previous arrangement. The output tank circuit, consisting of the outer cylinder, a hollow inner conductor, and a screening electrode, is separate from the tube. A solenoid placed around the tank circuit produces a focusing magnetic field in the axial direction. The focusing electrode, *f*, when operated at a suitable potential serves to concentrate the electron stream at the start and makes it possible to use a considerably weaker magnetic focusing field without absorption of current by the accelerating electrodes. The reason for using accel-

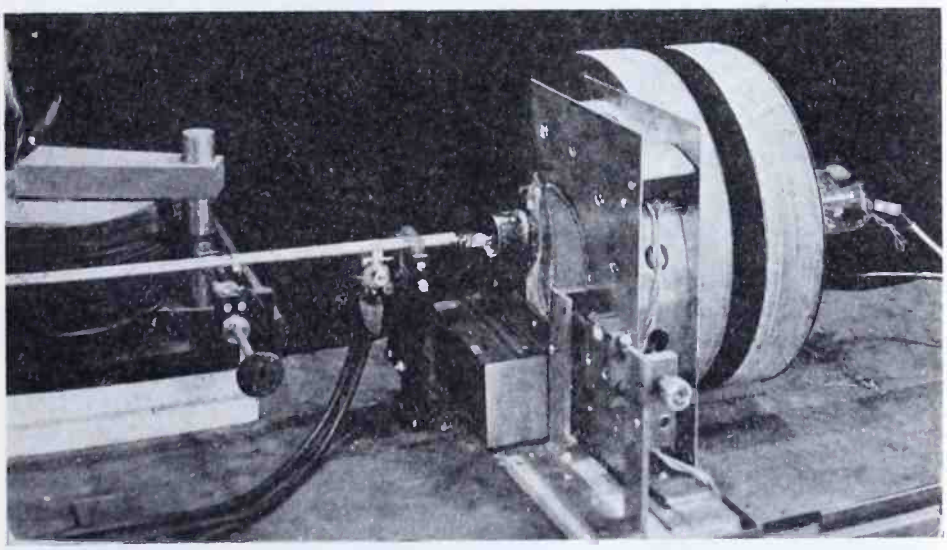


Fig. 6—The completed amplifier showing, from left to right, the driving magnetron, the input circuit, the housing for the load lamps, and focussing coils and the collector end of the tube.

erating electrodes inside the glass envelope is to avoid the undesirable effects of charges on the glass walls due to bombardment by stray electrons. These electrodes are positioned at a suitable distance from the gap *a-b* of the output tank circuit so that the high-frequency electromagnetic field from the resonant space of the tank does not reach them. Thus, these electrodes are not a part of the output circuit and do not carry circulating currents.

Figure 6 shows this tube mounted in the circuit. In this arrangement a small magnetron was used as a driving oscillator. The photograph shows the magnetron with its tuning mechanism, the parallel-wire input line, and a push-pull voltmeter which was used for measurements of driving power. The load consisted of eight incandescent lamps, fed in parallel from the coupling loop and mounted in

a shielded compartment to prevent radiation losses and coupling to the input line. Fine adjustment of load was obtained by moving the load coupling loop with a micrometer screw. Behind the focusing solenoids the collector end of the tube can be seen. Figure 7 shows the tube in operation with the load lamps brilliantly lighted and visible through the observation openings.

As an example of performance of the new amplifier, the following data for the developmental tube shown in Figure 5 may be of interest.

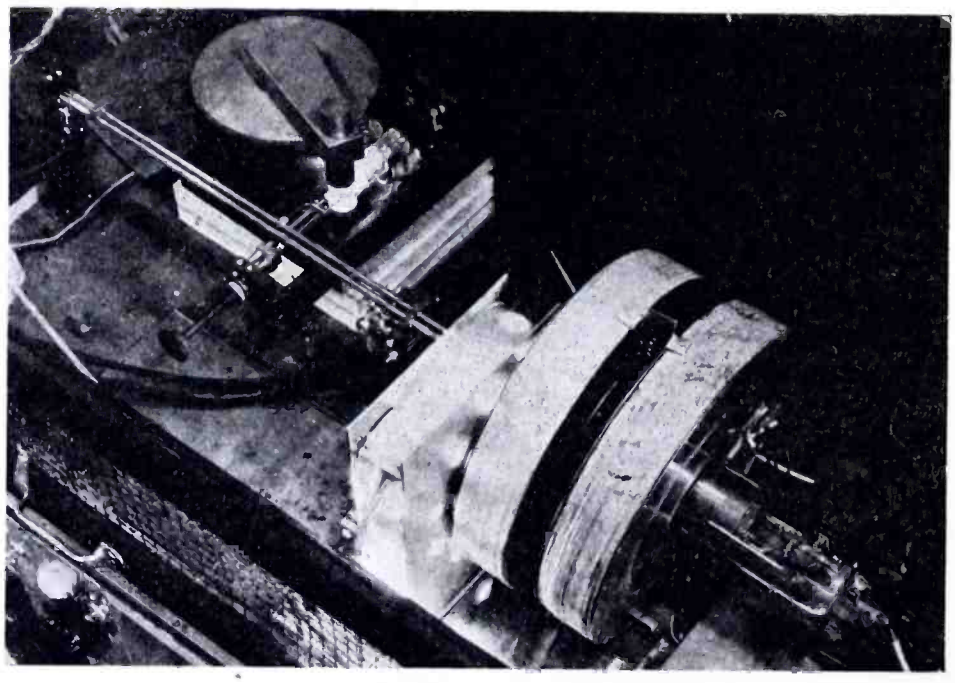


Fig. 7—The u-h-f amplifier in operation, and brilliantly lighting the incandescent lamp load circuit, contained in the lamp housing adjacent to the focussing coil.

Operating frequency	450 megacycles
Power output	110 watts
Total driving power (approx.)	10 watts
Efficiency (approx.)	35%
Accelerating voltage	6000 volts
Collector voltage	2000 volts
Collector current	150 milliamperes
Current to accelerating electrodes	less than 0.1 milliampere

We may summarize the advantages of the new amplifier as follows:

1. Reduction of electron-transit-time effects by utilization of electrons of high velocity.

2. High efficiency due to collection of electrons at low velocity.

3. High power output because the collector may be made of adequate size without influencing the performance of output circuit.

4. Non-regenerative amplification through reduction of output-input coupling to a negligible value.

5. Low circuit losses because circulating currents flow in electrodes of large periphery.

The author wishes to acknowledge the valuable help of Dr. L. S. Nergaard in making measurements in connection with this work.

DEVELOPMENT OF TRANSMITTERS FOR FREQUENCIES ABOVE 300 MEGACYCYLES*

By

N. E. Lindenblad

R.C.A. Communications, Inc., Rocky Point, L. I., N. Y.

Summary—*The fundamental functions of the electrons and their work cycle, in the interelectrode space of high vacuum tubes, are discussed. It is shown how the triode feed-back circuit becomes inoperable at very high frequencies due to space-time and reactance characteristics. It is further shown how the space-time conditions can be organized to benefit the maintenance of oscillation instead of becoming a detriment. Some of the more familiar arrangements, such as the Barkhausen and the magnetron circuits, which are based on this principle, are discussed in some detail. With these illustrations as a background the author describes a new method of frequency multiplication at very high frequencies. This method yields much greater power outputs than hitherto possible and promises to become very useful.*

Various means for frequency stabilization are referred to and the merits of frequency controlling devices, such as crystals and low power factor circuits, are compared.

Special problems encountered in the application of modulation at very high frequencies are described and reference is made to methods developed to meet these problems.

Practical considerations of circuit arrangements are described in some detail. Several examples of transmitter design are given. These sections are illustrated with photographs.

Important points in connection with antennas and transmission lines are discussed and the results of some measurements are given.

The paper ends with a brief reference to some propagation results obtained by RCA Communications engineers and others.

Introduction

THE purpose of this paper is to report progress in theoretical conception as well as in practice pertaining to the application of the frequency band between 300 and 1000 megacycles to radio communication.

The Electron Performance in High Vacuum

Since electrons are negative electric charges, their presence in the interelectrode space of a vacuum tube causes, by virtue of electrostatic induction, positive charges to be distributed over the electrodes which

* Decimal classification: R355.5. Presented before I.R.E. Ninth Annual Convention, May 30, 1934, Philadelphia, Pa.

Reprinted from *Proc. I.R.E.*, September, 1935.

will vary in accordance with variation in position of the electrons. If the external circuit consists of a resistance it can be seen that the current produced in this resistance, by virtue of the electron motion in the interelectrode space, always causes a voltage drop on the electrodes which retards the electron motion and decreases the rate at which electrons enter the interelectrode space.

When the electrons land on the electrodes, the positive charges in the electrode meet the corresponding electron charges and are cancelled.

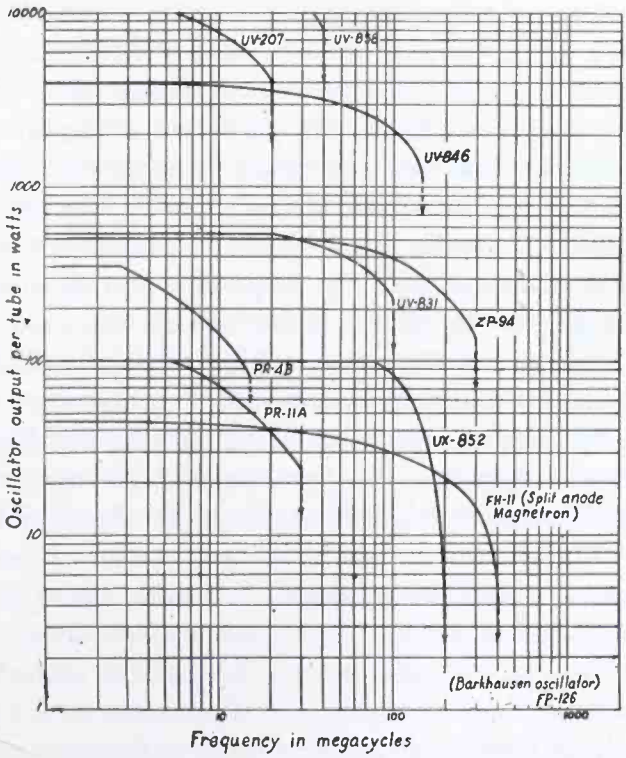


Fig. 1—Frequency versus power output at rated anode dissipation of some American transmitting tubes. Curves not otherwise indicated refer to performance in triode feed-back oscillators. These curves were originally published by W. C. White in the *General Electric Review*, September, (1933).

The current caused by the electron motion therefore ceases at the moment the electrons land. Whatever kinetic energy the electron possesses at the time it lands is lost in the form of heat.

The Electron Performance in High Vacuum at High Frequencies

When the well-known, triode feed-back oscillator is adjusted for higher and higher frequencies, a frequency will eventually be reached beyond which the device fails to perform. (Figure 1.) It is usually assumed that the increased circulating currents necessary to maintain proper electrode potential across the decreasing capacitive reactance

of the interelectrode space cause prohibitive losses. This is, however, not the major factor in well-designed circuits and the chief limitations are instead to be found in the vacuum tube itself.

Hitherto it has been possible to neglect the interelectrode transit time of the electrons in vacuum tube phenomena. The finite velocities of the electrons introduce phase lags in the electron motion which are unsuitable to the triode feed-back method and results in a reduction of efficiency. Accompanying this phase lag electrons are trapped in the interelectrode space. As the transit period terminates, the electrons in the grid-cathode space come to a stop. The grid potential is rapidly becoming very negative and assumes a controlling effect upon the field in the grid-cathode space. This field therefore changes direction and the electrons in the grid-cathode space are thrown back into the cathode at high velocity, resulting in high kinetic loss. Since the plate becomes more positive, while the grid grows negative, the field in the grid-anode space is increased. The electrons in this space therefore receive additional impetus in the direction of the anode and arrive there at high velocity and thus with a high kinetic loss. At the lower frequencies the interelectrode spaces are "cleaned out" before the potentials have had time to reach excessive values. The higher the frequency the greater becomes the number of electrons which fail to accomplish the transfer or which transfer under field conditions which cause excessive kinetic losses. The existence of these losses has previously been referred to but not explained.[1] Since a great portion of the grid input energy, due to trapped electrons, appears at the cathode, its value can be observed by noticing the increase in cathode resistance from the increased cathode temperature. Estimates of the loss obtained in this way indicate that it is a major source of frequency limitation in conventional transmitting tubes.

In order to reduce the losses during the "cleaning-out" period, the interelectrode space must be made small in volume so that it contains a small number of electrons in transit. If the power output is to be reasonably retained and since the cathode emission at the present cannot be increased, the cross section of the interelectrode space cannot be excessively reduced. The only way to obtain substantial decrease of volume is thus to reduce the length of the interelectrode space. While this on one hand results in increased capacity with the handicap of higher circulating currents it improves the phase relation between the moving electrons and the electrode potentials. Thompson and Rose[2]

[1] F. B. Llewellyn, "Vacuum Tube Electronics at Ultra-High Frequencies," *Proc. I.R.E.*, Vol. 21, pp. 1532-1574; November, (1933).

[2] B. J. Thompson and G. M. Rose, Jr., "Vacuum Tubes of Small Dimensions For Use at Extremely High Frequencies," *Proc. I.R.E.*, Vol. 21, pp. 1707-1722; December, (1933). (See Page 334.)

have had rather outstanding success in compromising these factors in the design of receiving tubes for very high frequencies.

Generation of Higher Frequencies

In the triode feed-back circuit the electron flow takes place during a very short favorable portion of the oscillation cycle. Due to the finite velocity of the electrons it has been shown that it is not possible to confine the existence of electrons in the interelectrode space to such short portions when the frequency is increased. As a result, and because of the nature of the circuit, the oscillating power created in the external circuit is returned to the electrons and lost in the form of heat. It is, therefore, necessary to choose methods in which the electrons may perform usefully during more prolonged portions of the oscillation cycle; in which the time of travel of the electrons and the electric fields produced by the electrons themselves contribute to the condition of oscillation. In other words, the electrons themselves and their motions constitute the whole oscillator. All such oscillators require electric, and often also magnetic, field conditions such that the electrons as a group are subjected to unstable conditions which can produce whistle effects in the interelectrode space. The periodic pressure effects in an air whistle would thus correspond to the potential effects set up by the periodic formation of concentrations in space charge.

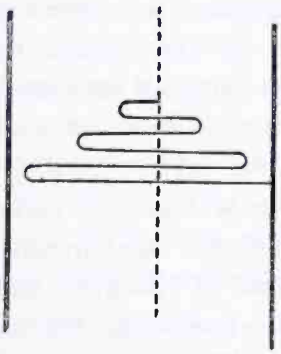
These conditions may be obtained if the electrons are given an opportunity to miss an anode as they are accelerated toward it. After missing the anode and as the electrons are thus carried away from it by their own momentum, they will be subject to a retarding force instead of an accelerating force from the anode. They will eventually come to a stop and again be accelerated toward the anode. The positive direct-current potential on the anode thus makes the electrons describe a pendulum motion. Since there are many electrons and thus many such pendulums they cannot be permitted to oscillate at random phase since they will then cancel each other's effect upon the external circuit. The pendulum motions must, therefore, be organized to operate in unison. By comparison with the traffic congestion on a highway which occurs at points where the traffic speed is reduced, it is easily seen that congestion of electrons will arise in the regions of the interelectrode space where the electrons turn around. As these accumulations form, the resultant electric field in the interelectrode space is gradually being altered. This alteration influences the motion of the electrons. Also, electrons arriving later at the turning region, retard the turning around of the earlier ones, while, on the other hand, the earlier electrons speed up the turning around of the later ones. This condition is

therefore conductive to synchronization of the oscillating electrons so that they form into groups. When the oscillating electrons form into groups they will thus no longer cancel each other's influence upon the external circuit. The losses from the currents induced in the external circuit by the group motion will, therefore, introduce a load upon this motion, and the motion of the individual members of the oscillating electron group will become attenuated. This attenuation takes place in the direction toward the prime mover, the anode, where the electrons are ulimately deposited. Since they are continually being replaced by newly emitted electrons of high momentum, the motion of the group as a whole is not attenuated but may be represented by an average of the motion of its continually changing members. It should also be clear that many of the electrons will be subject to accidents so that they will prematurely collide with the anode. There is no electrode system known where this can be prevented. The fact should, however, be noted that while such collisions represent very great losses the kinetic energy so spent is not taken from the oscillating energy in the external circuit as the case happened to be at the end of transit in the triode feed-back oscillator. Practically all of the momentum possessed by the electrons has been derived from acceleration by the interelectrode, direct-current field. Since the cathode is inherently located in a region where the electrons reverse their motion and since the space charges, there forming, appear periodically, it is clear that the emission from the cathode will also be subject to periodic fluctuations. This phenomenon and the synchronizing tendency between the individual electron pendulums as they approach the turning regions amplify one another and establish a reasonably substantial tendency for the electrons to form a whistle effect in the interelectrode space.

The best known methods utilizing these principles are the Barkhausen and the magnetron methods. In both these arrangements the electrons are made to turn around before reaching the plate. In the Barkhausen method an ordinary three-element vacuum tube may be used. The grid is highly positive and acts alternately as accelerator and decelerator for the electrons which pass through. The plate is mostly operated at a slightly negative potential to facilitate the reversing of the electron motion before the electrons reach the plate. The path of an individual electron is shown in Figure 2.

The magnetron, as is well known, consists in its most common form of a centrally located cathode which constitutes the axis of a cylindrical anode. The anode has a high positive potential and the electrons are made to miss it by virtue of the deflecting properties of an axial magnetic field. The path of an individual electron is shown in Figure 3.

One very interesting phenomena in common for both the Barkhausen and the magnetron methods is the difference in the behavior of the oscillations when the tuning of the external circuit is approached from a state of lower or higher circuit tuning. This phenomenon is due to the fact that the voltage inducing effect from the electrons is twofold. The electrons accumulating near the plate or near the cathode cause field variations similar to such as would be produced by variations in electrode potential. The voltage drop across the resistive component of the external circuit is, however, a maximum when the electrons are in a state of highest velocity. This voltage is therefore



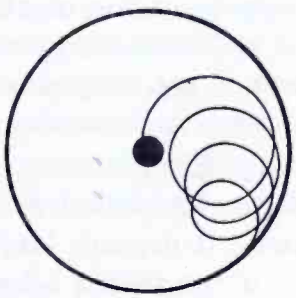


Fig. 2—Attenuated path of the individual electron when partaking in organized group motion in a triode pendulum oscillator.
Fig. 3—Attenuated orbit of the individual electron when partaking in organized group motion in a magnetron oscillator.

ninety degrees in phase lead of the voltage set up by the electron accumulations. The combined voltage on the electrodes therefore tends to make the electrons turn around sooner. The frequency is increased. Since the frequency of oscillation thus increases as the resistance between the electrodes is increased, a peculiar effect occurs when the tuning of the interelectrode circuit is varied. If the circuit resonance is made to approach the electron oscillation frequency from a lower value, the electron oscillation in the tube will recede upward and will thus have to be trailed by the circuit tuning. If the circuit tuning is adjusted above the electron oscillation frequency and then lowered, the oscillation frequency moves up to meet it. In the Barkhausen case this phenomenon has been called Gill and Morell oscillations.

Effects in the Vicinity of Electrodes

In addition to the effects of the electron motion so far considered it may also be of interest to consider the local effects in the close vicinity of an electrode.

The electric field from the electron, which at a distance covers the electrode fairly uniformly and causes a current of no definite origin

to flow through the electrode, as the electron moves, becomes more concentrated as the electron nears the electrode. The current origin in the electrode becomes more and more defined into a spot under the electron where it becomes very concentrated. The direction of this current is toward the spot if the electron is in an approaching state and away from the spot if the electron is in a receding state. These considerations, of course, do not apply to electrons moving parallel with the electrode surface. This phenomenon is naturally extremely rapid in that such concentrations do not become noteworthy until the electron is fairly near the electrode. It takes place during a very small fraction of the total transit time of the electron. Its period is therefore greatly in excess of that represented by the transit time and represents real ultra-high frequencies. These "surface oscillations" are independent of the frequency at which the device operates and depend only on number and velocity of the electrons.

Carrying the discussion a little further it becomes increasingly difficult to see where to draw a line between these "spot impulses" and heat quanta. It depends largely upon the size of the area under consideration if the period belongs to the radio-frequency region or the heat region. If the electron is headed for a landing on the electrode the spot becomes smaller and smaller until we reach the molecular and atomic structure of the electrode where the remaining kinetic energy is interchanged.

If the electron does not approach the electrode quite so close, like for instance when an electron passes through a grid structure the frequency produced, while high, is definitely one far below that of heat.

Not being organized these impulses cannot be shown in the external circuit under ordinary conditions. It may, however, be possible, by special methods to set up conditions, by using a very restricted number of electrons, whereby these oscillations may be shown.

Negative Resistance

The expression "negative resistance" is very commonly used in explaining oscillatory phenomena. Reference to it in the previous discussion has been avoided until sufficient background could be established for a clear understanding of its nature. It has been seen how in the triode feed-back circuit maximum electron transit is obtained when the anode potential is at a minimum. As the anode voltage decreases the current through the tube increases. In an electron pendulum device the electrode toward which electrons are moving has its highest negative tendency as the electrons possess their greatest radial velocity. Therefore, as the current through the tube increases the electrode

voltage decreases. In the so-called dynatron method, similar current-voltage conditions are obtained with the aid of secondary emission. The electrons are usually made to pass through a grid of high positive potential toward a plate of less positive potential. The electrons will, therefore, land on the plate with considerable velocity, causing other electrons in the plate to bounce off and be attracted by the grid. As the positive potential on the plate is increased the oncoming electrons will have a higher velocity and thus cause more electrons to bounce off and travel in the opposite direction. The total current will decrease because the ratio between electrons coming to the plate and leaving it has been decreased. All these methods, therefore, have the one characteristic in common that current and voltage vary inversely. This is characteristic of negative resistance. Since the operation of the dynatron is not based on time delay in electron transit this method, in its fundamental principle, is not adaptable to generators of extremely high frequencies. More or less developed dynatron action is, however, often obtained in conjunction with other methods whenever electrodes are bombarded by high velocity electrons.

Frequency Multipliers

Since the electron pendulum methods are critical to fields and space charges, each type of tube has a fairly well established maximum output at a particular frequency. The outputs from conventional sizes of tubes are limited to a very small portion of the power output of which they are capable when operating at frequencies where the time of electron transit is of no significance. It was, therefore, considered that if the frequency of the greater amounts of power possible to produce at the lower frequencies could be multiplied efficiently, greater power may also be realized at the higher frequencies.

The commonly known vacuum tube frequency multipliers, which are widely used for various purposes, consist of a triode, a circuit connected to the grid which is tuned to the fundamental frequency and a circuit connected to the plate which is tuned to a harmonic frequency. The grid has a negative bias and the plate has a positive potential. The frequencies used are usually well below the values at which the time of electron transit assumes significance. This shortness of the transit time is in fact an asset since the production of harmonics depends on the immediate establishment and discontinuation of electron current as the grid potential passes above and below the cut-off value. By virtue of the plate power the device also operates as an amplifier.

As the input frequency is increased, the electrons will remain in the process of transit while the electrode potentials may vary consider-

ably. Like in the triode feed-back circuit this is very detrimental. As the grid potential goes below cut-off value, it first retards the electrons moving toward it from the cathode. The electrons therefore deliver the energy they chiefly received from the direct-current source into the external oscillating circuit and the grid receives an impetus in accord with its negative swing. As the grid continues to become more negative the electrons will finally come to a standstill whereafter they will be accelerated back toward the cathode. They will then draw power from the external circuit. The acceleration received in this direction is greater since the grid is now very negative. The total effect of the motion from the cathode and back again is, therefore, a loss. The electrons in the grid-anode space received added acceleration in their original direction as the anode potential swings in a more positive direction. This accelerating power is thus also obtained from the oscillating energy in the external circuit. Both the fundamental frequency circuit and the harmonic circuit, therefore, lose more and more of their power as the frequency is increased.

Due to the increase in potential gradient toward the central cathode in cylindrical element tubes the greatest effects upon the motion of the electrons are, of course, obtained in this region. Electrons which emerge from this region have a more predetermined motion for the rest of their path than electrons in flat element tubes. This condition makes the electrons passing through the interelectrode space of cylindrical tubes less sensitive to variation in electrode potentials, during a great portion of their journey, than in flat electrode tubes with uniform fields. This fact, of course, holds regardless of the methods of oscillation employed.

At the lower frequencies the grid simply acts as a throttle which lets electrons through to the anode during a short portion of the cycle. The electron accelerating power required for the production of this motion and the subsequent harmonic it induces is thus chiefly taken from the direct-current plate supply. As the frequency is increased the number of electrons which do not complete their transit from cathode to anode becomes a greater and greater portion of the total number of electrons which enter the interelectrode space. The power stored in these electrons, as they are pushed toward the electrodes, is taken from the alternating-current component of the external circuit. It is, therefore, clear that as the frequency is increased, the frequency multiplier becomes less of an amplifier. The fundamental frequency input, instead of being only a guide to the electron performance, becomes a driving force. Since the frequency multiplier outputs at higher frequencies must derive the driving power from the fundamental frequency input it was necessary to find a way by which the motion of the electrons,

thus driven, could be organized to converge, so that they would, during a certain phase of their motion, fall through an appreciable portion of the interelectrode space in the short time corresponding to a harmonic current. It was found possible to accomplish this by introducing a magnetic field which is perpendicular to the electric field. In a cylindrical tube this then becomes an axial magnetic field. The effect of the magnetic field may be most easily understood by referring to the curves in Figures 4, 5, and 6. For the sake of simplicity a tube with two

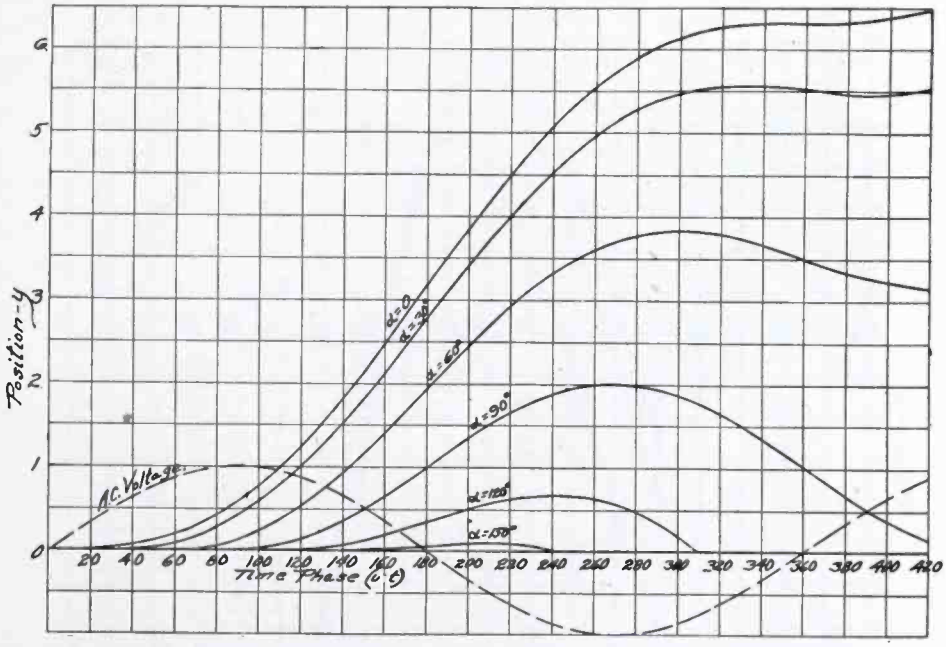


Fig. 4—Position versus time of electrons in an alternating-current field between flat electrodes. Each curve represents electrons which have left the cathode during a certain phase of the alternating-current cycle.

parallel plane electrodes is being considered. Alternating voltage is applied of such frequency that the time of transit is too slow for crossing the interelectrode space. The positions of the electrons leaving the cathode at different times, under the influence of the positive half cycle of the voltage on the other electrode, will be distributed throughout the space as shown in Figure 4. From these curves it can be seen that there is a tendency for the electron positions to become more and more divergent with time. Even a rectifying action can be observed in the motion of the electrons which leave the cathode during the first half of the positive half cycle. If a magnetic field, perpendicular to the electric field, is introduced it can be given such a strength that the electrons which leave the cathode successively, during the generous time period of a half cycle of the fundamental frequency, will all com-

plete their return journey to the cathode within a very definite time limit. (See Figure 5.) By integrating the velocities, the curve shown in Figure 6 is obtained. It represents the shape of the resulting electron current curve. A transformation of the fundamental frequency power directly into a harmonic component thus becomes possible. The electron stream exhibits a positive resistance to the fundamental frequency current and a negative resistance to the harmonic. As the electrons return to the cathode they still possess considerable kinetic energy. If it is attempted to lower this kinetic energy by further

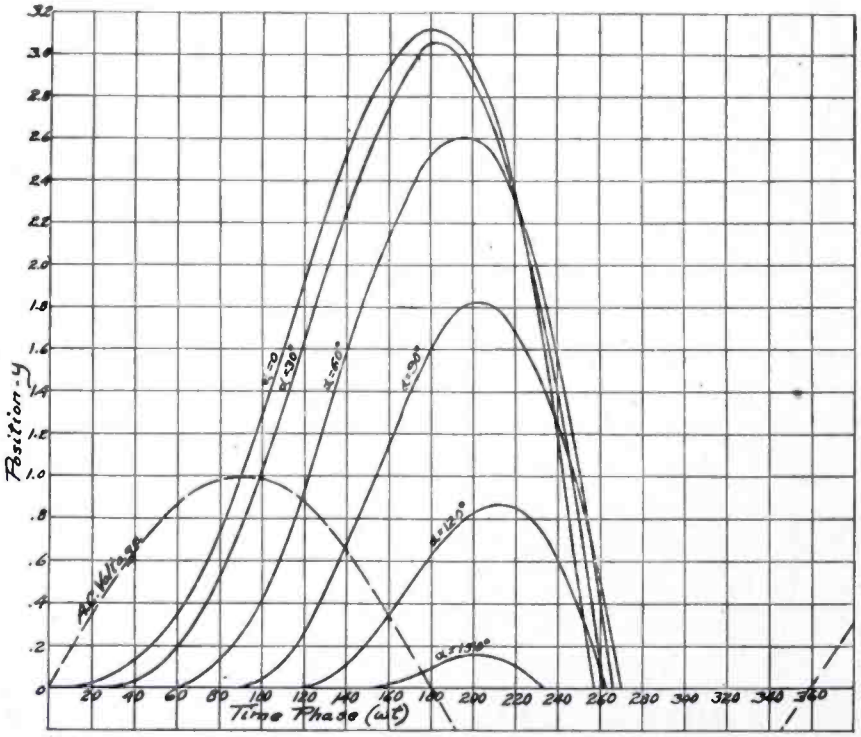


Fig. 5—Organization of positions shown in Fig. 4 when introducing a steady, transverse magnetic field. Note how the electrons, returning to the cathode, converge in respect to time.

circuit loading, the disorganizing effect from the load prevents further gain in efficiency. The efficiency also depends largly upon which harmonic is chosen. While the efficiency obtained when producing the second harmonic is high, the efficiency of the generator of the fundamental frequency oscillations is low since it has to operate nearer the border limits of the triode feed-back circuit than when a higher harmonic ratio is used. The multiplier efficiency, however, drops very rapidly as the harmonic ratio is increased. Fewer electrons become subject to complete organization. When odd harmonics are produced, two tubes can be operated in push-pull fashion. Very simple circuit arrangements are then obtained. From these considerations, the third harmonic has often been chosen as the most satisfactory compromise.

Among the standard tubes available, some three-element tubes happened to be the most satisfactory for frequency multiplication. Examples of such tubes are shown in Figure 7. Although it has been found most efficient to let all the elements become part of the circuit either for bias purposes alone, or for tuning purposes as well, the cathode and the grid are, of course, the essential elements between which the major electron performance takes place. The grid is negatively biased and the plate is given zero or negative potential. The input and the output circuits can be applied to any one of the elements as long as the combined circuit tuning produces potentials in coordination with the desired electron motion. Different values of magnetic

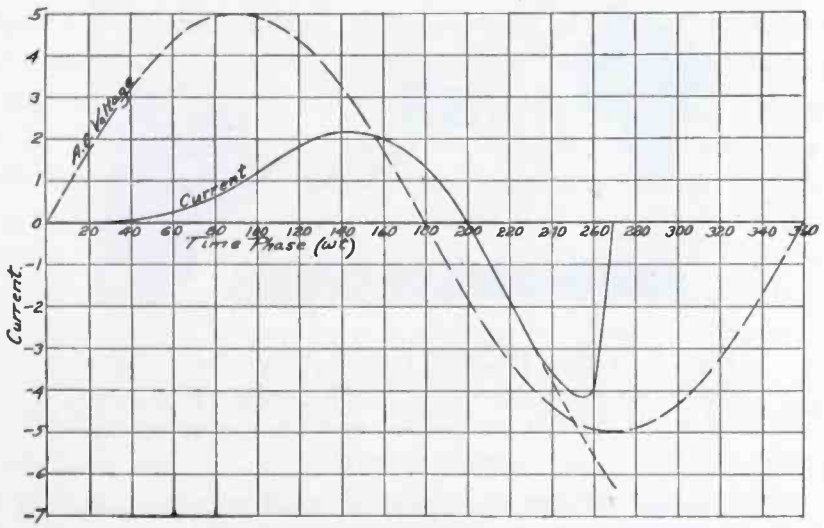


Fig. 6—Electron current curve corresponding to conditions shown in Fig. 5.

field strength are sometimes necessary when choosing various modes of connection.

The efficiency obtained with standard, cylindrical element triodes as triplers is about ten per cent. The efficiency gain by using a magnetic field varies but has so far been found to be at least three times that obtained without magnetic fields.

Two advantages result from the use of frequency multipliers. The power obtainable with a tube used as a frequency multiplier is many times greater than that which it would deliver as a pendulum oscillator. Frequency control circuits of great accuracy may be employed.

Frequency Control

One of the advantages of very high frequencies for communication purposes is the insignificant width of the modulation band as compared with the frequency of the carrier. This makes it possible to consider a

great number of communication channels within frequency limits only a few per cent apart. In order that full advantage may be taken of this situation it is, however, clear that as the carrier frequency is increased its relative stability must also be increased. Aside from this consideration, very accurate frequency control is required to permit maximum selectivity in the receiver in order to obtain optimum signal-to-noise ratio. The transmitters have been required to provide stability sufficient for reception by means of very selective superhetrodyne receivers, without loss of signal quality. Special efforts have been made to produce oscillator outputs as free as possible from undesirable ampli-

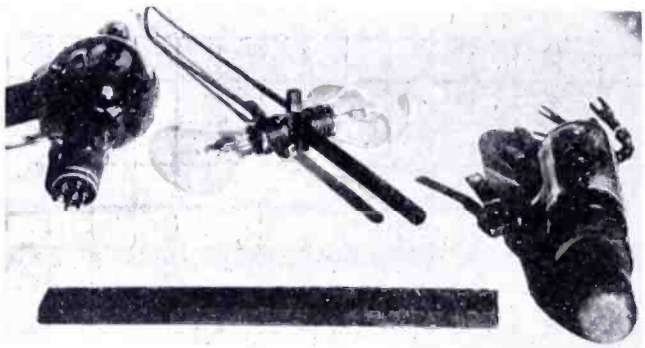


Fig. 7—Left: UX-852 transmitting tube; very useful all-around tube for triode feed-back circuits (Fig. 1), triode pendulum, and frequency multiplier circuits.

Middle: 300-600 megacycle wavemeter and output indicator. Note the two 50-watt load lamps.

Right: RCA-846 (UV-846) water-cooled tube, useful for triode feed-back circuits (Fig. 1) and for frequency multipliers.

tude and frequency variations. Since the difficulties of eliminating these disturbances increase in proportion to the oscillator frequency, the utmost of every available means for frequency stabilization had to be mobilized. Vital circuits must be well shielded against electric or magnetic influences of variable character. It is also very important that the frequency determining parts of the circuit be mechanically separated from their surroundings in order to avoid disturbance from vibrations. In coupling up with output circuits such as transmission lines and antennas, great care must be taken that a minimum of coupling is being used. The reaction from variation in the constants of such circuits, due to wind, precipitation, etc., will otherwise become very disturbing. Self-bias, by means of leak resistances and self-rectification, often results in a tendency to maintain constant tube

impedance if introduced in correct proportions in various branches of the tube circuit.

It has often been found necessary to provide smoother anode potentials than normally required at the lower frequencies. Variations in the direct-current supply have been eliminated by means of a regulator consisting of a combination of resistors with a vacuum tube, Figure 16. The grid of the regulator tube is connected to a potentiometer inserted between the direct-current power supply and ground through a source of constant negative bias such as a battery. The function of this bias source is to overcome the positive voltage drop in the potentiometer so that the grid may be maintained at proper operating potential. The anode of the regulator tube draws current from the direct-current power supply through a resistor. The oscillator power is supplied through the same resistor. As an example, an increase in voltage across the direct-current power supply makes the grid of the regulator tube less negative. This increases the current through the common regulator resistor through which both the anode of the regulator tube as well as the anode of the oscillator are supplied. An increase in the voltage drop across the regulator resistor thus takes place. If the device is properly adjusted, the rise in voltage is completely compensated by the drop across the regulator resistor. A front view of such a regulator device is shown at the right in Figure 14.

In addition to such measures it is very desirable to incorporate special frequency controlling devices such as crystals and low power factor tank circuits. Since crystals, of reliable performance above a few megacycles, are difficult to obtain, it is necessary to use many stages of frequency multiplication in conjunction with this type of master oscillator. Low power factor tank circuits can be applied more directly and, since they otherwise compare favorably with crystals, they are more practicable as frequency stabilizers at very high frequencies.

Modulation

In modulating an oscillator it is sometimes quite difficult to obtain pure amplitude modulation which is not distorted by variations in phase or frequency of the oscillation. This difficulty is rather outstanding with the electron pendulum type oscillators.

Higher degrees of purity may, however, be obtained by resorting to compound modulation, i.e., by applying the modulation energy in at least two different ways and in such a manner that components representing undesirable modulation effects are cancelled. In a Barkhausen oscillator it is, for instance, possible to obtain amplitude modu-

lation substantially free from frequency variations by modulating the plate and grid cophasially with correct proportions of modulating voltage. An increase in the positive grid potential will tend to increase the electron velocity while a simultaneously applied decreased negative potential on the plate will increase the length of path. The frequency of the electron oscillation may thus be kept constant while the flow of the electron current is being modulated. In order to obtain linear amplitude modulation over as large a range as possible, the power output of the nonmodulated carrier must, of course, be reduced since the maximum output of electron oscillators is very definite.

In master oscillator frequency multiplier types of transmitters, undesired frequency modulation is less prevalent, but may sometimes still be desirable to eliminate inasmuch as any variation in the load on the master oscillator will influence its frequency stability. Compound modulation, of various types, may thus be used to advantage.

One very effective way of producing amplitude modulation is to control the power output by interposing a modulator between the transmitter and the antenna. For producing a steady tone, such as required for interrupted continuous wave telegraphy, a commutator across the transmission line is very effective. For more universal purposes a vacuum tube circuit is, however, required. At frequencies above 200 megacycles most vacuum tubes now available are not capable of furnishing efficient conditions for electron transit. In the description of electron pendulum oscillators it was pointed out that the orbital motions of the individual electrons are subject to decay as they deliver their energy to the external circuit. In the case of an absorber the electrons must instead absorb energy from the external circuit and deposit the so acquired kinetic energy in the form of heat on the electrodes or the electrons must be so organized that they will serve efficiently as intermediary links by which the energy is efficiently directed into resistive load circuits. If the electrons themselves are to be made to absorb energy, conditions must be set up which give to the electrons a natural period, or the tendencies for such a period, near the period of the energy to be absorbed. In such a case, the orbital motion of the electrons will be of increasing amplitude so that its kinetic energy will increase. This principle has also been verified in the development of apparatus for the production of high speed ions for the purpose of bombarding atomic nuclei. Periodic oscillations of increasing amplitude are then given to ions instead of electrons.[3]

[3] E. O. Lawrence and M. S. Livingston, "The Production of High Speed Light Ions Without the Use of High Voltages," *Phys. Rev.*, April 1, (1932).

Circuit Considerations at Ultra-High Frequencies

If the electromagnetic field, established around an electrical circuit carrying oscillating energy, has dimensions comparable to the wavelength, energy will be radiated.

Radiation may be reduced by using a balanced or symmetrical circuit, one whose adjacent, corresponding parts give rise to electromagnetic fields of equal amplitudes at opposite phase. The circuit can actually be built symmetrical as is done in push-pull circuits or as in two-wire transmission lines. An originally nonsymmetrical circuit may be located near a conductive surface, Figure 8.

In the ultra-high-frequency technique, linear conductors are often used for tuning instead of circuits with lumped constants. The relative merits of various combinations of tuning circuits may sometimes be judged by the total number of resonance points obtained. The fewer

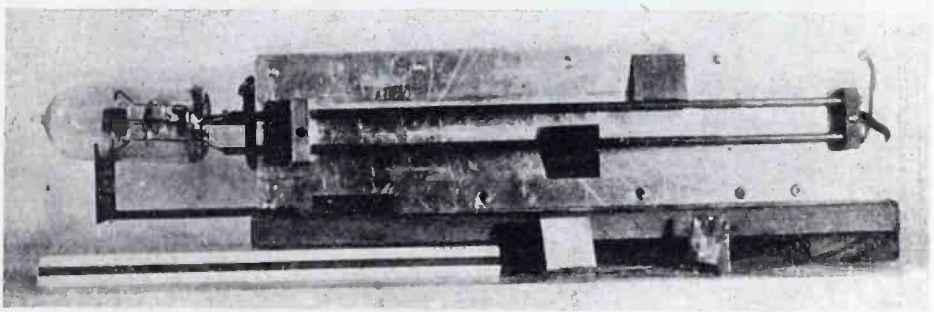


Fig. 8—FP-126, single tube, triode pendulum oscillator. Note the arrangement of the nonsymmetrical circuit near a shield and the wedge-shaped blocking capacitors which serve as tuning sliders. They are wedged between the conductor and the shield. The surface facing the shield is lined with mica.

such degrees of freedom, the better is the circuit. These resonance points arise from many conditions. At very high frequencies, the tube elements and the internal leads of a vacuum tube form an appreciable portion of the total tuned circuit. When continuing these elements and their leads with linear conductors, of dimensions giving constants which match, no additional degree of freedom is obtained. When using external circuits with lumped constants, the total circuit does, so to speak, more definitely consist of several distinctly separated reactance components. These components may neutralize each other in various combinations, creating several resonance points. An additional source of multiple degrees of freedom is also provided by the coupling phenomena created between the various circuits by the capacity combinations between the vacuum tube elements. This effect adds resonance points to linear tuning systems as well as to tuning systems with lumped constants. The resonance points so far referred to, are usually

quite close in frequency. The linear type of tuning circuit may, in addition to these, respond to octaves or harmonics of the frequency in which they are fundamentally tuned.

The choice of tuning method is thus determined by many factors. For low frequencies, the linear tuning method becomes impracticable on account of the large physical dimensions required. As the frequency is increased and it becomes possible to use linear tuning, a great number of harmonics may be obtained if the oscillator is still capable of producing an output at the harmonic frequency. By using circuits with lumped constants the harmonic feature may be more easily avoided. The trouble with coupling frequencies may be minimized, or avoided, by so adjusting the circuit that only one of these resonant

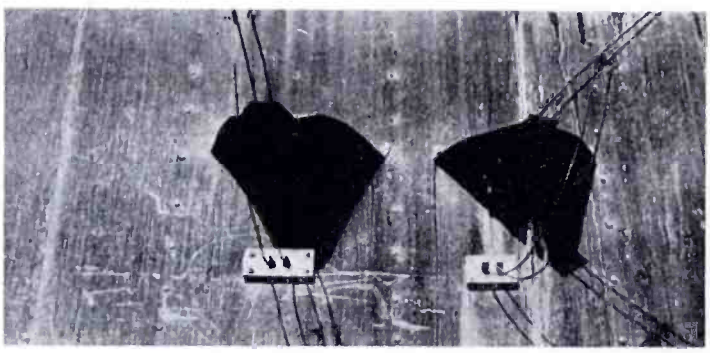


Fig. 9—Transmission line lead-in detail, showing high impedance metallic suspension links sometimes called "metallic insulators."

frequencies conditions, such as voltage and phase, become of controlling influence. At still higher frequencies, where special methods are used and in which the direct-current potential of the electrodes, in a larger measure, controls the frequency, the choice of tuning methods becomes more arbitrary. It may then be determined by other factors such as mechanical features, cost, space required, etc. When the frequency is so high that the first voltage nodal point appears on the lead inside the glass envelops of a vacuum tube, or in some other inaccessible place, the external circuit must be given characteristics equivalent to the addition of a conductor extending half a wave, or a multiple thereof, from this nodal point. Such arrangements are especially successful in master controlled circuits which are void of the ability of self-oscillate.

Linear conductors may, at very high frequencies, replace insulating supports. Such conductors must then be given a certain length so that their impedance is very high at the point of support. For this reason they must also be arranged to be nonradiating in the same manner as linear tuning circuits. A quarter-wave conductor with one end grounded

will have high impedance at the other end. For the support of a push-pull circuit a pair of such conductors may be used. Such a pair then corresponds to a U-shaped conductor half a wave long from end to end, with an electrically neutral point at the bend. The losses incurred from the use of these metallic insulators are usually much lower than obtainable with insulating material. Their mechanical strength is superior and their disturbing effect upon the electrical circuit is

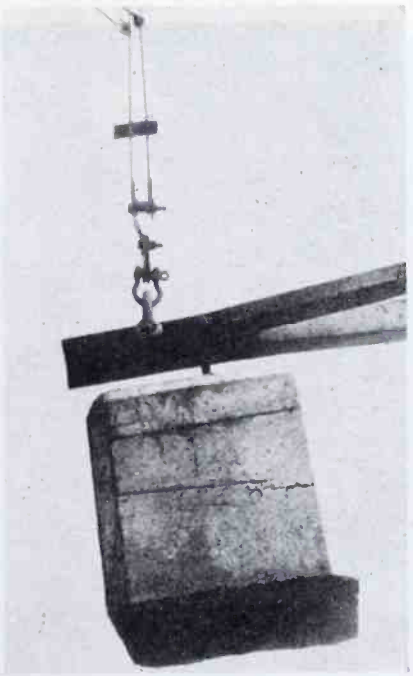


Fig. 10—Bend in a transmission line showing application of metallic insulator.

smaller. They have been found particularly useful in transmission-line work down to frequencies as low as fifty megacycles. Applications are illustrated in Figures 9 and 10.

It has been found that stranded wires cause very high losses. The reason is that the twisting increases the inductance of the individual strands. At very high frequencies a considerable portion of the current will, therefore, force its way over the shorter and lower reactance path from strand to strand either capacitively or conductively. Since such conductors are seldom clean, the losses at such crossovers are very high. This point requires special consideration in connection with some of the standard transmitting tubes which are equipped with stranded connecting leads. It is best to eliminate these leads whenever possible.

For obtaining a direct grounding effect on the filaments of a tube it is sometimes necessary, for reasons already mentioned, to extend the length of these leads to make them one-half wave long. In less conven-

tional circuits certain improvements in the phase and voltage conditions between the electrodes may be obtained if it is possible to tune the filament circuit. Whether linear conductors or circuits with lumped constants are used, the easiest way of supplying the filament heating current is to make the radio-frequency conductor of tubular material and to locate the leads for the heating current inside this conductor. Small blocking condensers should then be used at the filament end to

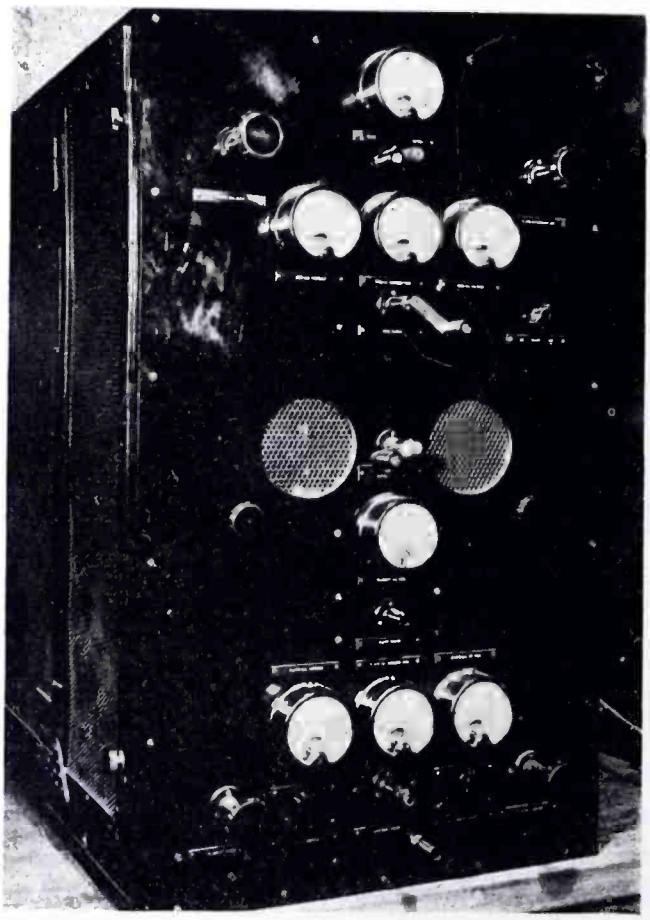


Fig. 11—Front view of 462-megacycle, 6-watt output, UX-852 tube, push-pull transmitter of the triode pendulum oscillator type.

prevent radio-frequency currents from entering the heating circuit. Such arrangements are illustrated in Figure 18.

When using larger tubes, of the water-cooled type, the water can be supplied to the tube jackets through metal tubing which also serves as tuning conductor. The supply and return water may be handled either through concentric or independent tubings. The latter is sometimes convenient in that the two parallel tuned circuits thus obtained allow independent branches for tuning and coupling. (See Figures 17

and 18.) If these leads also carry high direct-current potential the water circuit is continued through a rubber hose of adequate length. This hose is usually attached at a point neutral to radio-frequency voltages.

During experimental work it may often be found difficult to get a new circuit going. Since it is much easier to study the sources of trouble if a circuit is oscillating, it is sometimes valuable to use an auxiliary trigger or a "catalyst" circuit. The functions of such a circuit are manifold. The conductive losses of the oscillator circuit may be too high. There may be an excessive radiation resistance or some phase and voltage condition may need correction. A catalyst circuit

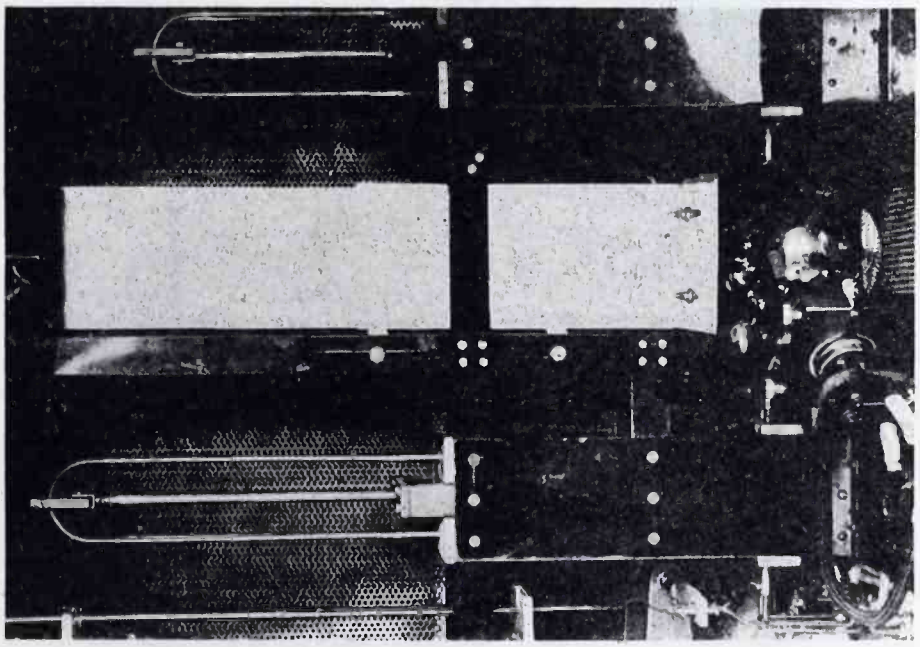


Fig. 12—Side view of 462-megacycle, 6-watt output, UX-852 tube, push-pull transmitter of the triode pendulum oscillator type, showing "catalyst" circuit.

has, therefore, no definite form but must be made up to suit the suspected condition. A wide strip of copper, a half wave long, may be bent around in different ways and placed in various positions near the tubes or the associated circuits, Figure 12. If, for instance, the trouble consists of too high resistance in the flexible tube leads, some of the external circuit tuning may be taken over by the catalyst circuit since it may be placed in such a position that it is capacitively coupled to the tube elements directly through the glass envelops without the aid of the leads. Phase conditions may sometimes be corrected by means of tuned loops. Most forms of catalyst circuits are applicable in reducing radiation. The location of shields at certain distances from the oscillator may also sometimes produce desired results.

Example of Transmitting Equipment

In order to obtain as complete design data as possible the experimental models have in some cases been developed to such a point that they would, as nearly as possible, conform with the exacting requirements of commercial operation. Many of the considerations in such designs have already been discussed in a general way. Examples of such equipment need, therefore, only a brief description.

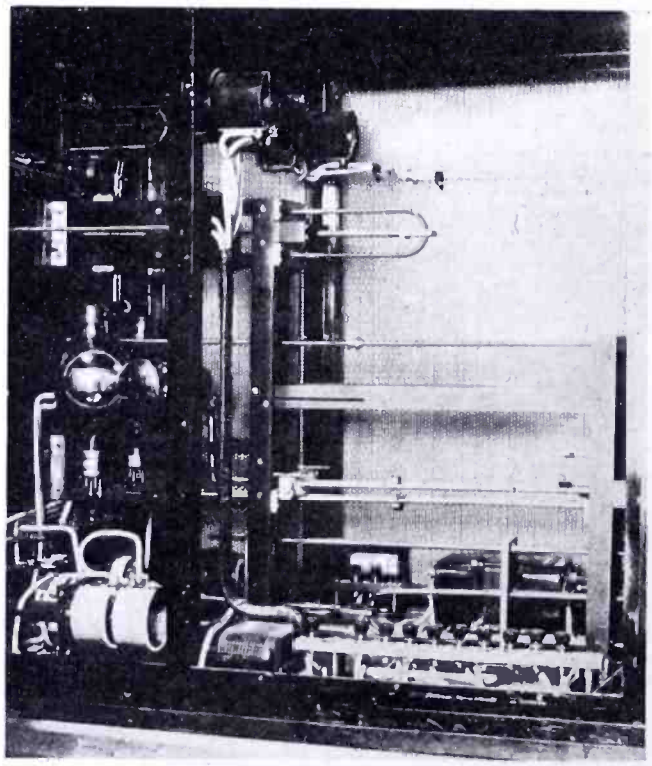


Fig. 13—Side view of 462-megacycle, 6-watt output, UX-852 tube, push-pull transmitter of the triode pendulum oscillator type, cathode circuit tuning is here being used.

1. UX-852 Barkhausen Transmitters

The first two transmitter units were built to operate at a frequency of 462 megacycles and a wavelength of 65 centimeters. For about a year they were used for two-way telephony between Rocky Point and Riverhead. These transmitters, which are shown in Figures 11, 12, and 13, are of the triode electron-pendulum type. Each transmitter consists of a pair of UX-852 tubes, operated in push-pull. The tuning circuits are of the linear conductor type and of the trombone variety. In one transmitter only the plate and the grid circuits were tuned. The filament heating current was supplied through choke coils. The oscillator was equipped with a catalyst circuit, Figure 12. This circuit consisted of sections of three-inch wide metal strips forming a U of

variable length similar to a trombone circuit. At the end of this U were flanges, each facing one of the oscillator tubes. As an alternative the other transmitter was equipped with a tuned cathode circuit, Figure 13. On account of the heating current leads inside, the filament tuning conductors cannot conveniently be of the trombone type but must instead be tuned by a sliding connector. It made little difference which circuit was coupled to the load. The power was taken from the plate circuit in this case. Potentiometers for individual control of the direct-current electrode potentials of the two tubes were also introduced. All tuning, coupling, and potential regulation was performed

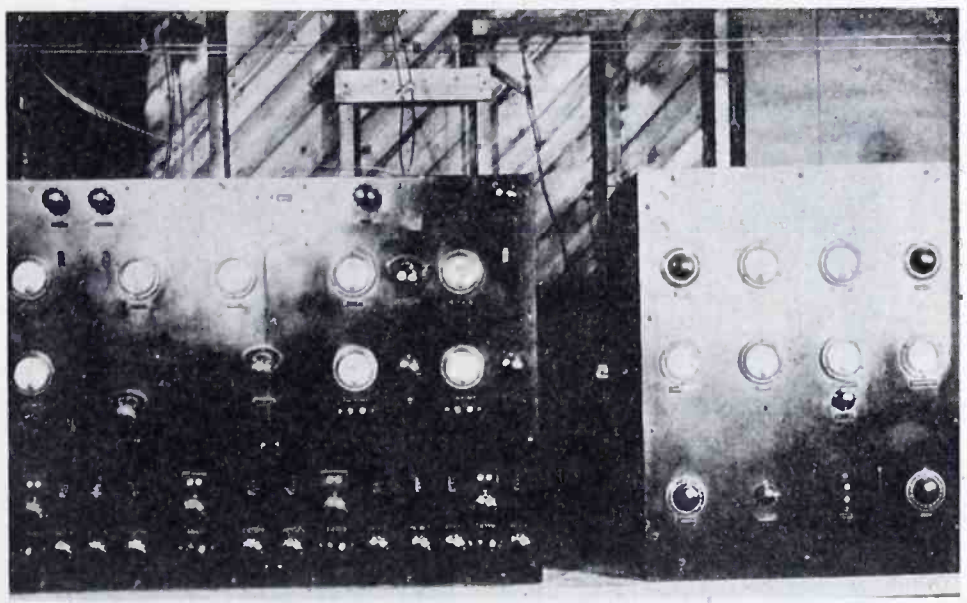


Fig. 14—Left: Front view of 432-megacycle, 15-watt output transmitter, consisting of a 144-megacycle oscillator, a buffer amplifier, a frequency tripler, and an amplitude modulator of the absorber type. All stages are of the push-pull type, using UX-852 tubes.
Right: Front view of plate voltage regulator for the plate supply of the oscillator stage and the buffer stage.

from the front of the panel. Radio-frequency ammeters were inserted at points neutral to radio-frequency voltages. It was due to the desirability of keeping these meters in a fixed position and on the front of the panel, that the circuits had to be looped over and be tuned by trombones.

This transmitter was compound modulated. A special modulator, shown in Figure 23, was built for this purpose.

The positive grid voltage required at 462 megacycles was 500 volts. The negative plate voltage was 125 volts. The normal filament voltage is ten volts for the UX-852 tubes. In order to obtain suitable emission of 250 milliamperes per tube it was necessary to cut the filament volt-

age down to between seven and nine volts. Due to the great heating gradient from the grid, the filament emission would at times become unstable. Some tubes would eventually become stable at about nine volts. This phenomena was credited to the sensitivity of the thoriated filaments used in this type of tube. A power output of six watts was obtained from these transmitters.

2. UX-852 Frequency Multiplier

As a result of the experiments with frequency multipliers, a transmitter as shown in Figures 14, 15, and 16 was built. It is of the push-pull type throughout. The first stage may either self-oscillate or be driven at about 144 megacycles. The output from this circuit drives

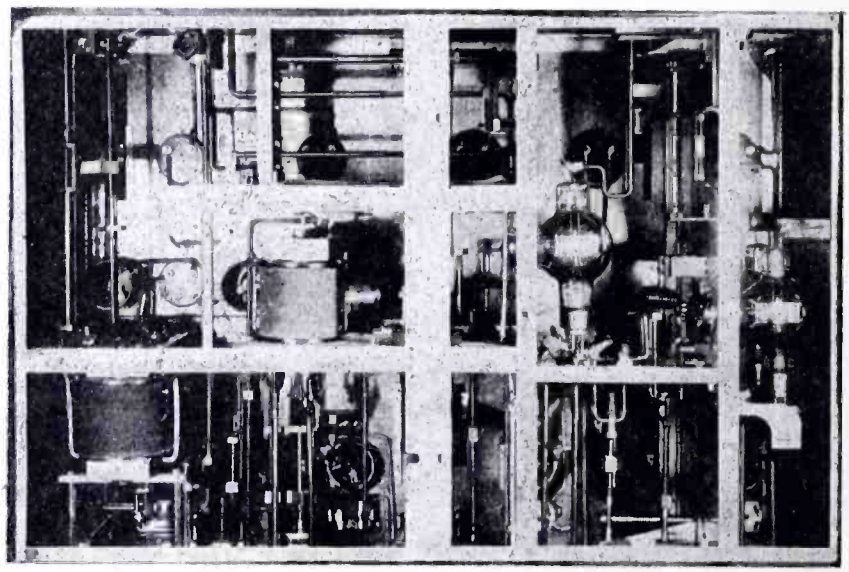


Fig. 15—Rear view of the 432-megacycle, 15-watt output, multiplier type transmitter shown on the left in Fig. 14.

an amplifier, or buffer stage, which gives a power amplification of two to one. This is the best amplification that could be obtained with the UX-852 tubes at this frequency. The power from this buffer is fed into a tripler circuit which delivers 432 megacycles output at about ten per cent efficiency. The tripler is equipped with means for producing an axial magnetic field of about 150 gauss. The power output obtainable from this transmitter is fifteen watts. All the tube elements are tuned and the tuning circuits are of the linear conductor type and of both the trombone and slider variety, depending upon the need.

One of the reasons for studying frequency multipliers has been the need of greater frequency stability and the particular adaptability of this method to frequency control. Crystal control was one of the first methods to be tried. An auxiliary control circuit was built which con-

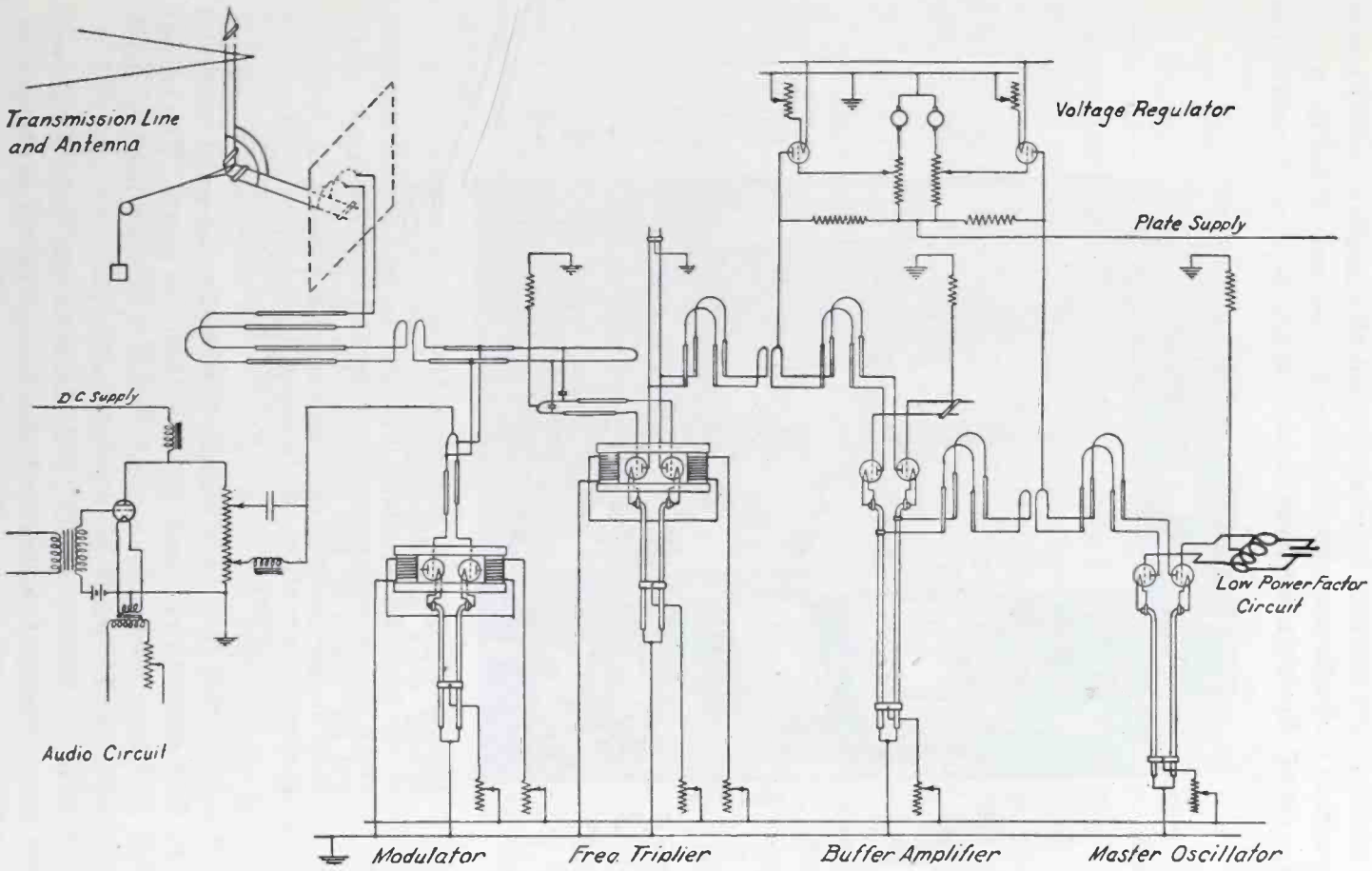


Fig. 16—Schematic wiring diagram of the 432-megacycle, 15-watt output, multiplier type transmitter shown in Figs. 14 and 15.

sisted of a crystal oscillator operating at about two megacycles followed by six stages of frequency multipliers. The crystal circuit and the first four stages were single sided circuits employing UX-210 tubes. The last of these stages fed into the combination of a single UX-860 buffer and a single UX-852 doubler driving a push-pull UX-852 tripler circuit, raising the frequency from 48 to 144 megacycles. It was found

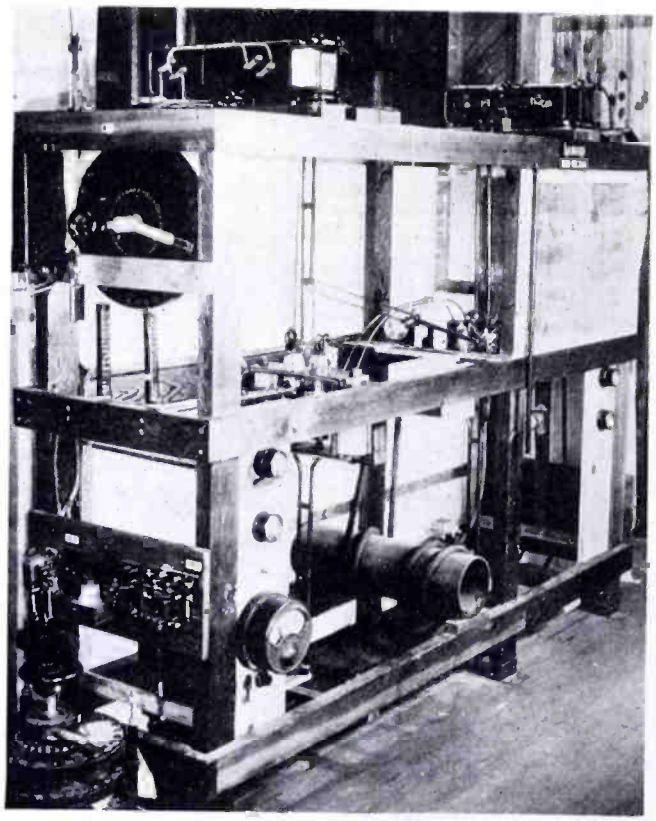


Fig. 17—Frequency multiplier transmitter, consisting of a 137-megacycle, 1500-watt output, triode feed-back oscillator (in the background) and a tripler, with organizing magnetic field, giving an output of 115 watts at 411 megacycles. Both stages use RCA-846, water-cooled tubes in push-pull formation. Note lamp loaded wavemeter (same as in Fig. 7).

beneficial, even in this frequency region, to introduce an axial magnetic field. The field strength required was less than 100 gauss. The output from this auxiliary frequency control circuit was then used to drive the first 144-megacycle stage of the transmitter.

For modulation, a push-pull UX-852 absorber was used which was connected across the antenna transmission line as shown in Figure 16. The length of the connecting leads to the absorber were such that optimum modulation effect would be obtained. These leads were connected to the grids of the absorber which act as modulation terminals.

The plates may be left floating since they have no part in the circuit. The grid and the filament circuits are tuned. The grids are given a positive bias of about sixty volts, which causes them to draw a current of about five milliamperes each at an axial magnetic field of 200 gauss. The modulation voltage as well as the radio-frequency carrier to be modulated were then superimposed upon this grid condition. In this way it was possible to obtain linear amplitude modulation of about 35 per cent.

Numerous arrangements for replacing the complicated crystal drive, with its many multiplier circuits, have been tried. These arrangements consisted of various forms of low power factor circuits connected to the grids of the 144-megacycle oscillator. Some of these arrangements have proved very successful in that stability approaching that of crystals was obtained. Data on such circuits have already been published and further information may be published later.

3. RCA-846 Frequency Multiplier Transmitter

Although the design of a semicommercial model of this apparatus has not been finished, at the time of writing this paper, it appears to have sufficient points of interest to warrant being included among the transmitter examples in its experimental form, Figures 17 and 18. It consists of a push-pull oscillator operating at 137 megacycles furnishing power to a push-pull frequency tripler. The tripler is provided with an axial magnetic field. All the tubes are RCA-846 water-cooled tubes. The tuning circuits are of the linear conductor type and arranged similarly to the circuits for the UX-852 multiplier. The tubular plate tuning conductors also carry the water supply to the plate jackets.

Several optimum circuit combinations, each requiring a different magnetic field, could be obtained. With an input to the tripler of 1200 watts the best output obtained when using a magnetic field was 115 watts at an output frequency of 411 megacycles. The magnetic field required for various circuit arrangements varied between 100 and 300 gauss. When using no magnetic field the best circuit arrangement would give an output of 35 watts.

4. Triode Feed-Back Circuits

Great progress has been made recently in the development of triodes for operation, according to the conventional feed-back method, at very high frequencies. Transmitting tubes have been developed which will give good output up to as high frequency limits as 300 megacycles. (ZP-94 curve in Figure 1.) It is therefore no longer necessary to resort to special methods for frequencies of this magni-

tude. These developments have, however, also resulted in a proportional increase in the frequency borders of the frequency multiplier method and it is now possible to consider this very stable method in the 1000-megacycle region.

Antennas and Transmission Lines

In designing directional antenna systems, choice can be made between two general principles. The first one, which may be called the Hertzian method, consists in redirecting the radiated energy from a

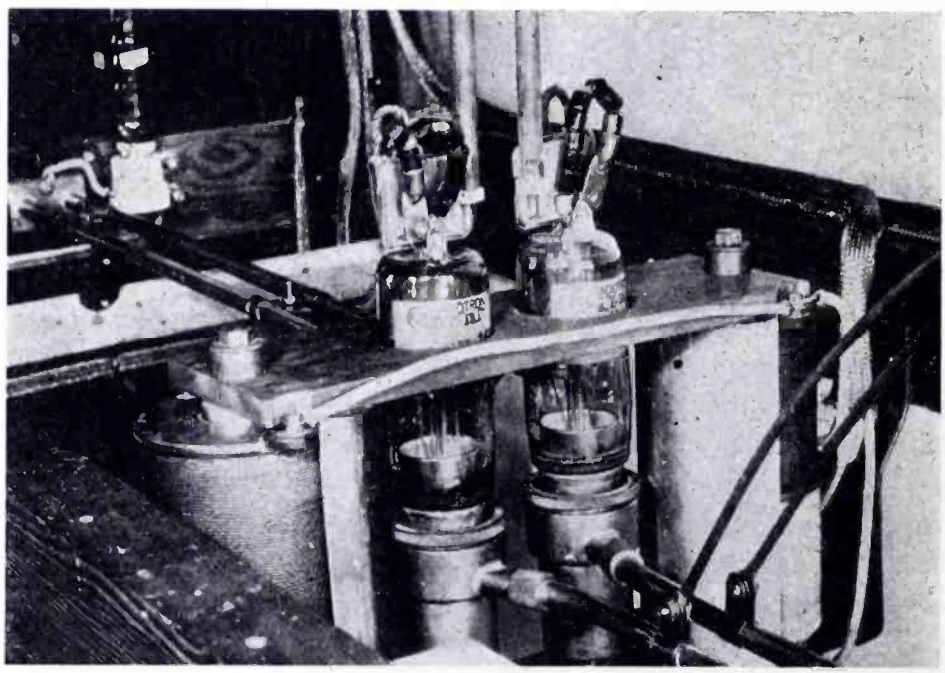


Fig. 18—411-megacycle, 115-watt output, RCA-846 water-cooled tube, push-pull tripler with parallel coil system magnet for producing organizing magnetic field.

single nondirective radiator into one direction. This is done by means of a parabolic, metallic reflecting surface or by means of conductors arranged in a parabolic array. The perfection of the method depends upon the size of the reflector system, measured in wavelengths.

In the second method no reflecting surfaces need be employed. The origin of radiation is instead distributed. A number of radiating elements, such as dipole antennas may be so spaced and phased that they add up very efficiently only in diametrically opposite directions. By further combination, radiation in one of these directions may also be eliminated.

Several years ago the method of combining straight harmonic radiators instead of dipoles was introduced by R.C.A. Communications,

Inc.[4] This principle has greatly simplified the details of directive antenna design. Detailed descriptions of such systems have been published previously in the *Proceedings*.[5] As the name indicates, the harmonic wire antennas are based on the standing wave principle. When the frequency is sufficiently high so that it becomes possible to make the length of the radiators on the order of 50 to 100 wavelengths, the radiation attenuation becomes so great that very little energy reaches the far end of the wires. The antenna then becomes unidirectional and aperiodic. This is a great simplification since it makes reflector systems and tuning arrangements unnecessary. An antenna of this type is shown in Figure 19. It is commonly known as a V antenna.

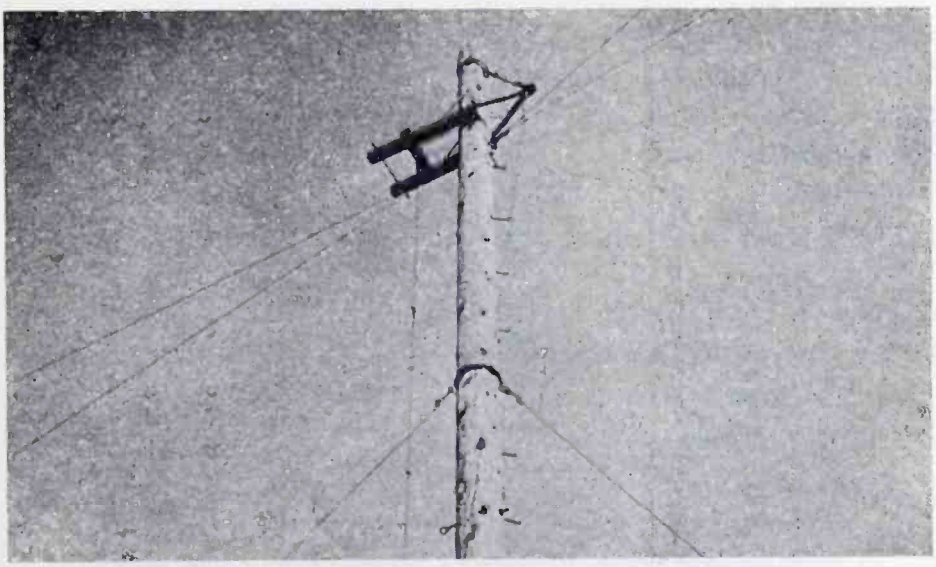


Fig. 19—Feed end of the V antennas.

A V antenna 100 wavelengths long should have an angle of 9.8 degrees between the wires. The power gain, over a doublet, is more than 100.

It may at times be desirable to provide an antenna entirely free from back radiation. Such an occasion may arise when the transmitting and receiving antennas have to be closely located and operate on near-by frequencies. An antenna having a metallic sheet reflector is then desirable since it is more perfectly free from back radiation than other antenna types. In order to avoid curved reflector surfaces, antennas as shown in Figure 20 were designed. This antenna consists of rectangularly bent wires with half-wave distances between the bends. By combining several such wires an arrangement is obtained in which

[4] Lindenblad, U. S. Patents No. 1,884,006 and No. 1,927,522.

[5] P. S. Carter, C. W. Hansell, and N. E. Lindenblad, "Development of Directive Transmitting Antennas by R.C.A. Communications, Inc," *Proc. I.R.E.*, Vol. 19, pp. 1773-1843; October, (1931).

all the vertical elements are free to radiate and are of the same phase, whereas all the horizontal ones have their radiation cancelled by a horizontal member of another wire in juxtaposition. The various bent wires are attached to a common feeder line in pairs, at half-wave intervals. The wires are supported by means of insulators or metal columns at the voltage nodal points. The whole system is mounted in front of either a metal screen or a solid metal sheet at a distance of about a quarter of a wave or an odd multiple thereof.

Some workers have followed the practice of building the transmitter and antenna system together as a single unit. This is undesirable in many cases because the antenna must be at a great height to obtain

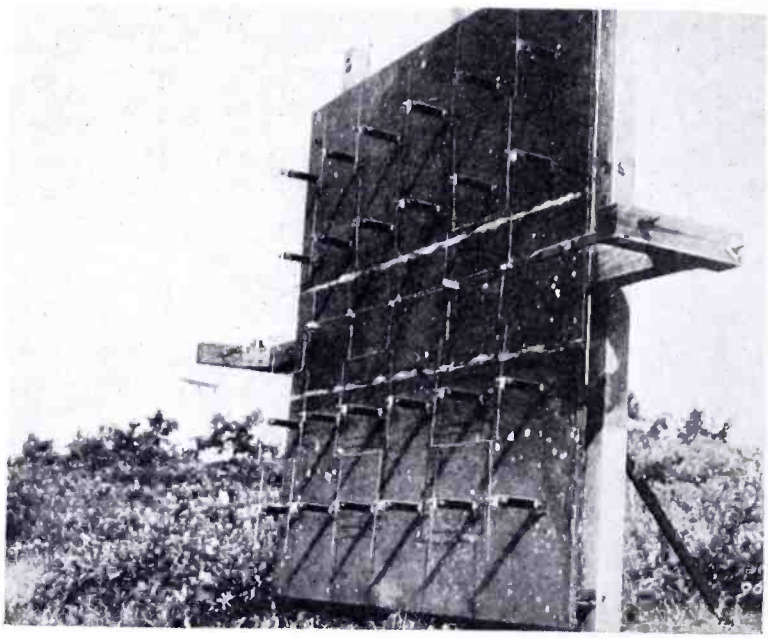


Fig. 20—"Billboard" type, directive antenna.

great range. A transmitter located at the antenna cannot be serviced conveniently. Consequently, it is often desirable to locate the antenna and transmitter some distance apart, requiring the use of transmission lines of considerable length. On most occasions this has been the practice. A study of the characteristics of such lines had therefore to be made. Balanced lines using No. 4 B&S and No. 6 B&S wire were investigated. Inductive coupling to the line under test was employed to minimize the effect of reflections and unbalance in the connecting lines from the transmitter. The measurements on the lines were made with the aid of a sensitive thermogalvanometer which made it possible to use very small capacitive coupling to the wire system, thus avoiding disturbance of the natural capacity and resistance of the line. The relatively large amount of power (100 watts) available from the fre-

quency multiplier transmitter was an important factor in making it possible to use very light coupling between the measuring instrument and the wires. The measurements included determining losses and also velocity variations due to moisture on the wires. The loss measurements were carried out at a frequency of 405 megacycles and the velocity measurements at 422 megacycles.

For No. 4 B&S wires spaced one inch apart the calculated value of the attenuation constant per 100 feet is 0.037. The measured average is 0.045. This corresponds to an efficiency of 91.4 per cent for a 100-foot line and 40.7 per cent for a 1000-foot line. The theoretical value of the attenuation constant for a No. 6 B&S wire is 0.042 per 100 feet and the measured value is 0.050. This corresponds to an efficiency of 90.5 per cent for a 100-foot line and 36.8 per cent for a 1000-foot line. It was found that the difference between losses on clean and weathered wires was too slight to be measured. These measurements indicate that it is practical to use balanced, closely spaced, two-wire transmission lines up to several hundred feet in length. For longer lines, the more expensive large diameter concentric conductor type of line must be chosen if good efficiency is to be maintained.

The presence of spacers, if made of good material and if used sparingly, usually add very little to the loss in dry weather. In wet weather, however, the loss and the reflections caused by such spacers are very serious handicaps. They are avoided entirely by using U-shaped half-wave conductor suspension loops, or "metallic insulators."

During rain the increase in loss from the presence of a water film on the wire was found negligible. The wave velocity along the line did, on the other hand, show a very marked decrease. Weathered and polished wires also showed a very marked difference in their ability to hold water. A horizontal line made of new No. 6 B&S wires with one-inch spacing was found to hold enough water to reduce the velocity by three per cent. The distribution of the water was very beady. The thickness of an evenly distributed water film required to cause this amount of variation in velocity was found by calculation to be 0.008 inch. The distribution of the water on a weathered wire was much less beady and was more evenly distributed. The velocity change for maximum water condition was 0.6 per cent, corresponding to a uniform film thickness of 0.002 inch. The maximum reduction in velocity for a weathered No. 4 B&S wire was 0.5 per cent.

From these values, the velocity change on a more open-wire system such as a V antenna can be calculated. For such a system the velocity change for weathered wires amounts to about 0.2 per cent or less. This is fortunate since the change will ordinarily not have sufficient influence upon the characteristics of a hundred-wave antenna to be serious.

Such variations and even greater ones can, however, be expected during sleet storms. In some commercial installations it will be necessary to provide for sleet melting. There are a number of ways in which this can be done with long wire antennas.

Propagation Tests

The chief purpose of the transmitter equipment which has been developed in accordance with the principles just outlined has been to study propagation in new frequency regions. Since the results of some of these tests are quantitatively described in detail in a paper by B. Trevor and R. George,[6] only some of the results obtained will be given in this paper.

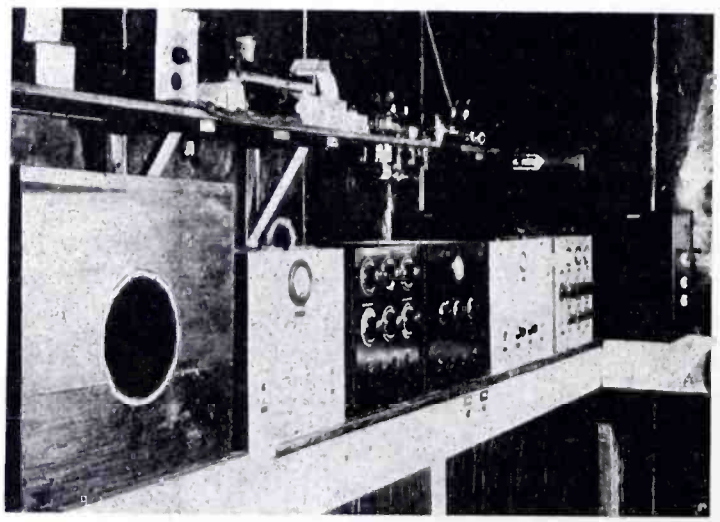


Fig. 21—Rocky Point receiving equipment for 200- to 500-megacycle signals.

After preliminary local tests, communication between Rocky Point and Riverhead, a distance of fourteen miles, was established. The 6-watt, 462-megacycle Barkhausen transmitter shown in Figures 11, 12, and 13 was used. This equipment was duplicated so that two-way telephony could be carried on. The Rocky Point receiving equipment is shown in Figures 21 and 22. Both transmitting and receiving antennas, which were of the billboard type shown in Figure 20, were mounted on supports fifty to seventy-five feet high in order to be in line of sight. As is usual when there is line of sight between transmitting and receiving antennas, no fading was observed. Rain, snow, and dense clouds of smoke from forest fires seemed to have no effect on the signals. The absence of fading on this circuit was a desirable

[6] B. Trevor and R. George, "Notes on Propagation at a Wavelength of Seventy-Three Centimeters," *Proc. I.R.E.*, Vol. 23, pp. 461-470; May, (1935).

feature since its purpose was to test the stability of transmitters and receivers. The equipment was used later for propagation studies under more difficult space-circuit conditions.

The Barkhausen transmitter, at Rocky Point, Figures 11, 12, and 13, was eventually replaced by a frequency multiplier unit, Figures 14, 15, and 16, capable of delivering up to 15 watts at 432 megacycles. The Rocky Point-Barkhausen and low power multiplier set-up is shown in

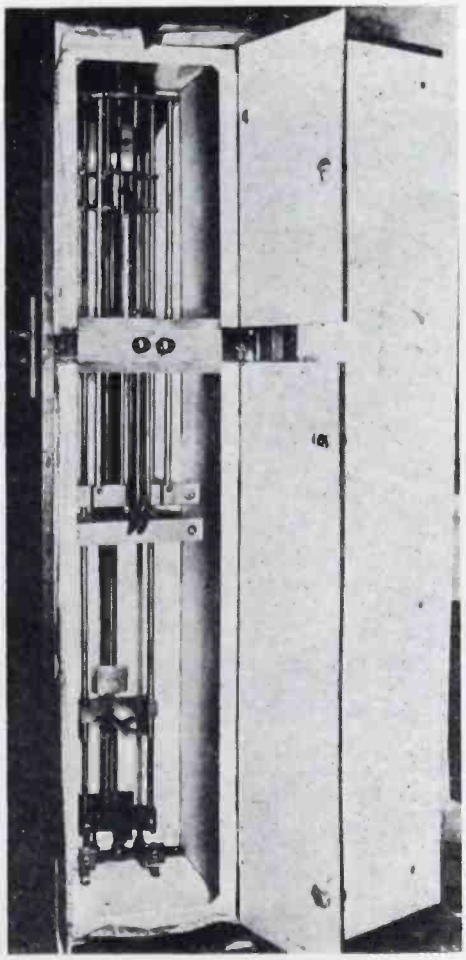


Fig. 22—Triode pendulum detector with trombone tuning circuits. Detail of unit in the background of Fig. 21.

Figure 23. Two horizontal V antennas were also erected. They were both seventy feet above the ground. One was 70 waves long and directed eastward on Riverhead. The other was 100 waves long and directed westward on the Empire State Building in New York City, fifty-six miles away. Signals from these antennas were observed with a portable receiver and a small directive antenna mounted on top of a car. Signals without fading were received at distances up to thirty

miles, in many localities considerably below the line of sight. By locating the receiving system on top of high buildings in New York and in airplanes it was found that signals without fading could be received at distances of from fifty to sixty miles when received as much as 500 feet below the line of sight. The distances below the line of sight given here represent actual values taken from contour maps of the terrain between transmitter and receiver. They will, therefore, check only approximately with simple geometric considerations based on the average curvature of the earth.

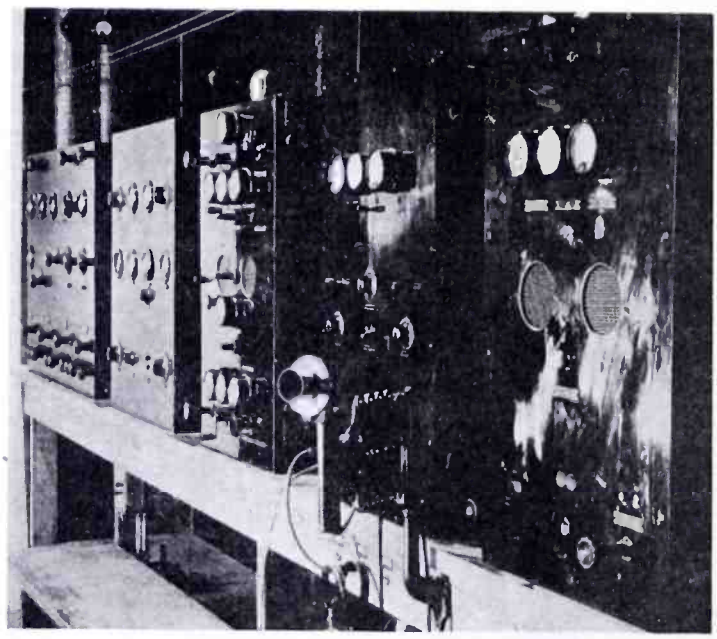


Fig. 23—Rocky Point low power ultra-high-frequency transmitter set-up. From left to right: 15-watt, 432-megacycle frequency multiplier; voltage regulator; 6-watt, 462-megacycle transmitter; compound modulator; and rectifier for modulator.

During these cruises with car and airplane, the unusual fact was observed that the apparent width of the beam was considerably smaller than the calculated width. For example, the beam from the 100-wave V antenna, which should have been about five miles wide over New York, seemed to be only two miles wide at the most.

After the 100-watt frequency multiplier transmitter became available it became possible to increase the range of observation. The frequency used was 411 megacycles. For the approximate distance of sixty miles, signals at ground level, 2000-3000 feet below the line of sight, would fade when completing their last lap over a very variable land and water route. When the receiving end of the route was over an uninterrupted body of water, there was very little fading at the

time of observation. At a point 113 miles away, with only a relatively small stretch of water midway, only violently fading signals could be observed, even during prolonged observation periods. These observations were made at an elevation of about 400-500 feet above sea level and 8000 feet below the line of sight. It was also determined, that at least for this particular space circuit, horizontally polarized waves were considerably superior to vertically polarized waves. The fact was also observed that the plane of polarization at the receiver was at all times identical with that at the transmitter.

With the transmitting antenna 70 feet above the surrounding country and 170 feet above sea level, badly fading signals were observed in an airplane, 165 miles away, flying at a height of 7500 feet. In a later flight, after the height of the transmitting antenna had been increased to 120 feet above the surrounding country and after the receiver and the ignition system of the plane had been very substantially improved, surprisingly little distance was added to that obtained in the previous flight. Badly fading signals could still be observed at a distance of 172 miles when flying at an altitude of 7500 feet and two miles below the line of sight. The routes of these signals were chiefly over land.

It appears from these tests that the distance of reliable as well as unreliable signals increases with the increase of power at the transmitter. It is unsafe to predict whether this proportionality will remain when still higher powers are applied. The distance beyond line of sight, for reliable signals, will, of course, also depend upon the nature of the route.

If improved facilities result in much greater distances, no doubt the electrical conditions of the portions of space more remote from the earth's surface will play an important part in determining the propagation characteristics. The indications are that, at the present, most of the phenomena observed are obtained in the region of space near the earth's surface. No doubt such phenomena as nonuniform humidity and the surface condition of water bodies have a great influence upon the characteristics of the propagation along the earth's surface. The shape and nature of the landscape, needless to say, has an enormous influence.

In Table I, a chart of propagation results, obtained at different frequencies by various investigators, is shown. In comparing these results the method referred to by Jouaust[7] has been used. A factor m is used by which the radius of curvature of the earth must be multiplied to give the curvature of the propagation path. Thus if m assumes its smallest value, that of unity, the curvature of the propagation path

[7] R. Jouaust, "Some Details Relative to Propagation of Very Short Waves," *Proc. I.R.E.*, Vol. 19, pp. 479-489; March, (1931).

TABLE I

EXAMPLES OF ULTRA SHORT-WAVE TRANSMISSION

Frequency Megacycles	Wavelength Centimeters	Distance		Elevation Meters		Value of m	Terminals of Test Circuit	Investigators
		km	miles	a	b			
60	500	205	128	700	530	3.87	Nice-Corsica	Jouaust, Ferrié Milan
40	750	456	284	1462	0	1.1	Hawaii-Kauai	Beverage, Hansell
40	750	145	90	518	0	1.446	Kauai-Oahu	Peterson, Fifield, Matthews
44 61	682 492	446	278	1916	387	1.32	Empire State Building—Mt. Washington	Jones
527	57	270	168	750	340	1.58	Rocca di Papa—Cape Figari	Marconi, Mathieu
527	57	85	53	750	0	3.0	Rocca di Papa—Yacht Elettra	" "
480	70	90	56	304	51	2	Rocky Point—Empire State Building	Lindenblad, Dow, George, Trevor
411	73	96.5	60	52	18	1.22	Rocky Point—Montauk Point	"
406	74	180	112	70	116	1.16	Rocky Point—Arney's Mount	"
406	74	105	65	70	65	1.43	Rocky Point—Atlantic Highlands	"
406	74	274	170	70	2287	1.88	Rocky Point—Airplane	"

will be equal to the curvature of the earth. According to Humphrey the value of m for light is 5.7. Others have given it values as high as 10. The results given in the table may be judged by the smallness of the factor m. It can be seen that for frequencies above 30 megacycles, the Hawaii-Kauai telephone circuit has a very highly curved propagation path. At frequencies above 300 megacycles, very high curving was obtained on the test circuit between Rocky Point and Arney's Mount.

ACKNOWLEDGMENT

The research and development outlined in this paper naturally represent the combined efforts of many individuals. Messrs. O. E. Dow and Bernard Salzberg have contributed much to the transmitter development. Messrs. R. W. George and Bertram Trevor have developed the receiving equipment and made the propagation observations. Mr. E. E. Spitzer and associates at the RCA Radiotron Company and Mr. P. S. Carter of R.C.A. Communications, Inc., have rendered very valuable assistance. The author wishes to express his thanks to Messrs. H. H. Beverage, C. W. Hansell, and H. O. Peterson for their many valuable contributions and for encouragement given.

TRANSMISSION OF 9-CM. ELECTRO-MAGNETIC WAVES

By

Irving Wolff and E. G. Linder

Research Division, RCA Manufacturing Company, Inc., Camden, N. J.

Summary—*A 9-cm beam transmitter and a receiver are described for making tests of the transmission of electromagnetic waves of this wavelength.*

Using this transmitter with the receiver placed at a series of locations from ½ up to a line of sight distance of 16 miles, measurements of field strength failed to show any noticeable attenuation due to the atmosphere. A series of more accurate tests comparing the transmission over a 2-mile distance during alternate clear and rainy periods indicated that the attenuation due to rainfall was less than .1 db per mile.

A NUMBER of tests have been carried out in different laboratories on the generation and reception of waves 10 cm. in length and shorter. However, for their practical application it is not only necessary to know that they can be generated and received, but we must also have data regarding the extent to which they are attenuated in the atmosphere, and the effect of weather conditions such as rain, snow and fog on their transmission. A series of tests to determine the attenuation of normal atmosphere were undertaken during the summer of 1934 at Atlantic Highlands, New Jersey, in cooperation with the U. S. Signal Corps, at Fort Monmouth and during the Spring of 1935, a system was placed in continuous operation between the laboratory at Camden and one of the tall buildings in Philadelphia, for the purpose of determining the effect of rain on the transmission.

The apparatus which was used in both of these tests was similar to that described in an article published in the Proceedings of the Institute of Radio Engineers in the January 1935 issue. A photograph of the transmitting apparatus is shown in Figure 1.

The transmitter consisted of a specially constructed magnetron connected to a half-wave antenna in the focus of a 4 ft. parabolic reflector, with appropriate voltage regulation of the supply circuits so that the output of the transmitter would remain reasonably constant without adjustment. Only a brief description of the tube and accompanying circuits will be given here. For more details, reference can be made to the article which appeared in the Proceedings of the Insti-

Reprinted from *Broadcast News*, December, 1935.

tute and an additional article which will appear shortly. A diagram of the tube is shown in Figure 2. This tube has for its basis the split anode magnetron which is shown diagrammatically in Figure 3. The split anode magnetron consists of two separated halves of a cylinder, whose axis is concentric with the filament, which are individually attached to the two halves of a two-wire balanced transmission line. The other end of the transmission line is terminated by a one-half wave antenna. This differs from the magnetron which is used at lower frequencies in having the anode in two parts, whereas, in the usual magnetron the anode is a continuous cylinder about the cathode and the oscillations are taken off between the cathode and the anode. The

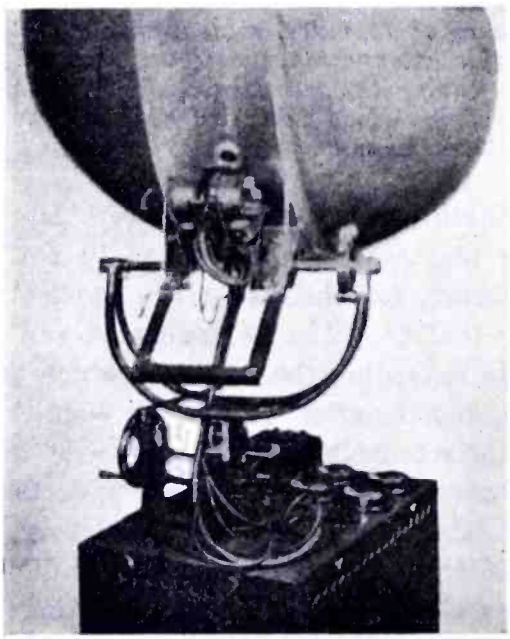


Fig. 1

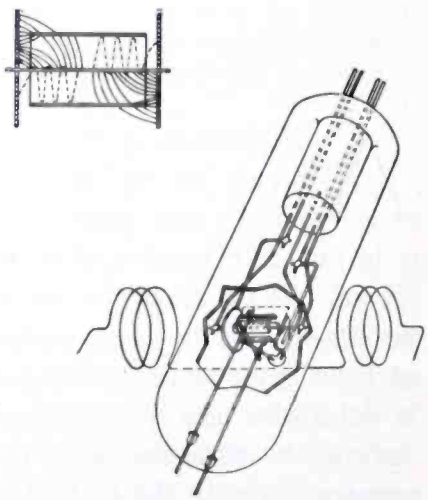


Fig. 2

high frequency split anode magnetron differs also in its mode of operation from the ordinary magnetron in the adjustment of the magnetic field.

When a magnetron is used at the lower frequencies the magnetic field should be just strong enough to prevent the electrons from reaching the anode to put the tube in an oscillating condition. As the potential applied to the anodes is increased a stronger magnetic field is required to do this. However, oscillations at a particular frequency can be obtained with a wide variety of electric and corresponding magnetic fields, the frequency being determined by the external circuit. As we attempt to continuously raise the frequency, we find that oscillations no longer take place for all adjustments of the plate potential and magnetic field in which the electrons just fail to reach the plate,

even though an external circuit is provided which could oscillate at the correct frequency. A further study shows the reason for this. The time taken for the electrons to go from the filament to the plate becomes an appreciable part of the cycle so that the phase relations, between current and voltage, which are required to deliver energy to the oscillating circuit no longer hold. Under such conditions it is necessary to time the arrival of the electron at or close to the anode in such a way that the correct phase between current and voltage will continue to be provided. This requires that the speed of travel of the electron across the tube be taken into account, a factor which is almost negligible in a consideration of oscillation, even in the so-called ultra short wave band between 3 and 10 meters. It has been found that the tube will oscillate at high frequency if the time it takes the electrons

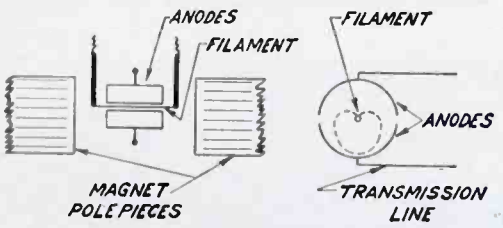


Fig. 3

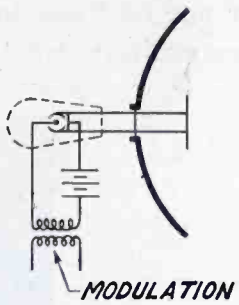


Fig. 4

to make a complete circuit in the tube from cathode in the direction of the anode and back to cathode again, is equal to the period of the wave it is desired to produce. This adjustment is quite critical. The oscillations which depend on adjusting the electron speed to the period of the wave have been called "electronic oscillations." The time that it takes the electron to go from cathode to anode has been called the "transit time."

The tube which was used for the transmission tests was an improved type of split anode magnetron in which an electric field in the direction of the magnetic field was added in addition to the transverse electric field between cathode and anode. This electric field was obtained by placing two metallic discs at the open ends of the cylindrical anodes and supporting them so that they are insulated from the cathode. They were both operated at a potential approximately three-quarters that of the anode potential. Their function was to draw electrons from the inside of the cylinder. A theoretical and experimental consideration had shown that a tilting of the tube without end plates was required to allow electrons to spiral out from the region inside the anodes, so as to obtain proper space charge conditions for

maximum oscillation. In the case of the ordinary split anode magnetron, the relative adjustment of angle of tilt and plate voltage is very critical and if either one is changed the other must be changed to some new value, in order to continue oscillation. In a tube using the end plates, stability of operation is much better. This is due to the fact that the anode and end plate potentials may be taken from the same voltage supply and therefore will vary proportionately. It has been found that when the anode potential is varied the end plate potential, which is required in order to maintain oscillation at the maximum value, is very nearly that which keeps this proportionality constant.

Since the output of the tube depends on the ratio between end plate and anode potential, modulation can be easily obtained by varying either one of these independently of the other. In order to modulate the transmitter, we therefore place the secondary of the modulation

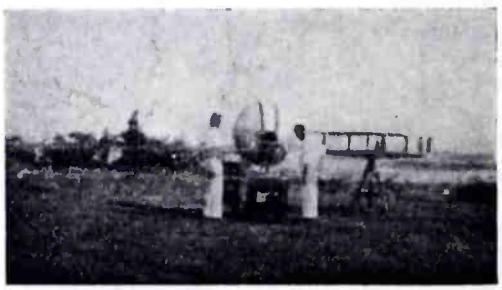


Fig. 5

transformer in series with the anode supply, and adjust the anode potential so that oscillation amplitude for no modulation is approximately one-half the maximum. This allows the output of the tube to swing from close to zero to maximum for full modulation. A diagram showing the modulation system, the antenna, and the reflector used with the transmitter is shown in Figure 4. It will be noted on this diagram that the anode supply is brought in across a line which appears to be shorted directly across the transmission line. This bar has two purposes. In the first place, it acts as a short circuit for long wave parasitic oscillations which the tube would like to generate, if it were not prevented from doing so. In the second place, its position is adjusted so that the capacity between the halves of the anodes and the inductance of the small loop circuit is correct to tune to the frequency of oscillation.

The receiver consisted of an iron pyrites crystal attached to a small loop, which was placed in the focus of a 4-foot reflector and does not require further explanation. The output of the crystal was attached

to an audio amplifier which in turn was fed into a tube detector and microammeter for the first tests, and into an Esterline recording meter for the tests which were made to determine the transmission through rain.

In the tests that were made at Atlantic Highlands, the transmitter was placed close to Navesink Light House which is on a hill about 200 feet above sea level.

The transmitter as set up for the tests is shown in Figure 5. The receiver, mounted on the stern of the test ship is shown in Figure 6.

The apparatus to the right of the transmitter in Figure 5 has no connection with these tests. These pictures are supplied through the courtesy of the Signal Corps Laboratories at Fort Monmouth. The

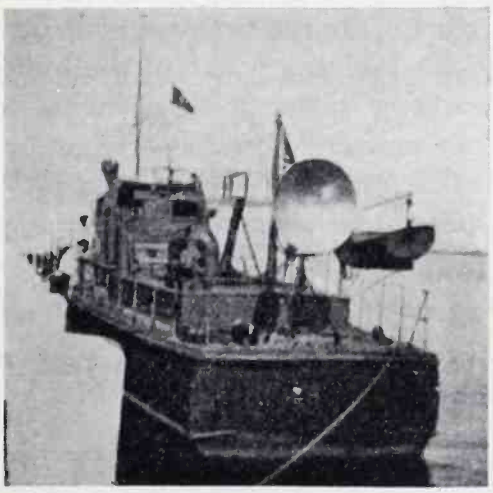


Fig. 6

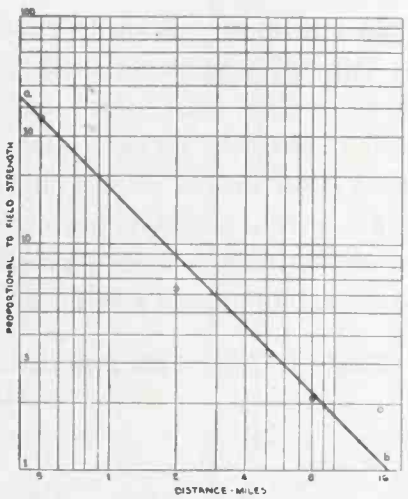


Fig. 7

line of sight range to the horizon from this point was approximately 17½ miles. A number of readings were taken at different distances from the transmitter. The transmission distances were from one-half to five and one-quarter miles over land, and eight to sixteen miles over water. The signal which was received at the 16 mile distance was sufficiently strong so that it would have been possible to move the receiver out to 40 miles before the signal intensity became equal to the amplifier noise. Some previous measurements had shown that the output of the crystal detector was proportional to the square of the signal strength. This fact, along with the amplifier gain control setting, made it easy to calculate the relative field strength at each one of the points where the receiver was set up. In Figure 7 the relative field strengths are plotted against the distance on log-log coordinate system. If attenuation in the atmosphere is negligible the signal strength should decrease inversely as the distance from the trans-

mitter, since the radiation is in the form of a cone starting at a relatively short distance from the transmitter. The points should lie along a straight line of slope —1. The line ab has been drawn through the half-mile point with this slope. There is no indication that there is any attenuation other than that due to the spreading of the energy. There was considerable uncertainty in the readings since the crystal had to be readjusted at each point, and the transmitter output may have shifted somewhat. Nevertheless, we can safely say, that attenuation in the clear atmosphere for 3000 megacycle electromagnetic waves is negligible up to distances of 16 miles, and probably more.

The measurements which were made at Atlantic Highlands were all conducted when the weather was clear, although at the time of the 16-mile test there was sufficient haze so that the transmitting point could not be seen from the receiving point. In the next series of tests, an attempt was made to determine whether the water in the atmosphere during heavy rain or fog would be sufficient to attenuate the 3000 megacycle signal. Although the amount of water present in a heavy rain storm over a distance of 20 miles, if concentrated into a single sheet, would definitely be sufficient to affect the transmission, we should expect a smaller effect to be caused by the rain because of the relatively small size of the rain drops compared to the wavelength.

TABLE I. DROP SIZE AND SPACING FOR DIFFERENT TYPES OF RAINFALL, MIST AND FOG.

(1)	(2) Precipitation mm/hr.	(3) Diam. of drop mm.	(4) Mass of drop grams	(5) Velocity of fall m per sec.	(6) Water mg. per m^3 air	(7) No. of drops per m^3	(8) Drop spacing (cm)	(9) Frac. of water by vol. in air	(10) Thickness of water layer in cm. in 2 mi. path
Fog	..	.01	$5.2 \cdot 10^{-10}$	.003	6	$1.2 \cdot 10^{7}$	.43	$6 \cdot 10^{-9}$	$1.9 \cdot 10^{-3}$
Mist	.05	.1	$5.2 \cdot 10^{-7}$	.25	55.5	$1.1 \cdot 10^{5}$	2.1	$5.5 \cdot 10^{-8}$	$1.8 \cdot 10^{-2}$
Drizzle	.25	.2	$4.2 \cdot 10^{-6}$	.75	92.6	$2.2 \cdot 10^{4}$	3.6	$9.3 \cdot 10^{-8}$	$3.0 \cdot 10^{-2}$
Light Rain	1.00	.45	$4.8 \cdot 10^{-5}$	2.00	138	$2.9 \cdot 10^{3}$	7.0	$1.4 \cdot 10^{-7}$	$4.4 \cdot 10^{-2}$
Moderate Rain	4.00	1.0	$5.2 \cdot 10^{-4}$	4.00	277	$5.3 \cdot 10^{2}$	12.3	$2.8 \cdot 10^{-7}$	$8.9 \cdot 10^{-2}$
Heavy Rain	15.00	1.5	$1.8 \cdot 10^{-3}$	5.00	833	$4.6 \cdot 10^{2}$	13.0	$8.3 \cdot 10^{-7}$	$2.7 \cdot 10^{-1}$
Excessive Rain	40.00	2.1	$4.9 \cdot 10^{-3}$	6.00	1850	$3.8 \cdot 10^{2}$	13.8	$1.8 \cdot 10^{-6}$	$5.9 \cdot 10^{-1}$
Cloudburst	100.00	3.0	$1.4 \cdot 10^{-2}$	7.00	5400	$3.9 \cdot 10^{2}$	13.7	$5.4 \cdot 10^{-6}$	1.7
Snow (Heavy) estimated melted	4.00			.25 to 1.00	1100 to 4500				

In Table 1, various data on fog and rain are given. Columns 1, 2, 3, 5, and 6 were taken from Humphreys' "Physics of the Air." Columns 4, 7, 8, 9 and 10 have been computed from Humphreys' data. The data on snow are our own estimates.

Since the effect of water in the atmosphere may be considered as an attenuation phenomenon, and therefore exponential with respect to distance, the measurements had to be made over as large a distance as

possible, in order to obtain sufficient accuracy. In view of the fact, that any surfaces when wet might have reflected the waves differently than when dry, it was also desirable that not only the test location be line-of-sight from the transmitter, but that there be a minimum of buildings or other objects on the side of the beam, particularly near the line joining the receiver and transmitter. This meant approximately that there should be as few objects as possible in the circular cone of vertex angle 10 degrees, whose vertex was on the transmitter, and whose axis was the line joining the receiver and transmitter.

An inspection of the available locations in the neighborhood of

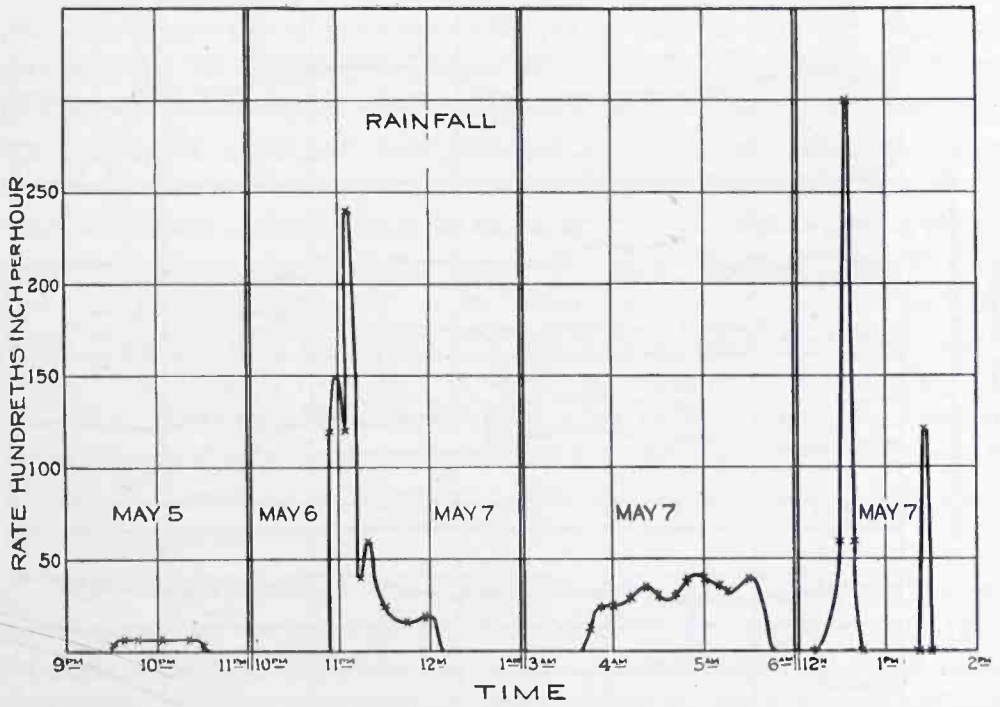


Fig. 8

Camden showed that these conditions could be most effectively and conveniently met by installing the transmitter on one of the upper floors of the engineering building, and the receiver in one of the tall buildings in Philadelphia, as far from the transmitter as possible. A location on the 34th floor of 12 S. 12th St., Philadelphia, was found suitable for the receiver. This was two miles from the transmitter.

The transmitter-receiver system was first maintained in operation during several clear days, so that an idea of the constancy of output to be expected could be measured. Having obtained these data, we hoped that several rainy periods would occur and permit a determination of the difference in transmission, after which an additional calibration could be made in clear weather to determine whether any change had taken place in the equipment.

On May 6 and 7, 1935, conditions were very favorable for determining the effect of rainfall on the transmission. Some very heavy short showers took place, followed by periods of no rainfall. The chart of rainfall for these days taken from the tipping bucket chart of the U. S. Weather Bureau at Philadelphia through the courtesy of Mr. Bliss is shown in Figure 8. The place where the rainfall was recorded was very nearly on the straight line joining the transmitter and receiver points, and about one-half way between them. It was also fortunate that the transmitter and receiver had been in continuous operation for several days previous to these dates.

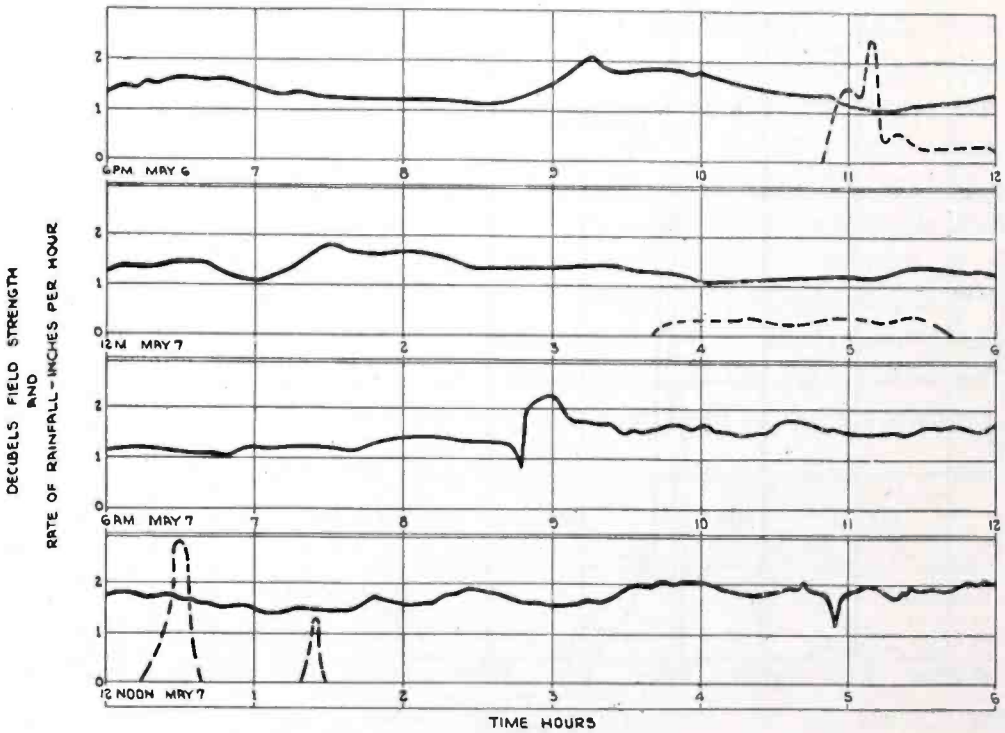


Fig. 9

The meter record from 6 P. M., May 6 to 6 P. M., May 7, is shown on Figure 9. The heavy line represents the field strength, the dotted line the rainfall. The maximum fluctuation in output was 1 db. This, however, was not connected with rainfall and also was recorded on clear days. The sharp rise at 8:45 A. M. was caused by a change which was made in the transmitter magnetic field. The first rainfall took place at about 11 P. M., reaching a maximum intensity of 2.4 inches an hour, which is what Humphreys calls an excessive rain. There were other showers from 3:45 to 5:30 A. M., and shortly after noon. The one at 12:30 reached almost cloudburst intensity. With the exception of the early morning shower the short duration of the rainfall makes it improbable that the whole region between the transmitter and re-

ceiver was filled with rainfall of maximum intensity at any one time. Comparison of the curves for field intensity and rainfall show that the maximum amplitude attenuation, if any, is less than .1 db. per mile.

On the basis of the data which were obtained, and Table I we are able to decide whether the lighter forms of rain and fog, mist, or snow should affect the transmission. To do this we will consider the amount and distribution of water in the air separately. Table I shows that the mass of water per unit volume was greater for the rains which were tested than for any other form of precipitation and therefore on this basis alone should have caused the greatest effect. As the rain becomes lighter the mass of water decreases and the drop size becomes smaller, but the spacing decreases also. Conditions approach closer to that of water vapor. It is, therefore, interesting to compare the mass of water vapor per cubic meter in saturated air with the water in the air due to the rainfall. (It seems fair to assume that the air is near saturation during all forms of rainfall.) The Smithsonian tables give the water vapor as 22 gms. per cu. meter at 20 degrees C. and 760 mm. mercury pressure. It is rather surprising that this is four times the amount of water due to a heavy rain, and almost 4000 times that of the droplets in a fog. It therefore appears as if the effect of the water vapor should be greater than that of the free water when it is very finely divided and closely spaced and that, therefore, by making the measurement in a heavy rain, the most severe conditions have been encountered.

The comparatively large amount of water in a saturated atmosphere points to the possibility that more effect on transmission might be expected under conditions where the path traverses regions of variable vapor content, than during times of precipitation when humidity is relatively constant. A calculation of the relative index of refraction of moist and dry air, assuming that there is no anomalous change in the dielectric constant at 3000 megacycles, compared to lower frequencies, indicates that it should be of the order of 1.0001. The figure of 1.0001 requires either many transitions, or interfaces at very glancing angles to cause noticeable variations in transmission.

In conclusion we may say that the tests, both at Sandy Hook and at Camden, have shown for 3000 megacycle electromagnetic waves: (1) that the attenuation in a clear atmosphere is negligible up to a distance as great as 16 miles, and (2) that the attenuation of the field caused by heavy rain, is less than .1 db. per mile; (3) that an investigation of the water content and drop size of fog, mist and light rain show that their effect on the transmission should be less than that of heavy rain.

Papers Published in Summary Form

Part I—Ultra-High Frequencies Below 300 Mc.

FREQUENCY ASSIGNMENTS FOR TELEVISION

By

E. W. Engstrom and C. M. Burrill
RCA Manufacturing Company, Camden, N. J.

Summary

This article is not a report of original work, but is a correlation or synthesis of information pertinent to the subject, available to the authors within the RCA Services or through published papers. Since the results of all have been taken into account it has not seemed feasible or desirable to give credit to individual sources except to mention the article by H. H. Beverage entitled "Some Notes on Ultra Short Wave Propagation" appearing in this number of RCA Review, and the bibliography forming a part of that article. Much credit is due collectively to the many workers in this field, who have made possible the drawing with reasonable certainty of the conclusions here stated. The basic plan of any new service must always be determined by the work of such pioneers, before commercial experience has made everything plain. Because fundamental plans for broadcast television are now in the making, it is hoped that this brief article will be found both timely and interesting.

Published in *RCA Review*, January, 1937.

TELEVISION TRANSMITTERS OPERATING AT HIGH POWERS AND ULTRA-HIGH FREQUENCIES

By

J. W. Conklin and H. E. Gihring
RCA Manufacturing Company, Inc., Camden, N. J.

Summary

A general review of some of the unusual problems encountered in the development of high-power ultra-high-frequency transmitters and particularly television power amplifiers. Tube and circuit problems are discussed and also some of the difficulties encountered with associated components.

Published in *RCA Review*, July, 1937.

THE REQUIREMENTS AND PERFORMANCE OF A NEW ULTRA-HIGH-FREQUENCY POWER TUBE

By
W. G. Wagener
Formerly with
Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

Large values of power are difficult to obtain in the ultra-high-frequency region. At the limiting frequencies it is increasingly more difficult to find vacuum tubes that will deliver such power and perform efficiently. The principal factors that affect the design and performance of the tubes are those involving the electrical circuit, the size requirements for the power desired, and the transit time of the electrons within the evacuated space of the tube.

The design principles that result from a consideration of these factors have been used in the development of a new ultra-high-frequency triode. A triode capable of delivering approximately 700 watts at 100 megacycles is described. This tube, which is cooled by water and air, is capable of operation as a neutralized power amplifier up to 200 megacycles with an output of approximately 500 watts.

Published in *RCA Review*, October, 1937.

EFFECT OF ELECTRON TRANSIT TIME ON EFFICIENCY OF A POWER AMPLIFIER

By
Andrew V. Haeff
Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

Measurements of the plate efficiency of a neutralized triode amplifier operated at high frequencies are reported. The results are compared with those obtained for an oscillator operated at the same frequency. At a frequency at which the oscillator efficiency approaches zero the amplifier efficiency is found to be reduced to only 50 per cent of the low-frequency value. It is shown that the difference in efficiency is primarily due to a large phase angle between the plate current and the grid voltage produced by the electron transit time.

Published in *RCA Review*, July, 1939.

ON THE OPTIMUM LENGTH FOR TRANSMISSION LINES USED AS CIRCUIT ELEMENTS

By
Bernard Salzberg
Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

The existence of an optimum length for transmission lines which are tuned by low-loss capacitor to give maximum sending-end impedance is discussed. This optimum length is found to be 0.185λ for a shorted line and 0.472λ for an open-circuited line, resulting in impedances 14 and 3 per cent higher, respectively, than can be obtained from lines without tuning condensers.

Published in *Proc. I.R.E.*, December, 1937.

A STUDY OF THE PROPAGATION OF WAVELENGTHS BETWEEN THREE AND EIGHT METERS

By

L. F. Jones

RCA Manufacturing Company, Inc., Camden, New Jersey

Summary

A description is given of the equipments used in an airplane, dirigible, automobile, and indoors to measure the propagation characteristics of wavelengths between about three and eight meters. The majority of observations were of television transmissions from the Empire State building.

The absorption of ultra-short-waves traveling through or around large buildings is shown to be in terms of amplitude about 50 per cent every 500 feet for seven meters and 50 per cent every 200 feet for three meters. A number of reflection phenomena are discussed and the influence of interference patterns on receiving conditions is emphasized. It is shown that any modulation frequency is partly or completely suppressed if propagation to the receiver takes place over two paths differing in length by half of the hypothetical radio wavelength of the modulation frequency. For a good television picture this corresponds to a difference of about 500 feet.

Various types of interference are mentioned. There are maps of the interference patterns measured in a typical residential room. The manner in which traffic movements cause severe fluctuations in ultra-short-wave field strengths at certain indoor points is shown by recorded field strengths.

It is shown that the service range of the Empire State transmitters includes most of the urban and suburban areas of New York, and that the interference range is approximately 100 miles. Variations of field strength with altitude, beyond line of sight, are shown. Observations made at a distance of 280 miles are described.

An empirical ultra-short-wave propagation formula is proposed. Curves are then calculated showing the relations between wavelength, power, range, attenuation, and antenna height.

Published in *Proc. I.R.E.*, March, 1933, and *Television*, Vol. I, 1936.

AN URBAN FIELD-STRENGTH SURVEY AT THIRTY AND ONE HUNDRED MEGACYCLES

By

R. S. Holmes and A. H. Turner

RCA Manufacturing Company, Inc., Camden, New Jersey

Summary

A description is given of the transmitter and receiver equipment used in making field strength surveys in the Camden-Philadelphia area for a low power transmitter whose antenna is 200 feet above the ground, at frequencies of thirty and one hundred megacycles.

Field strength contour maps for the area within approximately ten miles of the transmitter are given. From these maps the average field strength obtained at various distances from the transmitter was determined, and the attenuation of the signal was found to be proportional to the 1.84 power of the distance for thirty megacycles and the 2.5 power of the distance for one hundred megacycles for the region between one and ten miles from the transmitter.

Curves showing the variation from the average field strength of the signal along three routes radiating fifteen miles from the transmitter are given, and these variations are compared with the elevation profiles of the respective routes. It is shown that the signal is usually strongest on the brows of hills facing the transmitter.

Measurements were made in three representative residences, and from these data, curves showing the power required at the transmitter to furnish one hundred microvolts input to receivers with short indoor antennas located in houses at various distances up to ten miles from the transmitter were computed for the two frequencies.

Published in *Proc. I.R.E.*, May, 1936, and *Television*, Vol. I, 1936.

OBSERVATIONS ON SKY-WAVE TRANSMISSION ON FREQUENCIES ABOVE 40 MEGACYCLES

By

D. R. Goddard

R.C.A. Communications, Inc., Riverhead, L. I., New York

Summary

The results of daily observations at Riverhead, L. I., N. Y., since September, 1937, of European 40-to-45 megacycle transmitters are reported. Measurements of field strength were made on English, French, and German television signals. Multipath propagation of the English video-frequency channel was observed optically and the difference in path length determined.

Published in *Proc. I.R.E.*, January, 1939.

NOTES ON THE RANDOM FADING OF 50-MEGACYCLE SIGNALS OVER NON-OPTICAL PATHS

By

K. G. MACLEAN AND G. S. WICKIZER
R.C.A. Communications, Inc., Riverhead, L. I., N. Y.

SUMMARY

To obtain data on the variation of field strength beyond the horizon, simultaneous recordings were made at three locations, one within the optical path of the transmitter, one 700 feet below the line of sight, and one 11,400 feet below the line of sight. All three locations were on the same line from the transmitter. Recordings extended over a two-week period, chosen at a time when atmospheric refraction was likely to be favorable. Analysis of the recorded data indicates several things of interest. The variation of field strength at each location was random and showed no correlation with any other location; the range of field-strength variation exceeded 49 decibels at the most remote location; maximum fields generally occurred at night; and previous data on the rate of attenuation beyond the horizon were confirmed.

Published in *Proc. I.R.E.*, August, 1939.

APPLICATION OF ABELIAN FINITE GROUP THEORY TO ELECTROMAGNETIC REFRACTION

By

R. A. WHITEMAN
Instructor of Physics, R.C.A. Institutes, Inc., New York

SUMMARY

This paper presents the direct relationship between the group theory and indices of refraction encountered in the solution of optical problems. The properties and advantages of the group theory are made available by using the double-subscript notation when designating the indices of refraction. The notation is developed from fundamental concepts of refraction at a plane surface and applied to two specific refraction problems in the following order:

a. refraction in the Kennelly-Heavside layer (radio waves)
b. electronic refraction (electron optics)

The criterion to be used in determining the applicability of the method developed in this paper to any refraction problem is reduced to a simple equation.

Published in *RCA Review*, January, 1940.

THE USE OF GAS-FILLED LAMPS AS HIGH-DISSIPATION, HIGH-FREQUENCY RESISTORS, ESPECIALLY FOR POWER MEASUREMENTS

By

Ernest G. Linder

RCA Manufacturing Company, Inc., Camden, N. J.

Summary

A type of hydrogen-filled lamp suitable for use as a high-frequency resistance is described, which possesses unusually great heat-dissipation ability. This dissipation may be several hundred times that obtainable with vacuum lamps, and the gain is greatest for filaments of the smallest diameter. Other advantages are pointed out. The theory of heat loss in a gaseous atmosphere is summarized. Details of experimental lamps are given. Design data are presented in the form of a chart which includes, watts dissipated, resistance, temperature, and filament diameter.

Published in *RCA Review*, July, 1939.

MEASUREMENTS OF ADMITTANCES AT ULTRA-HIGH FREQUENCIES

By

John M. Miller* and Bernard Salzberg

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

A substitution method is described for determining admittances at ultra-high frequencies. Typical measurements of resistors and insulators are given.

The method employs a short transmission line excited by an ultra-high-frequency oscillator. The receiving end of the line is short-circuited and the sending end is shunted by a variable capacitor and the unknown admittance. A vacuum-tube voltmeter indicates the resonant voltage at a point on the line. The unknown admittance is removed, resonance restored by the variable capacitor, and the same voltage obtained by sliding a known resistor along the line. The unknown admittance is then determined by the frequency, capacitance change, and the position of the known resistor. With proper line constants, the equivalent sending-end resistance of the resistor is equal to the product of its resistance and the square of the ratio of the total line length to the distance of the resistor from the receiving end.

* Formerly with RCA Manufacturing Company.

Published in *RCA Review*, April, 1939.

ELECTRICAL MEASUREMENTS AT WAVE LENGTHS LESS THAN TWO METERS

By

LEON S. NERGAARD

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

SUMMARY

In this paper the measurement of power and voltage at ultra-short waves is considered. A signal generator delivering adequate power output with satisfactory stability over the wave-length range from twenty to 200 centimeters is described. The requirements of thermocouples satisfactory for the measurement of power are considered and a set of thermocouples covering the power range from 0.1 milliwatt to fifty watts is described. A study of vacuum tube voltmeters has shown that diode voltmeters have very small loading (of the order of 10^5 ohms) on the circuits to which they are connected, whereas conventional triode voltmeters using RCA-955's have an input impedance of about 10^4 ohms at a wave length of one meter. The errors of diode voltmeters at ultra-short wave lengths have been studied and found to be of two kinds:

1. An error due to partial series resonance between the lead inductances and the interelectrode capacity.

2. An error due to a transit time effect. This has been called "premature cutoff." These errors have been studied experimentally and theoretically, the results being in qualitative agreement. In the course of the experimental work very small diodes have been built, the smallest diode having an anode diameter of only 0.0065 inch. Calibrations for these diode voltmeters have been obtained. The calibration of the diode having an anode diameter of nine mils at a wave length of 100 centimeters differs from the sixty-cycle calibration by less than one per cent due to premature cutoff and by about six per cent due to resonance error. Voltmeters of this type have been supplied to other workers and have been found to give very useful and consistent results.

The RCA-955 acorn tube with grid and plate tied together has been studied as a diode voltmeter. From this study it is concluded that the RCA-955 is useful as a voltmeter down to the sixty-centimeter wave length, but that small diodes are necessary for precise measurements below 150 centimeters and for general measurements below sixty centimeters.

Published in *Proc. I.R.E.*, September, 1936.

RADIO-FREQUENCY GENERATOR FOR TELEVISION RECEIVER TESTING

By

A. H. Turner

RCA Manufacturing Company, Inc., Camden, N. J.

Summary

An automatic recording signal generator is described which has proven extremely useful in the development of television receivers. The more important specifications and characteristics of this signal generator are as follows:

1. Consists of push-pull r-f oscillator and push-pull modulated r-f amplifier.

2. Frequency range (motor-driven linear control) 40 to 100 megacycles.

3. R-F output (motor-driven continuously variable control and uniform logarithmic scale) 1 to 100,000 microvolts.

4. Modulation (from external source or 400 cycles internal) 30 to 3,000,000 cycles for both side-bands or may be altered for 30 to 5,000,000 cycles for single side-band.

5. R-F output impedance—50 ohms approximately.

6. Slow, normal, and fast curve-paper frequency scales selected by gear shift.

7. Receiver output may be set and held constant at any level between 0.5 volt and 25 volts by the automatic mechanism.

Published in *RMA Engineer*, May, 1939.

INPUT RESISTANCE OF VACUUM TUBES AS ULTRA-HIGH-FREQUENCY AMPLIFIERS

By

W. R. Ferris

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

Vacuum tubes which when operated as voltage amplifiers at low frequency require no measurable grid input power have been found to take very serious amounts of power at ultra-high frequencies. The grid input conductance is shown to be very accurately represented for electrodes of any shape by the expression

$$g_g = Ks_m f^2 \tau^2$$

where g_g is the input conductance, s_m the grid-plate transconductance, f the frequency, and τ the electron transit time. K is a parameter which is a function of the geometry of the tube and the voltage distribution. A physical picture of the effect, a simple theoretical derivation, and experimental proof with conventional tubes are given.

The magnitude of g_g is such that it is the principal limitation for amplifiers at frequencies of the order of 100 megacycles, and it seriously affects the amplification at frequencies as low as fifteen megacycles. The input resistance of a typical commercial tube, the RCA-57, is approximately 20,000 ohms at thirty megacycles. Other commercial tubes, being of the same general construction and size, have input resistances of the same order. The use of very small tubes, such as the RCA-954, with correspondingly short transit times is shown to be a practical means of increasing the amplification obtainable with conventional circuits.

Input capacity variation with frequency is found to be negligible with the RCA-57 even up to eighty megacycles and higher. However, the grid-cathode capacity is a function of the applied voltage; the ratio of this capacity under operating conditions to that with the tube cold having a value of four thirds for its minimum. The change in input capacity from cold to hot is of the order of one micromicrofarad for the RCA-57. No change in grid-screen capacity is indicated.

The plate resistance of screen-grid tubes is found to vary with frequency but with the RCA-57 at eighty megacycles it is over twenty times the grid resistance and thus constitutes a negligible amount of the total loss in the circuit.

Published in *Proc. I.R.E.*, January, 1936.

ANALYSIS OF THE EFFECTS OF SPACE CHARGE ON GRID IMPEDANCE

By

D. O. North

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

Previous theory of transit-time phenomena in high vacuum diodes is extended and augmented to provide an explanation of the high-frequency behavior of high-μ amplifiers with parallel plane electrodes. For mathematical reasons the analysis is restricted to triodes with plate at alternating-current ground and to tetrodes with screen grid at alternating-current ground. Expressions for internal input loading and capacity are derived, showing the dependence upon frequency, voltages, and tube dimensions, and it is shown how the theory in its present form can be made quantitatively applicable to many commercial tubes of cylindrical design.

In agreement with both elementary theory and observation, the theory shows that at the threshold of the effect the input loading varies as the square of the frequency. For the RCA-57 there is calculated an internal input resistance of 21 megohms at one megacycle, dropping to 2100 ohms at 100 megacycles. These figures are in excellent agreement with actual measurement, and illustrate the tremendous importance of transit times in the design of tubes for ultra-high frequencies. It is likely that internal input power losses of this character, together with closely allied losses in transconductance, are primarily responsible for high-frequency failure of both amplifiers and oscillators.

"Hot" input capacity exceeds the "cold" value. The magnitude and dependence upon tube parameters is given, the increase is primarily

due to space charge but also depends upon τ_2/τ_1, the ratio of the electron transit time between control grid and plate to the transit time between cathode and control grid. In agreement with observation, the theory indicates very slight frequency dependence.

There is included a brief account of temperature-limited diodes, illustrating their possibilities as a source of high-frequency negative resistance.

Published in *Proc. I.R.E.*, January, 1936.

MULTI-TUBE OSCILLATORS FOR THE ULTRA-HIGH FREQUENCIES

By

PAUL D. ZOTTU

Formerly with
Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

SUMMARY

An important limitation on the effective use of the ultra-high frequencies is the fact that with present methods the output power that can be developed is small. With the feedback type of oscillator, output decreases approximately inversely as the square of the frequency. Although the use of two tubes in push-pull permits doubling the output of a single-tube oscillator, in the region of 300 Mc this increase amounts to only a few watts with commercially-available tubes. Paralleling of tubes is out of the question because this method causes tube capacities to add and therefore requires reduction of inductance to maintain the same frequency. At the frequencies considered, inductance is already at a premium and further reduction necessitates making the oscillatory circuit inside the tube. Another disadvantage of the direct-parallel system is that the generated frequency changes with the addition or subtraction of a unit. It is evident, therefore, that a means of combining the output of two or more independent oscillators in such a way that these disadvantages are overcome would be highly desirable.

By proper utilization of two well-known effects in the operation of oscillators, a method for combining several oscillators in the desired manner becomes possible. Consider first a simple oscillator coupled to a tuned circuit which will be termed the secondary circuit. If the coupling is very loose, the secondary tuning will have negligible effect on the oscillator frequency, except in a small region near the point where the two circuits are in resonance, where a slight change in oscillator frequency will occur. With closer coupling, a considerably greater change in frequency results; as the coupling is still further increased, a "jump" in frequency as the secondary is tuned through resonance will occur. The behavior is modified somewhat by the presence of a load on the secondary, although the general effect is the same. The interesting point is that, depending upon the coupling and load, the oscillator frequency is seriously affected by the tuning of the secondary circuit.

The power output to the secondary circuit depends upon the coupling and tuning; with small values of coupling the output rises as the secondary is tuned into resonance, decreases on tuning away. With coupling values greater than the critical, two points of maximum power output will appear, with a minimum point occurring at resonance. The important thing is that maximum output of the oscillator can be obtained with coupling values equal to or greater than critical.

Published in *QST*, October, 1936.

RECENT DEVELOPMENTS IN MINIATURE TUBES

By

Bernard Salzberg and D. G. Burnside

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

The development of two indirectly heated miniature tubes, a triode and a sharp cut-off amplifier pentode especially suited for use at high frequencies, is described. The electrical and mechanical factors involved in the design and application of these tubes are discussed and their novel structural appearances is described.

Because of their decreased lead impedances, interelectrode capacitances, and transit times, these miniature tubes allow considerable improvement to be made in high-frequency receiving equipment. It is possible to operate the triode as an oscillator in a conventional circuit down to a wavelength of approximately 40 centimeters. The pentode can be operated as a radio-frequency amplifier down to a wavelength of approximately 70 centimeters. It is practicable to obtain stable gains with it of from ten to fifteen at three meters, a wavelength at which standard tubes are almost entirely ineffectual. Both tubes can be used, down to much lower wavelengths, in exactly the same manner and for the same applications that the corresponding conventional tubes are used; i.e., as oscillators, amplifiers, detectors, converters, and as negative-resistance devices.

The small size of the tubes and their novel structural design allow compact and convenient receiving equipment to be built. Even at the longer wavelengths, they are applicable to a large number of uses for which their excellent characteristics, small size, and low weight make them particularly useful.

Published in *Proc. I.R.E.*, October, 1935.

EFFECTS OF SPACE CHARGE IN THE GRID-ANODE REGION OF VACUUM TUBES

By

BERNARD SALZBERG AND A. V. HAEFF

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

SUMMARY

The effects of space charge in the region between grid and anode of a vacuum tube, for the case where the planes of the grid and plate are parallel, are determined from the results of a simple analysis. The main effects of the space charge are (a) to introduce departures from the linear potential distribution of the electrostatic case; (b) to set an upper limit, under certain conditions, for the anode current; (c) to introduce instabilities and hysteresis phenomena in the behavior of the tube; and (d) to increase the electron transit time in this region.

Four modes of potential distribution which may exist in this region are treated: (1) Neither potential minimum nor virtual cathode exist; (2) potential minimum exists; (3) space-charge-limited virtual cathode exists; and (4) temperature-limited virtual cathode exists (negative anode potentials). For each of the various states of operation, expressions are derived for the distribution of potential and electric intensity throughout the region; the time of flight of electrons from grid to anode, and from grid to the plane of zero potential; and the location and magnitude of the minimum potential. An expression is also derived for the dependence of the anode current on the space current, grid-anode distance, grid voltage, and anode voltage. Curves are plotted from these expressions, and it is shown how the behavior of a large variety of practical tubes can be predicted and explained with their aid. The assumptions which underlie the theory are stated, and the effects of the neglected phenomena are discussed qualitatively.

Anode-current vs. anode-voltage and anode-current vs. space-current curves representing observations made on a specially constructed tetrode are presented by way of experimental verification of the theoretical results.

For purposes of illustration, application is made of these results to elucidate the theory of the type of power-amplifier tube which employs a minimum potential, formed in front of the anode as a result of the space charge of the electrons, to minimize the passage of secondary electrons from anode to grid. In addition, it is shown how the decrease of anode current with increasing space current, which occurs when a space-charge-limited virtual cathode is formed in the grid-anode region, may be utilized to provide negative transconductance amplifiers and oscillators.

Published in *RCA Review*, January, 1938.

AN ELECTRON OSCILLATOR WITH PLANE ELECTRODES

By

B. J. Thompson and P. D. Zottu*

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

This paper describes a new type of thermionic tube capable of producing ultra-high frequencies by means of electron oscillations. Tubes of this type are characterized by having parallel plane electrodes, instead of cylindrical electrodes as in the conventional Barkhausen-Kurz tubes, and a fourth element called a backing plate.

The relations between wavelength and amplitude of oscillation and the various electrode potentials are shown by measurements on a typical tube. It is found that in these tubes the filament voltage is not critical, space-charge-limited operation being satisfactory, and that only one mode of oscillation is obtained. Both of these factors appear to give these tubes an advantage in stability over cylindrical Barkhausen-Kurz tubes.

A tube of the flat type is described which has produced oscillations at a wavelength of less than 10 cm in the fundamental mode with a positive grid potential of 150 volts.

* Formerly with RCA Manufacturing Company.

Published in *Proc. I.R.E.*, December, 1934.

THE DEVELOPMENTAL PROBLEMS AND OPERATING CHARACTERISTICS OF TWO NEW ULTRA-HIGH-FREQUENCY TRIODES

By

Winfield G. Wagener

Formerly with

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

Large values of power are difficult to obtain in the ultra-high-frequency region. At the limiting frequencies it is increasingly more difficult to find vacuum tubes that will deliver such power and perform efficiently. The principal factors that affect the design and performance of the tubes are those involving the electrical circuit, the size requirements for the power desired, and the transit time of the electrons within the evacuated space of the tube.

The design principles that result from a consideration of these factors have been used in the development of two new ultra-high-frequency triodes. A triode capable of delivering approximately 700 watts at 100 megacycles is described. This tube, which is cooled by water and air, is capable of operation as a neutralized power amplifier up to 200 megacycles with an output of approximately 500 watts. A second triode is described which is a radiation-cooled glass tube with a 300-watt plate-dissipation rating. Normal efficiency is obtained up to 40 megacycles and operation as a neutralized power amplifier is possible up to 100 megacycles. The efficiency at 100 megacycles is approximately 60 per cent.

Published in *Proc. I.R.E.*, April, 1938.

A PUSH-PULL ULTRA-HIGH-FREQUENCY BEAM TETRODE

By

A. K. Wing

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

The design of a vacuum tube capable of delivering 10 watts useful power output at frequencies of the order of 250 megacycles and with a d-c plate voltage of 400 volts and good economy of space and cathode power, is discussed. In order to keep the physical dimensions of the tube small and to make it adaptable to straightforward circuit arrangements, the tube was designed as a push-pull beam tetrode. Unusual constructional features include the use of short, heavy leads sealed directly into the moulded glass bulb.

Characteristics of the tube are given. Tests show that the tube will operate as a stable Class C amplifier at frequencies up to 250 megacycles. At that frequency a power output of the order of 13 watts with an efficiency of 45 per cent has been obtained. Satisfactory operation as a frequency multiplier is possible in the same frequency range. Oscillator operation has been obtained at considerably higher frequencies. The variation of output and efficiency with frequency is shown.

Published in *RCA Review*, July, 1939.

NEW TELEVISION-AMPLIFIER RECEIVING TUBES

By

A. P. Kauzmann

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

Television circuits require amplifying tubes of high grid-plate transconductance and low input and output capacitances to realize a voltage gain per stage sufficient to keep the amplifying stages at a reasonable number. To this end the 1851 and 1852, sharp cut-off, 9000-micromho grid-plate transconductance tubes—and the 1853, semi-remote cut-off, 5000-micromho grid-plate transconductance tube were developed. The improvements are the result of decreasing the control-grid-to-cathode spacing, and at the same time decreasing the pitch and diameter of the control-grid wires.

The maximum allowable resistance in the control-grid circuit is determined from the grid-plate transconductance of the tube, the cathode-bias resistor, and the screen voltage-dropping series resistor. Also, the use of a small un-bypassed resistor in the cathode circuit to neutralize the changes in input capacitance and input loading with varying plate current is presented.

The 1851 and 1852 have the highest ratios of grid-plate transconductance to plate current of any commercially available tubes with the result that they have high signal-to-noise ratios. The high grid-plate transconductances also result in high-conversion transconductance when these tubes are used as mixed tubes; the 1851 and 1852 give a maximum of 3500 micromhos, and the 1853 a maximum of 1500 micromhos. With practical circuits the 1851 and 1852 have produced gains per stage of 3.5 to 7.0 at 50 megacycles, and of 20 to 45 at 11 megacycles. Similarly, the 1853 has produced gains per stage of 2 to

4 at 50 megacycles, and of 6.5 to 13 at 11 megacycles. All of these values are for a band-pass of 2.5 megacycles.

Published in *RCA Review*, January, 1939.

DESIGN AND USE OF "ACORN" TUBES FOR ULTRA-HIGH FREQUENCIES

By

BERNARD SALZBERG

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

SUMMARY

Early work on the extension of the high-frequency limit of receiving equipment indicated that in conventional circuits employing standard tubes, the tubes became less and less effective as the frequency was increased and that ultimately, at frequencies of the order of 200 megacycles, tubes were practically useless. B. J. Thompson and G. M. Rose, Jr., conclusively demonstrated that the limitations to the successful operation of vacuum tubes at the higher frequencies may be overcome by reducing the dimensions of the tubes. The small tubes, which they built, being intended for a research study, were not suited for manufacture. However, the possibilities of these tubes aroused an interest sufficiently wide-spread to warrant their further development for use by experimenters. This article deals with the factors which were involved in the design and application of the first of such tubes, a Lilliputian triode.

Published in *Electronics*, September, 1934.

REVIEW OF ULTRA-HIGH-FREQUENCY VACUUM-TUBE PROBLEMS

By

B. J. THOMPSON

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J,

SUMMARY

The effects of electron transit time and of lead inductance and interelectrode capacitance which become of importance at ultra-high frequencies are reviewed. It is shown that, in radio-frequency amplifier tubes for use in receivers, the most serious effect is the high ratio of input conductance to transconductance, which is independent of the transconductance and is proportional to the square of the frequency. This ratio may be reduced by reducing the electron transit time, the lead inductances, and the interelectrode capacitances. In power amplifiers and oscillators for transmitters, the important effects are much the same as in receiving tubes. The solution is complicated, however, by the problem of obtaining sufficient power output in a small structure.

Published in *RCA Review*, October, 1938.

CONSTRUCTION AND ALIGNMENT OF THE TELEVISION RECEIVER

By

C. C. Shumard

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

The design of coupling circuits which will pass the extraordinarily wide frequency band necessary for good picture reception is one of the most critical features of a television receiver. In this article the construction of these circuits is discussed in detail. The exact dimensions given, as well as the carefully worked out alignment procedure, should enable the constructor to get the receiver into operation with a minimum of trouble. In conjunction with the deflection circuits for small cathode-ray tubes described in October *QST*, it constitutes a complete television receiving system.

Published in *QST*, January, 1939.

Part II—Ultra-High Frequencies Above 300 Mc.

APPLICATION OF FREQUENCIES ABOVE 30,000 KILOCYCLES TO COMMUNICATION PROBLEMS

By

H. H. Beverage,[1] H. O. Peterson,[2] and C. W. Hansell[3]

[1]R.C.A. Communications, New York City; [2]Riverhead, L. I., N. Y.; [3]Rocky Point, L. I., N. Y.

Summary

The authors briefly describe the results of a number of experiments with frequencies above 30,000 kc covering a period of several years. Since the major interests of radio communication companies has been in long-distance communications, this paper includes some qualitative data covering propagation beyond the optical, or direct vision, range. The authors have found that the altitude of the terminal equipment location has a marked effect on the signal intensity, even beyond the optical range.

Frequencies below about 43,000 kc appear to be reflected back to earth at relatively great distances in the daytime in north-south directions, but east-west transmission over long distances is extremely erratic.

Frequencies above about 43,000 kc do not appear to return to earth beyond the ground wave range, except at rare intervals, and then for only a few seconds or a few minutes. These frequencies, which do not return to earth, also appear to be free of echoes and multiple path transmission effects. Therefore, they are free from distortion due to selective fading and echoes. The range is also limited to the ground wave range, so these frequencies may be duplicated at many points without interference. As the frequency is raised, the range tends to approach the optical distance as a limit. Experiments with frequencies above 300,000 kc have, so far, indicated that the maximum range is limited to the optical distance. A number of possible applications are suggested, based on the unique properties of these frequencies. A specific application to telephony between the Islands of the Hawaiian group is briefly described.

Published in *Proc. I.R.E.*, August, 1931.

NOTES ON PROPAGATION AT A WAVELENGTH OF SEVENTY-THREE CENTIMETERS

By

B. Trevor and R. W. George

R.C.A. Communications, Inc., Riverhead, L. I., N. Y.

Summary

Quantitative field strength observations have been made on a wavelength of 73 centimeters with improved apparatus. The methods are described. Propagation data have been obtained with the receiver installed in an automobile and an airplane. Further observations were made on the ground at a distance of 113 miles, 8,000 feet below the line of sight from the transmitter. The results show the nature of the propagation of 73-centimeter waves over distances up to 175 miles. Below the transmitter's horizon, rapid attenuation occurs with increase in distance from the transmitter, the plane of polarization of the transmitted signal remains unchanged, and various types of fading are observed.

Published in *Proc. I.R.E.*, May, 1935.

DESCRIPTION AND CHARACTERISTICS OF THE END-PLATE MAGNETRON

By

Ernest G. Linder

RCA Manufacturing Company, Inc., Camden, N. J.

Summary

A new type of magnetron is described which is especially adapted to the generation of centimeter waves. It possesses several advantages over the simple magnetron, namely: (1) greater stability with respect to fluctuations of supply voltages, (2) less tendency to oscillate at undesired long wavelengths, (3) greater efficiency, (4) greater output, and (5) greater ease of adjustment. Static and dynamic characteristics are discussed. The effect of space charge on electron motion and tube performance is treated mathematically, and supporting experimental data are presented. Evidence is given that for best operation an optimum space-charge condition is required, which can conveniently be established and maintained by the use of end plates. Power output is limited by a type of instability involving electron bombardment of the filament, and apparently initiated by excessive space charge.

Published in *Proc. I.R.E.*, April, 1936.

THE MAGNETRON AS A HIGH-FREQUENCY GENERATOR

By

G. Ross Kilgore

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary

The search for methods of generating ultra-high frequencies has followed two main paths. One has been the improvement of the negative grid tube as an oscillator and amplifier so as to extend its upper frequency limit. The other has been the investigation of less conventional vacuum tubes that appear to be applicable at high frequencies. Of the latter group the magnetron now stands out as one of the most promising tubes.

The device first called a "magnetron" by Hull in 1921 is well known as a vacuum tube having a cylindrical plate structure and coaxial filament, with a uniform magnetic field directed along the electrode axis. Its use as a generator of high frequency currents has developed mainly in the past decade and its investigation has claimed the attention of a large number of research workers in many countries.

Magnetrons in a variety of structural forms and in a number of operating modes have been used for oscillation generation. Broadly speaking, magnetron oscillators can be divided into two classes; one using an alternating magnetic field, and the other using a constant

magnetic field. The alternating-field type described by Elder, in which the field coil is part of the oscillating circuit, is obviously limited to low frequencies and need not be considered in the present discussion. The constant field types which are useful in generating ultra-high frequencies can be subdivided as negative resistance oscillators and transit time oscillators.

The negative resistance magnetron oscillator may be defined as one which operates by reason of a static negative resistance between its electrodes and in which the frequency is equal to the natural period of the circuit. In Europe this type is sometimes called "dynatron magnetron" or Habann oscillator.

The transit time magnetron oscillator may be defined as one which operates by reason of electron transit time phenomena and in which the frequency is determined by the electron transit time. This type is often referred to as an "electronic oscillator" and sometimes as a "magnetostatic oscillator."

It is the purpose of the present paper to survey the various types of magnetron generators with particular reference to their performance and limitations at ultra-high frequencies.

Published in *Jour. of Applied Physics*, October, 1937.

MICRO-WAVES IN NBC REMOTE PICK-UPS

By

ROBERT M. MORRIS

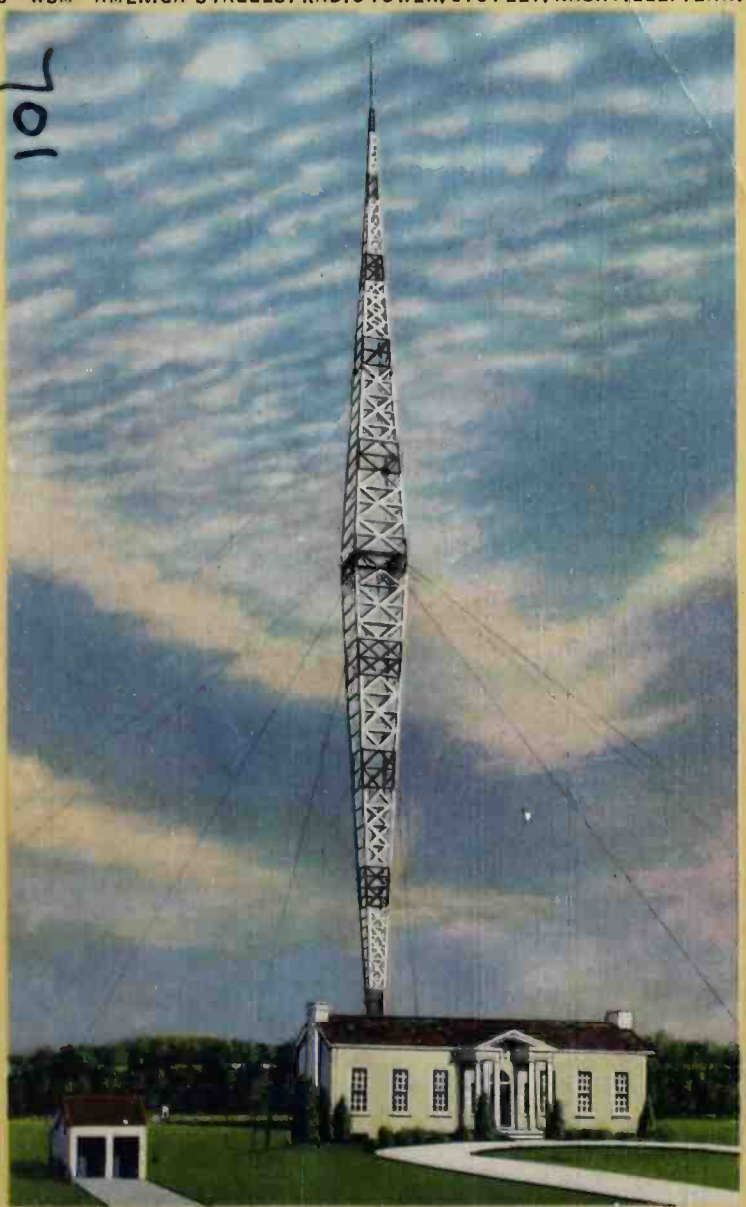
Development Engineer, National Broadcasting Company, New York

SUMMARY

The National Broadcasting Company in its efforts to best serve the public interest has constantly endeavored to take its microphones to more and more remote points. To do this has required the use of radio to an increasing extent. In many cases wire lines have not been available in time to do the job. In other instances the very nature of the pick-up point (ship, airplane, balloon, etc.) has necessitated the use of radio circuits. Radio apparatus for these special events or spot news broadcasts has in general fallen into two classifications; first, portable units of from 10 to 100 watts and second, self-contained packsets of less than 1 watt, and light enough to be carried by one person. The extreme mobility of this latter type of equipment has made it of special interest for unusual or difficult broadcasts.

Published in *RCA Review*, July, 1936.

8—WSM—AMERICA'S TALLEST RADIO TOWER, 878 FEET, NASHVILLE, TENN.



5A-H430

323 FEET HIGHER THAN THE WASHINGTON MONUMENT

RESTRICTIONS:
CONTENT:
Mail, cash or cash equivalents and all
s are prohibited.
ERNATIONALLY 20 POUND WEIGHT

specific prohibitions/restrictions may
ional Mail Manual (IMM) country pages for

e affix customs declaration PS Form 2976-A
6-E customs form pouch.

for Destinataire

Country of Destination Pays de destination

PRIORITY®
MAIL
ATES POSTAL SERVICE

n Flat Rate Box
c and International Use

packaging products have been awarded Crad
dle Certification℠ for their ecologically intelli
n. For more information go to mbdc.com/usps
Cradle Certified℠ is a certification mark of MBDC.