PRINCIPLES

OF

RADIO COMMUNICATION

WORKS OF

THE LATE J. H. MORECROFT

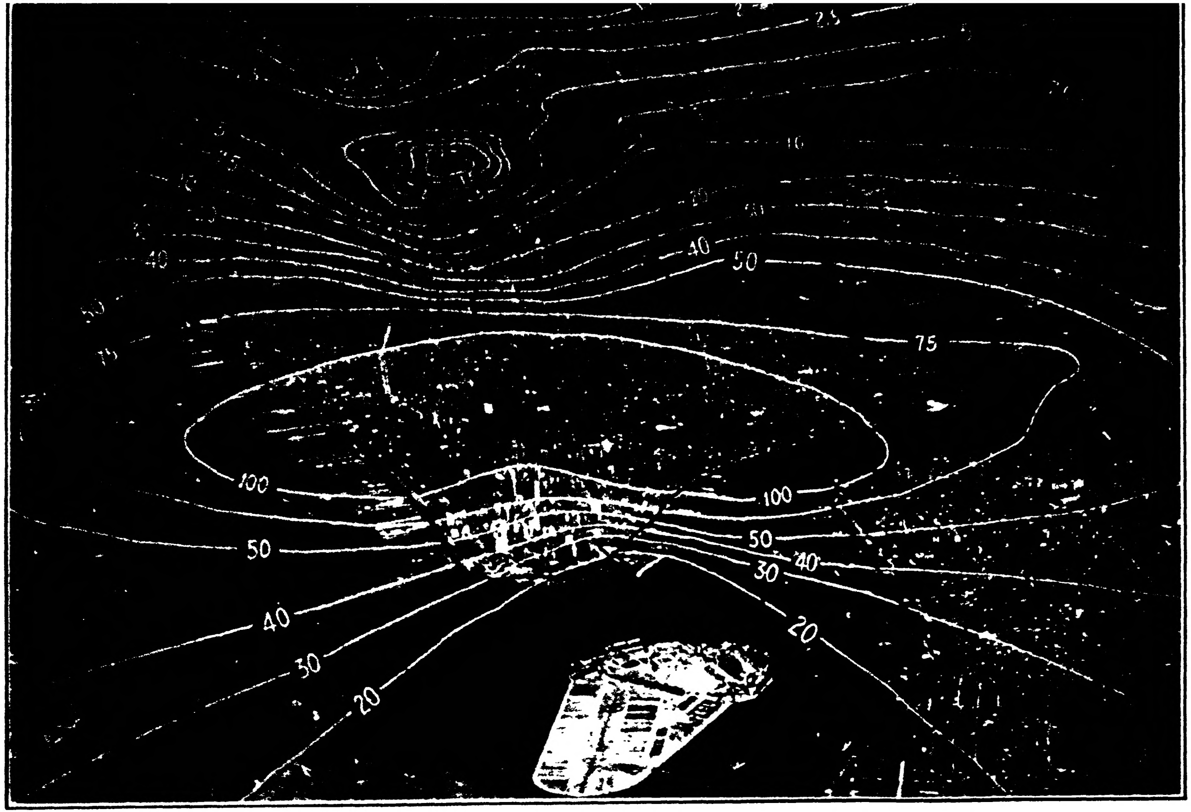
Electron Tubes and Their Application
A text and reference book on the theory and applications of electron tubes. By J. H. Morecroft. Second Printing, Corrected. 458 pages. 6 by 9. 539 figures. Cloth.

Elements of Radio Communication
An Introductory volume to the more comprehensive text "Principles of Radio Communication," by J. H. Morecroft. Second Edition. xii + 286 pages. 6 by 9. 241 figures. Cloth.

Principles of Radio Communication
A text dealing with all phases of the radio art. By J. H. Morecroft, Assisted by A. Pinto and W. A. Curry. Third Edition, thoroughly revised. xviii + 1084 pages. 6 by 9. 1090 figures. Cloth.

PUBLISHED BY

JOHN WILEY & SONS, INC.



A radio topographic map of New York City. The figures give the "millivolts per meter" of the carrier wave when not modulated. The station was located among the tall steel buildings in lower Manhattan. (For other radio topographic maps see Chapters IV and IX.)

Frontispiece.

PRINCIPLES

OF

RADIO COMMUNICATION

BY

JOHN H. MORECROFT, D.Sc.

Late Professor of Electrical Engineering, Columbia University
Past President Institute of Radio Engineers

ASSISTED BY

A. PINTO

AND

W. A. CURRY

THIRD EDITION

NEW YORK

JOHN WILEY & SONS, Inc.

London: CHAPMAN & HALL, Limited

IN THE REPRINTING OF THIS BOOK, THE RECOMMENDATIONS OF THE WAR PRODUCTION BOARD HAVE BEEN OBSERVED FOR THE CONSERVATION OF PAPER AND OTHER IMPORTANT WAR MATERIALS. THE CONTENT REMAINS COMPLETE AND UNABRIDGED.

THIRD EDITION
Fifth Printing, September, 1944

Printed in U. S. A.

PREFACE TO THE THIRD EDITION

The five years which have elapsed since the second edition of this text was prepared have seen great advances in the radio art, not only in the types of apparatus, amounts of power radiated, etc., but also in the popular appreciation of the value of radio communication. The radio receiving set is now as much a part of the home equipment as are the electric light and automobile. It is no longer a question as to whether the home shall have a radio set—it is merely a question as to what type it shall be.

During the last five years the battery-operated set has virtually disappeared from the market; the sets of today are practically always operated from power obtained by rectifying the alternating current house supply. In this edition, therefore, more material has been added on the subject of rectifying apparatus and circuits, and the action of filters.

With increasing amplifications the shielding of radio sets has become more important, and so more material in this field has been added. The electrolytic condenser is finding more extensive application and has therefore been analyzed with reasonable detail, and its performance shown. The newer types of tubes, and wider fields of use of the older ones have received proper notice.

The action of piezo active quartz and its use in frequency control has been explained more in detail, and the use of crystal oscillators, with degree of frequency control available, is shown.

The action of microphones, with actual calibration data on the different types, is taken up and the question of harmonics introduced by them into the modulation is discussed. The action of modulators is more thoroughly explained together with experimental data on their performance.

Radio circuits have become sufficiently standardized by now, to warrant giving some reasonably complete circuit layouts, and this is done, for both telegraphy and telephony channels.

An attempt has been made to introduce the more important findings regarding the action of short waves and their reflection from the ionized layers of the atmosphere; the action of this reflecting atmosphere, its location and movement, and its effect in changing the apparent direction of radio waves are given more in detail.

There is much new material dealing with the attenuation of radiated energy and the fields covered by a given amount of power under different conditions. The practicability of directed radiation and reception has been analyzed, and its use in commercial channels explained and illustrated.

Radio beacons for airplane guidance have been greatly developed during the last five years, due largely to the activity of the Radio Section of the Bureau of Standards; a résumé of their findings, and the present types of beacons used, is introduced at suitable places in the text.

Still more of the chapter on spark telegraphy has been deleted, as this type of radio communication finds less and less application. However, such material as may serve for analysis of shock excitation, etc., has been retained, as this will probably have permanent value.

J. H. M.

June 1, 1932.
Columbia University,
New York City.

PREFACE TO SECOND EDITION

The very cordial reception accorded the first edition of this book has encouraged the author to attempt a thorough revision with the idea of bringing it completely up to date. As in the previous edition, however, no pretense is made that the book is a treatise on radio practice; in general only the principles involved in the operation of radio apparatus have received attention. Whatever radio apparatus is discussed is dealt with only to illustrate those principles the text is intended to elucidate.

When the first edition of the book was in preparation there was very scant radio literature from which to draw so that most of the data submitted in verification of the theory was laboriously obtained by the author himself. In the intervening years prolific publication on radio subjects has taken place with the result that this edition is much more thorough and complete than was the first. The data on the various phases of the radio art have become so plentiful that but a small part could be incorporated in this volume; to make the text as useful as possible, however, references have been given to most of the outstanding articles which have appeared in English during the past decade.

The new material incorporated in this edition so increased the size that it was thought advisable to delete much of the first edition. A considerable part of the chapter on Spark Telegraphy has been taken out, therefore, and two of the chapters of the earlier edition have been deleted. The chapter on radio measurements, and that on experiments, have been omitted; if a demand for the material of these two chapters appears it will be published as a separate volume.

Notable additions to the older edition occur in Chapters II, IV, VIII and X. In Chapter II many new data on coils and condensers at radio frequencies are given. In Chapter IV, dealing with the general features of radio transmission, new material on field strength measurements, reflection and absorption, fading, short-wave propagation, etc., has been introduced. In Chapter VIII (radio telephony), a great deal of material on voice analysis has been added; the performance of loud-speaking telephones, frequency control by crystals, etc., has been discussed. In Chapter X, dealing with amplifiers, the question of distortionless amplification has been thoroughly dealt with, some of the material being given for the first time. The question of radio-frequency amplification, balanced circuits, push-pull arrangements, etc., have been explained.

No attempt has been made to give credit to the various inventors of radio apparatus, because nearly all of the important radio patents are at present in litigation and a text-book writer does not have available the material on which to base a reasonable conclusion as to whom the credit belongs. Only after the courts have given their final decisions can the various radio developments be ascribed properly to the different pioneer workers.

J. H. M.

April 13, 1927.
COLUMBIA UNIVERSITY,
NEW YORK CITY.

PREFACE TO FIRST EDITION

The student desiring to familiarize himself with the theory and practice of Radio Communication should be thoroughly grounded in the ordinary laws of continuous and alternating-current circuits; he should also have a clear physical conception of the transient conditions continually occurring in such circuits. These elementary ideas are best obtained by considering the electric current from the electron viewpoint, i.e., as a comparatively slow drift of innumerable minute negative electric charges, which, at the same time they are drifting through the substance of the conductor, are executing haphazard motions with very high velocities, continually colliding with each other and with the molecules of which the conductor is composed.

Due to the extremely high frequencies encountered in radio practice it is necessary to expand somewhat one's ideas of resistance, inductance, and capacity, the so-called constants of the electric circuit. As a result of the non-uniformity of current distribution the resistance of a conductor at high frequency is generally much higher in a radio circuit than it is at ordinary engineering frequencies; due to non-penetration of magnetic flux and hysteretic lag, the apparent permeability of an iron core is much less at radio frequencies than at the customary sixty cycles; due to imperfect polarization of dielectrics the apparent specific inductive capacity of an insulator may be much decreased at radio frequencies and the heating due to dielectric losses may be thousands of times as great as is the case in ordinary engineering practice. Furthermore, due to the unavoidable internal capacity, the apparent inductance of even an air core coil may be expected to vary at high frequencies; in fact, a piece of apparatus which is physically a coil, when used at radio frequencies, may, by electric measurement, be found a condenser.

All of the effects indicated above are treated in the early chapters of the text, not in as comprehensive manner as is possible, to be sure, but with sufficient thoroughness to open the student's eyes to the possible peculiar behavior of circuits when excited by the very high frequencies of radio practice.

Because of its importance to the radio art a considerable part of the text is given over to the theory and behavior of the thermionic three-given for the first time. The question of radio-frequency amplification, balanced circuits, push-pull arrangements, etc., have been explained.

No attempt has been made to give credit to the various inventors of radio apparatus, because nearly all of the important radio patents are at present in litigation and a text-book writer does not have available the material on which to base a reasonable conclusion as to whom the credit belongs. Only after the courts have given their final decisions can the various radio developments be ascribed properly to the different pioneer workers.

J. H. M.

April 13, 1927.
Columbia University,
New York City.

PREFACE TO FIRST EDITION

THE student desiring to familiarize himself with the theory and practice of Radio Communication should be thoroughly grounded in the ordinary laws of continuous and alternating-current circuits; he should also have a clear physical conception of the transient conditions continually occurring in such circuits. These elementary ideas are best obtained by considering the electric current from the electron viewpoint, i.e., as a comparatively slow drift of innumerable minute negative electric charges, which, at the same time they are drifting through the substance of the conductor, are executing haphazard motions with very high velocities, continually colliding with each other and with the molecules of which the conductor is composed.

Due to the extremely high frequencies encountered in radio practice it is necessary to expand somewhat one's ideas of resistance, inductance, and capacity, the so-called constants of the electric circuit. As a result of the non-uniformity of current distribution the resistance of a conductor at high frequency is generally much higher in a radio circuit than it is at ordinary engineering frequencies; due to non-penetration of magnetic flux and hysteretic lag, the apparent permeability of an iron core is much less at radio frequencies than at the customary sixty cycles; due to imperfect polarization of dielectrics the apparent specific inductive capacity of an insulator may be much decreased at radio frequencies and the heating due to dielectric losses may be thousands of times as great as is the case in ordinary engineering practice. Furthermore, due to the unavoidable internal capacity, the apparent inductance of even an air core coil may be expected to vary at high frequencies; in fact, a piece of apparatus which is physically a coil, when used at radio frequencies, may, by electric measurement, be found a condenser.

All of the effects indicated above are treated in the early chapters of the text, not in as comprehensive manner as is possible, to be sure, but with sufficient thoroughness to open the student's eyes to the possible peculiar behavior of circuits when excited by the very high frequencies of radio practice.

Because of its importance to the radio art a considerable part of the text is given over to the theory and behavior of the thermionic three-electrode tube; at the time this material was compiled there was no comprehensive treatment of the subject anywhere, but there has recently appeared an excellent volume on Vacuum Tubes (by H. J. Van der Bijl) which every student of radio should carefully peruse. It is hoped that the subject matter presented in this text may supplement, rather than duplicate, that given in the above-mentioned volume; the actual behavior of tubes in typical circuits is covered in this text in a more thorough manner than has been attempted in other texts, and practically all the theoretical deductions are substantiated by experimental data, much of which has been obtained in the author's laboratory.

A chapter has been devoted to each important phase of the radio art; there is also incorporated a short course of elementary experiments which may well be carried out by electrical engineering students especially interested in Radio. For those desiring to specialize in Radio, the material given in the body of the text will furnish ideas for unlimited further experimentation.

On certain parts of the text very valuable assistance has been given by the author's former colleague, Mr. A. Pinto, and by Mr. W. A. Curry, who is at present associated with him in radio instruction; due credit is given to them on the title page of the text.

J. H. M.

Columbia University,
April, 1921.

CONTENTS

CHAPTER I

Fundamental Ideas and Laws

CHAPTER II

Resistance—Inductance—Capacity—Shielding

CHAPTER V

Spark Telegraphy

CHAPTER VI

Vacuum Tubes and Their Applications

CHAPTER IX

Antennas and Radiation

CHAPTER X

Amplifiers

PRINCIPLES

OF

RADIO COMMUNICATION

CHAPTER I

FUNDAMENTAL IDEAS AND LAWS

Nature of Electricity.—Everyone is more or less familiar with elementary experiments having to do with electrically charged bodies. Fur, if rubbed on a dry day, crackles and gives off minute sparks; a glass rod rubbed with a cloth becomes electrified and will attract small bits of paper, cotton, etc.; due to wind friction, and other causes, clouds become intensely electrified and are able to break down the insulating strength of the air and produce sparks thousands of feet long.

In what way does an electrified body, or electrically charged body, differ from one in the uncharged, or neutral, state? A reasonable answer to this question is found in the modern conception of the constitution of matter.

Electrons.—It has been firmly established that every atom of matter is charged with minute particles [1] of negative electricity, so-called *electrons.* An electron, when detached from the atom of matter with which it was

[1] It may seem difficult at first to think of electricity as made up of separate, discrete quantities instead of a continuous distribution of electric charge, but it is pointed out that according to modern concept *energy itself* is always present as a certain number of *unit quantities*; that is, energy itself is to be "counted" in terms of the smallest possible quantity, called a "quantum."

During the last year or two, scientists of the Bell Telephone laboratory have carried out certain experiments which lead to the conclusion that electrons are waves, that is, definite amounts of energy in the form of wave trains. The experiment consisted essentially of shooting beams of electrons through crystals onto a photographic plate; the patterns which the emerging electrons produced on the plate gave almost conclusive proof that the electrons, at least during the time they were traversing the crystal, were groups of electromagnetic waves.

associated, shows none of the properties of ordinary matter. It does not react chemically with other electrons to produce some new substance; moreover, all electrons are similar, no matter from what type of atom they have been extracted. Thus an electron from the hydrogen atom acts precisely the same as the electrons from atoms of oxygen, iron, chlorine, or any other substance. It seems that the *electron is nothing but electricity*. It is definite in amount, always being exactly the same, and is generally believed to be the smallest possible quantity of electricity, i.e., electricity cannot be subdivided into quantities smaller than the electron.

The constants of the electron are: Radius[1] $=2\times10^{-13}$ cm.; mass $=8.99\times10^{-28}$ gram; charge $=1.59\times10^{-19}$ coulomb.[2] The mass of the electron depends upon the velocity with which it is moving; the value given here holds good only if the electron is traveling at velocities considerably less than the velocity of light, say less than 10^9 cm./sec.

The mass of the electron is determined from the curvature of its path as it travels at high velocities through a transverse electric or magnetic field.

It is found from the curvature of these paths that the low velocity mass of the electron, namely, 8.99×10^{-28} gram, must be multiplied by the fraction $\left(\frac{1}{1-\left(\frac{v}{V}\right)^2}\right)$ to represent its mass when traveling with a velocity commensurate with that of light. In this expression v is the velocity of the electron (with respect to the observer) and V is the velocity of light (practically 3×10^{10} cm. per sec.).

For many years it has been the custom for physicists to speak of positive electricity and negative electricity; from this standpoint the electron is negative electricity. All electrons are the same kind, or polarity, hence follows our present conception that *the electron is the smallest possible quantity of negative electricity.*

Charged Body.—From the electron viewpoint a negatively charged body is one having more than its normal number of electrons and a positively charged body is one having less than its normal number of electrons. Let the circular shape in Fig. 1 represent an atom of helium;[3] the small

[1] The radius of an electron is a rather arbitrary dimension, and is not a definitely measurable quantity such as the radius of a metal ball; it is obtained from the results of certain experiments on the collision of electrons. Rutherford estimates the radius of the nucleus of the very heavy elements (e.g., uranium) as 10^{-12} cm.

[2] The student who is particularly interested in the theoretical and experimental work from which these values are obtained is referred to "Conduction of Electricity through Gases," by J. J. Thomson.

[3] In recent years much work has been done in investigation of the structure of

circles with the minus sign in them represent the electrons associated with the normal helium atom. The normal atom is not charged; it does not exert any attractive or repulsive force on the other atoms, due to its electrical state.

The structure of the atom of nearly every substance is now known with a remarkable degree of certainty, this knowledge having been gained for the most part by experiments with the reflection and transmission of X-rays by the substance in question. In the center part of the atom are grouped some electrons and some positive charges, these being frequently called *positive electrons*; the amount of positive electricity on the positive electron is just equal to the amount of negative electricity on the negative electron, so that an atom having an equal number of positive and negative electrons shows no electric charge, the negatives and positives just neutralizing each other.

The center part of the atom, consisting of closely grouped numbers of positive and negative electrons, always has more positives than negatives, so that this group (called the *nucleus* of the atom) always exhibits a net positive charge; the nucleus is thus said to be positively charged.

Grouped around the nucleus (possibly as the planets are grouped around the sun) are numbers of negative electrons, this number being just sufficient in the uncharged atom to neutralize the excess positive charge of the nucleus. The number and arrangement of these outer electrons determine what the atom is, whether hydrogen, oxygen, copper, gold, chlorine, etc. All substances are made up of the two elementary quantities, positive and negative electrons, and the nearly one hundred different elements we know differ only in the number and arrangement of the two kinds of electrons.

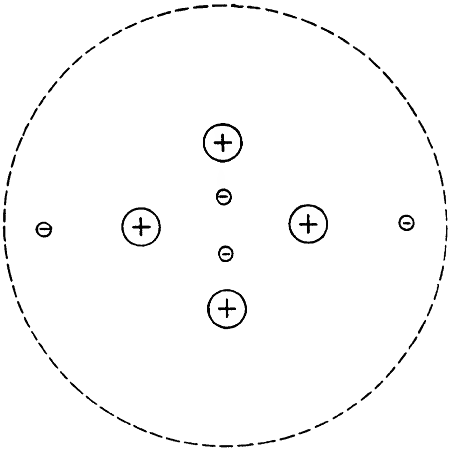


Fig. 1.—Conventional model of a simple, neutral, atom.

The heavier the substance the more complicated is the atom; thus the hydrogen atom has only two electrons, one positive and one negative, whereas such elements as tungsten and mercury have very many electrons, placed in certain well-known complex arrangements.

the atom; an interesting and elementary exposition of some of the modern views is given in "The Nature of Matter and Electricity," by Comstock and Troland.

If one electron is removed from the atom by some means or other (represented in Fig. 2) the balance between positive and negative charge is destroyed; an excess of positive charge exists on the atom and the atom is positively charged. The electron which has been removed from the atom constitutes a negative charge. If the electron is allowed to go back to the atom the balance of charge is restored and the atom is again uncharged, or neutral.

A positively charged body, therefore, is one which has been deprived of some of its normal number of electrons; a negatively charged body is one which has acquired more than its normal number of electrons. Thus when a piece of sealing wax is rubbed with dry flannel the wax becomes negatively charged and the flannel becomes positively charged. The friction between the wax and the flannel must have rubbed some of the electrons off the flannel molecules and left them on the surface of the wax.

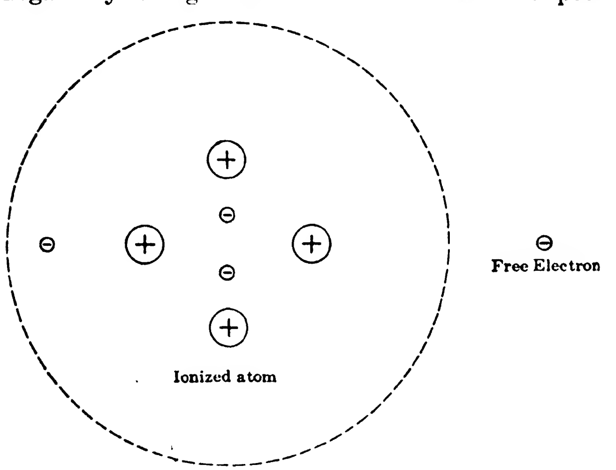


Fig. 2.—Conventional model of a simple atom charged positively, one of its electrons being free.

The extra electrons on the wax are attracted by the deficient molecules of the flannel (positive and negative charges attract each other) and if the flannel and wax are left together after being rubbed they soon lose their charges; the molecules of the flannel regain their proper number of electrons.

Number of Electrons Removable from an Atom.—Although there may be a great number of electrons associated with an atom or molecule it is generally not possible to remove more than one; in a body which is positively charged most of the atoms are neutral, having their proper complement of electrons; others have had one electron removed. If but few of the atoms of a body have had an electron removed the body has a small charge; the more highly the body is charged the more of the deficient atoms there are on it.

The reason that only one electron is generally removable becomes evident when we think of the electric forces at play in the atomic structure. With its proper complement of positive and negative electrons the normal atom (uncharged) exerts no force on charged bodies in its vicinity; thus there is practically no force tending to make some free wandering electron move to a neutral atom. But if one electron has been removed from an atom, the free electron will be strongly attracted thereto and will fall into the atomic structure to fill the place of the missing electron.

As a new electron joins the atomic structure of a deficient atom it does not generally fall quietly into place, but executes some type of oscillatory motion while it is "settling down." These oscillatory motions occur with the frequency of light waves, and in fact if they happen to fall within the range of visible light, they become evident to the eye. The frequency of these oscillatory motions (hence color of the light) depends upon the electric structure of the atom, that is, upon the substance itself.

After one electron has been removed from an atom it is more difficult to remove another because of the positively charged condition of the deficient atom; evidently it becomes increasingly difficult to remove others. In certain experiments it has been possible to remove several electrons, but to accomplish this excessive forces are required. Sometimes it is done by intense heat and sometimes by excessive energy of bombardment by high-speed electrons.

The breaking up of various atoms by taking away some of the outer electrons has extended greatly in recent years; it is now known that the transmutation of one element into another is quite possible, and such change is occurring in radioactive substances all the time. Radium, a very heavy and complex atom, is gradually disintegrating (by throwing off positive and negative electrons) into lead; lead itself may be changing into some other substance, but so slowly that within the short time man has experimented, it may not be possible to detect the change.

On the other hand, the "cosmic rays," about which so much has been written in recent years, seem to prove that in interstellar space elements are being formed, instead of broken down. When the lighter elements combine to form heavier ones, such as iron and nickel, energy is presumably given off in the form of these very short electromagnetic waves, the presence of which the experimenter determines by their ionizing action in the air. Their wave length is much less than that of X-rays, hitherto regarded as the shortest waves generated.

From this modern viewpoint, therefore, it seems that the amount of charge on a body should be counted; the charge consists of discrete things. Instead of saying that a body has a certain amount of negative electricity on it, we might more reasonably say that a certain number of electrons has been deposited on it.

Electric Fields.—If a light substance, such as a pith ball, is touched to a charged body, it becomes charged with electricity of the same polarity as that on the body itself; as like charges repel one another the pith ball will be repelled from the charged body. By experimenting it may be found that the repulsive force between the pith ball and the original charge exists even when there is considerable distance between the two. The space surrounding a charged body is evidently under some kind of strain which enables it to act upon a charged body with a force, attractive or repulsive, according to the relative polarities of the two charges. This space surrounding a charged body, in which another charged body is acted upon by a force tending to move it, constitutes an **electric field,** sometimes called an electrostatic field.

Such an electric field surrounds every charged body; it really extends to infinity in all directions from the charged body, but as the force becomes very small as the distance is increased it is generally considered that the electric field due to a charge extends but a short distance from the charge. For example, the field due to a piece of charged sealing wax is negligible at a point a few feet distant from the wax, so we say that the field of this charge extends but a few feet from the wax. On the other hand, the electric field produced by a large, highly charged, wireless antenna may extend several thousand feet from the antenna.

Electric Fields Represented by Lines.—In diagrams the electric field surrounding a charge is most easily depicted by drawing lines from the charged body into the surrounding space. The direction of the lines, properly drawn, gives the direction of the electric force and the relative closeness of the lines in various parts of the diagram shows the relative strengths of the field at these points, the closer the lines the more intense the field. A line of force originating on a positive charge is generally shown ending on an equal negative charge. In diagrams it is not always convenient to so represent them; they may be shown as discontinuous. It must not be supposed, however, that the electric force itself is discontinuous; it always continues from a positive charge to a negative charge.

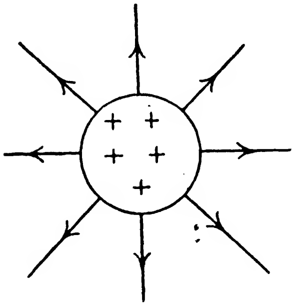


FIG. 3.—Electric field around a charged, isolated sphere represented by radial lines.

Fig. 3 shows how lines may be used to represent the electric field; it shows a positively charged metal ball supposedly far enough away from other bodies to be considered as by itself. The lines of force originate on the surface of the sphere and extend as radii in all directions. The arrow head on the lines indicates the direction in which a *positive charge* would be urged if placed in that part of the field.

The lines are closest together at the surface of the sphere, indicating that the force is greatest at this point, a fact easily proved experimentally.

Although the lines are shown as discontinuous, ending in uncharged space, each line really extends in some direction until it encounters a negative charge. In the case of a metallic sphere, suspended in the air distant from other bodies, the lines should all be shown as ending on the earth's surface as suggested in Fig. 4. Fig. 5 represents the electric field between two parallel metallic plates, one of which has been charged positively and the other negatively. Moreover, as all the lines originating on the positive plate are shown as ending on the negative plate, it shows that the two plates have been given equal charges. The field is properly shown as very intense between the two plates, weaker towards the edges, and very weak in the space not directly included between the two plates.

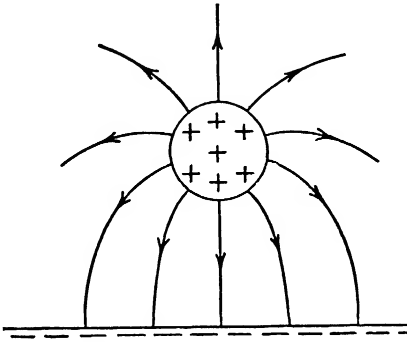


FIG. 4.—Charged body near the earth has its electric field radial near the body, all lines of force, however, bending over so that they end on the earth.

The Ether.—The actual electric field, at any point in space, is due to the combination of the individual fields of all the electric charges in the universe. Thus in Fig. 6 a negative and positive charge are shown close together; the electric field of each, individually, has a radial distribution, as indicated by the light radial lines. At any point in space the actual electric field, in so far as these two charges are concerned, is obtained by combining, vectorially, the two separate radial fields. The actual field is thus constructed at two points in Fig. 6. In a similar manner all other charges in the universe could be cared for and so the actual field at any point in space determined.

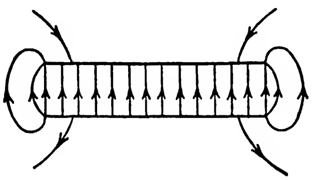


FIG. 5.—Two metallic plates, close to one another, one charged positively and the other negatively, have an intense electric field between the plates, and weak field elsewhere.

But instead of combining these different radial fields they may be kept distinct, at least in the imagination, and in some problems it is very necessary to do so, if a reasonable explanation is to be obtained.

As an illustration of this statement, the magnetic field induced inside a tubular conductor, carrying changing current, may be considered. As indicated later in the text, a magnetic field is nothing but a moving electric field, and so, from this concept, there can be no magnetic field induced inside the tubular conductor because there is no resultant electric field there. But if the electric fields of the individual electric charges are kept in mind, and their relative directions and motions considered, it will be seen that there is a magnetic field, *caused by a moving electric field, although there is no resultant electric field there.*

Most of the phenomena such as radio waves, propagation of light, etc., which apparently require for their explanation an ether with impossible properties, can be reasonably well pictured if we consider the ether as the superimposed electric fields of all the electric charges in the universe. The straight-line propagation of light, the e.m.f. of mutual induction set up at a distance from the inducing circuit, the radiation of energy from an accelerating electron, etc., may all be easily pictured from this concept of the ether. Of course it is to be remembered that building up a reasonable picture of a pheomenon is by no means the same as explaining it.

Some applications of this concept of the ether will be given in the chapter on radiation; whereas, as stated above, an imaginative picture of the ether, even if it offers a reasonable concept of such phenomena as radiation and induced voltages, is by no means an explanation of the ether, still it is well worth while to the average reader, who needs all the visualization possible in studying radio communication.

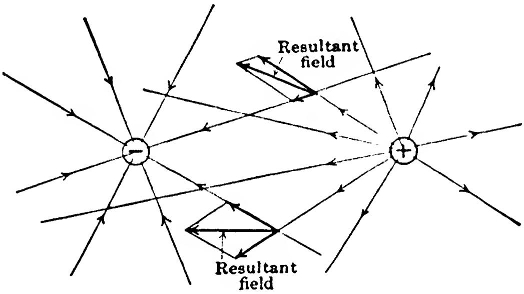


FIG. 6.—The electric field at any point is the vector resultant of the separate fields produced by the two charges.

Closed and Open Electric Systems.—In Fig. 5 most of the electric field is shown directly between the plates on which the charges are situated; such distribution of lines indicates a nearly closed electric system. The field illustrated in Fig. 4 is a comparatively open one; the distinction between open and closed field is not a very sharp one, but is nevertheless a very important one for the radio engineer.

Fig. 7 represents a vertical wire antenna, such as Marconi used in his early experiments; the electric field when the antenna is charged has the form shown. If the antenna is bent over in the form of an inverted L, the field has the form shown in Fig. 8. With the antenna in this form the most intense part of the electric field is evidently included directly between the earth's surface and the antenna wire, so the field is a closed one as contrasted with that of Fig. 7, which is regarded as an open field. The operating characteristics of the two antennas shown are quite different, the difference being due to the different distribution of the field in the two cases.

Induced Charges.—Suppose a charged metal ball is brought close to another conducting body as a metal rod, the rod being uncharged. Experiment shows that as the rod is brought into proximity of the brass ball the rod itself becomes charged in a peculiar way. If the ball is positively charged that end of the rod nearer to it becomes charged negatively and the farther end becomes positively charged as indicated in Fig. 9. As a whole the rod is not charged, there being as much negative charge as there is positive charge. These charges which have been produced on parts of the rod through the action of the charged ball are called *induced charges.*

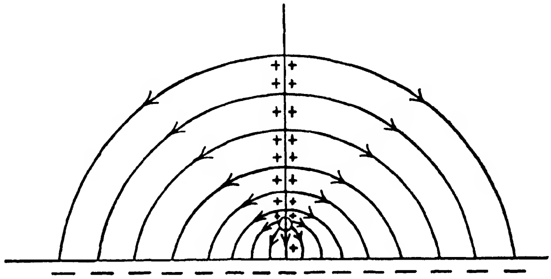


Fig. 7.—The electric field around a charged vertical wire.

Charges induced on a body are always double in kind; as much positive charge appears as does negative. However, if in Fig. 9 a wire having one end connected to the earth is touched to the end of the rod marked *C*, the positive charge which has been induced at this end of the rod will "run off to the earth," in the words of the ordinary text book on physics, and when the wire is removed there will be left on the rod only the negative charge.

Bound and Free Charges.—In the case considered above the positive charge runs off to the earth because there is no force tending to hold it on the rod; on the contrary, it is being repelled by the positive charge on the ball. The negative charge at *B* is held from running off to earth by the attractive force of the positive charge on the ball. The negative charge on the rod is called a *bound* charge and the positive charge, which runs away if given the opportunity, is called a *free* charge.

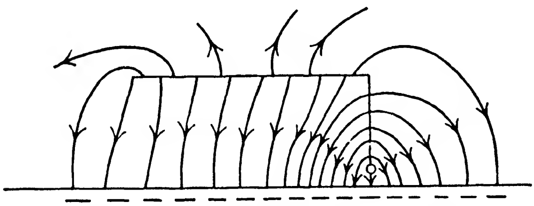


Fig. 8.—The electric field around an ordinary antenna.

An illustration of the way in which this method of producing charges is useful in radio circuits is shown in Fig. 10. The charge on ball *A* is to produce a charge of the opposite kind on the conductor *F* through the two condensers *BC* and *DE*. When *A* comes in contact with *B*, this becomes positively charged. A negative charge appears at C due to the inducing action of B. An equal positive charge must appear at D and this must induce a negative charge on E. But if a negative charge appears at E there must be an equal positive charge induced on F. If now the conductor F is connected to the ground this positive charge will run off to earth and there will be left on the conductor EF a negative charge.

Now the ground connection to the conductor F is removed; if then B is grounded (connected to earth), a part of its positive charge will run

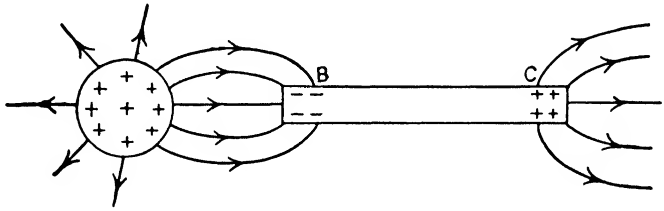


FIG. 9.—A charged body inducing charges on a metal rod.

off, thus releasing part of the negative charge on C to neutralize part of the positive charge in D. But if the positive charge on D is diminished, some of the bound negative charge in E will be released and so will spread out over the conductor EF. Thus a negative charge has been produced on F even though the action had to take place through two perfectly insulated condensers BC and DE. How much negative charge is thus produced on F depends upon the capacity of plate F to ground compared to that of plate B to ground.

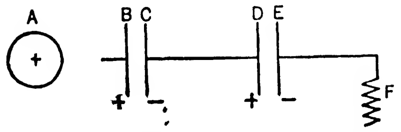


FIG. 10.—A charged body inducing charges on conductor F, acting through two condensers.

Induced Charges from the Electron Viewpoint.—As will be explained later, the electrons in a metallic conductor are more or less free to pass from one atom of the substance to another; they are continually moving around the complex molecular structure of atoms comprising the metal. When the rod of Fig. 9 is brought into the neighborhood of the charged ball the electric field due to the charge on the ball acts on the free electrons of the rod, attracting them. Hence the free electrons of the rod tend to congregate at that end of the rod which is nearest to the ball; they constitute the negative charge at this end of the rod.

But if the rod was uncharged before coming into the influence of the charged ball there must be just enough electrons on it to neutralize the positive charges of the atoms. If more than a proper portion of the electrons gather at one end of the rod there must necessarily be a shortage of them at the other end. *This shortage of electrons at the end C of the rod constitutes the positive charge at this end.*

When the end C is grounded, the positive atoms of the rod cannot leave the rod and go into the earth, but electrons from the earth can run up into the rod and they do so, being attracted by the deficient atoms at C. These electrons from the earth appear in sufficient quantity to make the atoms at C neutral. When the wire connecting the rod to the earth is removed and the charged ball is also removed the rod has on it a free negative charge, the quantity of charge being equal to the number of electrons which came from the earth into the rod.

An Essential Difference between Positive and Negative Charge.—As before stated, the electrons from all substances are the same; the electrons have none of these qualities by which we distinguish and classify matter. It is possible to have electrons in space entirely devoid of matter; a negative charge can exist in a perfect vacuum.

The question may be raised—How can it be a perfect vacuum if there are electrons present? By a vacuum we mean a space in which there is no material substance, solids which can be bodily removed, liquids which can be poured out, or gases which can be pumped out. A glass vessel which has been evacuated as perfectly as modern pumping methods can accomplish may nevertheless be filled with millions of electrons.

From our present conception of matter the positive charge is always associated with the nucleus, a group of positive and negative electrons. When an electron is taken away from an atom of oxygen, for example, what is left of the oxygen still reacts chemically as oxygen. It is, to be sure, charged (or *ionized*) oxygen, but this group of positive and negative electrons still has the properties of oxygen in spite of the fact that one electron has been taken away. But the electron which has been taken away shows none of the characteristics of oxygen whatsoever.

Thus positive charge is always associated with what we call matter; a nucleus with at least part of its ordinary complement of encircling electrons is always present where there is matter. To the same extent, for example, as we can show the existence of positive charge inside a vacuum tube, we know that to just that same extent, the gas which has supposedly been removed by the evacuation process is still present.

The Electric Current.—The electric current is more familiar to everyone than the electric charge. The current manifests itself in various ways, by generating heat and light, by producing mechanical forces such as those required to ring a doorbell or pull a subway train, by producing chemical changes such as occur in the production of aluminum, or electroplating, by producing death if it flows through a living organism with sufficient intensity, etc.

Older conceptions of the electric current made it a peculiar fluid of some kind, others made it consist of two fluids with different properties. From the electron standpoint the conception of the electric current is easy to comprehend and enables one to give a fairly logical explanation of the various actions of the current.

Nature of the Electric Current. *An Electron in Motion Constitutes an Electric Current.*—The amount of electricity on one electron is so small that the current produced by one electron in motion would not be detectable by the finest current-measuring instrument, even the most sensitive. To produce currents of the magnitude occurring in every-day experience requires the motion of electrons measured in billions of billions per second.

An ordinary incandescent lamp requires a current of about 1 ampere; such a current requires that about 10^{19} electrons flow past any point in the circuit each second. This large number per second might be brought about by a comparatively few electrons moving rapidly or by a great many moving more slowly. Contrary to what one would naturally think, the progressive movement of the electrons is very slow. To produce a current of 1 ampere in a copper wire 1 mm. in diameter requires that the average velocity of the electrons be only about 0.001 cm. per second, if we accept the assumption that there are as many free electrons in the copper as there are atoms.

Although the progressive motion of the electrons is very slow, as indicated above, it must not be thought that the actual velocity of the electrons is small. If we assume the " equipartition of energy " idea of thermodynamics and thus calculate the average velocity of the electrons in a copper wire, at ordinary temperature, we obtain a result of about 6×10^6 cm. per second. That is, even when no current is flowing in the wire the electrons have a haphazard motion, due to the thermal agitation of the atoms (or molecules), which gives them, on the average, a velocity of about 35 miles per second.

Now when current flows the required progressive velocity of the electrons is only a fraction of a centimeter per second; with a current so large that the copper wire is heated to the melting-point the velocity of drift of the electrons is less than 1 cm. per second. Thus an accurate concept of the electric current in a conductor shows it to be an inappreciable " drift " of the electrons which have, due to temperature effects, heterogeneous velocities millions of times as great as the velocity of drift.

The reason for the slow progressive motion of the electrons is to be seen in the tremendous number of collisions they have with the molecules of the substance. A given electron, acted upon by the potential gradient in the wire carrying current, accelerates very rapidly and would acquire tremendous velocities if it did not continually collide with the more massive molecules. One reason is that the mean free path of the free electrons in a copper wire is so small that, between successive collisions, the electron falls through a very small potential difference and hence gains a velocity (along the conductor) due to the current, which is extremely small.

It might seem that the electrons would gain considerable net velocity along the conductor, in spite of the numerous collisions with atoms, but such is not the fact. The mass of an atom or molecule is many thousand times as great as that of an electron; in copper, for example the mass of the molecule is about one hundred thousand times as much as that of the electron. When a collision occurs therefore between an electron and a copper atom, the electron rebounds from the atom with practically the same velocity as that with which it approached the atom.

Suppose that we wanted to measure the rate of flow of people past a given point in a large city; the unit of flow might be 100,000 persons per hour. At any time there will be people going in all directions, some uptown, some downtown, and some crosstown. In the morning a million people pass a certain point where the flow is to be ascertained. If 200,000 move in the uptown direction and 800,000 move downtown, the net flow is 600,000 people. If this number of people pass in 1 hour the flow is 6 units downtown. At noon time again a million people pass the same place let us suppose; 400,000 move uptown, 400,000 move downtown and 150,000 move crosstown west and 50,000 move crosstown east. The net flow is now 100,000 people west and if this number pass in one hour the flow is one unit west. Some of the people would be moving rapidly and others going more slowly and some might, at times, be standing still.

The picture suggested by the above traffic analysis probably gives one a reasonable idea of the motion of electrons in a conductor carrying current; it is of course too simple, because of the immense number of electrons in a conductor and the tremendous number of collisions occurring between the electrons. When a conductor is carrying no current the motion of the electrons resembles that of the individuals in a stationary crowd; there is a deal of agitation among the electrons, but they, on the whole, show no progress along the conductor.

Electromotive Force.—Suppose a copper rod, having in itself the heterogeneously moving electrons suggested above, is connected at its two ends to a battery as shown in Fig. 11. The end A of the rod becomes positive with respect to end B and the electrons, instead of moving backwards and forwards to the same extent, progress slowly towards A. When they arrive at A they leave the copper rod, move down the connecting wire, through the battery, through the other connecting wire, and so back to the rod. As long as the circuit remains closed as shown the electrons will continue to move around the circuit, bounding backward, forward, and across the conductor, but on the whole progressing gradually around the circuit; this progression of the electrons constitutes the electric current. The cause of the flow is the battery; it holds one end of the rod positive with respect to the other and so maintains the flow of electrons. The maintenance of this difference of electric pressure (or difference of potential) across the rod is due to chemical changes going on inside the battery.

A piece of apparatus which has the ability to maintain one of its terminals at a higher potential than the other, even though current is allowed to flow through it, is said to develop an *electromotive force.* As sources of electromotive force for the production of currents on a commercial scale we have at present only the battery and the electric generator. The battery depends upon chemical action for maintaining its difference of potential and the generator depends upon the conductors of its armature being driven through the magnetic field produced by its field poles.

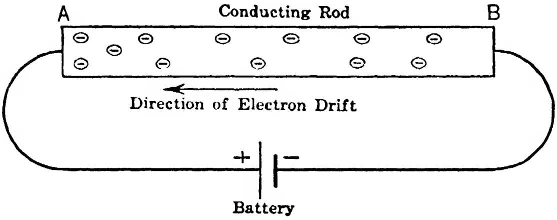


FIG. 11.—Electric current caused by flow of free electrons.

Electromotive Force Due to Thermal Agitation.—We have mentioned previously that the molecules and electrons composing a metal are in a state of continual motion, due to what is called thermal agitation. Unless the metal is at absolute zero temperature this haphazard motion exists, being greater the higher the temperature.

Now if one pictures the electrons in a short piece of wire, let us say, being bumped back and forth by the molecules, it seems reasonable to suppose that at certain times there may be more than the average number of electrons at one end or the other. The idea is simply depicted in Fig. 12. In diagram *a* the electrons are shown uniformly distributed throughout the length of the wire; they are in violent motion (thermal agitation), and the picture shows their distribution at a certain instant. Now an instant later this distribution might be changed to that shown in diagram *b*; extra electrons have accidentally been bumped toward the right end of the wire, so there are now four at the right end. Of course this means that there is probably a deficiency of electrons at the lower end, at this instant.

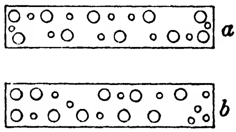


FIG. 12.—Thermal agitation of electrons, resulting in haphazard changes in the electron distribution, generates a voltage in a conductor.

With normal electron distribution (same number of electrons per cubic millimeter as there are positively charged molecules), there is no difference of potential between one part of the wire and another, but with the uneven distribution shown in diagram *b* the right end of the wire is negative with respect to the left end which is positive. In other words, the thermal agitation of the electrons and molecules has resulted in the production of a potential difference between the two ends of the wire.

Now during a succeeding instant it will happen that excess electrons appear at the left end of the wire, thus reversing the potential distribution from what it has in diagram *b*. As the clouds of electrons gyrate back and forth in the wire there will result a varying voltage across its terminals, the *thermal voltage* of the wire. Johnson has investigated this possible source of voltage and reports[1] that there is such an effect and that it is expressible by the relation

$$V^2 = WR, \qquad (1)$$

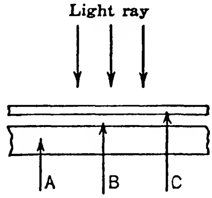


FIG. 13.—Light rays may cause the emission of electrons from certain surfaces; thus light energy generates directly an electromotive force.

in which V represents the integrated value of voltage for all frequencies in the audible spectrum, R is the resistance of the wire, and W is a constant depending upon the temperature and is about 1×10^{-16} watt at room temperature. If the wire has a resistance of say 0.5 megohm, calculation gives a voltage of about 7 microvolts. This voltage has no definite frequency but represents voltages distributed at frequencies from zero to 10,000 cycles per second. This thermal voltage produces the same power output from an audio-frequency amplifier as would a signal of 7 microvolts of some definite frequency in the audible range.

This thermal voltage sometimes results in *noise* from an amplifier designed for high gain; it gives the kind of noise that a radio operator calls "slush." At the present time this voltage is not applied to any useful purpose; on the contrary, it often gives trouble in some forms of apparatus.

Electromotive Force Due to Light.—If light shines through certain translucent substances onto a metallic surface it appears that the impinging light energy can pull electrons out of the metal and thus actually set up a voltage depending for its magnitude upon the intensity of the light. The idea is illustrated conventionally in Fig. 13. The iron disc A is coated with a compound B (iron selenide has been used) which itself is coated with a thin translucent metal film C. A galvanometer connected between disc A and film C will give reading varying with the light intensity. There is no battery or other source of energy in the circuit. The voltage produced is of course small, being generally only a few millivolts. It seems that in this piece of apparatus the light energy is actually changed directly into electrical energy. At this time there are but scant experimental data on the operation of the device; a minute motor has been run from its power, however.

[1] Physical Review, July, 1928.

Electromotive Force and Difference of Potential.—It is well to distinguish between electromotive force and difference of potential. Thus two brass balls, one charged positively and the other negatively, have a difference of potential between them and they will, if connected by a wire, cause a momentary flow of current through the connecting wire; when sufficient electrons have passed from the negatively charged ball to neutralize the positive charge on the other the current will cease. *There is no action taking place which tends to maintain the difference of potential between the two balls;* such a combination does not generate an electromotive force (hereafter abbreviated e.m.f.).

In the case of the battery or generator, however, when the two terminals are connected by a wire a current flows and continues to flow until the battery is worn out or the generator is stopped; such devices develop or generate an e.m.f. These ideas are depicted in Fig. 14.

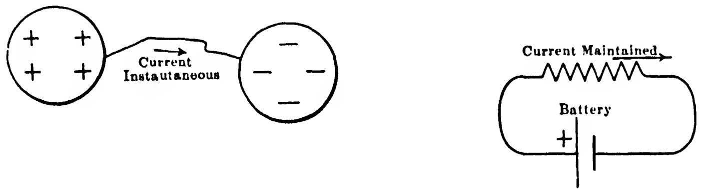


FIG. 14.—Illustrating difference between electromotive force and potential difference.

Direction of Flow of Current.—It has been accepted as convention that in a wire connecting the poles of a battery the current flow is from the positive pole of the battery to the negative. But by reference to Fig. 11 it is evident that in the connecting wire the electrons flow from the negative pole of the battery to the positive. Hence it must be remembered that although we shall talk of the current flowing from the positive terminal to the negative terminal of a battery or generator, the electrons (which really are the current) are flowing in the opposite direction. In dealing with currents through vacua the motion of the electrons themselves is generally had in mind and we often say that the electron current flows from the negative to the positive terminal of the vacuum tube. Although this sounds anomalous it is a correct statement of the facts.

Conductors and Insulators.—Roughly speaking, a conductor is a body which readily permits the passage of an electric current and an insulator is a body which offers a very high resistance to the passage of the current. There is no sharp distinction between conductors and insulators, however; a material which for some cases would be regarded as an insulator would, in other circumstances, be regarded as a conductor. Also a substance which is a good insulator at low temperatures may be a fair conductor at high temperatures.

Glass is the most striking illustration of this change of character with change of temperature; at ordinary temperature it ranks high with the very best insulators, but if it is heated in some way to a red heat it becomes a fair conductor and will permit the passage of enough current to melt itself.

Difference between Conductors and Insulators from the Electron Viewpoint.—When a conductor is carrying an electric current the electrons throughout the substance of the conductor are moving gradually along through the substance of the conductor. Now in a solid body, such as a metallic conductor, the atoms or molecules comprising the substance are practically fixed in position. They are not actually stationary in space at ordinary temperature of course; as a matter of fact the atoms have an irregular to-and-fro motion similar to that of the electron. But *there cannot be a progressive motion of the atoms as there may be of the electrons.* The reason for this is more or less evident. Suppose a copper wire is fastened to the terminals of a battery and that current is flowing as indicated in Fig. 11. The electrons move all the way around the circuit through the wire, connections, solution in the battery, etc.

As the atoms of copper are charged positively after an electron has left them it might seem that as the electrons move from B to A through the wire the atoms would move from A to B, then into and through the battery and so back to the wire. But the atoms are the real substance of the wire, and hence if the atoms should progress one way or the other it would result in the copper itself being carried from one end of the wire to the other and then through the battery. This state of affairs is not possible in solid bodies like metals, it would result in the mixing of metals wherever a current left one metal and went into another.

In chemical solutions, e.g., copper sulphate in water, the salt molecule breaks up into two parts, one of which has one electron more than its proper number, the other part lacking one electron. The two parts of the molecule are called *ions*; the metallic ion (in above case, copper) lacks one electron and so is charged positively. If now a current is passed through such a solution the metallic ion does move through the solution and is carried from the solution to one of the wires by which the current is led into the solution. Here copper itself is transported by the current and we have the process of electroplating.

From what has been said it follows that if the molecules of a body cling to the electrons so tightly that none of them are free to move away from the molecule, there can be no current in such a substance. As long as the molecule keeps all its electrons it remains electrically neutral, and so has no tendency to move when in an electric field. This is the essential difference between insulators and conductors; in the one the electrons cannot move from the atom or molecule and in the other the electrons are perfectly free to leave the atom.

Disruptive Strength of an Insulator.—With the above idea in mind the possibility of breakdown of an insulator, due to high voltage, becomes apparent. For low voltage the force tending to move the electron is not sufficient to break it loose from its atom. But it is reasonable to believe that, if the voltage gradient is made sufficiently high, any atom can be forced to let go of one electron, and such is the case. Such fine insulators as glass and mica break down and carry current when a great enough voltage is employed.

When the molecules of even a good insulator are acted upon by an electric field, there is a motion of the electrons due to this field, even though there is no continuous flow of electrons as in a conductor. When no electric field acts on the molecule, the electrons have a certain normal position with respect to the nucleus as suggested in *a* of Fig. 15, but when the molecule is placed in an electric field the electrons try to move and do so for an instant, until the elastic force of the nucleus tending to pull them back to their normal configuration is just equal to that exerted by the externally applied electric field.

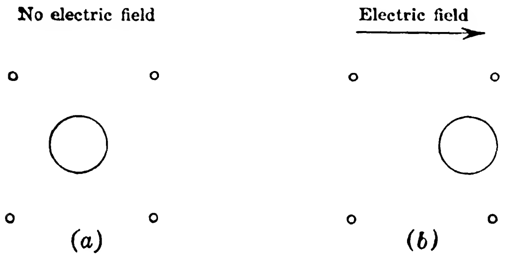


FIG. 15.—In an insulating material the electrons are not free to leave the atom, but an electric field distorts such an atom, causing the electrons to move one way and the internal positive charge the other way.

Thus we may say that the molecule of an insulator, placed in an electric field, is *distorted*, the amount of distortion depending upon the electrical rigidity of the molecule and the intensity of the impressed field. If the impressed field is too strong, one of the electrons pulls free from the molecule, moves through the substance of the insulator, which thus carries current and so is partially "broken down." Actually if only a few of the insulator's molecules do thus release one electron each, the insulator is likely to break down completely at once and become a conductor; the freed electrons bumping into other severely strained molecules knock other electrons loose and this cumulative effect results in complete disintegration of the insulator at the point where the breakdown occurs.

Effect of Temperature on the Disruptive Strength of an Insulator.—Imagine a good insulator heated by some outside source of power. The rise in temperature increases the to-and-fro motion of its molecules with the result that the collisions between the various molecules become more frequent and violent as the temperature is raised. As these collisions occur the resulting disturbances in the molecular structure tend to weaken the hold of the molecule on its electrons. Hence if an electric force impressed and maintained as an insulator is heated, the combination of electric force and weakening of the molecular holding power will result in some electrons leaving their molecules; the electric force then urges them along through the substance of the insulator with the result that a small current flows. This would be interpreted by the man testing the insulator as a weakening of the insulating power of the substance.

Generally the partial breakdown of an insulator as described above is rapidly followed by the giving away of the insulator completely; as current, even though small, flows through the insulator it generates more heat, thus still further decreasing the disruptive strength.

This effect of temperature upon the disruptive strength of an insulator is very important to the radio engineer. A glass or mica condenser, properly designed to operate in a radio circuit at 15,000 volts may, by improper use, be broken down when operating at only 5000 volts. Condensers heat up, when being used, due to various causes; in normal operation the condenser may be excited only a small fraction of the time as the sending key is opened and closed. In the intervals when the key is open the cause of the heating is removed and the condenser has a chance to cool off; this alternate heating and cooling results in a certain mean temperature at which temperature the condenser has sufficient disruptive strength to withstand the voltage employed.

If now the normal operating voltage is put on the condenser and maintained continuously, the heating action is much greater than when the voltage is applied intermittently (normal operation) and in a short time the dielectric is likely to puncture. Condensers which are designed for operation at a certain voltage with spark telegraphy (intermittent excitation) will nearly always fail if operated at the same voltage for undamped wave signaling (continuous excitation).

Resistance.—In a conductor where the electrons are free to leave the atom their progressive motion is hindered by collisions with the atoms of the substance. *This hindrance to their free progress constitutes the electrical resistance of the conductor.* It differs, as might well be expected, in different metals, and it varies with the temperature. As the temperature of a metal increases, the agitation of its atoms or molecules increases, and this results in more hindrance to the progressive motion of the electrons because of the more frequent collisions between the electrons and the atoms.

The increase in number of collisions between the electrons and atoms with increase in the flow of electrons (more current) gives the atoms themselves an increased agitation, which really means a higher temperature; this accounts for the well-known fact that when a conductor carries current it always heats to some extent and heats more with large than with small currents.

Continuous Current and Alternating Current.—If the electrons in a conductor continually progress in the same direction, the flow is called a *continuous current*, or *direct current*. Such is the current supplied by an ordinary battery.

If the battery is connected to the conductor first in one direction and then in the reversed direction, by some sort of a commutator, Fig. 16, the progressive motion of the electrons will reverse with every reversal of the battery connection. If this reversal of flow takes place at regular, short periods of time the alternate ebb and flow of the electrons constitute an *alternating current*. In ordinary power circuits supplied with alternating current this reversal takes place about 100 times per second; the alternating

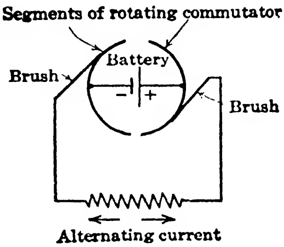


FIG. 16.—A battery in combination with a rotating commutator may produce an alternating current.

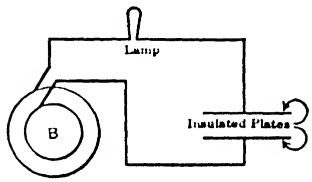


FIG. 17.—The lamp may burn even though there is a perfect insulator in series with the circuit.

currents used in radio circuits reverse much more rapidly, perhaps several million times per second.

Possibility of Alternating Current Flowing in a Circuit in Series with Which There is a Perfect Insulator.—Suppose a circuit connected as indicated in Fig. 17; B is a source of alternating e.m.f. and A consists of two metal plates separated by paraffined paper or mica. The disruptive strength of the insulator is such that for any voltage that B can give the insulation is perfect. A small incandescent lamp is inserted in the circuit to detect the current which may be flowing. The lamp will burn as soon as machine B is excited. Now if the lamp and *condenser* (the combination of two conducting plates and separating insulator) is connected to a battery which gives about the same voltage as machine B gives, the lamp will not burn, showing that there is no current in the circuit. Hence this circuit which is open for continuous current (i.e., it will not pass current) does permit the flow of alternating current.

The alternating current is possible because of the number of electrons required to *charge the condenser*. As the voltage of the alternator reverses in direction the condenser charges first in one direction and then in the other; this alternating charge and discharge requires the alternating flow of electrons throughout the whole circuit.

This flow, it is to be noted, is made possible by the distortion of the molecules of the insulator in the electric field between the two plates of the condenser. If the molecules of the insulator were perfectly rigid, so that the effect noted in Fig. 15 could not take place, the insulator would permit no alternating current to flow.

Of course we know that even though the insulated plates, constituting the condenser, are placed in a vacuum (so that there are no molecules of insulator between them), some current would still flow through the lamp of Fig. 17. The flow would be only a fractional part of what it would be if glass, mica, etc., were used, but still there would be some current. This current is really due to the compressibility of the electrons in the conductor of which the circuit is composed; when the electromotive force of the alternator reverses direction, a slight to-and-fro surging of the electrons in the circuit conductor occurs, even though there is no possible movement of electrons in the space between the condenser plates.

In such a circuit as that of Fig. 17 the average to-and-fro motion of the electrons in the circuit conductor is to be measured in *millionths of one centimeter or less*, depending upon the frequency of the current and capacity of the conductor used.

A simple analogy to the flow of alternating current in the circuit of Fig. 17 is shown in Fig. 18. Suppose a cylindrical chamber A, divided in the middle by a thin rubber diaphragm B, connected to a reciprocating action, valveless, pump C. As the pump works back and forth, water will circulate back and forth in the connecting pipes, constituting an a.c. flow of water. The diaphragm B will bend first in one direction and then in the other as the water reverses its flow.

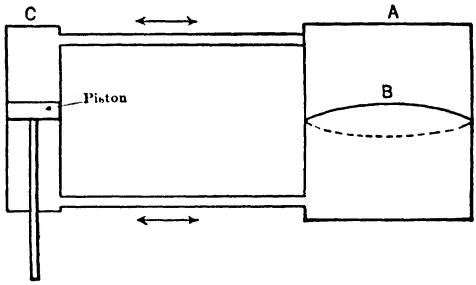


FIG. 18.—Hydraulic analogue of an alternating-current circuit containing a condenser.

Now suppose that a centrifugal-action pump be substituted for the reciprocating pump (Fig. 19). This type of pump tends to force water always in the same direction. If the pump is so connected as to force water into the bottom of A and suck it out of the top of A, the flow of water will last long enough to stretch the diaphragm into some such position as B', and then the flow will cease. At this position of the diaphragm the backward pressure of the stretched rubber will be just great enough to balance the pressure generated by the pump. In this water system the water would flow while the diaphragm was being displaced from its normal central position to position B', and then the flow would cease because the pump would not be able to further displace the diaphragm.

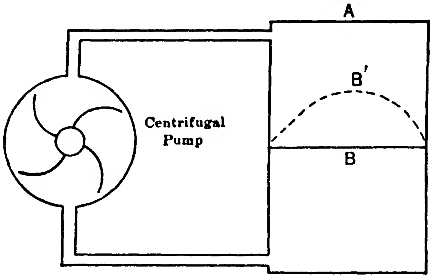


FIG. 19.—Hydraulic analogue of a direct-current circuit containing a condenser.

The water system corresponds very closely to the electrical circuit having a condenser in series with it and excited by a continuous e.m.f.; in such a circuit the current flows long enough to charge the condenser to such an extent that its back pressure (pressure tending to discharge the condenser) is just equal to the impressed e.m.f. and then the current ceases. It is to be remembered, however, that if the pressure is alternating there will be a flow in the system all the time, the current being an alternating one.

In electric circuits, therefore, it is possible to send an alternating current through a circuit in which continuous current cannot flow. Such use of a condenser occurs frequently in radio; the condenser so used is called a stopping or " blocking " condenser.

The Electric Generator.—Except for very small sets and emergency outfits the power for a radio set is obtained from a generator of either the continuous or alternating-current type. The c.c. generator is equipped with a commutator and supplies a continuous e.m.f.; that is, the e.m.f. impressed on the connected circuit is always in the same direction and practically constant in value. There are slight pulsations in the value of the voltage, perhaps a fraction of 1 per cent, at the frequency of commutation; this frequency is in the neighborhood of 1000 cycles per second. Although these pulsations are so small, they have a deal of importance in certain radio sets using vacuum tubes for the generation of high-frequency currents.

The a.c. generator (or simply *alternator*) has no commutator, but generally has slip rings on which its brushes make contact. The e.m.f. furnished by such a machine alternates in direction many times per second; for spark radio sets the generators ordinarily employed give several hundred complete reversals of voltage per second.

The number of complete reversals per second is called the *frequency* of the generator; thus a 500-cycle generator is one giving 500 complete reversals of e.m.f. per second.

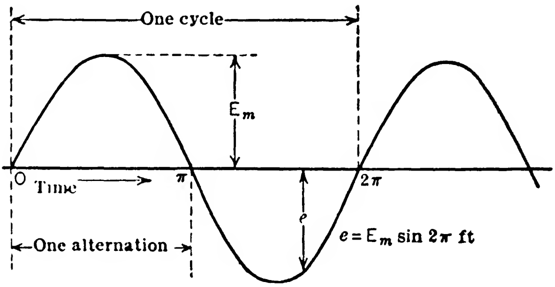
Wave Shape and Effective Values.—The form of voltage wave generated by a well-designed alternator is such that it can be closely represented by a sine curve as shown in Fig. 20. Expressed in the form of an equation,

$$e = E_m \sin \omega t, \quad \ldots \ldots \ldots \quad (2)$$

where

e = the value of voltage at any instant of time;
E_m = the maximum value of the voltage generated;
$\omega = 2\pi f$, f being the frequency of the voltage.

The same units are used for measuring alternating voltage and current as are used for continuous voltage and current. But as the voltage and current of an a.c. circuit are continually changing in value and reversing in direction, some value intermediate to the maximum and minimum value must be chosen as the unit. It is shown in all elementary texts on alternating currents that when the current flows according to the law of a sine curve *the alternating current will produce heat at the same rate as 1 ampere continuous current if the maximum value of the alternating current is 1.41 amperes.*



Form of alternating emf.

FIG. 20.—Sine wave of e.m.f.

To get the value of that continuous current which will give the same heating effect as a certain alternating current, therefore, we take 0.707 of the maximum value of the alternating current. That value of continuous current which will produce the same heating effect as the alternating current in question is called the *effective* value of the alternating current. It is approximately 0.7 of the maximum value.

In the same way the effective value of an alternating e.m.f. (sine wave shape assumed) is 0.707 of its maximum value. Thus, if a sine wave of voltage has a maximum value of 141 volts its effective value (or equivalent continuous voltage as far as producing the current, which causes heating, is concerned) is 100 volts.

Magnetic Field.—The action of the magnet is familiar to everyone. If a piece of iron is placed in the vicinity of the magnet a force of attraction is set up between the two and the piece of iron will, if free to move, be drawn to the magnet.

All the region surrounding a magnet, in which the magnet is able to exert a force on pieces of magnetic material, is said to be filled with the field of the magnet. Thus the magnetic field is exactly analogous to the electric field surrounding an electrically charged body.

The magnetic field is represented by lines in just the same way as the electric field; the direction of the lines indicates the way in which the north pole of a compass would be urged if placed at that point of the field, and the proximity of the lines to each other serves to show the relative intensity of the magnetic force at various points of the field.

Magnetic Field Set Up by an Electric Current.—The field of the permanent steel magnet is interesting historically, but it plays very little part in the electrical engineering of to-day. When an electric current flows through a conductor a magnetic field is set up around that conductor; such a field is frequently called an *electromagnetic field.* The magnetic fields used in modern apparatus are practically all of this type.

The strength of a magnetic field set up by an electric current depends upon the strength of the current, in general being directly proportional to the current strength. The direct proportionality holds good for magnetic fields without iron; use of iron in the magnetic circuit makes the relation between current and strength of field a complex one.

Ampere-Turns.—When the magnetic field is produced by a coil of several turns, its intensity is much greater than if only one turn were used. The magnetizing effect of a current depends not only on the strength of current, but also on the number of turns through which the current flows. In fact the magnetizing effect of a coil is proportional to the product of the current strength and the number of turns in the coil; this product is called the *ampere-turns* of the coil. If a coil consists of one turn and is carrying a current of one ampere it has one ampere-turn; a coil of twenty turns carrying 2.7 amperes has fifty-four ampere-turns.

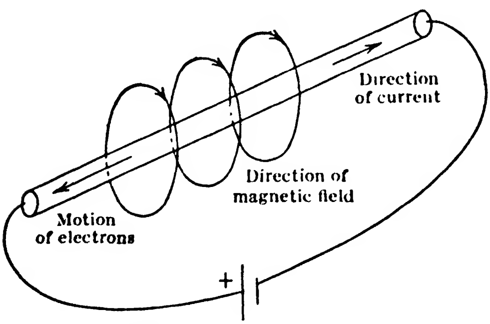


FIG. 21.—Direction of the magnetic field produced by a current.

Direction of the Magnetic Field Produced by a Current.—The direction of magnetic field around a conductor carrying a current may be easily determined by the application of the following rule. Imagine the conductor grasped in the right hand, fingers around the conductor, with the extended thumb pointing along the conductor in the direction in which the current is flowing; the fingers then point in the direction of the magnetic field. This is illustrated in Fig. 21. It is to be remembered that this rule assumes the commonly accepted direction of flow of current; in Fig. 21 the electrons are flowing in the opposite direction to that marked current.

It follows at once from the rule given above the direction of magnetic field reverses when the current reverses. If an alternating current is passed through a wire or coil the magnetic field produced will also be an alternating one, having the same frequency as the current, and *reversing simultaneously with the current.*[1]

Iron in the Magnetic Field.—In most electrical apparatus depending on the magnetic field for its operation the field is produced by currents flowing in coils. But the coils are usually fitted with iron cores so that the magnetic circuit consists partly of iron and partly of air. The reason for the use of iron in magnetic fields of electrical devices lies in its high permeability, i.e., the relatively high flux density produced by a given coil in iron compared to what it would produce if only air were used in the magnetic circuit.

The magnetic permeability of a substance is the ratio of the flux density produced in this substance by a certain magnetomotive force (magnetizing force) compared to the flux density the same magnetomotive force would produce in air.

For most substances the permeability has a value of unity; nickel, cobalt, and iron are the notable exceptions. Of these three iron is by all means the most important, not alone because of its comparative cheapness (and hence utility for electrical apparatus), but because of the high value of the permeability. For good magnetic iron it may be as high as several thousand; that is, if a given coil produces 500 lines of flux with a magnetic path of air, it will produce perhaps a million lines of flux if iron is used for the whole magnetic circuit.

When the magnetic circuit of a device is made up partly of air and partly of iron, the flux produced by a given coil is intermediate to that which would be produced in a complete iron path, and that which would be produced in a complete air path. The shorter the part of the path through air compared to that through iron the higher will be the flux set up. The permeability of iron varies greatly with the treatment it received during manufacture; also for a given specimen it varies greatly with the magnetizing force used. This point will be taken up in more detail in the next chapter, under the head of *self-induction* and its variations.

Units of Current, E.M.F., Resistance, etc.—The unit of current is the ampere; it is that flow of electrons which will deposit 1.118 milligrams of silver per second from a silver nitrate solution in a standard voltameter.

The unit of e.m.f. is the volt; it is generally defined in terms of the

[1] This statement is strictly accurate only for the magnetic field in the immediate neighborhood of the conductor; for more distant points the magnetic field reverses somewhat later than the current. This idea is taken up in more detail in Chap. IX.

voltage of a standard Weston cell, which gives an e.m.f. of 1.0183 volts. The volt is therefore defined as 1.0000/1.0183 of the voltage generated by a standard Weston cell.

The unit of resistance is the ohm; it is really defined already [1] when the ampere and the volt have been defined because the three units are directly connected by Ohm's law. However, it is also defined as the resistance of a column of pure mercury weighing 14.4521 grams at 0° Centigrade and having a height of 106.3 cm., the cross-section being uniform.

The unit of quantity is the coulomb; it is the quantity of electricity transported by a current of 1 ampere flowing for 1 second. Another way of defining it is in terms of electrons; it is the amount of electricity contained on 6.28×10^{18} electrons.

The unit of work is the joule; [2] it is the amount of energy required to transport 1 coulomb of electricity through an opposing potential difference of 1 volt. It is also the amount of work done in 1 second by a current of 1 ampere flowing against a pressure of 1 volt.

The unit of power is the watt; it is the rate at which work is done by a current of 1 ampere flowing against a pressure of 1 volt. It is therefore a rate of work equal to 1 joule per second.

Resistance of a Conductor.—The resistance of a circuit depends first of all on the kind of material used in making up the circuit; it depends upon the length of conductor used in making the circuit and upon the cross-sectional area of the conductor. This relation may be expressed by the equation,

$$R = \rho l/a, \qquad (3)$$

where ρ = the specific resistance of the material used;
l = the length of the conductor;
a = the cross-sectional area of the conductor.

When the length of the conductor is 1 cm. and the area is 1 sq. cm. the value of R is the specific resistance per centimeter cubed, and when the conductor has a length of 1 foot and an area of 1 circular mil the value of R is the specific resistance per mil-foot. In engineering the latter specification is more frequently used.

The specific resistance of some of the more common conductors is given in the accompanying table:

[1] The ohm and the ampere are really the two fundamental units of the practical system. See "American Handbook for Electrical Engineers."

[2] For small quantities of energy Rutherford has suggested the "electron-volt" as a suitable unit; it is the amount of work done in carrying one electron through an opposing potential difference of 1 volt.

SPECIFIC RESISTANCE OF COMMON METALS, ALLOYS AND SOLUTIONS

Substance	Composition	Microhms per cm³. at 0° C.	Temperature Coefficient Referred to 0° C.
Advance	Copper-nickel	48.8	0.000018
Aluminum	Pure	2.62	.00423
Brass	66 Cu+34Zn	6.29	.00158
Calido	Ni+Cr+Fe	100.0	.00034
Carbon	Lamp filament	4000.0	− .0003
Constantan	60Cu+40Ni	49.0	.00000
Copper	Standard	1.589	.00427
Copper	Electrolytic	1.56	.00428
German Silver	18Ni+Cu+Zn	33.1	.00031
Ia Ia soft	Cu+Ni	47.1	.00000
Iron	Pure	8.85	.00625
Iron	Hard steel	45.0	.00161
Manganin	Cu+Mn+Ni	{40.0 70.0	{.00001 .00004
Nickel	Electrolytic	6.93	.00618
Silver	Electrolytic	1.47	.00400
Tungsten	Hard	5.42	.0051
	Per Cent Solution	Ohms per Cm³.	
H_2SO_4	5	4.80	−0.012
	10	2.55	.013
	20	1.50	.014
	30	1.35	.016
	50	1.85	.019
	70	4.68	.026
KOH	20	2.01	− .020
HCl	20	1.31	− .015
HNO_3	20	1.41	− .014
NaCl	2	37.0	− .023
	5	14.9	.022
	10	8.2	.021
	20	5.1	.022

Practically all solutions have a minimum resistance with a density of solution between 20 and 30 per cent.

The resistance of a metal varies with the temperature, in general being directly proportional to the absolute temperature. This relation is approximately expressed for all *pure* metals by the equation,

$$R_t = R_0(1+at), \qquad (4)$$

where R_t = the resistance at t degrees Centigrade;
R_0 = the resistance at zero degrees Centigrade;
t = the temperature at which the resistance is desired;
a = the temperature coefficient of resistance.

When absolute zero temperature is approached the resistance changes no longer follow the simple law given in Eq. (4); the changes become more irregular.

The value of a is very nearly 0.004 for all pure metals, for copper it has been decided to take a as 0.00427, at 0° C.

A statement which gives the above rule in words is as follows—the resistance of a pure metal increases approximately 1 per cent for each 2.5° rise in temperature above 0° C.

The resistance of a field coil of a generator which has a resistance of 25 ohms at ordinary temperature might have a resistance of 30 ohms after the machine had been operating a few hours; the rise would be due to the heating of the coil. A tungsten lamp has when hot about nine times as much resistance as it has at room temperature.

In certain materials the resistance may show considerable departure from the rule given above, thus in carbon an increase in temperature brings about a decrease in resistance. In a certain alloy of nickel and copper there is practically no change in resistance with ordinary temperature changes.

There are very strange resistance changes in certain substances, e.g., a large change in resistance takes place in selenium, according to the amount of illumination it receives; bismuth shows a large change in resistance, as it is introduced into a magnetic field and is sometimes used to measure the strength of magnetic field by the determination of its resistance.

The resistance of a salt or acid solution such as we have in primary or secondary batteries depends among other things upon the strength of the solution. This variation does not follow a simple law; there is a certain strength of solution which gives minimum resistance. For sulphuric acid solution such as is used in lead storage batteries the mixture which gives minimum resistance is made up with 30 per cent (by weight) acid.

The effect of temperature on the resistance of electrolytes is to give a decrease of resistance with an increase of temperature for those changes in temperature to which they are ordinarily subjected. The resistance decrease is about 2 per cent per degree Centigrade.

In case a circuit is carrying an alternating current the resistance may show all sorts of variations; it may be changed by bringing a piece of iron, or another circuit, into its magnetic field, by varying the frequency or strength of current. These changes of resistance in so far as they have importance in radio work will be considered in the next chapter.

Induced Electromotive Force.—When current passes through a coil of wire it sets up a magnetic field in the coil and the strength of this field varies as the current varies. Now it is a fundamental law of the electric circuit that when the strength of magnetic field through a coil is varied, an *e.m.f. is induced in the coil*; this law, which was discovered by Faraday, is called the law of induced e.m.f. *The application of this law underlies the design and operation of nearly all electrical machinery and circuits.*

Magnitude of Induced E.M.F.—The magnitude of the induced voltage depends upon the rapidity with which the magnetic field is changing and upon the number of turns in the coil, it being directly proportional to each of these factors. It is written

$$e = -N \times \frac{d\phi}{dt}, \quad \ldots \ldots \ldots \ldots \quad (5)$$

in which e = the voltage induced at any instant of time, in abvolts;
N = the number of turns in the coil;
ϕ = the flux through the coil.

The minus sign is necessary because of the relation between the direction of the induced e.m.f. and the change in magnetic field, i.e., increase or decrease.

Direction of Induced E.M.F.—The change of flux is of course produced by a change of current; if the flux is decreasing it must be that the current in the coil is decreasing. *The direction of the induced e.m.f. is always such as to prevent the change of current which is producing the induced voltage.* Hence when the current (or flux) is decreasing, the direction of the induced e.m.f. is such as to prevent the decrease of current.

Suppose a circuit arranged as shown in Fig. 22; A is the battery, B is a coil, C is a switch, across which is connected a resistance D. With the switch closed current will flow in the direction of the arrow and will be fixed in magnitude by the voltage of the battery and the resistance of the coil. The resistance D will play no part in fixing the value of the current, because with the switch closed this resistance is cut out of the circuit, or short-circuited.

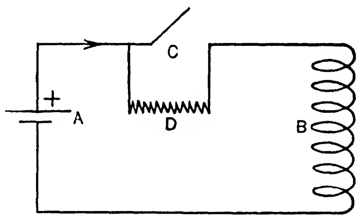


FIG. 22.—Opening the switch will reduce the current in the circuit.

A certain flux ϕ will be set up in the coil, the value of this flux being fixed by the current. If now the switch is opened the current must change to some lower value because of the added resistance D. This lower current will produce a lower flux ϕ_2. While the flux is changing from ϕ_1 to ϕ_2 an e.m.f. will be set up in the coil B and the direction of the e.m.f. will be the same as the battery e.m.f., i.e., it will assist the battery e.m.f. in tending to maintain the current at its original value.

In Fig. 23 the switch is supposed to be closed until time A and here it is opened. The flux will decrease from the value AE to BF, the time taken for the change being that shown on the diagram between A and B.

The decreasing flux generates a voltage in the coil shown by the curved line *AIB*, and this is in the same direction as the battery voltage, hence the total voltage acting in the circuit during the time *A-B* is shown by the curved line *GJH*.

When the switch is closed again at time *C* the flux increases from ϕ_2 to ϕ_1; the induced voltage is now in the opposite direction and is shown by the curved line *CKD*; it results in a total circuit voltage less than the battery voltage, as shown by the curved line *MNO*. (The shape of the induced voltage will not be exactly that shown by the lines of Fig. 23; these curves are only approximate indications of the actual form of the induced voltage. The exact form will depend upon the sparking taking place at the switch, etc.)

Summarizing the facts brought out by Fig. 23 and its explanation we have the proposition that when the current in an inductive circuit is

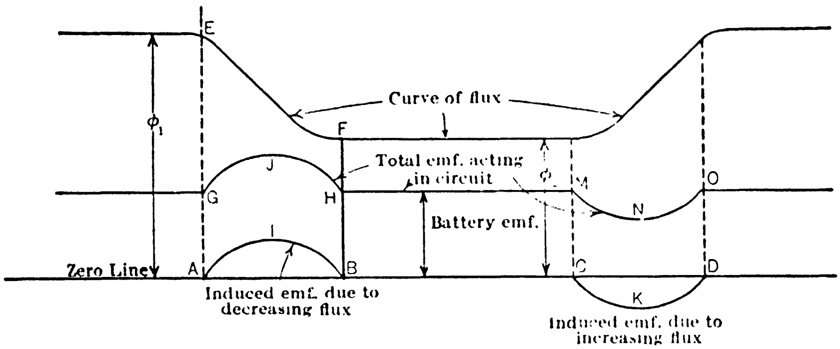


FIG. 23.—Curves showing direction of induced e.m.f.s when current is increasing and when decreasing. No attempt has been made to give the curves of current and flux changes their correct shapes. Later in this chapter the shape of current and flux changes will be discussed more in detail.

decreasing the induced voltage acts to increase the total voltage of the circuit; when the current is increasing the induced voltage is in such a direction that the total voltage acting in the circuit is decreased.

Illustrating the above ideas there is a certain circuit used in radio in which a continuous voltage of 1200 volts is applied through a coil to the plate of a vacuum tube; as the current in this circuit pulsates, alternately increasing and decreasing from its normal value, the induced voltage in the coil has a maximum value of 1100 volts. When the current is increasing this induced voltage acts in the opposite direction to that of the generator furnishing the 1200 volts, so that the total voltage effective in maintaining current through the resistance of the circuit is only 100 volts.[1] When the current is decreasing the induced voltage assists

[1] The reason for this statement regarding the magnitude of voltage in the coil will

the generator voltage and the total effective voltage in the circuit is 2300 volts. The effect of induced voltage in this special circuit is to produce a pulsating voltage, between 100 volts and 2300 volts, although there is in the circuit a generator to supply the current, which furnishes a continuous voltage of 1200 volts.

This voltage set up in a coil by the changing flux in the coil (the flux being caused by current in the coil itself) is called the *e.m.f. of self-induction.*

Coefficient of Self-induction.—Instead of expressing the magnitude of the induced voltage in a coil in the form given by Eq. (5) we may write

$$e = -L\frac{di}{dt}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

in which i = the current in the coil;

e = the instantaneous value of the induced voltage, due to the changing current, i,

L = *the coefficient of self-induction.*

The coefficient of self-induction of a coil varies with the square [1] of the number of turns in the coil and inversely as the reluctance of its magnetic circuit.

If a given air-core coil has an L of two units and the number of turns is doubled, the value of L is increased four times so it becomes eight units. If the magnetic circuit is changed from air to iron, the permeability of which is 1500, the L will be further increased 1500 times and so becomes 12,000 units. This increase is due to the iron decreasing the reluctance of the magnetic circuit 1500 times. If the iron core does not completely close the magnetic circuit, so that part of the magnetic path is still through air, the value of L is not increased to the extent stated above. For example, if the path through iron is 15 inches and the air part of the path is 0.01 inch long, then the value of L is increased 750 times, instead of 1500 times as stated.

The great increase in L produced by the use of iron for the magnetic circuit explains the almost universal use of iron cores (completely closed when possible) in coils which perform their function owing to the value of their self-induction.

It is to be especially noted that Eq. (5) is the real equation for giving the amount of e.m.f. induced in a circuit, and not Eq. (6). It is true that

not be apparent until the reader has studied the action of the vacuum tube, as given in Chapter VI.

[1] This law holds good for any shape of coil if the magnetic circuit is a closed iron core the permeability of which is constant; this requires that as the number of turns on the coil is increased the current through the coil is correspondingly diminished, so that the flux density remains constant. For an air-core coil the law is approximate only; it is more nearly true as the turns of the coil are placed closer together.

Eq. (6) is used more generally than is Eq. (5), but nevertheless Eq. (6) may sometimes give most ambiguous results, whereas Eq. (5) is always correct.

If a circuit has no iron to carry its magnetic flux, both Eqs. (5) and (6) always give correct results, but when the coil under consideration has an iron core, with accompanying variable permeability, Eq. (6) cannot be used in getting the induced e.m.f. This point is very important in connection with the use of iron-core coils in radio apparatus.

We frequently want to know how a coil will impede a certain change in current; thus in filters, iron-core coils are used to prevent fluctuations in a unidirectional current. Now whereas the coil may have a high self-induction for the current flowing through it, *it may have practically no self-induction for preventing changes in that current.* We might, for example, have a coil which with 4 amperes shows a self-induction of one henry and with 5 amperes shows only 0.8 henry. One would naturally conclude that if the current changes quickly from 4 to 5 amperes the coil would show for this change in current an average self-induction of 0.9 henry. Actually the coil may have practically no self-induction in so far as this current change occurs. This point is taken up in detail in Chapter II.

The unit of self-induction is defined by Eq. (6); if a rate of current change of 1 ampere per second gives an induced voltage of 1 volt, the coil has a self-induction of 1 unit. This unit is called the *henry*; the henry is, however, too large a unit for most of the coils used in radio work, so that subdivisions of the henry are used. The milli-henry is one thousandth of a henry and the micro-henry is the millionth part of a henry. Sometimes a still smaller unit is used, the centimeter, which is the billionth part of the henry. It may seem strange that the unit of length is also the unit of self-induction, but such is the fact; the derivation of the dimensions of the various units is outside the scope of this text. The coils used in " tuning " radio circuits vary from a few micro-henries to several millihenries, according to the frequency of the current being used.

Energy Stored in a Magnetic Field.—It requires work to set up a magnetic field just the same as it requires work to set into motion a heavy body. The greater the self-induction of a coil the greater is the work required to start current flowing in the coil; similarly the greater the mass of a body the greater is the work required to start it in motion.

The amount of work required to give a mass m, a velocity v, is measured by $\frac{1}{2}mv^2$, as shown in all texts on mechanics.

The amount of work required to set up, in a coil of self-induction L, the magnetic field caused by a current I, is,

$$\text{Energy, or work} = \tfrac{1}{2}LI^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where L is measured in henries and I is measured in amperes and the energy is measured in joules.

The field coil of a large generator may have many joules of energy stored in its magnetic field; in radio circuits the amount of energy in the coils of a transmitting set is variable because the current is variable. The maximum value of the energy in the coils of the ordinary transmitter is about one joule per kilowatt capacity of the set.

Mutual Induction.—When the flux through a coil varies an e.m.f. is set up in it; if the flux is produced by current in the coil itself the e.m.f. is spoken of as the e.m.f. of self-induction, but if the flux is due to some other coil, in proximity to the one in which the voltage is being induced, the e.m.f. is spoken of as the *e.m.f. of mutual induction.* The voltage induced in the second coil is proportional to the rate of current change in the first coil (the one producing the flux) and the mutual induction of the two coils. The relation is expressed in the form of an equation

$$e_2 = -M\frac{di_1}{dt}, \qquad (8)$$

where e_2 = voltage induced in the second coil;
i_1 = the current in the first coil;
M = *the coefficient of mutual induction of the two coils.*

It is to be noticed that Eq. (8) is written in the same form as Eq. (6), the minus sign in Eq. (6) signifying that if the current in the circuit is diminished (di/dt negative) the e.m.f. of self-induction thereby set up will be in such a direction as to prevent the decrease in current. The e.m.f. of self-induction always has this phase, tending to prevent whatever change of current is taking place. The use of the minus sign in Eq. (8), however, is not justified from such a consideration, because the voltage e_2 may be of the same sign as di/dt or of opposite sign, depending upon how the two coils are wound.

In the ordinary radio coupler, for example, the voltage induced in the movable coil may have its phase reversed by rotating the movable coil sufficiently; from this standpoint M may be considered as having *polarity* and is so considered in discussing the action of vacuum-tube circuits. Thus we may say M is positive if e_2 has a phase opposite to that of di_1/dt.

This e.m.f., due to mutual induction, is set up whether the second circuit (in which it is induced) is open or closed. If the second circuit is a closed one, then current will flow in it, due to the induced voltage, and the direction of this current flow serves to justify the use of the negative sign of Eq. (8).

If the coil of the second circuit is closed (its two terminals being connected together), current will flow whenever the flux through the coil is changing. Thus, suppose the flux in the second coil, set up by current flowing in the first coil, is increasing. Then the current which flows in the second coil will be in such a direction that the coil's magnetomotive force will act *against the magnetic field* and hence the magnetomotive force of the second coil will tend to prevent the increase of flux through the second coil.

In the same way if the flux through the second coil is decreasing the current set up in this coil by the e.m.f. of mutual induction will be in such a direction as to increase the flux. That is, the current set up in the second coil again tends to make the flux through the coil remain constant. Of course this tendency of the second coil to hold the flux through itself constant cannot actually accomplish this result. The secondary current will flow only as long as there is voltage induced in the second circuit, and there will be a voltage in the circuit *only if the flux through the coil is changing.*

If the e and i of Eq. (8) are measured in volts and amperes respectively, then M is measured in henries, the same unit as is used for L. For smaller values of M the same fractional parts of the henry are used as are used for L. M depends for its value upon the number of turns in each of the coils and upon their relative position; as the number of turns in either coil is decreased the value of M is correspondingly decreased and as the distance between the two coils is increased the value of M is again decreased. M may also be decreased by properly orienting the two coils with respect to one another. Imagine two cylindrical coils, shown in plan as rectangles in Fig. 24; M will have a relatively high value for the position shown in *a* and will have a smaller value for either position *b* or position *c*. The scheme of rotating one of the two coils to diminish M has the advantage over the other method that it is compact and so permits the design of a set to be kept to smaller dimensions, a very important point if the sets are to be portable.

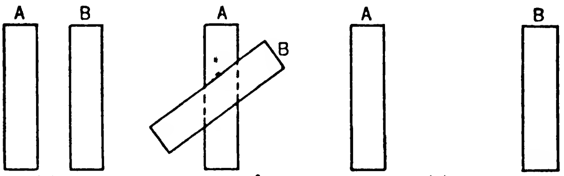


FIG. 24.—Variation of mutual inductance between two coils.

Coefficient of Coupling.—If all the flux produced by one coil threads with all the turns of the other, the coils are said to have 100 per cent coupling; if but a small fraction of the flux produced by the first coil threads the turns of the second, the coupling is weak. Also if all the flux of the first coil links with but a few turns of the second, the coupling is again weak.

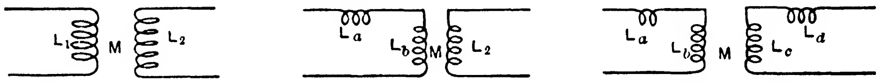
The coefficient of coupling [1] between the two circuits is given by the relation

$$k=\frac{M}{\sqrt{L_1L_2}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

where k = coefficient of coupling, always less than unity;
M = mutual induction between the two circuits;
L_1 = the total self-induction of the first circuit;
L_2 = the total self-induction of the second circuit.

M, L_1, and L_2 must all be expressed in the same units.

As examples of the proper use of Eq. (9) in determining k, Fig. 25 is given; it is to be especially noted that if there are two or more coils in series and only one of them is used to couple the two circuits the total L of the circuit must be used and not the L only of that coil used for the



L_1=.08 L_2=2.0 M =0.1

$K=\frac{0.1}{\sqrt{.08 \times 2.0}}=0.25$

L_a=.06 L_b=.08

L_2=2.0 M =0.1

$K=\frac{0.1}{\sqrt{(.06+.08)\times 2.0}}=.189$

L_a=4.0 L_b=2.4

L_c=16 L_d=20 M=5.0

$K\ \frac{5.0}{\sqrt{(4.0+2.4)(16+20)}}=.329$

FIG. 25.—Examples of coefficient of coupling.

coupling. Thus if two circuits are coupled to a certain extent by two coils in a certain position with respect to one another and another coil is added in series with one of these, *leaving the two original coils exactly as they were*, the coefficient of coupling of the two *circuits* has been lessened.

Mutual induction is, as the name indicates, a mutual characteristic of the two circuits, and has the same value whichever of the two circuits is considered as the primary, or inducing, circuit. This is more or less evident if the two coils have the same number of turns, but possibly not so apparent when the two coils have different numbers of turns. Thus in the first diagram of Fig. 25 the mutual induction is 0.1 henry whichever coil is considered as the primary. If, for example, the current in the 0.08 henry coil is changing at the rate of 1000 amperes per second, the voltage induced in the L_2 coil is $1000\times 0.1=100$ volts. But also if the current in the 2.0 henry coil is changing at the rate of 1000 amperes per second, the voltage induced in the L_1 coil is also $1000\times 0.1=100$ volts.

[1] A more detailed discussion of coefficient of coupling is given on p. 107.

Practical Uses of Mutual Induction.—In general, whenever energy is to be transferred from one circuit to another, insulated from the first, the transfer occurs across either a mutual electric or magnetic field, and generally this transfer utilizes a mutual magnetic field. That is, the energy flows from one circuit to the other because of the mutual induction of the two circuits. In a radio spark transmitter mutual induction is used between the two coils of the power transformer where the coupling is above 90 per cent; in the high-frequency oscillation transformer the coupling is about 20 per cent; in the coupler of the receiving set the antenna is coupled to the local tuned circuit with a coupling of perhaps 2 to 10 per cent.

Effect of a Short-circuited Coil on the Self-induction of a Neighboring Coil.—Suppose a coil A has a certain self-induction by itself; it will be found that if another coil B is brought close to A, and in such a position that M is not zero, the effective L of coil A is decreased, if the second coil is connected to form a closed circuit so that current can flow in it. The amount of decrease in L depends upon the coupling between the two coils, upon the frequency, and upon the resistance in the circuit of the second coil.

This effect is likely to occur in certain variable coils used in radio circuits; in the type of coil referred to the change in the self-induction of the coil is accomplished by using more or less turns of the coil by means of a sliding contact as indicated in Fig. 26. If the sliding contact B happens to make contact with two adjacent turns at once (quite a normal occurrence), there is one turn of the coil short-circuited, and this short-circuited turn is quite closely coupled with that part of the coil which is being used. The effect of this turn is to decrease very much the effective self-induction of the part of the coil $A-B$, which is being used. Now as the slider is being adjusted it will, with very little movement, make contact with two turns or with only one turn; a signal may come in very strong at a certain setting of the slider and the slightest movement of the slider one way or the other will make the signal disappear. This is due to the large change in the self-induction of the coil as the slider makes, or does not make, the double contact.

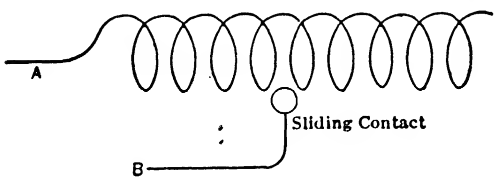


FIG. 26.—Variable inductance with sliding contact.

A short-circuited turn in a coil not only produces a decrease in the L of the coil, but it also increases very materially the resistance of the coil, and this is detrimental to the proper operation of the set; these two points will be taken up more in detail later in this chapter.

Capacity—Charging a Condenser.—Suppose a battery is connected through a switch to a condenser as indicated in Fig. 27. The condenser C consists of two metal plates a and b, close together, but perfectly insulated from one another by the layer of air between them. When the switch B is closed the plate b is made negative with respect to a, by an amount equal to the e.m.f. of the cell, perhaps 1.5 volts; that this must be so follows from the fact that, when the switch is closed, b is connected to the negative end of the cell and a is connected to the positive end of the cell.

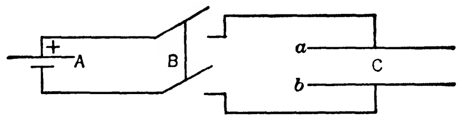


FIG. 27.—Charging a condenser.

As the two plates a and b were at the same potential before the switch was closed, and after the switch is closed b is 1.5 volts lower in potential than a, the closing of the switch must have been followed by a flow of electrons in the direction from a to b. This redistribution of the electrons in the circuit, which serves to bring the condenser plates to the same difference of potential as are the terminals of the cell to which they are connected, is called *charging the condenser*. A current flows during the short interval of time required for the redistribution of the electrons; this current is called the charging current of the condenser.

It is more or less evident that the condenser will take sufficient charge to bring its potential difference equal to that of the battery; as long as the condenser is at a lower potential difference than the terminals of the battery, the e.m.f. of the battery causes more electrons to flow; if, by any chance, so many electrons accumulate on the b plate of the condenser that potential difference of the condenser is greater than that of the battery, the excess of potential difference would so act as to make the condenser discharge itself until it was at the same potential difference as the terminals of the battery.

This charging process is often explained by stating that electrons leave one plate of the condenser and flow around the circuit, through the battery, to the other plate. It must not be thought, however, that such is actually the fact. If, for example, the condenser has a capacity of 1 microfarad (an ordinary telephone condenser has this capacity) and the connecting wire is a small one, such as is used in connecting electric door bells, etc., the electrons will move along in the connecting wire, in the average, only about 10^{-9} cm. to charge the condenser to 1.5 volts potential difference. Even if it is changed to 1000 volts they will move through the connecting wire only about one millionth of a centimeter.

Capacity of a Condenser.—Suppose the amount of electron flow necessary to charge two different condensers to a certain potential difference is measured by a ballistic galvanometer or similar device. It will be found in general that the different condensers require a different amount of charge to bring them to the same difference of potential. For example, if two condensers are made of the same sized metal plates, but in one the plates are only half as far apart as in the other, it will be found that the one with closer plates requires twice as much charge as the other; if two condensers have the same spacing for the plates, but one has larger plates than the other, again it will be found that one requires more charge than the other, in this case the one with the larger plates.

That characteristic of a condenser which determines how many electrons it takes to bring the condenser plates to a certain potential difference is called its *capacity*. A condenser which requires 1 coulomb of electricity to bring its plates to a potential difference of 1 volt, has a capacity of *1 farad*. Such a condenser would require immense plates very close together; the unit is altogether too large to represent the capacity of ordinary condensers. In ordinary engineering practice, such as telephone circuits, the microfarad is used as the unit of capacity. A condenser of 1 microfarad requires a charge of one millionth of a coulomb to charge it to 1 volt. Stated in another way, a current of 1 ampere would have to flow only one millionth of a second to charge the condenser to 1 volt potential difference; or 1 microampere, flowing for 1 second, would charge it to the same extent.

In radio circuits the microfarad is too large a unit to be conveniently used; a more suitable unit is the milli-microfarad, which is the thousandth part of a microfarad. Another unit is the micro-microfarad, which is one-millionth of a microfarad. Still another unit is the centimeter; which is one nine hundred thousandth of a microfarad. The micro-microfarad and the centimeter are nearly the same-sized units, the centimeter being about 1.1 of a micro-microfarad.

The capacity of a standard Leyden jar formerly much used in radio transmitting sets is 2 milli-microfarads. The variable condensers used for tuning a receiving set have a maximum capacity of one milli-microfarad or less. Certain condensers used with vacuum-tube detectors have a capacity of 100 micro-microfarads. Antennas, such as are used on small vessels, have a capacity of about 0.5 milli-microfarad; large land stations designed for transoceanic communication may have antennas of as much as 10 milli-microfarads capacity.

Specific Inductive Capacity.—Suppose a condenser made of two metal plates separated by $\frac{1}{8}$ in. of air and let the quantity required to charge it to one volt be measured. Then let a $\frac{1}{8}$-in. glass plate be slipped in between the two plates of the condenser and let the quantity be again measured; it will be found to be about six times as much as when air was used to separate the plates. If various other materials are used as dielectric it will be found that they all take more charge than the air condenser; in other words, when such insulators as glass, mica, rubber, etc., are used for the dielectric instead of air, the condenser has more capacity, its dimensions being the same in each case. The ratio of the capacity of a condenser in which some dielectric other than air is used to that it would have if air were used, is called the *specific inductive capacity* of the dielectric.[1] The values of this constant for some of the more common insulators are given in the table on page 253.

Energy Stored in a Charged Condenser.—It takes work, or energy, to charge a condenser; the amount of this work depends upon the capacity of the condenser and upon the voltage to which it is charged. The problem is analogous to the " pumping up " of a tire; the amount of work done in this case is evidently proportional to the size of the tire and depends in some way upon the pressure to which the tire is pumped. Actually the amount of work required increases with the square of the pressure to which it is pumped; pumping a given tire to 40 lb. pressure requires four times as much work as is required to pump it to 20 lb pressure.

The energy used in charging a condenser, and stored in the electric field, between the plates of the condenser, is

$$\text{Work} = \tfrac{1}{2}CE^2, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where C = capacity of condenser in farads;
E = voltage to which condenser is charged, in volts, and the work is given in joules.

A condenser of 0.002 microfarad, charged to 15,000 volts difference of potential, has stored in its field 0.225 joule of energy. If the energy stored in this condenser is discharged to produce the oscillatory currents required in the radio transmitter, it may be used to supply about 100 watts of power, with a suitable charge and discharge frequency.

Suppose sixteen such condensers are connected in parallel, so that each is charged to the same voltage, 15,000 volts. There will be stored in this battery of condensers 16×0.225 joule, or 3.6 joules. If the condensers discharge through a spark gap which operates 1000 times a second (a common spark frequency) there will be transformed into oscillatory current 3600 joules per second, that is, 3600 watts of power. Hence sixteen such jars, good to operate at 15,000 volts, would be sufficient for generating about $3\frac{1}{2}$ kilowatts of high-frequency power, for a spark frequency of one thousand.

[1] Actually the condenser used for reference should have its plates separated by vacuum; the specific inductive capacity of air is however so nearly the same as that of vacuum that in laboratory measurements an air condenser may be used for the reference condenser.

Current Flow in a Continuous Current Circuit Containing Resistance Only.—If a continuous e.m.f., such as that from a battery, is impressed upon a circuit containing resistance only, a continuous current will flow and its value is given by Ohm's law,

$$I=\frac{E}{R}, \qquad (11)$$

where I = current in amperes;
E = e.m.f. of the battery, in volts;
R = resistance, in ohms, of the entire circuit.

The current will have this value from the instant the switch is closed, and will be as continuous (constant in magnitude) as is the e.m.f. of the battery.

Current Flow in an Inductive Circuit.—If the circuit to which the battery is connected contains inductance as well as resistance, the current flowing will have the value given by Eq. (11) only after the switch has been closed for some instants; it does not rise to the value predicted by this equation for quite some time after the switch has been closed. The fact that there is inductance in the circuit as well as resistance does not affect the final value of current, but it does affect the current for a short time after closing the switch.

In an inductive circuit the current cannot at once rise to its steady value; it takes an appreciable time to reach the final value predicted by Ohm's law. The length of time taken depends upon the ratio of the inductance to the resistance of the circuit. The value of current is expressed at any time after closing the switch by the equation

$$i=\frac{E}{R}(1-\epsilon^{-\frac{Rt}{L}}), \qquad (12)$$

in which i = the current in amperes at time t after closing the switch;
E = the e.m.f. of the battery;
R = the total resistance in the circuit, including that of the battery, in ohms;
L = the coefficient of self-induction of the circuit, in henries;
t = the number of seconds elapsing after the switch is closed;
ϵ = the base of natural logarithms = 2.718.

This equation defines the expression "logarithmic rise of current." If a circuit has a very large value of inductance compared to its resistance, the rise of current may be so slow that it can actually be observed by means of an ammeter in the circuit. This is very easy to observe, for example, in the field circuit of a large generator, in which the current may take several seconds before it approximates its final value.

Time Constant of an Inductive Circuit.—When the time elapsed after the switch is closed is equal to the L/R of the circuit the current has risen to $(1-1/\epsilon)$ of its final value, or to about 63 per cent of its final value. The time taken for the current to reach this fraction of its final value is called the *time constant* of the circuit; in most inductive circuits it has a value only a small fraction of a second, but it may in special cases be considerably larger than this.

A point which is apparently but little appreciated by text-book writers is the fact that the time constant of an inductive circuit is intimately connected with its physical dimensions. Thus it is simple to "assume a coil of 1 henry inductance and 1 ohm resistance," as is suggested in one problem. But this means a time constant of 1 second. Now iron-core coils should not be assumed in such a problem, because the inductance of an iron-core coil is an indefinite quantity, being different for every different current. And an air-core coil to have a time constant of one second would require *many hundred pounds of copper wire;* thus a well-designed coil, employing 75 lb. of stranded copper cable, has a time constant of 0.06 second. The weight of copper required for greater time constants varies roughly with the square of the time constant desired.

The Oscillograph.—In investigating problems to-day the electrical engineer uses very extensively an instrument called the Duddell oscillograph. It receives its name from the fact that its essential part consists of a small mirror mounted on some fine wires, through which wires a current may be passed. The wires are mounted between the poles of a powerful magnet, and, due to the force acting between the magnetic field and the current in the wires, the mirror is caused to oscillate back and forth as the current in the wires changes its direction. This part of the instrument is really a small galvanometer so constructed that it can move very quickly a beam of light shining on the mirror and which, reflected therefrom, acts as a pointer to indicate the motion of the mirror. By suitable devices the motion of this beam of light may be either thrown on to a translucent screen and so serve for visual work, or it may be thrown to a rapidly rotating film and so give a permanent record of the excursions of the mirror. These films, showing how current varies with respect to time, are called *oscillograms;* such records will be frequently used in this text to illustrate phenomena being analyzed.

Such records are extremely valuable, as there are many rapid changes of current taking place in circuits which can be examined only in this fashion. Changes of current which are so rapid that they occupy only one-thousandth of a second are truthfully recorded by a properly used oscillograph; currents which alternate many hundred times a second are correctly shown by an oscillogram. Not only will the oscillogram show the number of times a second the current alternates, but it will also show

how closely the current approaches a sine wave in form and similar effects.

The ordinary oscillograph of the Duddell type requires about 0.1 ampere through its vibrator to give full-scale deflection. Its records are faithful reproductions of current forms for frequencies up to about one thousand cycles per second; the vibrator will respond to frequencies higher than this, but the picture obtained is not of just the same form as the current, due to the inertia of the vibrating parts (principally the mirror). Thus if a picture of a 500-cycle current is taken and in the 500-cycle current there are "ripples" which represent frequencies of several thousand per second, these ripples will not show up in the picture with as great an amplitude as they have in the actual current.

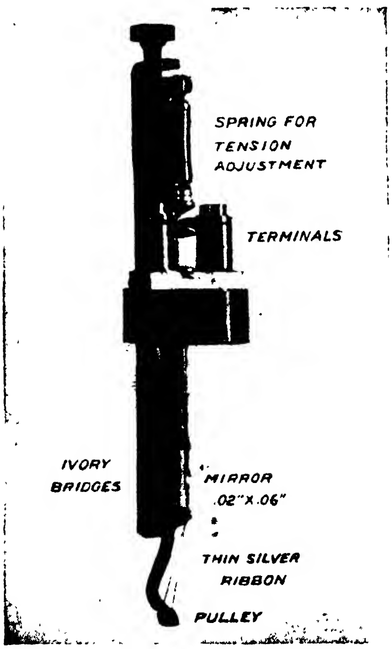


Fig. 28.—The vibrator of a Duddell oscillograph.

Because of the comparatively large current required and its frequency limitations, the Duddell oscillograph is hardly suitable for photographing the currents of radio circuits. Thus the current through the loud speaker of a broadcast receiver might be 0.005 ampere, consisting of a combination of frequencies from 100 to 5000 cycles per second; this current is too low in amplitude, and too wide in frequency range to be accurately photographed by a Duddell oscillograph. For ordinary engineering circuits, however, this type of oscillograph is most useful, as will be apparent from the illustrations given in later portions of this text.

The vibrator, or moving part of the Duddell oscillograph, is shown in Fig. 28. A very small mirror is cemented to two fine silver strips through which flows the current to be photographed. The silver strips are held taut in their proper place by being stretched over two small, slotted, ivory bridges. The natural frequency of oscillation of this bifilar suspension, with the mirror, is about 4000 per second. A system of this kind should never be used to photograph frequencies higher than its natural frequency; in fact, it is well not to use it for frequencies higher than about one half its natural frequency.

When properly mounted and ready for use, the vibrator is placed in a small oil-filled cell, oil being used to damp out any effects due to the tendency of the vibrator to oscillate at its own natural frequency. Through that part of the cell where the mirror is supported, a strong, controllable, magnetic field is caused to pass; the poles of a powerful electromagnet envelop the cell (which of course is made of brass, not iron). In one type of this oscillograph, commercially available, permanent magnets are used instead of electromagnets.

As generally used three vibrators (or more) are properly mounted so that their mirrors all receive light from the same intense source, and so that the beams from the separate mirrors can all be focused on the same

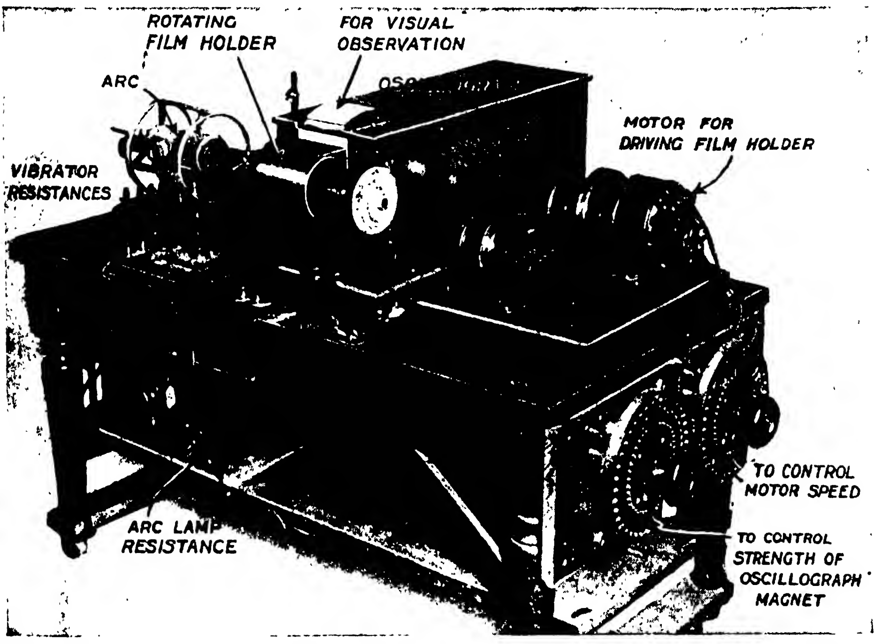


FIG. 29.—Three-vibrator Duddell oscillograph conveniently mounted on a movable table, with all required accessories.

photographic film. Thus three records of voltage or current, can be obtained simultaneously, to give not only forms and magnitudes of the various quantities, but also their relative phases. In Fig. 29 is shown a three-vibrator oscillograph conveniently mounted on a laboratory table, with all essential accessories. In a commercial oscillograph of another manufacturer, there are nine vibrators which may be used simultaneously.

There are other types of oscillographs which the radio engineer engaged in research work may employ. The Einthoven "string galvanometer" makes photographs of currents of amplitudes as small as 1 microampere, or even less. It has severe frequency limitations, however, and is generally not suitable for frequencies greater than a few hundred per second. Its resistance is ordinarily a thousand ohms or more, while the Duddell vibrator has a resistance of only 1 ohm.

For measuring the magnitude and forms of the currents of an audio-frequency amplifier, the string galvanometer is of great service, in fact it is the only device by which such currents can be readily photographed. The plate circuit currents of an amplifier may be as much as 1 milliampere (or somewhat larger) whereas the grid currents are measured in microamperes or fractions of 1 microampere. The frequencies may vary from 10 to 6000 cycles per second. Although the string gives rather limited response for the high frequencies, much less than for the lower ones, the photographic records obtained do show very well what is taking place in the various parts of the amplifier.

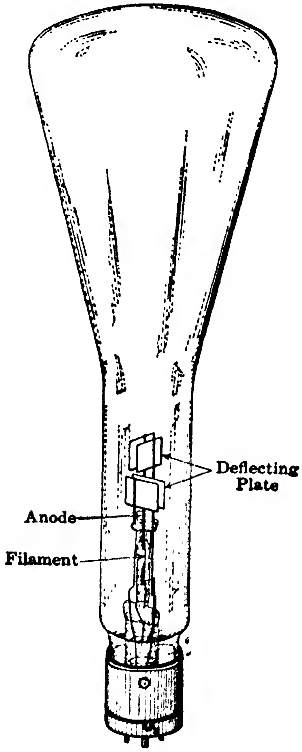


FIG. 30.—One form of cathode ray oscillograph.

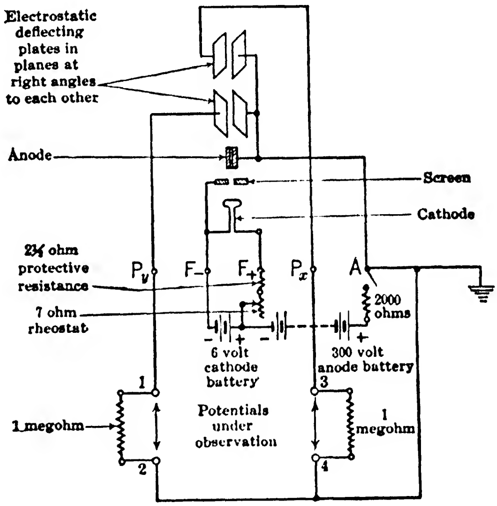


FIG. 31.—Connection scheme of cathode ray oscillograph arranged for electric field deflection of the beam.

The string is of quartz, silver coated, so small as to be invisible to the naked eye; it is suspended between two small microscope lenses in a powerful magnetic field. A compound optical system intensely illuminates the string, and throws its shadow upon a suitably moving photographic film. The optical system gives a magnification of 1000 times or more, so that a string motion of 0.001 cm. appears on the film record as a movement of 1 cm.

The comparatively small response of this instrument for frequencies higher than a few hundred per second is due to the relatively large mass of air which the string carries along as it oscillates.

In the *cathode ray oscillograph* a stream of electrons is used for the vibrator; this beam is made to affect a fluorescent screen or a photographic plate. In the former case a stationary figure is shown on the screen and the form of the figure can be used to calculate the shape of the radio wave used to deflect the electron ray. The fluorescent screen scheme requires that the radio current being investigated is periodic, that is, it repeats its form for several seconds.

In Fig. 30 is shown the form of one type of cathode ray oscillograph (developed by the Bell Telephone laboratories). A small filament furnishes the electrons which are accelerated by a plate potential of 300 volts and directed into a narrow beam by passing through a small hole in the plate. This beam (perhaps 1 mm. in diameter) shoots between the two sets of plates shown in Fig. 30 and impinges upon the flattened end of the tube. This end has been coated on the inside with some material which fluoresces when the electrons strike it, showing a greenish spot about 1 mm. in diameter when the tube is in proper adjustment.

A small amount of gas is left in the tube after evacuation, and the ionization of this permits a reasonably sharp focus for the impinging beam. For a given plate voltage a critical value of filament current gives the proper amount of ionization for producing a sharp spot on the fluorescing screen.

The voltage to be measured is impressed upon one or the other pair of plates and so deflects the beam of electrons in a direction perpendicular to the plane of the plates. The other pair of plates is used to draw the beam across the tube at a known speed, to give a time axis for the voltage wave. The IR drop across a resistance in series with a charging or discharging condenser serves well for this purpose. A few volts will give reasonable beam deflections; with 300 volts anode potential, 10 volts in the deflecting plates move the spot on the screen 1 cm.

In Fig. 31 is shown the electrical connection scheme for this cathode ray oscillograph, arranged to give deflections from both sets of plates by the drops across the megohm resistances.

Instead of using electric fields for deflection it is possible to use a magnetic field; the arrangement required is shown in Fig. 32. With two coils mounted outside the neck of the tube as shown here the spot on the screen is deflected 1 cm. for about 10 ampere turns in the two coils.

This type of oscillograph performs very important service in high-frequency circuits; the negligible mass of its vibrator (the beam of electrons) permits it to follow accurately frequencies measured in many millions per second. It is rather inconvenient to obtain photographic records from the oscillograph, and a further limitation arises from the fact that the phenomenon to be photographed must be periodic and must last several seconds.

In the most recently developed type of oscillograph (Dufour) the electron stream is made to impinge directly on the photographic plate, which is in the same evacuated vessel as the electron stream itself. The thin beam of electrons moves over the photographic plate and affects the sensitized film just as light does. Upon development, the path of the electron stream over the plate shows up as a narrow black line. This type of oscillograph will accurately picture the form of a current with a frequency as high as hundreds of millions of cycles per second or more, and does not require that the current be periodic.

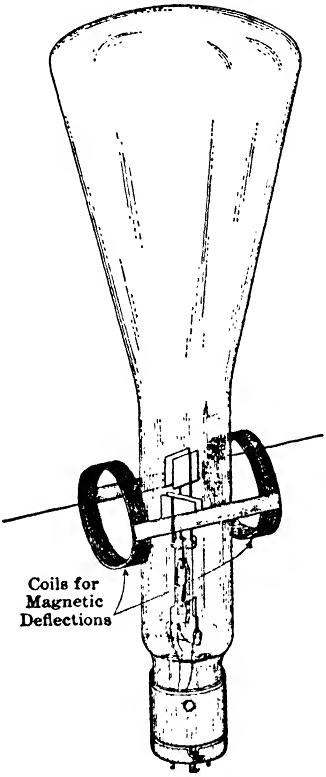


FIG. 32.—By using coils, of a few turns, mounted as shown here, the beam may be used to show the variation of a magnetic field, hence current.

Rise of Current in an Inductive Circuit.—In Fig. 33 is shown an oscillogram of the current rising in an inductive circuit; it will be seen that the current rises rapidly at first and gradually approaches its steady value. If the switch should be opened quickly in such an inductive circuit a large arc will form at the point of the switch where the circuit is opened. The energy stored in the magnetic field has to disappear when the current dies to zero because there can be no magnetic field without current.[1] The greater the self-induction of the circuit the greater is the amount of energy (for a given current) and the larger will be the arc when opening the circuit. The decay of current in an inductive circuit cannot be well examined therefore by opening the circuit, but it can be shown by short-circuiting the coil in which the current is flowing. In such a case the current dies away on a logarithmic curve quite similar to the curve of current rise. The equation of current decay is quite similar to that of the current rise and is

$$i = \frac{E}{R}\left(\epsilon^{-\frac{Rt}{L}}\right), \qquad (13)$$

where the letters have the same meaning as in Eq. (12).

[1] This statement of course neglects any residual field left in iron parts of the magnetic circuit when the current has fallen to zero.

Fig. 33 serves also to show this effect, the circuit having been arranged as shown in Fig. 34. The battery D was connected to the inductance C through a low resistance E and switch A. The oscillograph was connected in the circuit at the point O. A second switch B served to short-circuit the coil so that the decay of current in it could be shown as well as the rise.

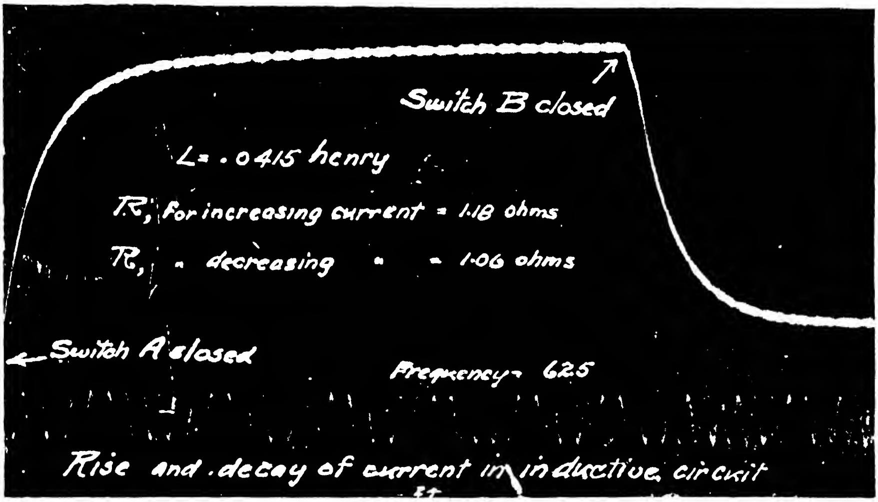


FIG. 33.—Oscillogram showing rise and fall of current in an inductive circuit.

With B open, A was closed and so the oscillograph recorded the rise of current; when the current had reached its steady state switch B was closed, and the decay of current in the coil was recorded. The resistance E was used in the circuit to prevent the short-circuiting of the battery when B was closed.

The curves of rise and decay are just as is predicted by Eqs. (12) and (13); the two curves show a slight difference in the rate of change of current, but this is to be expected, because the resistance was somewhat greater for the rise of current than it was for the decay, while the inductance was the same for both. The time constant was greater for the decaying current than for the rising current; the rising current had for its resistance that of the coil, that of the battery, and that designated by E, while the decaying current took place through the resistance of the coil only.

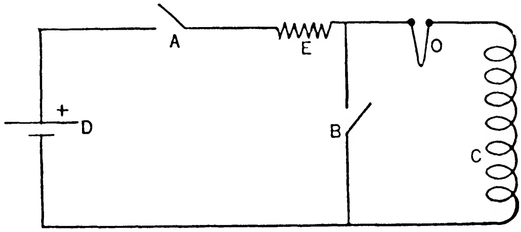


FIG. 34.—Circuit used to obtained oscillograms of growth and decay of current.

In inductive circuits it may take several seconds before the current reaches its steady state value given by the relation $I=E/R$. Directly after the switch is closed the current is of course zero; it starts to rise on a practically straight line, the slope of which gives the rate of change of current. This rate of change is given by the expression $di/dt=E/L$, where E is in volts, L is in henries, and the current change is in amperes per second. Thus the coil used in Fig. 33, connected to a 10-volt storage battery, would show a rise of current immediately after the switch is closed of $10/.0415=240$ amperes per second.

Effect of Rising and Decaying Currents on Neighboring Circuits.—As the current in the coil increases and decreases it must induce electromotive forces in any neighboring circuits which are so placed that they link with its magnetic field. If the neighboring circuit is closed, current will flow in one direction when the current in the first circuit is rising, and in the opposite direction when the current in the first circuit is falling. Hence when a circuit is closed and current starts to flow, all neighboring circuits, if closed, will have currents in one direction and in the opposite direction when the circuit is opened. The equations for currents in such an arrangement are cumbersome, but may be found in any good theoretical texts dealing with electric circuits.[1]

To bring out this fact a circuit was arranged as shown in Fig. 35; one oscillograph vibrator was introduced at C and the other at D. The currents which flowed in each circuit during the opening and closing of switch E is shown in the oscillogram given in Fig. 36. When the switch was closed current in coil B flowed in the opposite direction to that in coil A; when the switch was opened the current in B flowed in the reverse direction. The rather irregularly shaped curve of current at the time of opening the switch was due to the fact that an arc formed at the point of opening the circuit so that although the switch was open the circuit was not open, the arc serving to keep the circuit closed. As the resistance of the arc was indefinite and variable the current naturally followed no regular curve.

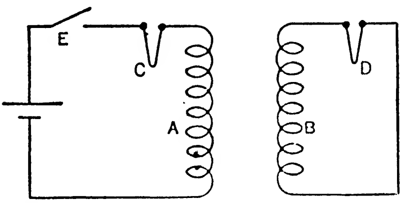


FIG. 35.—Circuit used to obtain oscillogram of currents in coupled circuits.

Magnetic Flux, not Current, Shows Inertia.—We quite generally say that in a coil the current cannot be quickly changed because of the self-induction of the circuit. The current rises on a logarithmic curve as given by Eq. (12), and decays according to Eq. (13), both of which lead one to believe that the current itself is what cannot change rapidly. A

[1] See Berg, "Electrical Engineering, Advanced Course," pp. 51, et seq.

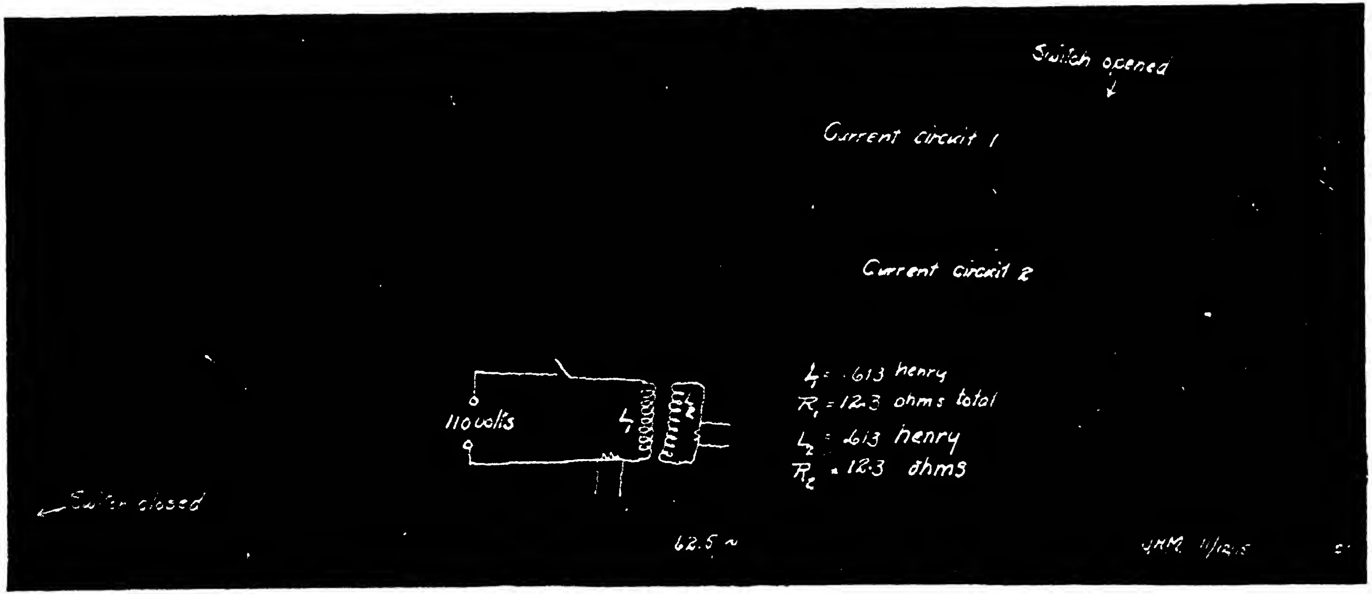


FIG. 36.—Oscillogram of currents in coupled circuits. Notice that when the switch is closed the currents in both circuits follow smooth curves, but that when the switch is opened both currents are irregular in form. This is due to the varying resistance of the arc across the switch opening; although the circuit is mechanically opened when the switch blade leaves the clips, the circuit is not opened electrically until the arc ruptures.

more correct analysis shows that the current in an inductive circuit resists change only in so far as the change in current is accompanied by a corresponding change in the magnetic flux of the circuit. If by some means we can change the current in a coil without changing the flux rapidly, then the current may rise and fall as suddenly as it does in a purely resistive circuit.

Now if we have two coils closely intermeshed on the same magnetic circuit, we get a phenomenon like that of Fig. 36, but the results are more striking. Fig. 37 shows what happens in this case. One coil was connected to a source of current through a switch and the other coil, identical with the first, was short-circuited. Each coil had about one henry of self-induction.

It will be seen that in closing the switch the current in the first coil immediately rose to practically its final value, whereas from the con-

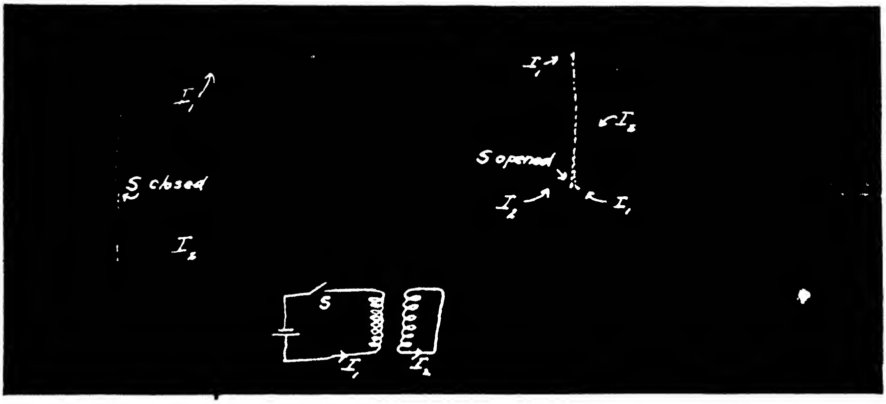


FIG. 37.—Showing that current in an inductive circuit can change very rapidly if another coil on the same iron core is short-circuited. The irregularity in both currents soon after the switch was closed is due to imperfect closure of the blade in its clips.

stants of the circuit it should have taken a second or more. But it will be noticed that as soon as the switch was closed a current flowed in the short-circuited coil, and this current was in the opposite direction to that in the first coil. Furthermore, directly after closing the switch the current in the second coil is practically as large as that in the first coil. As both coils have the same number of turns, and are on the same core, there are evidently two opposite, and practically equal, magnetomotive forces acting in the core, so that *no flux is set up in the core* in spite of the large current flowing in the coil connected to the power supply.

The current in the second coil now starts to decay, following almost exactly Eq. (13). As this secondary current dies down (primary current remaining practically constant) the *net* magnetomotive force acting in the magnetic circuit increases and the flux builds up, and it builds up in the logarithmic curve corresponding to that given for current in Eq. (12). We can see then that it is the flux, not the current, which shows inertia, or the tendency to resist change.

When an inductive circuit is opened we always expect to get an arc at the point of opening, and say generally that this arc is caused by the current in the coil refusing to change rapidly. But the oscillogram of Fig. 37 shows this is not so. When the switch was opened the current in the first coil dropped immediately to zero and there was no arcing at the switch. It is seen, however, that when the switch of coil No. 1 is opened, causing its current to drop to zero, that coil No. 2 immediately assumed a current just equal to that which had been flowing in coil No. 1 and in the same direction. In other words, circuit No. 2 immediately assumed the task of maintaining the flux of the circuit at its proper value. The current in coil No. 2 now dies down as given by Eq. (13), and as there is no current in coil No. 1 the flux of the core follows the current of coil No. 2, thus falling off in accordance with Eq. (13).

Current Flow on Connecting a Condenser to a Source of Continuous E.M.F.—When a condenser is connected to a source of continuous e.m.f. the condenser takes sufficient charge to bring its plates to a difference of potential equal to the e.m.f. of the source to which it is connected. This charging would take place instantaneously if there were no resistance in the circuit. But the generator or battery to which the condenser is connected always has resistance and the condenser itself has a kind of resistance due to the losses occurring in its dielectric, all of these resistance factors act in such a way that the condenser takes an appreciable time to charge itself.

A circuit was arranged as shown in Fig. 38; A is a 100-volt battery, B and D are switches, C is the condenser to be charged or discharged, O is the oscillograph vibrator, and R is a resistance which represents the total resistance of the circuit, battery, connections, condenser, etc.

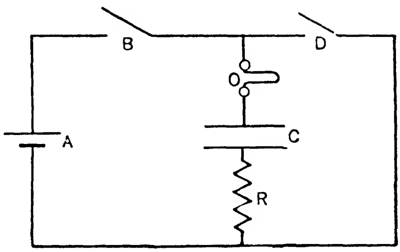


Fig. 38.—Circuit used to obtain oscillogram of charge and discharge of a condenser.

The equation for the current which flows in such a circuit is given by

$$i=\frac{E}{R}(\epsilon^{-\frac{t}{RC}}), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

where E = the battery voltage in volts;

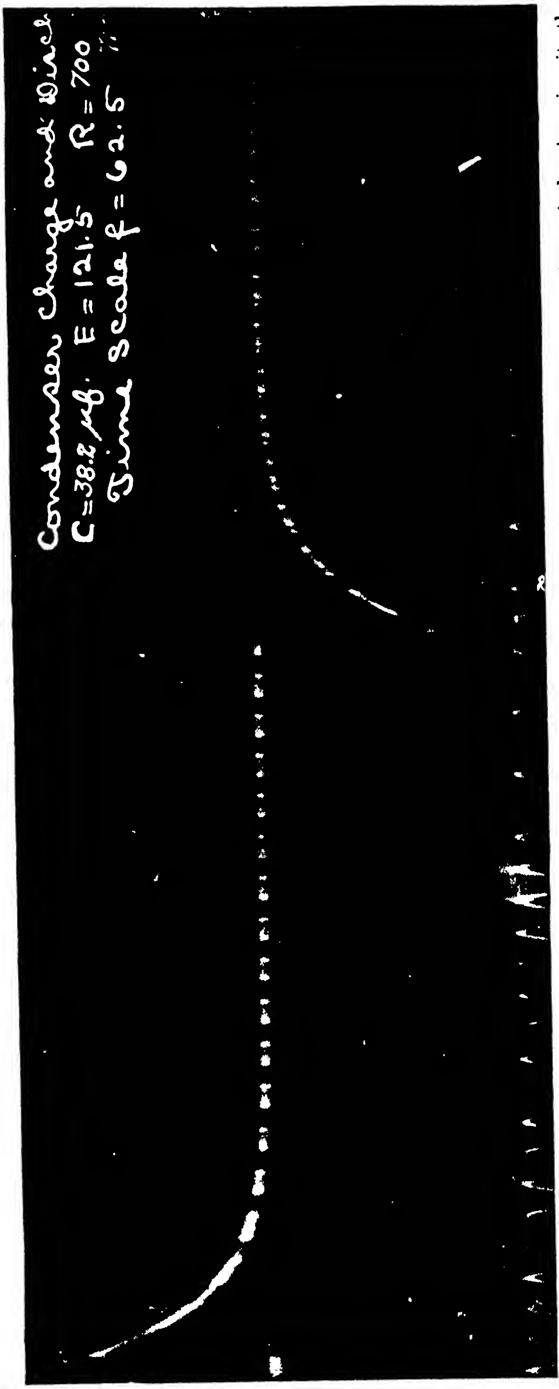


FIG. 39.—Oscillogram of charging and discharging current of a condenser. Unlike the case of current rise in an inductive circuit, the current here rises to its maximum value immediately the switch is closed.

R = the total resistance of the circuit in ohms. This resistance is exclusive of the infinite resistance between the plates of the condenser:

C = the capacity of the condenser in farads.

If now switch B is opened and switch D is closed, the condenser will discharge and the current will be given by

$$i=-\frac{E}{R}\left(\epsilon^{-\frac{t}{RC}}\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

where the letters have the same meaning as they have in Eq. (14). This current is evidently of the same shape as that taken by the charging operation with the exception that there is a minus sign before it; this signifies that the discharge current is of the same form as the charging current, but it flows in the opposite direction.

Time Constant of a Condenser Circuit.—The quantity RC is called the time constant of the condenser circuit; it is evidently the time taken for the current to fall from its maximum value to 37 per cent of this value. Another way of defining the time constant of a condenser circuit is in terms of the charge on the condenser; the time constant is the time required for the condenser to acquire 63 per cent of its final charge or, in the case of the discharging condenser, it is the time required for the condenser to lose 63 per cent of its charge.

Fig. 39 shows an oscillogram of charge and discharge which was taken from the circuit shown in Fig. 38. Some extra resistance must be necessarily added to the inherent resistance of the battery and condenser because the time constant of such a circuit is excessively small, too short for the oscillograph to function. Thus a 1-microfarad condenser in series with 2 ohms (a probable value for the battery) would have a time constant of 0.000 002 second, that is, the current would rise instantaneously upon closing the switch, to some value (depending upon the voltage used in charging) and in 0.000 002 second would have fallen to 37 per cent of this value, and in a correspondingly short time would have dropped to practically zero. Such an instantaneous occurrence is too rapid even for the oscillograph, hence to increase the time constant to a value suitable for the use of the oscillograph an extra resistance had to be introduced in the circuit.

The effect of adding resistance in series with a condenser to be charged is shown by the curves of Fig. 40; these were calculated from Eq. (14). They show that the initial current is cut down as the resistance is increased, in fact being equal to E/R, and that the duration of the current increases with the increase of resistance. The area between the X axis and any one of the curves is the same; this area represents the quantity of electricity on the condenser and so must be the same for all conditions, because the quantity of electricity on the condenser after the charging process is complete is the same no matter what the resistance of the circuit may be.

Power Expended in a Continuous-current Circuit.—If a current of I amperes is caused to flow through a circuit by an e.m.f. of E volts the rate of doing work in the circuit is

$$\text{Watts} = EI. \qquad (16)$$

If the circuit has a resistance R we know that $E = IR$ and so

$$\text{Watts} = IR \times I = I^2R \qquad (17)$$

from which we get

$$R = \frac{\text{Watts}}{I^2}. \qquad (18)$$

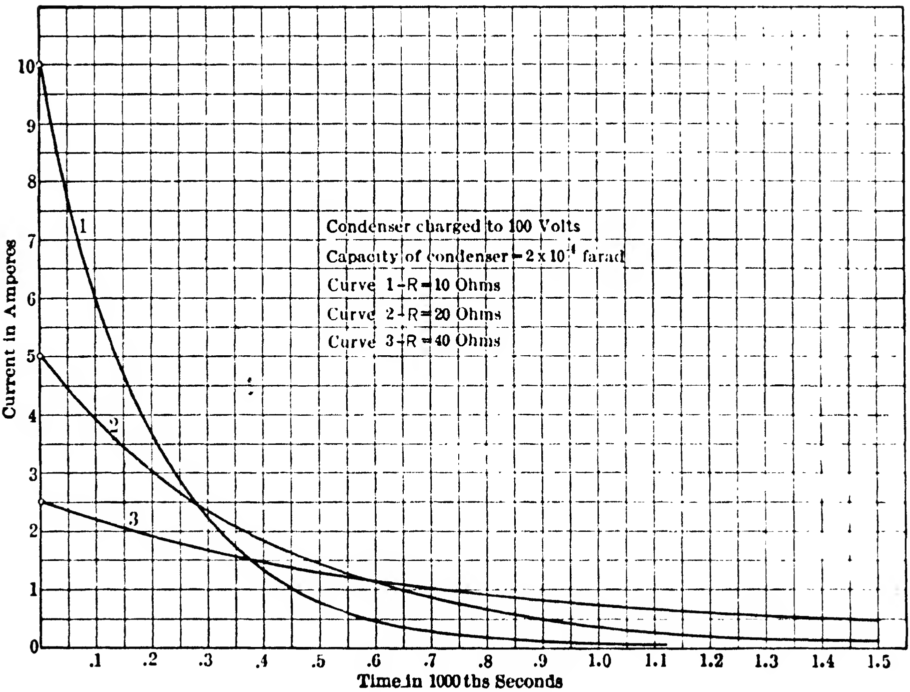


FIG. 40.—Condenser charging currents for different values of series resistance.

Eq. (18) is important; it is the broadest possible definition for the resistance of a circuit. This formula gives the resistance for any kind of current flow, whether continuous, pulsating, or alternating. In words it is stated thus: *the effective resistance of a circuit is equal to the amount of* *power consumed by the circuit divided by the square of the current required to supply this power.*

In simple c.c. circuits Ohm's law is sufficient to obtain the resistance of the circuit, but there are many cases especially in a.c. work, where Eq. (18) affords the only feasible means of determining the resistance of the circuit.

Power Consumed in a Circuit Excited by Pulsating Current.—In case the voltage or current of a circuit, or both of them, are pulsating the power consumed in the circuit cannot be obtained by using the product of the average voltage by the average current, as might at first seem correct; an error would be introduced generally making the calculated power consumed too low, the amount of this error depending upon the amount of fluctuation. The greater the amount of fluctuation or pulsation of the current or voltage, the greater is the error introduced.

The power is accurately obtained however by taking the product of the effective resistance of the circuit and the square of the effective value of the current. The derivation of the effective value of the current may be difficult; it can always be carried out graphically if the form of the pulsating current is accurately given, but is not easily calculated by ordinary arithmetic unless the form of the pulsation is very simple. Thus suppose that a pulsating current is simple enough to be represented by a continuous current, with a sine wave alternating current superimposed, as shown in Fig. 41. The actual pulsating current A is sufficiently well represented by the continuous current B, of amplitude I_1, and a sine wave current C, of maximum value I_2. The effective value of such a current is given by taking the square root of the sums of the squares of the effective values of the two components. The effective value of the continuous current is the same as its actual value, I_1; the effective value of the sine wave of current is the square root of one-half its (maximum value)^2. Hence the effective value of the pulsating current is $\sqrt{I_1^2+\frac{1}{2}I_2^2}$. The power used when such a current flows through a circuit of resistance R is

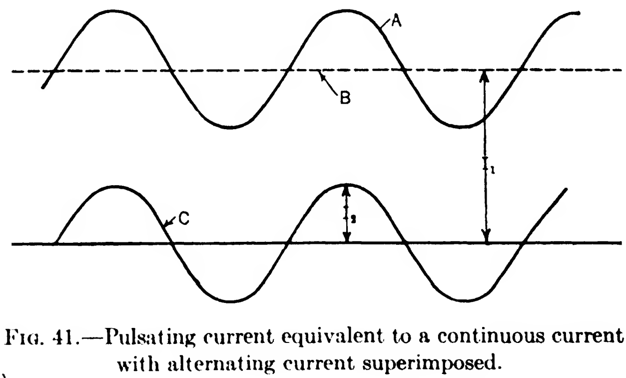


Fig. 41.—Pulsating current equivalent to a continuous current with alternating current superimposed.

$$\text{Watts used} = I_1^2R+\tfrac{1}{2}I_2^2R. \quad \quad (19)$$

If the average value of the current were used in calculating the power used, the power represented by the second term would be completely neglected, and so an error would be incurred equal to $\frac{1}{2}I_2{}^2R$. The amount of this error depends upon the amount of pulsation of the current. In such a circuit as the primary circuit of a spark-coil transmitting set excited by storage battery the error would be very large, and the power used in the circuit cannot be calculated at all accurately without knowing the form of the current flowing in the primary winding of the coil.

The above statement is made with the idea in mind that in such a circuit as this, excited by storage battery, a d.c. ammeter would be used in measuring the current. Now such an ammeter reads *average values* and so would read, when excited by such a current as sketched in Fig. 41, only the c.c. component. Hence the error pointed out would occur. If, however, an a.c. ammeter were used for reading the current, the error would not occur, because such an ammeter reads *effective values*, and not average values. If the power used in a pulsating-current circuit is to be

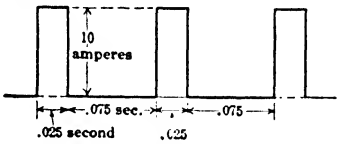


FIG. 42.—A unidirectional pulsating current, of rectangular form.

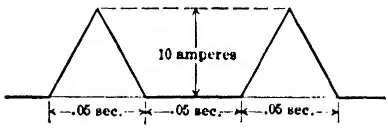


FIG. 43.—A pulsating current consisting of triangular pulses of current. The grid current of a triode may have nearly this form.

accurately determined, therefore, an a.c. ammeter must be used to measure the current.

In case the current is of the form of a rectangular pulse it is comparatively easy to calculate the average value, the effective value, and so the heating effect; this current is illustrated in Fig. 42. Here we have a rectangular pulse of current of 10 amperes amplitude, lasting one-quarter of the tenth second cycle. During three-quarters of the cycle the current is zero.

The average value of such a current is evidently one-quarter of 10 amperes, or 2.5 amperes; this is what a c.c. ammeter would read. The square root of the average square (effective value) is obtained by first finding the average squared current. As the squared value of the current is 100 and this lasts one-quarter of the cycle the average square is 25. The square root of this is 5, and this is the current an accurate a.c. ammeter would read. *One ammeter reads 2.5 amperes and the other reads 5 amperes—and both are correct.*

In the case of a triangular pulse of current, as shown in Fig. 43, the average value of current is readily calculated to be 2.5 amperes, the same as for the rectangular pulse. The average squared value of current is found after it is noticed that the triangular pulse becomes one with parabolic sides, after each ordinate is squared, and the average height of such a parabola-bounded area is one-third the maximum value. As the pulse lasts one-half the cycle the average value of the squared pulse is one-sixth the maximum, or 100/6. The effective value is the square root of this average square, or 4.08 amperes.

The two pulses of Figs. 42 and 43 will therefore give the same reading on a c.c. ammeter, whereas one of them gives a reading of 5 amperes and the other only 4.08 amperes, on an accurate a.c. meter.

The above analysis of the power used in pulsating-current circuits holds good only when *the resistance is constant throughout the cycle of current variation.* In many circuits this is not so, the resistance being a function of the current and so changing as the current changes. The calculation of the power used in such a circuit is not easily measured by ammeters and voltmeters; either a *wattmeter* or the oscillograph should be used. The wattmeter is an instrument having two windings in the same case, one corresponding to an ammeter and the other to a voltmeter. An analysis of its action and the way in which it is used will be taken up in a subsequent paragraph dealing with the power used in an a.c. circuit. The oscillograph, giving the form of voltage curve and current curve, makes it possible to calculate the power by graphical methods.

Current Flow in an Alternating-current Circuit Having Resistance only. Phase.—If an a.c. generator is connected to a circuit having resistance only the relation between current, resistance, and voltage is given by Ohm's law. It is, of course, impossible to construct a circuit "with resistance only"; a circuit must have some inductance and capacity no matter how it is built, but if the amount of inductance and capacity are so small that their influence upon the current is negligible compared to the influence of the resistance, the circuit may be considered to have nothing but resistance opposing the flow of current. The filament of an incandescent lamp is such a circuit. A rheostat constructed of high-resistance wire may be considered to have no inductance when being used in ordinary a.c. circuits, such as used for power and lighting, but such a rheostat would probably have such an amount of inductance that when used in a circuit of radio frequency it would be by no means negligible. It follows that a certain piece of apparatus might be considered free from inductance for some uses, but for other circuits the inductance might be of considerable importance.

In a circuit having resistance only the current and voltage have *the same phase* and are similar in form. A current and voltage are said to be *in phase* when they pass through their corresponding values simultaneously. The easiest point from which to judge the equality of phase is the zero value; if the two curves pass through their zero values, in the same direction, at the same instant they are in phase. In case the current passes through its zero value after the voltage has passed through its zero value it is said to be a *lagging current*; if it goes through the zero value before the voltage it is said to be a *leading current*.

In Fig. 44 are shown curves of current and voltage with (a) current and voltage in phase, (b) with current lagging behind the voltage by the angle ϕ, and (c) with the current leading the voltage by the angle ϕ.

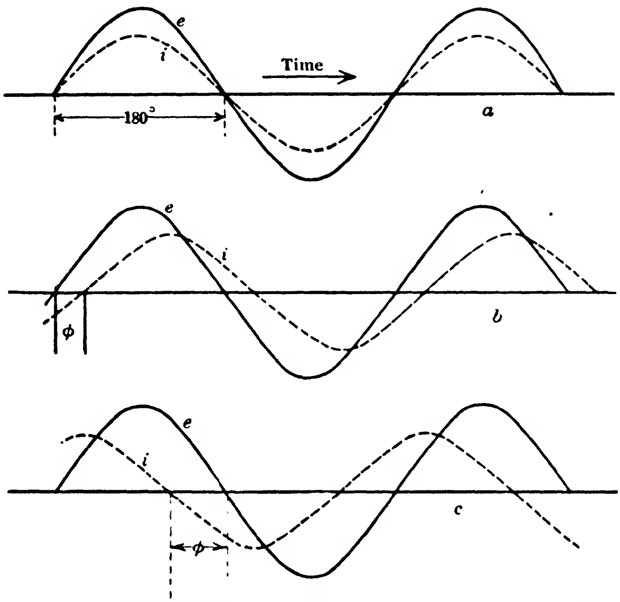


FIG. 44.—Phase difference of alternating current and voltage.

The magnitude of the angle of lag or lead may be easily approximated when it is remembered that the time from one zero point to the next zero point of the same curve is 180°; in curve b the current lags by about 30° and in curve c the angle of lead is about 70°.

In case the circuit has resistance only the relation between voltage and current is expressed by Ohm's law, whether instantaneous, maximum, or effective values are considered. Thus the equation for current flow in this circuit is

$$I=\frac{E}{R}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

Power Used in a Resistance Circuit.—The rate at which electrical energy is changed into heat by a current i flowing through a resistance R is i^2R, as has been shown for c.c. circuits. Or, as we know that for the circuit $e=iR$, we have,

$$\text{Rate of heat development} = \text{power used} = ei$$

The power curve has the form shown in Fig. 45; it is at all times positive, because although both e and i go through negative values they both reverse at the same instant; the product, therefore, is constantly positive. The maximum value of this power curve occurs when both e and i pass through their maximum values and is therefore equal to $E_m I_m$.

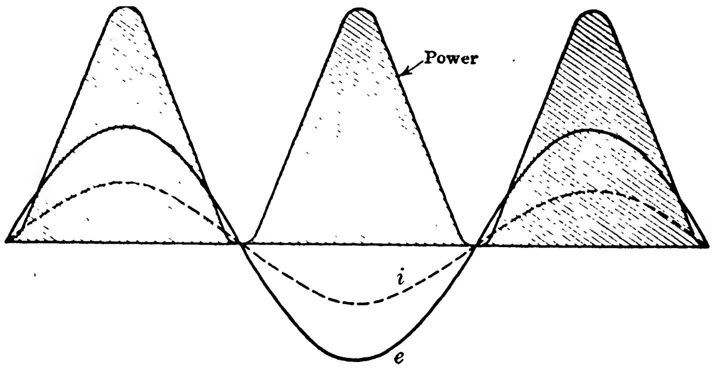


FIG. 45.—Power curve for an alternating-current circuit containing resistance only.

If the equation of current is $i = I_m \sin \omega t$ and the equation of voltage is $e = E_m \sin \omega t$, the equation of the power curve must be

$$p = E_m I_m \sin^2 \omega t$$

$$= \tfrac{1}{2} E_m I_m (1 - \cos 2\omega t). \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

The average value of $\cos 2\omega t$ is zero, hence the average value of power

$$= P = \tfrac{1}{2} E_m I_m = EI. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

It is seen therefore that the power (in watts) used in an a.c. circuit containing resistance only is the product of volts and amperes, as read by a.c. voltmeter and ammeter.

Power Used by Pulsating Current.—The idea of power used in a circuit with pulsating current and voltage is of great importance in the operation of vacuum tubes. It is frequently necessary to know the amount of power used in heating the plates of a vacuum tube, for example, and the current and voltage involved are unidirectional, but have wide fluctations. As the circuit is carrying a unidirectional current it would be natural to use c.c. ammeters and voltmeters to get readings and multiply the amperes and volts so measured to get the power in watts. This product would be by no means the power used on the plate.

A theoretical case (not quite obtainable in actual circuits) is shown in Fig. 46. The voltage on the plate has a sine-wave pulsation between zero and 4000 volts. A c.c. voltmeter would read the average value, namely, 2000 volts. The current has a sine-wave pulsation between zero and 1 ampere, having its maximum value when the plate voltage is a minimum. (This seems absurd, but is actually so, as will be explained in Chapter VI.) A c.c. ammeter reading this pulsating current would read the average value, or 0.5 ampere.

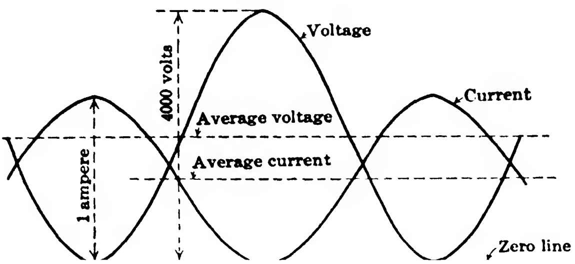


FIG. 46.—In vacuum-tube circuits currents and voltages approximating these forms may occur.

The product of ammeter and voltmeter readings would be 1000, so it would be concluded that 1000 watts of power were being used on the plate. Actually only 500 watts are being used.

We calculate first the amount of power used by the steady (continuous) components of voltage and current; this gives 1000 watts. Next we calculate the power due to the sine-wave components of voltage and current and notice at once that the sine-wave components of voltage and current are 180° out of phase. Then the power due to these must be reckoned as negative. Now the effective value of the sine-wave component of voltage is $\frac{2000}{\sqrt{2}}$ volts. The effective value of the sine-wave component of current is $\frac{1}{2\sqrt{2}}$ amperes. The product of these effective values is 500 watts, and because of the 180° phase relation we must reckon this as (−500 watts).

The total power used on the plate therefore is $1000-500=500$ watts. In this case the power indicated by c.c. meters is *twice as large* as the actual value.

If the phase of the sine-wave components of current and voltage should change 180°, bringing volts and current in phase, the true power would be 1500 watts, whereas the c.c. meter would again indicate 1000 watts.

Meters Used in A.C. Circuits.—It must be remembered that the ordinary c.c. instrument, ammeter or voltmeter, will not read at all if used in an a.c. circuit. Such instruments read the *average value* of voltage or current and, in an a.c. circuit the average values are zero. To read correctly on an a.c. circuit an instrument must give the same reading on a c.c. circuit, no matter which way the continuous current is flowing through it;[1] everyone familiar with the ordinary c.c. instrument knows that if the connection of the meter to the circuit is reversed the reading will reverse. Such an instrument, if actuated by an alternating

[1] This excludes the induction type of meter, which is practically never used in radio measurements.

current, would tend to oscillate between a certain direct reading and the equal reversed reading, but, as the alternating current reverses too rapidly for the needle of a meter to follow, it is evident that the meter would read zero no matter how much current was flowing through it.

Various types of meters are suitable for use on an a.c. circuit, the dynamometer type, the soft-iron vane type, the induction type, the thermocouple type and the hot-wire type. The last two types named are used almost exclusively for making measurements in radio circuits, as it is practically impossible to make the other types function properly at the very high frequencies used in radio work.

Transient Current on Switching a Resistance Circuit to an A.C. Line.—If a resistance circuit is switched to an a.c. line having no appreciable inductance the current rises instantaneously to the value it should have, depending upon the value of the voltage at the instant the switch is closed, as shown in Fig. 47. This condition of affairs is expressed by stating that there is " no transient current " or no transient condition, after closing the switch; the current rises at once to the value it would have had (at the time of closing the switch) in case the switch had been closed at some previous time.

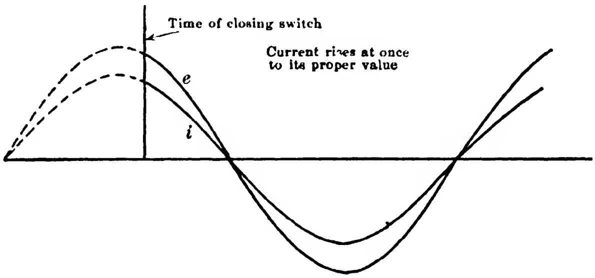


Fig. 47.—Current on switching a resistance circuit to an a.c. line.

Current Flow in an A.C. Circuit Having Inductance and Resistance.—Suppose that an inductance (without resistance) and a resistance, connected in series, are connected to an a.c. line so that a sine wave e.m.f. is impressed, as indicated in Fig. 48. Although the inductance must really have resistance, it is shown as resistanceless, all the resistance of the circuit being supposed concentrated in R. The current flowing in such a circuit depends upon four things, L, E, R, and the frequency of the impressed e.m.f. Provided that L and R are constant throughout the cycle (do not vary with the value of the current) it is a fundamental law of electrical circuits that the current will have

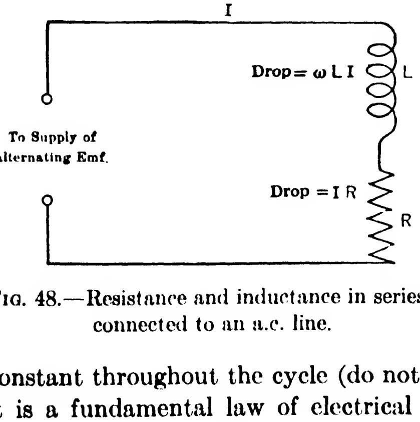


Fig. 48.—Resistance and inductance in series connected to an a.c. line.

the same form as the impressed force. We may therefore assume that the current is a sine wave and then find its magnitude and phase.

The impressed voltage must be equal to the sum of the drops in potential across L and R.

Suppose the current to be $i = I_m \sin \omega t$.

The drop across the resistance $= iR = I_m R \sin \omega t$.

The drop across the inductance $= L di/dt = \omega L I_m \cos \omega t$.

The impressed voltage must be

$$= I_m(R \sin \omega t + \omega L \cos \omega t) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (23)$$

In Fig. 49 these two component voltages are shown as curves; the impressed voltage e must be equal at all times to the sum of the resistance

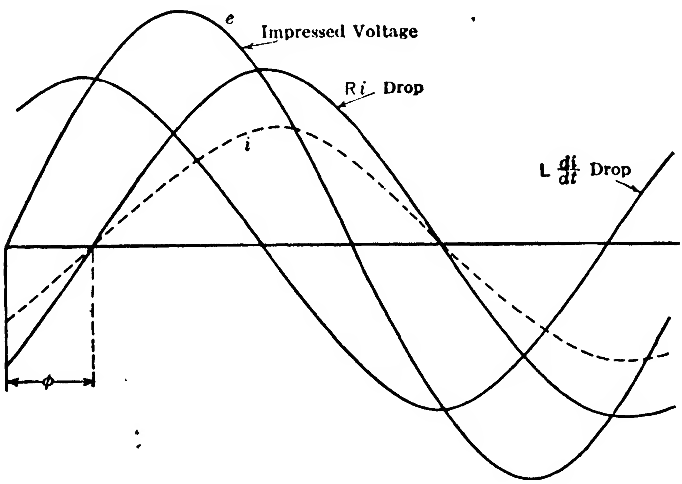


FIG. 49.—Voltage components in an a.c. circuit containing inductance and resistance.

drop and the inductance drop, and is so shown by the curve marked e in Fig. 49.

A vector diagram representing the curves of Fig. 49 is given in Fig. 50; effective values, instead of maximum values, are shown. From this diagram we have

$$E^2 = I^2(R^2 + (\omega L)^2)$$

or

$$I = \frac{E}{\sqrt{R^2 + (\omega L)^2}} = \frac{E}{\sqrt{R^2 + X_L^2}} = \frac{E}{Z} \quad . \quad . \quad . \quad . \quad . \qquad (24)$$

The quantity ωL, or X_L, is called the *reactance* of the circuit and the quantity Z is called the *impedance* of the circuit. The current lags behind the voltage by the angle ϕ, which is determined by the relation

$$\cos \phi = \frac{R}{Z} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

and

$$\tan \phi = \frac{\omega L}{R} = \frac{X_L}{R} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (26)$$

Power Used in an Inductive Circuit.—The power used at any instant in the circuit of Fig. 48, is obtained by multiplying the instantaneous value of e by that of i; it is shown by the power curve in Fig. 51. For this circuit it is evident that the power is sometimes negative, i.e., the circuit, instead of drawing power from the line, is actually furnishing power to the line. Energy which has been stored in the magnetic field of the inductance is flowing back into the source of power supply.

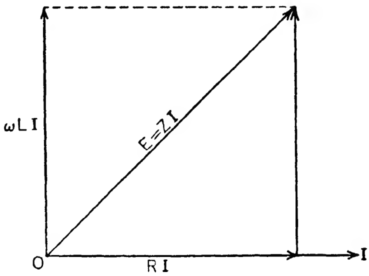


Fig. 50.—Vector diagram for an a.c. circuit containing inductance and resistance in series.

The expression for the power is,

$$p = E_m \sin \omega t \times I_m \sin (\omega t - \phi)$$
$$= E_m I_m \sin \omega t \times \sin (\omega t - \phi).$$

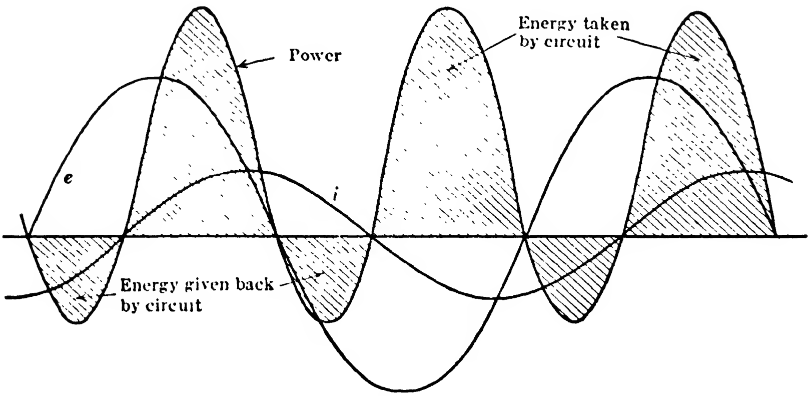


Fig. 51.—Power curve for an a.c. circuit containing inductance and resistance.

The average value of this expression is given by the average value of the expression

$$p = \frac{E_m I_m}{2} \cos \phi \, (1 - \cos 2\omega t), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

because the average value of $E_m I_m \sin \omega t \cos \omega t \sin \phi$ is evidently zero. So average power is

$$P = EI \cos \phi. \qquad (28)$$

The power in the circuit is equal to the product of the volts and amperes in the circuit and the quantity $\cos \phi$. For this reason $\cos \phi$ is called the *power factor* of the circuit; it may have any value between unity and zero. In ordinary power circuits it has a value between about 0.7 and 0.95, very seldom being unity. In some parts of efficient radio circuits the power factor may be as small as 0.005.

The power may be expressed in terms of current and resistance by changing the form of Eq. (28).

$$P = EI \cos \phi = EI \frac{R}{\sqrt{R^2 + (\omega L)^2}}$$

$$= I \times I \sqrt{R^2 + (\omega L)^2} \frac{R}{\sqrt{R^2 + (\omega L)^2}} = I^2 R. \qquad (29)$$

This equation for the power used in an a.c. circuit is really a definition of the effective resistance of the circuit; the resistance of the circuit, for alternating current, may be entirely different from the c.c. resistance of the circuit. There are many effects which combine to make the a.c. resistance sometimes several times as great as the c.c. resistance (or the a.c. resistance may be negative even, while the c.c resistance is positive) and the only way[1] of measuring this resistance is by use of Eq. (29). The power used in the circuit is measured by a wattmeter, the current by an ammeter, and the resistance found by the relation

$$\text{Effective resistance} = \frac{\text{watts}}{I^2}. \qquad (30)$$

Wattmeter.—It is generally not possible to measure the power used in an a.c. circuit by use of Eq. (28) because the phase difference of the voltage and current is not known and there is no easy method of measuring it directly. To get the power used in an a.c. circuit it is nearly always necessary to use a *wattmeter*. This is an indicating instrument, resembling an ammeter or voltmeter externally, but differing from these instruments in that it has two independent electrical circuits. Two coils inside the instrument, one in shunt with the circuit, and one in series with it, react on one another to produce the force which moves the indicating pointer.

[1] If the circuit is such that a measurement in an alternating current Wheatstone bridge is feasible, of course such method also is available. Even in the bridge determination the idea expressed in Eq. (29) is, however, involved.

The theory involved in its operation is explained in practically any text on a.c. measurements and will not be given here. The scale of the meter is calibrated directly in watts and, with a properly calibrated instrument, the reading of power is accurate no matter what the power factor may be; for very small power factors, and for circuits of frequency much higher than that for which the meter is intended, a correction may be necessary.[1]

The power factor of an a.c. circuit is then determined from the readings of three instruments, ammeter, voltmeter, and wattmeter. The power factor, cos ϕ, is the quotient of the wattmeter reading by the product of the readings of the other two instruments. If it is desired to know the angle ϕ itself, it is only necessary to consult a table of natural cosines.

It is generally not feasible to use a wattmeter to make measurements of power at radio frequency; the errors are too large. For the low-frequency power supply of radio transmitters, however, wattmeter measurements are always used. It is to be pointed out that the rectifying apparatus generally interposed betweeen radio sets and the a.c. power supply introduce distortions in voltage or current (or both) of the power supply. No matter how badly distorted the voltage and current may be, however, the wattmeter readings will be a correct measure of the power actually supplied.

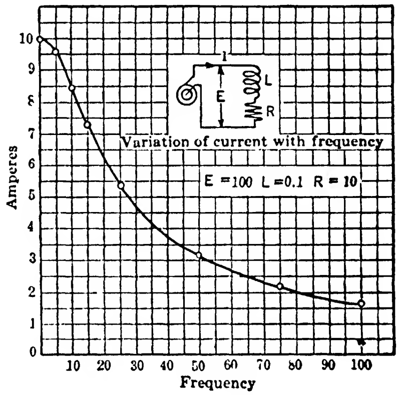


Fig. 52.—Current variation with frequency in an a.c. circuit containing inductance and resistance in series.

The power factor, as obtained by the ratio of watts to volt-amperes, cannot logically be interpreted as the cosine of an angle however. When distorted voltages and currents are involved, the power factor must be thought of as simply the ratio of the wattmeter reading to the product of the voltmeter and ammeter readings.

The effective resistance of the circuit is obtained by finding the quotient of the wattmeter reading and the square of the ammeter reading. As stated before, this resistance will generally be very different from the resistance measured by a c.c. test.

Variation of Current with Frequency in an Inductive Circuit.—The magnitude of the current flowing in a circuit consisting of a resistance and inductance in series evidently depends upon the frequency (see Eq. (24)).

At zero frequency (continuous current) this equation reduces to $I =$

[1] See Morecroft's "Laboratory Manual of Alternating Currents," p. 11.

E/R. This relation holds good only after the switch has been closed long enough for the transient condition to disappear (see Fig. 33).

At very high frequencies the resistance becomes negligible compared to the reactance, and so the value of the current is given, very nearly, by the equation $I=E/\omega L$. As the frequency varies between high and low values, voltage being held constant, the current varies as shown in Fig. 52; for frequencies sufficiently high that R is small compared to ωL, the curve approximates a hyperbola,

$$I\times f=\frac{E}{2\pi L}. \quad \quad (31)$$

Transient Current in a Circuit Having Inductance and Resistance.—After the switch has been closed for some time there is always a definite relation between the instantaneous values of the current and voltage; for every cycle the two go through exactly corresponding values. Thus in Fig. 53, when e has a maximum value AC, the current has the value AB, and whenever the voltage has the value AC the current will have the value AB. Now suppose the switch to be closed when the voltage has the value AC; the current should have the value AB, but in an inductive circuit the current cannot rise instantaneously; this was shown by the oscillograms in Figs. 33 and 36. The complete equation for the current in an inductive circuit must therefore include a transient term as well as the term for the steady state; it is properly written

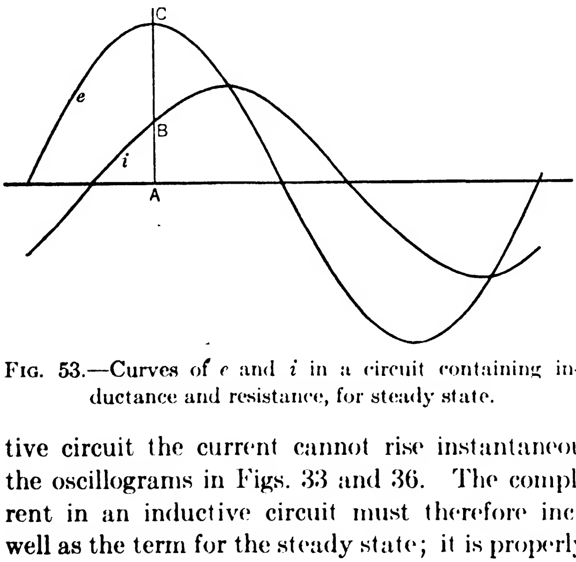


FIG. 53.—Curves of e and i in a circuit containing inductance and resistance, for steady state.

$$i=\frac{E_m}{\sqrt{R^2+(\omega L)^2}}\sin(\omega t-\phi)+K\epsilon^{-\frac{Rt'}{L}}. \quad \quad (32)$$

The second part of the current, $K\epsilon^{-\frac{Rt'}{L}}$ is determined in magnitude by the value of the current, in the steady state, at the time in the cycle corresponding to the time in the cycle that the switch is closed. Thus in Fig. 54, at the time of closing the switch the current should have the value AB; this fixes the value of K in Eq. (32). In Fig. 54 are plotted the steady value of the current i, the transient current $K\epsilon^{-\frac{Rt'}{L}}$, and the actual current for the first cycle after closing the switch; this actual current is the sum of the other two.

In Figs. 55 and 56 are shown oscillograms of the current flowing in an inductive circuit for the first few cycles after the switch had been closed; in one the switch was closed at the peak of the voltage and in the other it was closed when the voltage was very nearly zero. In Fig. 55 the effect of the transient term is plain; the current (steady value) has been plotted in dotted lines, as has also the transient term, the latter having been calculated from the value of the steady current at the time the switch was closed and the L and R or the circuit. It may be seen that the actual current is correctly given by Eq. (32). In Fig. 56 the switch was closed

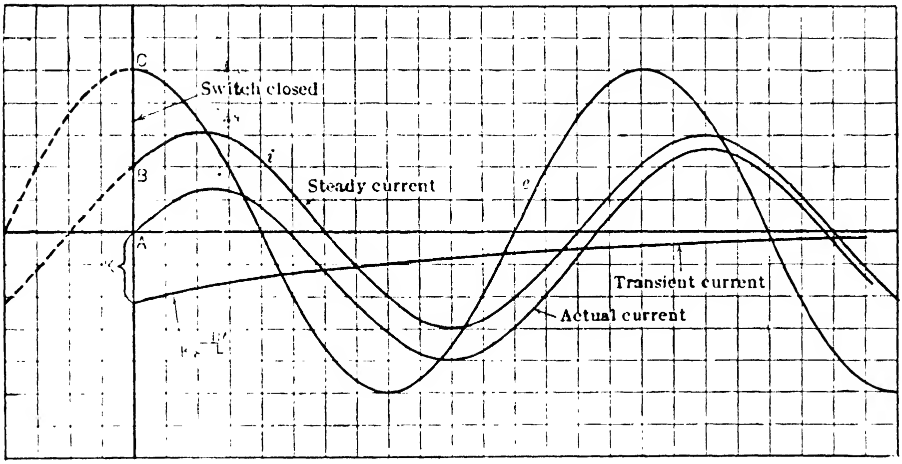


Fig. 54.—Curves of e and i in a circuit containing inductance and resistance for transient state.

at that part of the e.m.f. cycle which, in the steady state, is the proper time for the current to be zero; it is seen that for this case the transient term reduces to zero, and the actual current is represented completely by only the first term of Eq. (32).

Circuits Having Resistance and Iron-core Inductance.—In case the L of the circuit, Fig. 48, consists of an inductance having a closed iron path for its magnetic circuit, the conclusions deduced above will not be correct. The value of L in this case is not constant, but varies throughout the cycle, and for this reason the relation between the current and voltage is a complex one; the current in this case requires an equation with an infinite number of terms to express it accurately. The current, instead of being sinusoidal, has a decided hump, as shown by Fig. 57, which shows the magnetizing current of a closed-core transformer.

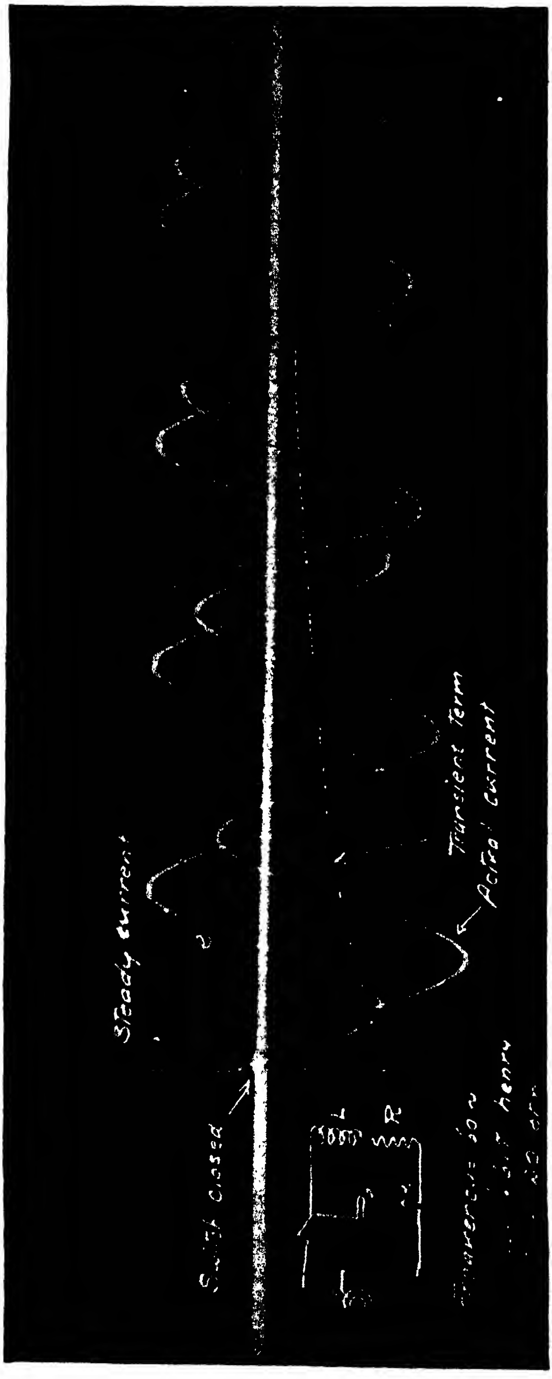


FIG. 55.—The ordinary form of current occurring when switching an inductive circuit to an a.c. supply line; analysis of the actual current into its two mathematical components is shown by the dotted curves.

Not only is the steady value of current in such a circuit irregular, but the transient current may show even greater irregularities. This irregularity may last for many cycles, depending upon the kind of iron used in the core and upon its condition of magnetization at the time the switch is closed, as well as upon the part of the cycle selected for the closing of the switch. Thus in Fig. 58 is shown the current in the primary circuit of a transformer for a few cycles after closing the switch; the transient current may be so large in this case that during the first cycle the current never reverses its direction.

The rise of current in such an inductive circuit as this is not as simple as that illustrated in Fig. 33; the analysis given in explaining this figure assumed constant L, so will not hold good if L varies during the rise of current. The actual form of rising current in such a circuit, when connected to a c.c. line, is shown in Fig. 59; it is quite evidently different from that shown in Fig. 33, which was for an air-core inductance.

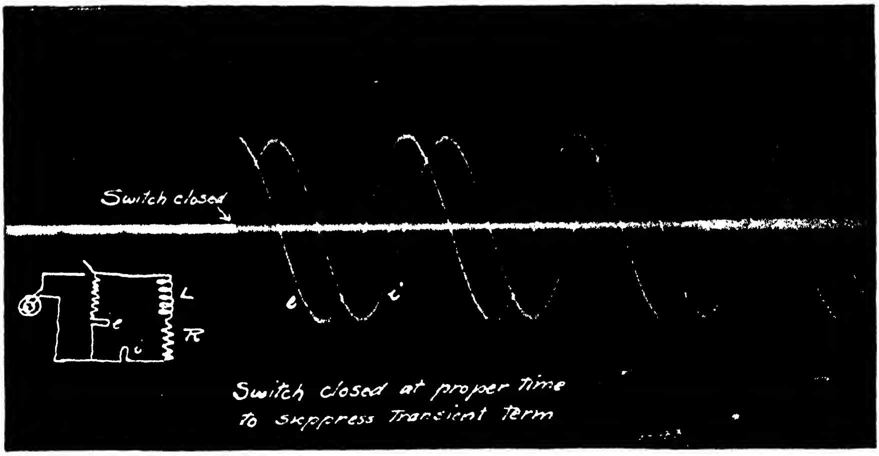


Fig. 56.—Oscillogram illustrating absence of transient current in an inductive circuit.

A little reflection brings out the reason for this peculiar form. For low flux densities the iron is not saturated and so shows a high permeability. But high permeability means high self-induction, and hence a comparatively low rate of current increase. Therefore, when the current is low, directly after the switch is closed, the current rises at a slow rate. After the current has risen to a value large enough to bring the iron near its saturation condition, the self-induction of the circuit, especially for *current changes*, is low and so the current rises rapidly.

Coefficient of Self-induction for Current Changes.—As has just been pointed out, the self-induction of an iron-core coil varies as the current rises, being large for the smaller currents and diminishing as the core becomes

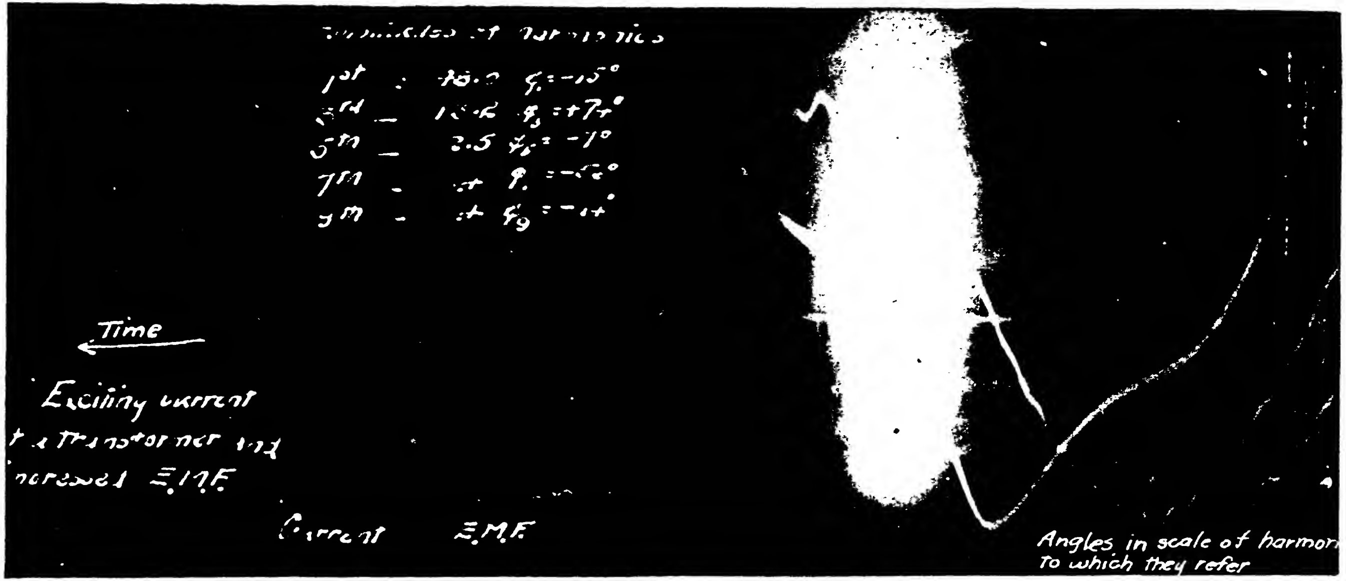


FIG. 57.—Curve of e.m.f. wave (with "ripples") and the current it produces in an iron core inductance. This complex current has been analyzed into its various harmonic components by the Fourier series method. It contains about 28 per cent of third harmonic, 5 per cent of fifth harmonic, and smaller percentages of the still higher harmonic frequencies.

magnetically saturated. An ordinary definition of L, the coefficient of self-induction, gives its magnitude in terms of the interlinkages per ampere,

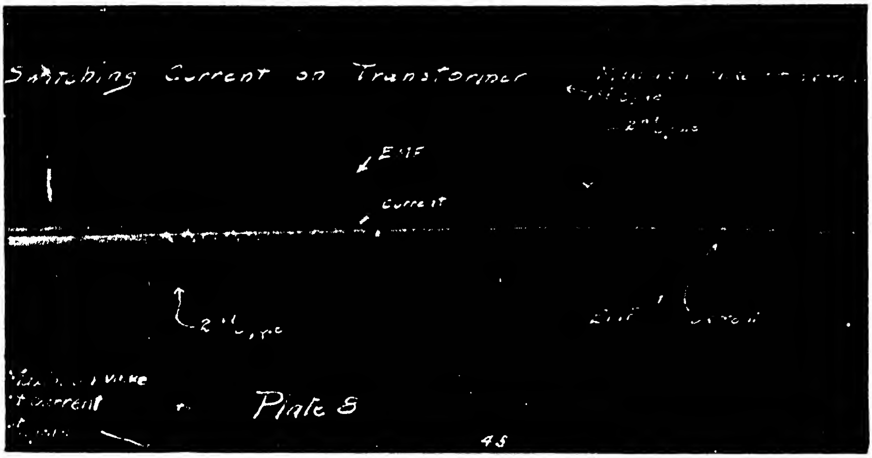


FIG. 58.—Oscillogram showing the transient current when switching an iron-core inductance to an a.c. line.

the coil having an inductance of 1 henry if 1 ampere through the coil sets up 10^8 interlinkages.

Let us imagine a coil of 1000 turns, giving a flux-current relation as

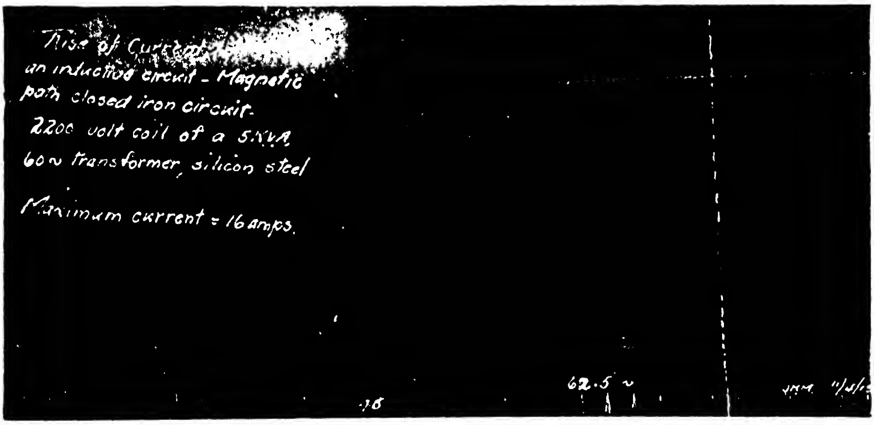


FIG. 59.—Peculiar growth of current when an iron-core inductance is switched to a source of continuous e.m.f.

shown in Fig. 60. With 1 ampere there are 2×10^5 lines of flux; this amount of flux, through 1000 turns, gives 2×10^8 interlinkages, so the coil has 2 henries of self-induction. With 2 amperes there are 3×10^5 flux lines and the self-induction is correspondingly 1.5 henries. In a similar way, at 3 amperes the coil shows 1.1 henries, at 4 amperes 0.85 henry, and at 5 amperes it shows 0.68 henry. These values of self-induction have been plotted in Fig. 60.

Now if such an iron-core coil is to be used as a " choke," its performance will be much different from what this current-self-induction curve would lead one to suppose. A choke coil is used to suppress the fluctuations, or ripples, in a unidirectional current; such coils are always necessary in the power pack of a radio set, in which alternating current is rectified into pulsating unidirectional current, and the pulses then smoothed out by a filter using iron-core coils.

Let us suppose such a current as that shown in Fig. 61, having an average value of 4.5 amperes, but fluctuating between 4 and 5 amperes. Now according to ordinary theory these ripples in the current will be opposed by an $L\,di/dt$ voltage set up in the choke coil, but, as the current is varying,

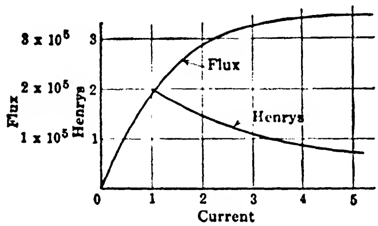


FIG. 60.—Flux-current curve for an iron core coil; the self-induction has been calculated, in terms of "interlinkages per ampere."

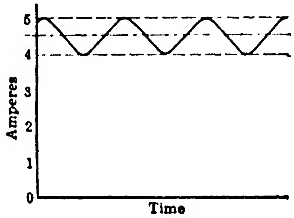


FIG. 61.—A possible current from a rectified power supply.

and correspondingly the self-induction is varying, the question arises—what value of L shall be used in the expression $L\,di/dt$? We might use the value of L at 4.5 amperes, but common sense would indicate that the average value of L should be used, which in this case would be 0.765 henry (slightly different from the value of L at the average current). Now such a value of L, substituted in the expression $L\,di/dt$, would indicate a lot of choking action, and yet it will be found experimentally that there is *practically no choking action.* This anomaly arises because of lack of appreciation of the derivation and significance of the formula used.

The real choking action is caused by *change in the flux interlinkages of the coil,* and the expression $L\,di/dt$ is derived from this fundamental concept. Now inspection of the current-flux curve of Fig. 60 shows that when the current is increased from 4 to 5 amperes there is practically *no change in flux interlinkages,* hence practically *no counter voltage* due to changing magnetic field is set up when the current varies between these two limits.

If now we put the counter voltage (due to changing magnetic field) as equal to $L\,di/dt$, we are forced to the conclusion that *for current changes*, between 4 and 5 amperes, the coil has *no self-induction.* It has a value of 0.85 henry for 4 amperes and 0.68 henry for 5 amperes, but for current changes between these limits it has no self-induction, when this is viewed from the standpoint of counter voltage set up by the changing current, that is, counter voltage $= L\,di/dt$.

Current Flow in a Condenser.—By the definition of a condenser no electrons can actually pass from one plate to the other; they are insulated from one another. If, however, a condenser is connected to a source of alternating e.m.f., current will flow in this circuit, as may be seen by the reading of an a.c. ammeter placed in series with the condenser.

Suppose a condenser of capacity C farads is connected to a line the e.m.f. of which is given by the equation $e = E \sin \omega t$. The condenser will, of course, take enough charge to bring the potential difference of its plates continually equal to that of the line to which it is connected. As this impressed e.m.f. continually varies in magnitude and direction, electrons must be continually running in and out of the condenser to maintain its plates at the proper potential difference. This continual charging and discharging of the condenser constitutes the current read by the ammeter. The electrons, the motion of which constitutes the current, do not actually pass from one plate of the condenser to the other through the dielectric; they merely flow in and out of the condenser. With this idea in mind it is easy to see why the charging current of a condenser increases with the capacity of the condenser, and with the frequency of the impressed e.m.f.

The magnitude of the charging current is obtained as follows:

The charge $q = Ce$ and the current $i = dq/dt$.

Now

$$q = CE_m \sin \omega t,$$

so

$$i = \omega C E_m \cos \omega t. \qquad (33)$$

This current is then of the same form as the impressed e.m.f. (a cosine curve is similar to a sine curve in form) but leads it by 90° as shown in Fig. 62; its maximum value, in amperes, is equal to $\omega C E_m$.

In effective values the relation between the impressed voltage and the charging current is

$$I = \omega C E = 2\pi f C E. \qquad (34)$$

It is evident that, other things being equal, the charging current of a condenser is directly proportional to the frequency of the impressed e.m.f.

This should be contrasted to the inductive circuit in which the current varies inversely as the frequency, if the resistance is small compared to the reactance.

The relation between the current and voltage may be written

$$I = E \div \frac{1}{2\pi f C} = \frac{E}{X_c}. \quad \text{(35)}$$

The quantity $\frac{1}{2\pi f C}$ is called *the reactance of the condenser*, generally specified as capacitive reactance to distinguish it from inductive reactance, $2\pi f L$.

Effect of Condensers and Coils on Wave Shape.—The voltage form of an ordinary alternator is never a pure sine wave. Its general form is that of a sine wave, but there are *ripples* in this form, as is well illustrated

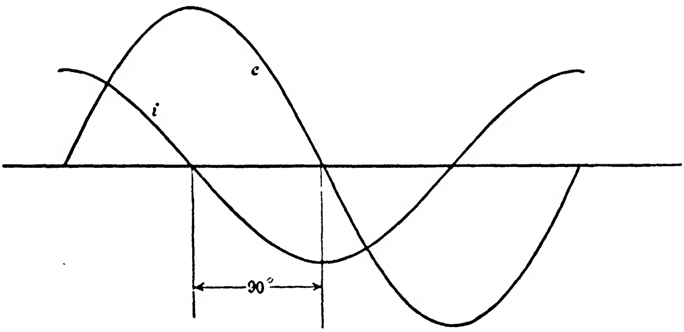


FIG. 62.—Current and voltage for a perfect condenser connected to an a.c. line.

by the voltage wave in Fig. 57. These ripples are due to the fact that the alternator generates voltages of other than its main, or supposed, frequency. Thus the voltage of Fig. 57 is that of a 60-cycle alternator, but a careful analysis of the wave shows that in addition to the 60-cycle voltage the machine was generating two other frequencies, namely, 660 cycles and 780 cycles. The magnitude of these other frequencies is small compared to that of the fundamental voltage, so that they appear as ripples on the 60-cycle sine-wave form.

Now an important question for the radio engineer to consider has to do with the shape of current which flows in a circuit connected to such an alternator. Will the current be more or less distorted than the impressed voltage? An examination of Eq. (31) shows that the current through an inductive circuit varies *inversely* with the frequency of the impressed voltage, while Eq. (35) shows that in a capacitive circuit the current varies *directly* with the frequency of the impressed voltage.

It is a fundamental property of electric circuits having constant L, C, and R, that if several voltages, of different frequencies, are simultaneously impressed, the current due to any one of these voltages can be calculated on the assumption that it alone is active in the circuit. In other words, each one of the component voltages can be treated alone, its current found, and the actual current will be the sum of the currents due to the individual voltages.

From this consideration, and from Eqs. (31) and (35), it therefore follows that, if a voltage having ripples is impressed on an inductive circuit, the current which flows will have ripples, but they will be *much less prominent* than they are in the voltage wave itself. That is, the current wave is smoothed out by the effect of inductance. In a condensive circuit, however, the ripples will be exaggerated and the current wave thus *much more distorted* than is the voltage itself.

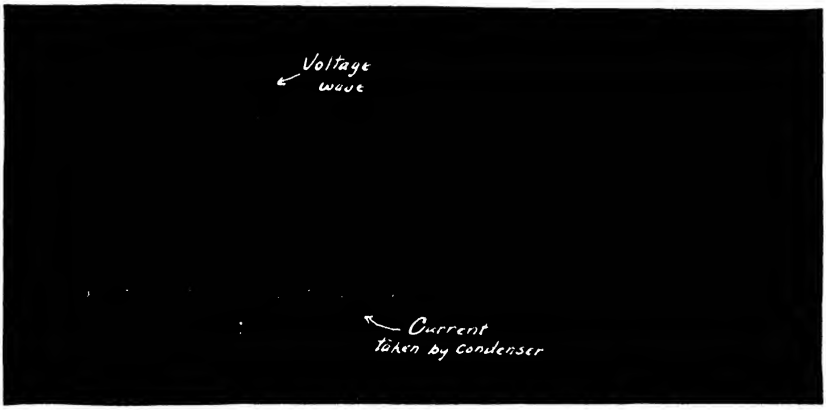


FIG. 63.—Showing the peculiar form of current in a condenser when connected to a certain commercial form of alternator.

If the voltage of Fig. 57 is impressed on an inductive circuit (having low resistance), the ripples in the current will be about one-twelfth as prominent as they are in the voltage. If the voltage is impressed on a condenser (having low effective series resistance), the ripples in the current will be about twelve times as pronounced as they are on the voltage wave. This fact is well brought out by the oscillogram of Fig. 63, which shows the charging current of a condenser, connected to a small alternator, of ordinary design. If the voltage is impressed on a resistive circuit (having neither inductance nor capacity) the current will be of just the same shape as the voltage, the ripples being just as pronounced as they are on the voltage, and no more so.

Condenser and Resistance in Series.—If a condenser and resistance are connected in series and a sine wave of voltage is impressed, a sinusoidal current will flow; its magnitude and phase depend upon the

R, C, E, and f of the circuit. Suppose this current to be given by $i = I_m \sin \omega t$.

The resistance drop $= I_m R \sin \omega t$.

The capacity reactance drop, in magnitude, is $\frac{I_m}{\omega C} \cos \omega t$. But as shown before, the current leads the voltage impressed on a condenser; the capacity drop is therefore properly written,

$$\text{Capacity drop} = -\frac{I_m}{\omega C} \cos \omega t.$$

The impressed voltage must be the sum of the drop over the resistance and that over the condenser and is so shown in Fig. 64. The current leads

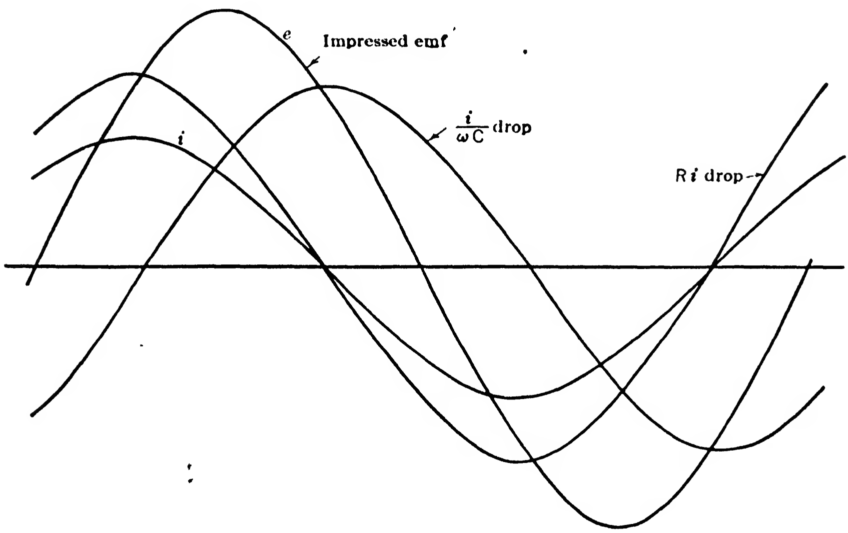


FIG. 64.—Voltage and current curves for circuit containing R and C in series.

the impressed voltage by the angle ϕ, the magnitude of which is fixed by the relative magnitudes of the reactance and resistance drops.

The three curves of Fig. 64 are represented vectorially in Fig. 65, effective values being used instead of maximum values. From this vector diagram we have

$$E^2 = (IR)^2 + \left(\frac{I}{\omega C}\right)^2,$$

or

$$I = \frac{E}{\sqrt{R^2 + \left(\frac{1}{\omega C}\right)^2}} = \frac{E}{Z}. \quad . \quad . \quad . \quad . \quad . \quad (36)$$

and

$$\tan \phi = \frac{\frac{1}{\omega C}}{R} = \frac{1}{\omega C R}. \qquad (37)$$

The current in the circuit, as shown in Eq. (36), evidently depends upon the frequency; its variation as the frequency is changed is shown in Fig. 66. At very high frequency the current approaches the value E/R, the capacity reactance being negligible, while at zero frequency the current is zero, the condenser being equivalent to an open circuit.

Transient Current in a Circuit Consisting of Resistance and Condenser in Series.—In general there will be a transient current when switching

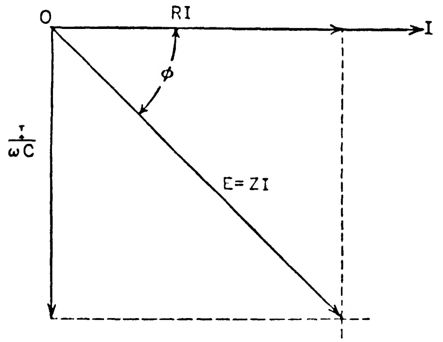


FIG. 65.—Vector diagram of voltages and current for circuit containing R and C.

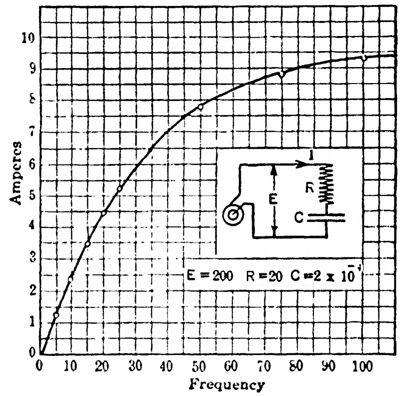


FIG. 66.—Variation of current with frequency in circuit containing R and C in series.

such a circuit to an a.c. line; the duration of the transient term is so short, however, on all commercial circuits that an oscillogram shows the current rising immediately to its proper value, this being fixed by the time on the e.m.f. cycle that the switch is closed.

Current Flow in a Circuit Having Resistance, Inductance, and Capacity in Series.—The current flowing in the circuit shown in Fig. 67 will require three components of e.m.f.; the resistance drop IR, the inductance drop $2\pi fLI$, and the capacity drop $\frac{I}{2\pi fC}$. The resistance drop is in phase with the current, the inductance drop is 90° ahead of the current and the capacity drop is 90° behind the current. These three components of the impressed e.m.f. are shown vectorially in Fig. 68. The two reactance drops evidently tend to neutralize one another.

The total reactance drop

$$=2\pi fLI-\frac{I}{2\pi fC}. \qquad (38)$$

The resultant required impressed voltage is seen to be

$$E=I\sqrt{R^2+\left(2\pi fL-\frac{1}{2\pi fC}\right)^2}, \qquad (39)$$

and the magnitude of the current may be written

$$I=\frac{E}{\sqrt{R^2+\left(\omega L-\frac{1}{\omega C}\right)^2}}=\frac{E}{Z}. \qquad (40)$$

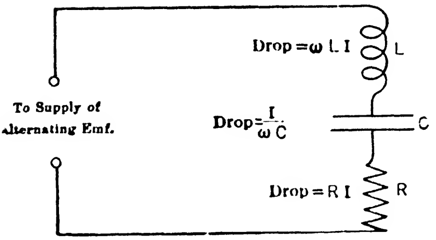


FIG. 67.—Circuit containing R, L, and C in series.

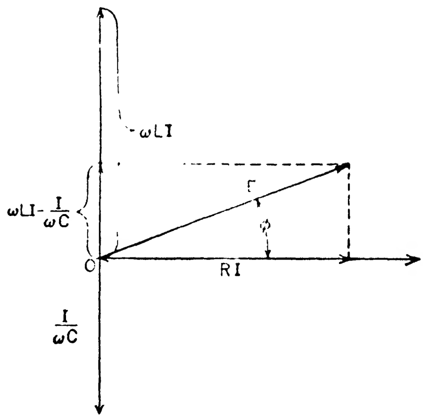


FIG. 68.—Vector diagram of voltages and current for circuit containing R, L and C in series.

The phase difference between impressed voltage and current is fixed by the equation

$$\cos\phi=\frac{R}{Z} \quad \text{or} \quad \tan\phi=\frac{\omega L-\frac{1}{\omega C}}{R}. \qquad (41)$$

The reactance of the circuit may either be positive or negative, according to which component of the reactance predominates. If $2\pi fL$ is greater than $\frac{1}{2\pi fC}$ the reactance is positive and the current lags, whereas if the capacity reactance is the greater, the current leads the impressed e.m.f.

The magnitude of the current will evidently depend upon the frequency and will have about the form shown in Fig. 69. At zero frequency the condenser offers infinite reactance so the current is zero; at infinitely high frequency the inductance reactance becomes so great that again the current approaches zero; at some intermediate frequency the inductance reactance just balances the capacity reactance so that *the total reactance is zero.* For this frequency the current has a maximum value, as shown for frequency f_r in Fig. 69. The form of this curve could have been predicted by considering the two curves given in Figs. 52 and 66.

Resonance.—For such a circuit as shown in Fig. 67 there will always be one frequency which will give a total reactance zero; this will be true no matter what values of L and C may be chosen. At this frequency the current will be in phase with the e.m.f. and its magnitude will be a maximum, being limited only by the resistance of the circuit, $I=E/R$.

The frequency at which this occurs is called the *resonant frequency* of the circuit; it is at this frequency that most radio circuits are operated.

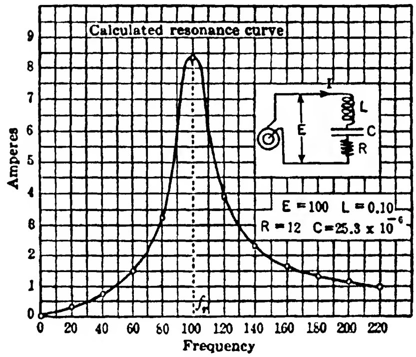


FIG. 69.—Variation of current with frequency in circuit containing R, L, and C in series.

Unless care is exercised when performing experiments on resonance the condensers used in the circuit will be spoiled by the puncturing of the dielectric at the resonant frequency. For any frequency whatever the drop across the condenser is fixed by the relation,

$$E_c=\frac{I}{2\pi fC}.$$

If we substitute in this equation the value of the current, in terms of impressed voltage and resistance we get, at resonance,

$$E_c=E\frac{1}{2\pi fCR}=E\frac{X_c}{R}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (42)$$

As the value of X_c/R may be much greater than unity, so the voltage across the condenser may be many times as great as the impressed voltage; in a certain laboratory circuit used in performing low-frequency resonance tests the drop across the condenser at resonant frequency is *eighteen times as great as the impressed voltage.* At this frequency the drop across the inductance is equal to that across the condenser, but this excessive voltage across the inductance coil will generally do no harm. In a very efficient radio circuit it is possible to have the drops across the inductance and condenser each 400 times as great as the impressed voltage.

Actually the circuit of Fig. 69 is generally made up of a coil in series with a condenser only; no resistance is actually added in the circuit. The resistance R, therefore, is to be regarded as the resistance of the coil itself. (The condenser always has some small value of equivalent series resistance, but at power frequencies it is so small compared to the resistance of the coil that it can generally be neglected.)

Then $\dfrac{R}{\sqrt{R^2+\overline{2\pi fL}^2}}$ is the power factor of the coil. Now in radio circuits the reactance of the coil, $2\pi fL$, is much greater than its resistance, R, on the average perhaps fifty to one hundred times as great. Then we may write, with negligible error

$$\frac{R}{\sqrt{R^2+\overline{2\pi fL}^2}}=\frac{R}{2\pi fL}=\cos\phi_L.$$

Even if $2\pi fL$ is only ten times as great as R, the error resulting from this approximation is only $\frac{1}{2}$ per cent.

At resonance $\dfrac{1}{2\pi fC}=2\pi fL$ so we may write, for Eq. (42)

$$E_c=E\times\frac{2\pi fL}{R}=E\frac{1}{\cos\phi_L}. \qquad (43)$$

The series resistance of the condenser has been considered negligible, as is generally permissible.

From this equation we get the convenient relation that at resonance the voltage across the condenser is equal to the voltage impressed on the circuit multiplied by the reciprocal of the power factor of the coil.

Resonant Frequency.—A circuit is said to be resonant when the reactance is zero. Therefore we have for the resonant frequency

$$2\pi fL=\frac{1}{2\pi fC}$$

from which we get the value of the resonant frequency

$$f=\frac{1}{2\pi\sqrt{LC}} \qquad (44)$$

In this equation L must be in henries, C in farads, and f will be in cycles per second. As the microfarad is the usual unit of capacity a more convenient form is

$$f = \frac{1000}{2\pi\sqrt{LC}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (45)$$

C being in microfarads. In determining this frequency the separate values of L and C do not matter; the product LC is the quantity which fixes the critical frequency. That is, a circuit having $L = 0.24$ henry and $C = 10$ microfarads will be resonant at the same frequency as one which has $L = 0.06$ henry and $C = 40$ microfarads.

In the accompanying table (pp. 82–83) is shown a set of inductance and capacity values which produce resonance, in the range of frequencies

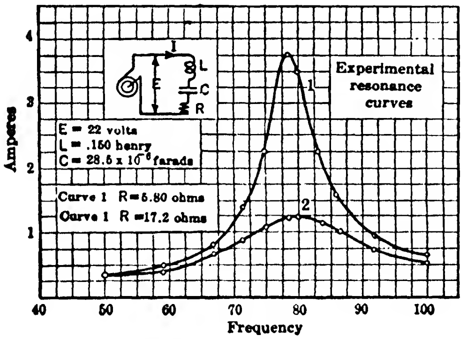


Fig. 70.—Effect of resistance on resonance curve.

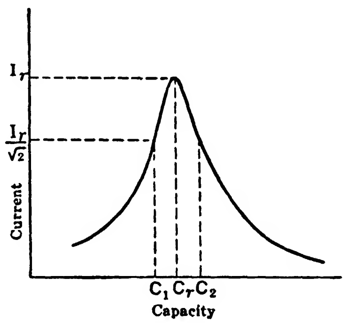


Fig. 71.—Variation of current with capacity in a resonant circuit.

generally encountered in radio practice. For ordinary tuned circuits most suitable values of capacity will be obtained by selecting such values of L and C that the inductance, in microhenries, is about the same as the capacity in micro-microfarads. Thus for a frequency of 500 kc., use 338 microhenries and 0.0003 microfarad.

The sharpness of the resonance curve is determined by the resistance of the circuit; the inductance being fixed, the less the resistance the more sharply defined is the resonant frequency and the larger is the current at the resonant frequency. In Fig. 70 are shown the resonance curves obtained for a circuit having $L = 0.15$ henry and $C = 28.5$ microfarads. The one curve shows the variation of current with a circuit resistance of 5.8 ohms and the other shows the same thing after the resistance had been increased to 17.2 ohms.

In a low-resistance circuit the resonance is said to be sharp and in a high-resistance circuit it is said to be flat or dull.

Frequency in Kilocycles per Second	Wave Length in Meters	Capacity in Microfarads									
		0.0001	0.0002	0.0003	0.0004	0.0005	0.0006	0.0007	0.0008	0.0009	0.001
30,000	10	0.282	0.141	0.094	0.070	0.056	0.047	0.0402	0.0352	0.0312	0.0282
20,000	15	0.633	0.317	0.211	0.158	0.127	0.106	0.0900	0.0790	0.0704	0.0633
15,000	20	1.13	0.563	0.375	0.281	0.225	0.188	0.161	0.141	0.125	0.113
10,000	30	2.53	1.27	0.845	0.633	0.507	0.422	0.362	0.317	0.281	0.253
7,500	40	4.50	2.25	1.50	1.13	0.901	0.751	0.644	0.563	0.500	0.450
6,000	50	7.04	3.52	2.35	1.76	1.41	1.17	1.01	0.880	0.782	0.704
5,000	60	10.1	5.07	3.38	2.53	2.03	1.69	1.45	1.27	1.13	1.01
4,286	70	13.8	6.89	4.60	3.45	2.76	2.30	1.97	1.72	1.53	1.38
3,750	80	18.0	9.00	6.00	4.51	3.60	3.00	2.57	2.25	2.00	1.80
3,333	90	22.8	11.4	7.60	5.70	4.56	3.80	3.26	2.85	2.53	2.28
3,000	100	28.2	14.1	9.38	7.04	5.63	4.69	4.02	3.52	3.12	2.82
2,000	150	63.3	31.7	21.0	15.8	12.7	10.6	9.00	7.90	7.04	6.33
1,500	200	113	56.3	37.5	28.1	22.5	18.8	16.1	14.1	12.5	11.3
1,000	300	253	127	84.5	63.3	50.7	42.2	36.2	31.7	28.1	25.3
750.0	400	450	225	150	112	90.1	75.1	64.4	56.3	50	45.0
600.0	500	704	352	235	176	141	117	101	88.0	78.2	70.4
500.0	600	1,010	507	338	253	203	169	145	127	113	101
428.6	700	1,380	689	460	345	276	230	197	172	153	138
375.0	800	1,800	900	600	451	360	300	257	225	200	180
333.3	900	2,280	1,140	760	570	456	380	326	285	253	228
300.0	1,000	2,820	1,410	939	704	563	469	402	352	312	282
250.0	1,200	4,050	2,030	1,350	1,010	810	675	579	507	450	405
214.3	1,400	5,520	2,760	1,840	1,380	1,100	920	788	690	613	552
187.5	1,600	7,210	3,600	2,400	1,800	1,440	1,200	1,030	901	801	721
166.7	1,800	9,120	4,560	3,040	2,280	1,820	1,520	1,300	1,140	1,010	912

150.0	2,000	11,300	5,630	3,750	2,820	2,250	1,880	1,610	1,410	1,250	1,130
136.4	2,200	13,600	6,810	4,540	3,410	2,730	2,270	1,950	1,700	1,510	1,360
125.0	2,400	16,200	8,110	5,410	4,050	3,240	2,700	2,320	2,030	1,800	1,620
115.4	2,600	19,000	9,520	6,340	4,760	3,810	3,170	2,720	2,380	2,120	1,900
107.1	2,800	22,100	11,000	7,360	5,520	4,410	3,680	3,150	2,760	2,450	2,210
100.0	3,000	25,300	12,700	8,450	6,340	5,070	4,220	3,620	3,170	2,820	2,530
85.71	3,500	34,500	17,200	11,400	8,620	6,900	5,750	4,930	4,310	3,830	3,450
75.00	4,000	45,000	22,500	15,000	11,300	9,010	7,510	6,450	5,630	5,010	4,500
66.67	4,500	57,000	28,500	19,000	14,300	11,400	9,500	8,140	7,130	6,330	5,700
60.00	5,000	70,400	35,200	23,500	17,600	14,100	11,700	10,100	8,800	7,820	7,040
54.54	5,500	85,200	42,600	28,400	21,300	17,000	14,200	12,200	10,600	9,460	8,520
50.00	6,000	101,000	50,700	33,800	25,300	20,300	16,900	14,500	12,700	11,300	10,100
46.15	6,500	....	59,600	39,600	29,700	23,800	19,800	17,000	14,900	13,200	11,900
42.86	7,000	....	69,000	46,000	34,500	27,600	23,000	19,700	17,200	15,300	13,800
40.00	7,500	....	79,200	52,800	39,600	31,700	26,400	22,600	19,800	17,600	15,800
37.50	8,000	....	90,000	60,000	45,000	36,000	30,000	25,700	22,500	20,000	18,000
35.29	8,500	....	102,000	67,800	50,900	40,700	33,900	29,100	25,400	22,600	20,300
33.33	9,000	....	114,000	76,000	57,000	45,600	38,000	32,600	28,500	25,300	22,800
31.58	9,500	....	....	84,700	63,500	50,800	42,300	36,300	31,800	28,200	25,400
30.00	10,000	....	....	93,800	70,400	56,300	46,900	40,200	35,200	31,300	28,200
27.27	11,000	....	....	114,000	85,200	68,100	56,800	48,700	42,600	37,900	34,100
25.00	12,000	....	....	....	101,000	81,100	67,600	57,900	50,700	45,000	40,500
23.08	13,000	....	....	....	....	95,200	79,300	68,000	59,500	52,900	47,600
21.43	14,000	....	....	....	....	110,000	92,000	78,800	69,000	61,300	55,200
20.00	15,000	....	....	....	....	....	106,000	90,500	79,200	70,400	63,300
18.75	16,000	....	....	....	....	....	....	103,000	90,100	80,100	72,100
17.65	17,000	....	....	....	....	....	....	....	102,000	90,400	81,400
16.67	18,000	....	....	....	....	....	....	....	114,000	101,000	91,200
15.79	19,000	....	....	....	....	....	....	....	....	113,000	102,000
15.00	20,000	....	....	....	....	....	....	....	....	....	113,000

Value of inductance, in microhenries, to produce resonance with various values of capacity, throughout the radio spectrum.

Series Resonance with Varying Capacity—Decrement.—If the frequency impressed on the circuit of Fig. 67 is held constant and the capacity or inductance varied, resonance curves similar to those in Fig. 69 will be obtained except that the variables will be different. Suppose such a curve has been obtained, as shown in Fig. 71. We shall now show how the shape of the curve depends upon the resistance and how to actually calculate the value of this resistance from the shape of the curve, provided that the value of L is known.

The quantity which is actually determined from the resonance curve is the ratio $R/2fL$, f being the resonance frequency of the circuit. This ratio is called the *decrement* of the circuit, for reasons which will be apparent when the subject of oscillations is discussed.

Referring to Fig. 71, let C_r be the capacity which gives resonance, the current for this value of capacity being I_r. Let C_1 and C_2 be the two values of capacity, one greater than C_r and the other less than C_r, which serve to reduce the current to $I_r \div \sqrt{2}$ or 0.707 I_r. When the capacity has the value C_r there is no effective reactance in the circuit, so we have, for

$$C = C_r, \qquad I_r = \frac{E}{R},$$

and for

$$C = C_2, \qquad I = \frac{E}{\sqrt{R^2+X_2^2}} = .707\ I_r$$

which can be true only on condition that $X_2 = R$, or

$$2\pi fL - \frac{1}{2\pi fC_2} = R. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (46)$$

For $C = C_1$,

$$I = \frac{E}{\sqrt{R^2+X_1^2}} = .707\ I_r$$

which can be true only if

$$X_1 = R \quad \text{or} \quad -\left(2\pi fL - \frac{1}{2\pi fC_1}\right) = R. \quad . \quad . \quad . \quad . \quad (47)$$

The capacity reactance is greater than the inductive reactance for C_1 and less than the inductive reactance for C_2, hence the reversal of the signs in front of the reactance terms in Eqs. (46) and (47).

Adding (46) and (47) we get

$$\frac{1}{2\pi fC_1} - \frac{1}{2\pi fC_2} = 2R.$$

Multiplying this equation through by $\frac{1}{2\pi fL}$, we get

$$\frac{1}{(2\pi f)^2LC_1}-\frac{1}{(2\pi f)^2LC_2}=\frac{2R}{2\pi fL}$$

or

$$\frac{1}{(2\pi f)^2L}\left(\frac{C_2-C_1}{C_2C_1}\right)=\frac{2R}{2\pi fL}. \qquad (48)$$

Now if C_2 and C_1 do not differ from C_r very much (say 10 per cent) we may put without appreciable error

$$C_2C_1=C_r^2. \qquad (49)$$

This is, of course, an approximation, and is more nearly true the sharper the resonance curve. We may now put

$$\frac{1}{(2\pi f)^2LC_r}\left(\frac{C_2-C_1}{C_r}\right)=\frac{2R}{2\pi fL}. \qquad (50)$$

But $(2\pi f)^2=\frac{1}{LC_r}$, as may be seen by writing the equation for resonance,

$$f=\frac{1}{2\pi\sqrt{LC_r}},$$

C_r being the value of the capacity which gives resonance.

So (50) becomes

$$\frac{C_2-C_1}{C_r}=\frac{2R}{2\pi fL}$$

or

$$\frac{R}{2fL}=\frac{\pi}{2}\frac{C_2-C_1}{C_r}. \qquad (51)$$

As an illustration of the application of this formula suppose that the resonant capacity for a certain circuit is 32 microfarads and that the values of C_2 and C_1 are 34 microfarads and 30.2 microfarads respectively. Then for this circuit the decrement, generally designated by the Greek letter δ, is

$$\delta=\frac{R}{2fL}=\frac{\pi}{2}\frac{34-30.2}{32}=0.187.$$

The decrement may also be calculated from a resonance curve plotted with frequencies as abcissas as given in Fig. 72; we have derived the formula when capacity is used for abscissas because such is generally the case in radio measurements. If, however, frequency is used as abscissas, the frequency having been varied in getting the resonance curve, L and C having been maintained constant, the derivation of δ from the half energy points of the resonance curve is as follows:

$$2\pi f_1 L - \frac{1}{2\pi f_1 C} = -R,$$

$$2\pi f_2 L - \frac{1}{2\pi f_2 C} = R.$$

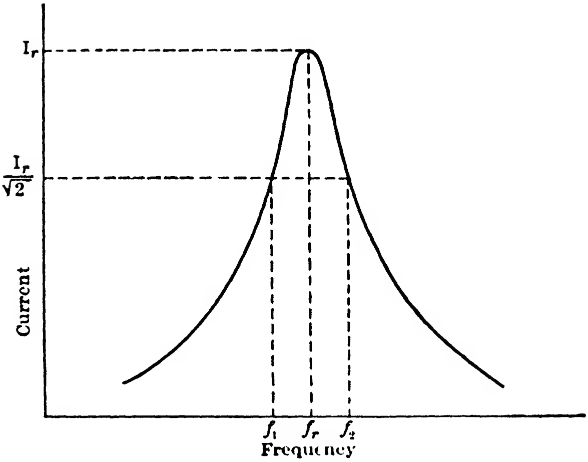


FIG. 72.—Variation of current with frequency in a resonant circuit.

To eliminate C from these two equations, multiply them by $2\pi f_1 C$ and $2\pi f_2 C$, respectively, and get the two equations

$$(2\pi f_1)^2 LC - 1 = -R2\pi f_1 C,$$

$$(2\pi f_2)^2 LC - 1 = R2\pi f_2 C.$$

Put these in the forms

$$C\left\{(2\pi f_1)^2 L + 2\pi f_1 R\right\} = 1,$$

$$C\left\{(2\pi f_2)^2 L - 2\pi f_2 R\right\} = 1.$$

Combining

$$(2\pi f_1)^2 L + 2\pi f_1 R = (2\pi f_2)^2 L - 2\pi f_2 R,$$

$$R(2\pi f_1 + 2\pi f_2) = L\left\{(2\pi f_2)^2 - (2\pi f_1)^2\right\}.$$

So

$$\frac{R}{2L}=\frac{2\pi(f_2{}^2-f_1{}^2)}{2(f_2+f_1)}=\pi(f_2-f_1).$$

Dividing by f_r, the resonant frequency,

$$\frac{R}{2f_rL}=\delta=\pi\frac{f_2-f_1}{f_r}. \qquad (52)$$

For a given circuit $\frac{f_2-f_1}{f_r}$ is approximately equal to $\frac{1}{2}\frac{C_2-C_1}{C_r}$. This follows from the relation between frequency and capacity; to produce a certain small percentage change in the natural frequency of a circuit it is necessary to change the capacity of the circuit by twice this amount, the frequency varying not with the capacity, but with the square root of the capacity.

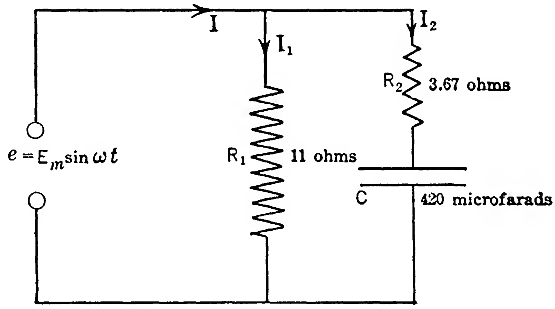


FIG. 73.—Parallel circuits.

Flow of Current in Parallel Circuits and Relation of Line Current to Branch Currents.—When a circuit consists of two or more branches in parallel the line current cannot be obtained by calculating the branch currents and adding them arithmetically as is done in continuous current circuits, because of the difference in phase of the various branch currents. The line current, instead of being equal to the arithmetical sum of the branch currents, may be even smaller than either of the branch currents and, in fact, is so in many radio circuits. It is necessary to calculate not only the magnitude of the different branch currents, but also their phase; these branch currents are then added *vectorially* to give the line current.

Suppose a circuit made up as shown in Fig. 73, the current I_1 being 10 amperes, in phase with the line voltage and the current I_2 being 15 amperes, leading the line voltage by 60°; the line current will be the vector sum of 10 and 15 as shown in Fig. 74. It proves to be 21.8 amperes. The angle of the lead is found by the relation of the reactive and active components of the line current (the active component of a current is that component which is in phase with the voltage and the reactive component is that which is 90° out of phase with the voltage). I_1 has no reactive component and so contributes 10 amperes to the active component of the line current only; I_2 has a reactive component equal to 15 sin 60° or 13 amperes, and an active component of 15 cos 60° or 7.5 amperes. The total active line current is therefore 17.5 amperes and the reactive component is 13 amperes. The angle of lead of the line current is then $\tan^{-1}$ 13/17.5 or 36.6°.

If the impressed voltage is 110 volts the impedance of the combined circuit is equal to 110/21.8 or 5.05 ohms.

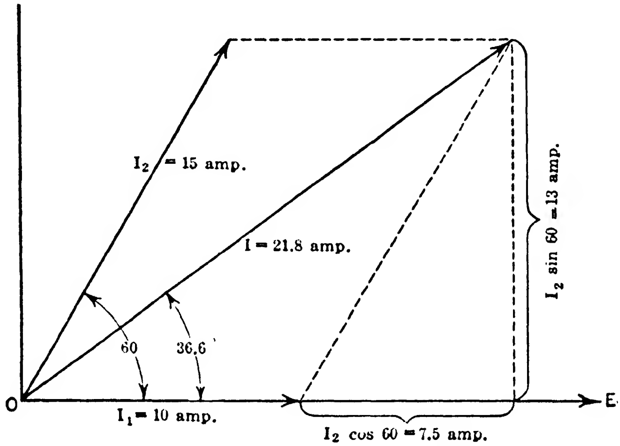


FIG. 74.—Vector diagram of currents in the parallel circuit shown in Fig. 73.

The equivalent resistance is $Z \cos \phi = 5.05 \cos 36.6° = 4.06$ ohms.

The equivalent reactance is $Z \sin \phi = 5.05 \sin 36.6° = 3.02$ ohms.

The equivalent series condenser of the combined circuit is found by putting the reactance equal to $\frac{1}{2\pi fC'}$. If the frequency of the supply is 60 cycles this gives

$$\frac{1}{2\pi fC'} = 3.02 \text{ ohms} \quad \text{or} \quad C' = 882 \text{ microfarads.}$$

Hence the branched circuit shown in Fig. 73 is exactly equivalent to the single circuit shown in Fig. 75, for the frequency assumed; for a different frequency other values of equivalent resistance and equivalent capacity would be obtained. A more detailed analysis of a branched circuit, using complex quantities, is given elsewhere.

In case the branched circuit is more complex than that given in Fig. 73, such as that given in Fig. 76, the branched part must first be replaced by its equivalent single circuit, calculated as shown for Fig. 73;

the resistance and reactance of this equivalent circuit must then be added to the resistance and reactance of R_1 and L_1. By vectorially combining this total resistance and reactance the impedance of the simple equivalent circuit is obtained.

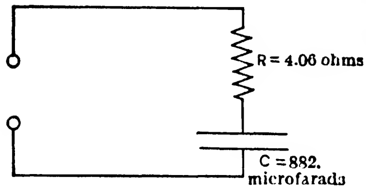


FIG. 75.—Simple series circuit equivalent to parallel circuit of Fig. 73.

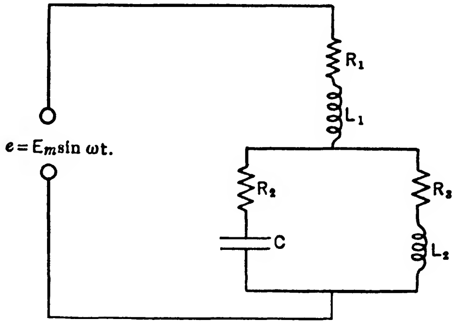


FIG. 76.—Series-multiple circuit.

Impedance of a Circuit Made Up of *L*, *R*, and *C*, in Series.—The reactance of this circuit is calculated by finding the sum of the inductive and capacitive reactances at all the frequencies necessary; the equivalent resistance of this circuit is independent of frequency and equal at all frequencies to the actual resistance, R. The several quantities are shown

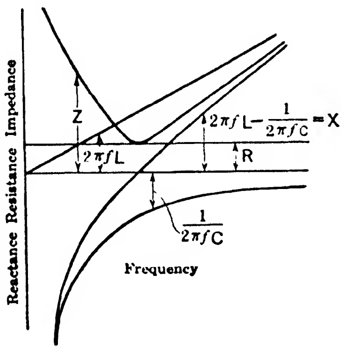


FIG. 77.—Variation of reactance, resistance and impedance with frequency in a circuit containing L, R and C in series.

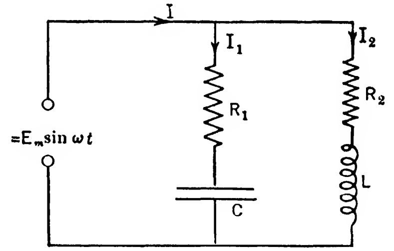


FIG. 78.—Branched circuit, having L and R in one branch and C and R in the other.

in the form of curves in Fig. 77. The reactance, $\frac{1}{2\pi fC}$, is shown negative; the total reactance, X, is negative at frequencies lower than the resonant value and positive above this value. The impedance is positive for all values of frequency, having its minimum value when the total reactance X, is zero, then being equal to R.

The current leads the voltage for frequencies lower than the resonant value and lags behind the voltage for higher frequencies.

Impedance of a Branch Circuit, having *L* and *R* in One Branch and *C* and *R* in the Other.—The simplest way of comprehending the impedance of this complex path, Fig. 78, is to calculate for each value of frequency the magnitude and phase of the current in each branch. The active and reactive components of the two branch currents are then calculated. The active component of line current is found by adding the two active branch currents, and the reactive component of the line current is found by adding the reactive branch currents. These additions are to be alge-

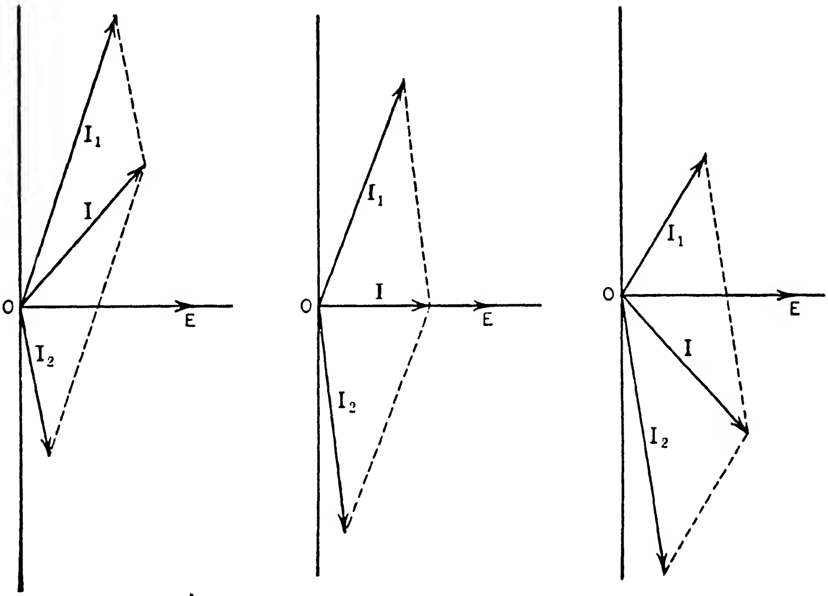


FIG. 79. FIG. 80. FIG. 81.

FIG. 79.—Vector diagram for circuit of Fig. 78, line current leading.

FIG. 80.—Vector diagram for circuit of Fig. 78, line current in phase with impressed e.m.f.

FIG. 81.—Vector diagram for circuit of Fig. 78, line current lagging.

braic; in the case of the active current the algebraic sum is the arithmetic sum, but the reactive current in the line is the *difference* of the reactive currents of the branches.

In Fig. 79 is shown the vector diagram for frequency above the resonant frequency of the circuit; the line current in this case leads the voltage so the equivalent simple circuit would consist of a condenser in series with a resistance, the two having such values that when the simple circuit was connected to a line voltage *E*, the current flowing would be equal, in magnitude and phase, to *I* of Fig. 79.

In Fig. 80 is shown the condition when impressed frequency is so adjusted that the reactive currents in each branch neutralize each other; in this case the simple circuit would consist of a resistance only. The resistance of the simple circuit would, in general, *be many times as great* as the resistance in the actual branched circuit.

In Fig. 81 the frequency is supposed lower than the resonant frequency, the current taken by the inductive branch being greater than that taken by the capacitive branch; the equivalent simple circuit for this case would consist of a resistance in series with an inductance.

The above simple analysis shows that the branched circuit of Fig. 78 may be represented by a single circuit, but the constants of this single circuit must be made to vary as the frequency is varied.

The equivalent R may be obtained by calculating the I^2R loss in each branch and adding to give the total loss in the circuit; this total loss, divided by the square of the line current (obtained vectorially as shown in Figs. 79-81), gives the equivalent resistance of the combination.

The equivalent inductance or capacity is obtained by calculating the reactive component of the impressed e.m.f.; this equals $E \sin \phi$ where E is the value of the impressed voltage and ϕ is the angle between the impressed voltage and the line current. This value of e.m.f., $E \sin \phi$, is put equal to $2\pi fL'I$, where L' is the equivalent inductance and I is the line current.

In case the line current is leading $\sin \phi$ is negative and the equivalent inductance would be negative. In this case the reactive component of the impressed voltage, $E \sin \phi$, is put equal to $\frac{I}{2\pi fC'}$, where C' is the equivalent series capacity of the circuit.

The circuit is analyzed exactly most easily by the use of complex algebra, a method of treatment explained in all standard texts on alternating currents.

Let Z_1 = impedance of branch $1 = R_L + j\omega L$;

Z_2 = impedance of branch $2 = R_C - j\frac{1}{\omega C}$;

Z = impedance of the joint path.

$$\frac{1}{Z} = \frac{1}{Z_1} + \frac{1}{Z_2} = \frac{1}{R_L + j\omega L} + \frac{1}{R_C - j\frac{1}{\omega C}}$$

$$= \frac{\left(R_C - j\frac{1}{\omega C}\right) + (R_L + j\omega L)}{(R_L + j\omega L)\left(R_C - j\frac{1}{\omega C}\right)}.$$

Hence

$$Z=\frac{(R_L+j\omega L)\left(R_C-j\frac{1}{\omega C}\right)}{(R_L+R_C)+j\left(\omega L-\frac{1}{\omega C}\right)}. \qquad (53)$$

Rationalize by multiplying numerator and denominator by

$$(R_L+R_C)-j\left(\omega L-\frac{1}{\omega C}\right).$$

Collecting terms we have

$$Z=\frac{\left\{\begin{array}{l}R_C(R_L^2+\overline{\omega L}^2)+R_L\left(R_C^2+\left(\frac{1}{\omega C}\right)^2\right)\\ \qquad +j\left[\omega L\left(R_C^2+\left(\frac{1}{\omega C}\right)^2\right)-\frac{1}{\omega C}(R_L^2+\overline{\omega L}^2)\right]\end{array}\right\}}{(R_L+R_C)^2+\left(\omega L-\frac{1}{\omega C}\right)^2} \qquad (54)$$

$$=\frac{\left\{\begin{array}{l}R_CR_L(R_C+R_L)+R_C\overline{\omega L}^2+R_L\left(\frac{1}{\omega C}\right)^2\\ \qquad +j\left[R_C^2\omega L-R_L^2\frac{1}{\omega C}-\frac{L}{C}\left(\omega L-\frac{1}{\omega C}\right)\right]\end{array}\right\}}{(R_L+R_C)^2+\left(\omega L-\frac{1}{\omega C}\right)^2}. \qquad (55)$$

Of this complex impedance the real part is the effective resistance of the branched circuit and the imaginary part is the reactance, or $\omega L'$, where L' is the effective inductance. So,

$$R'=\frac{R_CR_L(R_C+R_L)+R_C(\omega L)^2+R_L\frac{1}{(\omega C)^2}}{(R_L+R_C)^2+\left(\omega L-\frac{1}{\omega C}\right)^2}. \qquad (56)$$

$$\omega L'=\frac{R_C^2\omega L-R_L^2\frac{1}{\omega C}-\frac{L}{C}\left(\omega L-\frac{1}{\omega C}\right)}{(R_L+R_C)^2+\left(\omega L-\frac{1}{\omega C}\right)^2}. \qquad (57)$$

In case the resistance of the inductance is large compared to that of the condenser, Eqs. (56) and (57) may be simplified to the approximately correct forms

$$R'=\frac{R}{m^2R^2\frac{C}{L}+(m^2-1)^2}. \qquad (58)$$

Also

$$\omega L' = -\omega L \frac{R^2 \frac{C}{L} + (m^2 - 1)}{m^2 R^2 \frac{C}{L} + (m^2 - 1)^2}, \quad \ldots \ldots \quad (59)$$

and for resonance Eq. (58) gives

$$R' = \frac{L}{C}\frac{1}{R'}, \quad \ldots \ldots \quad (60)$$

where R' = equivalent series resistance; $\omega L'$ = equivalent series reactance; R = total actual resistance in the circuit, that is, resistance of the inductive branch plus that of the capacitive branch; C = actual capacity of the capacitive branch; L = actual inductance of the inductive branch; m = ratio of the impressed frequency to the resonant frequency of the circuit $= 2\pi f\sqrt{LC}$.

In case L' comes out a negative quantity it may be converted to its equivalent series capacity by the relation

$$C' = 1/(2\pi f)^2(-L'). \quad \ldots \ldots \quad (61)$$

An interesting condition obtains in a circuit having parallel resonance. Thus suppose that the values of L and C and the frequency of supply for the circuit of Fig. 82 have been so adjusted that for a voltage impressed across A–B the circuit shows no reactance; the power factor is unity and the circuit shows resistance only.

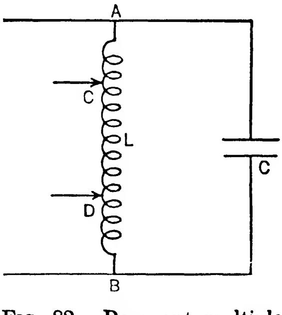


Fig. 82.—Resonant multiple circuit.

If the supply voltage is impressed *across any other two points in the circuit*, the circuit will be approximately in resonance for these points also; if, for example, the voltage is impressed across points C–D, the circuit will show resistance only.

The resistance will not be the same when measured between points C–D as it is for the points A–B. It may be proved that the resistance between any two points in the circuit is nearly proportional to the square of the reactance included between the two points, in either branch. The reactance in each branch of the parallel circuit will be the same, no matter where the two points are taken, but the reactance will be inductive in one branch and capacitive in the other.

There will generally be more or less mutual induction between the parts of the inductance which the new connection gives. Thus in Fig. 82, if the power supply is connected to points A–D, the inductance from A to D will be in one of the paths, while that between D and B will be in the other. The effect of this mutual induction will complicate somewhat the quantitative relations, so for the moment we neglect it.

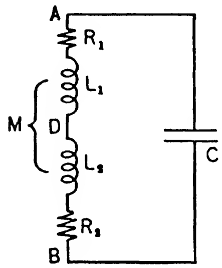


FIG. 83.—A tapped coil may be considered as two separate coils with mutual induction between the two parts.

On the basis of no mutual induction between the two parts of the coil, in the two separate paths, we can reach the general conclusion that as the points of power supply (Fig. 82), are moved closer together, the resistance, at resonant frequency, between these two points decreases. Quantitatively it decreases with the square of the amount of reactance between the points of power supply. Thus suppose in Fig. 83 $M \doteq 0$ and that $L_1 = L_2$. With the power supply at points A–B, and frequency at its resonant value (line power factor equal to unity) the line resistance becomes, from Eq. (56)

$$R'_{A-B} = \left(\frac{L_1 + L_2}{C}\right)\frac{1}{R_1 + R_2}. \qquad (62)$$

If the power supply is connected across only one coil, say, at A–D, the line resistance becomes

$$R'_{A-D} = R'_{A-B}\left(\frac{L_{A-D}}{L_{A-B}}\right)^2 = \frac{1}{4}R'_{A-B}. \qquad (63)$$

Effect of Mutual Induction in Parallel Circuits.—In general there will be mutual induction between coils L_1 and L_2 of Fig. 83, and this mutual induction will appreciably affect the resistance and resonant frequency of the circuit.

It is first to be noted that where the two inductances L_1 and L_2 are really two parts of the same continuous coil (point D merely being a tap on the coil), the effect of M is an additive one, that is, the total inductance from A to B is greater than the sum of L_1 and L_2. It might seem that when L_1 is in one path and L_2 in the other path, that the effect of mutual induction becomes complicated, but actually it is not so. Then it might seem that whereas M has a certain action when L_1 and L_2 are in series, in the same path, when they are in different paths (power supply at A and D), the action of M would be different because possibly of different magnitude and phases of currents in the two paths.

However, on the basis that the reactance of either path is large compared to its resistance, practically always so in radio circuits, the currents through L_1 and L_2 will be equal in magnitude when these are in different paths. The reactance of one path (directly from A to D) will be inductive, and the reactance of the other path (from A to D through C) will be capacitive, but these two reactances will be equal in magnitude. Furthermore the current in one path will lag behind the line voltage by nearly 90°, and in the other path it will lead by about 90°. This means that in so far as their magnetic action on each other is concerned, due to their mutual induction, the two coils, L_1 and L_2, will act practically the same whether they are in different paths or in the same path, hence the resonant frequency of the circuit stays practically the same for the connection A–D as for the connection A–B.

The resistance of the combination at the resonant frequency when connected across the taps A–B is evidently

$$R'_{A-B}=\frac{L_1+L_2+2M}{C}\,\frac{1}{R_1+R_2}. \qquad (64)$$

For connection to points A–D, at resonant frequency the line resistance approaches the value $R'_{A-B}\left(\frac{L_1}{L_1+L_2}\right)$ as the coupling approaches unity.

Fig. 84 illustrates another combination of inductance and condensers; such a circuit is used in one of the common forms of radio telephone apparatus. The frequency of current in the closed circuit is fixed by the resonant period of this circuit, that is, $f=\frac{1}{2\pi\sqrt{LC}}$ where $C=\frac{C_1C_2}{C_1+C_2}$. The alternating current supply for the circuit is furnished across the condenser C_1, and the power factor of this circuit (i.e., between points A and B) is unity; *the impedance offered to the supply circuit is resistance only.* If the point B is moved around the circuit so as to include part of the inductance L in either path, as shown at B', the impedance between the two points A and B' would still be resistance only.

It is often desired in radio circuits to alter the impedance of the circuit to which the power is supplied. Thus in certain vacuum-tube circuits a resonant circuit (as shown in Fig. 84) is used as load for the tube output and, to get the maximum output from the tube, the circuit must offer resistance only (no reactance), and this resistance must have a proper value. Evidently such a circuit as that shown in Fig. 84 offers such possibility; by properly adjusting the position of B' the desired resistance will be obtained.

We might keep B' fixed and vary the value of C_1; varying the value of C_1, however, has the disadvantage of changing also the frequency of the circuit. If C_1 is held constant and the point B is moved along the inductance, the effective resistance between points A and B will vary while the frequency is maintained practically constant. Such a connection scheme is generally used in practice.

The circuit of Fig. 83, when the mutual induction is taken into account, requires rather cumbersome equations for its analysis. This circuit is shown, in rearranged form, in Fig. 85. On the basis that the reactances

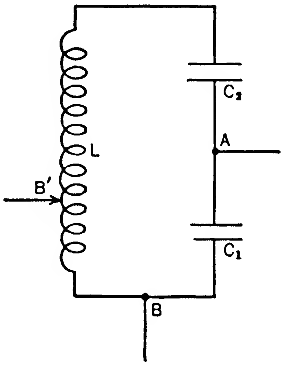


FIG. 84.—Resonant multiple circuit used in a certain radio-telephone set.

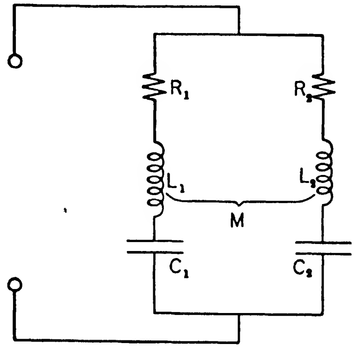


FIG. 85.—In the general case of branched circuits there may be mutual induction between the two branches.

of the coils are large compared to their respective resistances, it can be shown[1] that the line resistance is given by the equation

$$R'=\frac{\left\{\begin{array}{c}\left[R_1R_2-\left(\omega L_1-\dfrac{1}{\omega C_1}\right)\left(\omega L_2-\dfrac{1}{\omega C_2}\right)+\overline{\omega M}^2\right](R_1+R_2)\\ +\left[R_2\left(\omega L_1-\dfrac{1}{\omega C_1}\right)+R_1\left(\omega L_2-\dfrac{1}{\omega C_2}\right)\right]\\ \left(\omega L_1-\dfrac{1}{\omega C_1}+\omega L_2-\dfrac{1}{\omega C_2}+2\omega M\right)\end{array}\right\}}{(R_1+R_2)^2+\left(\omega L_1-\dfrac{1}{\omega C_1}+\omega L_2-\dfrac{1}{\omega C_2}+2\omega M\right)^2} \quad . \quad . \quad (65)$$

And if the impressed frequency is adjusted to give the circuit unity power factor for the impressed voltage this reduces to

$$R'=\frac{R_1R_2+\overline{\omega M}^2-\left[\left(\omega L_1-\dfrac{1}{\omega C_1}\right)\left(\omega L_2-\dfrac{1}{\omega C_2}\right)\right]}{R_1+R_2} \quad . \quad . \quad . \quad . \quad (66)$$

[1] See article by Heising in Journal A.I.E.E., May, 1920.

Eq. (66) may be used to get the resistance of two parallel paths which have no condensers by putting $\frac{1}{\omega C}=0$. If, in the resulting equation, we put $M=0$, we shall get an equation of form similar to Eq. (56). If in this Eq. (56) we suppose the condenser to be replaced by a coil L_2, the resulting equation will be just the same as that derived from Eq. (65) by putting $\frac{1}{\omega C_1}=\frac{1}{\omega C_2}=M=0$.

Experimental Tests of Parallel Resonant Circuits.—To illustrate this idea of parallel resonance with experimental data, using ordinary a.c. measuring instruments, a simple test was performed. An inductance coil, having a tap near its center, was connected to a condenser as shown in Fig. 86 and resonant frequency was impressed across A–C. The values of inductance and capacity were about as shown in the diagram, the resistance of the coil being about 10.8 ohms. A watt-meter, volt-meter, and ammeter were used to measure the input; the frequency was held constant at 45 cycles, which is the frequency to give resonance for $L=0.628$ henry and $C=20$ microfards. The effective resistance was calculated by dividing the wattmeter reading by the square of the ammeter reading. The results obtained are tabulated below:

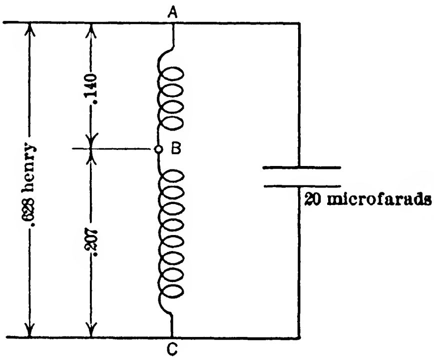


FIG. 86.—Experimental resonant multiple circuit similar to that of Fig. 84.

	Volts	Amperes	Watts	Effective Resistance, Ohms
Terminals C–A	105	0.050	5.0	2000
Terminals C–B	105	0.145	14.8	705
Terminals B–A	100	0.170	16.6	575

These values of resistance are nearly proportional to the square of the value of reactance between the respective terminals; better results cannot be obtained by this method, because the current taken by a condenser exaggerates the non-sinusoidal form of the impressed voltage and so may differ quite appreciably from the true sine form. In parallel resonance the very small minimum line current obtained is a result of the inductive and capacitive currents of the two branches neutralizing each other; if,

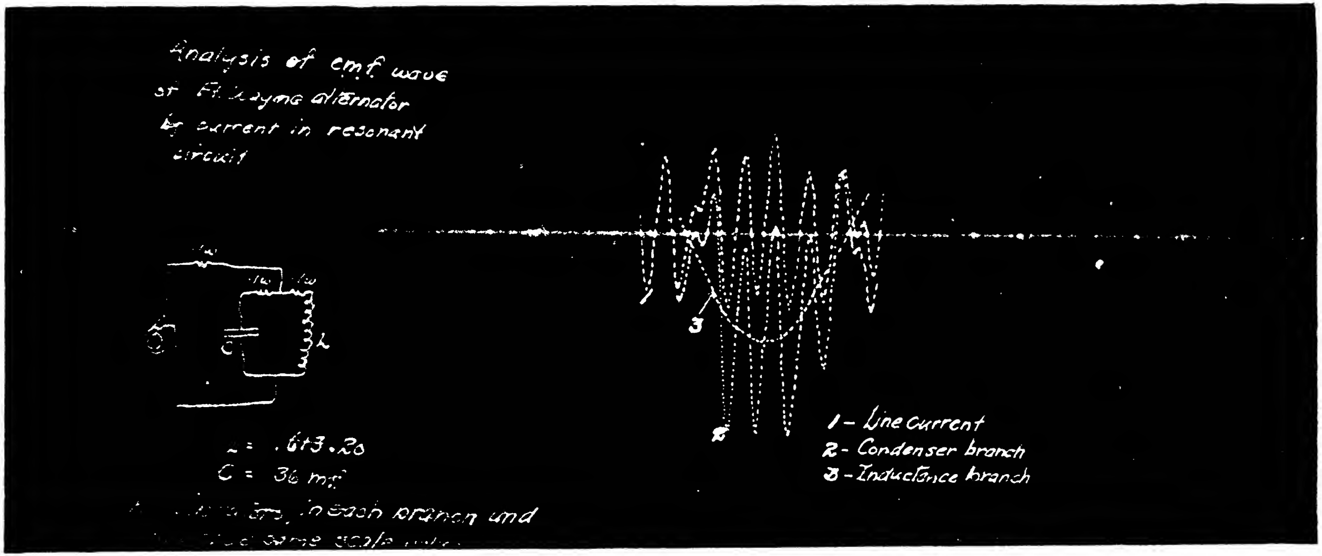


FIG. 87.—Oscillogram of currents in a multiple resonant circuit, showing effect of complex wave of e.m.f.; the current in the inductive branch is shown 180° out of its proper phase.

however, the two currents are not of the same form, it is evident that the neutralization cannot be very complete and the line current at resonance will not be as small as it should normally be.

A case of this kind is shown clearly in the oscilligram of Fig. 87; the generator supplying the power was of an ordinary commercial type, having, however, a rather smaller air gap than is usual. The inductance and capacity were connected in parallel and the impressed frequency was varied until the line current showed a minimum value. The form and phase of the currents in the two branches of the circuit are well shown on the film, and it is at once evident that the great difference in form of the two currents would prevent the resonance phenomena being very marked. Probably 50 per cent of the current flowing in the condenser circuit is of some frequency much higher than that for which the circuit was resonant and at least this much current would persist in the supply line no matter how carefully the circuit was adjusted for resonance.

This question of upper harmonics is often of much importance in the operation of radio apparatus; more specific mention of the occurrences will be made when discussing certain types of radio generators.

In another test the measurement of a circuit was made in a Wheatstone bridge, by which scheme of measurement a skilled observer may reach determinations independent of upper harmonic disturbances. A trained ear may balance the bridge for the fundamental frequency in spite of the fact that the upper harmonics are producing much noise in the telephones due to their unbalance. The circuit measured was like that pictured in Fig. 84, with the following constants: $L_1=181$ μh, $L_2=862$ μh, $R_1=0.65$ ohm, $R_2=1.97$ ohms, $M=267$ μh, $C=19.8$ μf, and $f=1000$ cycles. The value calculated from Eq. (67), after putting $\dfrac{1}{\omega C_2}=0$, gives 15.8 ohms, whereas the measured value was 16.1 ohms.

It is possible to move points A and B' (Fig. 84) to such a position that there is no reactance in either path. In this case we have a maximum possible line current (for a given impressed voltage) and the resistance of the combination is a minimum. It is equal to the resistance of one path divided by two, if the two paths have equal resistances; if not it is equal to the reciprocal of the sum of the reciprocals of the resistances in the separate paths.

Experimental Curves of Parallel Resonance.—In Figs. 88 and 89 are some experimental curves showing the characteristics of parallel resonance; they were obtained by ammeter, voltmeter, and wattmeter readings, frequency being varied and impressed voltage being held constant. The equivalent resistance was obtained directly by dividing the wattmeter reading by the squared value of the line ammeter reading; the equivalent inductance or capacity was found after calculating the reactive component of the impressed voltage and knowing the line current from the ammeter reading. The alternator used had a very pure sine wave of e.m.f. compared to that given by the average machine.

A rather extraordinary effect is seen in these curves; the equivalent series resistance at resonance is higher the lower the actual resistance of the circuit. Thus in the first case where the actual resistance was 6 ohms the equivalent resistance has a maximum value of 320 ohms; in the second case where the actual resistance has been increased to 16 ohms the maximum value of R is only 240 ohms. In neither case is the equivalent

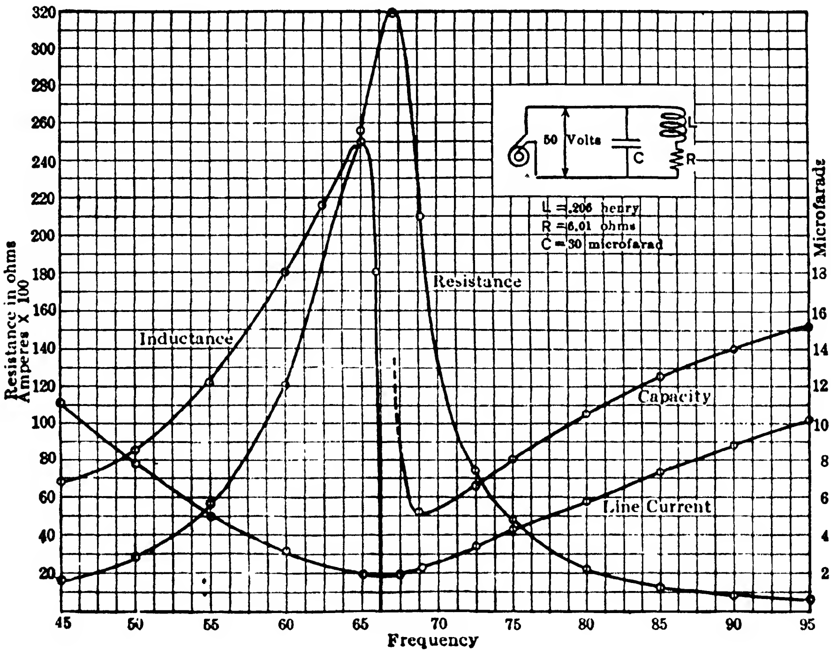


FIG. 88.—Reactance and resistance curves for a parallel resonant circuit having low resistance.

resistance nearly as great as calculation by Eq. (56) would indicate; the reason for this discrepancy lies in the method of measurement which involves an error depending upon the non-sinusoidal form of the voltage impressed on the circuit as outlined above.

It will be noticed from the curves given in Figs. 88 and 89 that the effective inductance of a coil may be increased by putting a condenser in parallel with the coil; the equivalent resistance of the coil also increases and this increase rapidly grows larger as the amount of capacity shunting the coil is increased.

For the frequencies far removed from the resonant frequency of the circuit $\left(\text{so that } (m^2-1) \text{ is large compared to } m^2R^2\frac{C}{L}\right)$ we get rather simple formulas for the equivalent inductance and resistance of the coil. Formulas (58) and (60) in this case reduce to the forms

$$R' = \frac{R}{(m^2-1)^2}. \qquad (67)$$

$$\omega L' = -\frac{\omega L}{m^2-1}. \qquad (68)$$

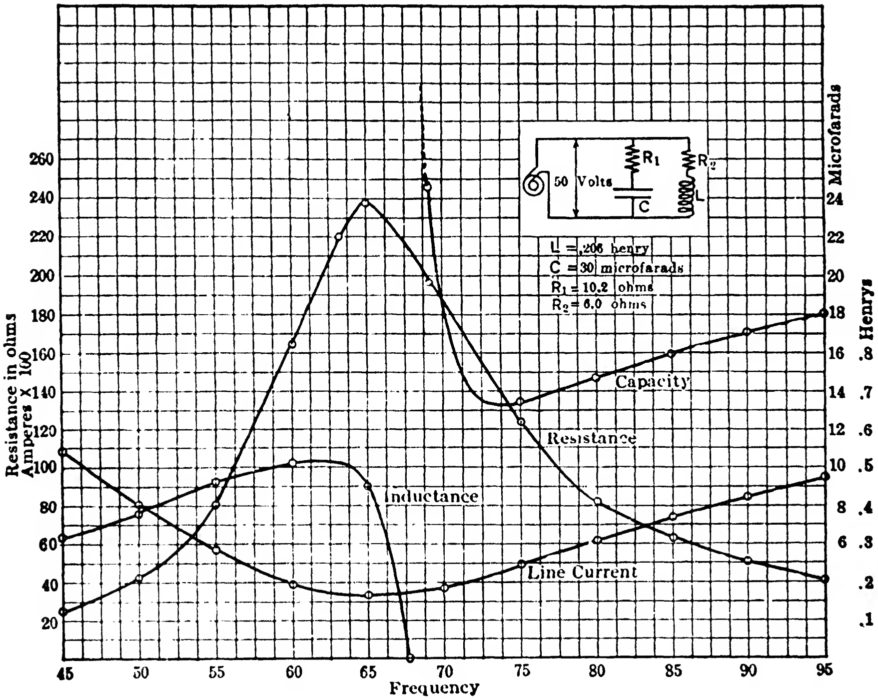


FIG. 89.—Effect of increasing the resistance in a parallel resonant circuit; compare with curves of Fig. 88.

Inspection of these equations shows that the effective resistance of the circuit rises more rapidly than does the effective inductance, especially as the resonant frequency is approached.

Effect of Harmonics on Parallel Resonance.—In attempting to check the theoretically derived formulas for parallel resonance by experimental results, such as those given in Figs. 88 and 89, it will be found that the discrepancy is very large, the difference between theory and measurement being not a few per cent, but frequently several hundred per cent; thus the resonant frequency resistance of the circuit shown in Fig. 88 should be (by Eq. 60) 1142 ohms, whereas it actually measured 320 ohms, about one-quarter the proper value.

This discrepancy is caused by slight deviations from sinusoidal form of the voltage wave of the alternator furnishing the power; the one used in getting the results of Fig. 88 had an almost perfect sine wave, so nearly so that careful observation of the oscillogram showed practically no upper harmonics in its form. The manufacturer guaranteed less than 1 per cent deviation from sinusoidal form.

To bring out the reason for the wide discrepancy between test results and theory we will assume that the alternator to be used in the test has 12 armature teeth per pair of poles, thus giving in its wave form an eleventh and thirteenth harmonic. We will assume that each of these harmonics is present with an amplitude equal to 1 per cent that of the fundamental frequency. Using the circuit given in Fig. 90 (practically the same as that used in Fig. 88), we calculate that each branch has a reactance of about 82 ohms at resonant frequency. As the condenser is assumed to have no series resistance, and as the coil has a resistance very low compared to its reactance, the impedance of each path at resonant frequency is practically 82 ohms. If now the fundamental voltage is 110 volts the current in each branch (of fundamental frequency) is $110/82 = 1.34$ amperes. The amount of fundamental current in the line is merely the active component of the coil current, or $1.34 \times 4/82 = 0.0655$ ampere. This is what ammeter A (Fig. 90) should read.

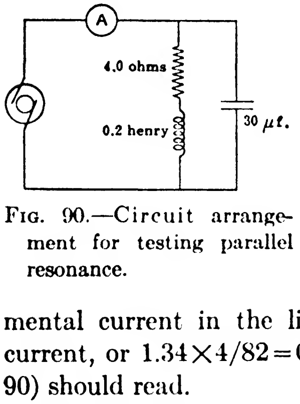


FIG. 90.—Circuit arrangement for testing parallel resonance.

Now for the eleventh and thirteenth harmonic voltages the current through the coil branch will be negligible, because the harmonic voltages are small, and the coil reactances are so high. In the condensive branch, however, the reactances are low, so appreciable harmonic current may flow in spite of the small value of the harmonic voltages; we will calculate how much it is.

Each harmonic voltage has been assumed to have an effective value of 1.1 volts (1 per cent of fundamental). The reactance of the condenser branch for the eleventh harmonic is $82/11 = 7.45$ ohms, and for the thirteenth harmonic it is 6.3 ohms. Therefore the respective harmonic currents in the condenser branch will be 0.1475 ampere and 0.1745 ampere; as these currents have no counterparts in the inductive branch (they are so small as to be negligible in that branch), they must come from the line and so flow through ammeter A (Fig. 90). This meter will then have

0.0655 ampere of fundamental frequency, 0.1475 ampere of eleventh harmonic, and 0.1745 of thirteenth harmonic. It will therefore read $\sqrt{0.0655^2+0.1475^2+0.1745^2}=0.238$ ampere. There is only 0.0655 ampere of fundamental frequency current in the line, yet the ammeter reads 0.238 ampere.

So far as fundamental frequency is concerned, the resistance of the parallel circuit is $110/0.0655 = 1680$ ohms (same as theory predicts), but the test results would show a resistance of $110/0.238 = 462$ ohms. The higher the frequency of the ripple on the voltage wave the greater is this difference between theoretical and test results.

This discrepancy can be eliminated by choking out the high-frequency currents which flow in the condensive branch. This is shown in Fig. 91, in which are given the experimental results for two circuit arrangements. In one the alternator (supposedly a sine wave) was connected directly to the resonant circuit, and in the other a large choke coil was connected in the line between circuit and alternator. With this latter connection only a few points were obtainable (owing to the large voltage drop through the choke coil), and these are plotted in the dotted line of Fig. 91.

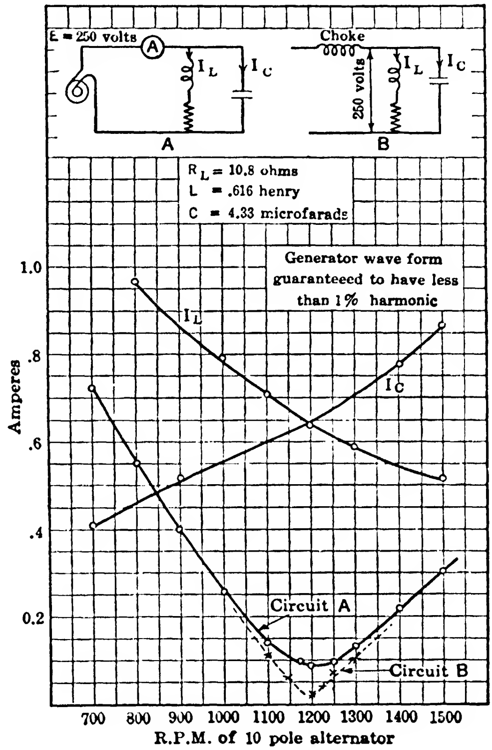


Fig. 91.—Showing how a large inductance in the power supply line of Fig. 90 diminishes very much the line current at resonant frequency; this is due to eliminating the upper harmonics of the current.

The theoretical value of circuit resistance at resonance is 13,160 ohms. The experimental data for circuit *A* give 2780 ohms, and for circuit *B* they give 11,900 ohms, practically the same as the theoretical value. It is thus evident that the upper harmonics of the alternator voltage may cause great difficulty in proving the theory of parallel circuits. At higher frequencies, where a vacuum-tube oscillator is used for power, even greater discrepancy may occur, because this type of power source gives a voltage wave having a great many harmonics. By using properly tuned circuits to eliminate them, the theory will be proved experimentally as well as it is in the alternator test described above.

Telephone Receivers as a Parallel Circuit.—The ordinary pair of head telephone receivers, such as used with many radio sets, is really a case of parallel circuits, although it is not at first evident why it should be so. Each receiver has an inductance of about one henry and a resistance of about 1000 ohms. The two connecting wires of the receivers are generally woven into one cord, perhaps 4 feet long. This cord really constitutes a condenser, not of the ordinary form, to be sure, but it is two conductors separated by an insulator, and this we know must act like a condenser. The capacity of the cord and the attached internal wires of the telephone receiver is about 10^{-10} farad, so that the telephone receiver must really be represented about as shown in Fig. 92. Such a circuit must of course have a resonant frequency, when the charging current taken by the condenser is equal to the lagging current taken by the coils of the receivers.

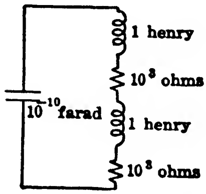


FIG. 92.—The approximate electrical equivalent of an ordinary pair of head phones.

The resistance and reactance of a typical pair of commercial head phones is shown in Fig. 93. According to Stowell[1] the various phones on the market to-day have their resonant frequencies between 9 and 15 kc., at which frequency the resistance may be as much as 200 kilohms. Above 50 kc. the phones act like a condenser of about 10^{-10} farad.

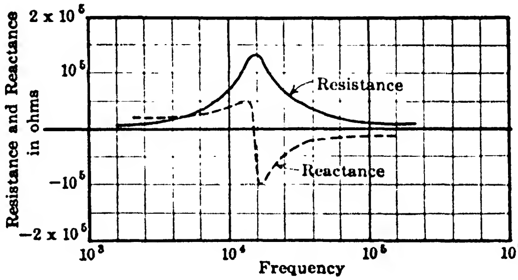


FIG. 93.—The electrical constants of the ordinary head phones at various frequencies.

A Peculiar Case of Parallel Resonance.—A very interesting case of resonance occurs if, with an inductance and condenser in parallel, the resistances in each path are properly adjusted. Thus suppose that the resistance in the two paths are equal and also equal to

$$\sqrt{\frac{L}{C}},$$

[1] Proc. I.R.E., April, 1925.

that is,

$$R_C = R_L = \sqrt{\frac{L}{C}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (69)$$

By inserting this condition in Eqs. (56) and (57) *it will be found that the reactance of the circuit is zero for all frequencies and that the resistance is constant for all frequencies and equal to the resistance of either path.*

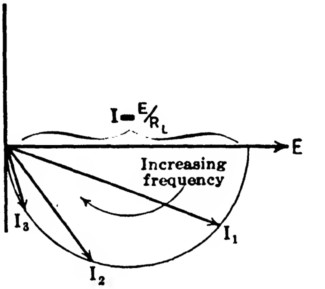


FIG. 94.—Locus of current in an inductive circuit with varying frequency.

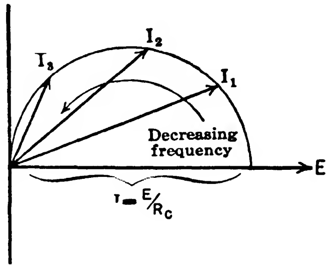


FIG. 95.—Locus of current in a capacitive circuit with varying frequency.

It is a well-known principle in a.c. theory that if a voltage of fixed magnitude and varying frequency is impressed in a circuit of resistance and inductance in series, the locus of current is a semicircle, as shown in Fig. 94. At zero frequency the reactance is zero and the current is equal to E/R_L, and is of course in phase with E. As the frequency increases the current takes the values I_1, I_2, I_3, etc., successively decreasing in magnitude and approaching a lag angle of 90°.

If a capacitance and resistance in series are considered, instead of inductance and resistance, similar conclusions are reached as to the locus of current; it is again a semicircle as shown in Fig. 95. At infinite frequency the reactance of the condenser is zero, so the current is equal to E/R_C. As the frequency diminishes the current decreases in magnitude, and leads more and more, at low frequencies, being very small and leading nearly 90°.

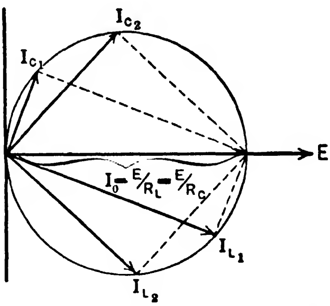


FIG. 96.—The line current in a peculiar form of parallel resonant frequency; its magnitude and phase are independent of frequency.

Now if the resistance of each circuit is equal to $\sqrt{L/C}$ and they are connected in parallel the current relations are as shown in Fig. 96. At very low frequencies the condenser current is negligible and the coil current is I_0. At very high frequencies the coil current is negligible and the condenser current is I_0. For two other frequencies the branch currents are shown in Fig. 96 by I_1 and I_2, respectively. The vector sum of the branch currents is always equal to I_0 and is always in phase with the line voltage E.

Resonant Frequency of Parallel Circuits.—If we define the resonant frequency of a parallel circuit as that frequency which makes the reactance of the circuit zero, thus making the power factor of the circuit unity, we find the resonant frequency by using Eq. (57), putting the numerator equal to zero. This gives the equation

$$R_C{}^2L - R_L{}^2\frac{1}{\omega^2 C} - \frac{L}{C}\left(L - \frac{1}{\omega^2 C}\right) = 0 \quad . \quad . \quad . \quad . \quad . \quad (70)$$

or

$$\frac{1}{\omega^2}\left(\frac{L}{C^2} - \frac{R^2{}_L}{C}\right) = \frac{L^2}{C} - R_C{}^2L,$$

from which we get

$$\omega = \sqrt{\frac{L - R_L{}^2C}{L^2C - R_C{}^2C^2L}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (71)$$

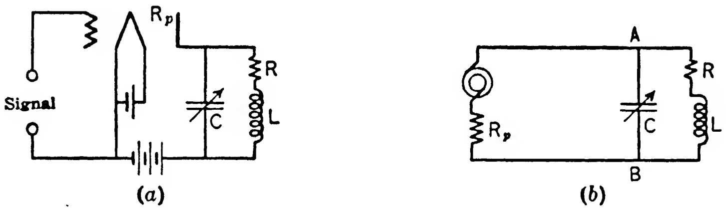


FIG. 97.—An ordinary type of vacuum-tube circuit and its simplified equivalent.

In case the resistance of the condenser arm is negligible a simpler form is obtained,

$$\omega = \sqrt{\frac{1}{LC} - \frac{R^2}{L^2}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (72)$$

It is to be noted that in parallel circuits the resistances in the circuit affect to some extent the resonant frequency of the circuit, whereas in the series circuit the resonant frequency is independent of the resistance.

The condition for resonance in parallel circuits (unity power factor) will in general not be the frequency which gives minimum line current. In case we had defined resonance as that condition which gave minimum line current, formulas somewhat different from Eqs. (71) and (72) would have been obtained.

Parallel Circuit in Series with a Resistance.—In the use of vacuum tubes the plate circuit sometimes contains a parallel resonant circuit as indicated in Fig. 97. We shall take up the action of the vacuum tube in Chapter VI; it here suffices to say that it is essentially a small a.c.

generator, the signal from the receiving antenna generally being the source of generated voltage. The actual circuit arrangement is shown in diagram (*a*), and the equivalent, simplified circuit in diagram (*b*).

The a.c. resistance R_p, between the filament and plate of the actual tube, is represented in diagram (*b*) as R_p, supposed to be the internal resistance of the generator. Under what conditions will the alternator deliver a maximum power to its load circuit, or, in other words, how should the plate circuit of diagram (*a*) be designed to get the maximum output from the tube?

It is shown in many engineering text books that the maximum output of a generator is obtained when the load circuit has a resistance equal to the internal resistance of the generator. It follows then that the resistance between points $A-B$ must be equal to the resistance R_p, generally several thousand ohms. We shall show later that if the vacuum tube is being used as part of an amplifier most of the resistance R_{A-B} should come from the input circuit of the next tube.

Coupling of Various Kinds—Coefficient of Coupling.—When two circuits are so placed or interconnected that energy may be transferred from one to the other they are said to be coupled. There are three types of coupling, resistive, inductive, or capacitive coupling. In the first (practically never used) that part of the network which is common to the two circuits is a resistance; in the second, part of the magnetic field generated by currents in the network is common to both circuits; in the third, a part of the electrostatic field set up in the network is common to both circuits. The coupling which uses the magnetic field is called inductive or magnetic coupling, and that which uses the electric field is called capacitive or static coupling. The magnetic coupling may be through an inductance common to both circuits called *direct*, or it may be through a mutual inductance in which case it is generally called *inductive* coupling.

The three principal types of coupling are shown in Figs. 98, 99, and 100, that of Fig. 98 being direct, that of Fig. 99 being inductive, and that of Fig. 100 being capacitive.

The extent to which circuits are coupled is given quantitatively by the *coupling coefficient* or *coefficient of coupling*. This is defined as *the ratio of the common reactance of the two circuits to the square root of the product of the reactances (of similar kind to that giving the coupling) of the two circuits.*

Thus if X_m = reactance common to both circuits;

X_1 = reactance of circuit 1;

X_2 = reactance of circuit 2;

k = coupling coefficient.

$$k=\frac{X_m}{\sqrt{X_1X_2}}. \qquad (73)$$

In Fig. 98 the total reactance of circuit 1 is $\omega(L_1+M)$, that of circuit 2 is $\omega(L_2+M)$, and the common reactance is ωM. Therefore

$$k=\frac{\omega M}{\sqrt{\omega(L_1+M)\omega(L_2+M)}}=\frac{M}{\sqrt{(L_1+M)(L_2+M)}}. \quad . \quad . \quad (74)$$

In Fig. 99 the total inductance of circuit 1 is indicated by L_a+L_b; part of this is in inductive relation to circuit 2 and part is not. Similarly the

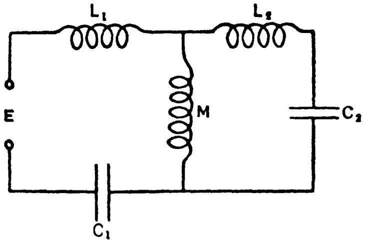


FIG. 98.—Direct coupling.

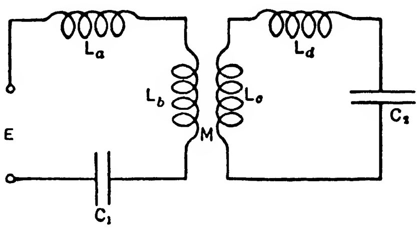


FIG. 99.—Inductive coupling.

inductance of the second circuit consists of two parts L_c and L_d, one part magnetically coupled to circuit 1 and the other part not so coupled. The common reactance is ωM. Hence for this case we have,

$$k=\frac{\omega M}{\sqrt{\omega(L_a+L_b)\omega(L_c+L_d)}}=\frac{M}{\sqrt{(L_a+L_b)(L_c+L_d)}}. \quad . \quad . \quad (75)$$

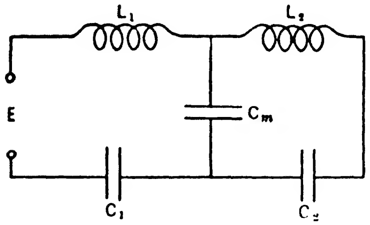


FIG. 100.—Capacitive coupling.

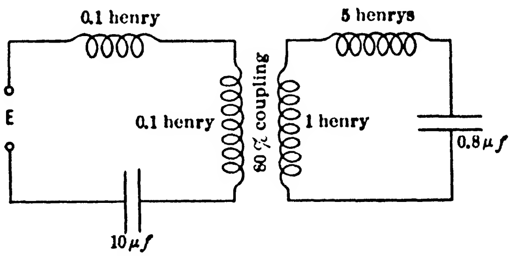


FIG. 101.—Inductively coupled circuits.

The inductively coupled circuit of Fig. 99 can always be considered as a direct-coupled circuit after the proper transformations have been made. The inductance of circuit 2 must be decreased in the ratio L_b/L_c and the capacity of circuit 2 must be increased in the ratio L_c/L_b. This transformation of the inductance and capacity of circuit 2 leaves the *oscillation constant* (LC) the same as it was with the original values of L and C.

The M of the equivalent circuit is obtained by multiplying the actual value of M by the ratio $\sqrt{L_b/L_c}$. Let us call these new values M', L'_c, L'_d, and C'_2. The inductively coupled circuit is now replaced by the direct-coupled circuit similar to that of Fig. 98. For the L_1 of Fig. 98 we use $(L_a+L_b)-M'$ and for the L_2 of Fig. 98 we use $(L'_c+L'_d)-M'$.

An actual inductively coupled circuit is shown in Fig. 101, the coefficient of coupling of the *two coils of the transformer* is 80 per cent, or

$$M=0.8\sqrt{1\times 0.1}=0.253 \text{ henry.}$$

In Fig. 102 the inductances of the second circuit have been decreased in the ratio 0.1/1 and the capacity has been increased in the ratio 1/0.1. The coefficient of coupling must remain as it was for Fig. 101, so we decrease M in the ratio $\sqrt{0.1/1}$, giving a value of 0.08 henry.

The direct-coupled circuit, which is the exact equivalent of the inductively coupled circuit of Fig. 102, is now given in Fig. 103. The total L of circuit 1 is the same as that of Fig. 101, 0.08 henry being coupled 100 per cent to circuit 2; similarly the total inductance of circuit 2 is the same as it is in Fig. 102.

The coefficient of coupling of the circuit of Fig. 103 is

$$k=\frac{0.08}{\sqrt{0.2\times 0.6}}=0.232,$$

and for the actual inductively coupled circuit of Fig. 101 it is

$$k=\frac{0.253}{\sqrt{0.2\times 6.0}}=0.232,$$

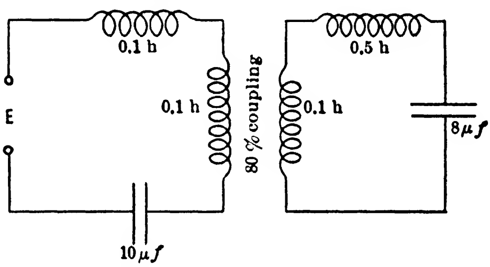


FIG. 102.—Circuit of Fig. 101 reduced to an equivalent 1 : 1 ratio circuit.

which is just the same as for the substituted direct-coupled circuit.

It is possible to replace an inductively coupled circuit by one directly coupled without making the transformations explained above. Calling the total inductance in primary and secondary of the inductively coupled circuit L_3 and L_4 respectively, and the mutual inductance M, then the direct-coupled circuit is written down at once as shown in Fig. 98 by making $L_1=L_3-M$ and $L_2=L_4-M$. The same values of M, C_1 and C_2 are used in the direct-coupled circuit as in the inductively coupled circuit.[1] The justification for making the change from one type of circuit to the other may be seen upon writing the equations for the reactive voltages of the two circuits of Figs. 98 and 99. For Fig. 98 we have,

$$\omega L_1 I_1-\frac{I_1}{\omega C_1}+\omega M(I_1-I_2)=E, \quad . \quad . \quad . \quad . \quad . \quad (76)$$

[1] See Bulletin 74 of the Bureau of Standards, p. 50.

and

$$\omega L_2 I_2 - \frac{I_2}{\omega C_2} + \omega M(I_2 - I_1) = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (77)$$

For the circuit shown in Fig. 99 we may put $L_a + L_b = L_3$ and $L_c + L_d = L_4$; the reactive voltages for these circuits then become,

$$\omega L_3 I_1 - \frac{I_1}{\omega C_1} - \omega M I_2 = E, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (78)$$

and

$$\omega L_4 I_2 - \frac{I_2}{\omega C_2} - \omega M I_1 = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (79)$$

Put $L_3 = L_1 + M$ and $L_4 = L_2 + M$ and these equations become

$$\omega(L_1 + M)I_1 - \frac{I_1}{\omega C_1} - \omega M I_2 = E, \quad . \quad . \quad . \quad . \quad . \quad . \quad (80)$$

and

$$\omega(L_2 + M)I_2 - \frac{I_2}{\omega C_2} - \omega M I_1 = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (81)$$

By collecting terms these may be changed into the forms,

$$\omega L_1 I_1 - \frac{I_1}{\omega C_1} + \omega M(I_1 - I_2) = E, \quad . \quad . \quad . \quad . \quad . \quad . \quad (82)$$

and

$$\omega L_2 I_2 - \frac{I_2}{\omega C_2} + \omega M(I_2 - I_1) = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (83)$$

But these equations, which are for an inductively coupled circuit, are identical with Eqs. (76) and (77), which are for the directly coupled circuit.

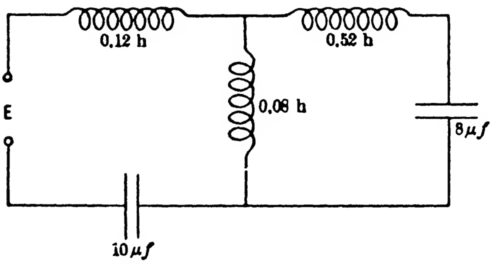


FIG. 103.—Direct-coupled circuit equivalent to inductively coupled circuit of Fig. 102.

The author does not believe that this method is as satisfactory a one as that using transformed L and C in the secondary because of certain ambiguities which may arise. As an illustration of the cases in which the method works out all right we take Fig. 104. For the L_1 of Fig. 98 we must put $0.2 - 0.1 = 0.1$ henry and for L_2 of Fig. 98 we put $0.4 - 0.1 = 0.3$ henry. M, C_1, and C_2 remain as in Fig. 104. The equivalent directly coupled circuit is given in Fig. 105; it is electrically equivalent to Fig. 104. Now suppose the circuit of Fig. 101 to be treated in this manner; for L_1 we obtain

$0.2-0.253=-0.053$ henry. This means that instead of putting in an inductance for the L_1 of Fig. 98 we must put a condenser, the capacity of which is such that its reactance is equal, in magnitude, to that given by 0.053 henry of inductance.

For the circuit shown in Fig. 100, we get the coupling coefficient from Eq. (73), in the following manner:[1]

$$\text{Mutual capacity reactance}=\frac{1}{\omega C_m}.$$

$$\text{Capacity reactance of circuit } 1=\frac{1}{\omega C_1}+\frac{1}{\omega C_m}=\frac{1}{\omega C_a}, \quad . \quad . \quad . \quad . \quad . \quad (84)$$

in which

$$C_a=\frac{1}{\frac{1}{C_1}+\frac{1}{C_m}}.$$

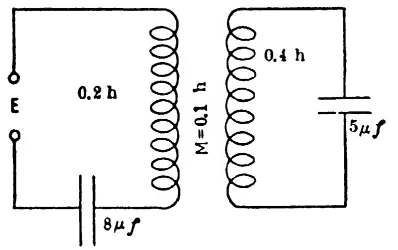


FIG. 104.—Inductively coupled circuit.

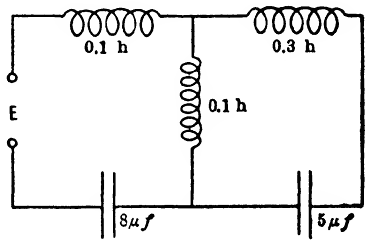


FIG. 105.—Direct-coupled circuit equivalent to circuit of Fig. 104.

$$\text{Capacity reactance of circuit } 2=\frac{1}{\omega C_2}+\frac{1}{\omega C_m}=\frac{1}{\omega C_b}, \quad . \quad . \quad . \quad . \quad . \quad (85)$$

in which

$$C_b=\frac{1}{\frac{1}{C_2}+\frac{1}{C_m}}.$$

Hence Eq. (73) becomes for this case,

$$k=\frac{\frac{1}{\omega C_m}}{\sqrt{\frac{1}{\omega C_a}\times\frac{1}{\omega C_b}}}=\frac{\sqrt{C_aC_b}}{C_m}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (86)$$

[1] For more complete analysis of capacitively coupled circuits and comparison of capacitive and inductive coupling see Cohen, "Electrostatically Coupled Circuits," Proc. I.R.E., Vol. 8, No. 5, Oct., 1920.

A rather more complicated case[1] of static coupling is given in Fig. 106.

We first replace the figure by a simpler one, equivalent to it. This is given in Fig. 107, in which

$$C_3=\frac{1}{\frac{1}{C'}+\frac{1}{C''}}=\frac{C'C''}{C'+C''}.$$

Now calculate the capacity of circuit A, remembering that C_3 and C_2, in series, are in parallel with C_1. This gives

$$C_A=C_1+\frac{1}{\frac{1}{C_3}+\frac{1}{C_2}}=\frac{C_1C_2+C_2C_3+C_3C_1}{C_2+C_3}. \quad . \quad . \quad . \quad . \quad (87)$$

The "mutual reactance" is now calculated by assuming 1 ampere to be circulating in circuit A, and calculating how much voltage is produced

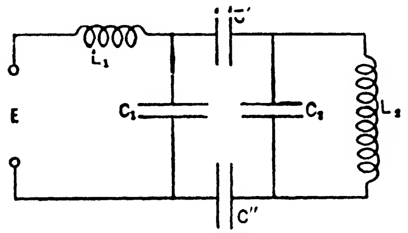


FIG. 106.—Complex capacitive coupling.

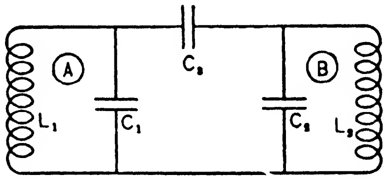


FIG. 107.—The circuit of Fig. 106 may be simplified to this form.

in circuit B. Per ampere of current in circuit A the voltage across C_1 is $\frac{1}{\omega C_A}$ and of this voltage a fraction is set up in circuit B across condenser C_2. This fractional part is

$$\frac{\frac{1}{\omega C_2}}{\frac{1}{\omega\frac{C_2C_3}{C_2+C_3}}}=\frac{C_3}{C_2+C_3}.$$

[1] In case it is not evident just what the mutual reactance of the two circuits is, it may be obtained by calculating the voltage generated in circuit 2 when a current of 1 ampere is flowing in circuit 1, or vice versa. This voltage is equal to the mutual reactance, in ohms.

So that the voltage across C_2 is equal to

$$\frac{1}{\omega}\left(\frac{C_3}{C_2+C_3}\times\frac{1}{C_A}\right)=\frac{1}{\omega}\frac{C_3}{C_1C_2+C_2C_3+C_3C_1}, \quad . \quad . \quad . \qquad (88)$$

which is thus the mutual reactance.

Now the capacitive reactance $\left(\frac{1}{\omega C_A}\right)$ of circuit A is

$$\frac{1}{\omega\dfrac{C_1C_2+C_2C_3+C_3C_1}{C_2+C_3}},$$

and of circuit B it is

$$\frac{1}{\omega\dfrac{C_1C_2+C_2C_3+C_3C_1}{C_1+C_3}}.$$

Then taking the quotient of the mutual reactance by the geometric mean of the self reactances (that is, using Eq. (73)) we get

$$k=\frac{C_3}{\sqrt{(C_1+C_3)(C_2+C_3)}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (89)$$

Combined Electric and Magnetic Coupling.—In certain radio receiving sets combined capacitive and inductive couplings are used, as indicated in Fig. 108, in this case the coefficient of magnetic coupling and that of static coupling are calculated separately, and the actual amount of coupling is the sum or difference of these two, according as the e.m.f.s induced in circuit 2 through the two types of coupling are in phase or 180° out of phase with each other.

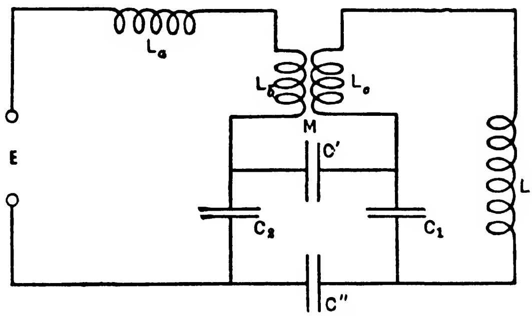


Fig. 108.—Combined capacitive and inductive coupling.

It may sometimes be desired to induce a constant voltage in the second circuit as the frequency of the impressed voltage is varied, the current in the first circuit being held constant. With the combined electric and magnetic coupling of Fig. 108 this is nearly possible.

We first notice that the two coupling condensers, C' and C'', may be replaced by a single one reducing Fig. 108 to the simpler one given in Fig. 109. Now we assume a current of one ampere to flow from the power supply and calculate what voltage is set up in circuit 2, as the frequency is changed.

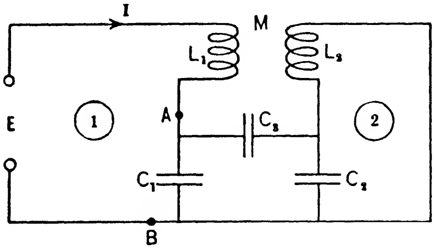


FIG. 109.—In some radio circuits both magnetic and capacitive coupling exist simultaneously.

There will be a voltage induced in coil L_2 due to the mutual induction between L_1 and L_2. This voltage, E_{L_2}, will be directly proportional to the frequency and to the value of the mutual induction.

$$E_{L_2} = 2\pi f M I_1 = 2\pi f M.$$

The current I_1 flowing through the condenser combination (C_1, in parallel with C_2 and C_3 in series) will set up across C_1 a voltage inversely proportional to the frequency, and a certain fraction of this will appear across C_2. Let us call the capacity between points A-B, C_A. Then the voltage across points A–B will be given by the expression

$$E_{A-B} = \frac{I_1}{2\pi f C_A} = \frac{1}{2\pi f C_A}.$$

The total reactance from A to B through C_3 and C_2 is

$$\frac{1}{\omega C_3} + \frac{1}{\omega C_2} = \frac{1}{\omega}\,\frac{C_2 + C_3}{C_2 C_3}.$$

The ratio of the reactance of C_2 to this total reactance is

$$\frac{\dfrac{1}{\omega C_2}}{\dfrac{1}{\omega}\,\dfrac{C_2+C_3}{C_2 C_3}} = \frac{C_3}{C_2 + C_3},$$

and this is the proportion of the voltage across points A–B, which is impressed in the second circuit. This gives us

$$E_{C_2} = \frac{1}{2\pi f C_A}\left(\frac{C_3}{C_2 + C_3}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (90)$$

The total voltage set up in circuit 2 is then the sum of E_m and E_{C_2}, due notice being taken of their relative phases.

Now the voltage E_{L_2} is 90° behind I_1 in phase and the voltage E_{C_2} is also 90° behind I_1. But it is to be noticed that by reversing the connection of coil L_2, that the phase of E_{L_2} (in so far as its relation to E_{C_2} is concerned) may be reversed, hence the total voltage set up in circuit 2 may be either the arithmetical sum or difference of E_{L_2} and E_{C_2}, as we please.

The two voltages, E_{L_2} and E_{C_2}, in their relation to frequency, are shown in Fig. 110. Their sum is also shown, and it can be seen that within a certain frequency range (for example, from f_1 to f_2) the voltage set up in circuit 2, by a current of fixed amplitude in circuit 1, is essentially constant.

By properly proportioning E_{L_2} and E_{C_2} it is possible to make the induced voltage set up in circuit 2 to vary (in a limited frequency range) in almost any desired manner.

In another coupling scheme (shown in Fig. 111) a so-called link circuit, untuned, is used to connect the other two circuits. In this case the coupling between circuits 1 and 2 is obtained by calculating the coupling of circuits 1 and 3 and then that of 3 and 2.

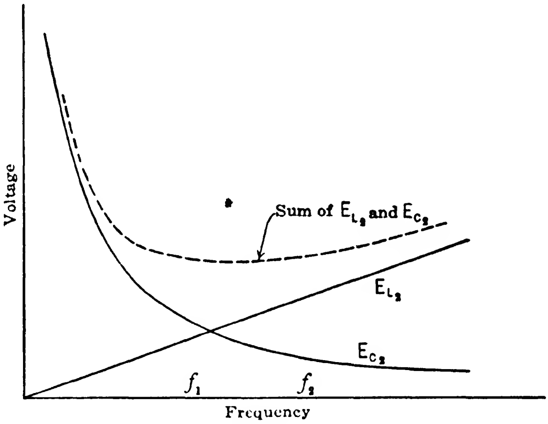


FIG. 110.—A possible action of the circuit of Fig. 109.

$$k_{1-3}=\frac{M_{1-2}}{\sqrt{L_1(L_2+L_3)}},$$

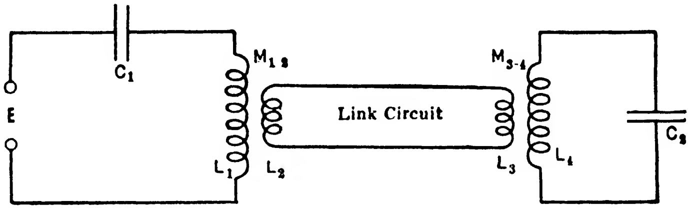
$$k_{3-2}=\frac{M_{3-4}}{\sqrt{L_4(L_2+L_3)}}.$$

Then

$$k_{1-2}=k_{1-3}\times k_{3-2}=\frac{M_{1-2}\times M_{3-4}}{(L_2+L_3)\sqrt{L_1L_4}}. \quad \ldots \ldots \quad (91)$$

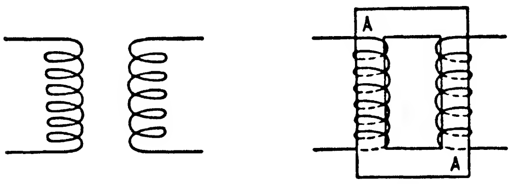
The link-circuit scheme is frequently very convenient in arranging laboratory circuits for making radio frequency measurements. It permits of getting a magnetic coupling without having at the same time a large and unknown electrostatic coupling. Circuit 1, for example, may be connected to a power tube, generally having an intense electric field associated with it. Circuit 2 is the measuring circuit, in which a small, known, voltage is desired. Generally if circuit 2 is brought into proximity with circuit 1, the electric fields will give large and unknown errors.

By enclosing circuit 1 in a copper box and using a link circuit as shown in Fig. 111, to get the requisite power to the measuring circuit and grounding the link circuit, errors due to electric fields can be avoided.



[FIG. 111.—"Link-circuit" coupling.

"Coefficient of Coupling" Sometimes Ambiguous.—Not only the relative positions of the two coils enter into the determination of their coefficient of coupling, but also the permeability of the medium used in their magnetic circuit, and the disposition of this medium, have their effects. Thus in Fig. 112(*a*) are shown two coils in air, having a coefficient of coupling of possibly 5 per cent. In Fig. 112(*b*) these same coils, in the same relative positions, have been threaded by an iron core *A–A*. The coefficient of coupling may now be 90 per cent, as measured with 60-cycle current. But if the measurement is made at 1000 cycles, the measured value of coupling will probably be found about 50 per cent and at 10,000 cycles it may be only 10 per cent, this variation depending upon the degree of lamination of the core.



(*a*) (*b*)

FIG. 112.—The magnetic coupling between two coils (*a*) can be greatly increased by the use of an iron core threading both coils (*b*).

When both condensers and coils enter into the action of the two circuits coupled together, certain other ambiguities arise in interpreting the meaning of the term, coefficient of coupling. Thus in Fig. 111 circuit 1 and circuit 2 are coupled magnetically, with a coefficient of coupling given by Eq. (91). Let us now suppose that the link circuit is opened and a variable condenser put in the circuit. Then let us suppose that this condenser is adjusted to such a value that in combination with L_2 and L_3 it produces resonance for the frequency impressed in circuit 1. As this condenser is varied through its resonance value the current in the link circuit varies greatly and correspondingly the voltage induced in circuit 2 varies. As the current in circuit 1 may be maintained at a fixed value, while the condenser in the link circuit is varied, we have the anomalous situation that with fixed coupling and fixed primary current (fixed in both magnitude and frequency) the voltage induced in the secondary circuit varies greatly.

The link circuit in this case corresponds somewhat to the iron core of Fig. 112(*b*); it serves to transfer energy from circuit 1 to circuit 2, so that energy may flow more readily when the link circuit is present than if not. But with the tuning condenser present in the link circuit energy flows more readily from circuit 1 to circuit 2 at one frequency (the resonant frequency of the tuned link circuit) than at any other.

But the coefficient of *magnetic coupling* between the two circuits (as ordinarily conceived) must evidently be independent of the adjustment of a tuning *condenser* in the link circuit. The link circuit in this case, however, acts somewhat like the iron core of Fig. 112(*b*), if we suppose that the iron has a high permeability for one frequency only.

Resonance in a Circuit to Which Another Circuit is Magnetically Coupled.—In discussing this question we shall calculate the effect of circuit 2 on the resistance and reactance of circuit 1. The method of analysis is somewhat more elementary than that ordinarily given (which depends upon the solution of simultaneous differential equations), and perhaps leads to a clearer insight into the mutual reactions of the two circuits. We shall assume unit current flowing in the primary (circuit 1), and get the voltage E_2 induced in the second circuit by this current. This voltage will produce current in the second circuit, which current will be divided into its active and reactive components (in phase with E_2 and 90° out of phase with E_2). The active component will be 90° behind the primary current and will produce a voltage back in the primary circuit which will be 180° out of phase with the primary current; from this voltage we calculate the effect of the second circuit on the resistance of the first.

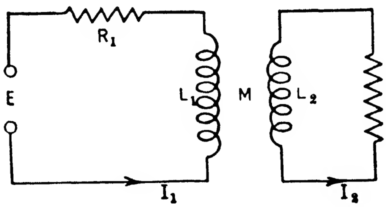


FIG. 113.—Inductively coupled circuits, neither circuit having a condenser.

The resistance of a circuit may be defined as the counter voltage set up in the circuit by a current of 1 ampere flowing, this counter voltage to be 180° out of phase with the current; in the same way *the reactance of a circuit may be considered as the counter voltage set up in the circuit by a current of 1 ampere, the counter voltage to be 90° out of phase with the current.*

In Fig. 113 suppose the current I_1 is 1 ampere at frequency $\omega/2\pi$. Voltage induced in the secondary

$$E_2 = \omega M I_1 = \omega M. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (92)$$

Current in circuit 2,

$$I_2 = \frac{E_2}{Z_2} = \frac{\omega M}{Z_2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (93)$$

and this current lags behind E_2 by an angle θ, defined by the equation $\tan \theta = \omega L_2/R_2$. The active component of I_2 is in phase with E_2 so we put

$$I_2 \cos \theta = \frac{\omega M}{Z_2} \times \frac{R_2}{Z_2}.$$

The voltage induced in circuit 1 by this current is

$$I_2 \cos \theta \times \omega M = \frac{\omega M}{Z_2}\frac{R_2}{Z_2}\omega M = \left(\frac{\omega M}{Z_2}\right)^2 R_2. \quad . \quad . \quad . \quad (94)$$

As this voltage lags 90° behind the inducing current $I_2 \cos \theta$, and as $I_2 \cos \theta$ lags 90° behind I_1, this voltage $\left(\frac{\omega M}{Z_2}\right)^2 R_2$ is 180° behind I_1 and *so is an IR reaction*. As we assumed unit current in circuit 1 this voltage $\left(\frac{\omega M}{Z_2}\right)^2 R_2$ is really *the increase in resistance of circuit 1, in ohms, due to the current in circuit 2*. Hence the apparent resistance of circuit 1 is evidently

$$R'_1 = R_1 + \left(\frac{\omega M}{Z_2}\right)^2 R_2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (95)$$

To those familiar with the algebra of complex numbers this detailed method of analysis may seem tedious and cumbersome; it does have the merit, however, of making the reader follow more closely the physical actions occurring in the circuit and so is possibly to be preferred to the shorter method of complex algebra.

Now the reactive current in circuit 2 is $I_2 \sin \theta$, and this current lags 90° behind E_2, which itself lags 90° behind I_1. The voltage induced in circuit 1 by this current $I_2 \sin \theta$ will be equal to $\omega M I_2 \sin \theta$, and this will lag 90° behind the inducing current $I_2 \sin \theta$, and hence will lag 270° behind I_1, that is, it leads I_1 by 90°.

Now the reactive voltage in circuit 1 due to L_1 is 90° behind the current I_1. This may seem incorrect at first glance, because it makes the current I_1 lead the reactive voltage by 90°, whereas we know that an inductive circuit draws a lagging current. It must be remembered that the component of the impressed voltage which overcomes the reacting voltage of the circuit must be 180° ahead of the reacting voltage itself; this makes the current in an inductive circuit lag behind the *impressed voltage*, as it should.

It appears then that the voltage induced in circuit 1 by the current $I_2 \sin \theta$ is 180° out of phase with the reactive voltage in circuit 1 due to L_1 of circuit 1, hence the total reactive voltage of circuit 1 will be less when circuit 2 is present than when it is not present.

The amount of voltage induced in circuit 1 by $I_2 \sin \theta$ is $\omega M I_2 \sin \theta$ and this is equal to $\left(\frac{\omega M}{Z_2}\right)^2 \omega L_2$.

So the total reactive voltage in circuit 1 when a current of 1 ampere is flowing is

$$\omega L_1 - \left(\frac{\omega M}{Z_2}\right)^2 \omega L_2 = \omega\left(L_1 - \left(\frac{\omega M}{Z_2}\right)^2 L^2\right),$$

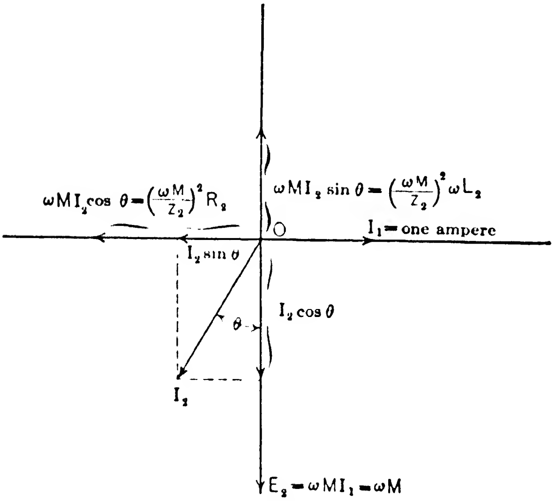


FIG. 114.—Vector relations of currents and voltages in the circuit of Fig. 113.

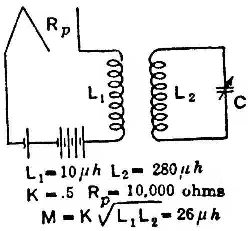


FIG. 115.—An actual radio circuit to which this theory applies; the resistance of the tuned circuit is materially increased by the plate circuit of the preceding triode.

and from this we get the equivalent self-induction of circuit 1,

$$L'_1 = L_1 - \left(\frac{\omega M}{Z_2}\right)^2 L_2, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (96)$$

and, of course,

$$X'_1 = X_1 - \left(\frac{\omega M}{Z_2}\right)^2 X_2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (97)$$

It is therefore seen that the effect of the current in circuit 2 is to *increase* the resistance of circuit 1 by the amount $\left(\frac{\omega M}{Z_2}\right)^2 R_2$ and to *decrease* its self-induction by an amount $\left(\frac{\omega M}{Z_2}\right)^2 L_2$.

The foregoing relations are shown vectorially in Fig. 114. The phase of I_1 is taken as reference and the diagram at once becomes clear when it is remembered that a sine wave of current induces (in its own circuit as well as in any other circuits coupled to it magnetically) a sine wave of voltage *90° behind the current.* The division of the current I_2 into the two components shown is for convenience only; by doing it the two components of the voltage induced in circuit 1, by the current I_2, have the phases desired.

An illustration of the application of the foregoing theory, insofar as it deals with increase of resistance due to the presence of an untuned secondary circuit, is given in Fig. 115, showing part of the circuit of a tuned radio-frequency amplifier. The L_2–C_2 circuit is tuned to select the desired signal from all the others present in the air; how well it does this is determined by the sharpness of tuning of the L_2–C_2 circuit. This circuit has a resistance of its own, and in addition, according to the theory above, the plate circuit of the vacuum tube will act to increase the resistance of the L_2–C_2 circuit. Between the plate and filament of the tube there is a resistance (of a peculiar sort) generally designated by R_p. This is so large compared to the resistance of coil L_1 that this latter may be neglected.

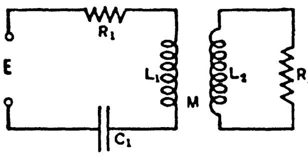


FIG. 116.—Inductively coupled circuits with a condenser in the primary.

In the table below are given the measured values of resistance of the L_2–C_2 circuit, with the L_1 circuit open; it shows that the resistance increases rapidly as the frequency increases. This effect is taken up in detail in Chapter II.

Now when the plate circuit of the tube (Fig. 115) is closed, the resistance of the L_2–C_2 circuit increased as shown in the table. The increase, for any one frequency, is obtained by subtraction, and evidently gives the resistance which the plate circuit introduces into the L_2–C_2 circuit.

In the table is also shown the calculated value of resistance increase, using Eq. (95). The results show quite reasonable agreement between test and theory; it is not possible to get the same accuracy of measurement at radio frequency as it is at 60 cycles.

Frequency	620 kc.	1000 kc.	1500 kc.
R_2 of L_2–C_2 circuit alone with plate circuit open	7.2	9.5	16.0
R'_2 of L_2–C_2 circuit when plate circuit is closed	7.9	11.7	21.3
Difference	0.7	2.2	5.3
$(\omega M)^2/R_p$	1.0	2.6	5.9

In such a circuit as that shown in Fig. 116 we can at once write the characteristics of circuit 1 by using Eqs. (95) and (96).

$$R'_1=R_1+\left(\frac{\omega M}{Z_2}\right)^2 R_2,$$

$$L'_1=L_1-\left(\frac{\omega M}{Z_2}\right)^2 L_2,$$

$$I_1=\frac{E}{\sqrt{\left[R_1+\left(\frac{\omega M}{Z_2}\right)^2 R_2\right]^2+\left[\omega\left(L_1-\left(\frac{\omega M}{Z_2}\right)^2 L_2\right)-\frac{1}{\omega C_1}\right]^2}}. \quad . \quad . \quad (98)$$

$$I_2=\frac{E_2}{Z_2}=\frac{\omega M E}{Z_2\sqrt{\left[R_1+\left(\frac{\omega M}{Z_2}\right)^2 R_2\right]^2+\left[\omega\left(L_1-\left(\frac{\omega M}{Z_2}\right)^2 L_2\right)-\frac{1}{\omega C_1}\right]^2}}. \quad (99)$$

These two equations may be written in a somewhat more convenient form for calculation, by combining terms,

$$I_1=$$

$$\frac{EZ_2}{\sqrt{\left[(\omega M)^2+R_1R_2-\omega L_2\left(\omega L_1-\frac{1}{\omega C_1}\right)\right]^2+\left[\omega L_2R_1+R_2\left(\omega L_1-\frac{1}{\omega C_1}\right)\right]^2}}. \quad (100)$$

$$I_2=$$

$$\frac{E\omega M}{\sqrt{\left[(\omega M)^2+R_1R_2-\omega L_2\left(\omega L_1-\frac{1}{\omega C_1}\right)\right]^2+\left[\omega L_2R_1+R_2\left(\omega L_1-\frac{1}{\omega C_1}\right)\right]^2}}. \quad (101)$$

In case the impressed frequency is adjusted to give resonance in the primary circuit (without the presence of the secondary) these equations reduce to the forms

$$I_1=\frac{EZ_2}{\sqrt{(\omega^2M^2+R_1R_2)^2+\omega^2L_2^2R_1^2}}. \quad . \quad . \quad . \quad (102)$$

$$I_2=\frac{E\omega M}{\sqrt{(\omega^2M^2+R_1R_2)^2+\omega^2L_2^2R_1^2}}. \quad . \quad . \quad . \quad (103)$$

if M is varied a maximum current will occur in the secondary when

$$\omega^2M^2=R_1\sqrt{R_2^2+(\omega L_2)^2}=R_1Z_2. \quad . \quad . \quad . \quad . \quad . \quad (104)$$

For this value of M the values of the two currents become

$$I_1=\frac{EZ_2}{R_1\sqrt{2(Z_2^2+R_2Z_2)}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (105)$$

$$I_2=\frac{E\omega M}{R_1\sqrt{2(Z_2^2+R_2Z_2)}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (106)$$

Fig. 117 shows a set of experimental curves to illustrate the relations given above; the circuits were arranged as shown in Fig. 116 and the frequency adjusted for the value which gave resonance in the primary alone; the

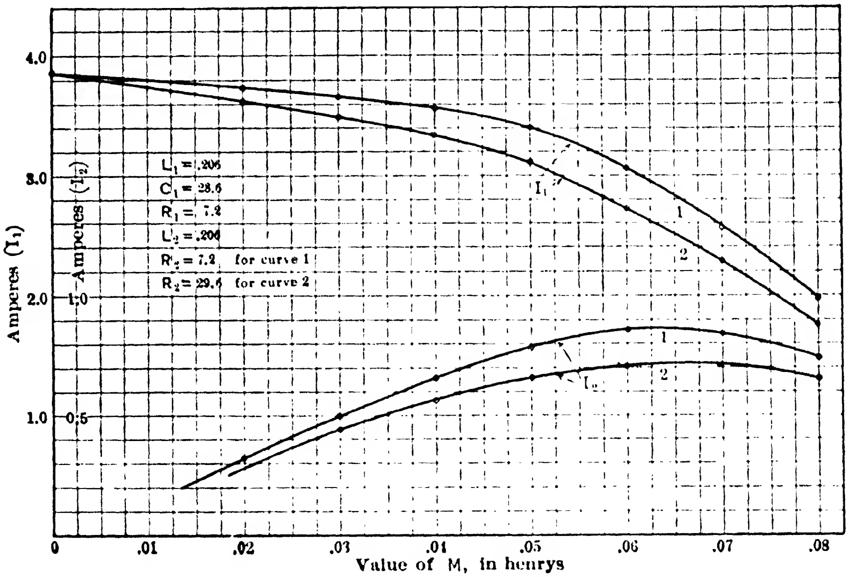


FIG. 117.—Variation of current with M in circuit of Fig. 116, for two values of secondary resistance.

coupling was then varied and the two currents went through variations as appear from the curves.

With the same value of frequency as used for the curves of Fig. 117 and that coupling which gave maximum secondary current (which value of coupling does not vary greatly as the secondary resistance is varied, so long as the secondary resistance is small compared to the secondary reactance) a series of readings was taken to show the effect of the secondary resistance on secondary current and so on the amount of power transmitted to the secondary circuit. The results are shown in Fig. 118; it is seen that the adjustments for maximum power of this circuit are not very critical.

By using the relation given in Eq. (96) we find the resonant frequency of the circuit of Fig. 116 is given approximately by

$$\omega=\sqrt{\frac{L_2}{C_1(L_1L_2-M^2)}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (107)$$

For very weak coupling, M approaching zero, it is seen that the value of ω approaches the value $\frac{1}{\sqrt{L_1C_1}}$, as we know it should.

The resonance curve for such a circuit as that shown in Fig. 116 differs from the curve of the primary alone in that the critical frequency is higher

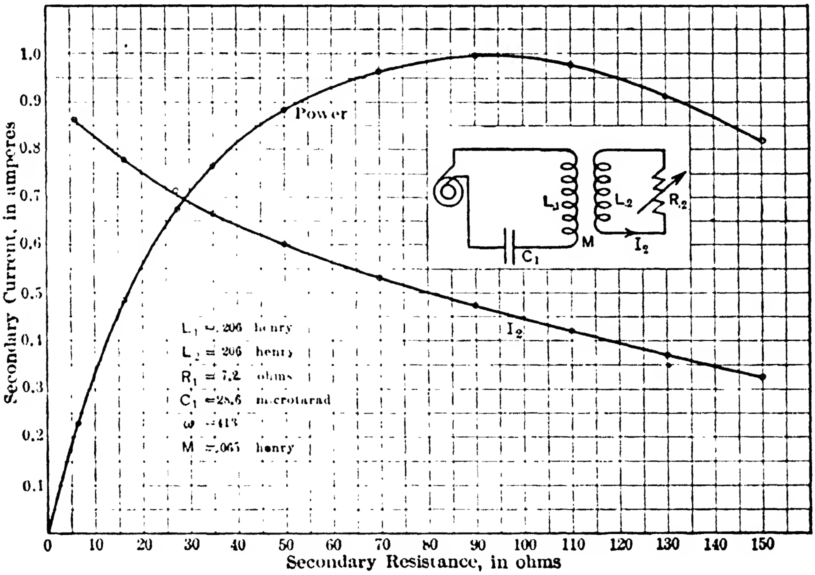


FIG. 118.—Variation of power and current (in Circuit 2) with secondary resistance in circuit of Fig. 116, coupling constant.

and the resonance curve is not so sharp. The resistance of circuit 1 is increased by the amount given in Eq. (95) and the inductance is decreased by the amount shown in Eq. (96). Fig. 119 shows the resonance curve of a circuit arranged like that of Fig. 116; in curve A is shown the resonance action of the primary without the presence of the secondary. The same voltage was applied to the primary circuit in getting the two sets of curves, hence the magnitudes of current for the two curves give an exact measure of the effect of the secondary circuit upon the first. The calculated R' and L' of the primary, using first the experimental data on the curve sheet of Fig. 119 and then Eqs. (95) and (96), agree within the precision of the experimental work.

The resistance of the secondary circuit was then increased by 12 ohms and another resonance curve taken; the results are shown in Fig. 120; the curves of Fig. 119 are shown in dotted lines for comparison. It is seen that the addition of resistance to the secondary circuit makes the sharpness of resonance less and the effect of the secondary in determining the resonant frequency of the primary is somewhat less than for the lower resistance secondary circuit.

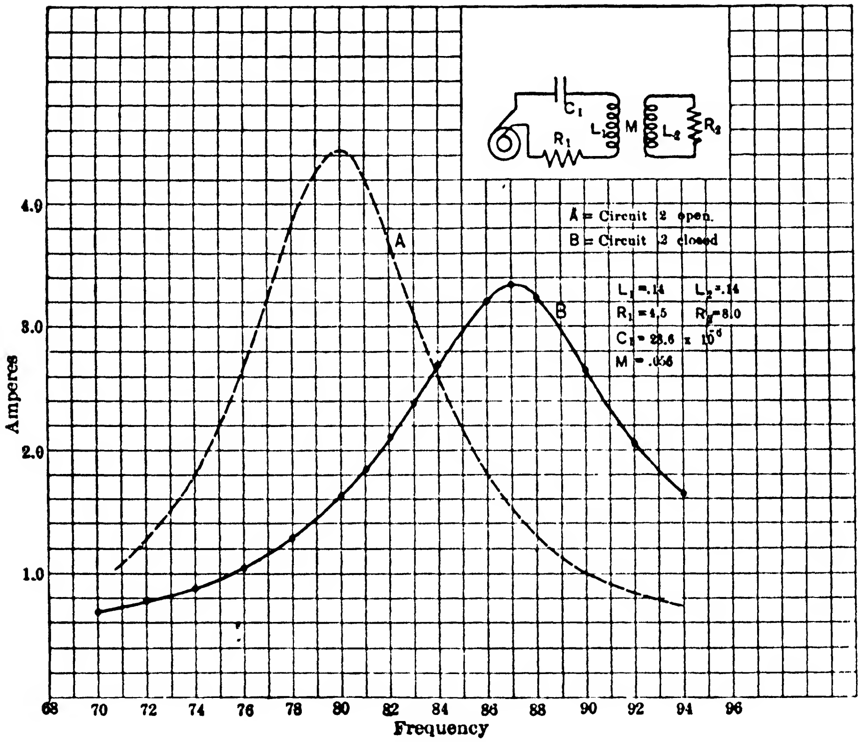


FIG. 119.—Current *vs.* frequency in primary circuit of Fig. 116 with secondary open and closed.

We will next consider the circuit shown in Fig. 121, the condenser now being in the secondary circuit instead of the primary.

In this circuit the resistance of the primary is always increased by the presence of the secondary, but the effect upon the inductance depends upon the frequency impressed on the primary circuit. If the frequency is such as to satisfy the condition for resonance in the secondary $\left(f=\dfrac{1}{2\pi\sqrt{L_2C_2}}\right)$, the apparent inductance of circuit 1 will be the same as

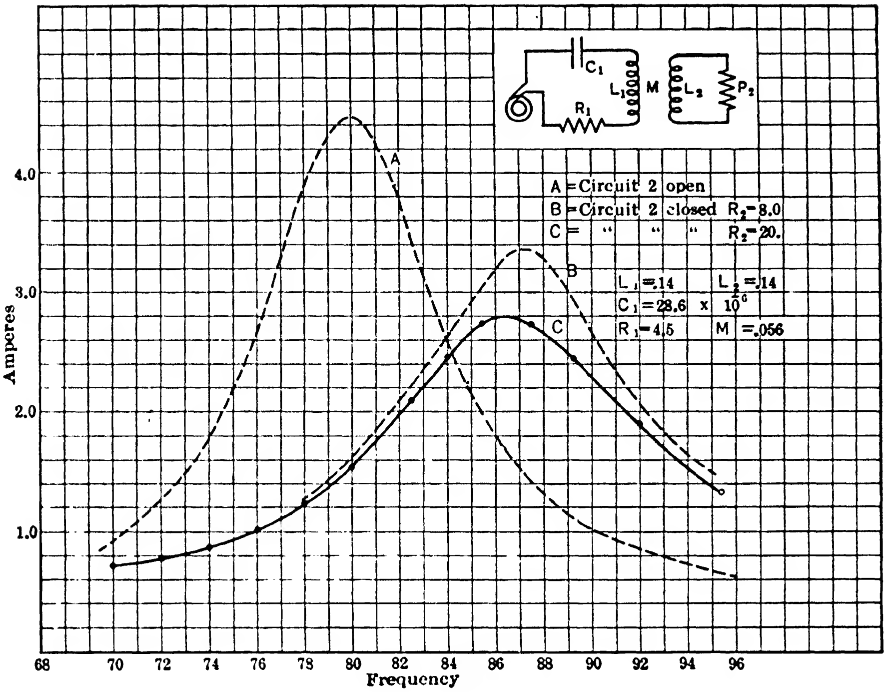


FIG. 120.—Current *vs.* frequency in circuit of Fig. 116 with added secondary resistance.

the actual inductance, that is, the presence of circuit 2 does not affect the inductance of circuit 1. With higher than resonant frequency the apparent inductance of circuit 1 is decreased by circuit 2 and with lower frequency the inductance of circuit 1 is increased. In other words, if I_2 lags behind E_2, the effect on circuit 1 is to reduce the apparent inductance, whereas if the current in circuit 2 leads the generated voltage in this circuit, the effect on circuit 1 is to cause an increase in the apparent inductance.

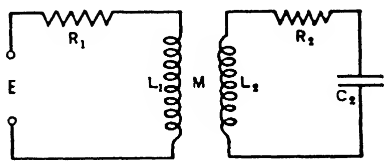


FIG. 121.—Inductively coupled circuit with condenser in secondary.

Applying Eqs. (95) and (96) to the circuit of Fig. 121 we get,

$$R'_1 = R_1 + \left(\frac{\omega M}{Z_2}\right)^2 R_2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (108)$$

$$L'_1 = L_1 - \left(\frac{\omega M}{Z_2}\right)^2 \left(L_2 - \frac{1}{\omega^2 C_2}\right), \quad . \quad . \quad . \quad . \qquad (109)$$

in which

$$Z_2 = \sqrt{R_2{}^2 + \left(\omega L_2 - \frac{1}{\omega C_2}\right)^2}.$$

It is seen that if $\frac{1}{\omega^2 C_2}$ is greater than L_2, L'_1 is greater than L_1; if $L_2 = \frac{1}{\omega^2 C_2}$, $L'_1 = L_1$; if L_2 is greater than $\frac{1}{\omega^2 C_2}$, then L'_1 is less than L_1.

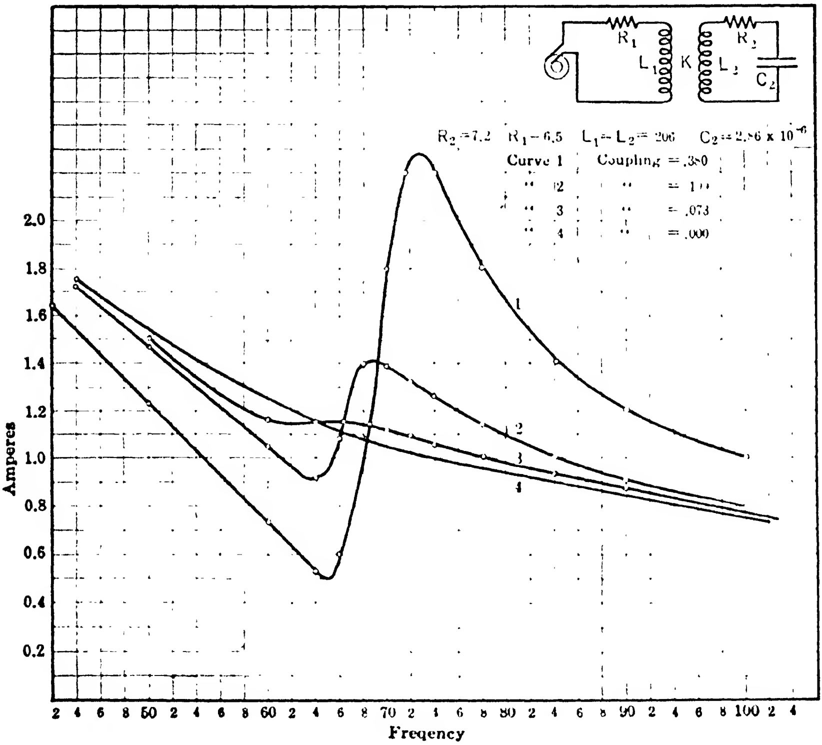


FIG. 122.—Current *vs.* frequency in primary circuit of Fig. 121.

Using the constants given in Eqs. (108) and (109) we can write at once

$$I_1 = \frac{E}{\sqrt{\left[R_1 + \left(\frac{\omega M}{Z_2}\right)^2 R_2\right]^2 + \omega^2\left[L_1 - \left(\frac{\omega M}{Z_2}\right)^2\left(L_2 - \frac{1}{\omega^2 C_2}\right)\right]^2}}. \quad . \quad . \quad (110)$$

$$I_2 = \frac{E\omega M}{Z_2\sqrt{\left[R_1 + \left(\frac{\omega M}{Z_2}\right)^2 R_2\right]^2 + \omega^2\left[L_1 - \left(\frac{\omega M}{Z_2}\right)^2\left(L_2 - \frac{1}{\omega^2 C_2}\right)\right]^2}}, \quad . \quad (111)$$

which may be somewhat simplified to the forms

$$I_1=\frac{EZ_2}{\sqrt{\left[\omega^2(M^2-L_1L_2)+\frac{L_1}{C_2}+R_1R_2\right]^2+\omega^2\left[L_1R_2+L_2R_1-\frac{R_1}{\omega^2C_2}\right]^2}}. \tag{112}$$

$$I_2=\frac{E\omega M}{\sqrt{\left[\omega^2(M^2-L_1L_2)+\frac{L_1}{C_2}+R_1R_2\right]^2+\omega^2\left[L_1R_2+L_2R_1-\frac{R_1}{\omega^2C_2}\right]^2}}. \tag{113}$$

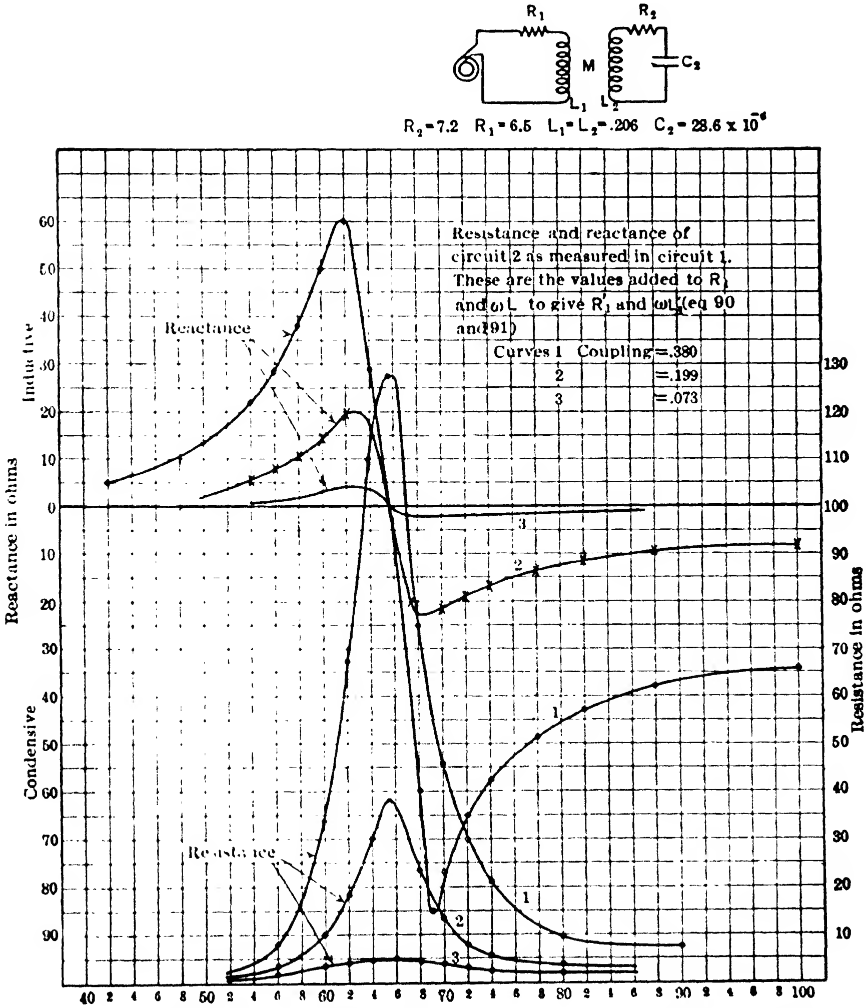


Fig. 123.—Change in primary resistance and reactance, due to presence of secondary circuit, for various frequencies.

In Fig. 122 are shown curves of primary current in such a circuit as given in Fig. 121; the voltage impressed on the primary circuit was held constant at 100 volts, and the frequency varied through a suitable range and current readings taken for three values of coupling between the two circuits. The value of primary current was also taken with secondary circuit open, and is shown in curve 4. From these curves it is seen that the effect of the secondary circuit may be either an increase or decrease in the primary current, depending upon the frequency used. In Fig. 123 are shown the values of change in primary resistance and reactance brought about by the action of the current in circuit 2; they were determined by subtracting from the apparent resistance and reactance of circuit 1 the values of these quantities when the secondary circuit was open.

A closer study of these curves will be worth while when analyzing the action of certain oscillating tube circuits. An oscillating tube may refuse to function if the resistance of the circuit to which it is connected is too high and it will be found that a tube may be made to stop oscillating by tuning to its frequency another circuit coupled to it. The reason is to be found in the extra value of the resistance added to the oscillating circuit by the second circuit *when this second circuit is brought into resonance with the tube circuit.*

Phase Relations in the Circuit of Fig. 121.—In Fig. 124 is shown the vector diagram of the reactions set up in the circuit of Fig. 121. Assuming a current of 1 ampere in circuit 1, there are developed in this circuit two reacting voltages, due to its resistance and self-induction. These are shown in Fig. 124 as OD and OC, respectively. Now the current I_1 induces the voltage $E_2(=\omega MI_1)$ in circuit 2, and this voltage lags 90° behind the inducing current I_1, and is so shown in Fig. 124.

The voltage E_2 sets up a current in circuit 2, the phase and magnitude of which depend upon the impedance of this circuit. For frequencies lower than that required for resonance in the second circuit, the current I_2 leads the voltage E_2, as shown at OA. This current induces a voltage back in circuit 1 in the phase shown at OB, 90° behind the current I_2. This voltage OB may be resolved into two components, OE and OF.

There is then set up in circuit 1, by a current of 1 ampere, a total voltage $OG(=OD+OF)$ 180° out of phase with the current I_1. This vector OG therefore is a measure of the effective resistance of circuit 1, and it is this magnitude, OG, which is specified by Eq. (108).

There is also set up in circuit 1 a reacting voltage of OH $(=OC+OE)$ in phase 90° behind I_1. This is then the magnitude to the total inductive reactance in circuit 1. It is to be noticed that the leading phase assumed for I_2 (that is, leading with respect to E_2) in this diagram results in an increase in the apparent inductance of circuit 1.

In case the capacity of circuit 2 is increased, or the impressed frequency is increased sufficiently that the impressed frequency is higher than the resonant value for circuit 2, then the current I_2 will lag behind E_2. This condition is shown in Fig. 125 in which the various vectors have the same meaning as in Fig. 124. The total effective resistance of circuit 1 is now OG, much larger than for the previous case. This is because Fig. 125 represents a condition closer to resonance in circuit 2 than does Fig. 124.

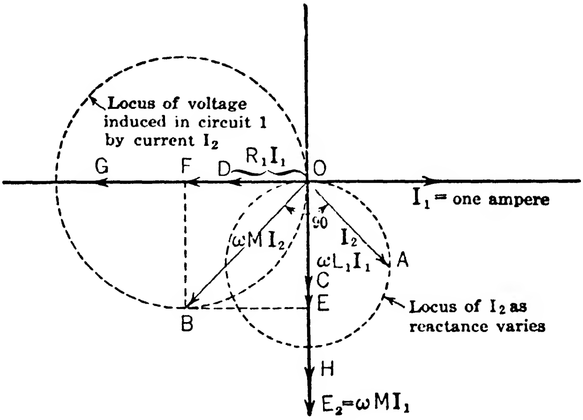


FIG. 124.—Vector relations for the circuit of Fig. 121.

The vector OE for Fig. 125 is subtractive when combined with the $\omega L_1 I_1$ vector, OC. The resultant effective voltage in circuit 1, in 90° phase relation with respect to I_1, is now OH, much smaller than OC. This means that the effective inductance of circuit 1 has been decreased by the effect of circuit 2.

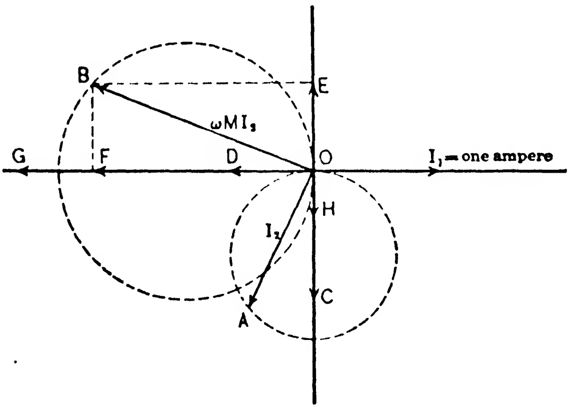


FIG. 125.—The relations of Fig. 124 may be changed to those shown here by increasing either secondary capacity or impressed frequency.

If we had assumed a slightly greater value for I_2 (in the same phase as in Fig. 125) or a little greater capacity in circuit 2 to make the current lag more, the vector OE would exceed in value OC, and the net vector would be leading I_1 by 90°. *This means that circuit 1, actually consisting of a coil and resistance, would draw a leading current from its power supply, acting as though it were a condenser.*

Application of Circuit of Fig. 121.—The circuit of Fig. 121 is used in practically every radio receiving set, and upon its correct design depends the efficiency of the amplifier. The elementary connection scheme of

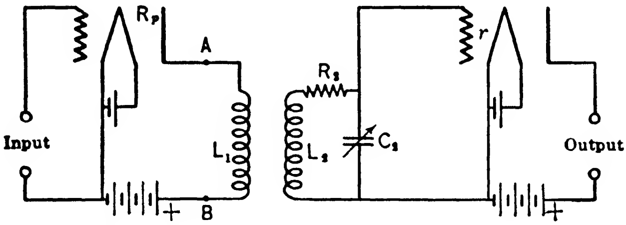


FIG. 126.—A common vacuum-tube circuit.

an amplifier is shown in Fig. 126, and this is reproduced in simplified form in Fig. 127.

The first tube of the amplifier contains in its plate circuit the coil L_1, and the second circuit, with the tuning condenser, feeds the power from the first tube into the second. The resistance R_p of the first tube (corresponding to R_1 of Fig. 127) is the a.c. resistance of the plate circuit of the vacuum tube, generally about 10,000 ohms or less. The resistance of coil L_1 is always negligible compared to R_p. The resistance R_2 is the resistance of coil L_2, generally a few ohms. The resistance, r, is the a.c. resistance of the input circuit of the second tube, generally 100,000 ohms or

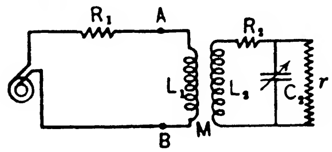


FIG. 127.—The simplified equivalent of Fig. 126.

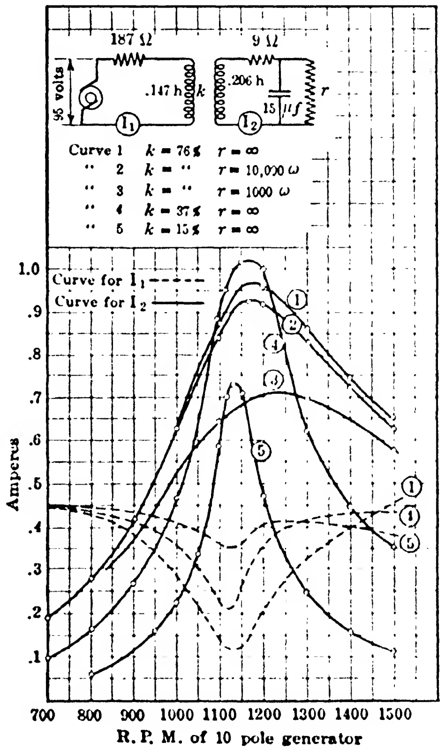


FIG. 128.—The frequency characteristics of the circuit of Fig. 126; in actual radio circuits the resonance phenomena are much sharper than those shown here.

more. The condenser C_2 is variable, having a useful capacity range of about 10 to 1; due to the circuit characteristics (stray capacity of coils, leads, etc.) its maximum capacity is not made less than about 300×10^{-12} farad.

Having the above-mentioned values, and neglecting the factor of selectivity for the moment, the circuit should be so designed that with a given voltage operating in the plate circuit of the first tube, there is obtained the maximum possible voltage across C_2, this controlling the strength of the amplified signal.

We first change the shunt resistance r into its equivalent series value. The basis for this equivalency is given on p. 255, where it is shown that we may write $R_3 = \frac{1}{(\omega C_2)^2 r}$, R_3 being the equivalent series resistance. As radio circuits are practically always used at resonance, we may write for ω^2 its equivalent, $\frac{1}{L_2 C_2}$, and so find the relation $R_3 = \frac{L_2}{C_2 r}$.

The total resistance of the second circuit is then $R_2 + \frac{L_2}{C_2 r}$. Assuming the second circuit resonant, we find the total effective resistance of the primary circuit $\left(\text{using Eq. (108) and putting } \omega L_2 = \frac{1}{\omega C_2}\right)$ to be

$$R'_1 = R_1 + \frac{\overline{\omega M}^2}{R_2 + \frac{L_2}{C_2 r}},$$

which can be simplified to the form

$$R'_1 = R_1 + \frac{M^2}{L_2 C_2 R_2 + L_2^2/r}. \quad . \quad . \quad . \quad . \quad . \quad (114)$$

To get the maximum power output from a tube, the load circuit (from A to B of Fig. 126) should be resistive only and should be equal to the tube resistance R_p. This is proved in Chapter VI. For best amplification, therefore, we may put, after neglecting, R_1, in comparison with the other term in Eq. (114).

$$R_p = \frac{M^2}{L_2 C_2 R_2 + L_2^2/r},$$

or

$$M^2 = L_2 C_2 R_p R_2 + \frac{L_2^2 R_p}{r}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (115)$$

In Fig. 128 are given the results of a test bringing out the relations analyzed above, and it is seen that the voltage across the load is a maximum at a certain frequency. The value of L_1, compared to the other constants of the circuit, was too large to have the test duplicate exactly, the relations existing in a radio amplifier circuit, but qualitatively the curves do illustrate the behavior of this type of apparatus.

It is to be remembered that when the optimum value of coupling is used (Eq. 115) the amount of resistance introduced into the plate circuit (Fig. 126) by the L_2–C_2 circuit is just equal to R_p, thus making the effective resistance of this circuit equal to $2R_p$. Also the effect of the plate circuit on the L_2–C_2 circuit is to give this circuit an effective resistance just twice its own true effective resistance. Thus each circuit acts to double the resistance of the other. The doubling of the resistance of the L_2–C_2 circuit has the undesirable result of doubling its decrement, thus materially diminishing the selectivity of the circuit.

Radio sets are seldom designed with a value of coupling as great as that given by Eq. (115); somewhat less than maximum possible amplification is obtained in the interest of selectivity. A typical circuit, tuned for 1000 kc., had constants as follows: $R_2=7$, $L_1=10\times10^{-6}$, $L_2=200\times10^{-6}$, $C=100\times10^{-12}$, $R_p=10^4$, $r=2\times10^5$, $k=50$ per cent. We find M to be $0.5\sqrt{10\times200}=22.4$ microhenries. By using Eq. (115) we find that for maximum amplification M should be equal to 58.4 microhenries.

Coupled Resonant Circuits.—We next consider the more general case of two coupled circuits, each of which has inductance, capacity, and resistance, as indicated in Fig. 129. The resistance and inductance of circuit 1 are obtained from Eqs. (95) and (96), as before.

$$R'_1=R_1+\left(\frac{\omega M}{Z_2}\right)^2 R_2. \qquad (116)$$

$$L'_1=\left(L_1-\frac{1}{\omega^2C_1}\right)-\left(\frac{\omega M}{Z_2}\right)^2\left(L_2-\frac{1}{\omega^2C_2}\right). \qquad (117)$$

Then we have,

$$I_1=\frac{E}{\sqrt{\left[R_1+\left(\frac{\omega M}{Z_2}\right)^2 R_2\right]^2+\omega^2\left[\left(L_1-\frac{1}{\omega^2C_1}\right)-\left(\frac{\omega M}{Z_2}\right)^2\left(L_2-\frac{1}{\omega^2C_2}\right)\right]^2}}. \qquad (118)$$

$$I_2=\frac{E\omega M}{Z_2\sqrt{\left[R_1+\left(\frac{\omega M}{Z_2}\right)^2 R_2\right]^2+\omega^2\left[\left(L_1-\frac{1}{\omega^2C_1}\right)-\left(\frac{\omega M}{Z_2}\right)^2\left(L_2-\frac{1}{\omega^2C_2}\right)\right]^2}}, \qquad (119)$$

in which

$$Z_2=\sqrt{R_2^2+\left(\omega L_2-\frac{1}{\omega C_2}\right)^2}.$$

In solving for resonant frequency we may assume that the resistance term of the impedance in Eq. (118) is nearly constant as the frequency is varied. The fraction $\omega M/Z_2$ is evidently nearly constant as ω is varied until ω approaches such a value that $(\omega L_2 - 1/\omega C_2)$ is nearly equal to zero. In this region of frequency variation the resistance R'_1 varies greatly as frequency is varied and any solution which we may reach on the basis of R'_1 remaining constant, therefore, will not be accurate for frequencies in the vicinity of the natural frequency of the secondary circuit.

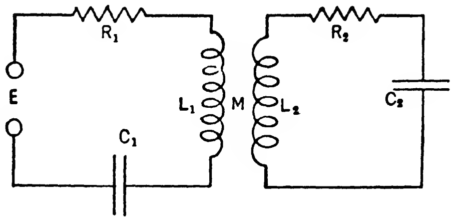


FIG. 129.—General case of inductively coupled circuits.

Resonant Frequencies in Coupled Circuits.—On the assumption that R'_1 is constant it is evident that I_1 will be a maximum for any frequency that makes the reactance term of the impedance equal to zero. Hence we write as the condition for resonance

$$L_1 - \frac{1}{\omega^2 C_1} - \left(\frac{\omega M}{Z_2}\right)^2 \left(L_2 - \frac{1}{\omega^2 C_2}\right) = 0. \quad . \quad . \quad . \quad (120)$$

If now we again neglect R_2 in comparison with $\left(\omega L_2 - \frac{1}{\omega C_2}\right)$ thus making it possible to replace Z_2 by $\left(\omega L_2 - \frac{1}{\omega C_2}\right)$, we have

$$L_1 - \frac{1}{\omega^2 C_1} - \frac{M^2}{\left(L_2 - \frac{1}{\omega^2 C_2}\right)^2}\left(L_2 - \frac{1}{\omega^2 C_2}\right) = 0. \quad . \quad . \quad . \quad (121)$$

Now Eq. (121) can be written,

$$L_1 L_2 - \frac{L_1}{\omega^2 C_2} - \frac{L_2}{\omega^2 C_1} + \frac{1}{\omega^4 C_1 C_2} - M^2 = 0,$$

which can be changed, by multiplying through by $\frac{\omega^4}{L_1 L_2}$, to the form

$$\omega^4 - \frac{\omega^2}{L_2 C_2} - \frac{\omega^2}{L_1 C_1} + \frac{1}{L_1 C_1 L_2 C_2} - \frac{\omega^4 M^2}{L_1 L_2} = 0. \quad . \quad . \quad . \quad (122)$$

If we now put

$$\omega_1 = \frac{1}{\sqrt{L_1 C_1}}, \quad \omega_2 = \frac{1}{\sqrt{L_2 C_2}} \quad \text{and } k = \frac{M}{\sqrt{L_1 L_2}},$$

we get

$$\omega^4(1 - k^2) - \omega^2(\omega_1^2 + \omega_2^2) + \omega_1^2 \omega_2^2 = 0. \quad . \quad . \quad . \quad (123)$$

The solution of this equation is obtained by dividing through by $(1-k^2)$, properly completing the square of the left-hand member and extracting the square root, which gives

$$\omega^2=\frac{\omega_1{}^2+\omega_2{}^2\pm\sqrt{(\omega_1{}^2+\omega_2{}^2)^2-4(1-k^2)\omega_1{}^2\omega_2{}^2}}{2(1-k^2)}.$$

By combining terms under the radical this becomes

$$\omega^2=\frac{\omega_1{}^2+\omega_2{}^2\pm\sqrt{(\omega_1{}^2-\omega_2{}^2)^2+4k^2\omega_1{}^2\omega_2{}^2}}{2(1-k^2)}.$$

The two real solutions for ω, which we call ω' and ω'', are

$$\omega'=\sqrt{\frac{\omega_1{}^2+\omega_2{}^2-\sqrt{(\omega_1{}^2-\omega_2{}^2)^2+4k^2\omega_1{}^2\omega_2{}^2}}{2(1-k^2)}}, \quad . \quad . \quad . \quad (124)$$

and

$$\omega''=\sqrt{\frac{\omega_1{}^2+\omega_2{}^2+\sqrt{(\omega_1{}^2-\omega_2{}^2)^2+4k^2\omega_1{}^2\omega_2{}^2}}{2(1-k^2)}}. \quad . \quad . \quad . \quad (125)$$

When k is large (approximately unity) the values of ω' and ω'' are nearly

$$\omega'=\sqrt{\frac{\omega_1{}^2\omega_2{}^2}{\omega_1{}^2+\omega_2{}^2}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (126)$$

and

$$\omega''=\sqrt{\frac{\omega_1{}^2+\omega_2{}^2}{1-k^2}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (127)$$

When k is small the values of ω' and ω'' approach the limits

$$\omega'=\frac{\omega_2}{\sqrt{1+k^2}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (128)$$

and

$$\omega''=\frac{\omega_1}{\sqrt{1-k^2}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (129)$$

In Fig. 130 are shown the relations between ω' and ω'' and k; for small values of k Eqs. (128) and (129) determine the values and for the large values of k Eqs. (126) and (127) are used.

In radio operation it is the practice to tune the primary and secondary circuits, that is, adjustments are made to make ω_1 equal to ω_2. In this case Eqs. (124) and (125) reduce to the very simple forms

$$\omega'=\frac{\omega}{\sqrt{1+k}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (130)$$

and

$$\omega''=\frac{\omega}{\sqrt{1-k}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (131)$$

in which $\omega=\omega_1=\omega_2$.

In using Eqs. (130) and (131) it is convenient to remember that, for small values of k, we may write $\dfrac{1}{\sqrt{1+k}}=1-k/2$ and $\dfrac{1}{\sqrt{1-k}}=1+k/2$. These approximations are quite accurate for the values of coupling generally used in radio circuits. Thus with a coupling of 10 per cent the error in the above approximations is less than $\frac{1}{2}$ per cent. With k as high even as 20 per cent the error is only about 1 per cent.

The curves of variation in ω' and ω'' as the coupling is varied for this case of tuned circuits are shown in Fig. 131. It is seen that for weak coupling both ω' and ω'' approach ω, the natural frequency of each circuit; however, it has been pointed out that the neglect of R_2 in obtaining the solutions of the resonant frequencies that the values of ω' and ω'' do not hold

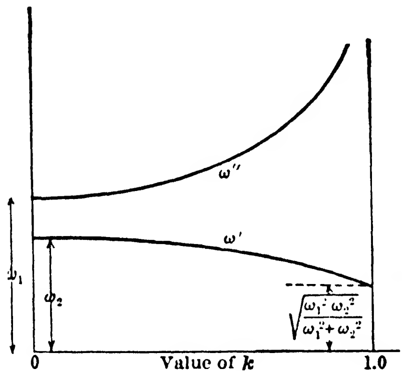


FIG. 130.—Variation of ω' and ω'' with k in coupled circuits, primary and secondary not tuned.

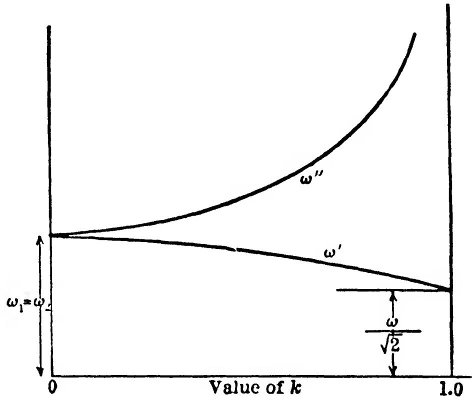


FIG. 131.—Variation of ω' and ω'' with k in tuned coupled circuits.

good when they have values in the vicinity of the natural frequency of the secondary circuit. Hence we can now see that *for weak couplings the solutions for ω' and ω'' do not quite hold good.*

Referring to Eq. (119) it is seen that $I_2=I_1\dfrac{\omega M}{Z_2}$, and hence in so far as the factor $\dfrac{\omega M}{Z_2}$ is independent of the frequency changes, *I_2 will have maximum values at the same frequencies as give maxima for I_1.* However, the factor $\dfrac{\omega M}{Z_2}$ is not independent of the frequency, and this is especially so in the region of frequency fixed by the relation $\left(\omega L_2-\dfrac{1}{\omega C_2}\right)=0$; for frequencies less than this the value of $\frac{\omega M}{Z_2}$ increases with the frequency and for values of frequency higher, the value of $\frac{\omega M}{Z_2}$ decreases with an increase of frequency.

We can then conclude that, for frequencies in the region of

$$\omega_2 = \frac{1}{\sqrt{L_2 C_2}}.$$

Eqs. (130) and (131), while slightly incorrect for primary current maxima, are somewhat more incorrect for the maxima of secondary current. In consequence of the changes in the value of $\frac{\omega M}{Z_2}$ noted above, we may predict that when ω' and ω'' are not in the region of ω_2 the calculated values of ω'

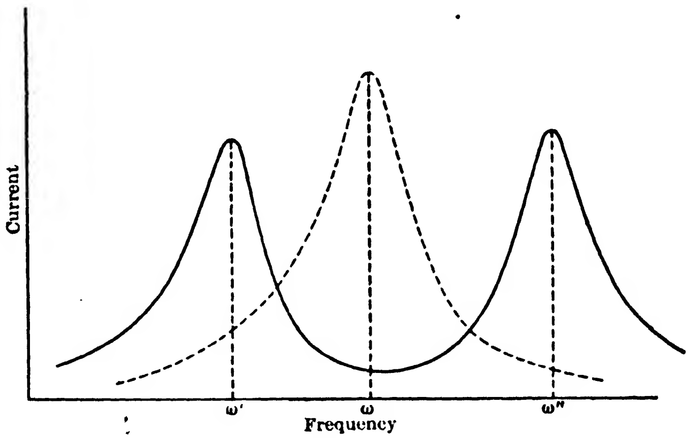


FIG. 132.—Resonance curves for circuit of Fig. 129.

and ω'' will be more accurate for the primary than for the secondary circuit, and that the actual value of ω' of the secondary circuit will be somewhat higher than that for the primary and that the actual value of ω'' for the secondary will be somewhat lower than ω'' for the primary current.

The general form of the resonance curve of the circuit shown in Fig. 129 is indicated in Fig. 132; the dotted curve shows the resonance for one circuit by itself.

The value of the coefficient of coupling can be calculated from the spacing of the resonance peaks of the current curves; thus from Eqs. (130) and (131) we get the relation

$$\frac{\omega'^2}{\omega''^2} = \frac{1-k}{1+k},$$

from which there is obtained

$$k=\frac{\omega''^2-\omega'^2}{\omega''^2+\omega'^2}. \quad \ldots \ldots \ldots \quad (132)$$

In case the resonance frequency of one circuit by itself is known, and assuming tuned circuits, the equation for coupling value becomes more simple in form, giving the closely approximate value

$$k=\frac{\omega''-\omega'}{\omega}, \quad \ldots \ldots \ldots \quad (133)$$

ω being the frequency of one circuit by itself.

In the foregoing discussion of resonant frequencies formulas have been derived using ω for frequency; it is of course to be remembered that ω is

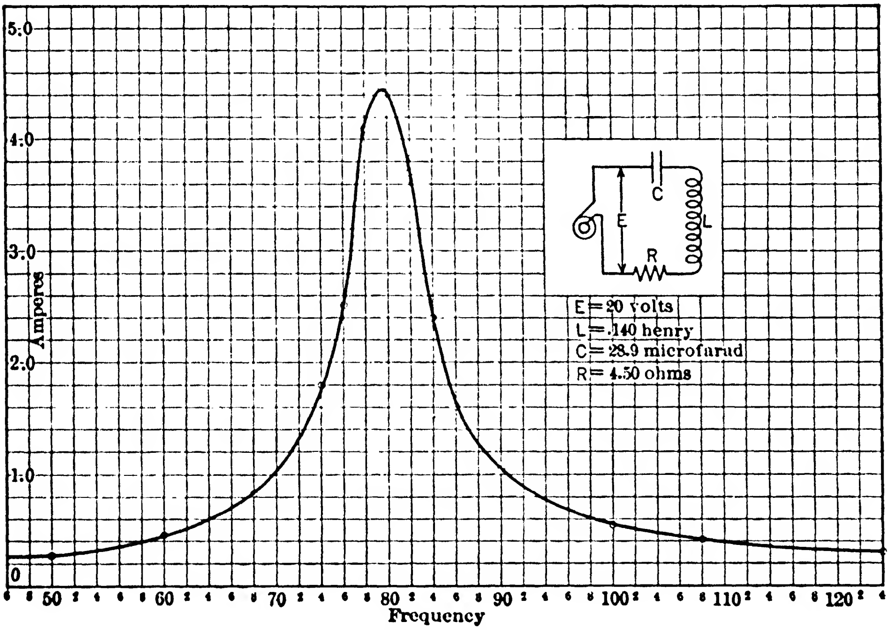


FIG. 133.—Experimental resonance curve for single circuit.

not frequency, but 2π times the frequency. The value of ω has been used rather than frequency itself to save the repeated writing of the quantity 2π throughout all the derivations.

In Figs. 133 to 140 are shown some experimental curves of resonance in coupled circuits for different conditions as regards coupling, resistances, tuning, etc.; Fig. 133 shows the resonance curve for a single circuit having $L=0.140$ henry, $C=28.9$ microfarads, and $R=4.50$ ohms.

Fig. 134 shows the resonance curves for two coupled circuits, each circuit had the same constants as those given for Fig. 133; the coefficient of coupling was 0.36. For these curves, as well as those of the two following figures, it was intended that the circuits should be tuned to each other, as indicated on the curve sheets. Actually there was a defective connection which slightly mistuned the two circuits; this accounts for one peak of current being higher than the other. The curve of primary current is shown by the full line and that for the secondary circuit by the dotted line. The two resonant frequencies check with those calculated from Eqs. (130) and (131) within the precision of the test.

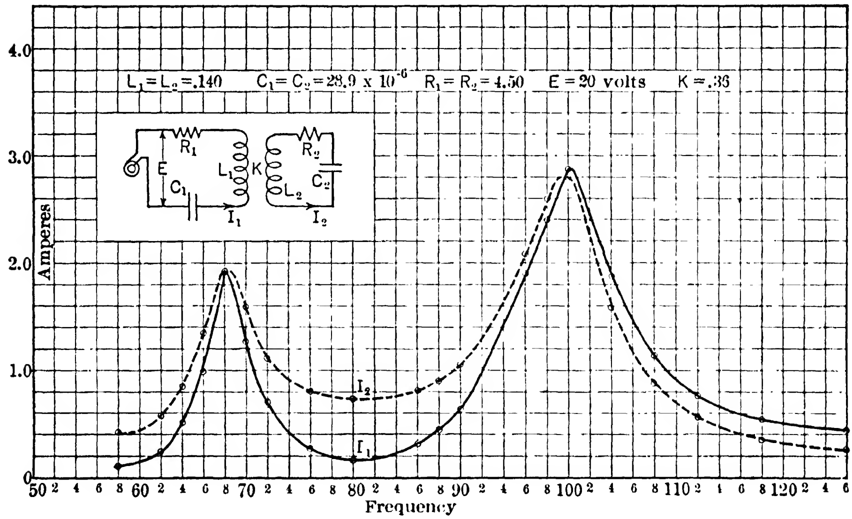


FIG. 134.—Resonance curve for coupled circuits, each circuit having approximately constants as in Fig. 133. $k = 0.36$.

In Figs. 135 and 136 are shown curves of current for the same two circuits as those used in Fig. 134 but with different values of coupling, this being 0.18 for Fig. 135 and 0.07 for Fig. 136. It may be seen that with small values of coupling the two frequencies merge into one another and Eqs. (130) and (131) do not predict accurately the resonant frequencies of the primary circuit and for reasons noted in the derivation of the formulas; the predicted values of ω' and ω'' for the secondary circuit differ from the actual values more than do those of the primary circuit.

A peculiarity of all these resonance curves is seen in the relative values of the primary and secondary currents; between the two resonant frequencies the secondary circuit carries a greater current than the primary, but for all other frequencies the primary carries a greater current. If a weaker coupling than that used in the adjustments for Fig. 136 had been used it would have been found that the primary current was greater than the secondary current for all values of frequency.

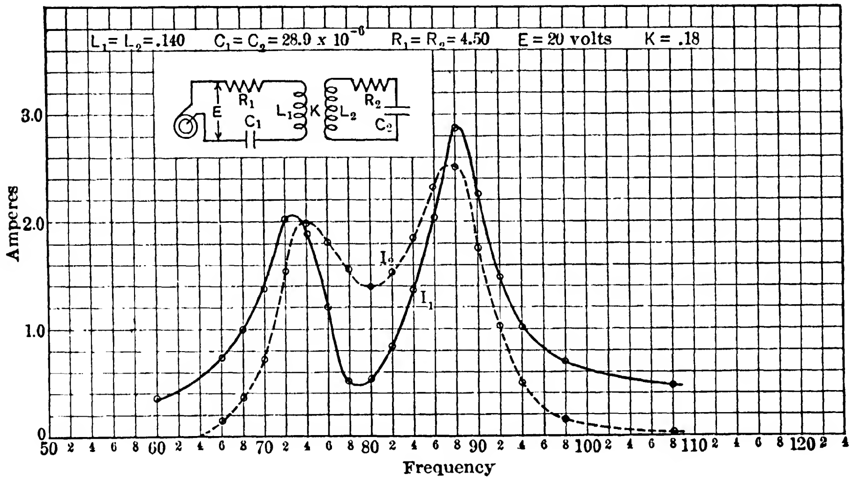


FIG. 135.—Resonance curves for circuit as shown in Fig. 134, $k = 0.18$.

In Fig. 137 is shown the result of increasing the resistance of the secondary circuit from 4.5 to 9.7 ohms; with this exception the circuits

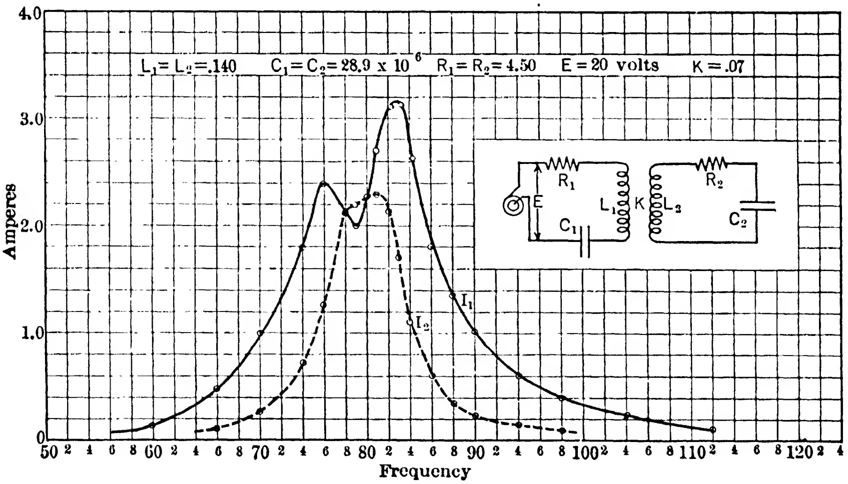


FIG. 136.—Resonance curves for circuit as shown in Fig. 134, $k = 0.07$.

were exactly the same as those used for Fig. 134. By comparison of the two sets of curves it will be seen that the two resonant frequencies are, within the precision of measurements, the same for the two conditions; the value of the current at resonance is, however, decreased in nearly the proportion predicted from the value of resistance, calculated from Eq. (116). The decrease in current, it will be noted, takes place in both circuits although the resistance of the secondary circuit only was increased. The resonance is much less marked than for the lower resistance used in Fig. 134.

Form of Resonance Curve.—The form of the resonance peaks is determined by the combined decrements of both circuits. For the simplest

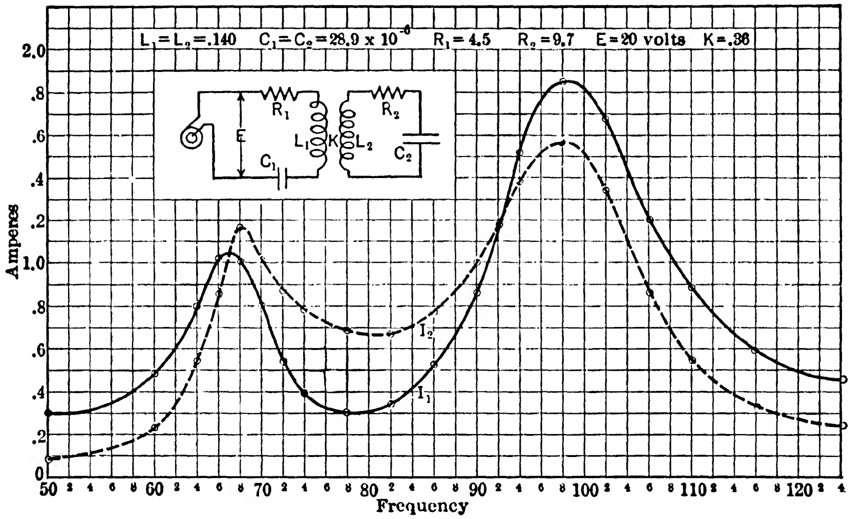


FIG. 137.—Resonance curves for circuit shown in Fig. 134 with added secondary resistance.

case, that of tuned circuits, it will be found that the decrements will be nearly given by the approximate formulas:

For the frequency ω'

$$\delta' = \frac{\delta_1+\delta_2}{2\sqrt{1+k}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (134)$$

and for the frequency ω''

$$\delta'' = \frac{\delta_1+\delta_2}{2\sqrt{1-k}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (135)$$

in which δ_1 and δ_2 are the decrements of circuits 1 and 2 when not affected by other circuits.

The decrements δ' and δ'', calculated from the shape of the curves of Figs. 134 and 135 by use of Eq. (52), check with the values given by Eqs.

(134) and (135) fairly well; it is noticeable that in all the curves given the width of the resonance curve is greater for the higher frequency than for the lower, indicating thereby a greater decrement. With weak coupling the form of the curves does not permit the calculation of δ' and δ'', because the two peaks merge into one.

Circuits not Tuned.—In Figs. 138 and 139 are shown the resonance curves for two circuits which are not tuned, that is, ω_1 is not equal to ω_2. For this condition the curves are not as symmetrical as for the tuned condition, and the currents in the two circuits are no longer nearly equal to each other at the two resonant frequencies. At one resonant

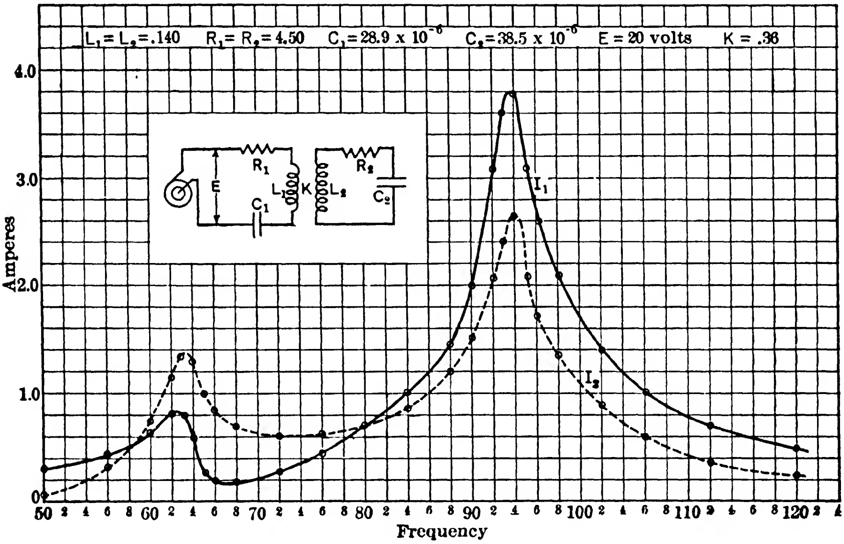


Fig. 138.—Resonance curves for coupled untuned circuits.

frequency the primary circuit carries more current than the secondary and at the other frequency the reverse is true. The difference in the two currents is greater the greater the difference in the natural periods of the two circuits. For Fig. 138 the natural frequency of circuit 2 is 15 per cent lower than that of circuit 1, and for Fig. 139 circuit 2 has a natural frequency 29 per cent lower than that of circuit 1.

Effect of Mistuning on the Shape of the Resonance Curve.—The data given in Figs. 134–136 state that the two circuits were tuned to the same frequency, but it was found in checking them that there was a slight discrepancy in the *LC* products. This inequality accounts principally for the fact that the current value at one resonance frequency is higher than at the other.

We will first assume tuned circuits, having equal inductances and capacities. At the two resonance frequencies, ω' and ω'', we find from Eq. (118) that, for low power factor coils, the primary current is given by $E/(R_1+R_2)$. At the two resonance frequencies, ω' and ω'', the circuits show unity power factor so that the primary current I_1 is limited only by the effective resistance of this circuit. Using Eq. (116), we find the effective resistance R'_1 to be equal to R_1+R_2 at each of the frequencies ω' and ω'', so this checks the result which we just stated could be obtained from Eq. (118).

At very low values of coupling we have shown that Eqs. (130) and (131) are not quite correct; in the same way the statement that

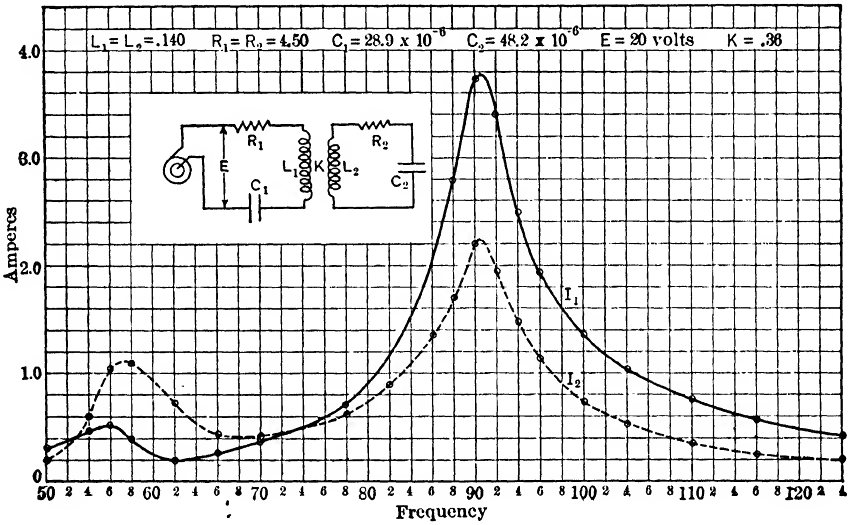


FIG. 139.—Resonance curves for coupled untuned circuits.

$I_1=E/(R_1+R_2)$ at ω' and ω'' is incorrect if very low values of coupling are assumed. For values of coupling equal to, or greater than, the power factor of the coil in the secondary circuit, the above approximations are quite accurate.

From the statement that $I_1=E/(R_1+R_2)$ at both ω' and ω'', it follows that the values of primary current at both resonance peaks should be the same. Further, as $I_2=\omega MI_1/Z_2$, it can be shown that the two peaks of secondary current will have the same magnitude.

If the condenser of either circuit is decreased, both ω' and ω'' are increased, whereas a decrease in ω' and ω'' takes place if either condenser is increased. The same idea holds good if either inductance is changed.

We can calculate the effect on I_1 of changing C_1 and C_2 with the help of Eq. (116), after making the approximation that at the frequencies ω' and ω'', Z_2 has the same value as $\omega L_2 - 1/\omega C_2$. Then we have

$$R'_1 = R + \left(\frac{\omega M}{\omega L_2 - \frac{1}{\omega C_2}} \right)^2 R_2. \quad . \quad . \quad . \quad . \quad . \quad (136)$$

By noticing the changes that take place in R'_1, as ω' and ω'' are changed by altering either C_1 or C_2, we come to the conclusion that if C_1 is decreased, I_1 has its greater peak at ω'', and if C_1 is increased, I_1 has its greater peak at ω'. The amount of change in peak values as the circuits are mistuned depends upon the value R_2, being greater as this is increased.

A series of calculations on two theoretical circuits shows results as follows:

I. Tuned circuits:

$L_1 = L_2 = 1$ henry. $C_1 = C_2 = 100$ microfarads.

$R_1 = R_2 = 1$ ohm. $E_1 = 20$ volts. $K = 0.40$.

(It is well for the student to appreciate the fact that there is a certain correlation between the inductance, resistance, and weight of a coil. Thus we have here assumed 1 henry inductance and 1 ohm resistance; such a coil could be built only by using many turns of large-sized wire to make a coil of 2 or 3 ft. diameter. The coil would weigh several hundred pounds!)

$\omega' = 84.6$.	R'_1 at $\omega' = 2$.	I_1 at $\omega' = 10$ amperes.
$\omega'' = 129.2$	R'_1 at $\omega'' = 2$.	I_1 at $\omega'' = 10$ amperes.

II. Mistuned circuits: C_2 increased to 156 microfarads.

$\omega' = 73.0$.	R'_1 at $\omega' = 4.8$.	I_1 at $\omega' = 4.17$ amperes.
$\omega'' = 119.0$.	R'_1 at $\omega'' = 1.5$.	I_1 at $\omega'' = 13.4$ amperes.

III. Mistuned circuits: C_2 back at 100 microfarads. $C_1 = 156$ mf.

$\omega' = 73$.	R'_1 at $\omega' = 1.23$.	I_1 at $\omega' = 16.3$ amperes.
$\omega'' = 119$.	R'_1 at $\omega'' = 2.90$.	I_1 at $\omega'' = 6.9$ amperes.

In all cases I_2 is given by the relation

$$I_2 = \frac{\omega M}{Z_2} I_1$$

so the peak values of I_2 at ω' and ω'' can readily be calculated.

Experimental results showing the effect of such a capacity change as that calculated above are given in Fig. 140. In curve *A* the two circuits were tuned and both peaks of primary current have essentially the same magnitude. Doubling the value of condenser used in either circuit decreased both ω' and ω'' by about *half* the percentage change which would have occurred if *both* circuits had been subjected to the same capacity change.

The new values of ω' and ω'' are the same no matter which circuit has its condenser changed. When the value of C_1 is increased the greater current peak occurs at the lower frequency, ω', and if the secondary

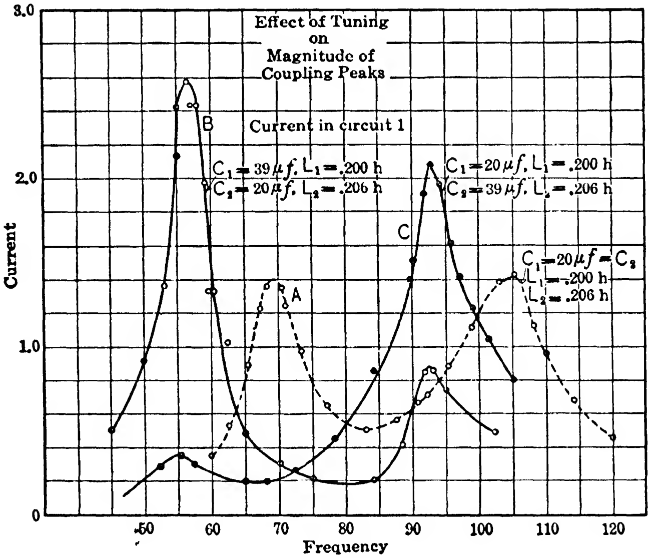


FIG. 140.—As the two circuits are mistuned, one peak becomes more prominent than the other; all three curves are for current in the primary circuit.

condenser is the one which is increased, then the greater current peak occurs at the higher frequency, ω''.

Variation of Coupling with Tuned Circuits.—In Fig. 141 is shown the effect of varying the coupling between circuits 1 and 2, they being tuned alike. A constant e.m.f. was impressed on circuit 1 and the coupling of the two circuits was gradually increased from zero to the maximum obtainable. It might seem at first sight that the secondary current would be greater the greater the coupling, as would occur in ordinary transformer tests, but with tuned circuits as used in radio this is not the case. For a given resistance of circuits there will be a certain coupling which gives the greatest secondary current and the lower the resistance of the circuits the less this critical value of coupling will be.

This might be predicted from Eq. (119) by differentiating I_2 with respect to M; it will be found that with tuned circuits having impressed on the primary a voltage of the same frequency as that for which the circuits are tuned, a certain value of M will produce a maximum secondary current and this value of M will depend upon the resistances in the two circuits. This condition for maximum secondary current proves to be fixed by the relation,

$$\omega^2M^2=R_1R_2. \qquad (137)$$

The curves of Fig. 142 were taken with the idea of proving this relation

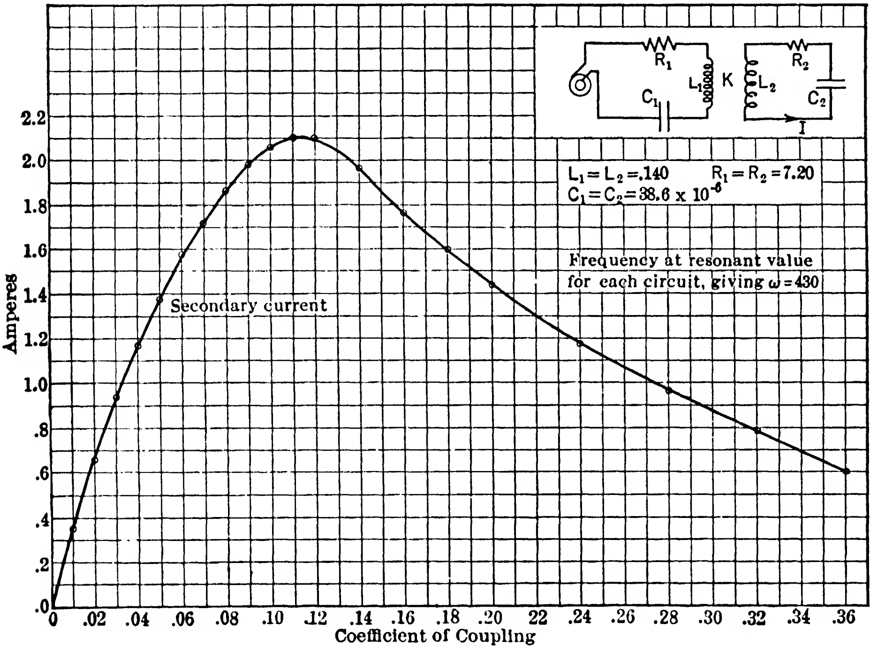


FIG. 141.—Secondary current *vs.* coefficient of coupling in tuned coupled circuits.

and also to show the effect of the secondary resistance on the sensitiveness of the adjustment for maximum secondary current. If the two circuits are tuned alike and the frequency of the e.m.f. impressed on the primary is the same as the natural frequency of either circuit the values of the primary and secondary current may be obtained by simplifying Eqs. (118) and (119) and are found to be

$$I_1=\frac{ER_2}{R_1R_2+\omega^2M^2}. \qquad (138)$$

and

$$I_2=\frac{E\omega M}{R_1R_2+\omega^2M^2}. \qquad (139)$$

The experimental curves given in Fig. 142 follow the values predicted from Eqs. (138) and (139) within the precision of measurement, that is, within less than 1 per cent.

Resonance in Circuits with Capacitive Coupling.—The equations for I_1 and I_2 are obtained for this case in a fashion exactly the same as that used for the magnetic coupling, and the conclusions reached are nearly the same. Using ω_1, ω_2, and k in the same sense as for the magnetically coupled circuits we get for the two resonant frequencies of the combination

$$\omega' = \sqrt{\frac{\omega_1^2+\omega_2^2+\sqrt{(\omega_1^2-\omega_2^2)^2+4k^2\omega_1^2\omega_2^2}}{2}} \quad . \quad . \quad . \quad . \quad (140)$$

$$\omega'' = \sqrt{\frac{\omega_1^2+\omega_2^2-\sqrt{(\omega_1^2-\omega_2^2)^2+4k^2\omega_1^2\omega_2^2}}{2}} \quad . \quad . \quad . \quad . \quad (141)$$

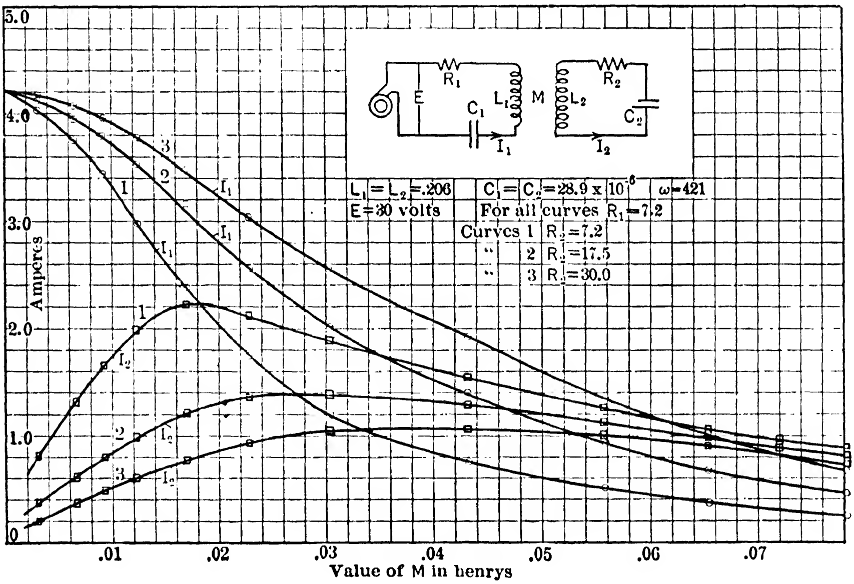


FIG. 142.—Current *vs.* mutual inductance in tuned coupled circuits, with different values of secondary resistance.

In applying these formulas the values of ω_1 and ω_2 must be calculated in a manner somewhat different from that used for the magnetically coupled circuits. It will be remembered that for magnetic coupling these two frequencies were fixed by the L and C of the circuit in question and were independent of the constants of the other circuit and of the coupling used. Such is not the case for capacitive coupling, however. The frequencies

ω_1 and ω_2 depend upon the capacity used in the other circuit and upon the coupling in the following manner.

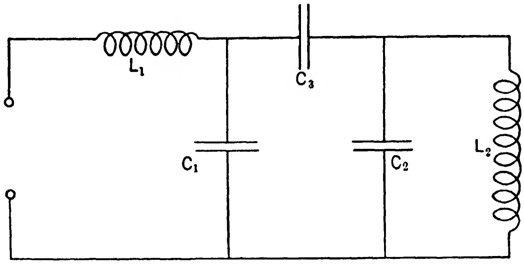


FIG. 143.—Capacitively coupled circuits.

In Fig. 143 the frequency ω_1 is fixed by L_1 and by the capacity C_1 in parallel with C_3 and C_2 in series. Thus ω_1 may be varied by changing either the coupling condenser C_3, or the capacity of the second circuit C_2. Hence we have the formulas

$$\omega_1=\frac{1}{\sqrt{L_1\left(C_1+\dfrac{1}{\dfrac{1}{C_2}+\dfrac{1}{C_3}}\right)}}=\frac{1}{\sqrt{L_1\left(C_1+\dfrac{C_2C_3}{C_2+C_3}\right)}} \quad . \quad . \quad (142)$$

and

$$\omega_2=\frac{1}{\sqrt{L_2\left(C_2+\dfrac{C_1C_3}{C_1+C_3}\right)}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (143)$$

For the value k we have

$$k=\frac{C_3}{\sqrt{(C_1+C_3)(C_2+C_3)}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (144)$$

In Fig. 144 are shown the resonance curves for a combination of circuits nearly like that shown in Fig. 143; the coupling condenser was in two parts as shown in the sketch on the curve sheet.

Using the values of L and C indicated on the curve sheet we have

$$\omega_1=\frac{10^3}{\sqrt{0.206\left(18.3+\dfrac{4.55\times18.3}{4.55+18.3}\right)}}=470$$

or $f_1=74.8$ cycles and the same value for f_2.

From the curve sheet $f''=82.5$ and $f'=67.0$, so we find k from the curve sheet to be 0.208, by Eq. (133) and estimating the frequency of one circuit alone as the average of f'' and f'. By using the known values of C_1, C_2, and C_3, and Eq. (144), we find k to be 0.207.

The value of k could have been calculated without knowing the constants of the circuits by using Eq. (132).

We have

$$k=\frac{82.5^2-67.0^2}{82.5^2+67.0^2}=0.207.$$

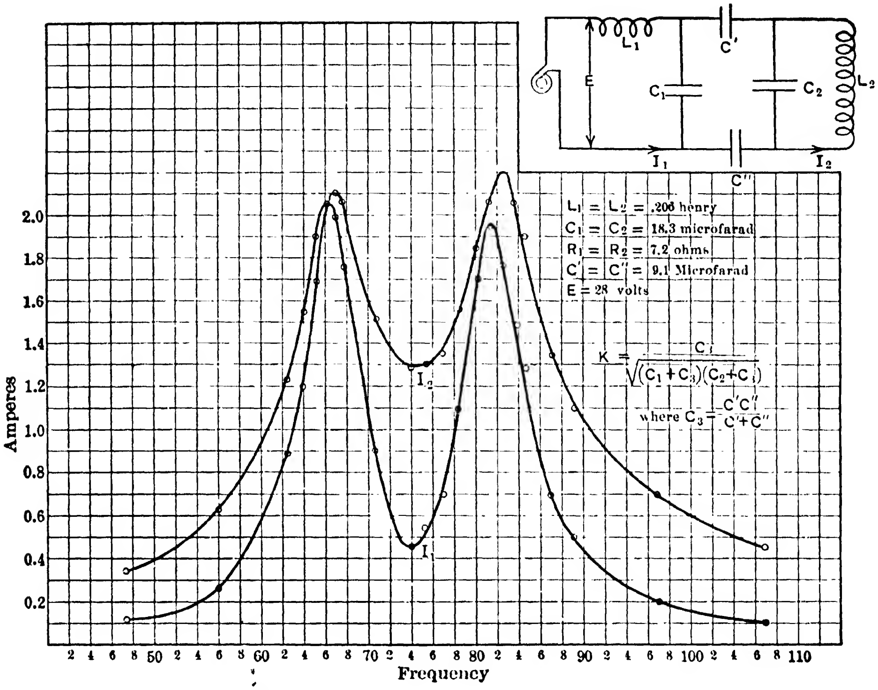


FIG. 144.—Resonance curves for capacitively coupled circuits.

When $\omega_1=\omega_2$, Eqs. (140) and (141) reduce to the simple forms

$$\omega'=\omega\sqrt{1+k}, \quad \ldots \ldots \ldots \quad (145)$$

$$\omega''=\omega\sqrt{1-k}, \quad \ldots \ldots \ldots \quad (146)$$

and if further $C_1=C_2$, then when C_3 is varied, thus varying the coupling, ω' stays constant and equal to $\dfrac{1}{\sqrt{L_1C_1}}$.

Special Forms of Coupled Circuits.—It has been shown in the previous paragraphs that two coupled circuits, each having coils and condensers, have two resonance frequencies, and that the values of both of these frequencies change as the coupling is altered, one increasing and the other decreasing as the coupling is tightened.

It is possible to design coupled circuits with such constants, however, that one of the two resonant frequencies remains constant as the coupling is changed.

In Fig. 145 is shown a conductively coupled circuit, the coil L_m being common to both circuits. If we choose $L_1 = L_2 = L$ and $C_1 = C_2 = C$, we find for the two resonant frequencies

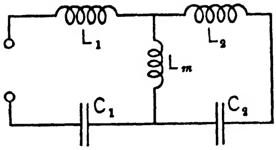


FIG. 145.—A simple conductively coupled pair of circuits.

$$\omega' = \frac{1}{\sqrt{(L+2L_m)C}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (147)$$

and

$$\omega'' = \frac{1}{\sqrt{LC}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (148)$$

It is evident that ω'' is independent of the coupling determined by the ratio of L_m to $L+L_m$.

As the coupling coil L_m is increased in value the two resonance frequencies do separate, as in other cases of coupled circuits, but all of

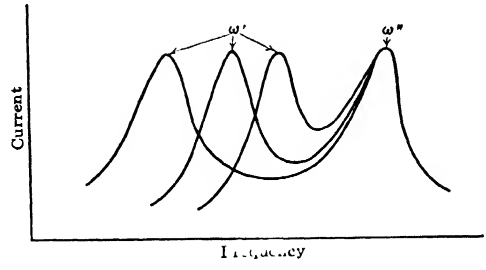


FIG. 146.—As the coupling coil of Fig. 145 is varied, one of the resonance peaks stays fixed at a certain frequency.

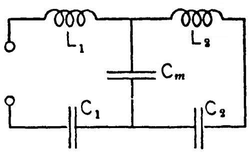


FIG. 147.—A simple capacitively coupled pair of circuits.

the frequency change takes place in the lower frequency. The resonance curves showing this effect are given in Fig. 146, for three values of the coupling coil L_m.

The same peculiar behavior of the resonance frequencies may be obtained with capacitively coupled circuits. With an arrangement as in Fig. 147, there are two resonance frequencies, the separation of which depends upon the degree of coupling. The condenser C_m affects the coupling, this being low if C_m is large and approaching unity as C_m approaches zero. For the circuit of Fig. 147 we have in general [1]

$$k=\sqrt{\frac{C_1C_2}{(C_1+C_m)(C_2+C_m)}}. \qquad (149)$$

$$\omega_1=\sqrt{\frac{C_1+C_m}{L_1C_1C_m}}. \qquad (150)$$

$$\omega_2=\sqrt{\frac{C_2+C_m}{L_2C_2C_m}}. \qquad (151)$$

And if $C_1=C_2=C$ and $L_1=L_2=L$ we have

$$\omega'=\sqrt{\frac{2C+C_m}{LCC_m}} \qquad (152)$$

and

$$\omega''=\sqrt{\frac{1}{LC}}, \qquad (153)$$

which latter frequency is evidently independent of coupling.[2]

Characteristics of Filters.[3]—From the material already given in this chapter it is evident that coils tend to pass low-frequency currents much more readily than high ones, and that the reverse is true for condensers. Further, that a coil and condenser in series pass one certain frequency and tend to block out all others, and that a coil and condenser in parallel tend to block one certain frequency and pass all others, either higher or lower. Also that by using two similarly tuned circuits coupled together loosely, a low impedance is offered to all frequencies within a certain band, and high to all outside this band, and that if the coupling is increased two certain frequencies may be passed readily and all others eliminated.

By inspection of curves, and formulas developed in the earlier part of this chapter, it may also be concluded that the discriminatory power of these circuits is determined primarily by the ratio of resistance to reactance of the devices used, being greater as this is less.

[1] For this and similar circuits see "Radio Instruments and Measurements," Bureau of Standards, Circular No. 74.

[2] For further discussion on coupled circuits see article by Purington, Proc. I.R.E., June, 1930, p. 983.

[3] See "Physical Theory of Electric Wave Filters," G. A. Campbell, B.S.T.J., November, 1922; "Theory and Design of Composite Electric Wave Filters," O. J. Zobel, B.S.T.J., January, 1923; "Theory of Electric Wave Filters Built up of Coupled Circuit Elements," L. F. Peters, Journal A.I.E.E., May, 1923; and "Mutual Inductance in Wave Filters with an Introduction on Filter Design," K. S. Johnson and T. E. Shea, B.S.T.J., January, 1925.

Circuits which are designed primarily to pass currents of certain frequencies and reject others are called filters, and in the succeeding sections we shall discuss a few of the simpler types.

Coil-resistance Filters.—If a coil of 0.1 henry is put in series with a resistance of 50 ohms (some of the 50 ohms will be in the coil itself) and 100 volts of varying frequency is impressed on the circuit, the current which

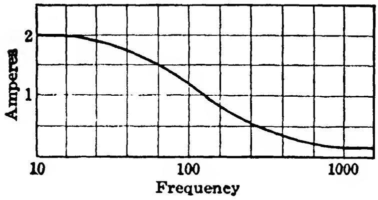


FIG. 148.—Action of simple coil filter.

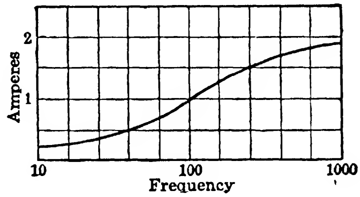


FIG. 149.—Action of simple condenser filter.

flows follows a variation as depicted in Fig. 148. At low frequencies 2 amperes flow, the coil offering practically no reactance; at 100 cycles only 1.2 amperes flow, and at 1000 cycles only a small fraction of 1 ampere can be forced through the circuit by the impressed voltage. This then is a *low-pass filter* of a simple sort.

If a condenser of 60 microfarads is put in series with 50-ohm resistance, and a 100-volt power supply of varying frequency is connected to the circuit, the current which flows at various frequencies is as shown in Fig. 149. At 10 cycles very little current flows, at 100 cycles less than 1 ampere, and at 1000 cycles nearly 2 amperes. For all frequencies above 1000 cycles, 2 amperes flow. This then is a circuit which readily passes high frequencies and eliminates the lower ones; it is called a *high-pass filter*.

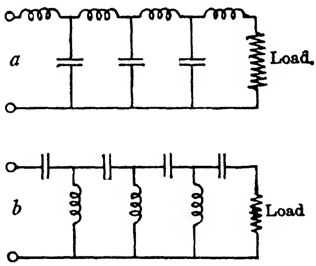


FIG. 150.—Two types of multiple-section filters; *a* is a low-pass filter and *b* is a high-pass filter.

Inductance-capacity Filters.—Neither of the two circuits discussed above shows a discrimination which varies rapidly with frequency; to obtain this characteristic it is necessary to use inductance as well as capacity in a network of some kind; of course, resistance also is necessarily present, although generally objectionable.

In Fig. 150 are shown two types of filter, using combinations of coils and condensers; the arrangement *a* is a low-pass filter, and arrangement *b* is a high-pass one. In each case the combination of one coil and its corresponding condenser is called a *section*; there are three sections in each of the two filters of Fig. 150. In general, the more sections there are the sharper is the discrimination between the desired and undesired frequencies.

There are two important characteristics to be considered in designing a filter for any particular problem, its *surge impedance* (sometimes called *iterative impedance*) and its *cut-off frequency*. By surge impedance is meant the impedance the filter would offer, at its input terminals, if it were very long, that is, if the filter were made up of a great many sections. The cut-off frequency, as the name implies, is that frequency at which the filter suddenly changes its transmitting ability, from high attenuation to low, or *vice versa*. The student especially interested in filter design and performance should consult some such book as Pierce's "Electric Oscillations and Electric Waves," or Johnson's "Transmission Circuits for Telephone Communication."

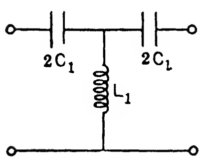


FIG. 151.—One typical section of a high-pass filter.

Having a high-pass filter (*b* of Fig. 150) with inductance per section L_1 and capacity per section C_1, its typical section is of the so-called T form, and is made up as shown in Fig. 151. There are two condensers per section, each of capacity $2C_1$. As these two condensers are connected in series the capacity per section is only equal to C_1, the same as one piece of the actual filter. For this section the cut-off frequency is given by the relation

$$F_0=\frac{1}{4\pi\sqrt{L_1C_1}}, \quad \ldots \ldots \ldots \quad (155)$$

and the surge impedance for a section made up as in Fig. 151 is

$$Z_0=\sqrt{\frac{L_1}{C_1}}. \quad \ldots \ldots \ldots \quad (156)$$

By combining these equations we get

$$C_1=\frac{0.0796}{Z_0F_0}\text{ farad} \quad \ldots \ldots \ldots \quad (157)$$

and

$$L_1=\frac{0.0796\ Z_0}{F_0}\text{ henry.} \quad \ldots \ldots \ldots \quad (158)$$

To make a filter perform properly the resistance of the load circuit must be the same as the surge impedance. Stating this condition differently we say that the filter must be designed with a surge impedance equal to the resistance of the load into which it is to work. When we state that the surge impedance must be equal to the load resistance we really state the condition that there shall be no reflection of energy from the end of the filter. The ratio of voltage and current of a pulse of electric energy traveling along the filter is given by Z_0, the surge impedance. When this pulse meets the load resistance $R(R=Z_0)$ all of its energy is at once absorbed by this load; the ratio of pulse voltage and pulse current is just right to suit the load resistance so none of the pulse energy needs to be reflected.

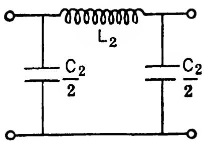


FIG. 152.—One typical section of a low-pass filter.

The low-pass filter (a of Fig. 150) is regarded as made up of sections having the same inductance as each actual coil of the filter, and two condensers, each having half as much capacity as the actual filter condenser. This type of section is called a π section, and is shown in Fig. 152. For such a section it can be shown that the cut-off frequency is given by

$$F_0=\frac{1}{\pi\sqrt{L_2C_2}} \qquad (159)$$

and surge impedance

$$Z_0=\sqrt{\frac{L_2}{C_2}}, \qquad (160)$$

from which we derive

$$C_2=\frac{0.318}{F_0Z_0}\text{ farad} \qquad (161)$$

and

$$L_2=\frac{0.318\,Z_0}{F_0}\text{ henry.} \qquad (162)$$

It will be noticed that at the cut-off frequency the high-pass filter has an inductive reactance $2\pi fL_1$, equal to half the load resistance, and a capacitive reactance equal to twice the load resistance. The reverse is true for the low-pass filter, $2\pi fL_2$ being twice as large as the load resistance, and the capacitive reactance, $1/2\pi fC_2$, being but half the load resistance.

In Fig. 153 are shown the experimentally determined curves for one section of low-pass filter designed to cut off at 1000 cycles and one section of high-pass filter designed to cut off at 2000 cycles. The approximate values of inductance and capacity used in these two sections is shown in Fig. 154. Of course by putting several similar sections in series the cut off takes place more abruptly; that is, the discrimination between desired and undesired frequencies is accentuated.

In Fig. 155 are shown the performance curves of three sections, in series, of the filters shown in Fig. 154. It can be seen that as the number of sections is increased the cut-off frequency becomes more sharply defined, also in the selected and rejected regions, the various frequencies in each region are treated more nearly alike.

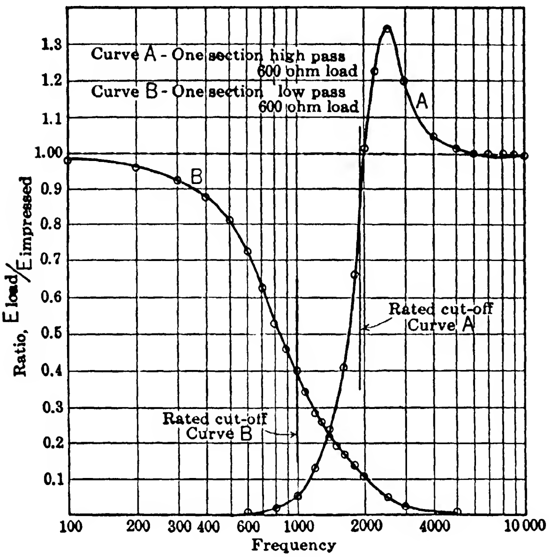


FIG. 153.—Experimentally determined performance of one section each of low-pass, and high-pass, filter.

It is possible to build filters with much sharper cut off than those shown in Fig. 153 by using somewhat more complicated types of section. In Fig. 156 is given a section of L.P. (low-pass) filter which shows a much sharper cut off than the curves of Fig. 153. When a load resistance of 125 ohms is connected to its output terminals it performs as shown in Fig. 157, curve *A*. In another section of the same type the line inductances were 0.336 henry with 5.13 ohms resistance, the shunt path had a coil of 0.191 henry and 3.74 ohms resistance, and the condenser had 12 microfarads of capacity. When supplying power to a 125-ohm load the section performed as in curve *B* of Fig. 157. The cut-off frequency has been advanced from 50 cycles to 80 cycles principally by the change in constants of the shunt path.

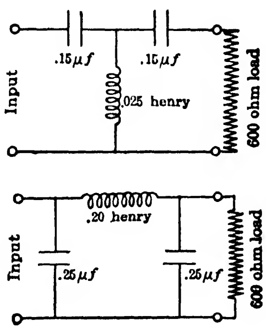


FIG. 154.—Circuit arrangements for getting the results of Fig. 153.

In this same diagram (Fig. 157), curve *C* shows the performance of a high-pass (H.P.) filter section made up as in Fig. 158. The load in this case also was 125 ohms.

Filters for Separating Two Bands of Frequency.—In Fig. 159 is shown a network for separating two ranges of frequencies coming over a telephone line. The L.P. section shows practically no attenuation for all frequencies below 2800 cycles, and the H.P. filter shows practically no attenuation for frequencies above 3200 cycles. The performance of these two filters, when each has a load circuit of 600 ohms, is shown in Fig. 160. The diagram gives in dotted lines the attenuation in decibels, a logarithmic measure of power ratios. The explanation of this scheme of amplification, or attenuation, measurement is taken up in the chapter dealing with amplifiers. Plotted in full lines is the performance

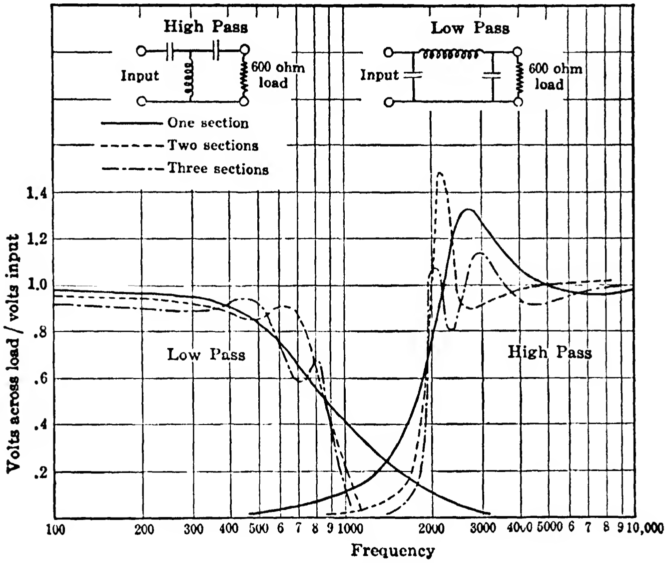


FIG. 155.—Actions of multiple section filters made up of sections similar to those of Fig. 154.

of the filters in terms of voltage at the load divided by voltage at input. These curves show the discrimination possible with commercial types of filters.

Band Pass Filters.—In certain kinds of telephone communication it is desirable to use a filter which will pass equally well all currents in a definite frequency range and reject all others; these are called band pass filters. As one might expect, they are somewhat more complicated in their make-up than the simpler L.P. or H.P. filters. In Fig. 161 is shown one band pass filter, and in Fig. 162 is shown its performance. In this curve also the ratio of output voltage to input voltage is given by the solid line curve, and the decibels attenuation is given by the dotted curve.

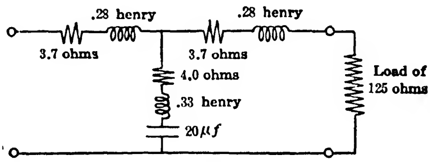


FIG. 156.—A low-pass filter section, with resonant shunt path.

Power Supply Filters.—The power supply for the plate circuit of the modern radio set is practically always obtained by rectifying an a.c. supply, thus getting unidirectional, pulsating current. The pulsations in this current are all multiples of the power supply frequency if a "single-wave rectifier" is used, and are all multiples of twice this frequency if a "two-wave rectifier," or "full-wave" rectifier is used. These pulsations must be ironed out before the power is put into the plate circuits of the radio set. In Fig. 163 is shown a filter for doing this; it is shown as having three sections. It is not generally necessary to use as many as three sections, two usually being plenty.

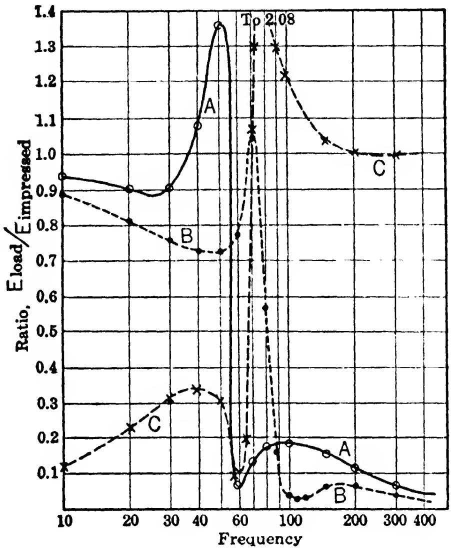


FIG. 157.—Characteristic curves (*A* and *B*), of filters similar to that of Fig. 156 and in curve *C* is shown the performance of the section shown in Fig. 158.

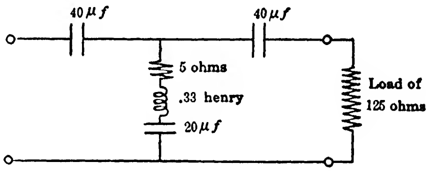


FIG. 158.—A high pass filter section, with resonant shunt path.

Assuming that one volt of various frequencies is impressed on the beginning of the filter, the amplitude of those frequencies which appear across the 5000-ohm load, for a one-, two-, and three-section filter, are as shown in the accompanying table.

Frequency	0	60	120	180	240	300
One section	0.98	0.35	0.075	0.032	0.018	0.011
Two sections	.96	.12	.0056	.0010	.00032	.00012
Three sections (as in Fig. 163)	.94	.048	.00042	.00003	.000006	.000001

It will be appreciated that there must be negligible mutual induction between the three coils if the high attenuation of the foregoing table is to be obtained. A coefficient of magnetic coupling between the first and third coils of the filter, of only a very small fraction of 1 per cent, would materially increase the amplitude of the higher harmonics at the load.

It is well to have each coil and its associated condenser in separate sheet-iron boxes, if best performance of the filter is to be expected. In the average radio set not more than two sections of filter are used, and sometimes only one gives sufficient filtering action.

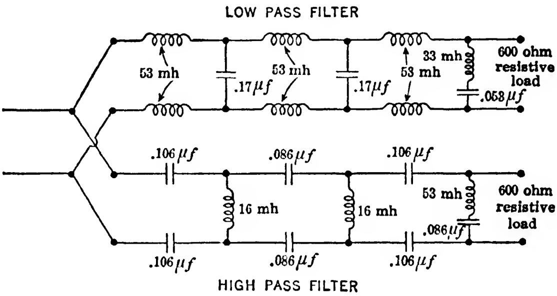


FIG. 159.—A network for separating two ranges of frequencies coming over a telephone line.

With a two-wave rectifier, and 60-cycle supply, there is no 60-cycle ripple to be eliminated by the filter; the 120-cycle frequency is the lowest one present in the rectified alternating current. The three-section filter of Fig. 163 would reduce this ripple, at the load, to only 0.042 per cent of what it has at the power supply end of the filter, whereas the c.c. component (zero frequency), which it is the function of the filter to deliver, has a voltage at the load 94 per cent of its value at the power supply end. In case electrolytic condensers are used, this figure is not quite correct because each of the condensers draws a small continuous current, sufficient to keep itself properly polarized.

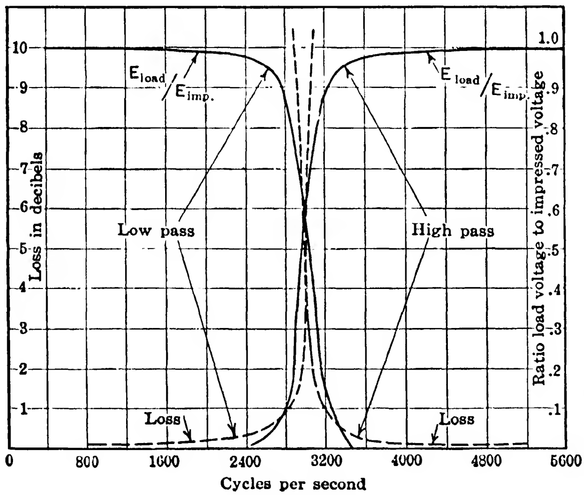


FIG. 160.—Action of the two filters shown in Fig. 159.

Characteristics of Circuits Having Distributed Inductance and Capacity.—Prior to the sections on filters the analyses given apply only to those circuits in which the inductance and capacity are concentrated; another way of specifying the circuits with which we have been dealing is to state that the current, at any given instant, is exactly the same at every point in the circuit. This condition holds for the majority of radio circuits, but there are cases where it evidently does not obtain, thus practically every antenna has zero current at its farther end, whereas the current entering it at the base may be many

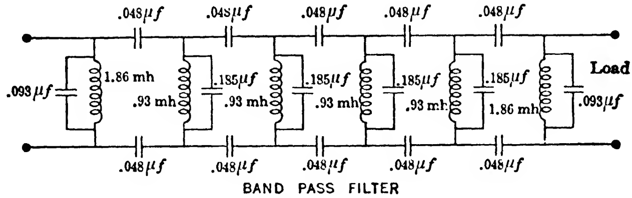


FIG. 161.—A typical band pass filter.

amperes. This is the most striking case of a circuit which has a current varying along its length, but there are others in which the same effect exists to a lesser degree. A coil, for example, may have an internal distributed capacity which appreciably affects its behavior.

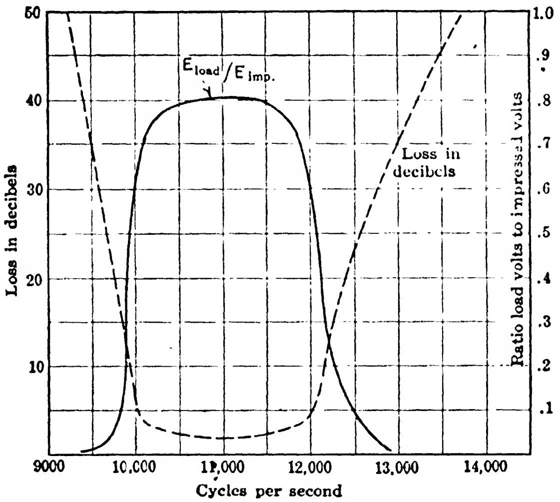


FIG. 162.—Performance of the filter shown in Fig. 161.

The change in current at successive points along an antenna is due entirely to the distributed capacity; each unit length contributes its share to the total capacity and of course requires its proportion of the total charging current. The current flowing in at the base of the antenna must be sufficient to charge the whole length, while that flowing past the middle point of the antenna must be sufficient to charge merely the upper half of the antenna, and so will be considerably less than the current at the base of the antenna.

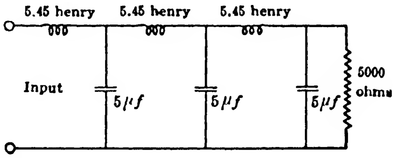
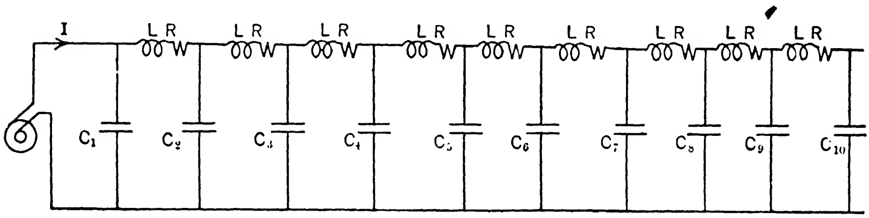


FIG. 163.—Filter for smoothing out the "ripples" of a power supply, obtained by rectifying alternating current. Each coil has a wire resistance of 100 ohms.

Every coil has more or less distributed capacity, every piece of the winding acting as one plate of a condenser for every other part because the various parts are at different potentials and so will have electric fields set up between them when the coil is excited. But, if when a coil is used, it sets up an electric field as well as a magnetic field, it must be considered as a combination of coil and condenser. This internal capacity varies in magnitude appreciably as the frequency, at which the coil is used, is varied and so cannot be treated correctly as a concentrated capacity. As ordinarily used a coil does not show much effect from this internal capacity because the condenser to which the coil is attached has so much more capacity that the internal capacity is completely masked; if, however, the coil is used for tuning a circuit and the tuning condenser used has a small capacity then the internal capacity may produce an appreciable effect on the tuning qualities of the circuit. The calculation of this internal capacity and its effect on the apparent inductance of a coil will be taken up in the next chapter.

Now the resistance and reactance of such a circuit (having distributed capacity and inductance) vary with the frequency through a very large



L = .0415 henry R = .702 ohm C = 36.6 microfarads
$C_1 = C_{10} = 18.3$ microfarads

FIG. 164.—A circuit having distributed inductance and capacity, similar to an antenna.

range of values; the resistance (as measured at the base of the antenna) goes from very small to very large values, while the reactance changes from a large inductive reactance to an equally large capacitive reactance. Moreover, these changes occur periodically as the impressed frequency is continually changed.

To demonstrate experimentally the peculiar characteristics of a circuit having distributed constants, the author built an artificial line having inductance, capacity, and resistance, as shown in Fig. 164; this line resembles somewhat a long antenna, having inductances and capacities, however, several hundred times as large as those of an actual antenna.[1]

A variable frequency was impressed on this artificial line and, by means of a wattmeter, ammeter, and voltmeter its resonance characteristics were determined. The impressed voltage was kept constant at 20 volts and

[1] See "Some Experiments with Long Electrical Conductors," by John H. Morecroft, Proc. I.R.E., Vol. 5, No. 6, Dec., 1917.

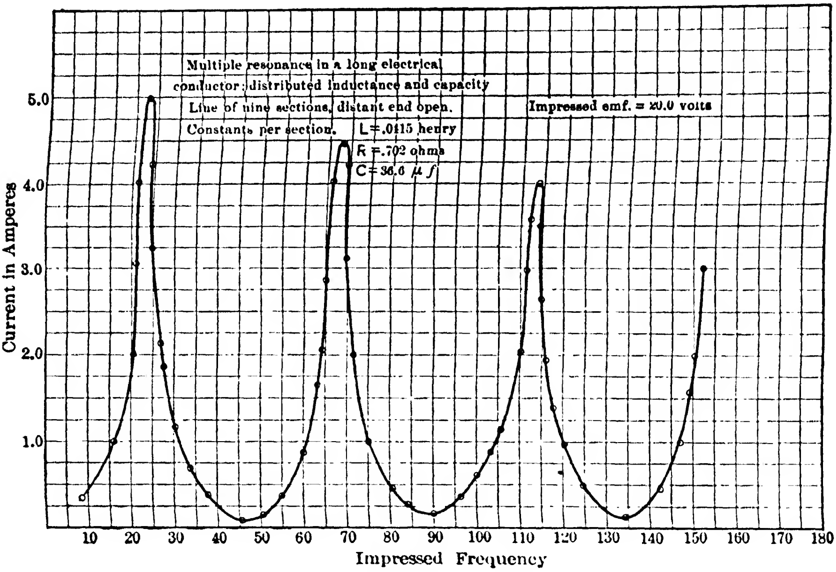


Fig. 165.—Current *vs.* frequency for circuit shown in Fig. 164.

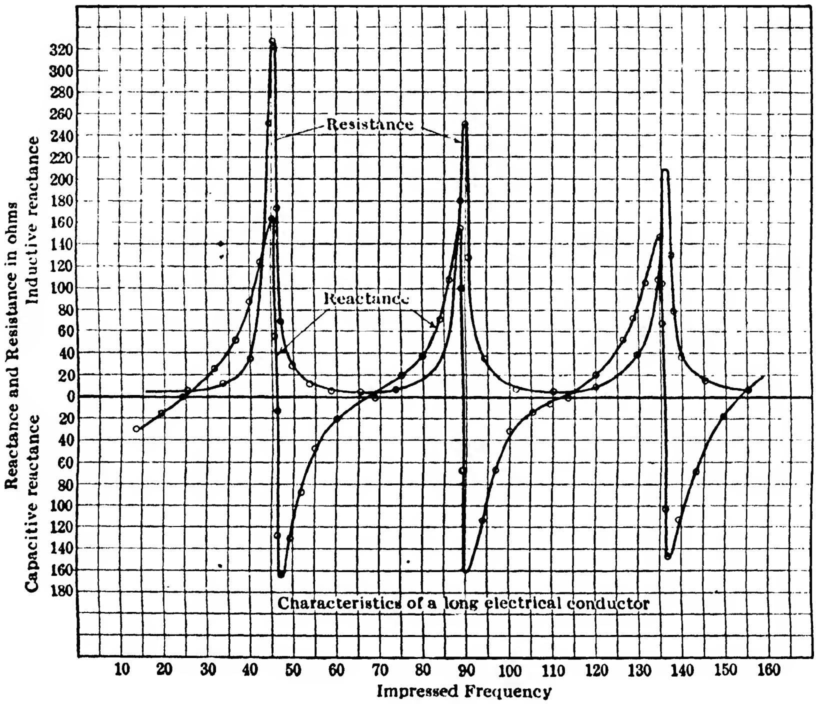


Fig. 166.—Resistance and reactance *vs.* frequency for circuit shown in Fig. 164.

the frequency varied in small steps, from 12 to 152 cycles per second. Fig. 165 shows the current which flowed from the generator into the line at the various frequencies, the line being open at its distant end. The line showed six frequencies in the range used, at which we can say the line was in resonance, meaning by the term resonance a frequency at which the power factor is unity, the line offering resistance reaction only. Three of the frequencies correspond to what we have called parallel resonance, the current being a minimum, and the other three to series resonance, giving large values of current.

The resistance and reactance of the line were calculated from the meter readings and are shown in Fig. 166. The resistance varies periodically from a low value, corresponding to series resonance, to a high value, corresponding to parallel resonance.

The reactance of the line varies periodically from inductive to capacitive reactance and takes all values between 165 ohms positive and 165 ohms negative. From the results shown in Fig. 166 it may be judged how indefinite are the so-called constants of such a circuit.

In a uniform line of this kind the resonance frequencies are multiples of one another. The lowest resonant frequency can be calculated from the formula $f=\frac{1}{2\pi\sqrt{LC}}$, using for L and C values equal to the total values of the line multiplied by the factor $2/\pi$.

As generally used, antennas show a capacitive reactance; they are operated at a frequency lower than their lowest natural frequency. This condition corresponds to frequencies lower than 25, in Fig. 166.

CHAPTER II

RESISTANCE—INDUCTANCE—CAPACITY—SHIELDING

General Concept of Resistance.—The elementary idea of resistance, obtained by a student analyzing c.c. circuits, must be very greatly enlarged and generalized when studying high-frequency circuits. In the c.c circuit, Ohm's law is in general a sufficient definition for the term resistance, that is, $R=E/I$. This definition presupposes that all of the voltage E is used up in overcoming the resistance reaction of the circuit; there must be no reaction such as the c.e.m.f. of a motor, or c.e.m.f. such as exists in a circuit in which storage batteries are being charged, or else the definition is ordinarily changed to the form

$$R=\frac{E_{imp}-E_c}{I},$$

where E_{imp} = the impressed voltage;
E_c = the counter voltage of motor, batteries, etc.

This restated definition must be still more generalized when the ordinary a.c. circuit is considered; in fact, a new concept of resistance must be obtained. It might seem that Joule's law would serve sufficiently to define resistance; this law states that the electrical power liberated as heat in a circuit is given by the equation

$$\text{Heat generated} = I^2Rt.$$

Certainly this law is a more general definition of resistance than Ohm's law because it automatically excludes the effects of counter e.m.f.s, etc.; thus a storage battery, being charged, might (if suitable precautions were taken) be immersed in a calorimeter while being charged and the heat produced be measured by the rise in temperature of the calorimeter water. This amount of heat, properly substituted in Joule's law, will determine the resistance of the circuit in so far as this resistance manifests itself in producing heat.

However, electric energy may be dissipated in forms other than heat; thus radiation of electromagnetic waves from an antenna dissipates energy from the circuit as truly as does the ordinary heating of the circuit.

These elementary considerations force us to adopt a new concept of resistance, it being based on the idea that any transfer of energy from (or to) that part of the circuit, the resistance of which is desired, must be considered in determining the resistance. Thus the resistance between two points in a circuit a–b, Fig. 1, is defined by the equation,

$$R=\frac{\text{power transferred between points } a \text{ and } b}{I^2}. \quad . \quad . \quad . \quad (1)$$

This " power transferred " between a and b may be *leaving* the electrical circuit between these two points or it may be *entering* the circuit between these points. If energy is leaving the circuit between these points, as heat or otherwise, the resistance is *positive*; if energy is entering the circuit between these two points the resistance is *negative*, and if energy is entering the circuit at the same rate as it is leaving then the resistance is zero. From this standpoint any electrical circuit carrying current, after reaching the steady state (no change in the amplitude of the current) has on the whole, zero resistance. Of course we know that the circuit does actually have resistance, in the ordinary meaning of the word, but we may consider the source of power supply as having as much negative resistance as the rest of the circuit has positive resistance. At the generator (or other source of power supply) energy is entering the circuit as fast as it is dissipated in other parts of the circuit. If the circuit, as a whole, has positive resistance the current must be decreasing in amplitude; this state of affairs occurs in the ordinary damped oscillatory discharge of a condenser, whereas a circuit which takes an appreciable time to build up its steady state has, during the time required to reach the steady state, on the whole a negative resistance because, considering the circuit as a whole, energy is entering at a rate faster than that at which energy is leaving.

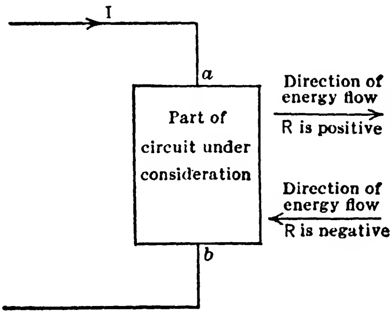


Fig. 1.—If energy is leaving the circuit the resistance is positive, so that if power is entering the circuit its resistance must be considered negative.

Various Factors Affecting the Resistance of a Circuit.—Among the factors contributing to the resistance of a radio circuit are to be considered (1) resistance of the conductor itself; (2) resistance of neighboring closed circuits and their proximity; (3) magnetic material close enough to the circuit to be magnetized by it; (4) losses in the dielectric of any condenser in the circuit; (5) corona losses from parts of the circuit; (6) radiation of electromagnetic energy. All of these factors vary with the frequency of the current in the circuit, some of them with the magnetic gradient set up by the circuit and some with the electric gradient set up by the circuit. Each will be taken up in turn and analyzed as much as seems suitable for a text of this kind.

Conductor Resistance.—The resistance of a conductor, in the form of a wire, to the flow of continuous current is given by the formula

$$R=\frac{\rho l}{a}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

in which ρ = the specific resistance of the material composing the conductor;
l = the length of the conductor;
a = the cross-sectional area of the conductor.

This formula assumes that all parts of the cross-section of the conductor carry the same proportion of the total current; in other words, that the current density is uniform throughout the section of the conductor. This assumption is true for continuous current or for alternating current of very low frequency. If the conductor is large in cross-section or the frequency is high, the inner sections of the conductor carry a relatively small part of the total current, the density of current being greatest at the surface of the conductor; in fact, for very high frequencies a comparatively thin skin on the surface of the conductor carries practically all the current, so much so that if the center part of the conductor were removed, leaving nothing but a thin-walled tube of the same diameter as the original wire, the resistance would be practically the same. This tendency of the current to concentrate on the outer surface of the wire at high frequencies is called the *skin effect*, the reason for the name being obvious. If there are no other conductors carrying current in the vicinity of the one in question this distribution of current will be symmetrical about the axis of the wire, but if there are other current-carrying conductors in the neighborhood the distribution of current through the cross-section of the wire may be irregular, perhaps only the surface part of the conductor on one side carrying an appreciable current.

Any distribution of current other than the regular distribution of equal current density throughout the section of the conductor will result in an increase in the resistance of the conductor; this increase may be so great that the resistance for a high-frequency alternating current may be many times as much as the resistance of the same wire for continuous current.

A simple illustration of this effect is given in Fig. 2, showing three 10-ohm resistances in parallel. Suppose the resistance of this combination is determined by the power loss instead of by the ordinary law for resistances in parallel. Let 3 amperes flow through each resistance, giving a line current of 9 amperes and a power loss of $3 \times (3^2 \times 10) = 270$ watts. Then the total resistance will be obtained by the equation $R = P/I^2$ or $270/9^2 = 3.33$ ohms, the same as we should get by the law for resistances in parallel.

Now suppose that for some reason or other the current redistributes itself so that the two outside paths carry 4 amperes each and the center one carries 1 ampere. (Such a redistribution might well occur if the combination were used in a high-frequency circuit.) The line current will again be 9 amperes and the loss will be $(2 \times (4^2 \times 10)) + (1 \times (1^2 \times 10)) = 330$ watts, which, divided by the square of the line current, gives a resistance of 4.08 ohms, a considerable increase over the value for a uniform distribution of current between the different paths.

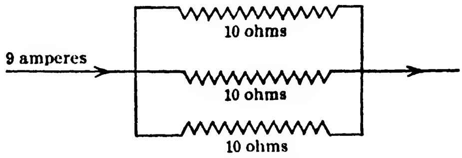


FIG. 2.—A solid conductor of 3.33 ohms resistance may be considered as three separate filaments each of 10 ohms resistance.

The paths shown in Fig. 2 might represent three of the imaginary filaments into which a wire may be supposed divided, and the calculation shows that any distribution of current between the filaments other than uniform distribution results in an increase in the resistance of the conductor; moreover, the greater the non-uniformity of current density the greater will be the corresponding increase in resistance.

Skin Effect in Continuous Current Circuits.—When we say that the conductor of a c.c. circuit has a uniform current density, we have in mind only the condition when the continuous current has reached its steady value. Directly after the switch of a c.c. circuit is closed it cannot be said that the current distribution throughout the conductor is uniform, because it is not. When the current is either rising or falling (see Fig. 33, Chapter I) in a c.c. circuit the outside of the conductor carries a proportionately larger part of the current, and then as the current approaches its steady value the current assumes a uniform distribution.

Even in a c.c. circuit, therefore, there is skin effect to consider when the current changes. If the current of Fig. 33, Chapter I, for example, is flowing through an air-core coil wound with, let us say, No. 4 copper wire, a proper grasp of what is taking place in the circuit can be obtained only if the resistance is considered about as shown in Fig. 3. The more rapidly the current changes the more does the resistance exceed its steady current value, which would be measured as the ratio of volts to amperes in a c.c. resistance measurement. This value of resistance is indicated as R_0 in Fig. 3. At the times T_1 and T_2 the resistance of the conductor might be many times as great as it is for the current in the steady state, the reason being that during the rapid current changes the inside part of the conductor is not as effective in carrying current as is the outer part.

A circuit in which a somewhat analogous situation exists is shown in Fig. 4. The circuit here shown has inductance in the center path and none in the two outside paths. Directly after the switch is closed the two outer paths each carry 2 amperes of current and the center path carries none. After the switch is closed, the currents have the form shown in the (*b*) diagram of Fig. 4; the currents in the outer branches maintain their values of 2 amperes and the center branch has a current rising in a logarithmic curve to a final value of 2 amperes.

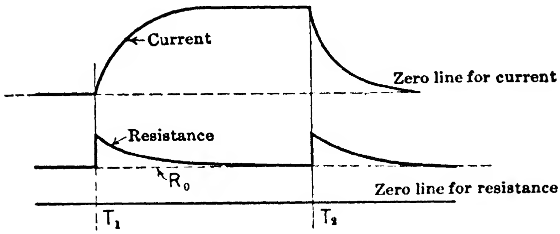


FIG. 3.—Even a continuous-current circuit shows skin effect for current changes.

Directly after the switch is closed, therefore, 4 amperes flow from the 6-volt source, and the apparent resistance of the circuit is 1.5 ohms. After the center branch has attained its final current, the 6-volt battery is delivering 6 amperes to the circuit and the resistance of the circuit

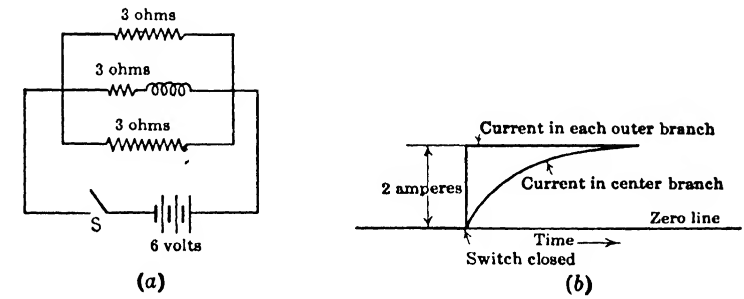


FIG. 4.—In such a divided circuit the currents rise as shown in the (*b*) diagram.

is only 1 ohm. The change from 1.5 ohms to 1 ohm follows a curve similar to that given in Fig. 3.

Skin Effect in Straight Wires.—The non-uniformity of current distribution referred to above occurs in every conductor carrying alternating current, the current density being greater at the surface than at the center of the wire, but this non-uniformity is not appreciable unless the wire is large in diameter, or the frequency is high; the increase in resistance due to skin effect depends upon the product of the cross-section and the frequency and for copper wires the general idea given by the following table is useful.

Frequency multiplied by the Cross-section in circular mils	Ratio of a.c. to c.c. resistance
10,000,000	1.003
20,000,000	1.012
100,000,000	1.30

As an example No. 10 wire has a cross-section of 10,000 circular mils; at a frequency of 2000 cycles its a.c. resistance is 1.2 per cent greater than its c.c. resistance, while at 10,000 cycles its resistance would have increased over its c.c. value by 30 per cent.

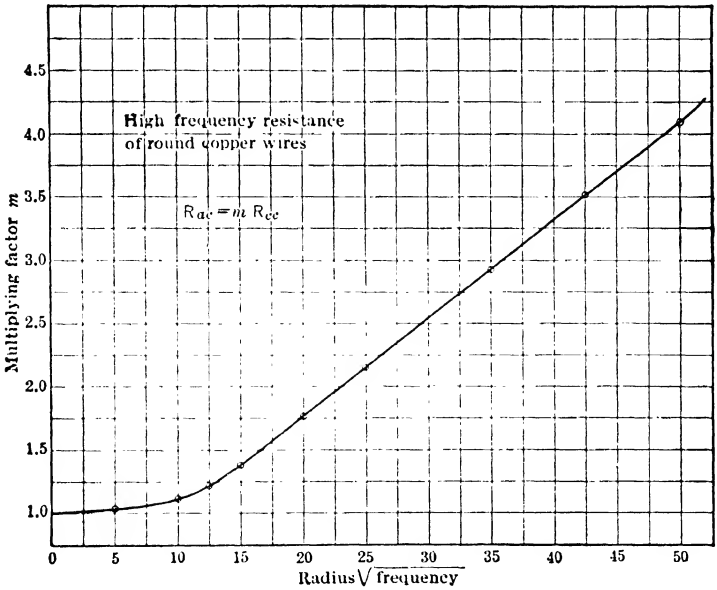


FIG. 5.—Variation of resistance of round, straight, copper wire with frequency and radius, this being measured in centimeters.

An exact analysis shows that the ratio of a.c. resistance to c.c. resistance may be expressed in terms of diameter, permeability, frequency, and resistivity; a correct expression involves an infinite series of terms, but these series have been summed so that accurate data are available for calculating the resistance of any round wire, the permeability and resistivity of which are known. For copper wire, in which the permeability is unity, tables have been compiled which present the data in convenient form In the curves of Figs. 5 and 6 is shown the factor, m, by which the c.c. resistance must be multiplied to give the resistance for alternating current. Plotted as abscissas are values of $r\sqrt{f}$, where r is the radius of the wire in cm. and f is the frequency of the current being used.

It is sometimes useful to know how large a wire can be used without having its a.c. resistance exceed its c.c. resistance by more than a specified amount. The data given in the accompanying table, compiled by L. W. Austin, may be useful for this purpose:

TABLE I

WIRE DIAMETERS

Largest wire (straight) which can be used without the high-frequency resistance exceeding the c.c. resistance by more than 1 per cent

Wave length in meters	Diameters Given in Millimeters			
	Advance	Manganin	Platinum	Copper
100	0.30	0.29	0.13	0.006
200	0.46	0.40	0.20	0.045
300	0.57	0.50	0.27	0.09
400	0.66	0.60	0.30	0.10
600	0.83	0.75	0.37	0.15
800	0.98	0.88	0.42	0.20
1000	1 10	0.99	0 50	0.21
1200	1.20	1.10	0.57	0.22
1500	1.30	1.21	0.63	0.26
2000	1.52	1.38	0.73	0.30
3000	1.82	1.62	0.80	0.33

Frequency $= 3 \times 10^8 \div$ wave length

As the current travels almost exclusively on the outer parts of a wire at high frequency, it is important that this outer part have as low a specific resistance as feasible. Thus it may well be that tinned copper wire will show a higher resistance than a wire which has no tinning. At a frequency of 10^7 cycles per second one experiment showed a resistance of tinned copper wire 30 per cent greater than for a corresponding untinned wire; nickel plating a copper wire increased its resistance, at 10^7 cycles, as much as four times.

Current Penetration in Conductors.—In the case of a wide, flat conductor, such as the earth's surface, the currents which are set up in the surface penetrate into the substance of the conductor according to the specific resistance of the material, permeability, and frequency. The relation between the density of current at the surface and the density at a point distant x below the surface is given by:

$$i = I_0 \epsilon^{-\left(\sqrt{\frac{2\pi\omega\mu}{\rho}}\right)x} \times \sin\left(\omega t - \left(\sqrt{\frac{2\pi\omega\mu}{\rho}}\right)x\right), \quad . \quad . \quad . \quad . \quad (3)$$

in which I_0 = current density at surface;

i = current density a distance x centimeters below the surface;

$\omega = 2\pi f$, where f is the frequency of current;

μ = permeability of the substance;

ρ = specific resistance of the substance, abohms per centimeter.[3]

Not only does the density of current decrease as the distance below the surface is increased but, as indicated by Eq. (3), it reaches its corresponding values at later time than at the surface, this amount of time lag

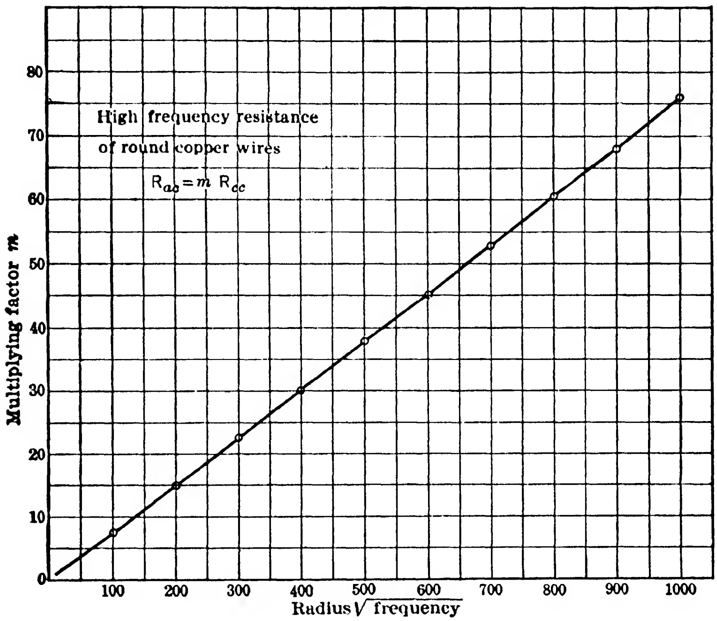


FIG. 6.—Variation of resistance of round, straight, copper wire with frequency and radius (radius in cms.).

increasing as the depth below the surface is increased. This really means that the current penetrates into the substance with a wave motion; the attenuation is, however, very high, so that probably only a fraction of a wave length is actually set up in the conductor with an appreciable amplitude.

In deriving Eq. (3) the displacement current which may occur in the material has been neglected. This is of course permissible in ordinary iron because of the very large value of the conduction current in comparison to the displacement current.

It may be, however, in very special cases that this assumption is not warranted by the facts, in which case Eq. (3) is to that extent incorrect.

Thus in iron-dust cores the specific resistance is increased enormously over that of ordinary sheet iron, and it may be that at very high frequencies the displacement current is not small enough, compared to the conduction current, to be neglected.

A Simple Analysis of Skin Effect.—Although an exact analysis of skin effect in a conductor requires the theory of wave propagation, and special mathematical series for a solution, a very good idea of its cause (and, what is much more important, its remedy) may be had from the ordinary laws of current flow in inductive circuits. The first thing to notice about the problem is the effect of frequency upon the division of current between two paths in parallel, as shown in Fig. 7, the two paths having equal resistance but unequal inductance. The formula for the current flow in each path is

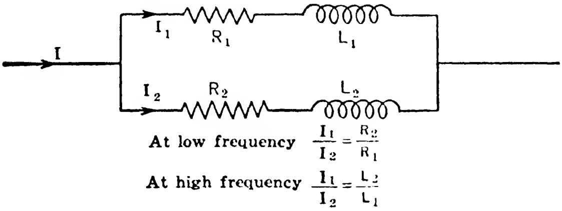


FIG. 7.—For branched circuits the resistance controls the division of current at low frequency whereas the reactance controls the division at high frequency.

$$I=\frac{E}{\sqrt{R^2+(\omega L)^2}}.$$

At very low frequency the ωL term is negligible, and so we have the currents dividing between the two paths inversely as the two resistances, that is, the two currents will be alike. At very high frequency, however, the resistance term becomes relatively negligible and the current divides inversely proportional to the inductance in the two branches. The same voltage is applied to both paths, so the sum of the resistance reaction and the inductance reaction (added vectorially) in each path must be the same. It is from this standpoint that we will investigate the skin effect in wires.

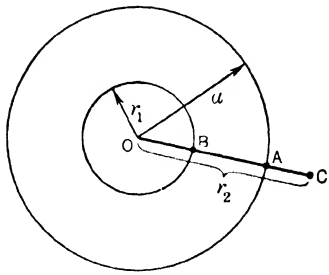


FIG. 8.—Cross-section of wire of radius a; magnetic field density to be calculated at points A, B, and C.

Imagine a round copper wire carrying current uniformly throughout its cross-section, Fig. 8. The density of magnetic flux at a point A on the surface of the wire is given by the formula

$$B_A=\frac{0.4\pi I}{2\pi a}=\frac{0.2I}{a}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where I = total current in amperes;
a = radius of the conductor in centimeters.

At a point B inside the conductor, the amount of current producing flux (for of course there *is* magnetic flux inside the conductor) is only that part of the total current which flows inside the circle inscribed through the point B. This amount of current is, for uniform current density, equal to $I \times r_1^2/a^2$. The magnetic flux density at B is, therefore, by Eq. (4)

$$B_B = \frac{0.2Ir_1^2}{a^2} \times \frac{1}{r_1} = \frac{0.2Ir_1}{a^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

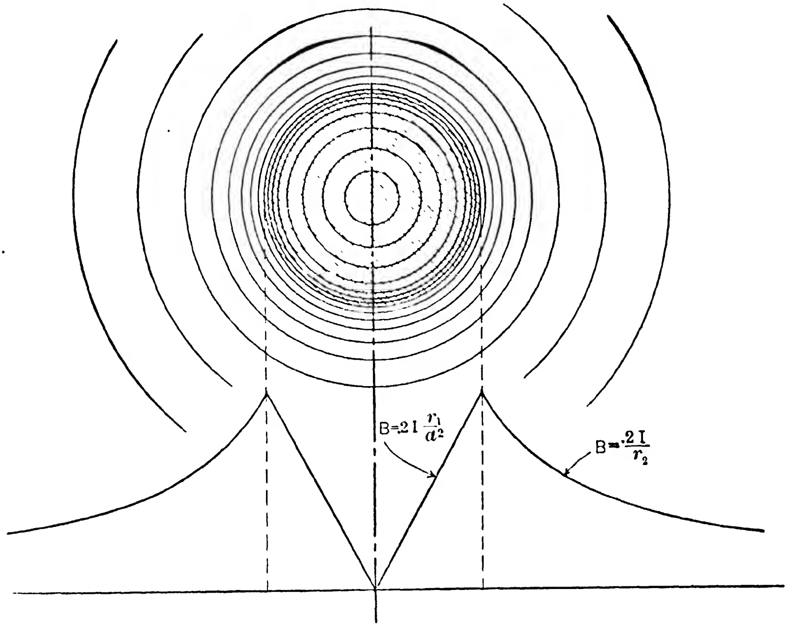


FIG. 9.—Closeness of circular lines show the density of magnetic field around a non-magnetic wire; in lower part of figure magnetic field density is shown by distance from reference line to curve marked B.

For a point C outside the wire, distant r_2 from the axis of the wire, the flux density is given by the equation

$$B_C = \frac{0.2I}{r_2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

From Eqs. (4), (5), and (6), the flux density throughout the cross-section of the wire and in the region surrounding the wire may be calculated; Fig. 9 shows the result of such a calculation. In the upper part of the figure the flux is shown in the form of circles concentric with the axis of the wire, the closeness of the circles representing the flux density, and in the lower part of the figure is shown a plot of the flux densities, ordinates being values of flux density and abscissas being distance from center of wire.

The total flux surrounding any point is obtained by adding the flux from a point infinitely distance from the wire, up to the point in question; a curve showing the value of this flux for different points inside and outside the wire is shown in Fig. 10. The ordinates of the curve are obtained by integrating the density curves of Fig. 9. The flux ϕ_1 is the total flux produced by the current in the wire, outside of the wire itself, whereas surrounding any point outside the wire as, e.g., C of Fig. 8, there is a flux equal to ϕ_3. There is a certain amount of flux inside the wire itself and the flux surrounding the innermost filament is obtained by adding to ϕ_1 this internal flux; it is shown by ϕ_2 in Fig. 10.

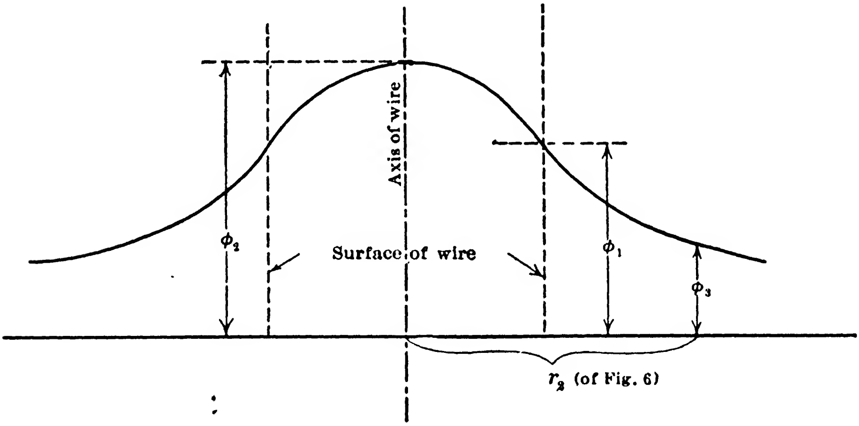


FIG. 10.—Curve showing *total flux* outside any point considered in Fig. 9.

Now let us consider the wire made up of a bundle of separate parallel filaments, such a wire as would be obtained by using a cylindrical bundle of very fine wires, each insulated from its neighbor, except at the ends of the wire in question, where they are all electrically connected together. Let the resistance of each of these filaments be R. If an e.m.f., $E \sin \omega t$, is impressed across the ends of this composite wire, all filaments will have the same impressed e.m.f., and it is therefore evident that the sum of the reactions in each filament must add up (vectorially) to equal this impressed force.

The resistance drop in each filament is IR and the inductance drop is $\frac{d\phi}{dt} = \omega\phi$, where ϕ is the maximum amount of flux surrounding the filament in question. Hence for two filaments, one at O and the other at A (Fig. 8), we must have

$$E^2=(I_1R)^2+(\omega\phi_1)^2$$

and

$$E^2=(I_2R)^2+(\omega\phi_2)^2,$$

where I_1 and I_2 are the currents in the two filaments considered.

In these equations E, I, and ϕ must have corresponding values, i.e., either effective values or maximum values.

At very high frequencies the resistance drop is negligible compared to the flux reaction drop, and hence $E=\omega\phi_1=\omega\phi_2$. We must therefore

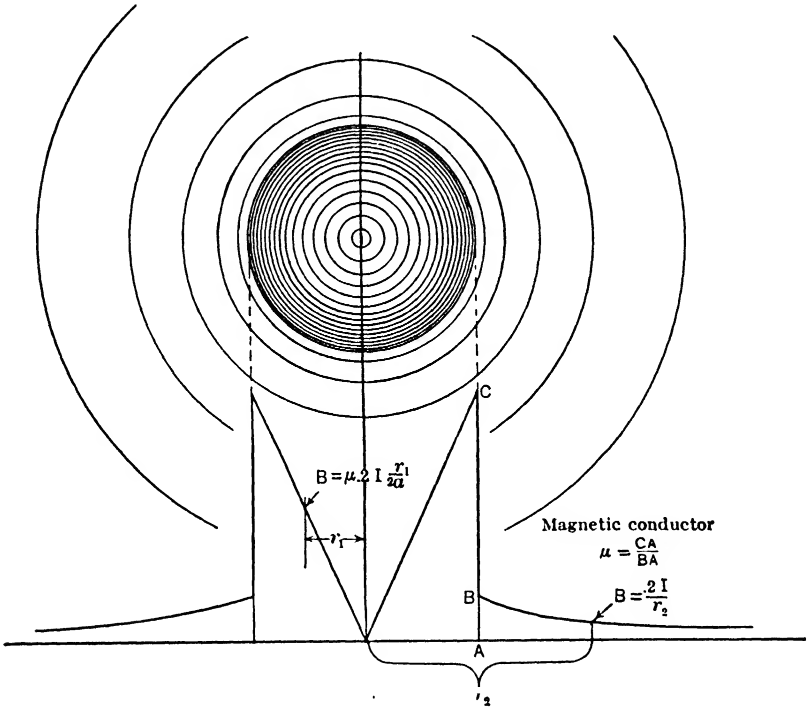


Fig. 11.—Flux density of the magnetic field in and around a wire of magnetic material.

conclude that, for frequencies which make IR negligible compared to $\omega\phi$, *the flux surrounding all filaments of the wire must be the same.* Referring to Fig. 10 this statement means that $\phi_2-\phi_1=0$, that is, there is no flux inside the wire itself. If there is no internal flux the flux density everywhere inside the wire must be zero, and as the flux density at any point in the wire distant r from the axis is equal to $0.2I/r$, where I now signifies *the current flowing in the wire inside of a circle through the point in question;* we must conclude that there is no current anywhere inside the wire.

At ordinary frequencies the resistance drop is not negligible in comparison with the reactance drop, so that the sweeping conclusion of the previous paragraph (no current anywhere inside the conductor) is not true, but it is evident that, as the frequency increases more and more the difference between ϕ_2 and ϕ_1 of Fig. 10 must continually decrease.

If instead of a copper wire an iron wire had been assumed, the internal flux density would have been very much increased so that Figs. 9 and 10 would have more nearly the appearance of Figs. 11 and 12. The value of the internal flux $(\phi_2 - \phi_1)$ would be very much increased, so that the frequency at which the IR drop becomes negligible compared to the $\omega\phi$ drop is much less than for copper wire.

Offsetting this effect to some extent, however, is the fact that the specific resistance of the iron is greater than that of copper; the result is that iron, while it has a greater skin effect than copper, does not have as much greater effect as the increased value of μ would indicate.

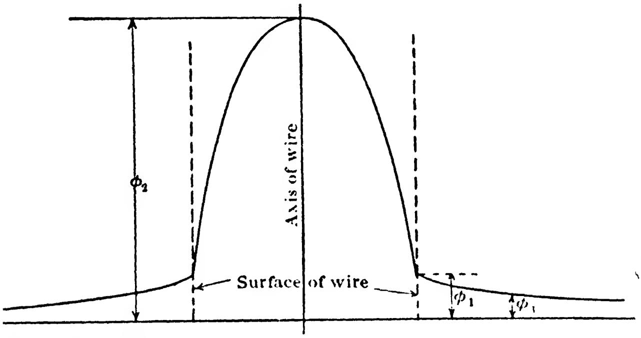


FIG. 12.—Total flux outside any point considered in Fig. 11.

The change in current density throughout the cross-section of wire due to the effect of the internal flux, is indicated (for a certain wire) in Fig. 13, the three curves showing how, as the frequency is increased, the current shifts more and more to the outer skin of the conductor. The current density at the surface of the conductors has been assumed the same for the three frequencies.

It is evident from the foregoing discussion that a substance having high specific resistance and low permeability will have the least skin effect; this is shown in Table I on p. 168. The wires used for resistance in making tests and measurements in high-frequency circuits should be made of small wires of the high-resistance alloys, practically all of which have unity permeability.

Elimination of Skin Effect.—One obvious remedy for skin effect is to so construct the conductor that there is no internal flux, or rather that the internal flux is negligible compared to the external flux, which of course produces no skin effect, as it affects all filaments of the wire equally. A conductor with no internal flux is impossible, but such a condition may be approximated by using a tubular conductor; such a construction is

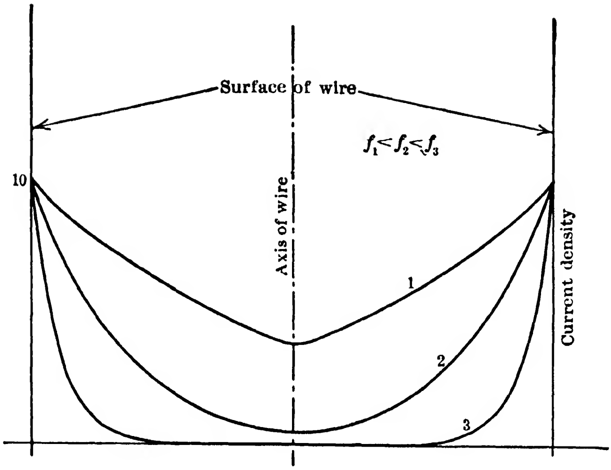


Fig. 13.—Current density in a solid round conductor at three different frequencies.

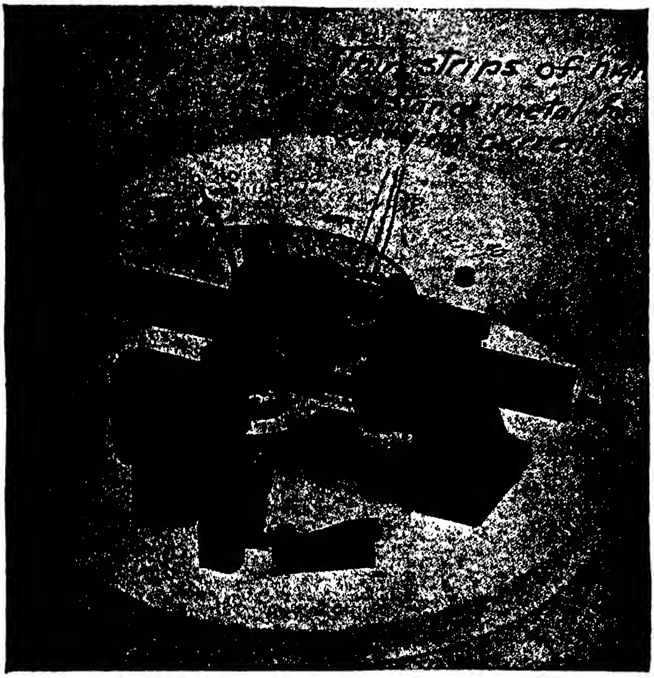


Fig. 14.—A hot-wire ammeter showing how the skin effect is minimized by special arrangement of very thin strips of high-resistance metal.

used for high-frequency ammeters designed to carry comparatively large currents (say 25 amperes or more). The current to be measured is carried from the connectors of the ammeter to two circular discs, and these discs are connected by a set of very thin high-resistance strips, the whole arrangement having the appearance of a barrel, the thin strips taking the place of the barrel staves. Such a construction is shown in Fig. 14, this showing the construction of a 40-ampere, so-called "hot-wire" meter. As the radial thickness of these strips is only about 0.004 cm., there is practically no internal cross-section to the conductor; it is all "skin."

In the scheme ordinarily employed for reducing skin effect the required cross-section of conductor (which depends upon the amount of current to be carried), is made up of a great many small wires, each completely insulated from all the rest; a common form of this cable used for winding radio coils consists of 48 No. 38 enameled wires properly woven together. In eliminating skin effect by this construction it is not sufficient to merely subdivide the conductor into small well-insulated strands; these strands must be so woven or twisted together that *each strand is as much on the outer surface of the cable as every other one.*

If 48 No. 38 insulated wires are laid loosely together, parallel to one another, it will be found that the increase in resistance due to skin effect is nearly as much as though a solid wire (of the same cross-section as that of the cable) were used. (The stranded cable would be somewhat better than the solid wire, because of its somewhat greater useful outer surface.) It is therefore important in getting the stranded wire (sometimes called *litzendraht*) to see that not only is it made up of a great number of well-insulated strands, but also that these strands are properly interwoven. A real braiding process will accomplish the result, but a suitably twisted cable is nearly as good. A properly twisted cable must be made up of several component twisted cables to be free from marked skin effect. For 48 No. 38 cable, e.g., three separately twisted cables, each of 16 wires, may be twisted together and the resulting cable will be nearly as good as braided cable.

It is important to note just what effect is to be obtained in making these high-frequency radio cables; the cable must be so constructed that each strand has, per unit length (say, per meter) the same flux surrounding it, when each strand is carrying the same current. When the strand is in the center of the cable it has more flux surrounding it than when it is on the periphery, hence each strand must occupy corresponding positions in the cross-section of the cable for equal distances, in order that it may have the same surrounding flux per unit length as all the other strands.

Even in suitably woven cable there is still some skin effect due to the finite size of the strands themselves, each strand in itself having an appreciable skin effect at very high frequencies.

It is important in purchasing radio cable of the kind just described to make tests for the c.c. resistance. In making this test the enamel must be removed carefully from each strand, at both ends of the piece to be tested, and the strands be well soldered together. The c.c. resistance should be calculated from the total cross-section of copper in the cable and the measured value should approach this very closely. In making cable a strand may break and the operator insert another and continue the process of weaving the cable. But such a broken strand is evidently of no use in carrying current, because one break opens that strand completely, the strand being insulated from its neighbors.

Specifications for radio cable should therefore state not only the size and number of wires to be used, quality of enamel, method of twisting, etc., but should also call for a c.c. resistance within a certain percentage of the theoretical value. The longer the pieces of cable called for, the more likely are breaks to occur; for this reason the cable is generally obtained in lengths of a few hundred feet only. If two pieces of radio cable are to be joined, the greatest of care must be exercised in making the joint; if only half the strands are soldered together (quite likely unless each individual wire is separated from the rest and properly cleaned before attempting to solder the joint) then the resistance of the whole cable is much more than it should be.

Too Much Conductor May Be Used.—It has been explained how the current leaves the center of a conductor with increasing frequency, thus making the metal on the inside of the wire practically useless in so far as conductivity is concerned. The simple explanation given neglects the possibility of wave propagation of the current into the conductor, as the relative phases of the currents in the several hypothetical filaments were not considered.

If, however, we considered the matter more exactly, we should find that the current actually does penetrate the conductor in the wave motion indicated by Eq. (3). On the surface of the wire the current density has its greatest magnitude; with increasing distance from the surface of the conductor the magnitude of current density decreases and its phase is continually retarded. After a certain depth of penetration the current may actually be flowing in a direction *opposite to that it has on the surface.* In such a case the resistance of the conductor would actually be decreased by removing the inside part of the conductor, that is, a copper tube would show less resistance than a solid copper rod of the same diameter.

This effect has been noticed even at engineering frequencies for bus bars of the extreme sizes being used in the power plants today. By taking out some of the copper strips forming the central portion of the bus bar, its resistance has actually been decreased.

This effect is indicated in Fig. 15, that current flowing in the central part of the copper bar is shown in opposite phase to the main current flowing in the copper near the surface. This internal reversed current of course contributes heat to the bus bar and actually decreases the net current flowing in the circuit. It is apparent then that if the central part of the conductor of Fig. 15 were removed, the resistance of the copper bar would be decreased.

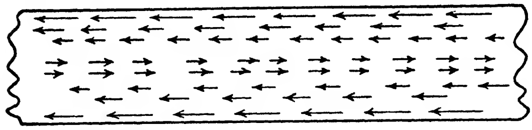


FIG. 15.—In a large copper bus bar the current in the center may be flowing in phase opposite to the current in the outer part of the bar.

Skin Effect in Coils.—With the foregoing analysis of skin effect in mind it is at once evident that the redistribution of current throughout the cross-section of a conductor will be greater if the conductor is used in making a coil than if it is used in the form of a straight wire. The distribution of magnetic flux inside a single layer solenoid is somewhat as shown on Fig. 16; the flux density is high just inside the solenoid and practically zero at the outer surface of the coil. Assuming that this density decreases to zero from the inner surface of the winding to the outer (nearly

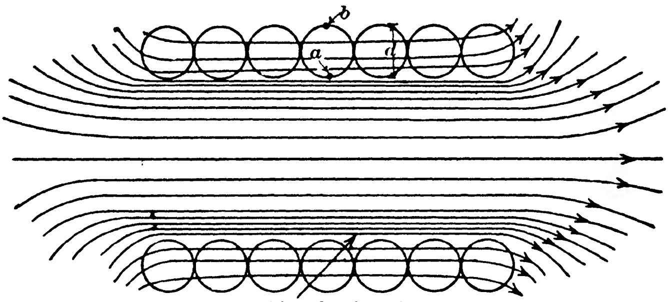


FIG. 16.—Approximate flux distribution inside a short solenoid.

the case for ordinary coils) it is evident that the outer filaments of the wire are linked with much more flux than are the inner filaments. Thus an imaginary filament on the outside of the wire as at *b*, Fig. 16, will be linked with a flux in excess of that linked with filament *a* by an amount equal to $(B/2)\times d$ (where B is the flux density at the inner surface of the winding and d is the radial depth of the winding) per unit length of the wire. It is apparent that the current will tend to crowd into that part of the wire which is on the *inside of the coil*, the inductance reaction being less for the filaments on the inner side of the winding than for those on the outer side.

But besides this tendency of the current to redistribute itself, there is also the tendency to redistribution about the axis of the wire, and also each conductor exerts a certain effect on its neighbor—these all combine to produce a current distribution about as indicated in Fig. 17, the density of current being indicated by the proximity of the dots.

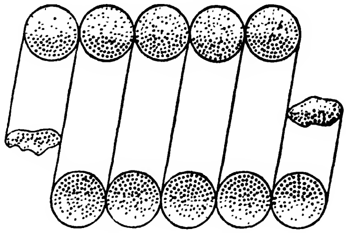


FIG. 17.—Distribution of current in the conductors of a short solenoid; density of shading corresponds to current density.

In constructing variable resistances for use in making radio measurements, skin effect must be carefully considered. The most convenient form of variable rheostat is a cylindrical one with a sliding contact, the almost universal form of laboratory rheostat for ordinary c.c. and a.c. measurements. But this type of winding is not satisfactory for high-frequency currents because of the extra skin effect caused by solenoidal winding and also because the amount of self-induction in such a rheostat is too great to be neglected in radio circuits. Radio cable cannot be used with sliding contact rheostats for evident reasons; solid wire must therefore be used and still the skin effect and self-induction be reduced to a minimum. This is done by winding on a porcelain tube a bifilar high-resistance solid wire; the two wires making the bifilar construction are wound around the cylinder in opposite directions, the two wires crossing each other twice per turn. Such a winding has a self-induction practically zero, and hence has a minimum skin effect.

The increase in resistance of coils, due to skin effect, is a very difficult problem to analyze mathematically; only the simplest cases have been considered, and even then assumptions have been made which make the validity of the equations obtained doubtful.

An experimental investigation of the skin effect in coils was carried out by the author, measurements being made on a Wheatstone bridge, and the results are given herewith; they serve to indicate how much increase in resistance from skin effect may be expected with coils similar in form. The single-layer coils were wound on dry wood reels with double cotton-covered wire, the wires being laid as closely together as possible. The length of the winding was 10 cm. and the approximate diameter (the cross-section was actually octagonal) was 10.5 cm. The data are given in Table II, p. 180, both self-induction and resistance being given, the results being probably accurate to within 1 per cent unless otherwise stated.

There are two effects which must be kept in mind when interpreting these results; there is an actual increase in resistance due to redistribution of current in the conductor of which the coil is made, and there is an

TABLE II

Coil		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Size wire		11	12	14	16	18	20	22	24	26	28	30	32	34	42/36s	48/38s
Number turns		37	41	54	64	83	96	115	138	188	237	290	345	361	85	97
Frequency in Kilocycles																
0.0	R	0.050	0.07	0.15	0.30	0.61	1.07	2.10	3.89	7.05	17.2	33.3	59.8	101.0	1.14	1.58
0.870	R	0.066	0.087	0.167	0.305	0.628	1.11	2.12	3.91	7.22	17.3	33.3	59.8	101.0	1.20	1.60
	L	0.0975	0.120	0.202	0.278	0.460	0.614	0.894	1.29	1.78	3.79	5.54	7.79	8.49	0.496	0.639
	R L	0.68	0.725	0.834	1.09	1.37	1.81	2.38	3.04	4.06	4.57	6.00	7.68	11.9	2.43	2.50
50.0	R	0.400	0.495	0.70	0.96	1.42	1.86	2.82	5.00	8.30	20.0	35.9	63.1	105.0	1.42	1.82
	L	0.0955	0.118	0.198	0.278	0.457	0.614	0.897	1.30	1.78	3.82	5.56	7.88	8.57	0.495	0.641
	R L	4.20	4.25	3.53	3.45	3.10	3.03	3.15	3.85	4.65	5.23	6.46	8.02	12.3	2.87	2.83
100.0	R	0.580	0.668	1.02	1.43	2.18	2.80	4.14	6.05	9.75	22.8	42.0	77.3	124.0	1.62	2.18
	L	0.0942	0.118	0.199	0.279	0.460	0.617	0.897	1.30	1.79	3.86	5.66	8.15	8.90	0.498	0.644
	R L	6.15	5.65	5.18	5.14	4.75	4.55	4.63	4.65	5.44	5.90	7.42	9.50	13.9	3.24	3.40
150.0	R	0.735	0.84	1.29	1.76	2.86	3.75	5.26	7.74	11.8	31.4	61.0	112.0	164.0	2.07	2.39
	L	0.0947	0.118	0.199	0.278	0.460	0.616	0.898	1.30	1.80	4.00	6.00	8.51	9.26	0.500	0.647
	R L	7.77	7.22	6.48	6.33	6.22	6.10	5.85	5.96	6.55	7.85	10.2	13.1	17.7	4.13	3.70
200.0	R	0.803	1.00	1.60	2.17	3.52	4.45	6.75	9.70	16.1	45.0	85.0	157.0	220.0	2.26	2.71
	L	0.0952	0.119	0.198	0.278	0.461	0.621	0.899	1.31	1.84	4.22	6.21	8.70	9.5	0.501	0.649
	R L	8.45	8.40	8.10	7.80	7.61	7.15	7.52	7.40	8.75	10.7	13.7	18.0	23.2	4.50	4.18

Size of wire to the nearest B. & S. gage number.
Inductance given in millihenries.
Resistance given in ohms and ohms per millihenry.

increase in the measured value of resistance due to the effect of internal capacity, explained in the previous chapter when analyzing resonance in parallel circuits.

Every coil has internal capacity due to one part of the winding being equivalent to one plate of the condenser, acting with every other part to form a condenser. It was shown that the apparent resistance of an inductance, shunted with a condenser, increases as the frequency is increased, in accordance with Eq. 56, p. 92. Although this equation is not directly applicable to these coils (the capacity of which changes with frequency changes) it indicates that the measured value of resistance may be expected to increase entirely aside from any skin effect which may be present. But this effect of capacity which gives an apparent increase in resistance produces at the same time an increase in the *apparent inductance of the coil*, so that in the results of the table any increase in resistance which occurs without a corresponding increase in inductance is due to redistribution of current in the conductor of the coil (i.e., real skin effect); for frequencies high enough to produce an increase in the apparent inductance the skin effect is not alone in producing the increase in resistance, the internal capacity contributing its effect also in increasing the apparent resistance.

It will be noticed that for the larger wires the inductance actually decreases as the frequency increases, for the lower values of frequency.

Illustrating this effect we take the data for coil No. 1 made of No. 11 wire. In the range of frequencies used the inductance decreased with increase of frequency, whereas the resistance increased from 0.050 ohm to 0.803 ohm. The radius of this wire is 0.114 cm. and so the factor, $r\sqrt{f}$ for 2×10^5 cycles, is 51. Referring to Fig. 5 the factor m is found to be 4.2. If, therefore, the wire had been used in the form of a straight conductor, we might have expected an increase in resistance from 0.050 ohm, the c.c. resistance, to $4.2\times0.05=0.21$ ohm. Actually it changes from 0.05 ohm to 0.80 ohm, thus showing how the skin effect is augmented when the wire is used in the form of a coil. The superiority of the radio cable, either 42/36's or 48/38's, is at once evident from the results given in the table.

If the coil used has more than one layer, the magnetic field density is much greater than it is for a single-layer coil, hence we should expect a much greater skin effect for multi-layer coils than for single-layer coils and the experimental results of Table III which were obtained with 10-layer coils prove the point. Thus the single-layer coil of No. 18 wire showed an increase in resistance of $\frac{2.18}{0.61}=3.6$ times as the frequency varied from zero to 100,000 cycles. This same wire wound in a 10-layer coil showed an increase through the same range of frequency of $\frac{84}{1.74}=48$ times so that the

TABLE III

Coil		16	17	18	19	20	21	22	23	24	25	26
Size wire		12	18	20	22	24	26	28	30	32	34	48/38s
Number turns		100	197	239	300	343	410	586	719	859	898	250
Frequency in Kilocycles												
0	R	0.12	1.74	3.30	6.7	11.5	21.3	49.0	96.0	172.	300.	5.1
1.2	R	0.48	2.48	3.40	6.8	11.7	21.6	50.0	96.5	173.	300.	5.2
	L	1.46	5.21	8.10	12.6	16.7	23.7	48.0	72.0	102.	117.	9.2
	R/L	0.33	0.48	0.42	0.54	0.70	0.91	1.0	1.3	1.7	2.6	0.56
10.5	R	3.85	6.27	7.80	11.4	16.2	27.0	64.0	122.	210.	345.	6.7
	L	1.42	5.18	8.10	12.6	16.8	24.0	48.0	73.0	104.0	120.	9.3
	R/L	2.7	1.2	0.96	0.91	0.96	1.1	1.3	1.7	2.0	2.9	0.72
15.4	R	5.50	10.2	11.3	15.4	20.2	32.0	71.0	126.	225.	370.	7.2
	L	1.39	5.20	8.12	12.7	16.9	24.0	49.0	74.0	105.0	124.0	9.3
	R/L	4.0	1.9	1.4	1.2	1.2	1.3	1.5	1.7	2.1	3.0	0.77
25.0	R	7.40	21.0	22.3	25.7	28.5	42.0	97.0	194.	360.	600.	8.5
	L	1.37	5.20	8.14	12.8	17.1	24.5	51.0	78.0	116.	138.	9.3
	R/L	5.4	4.0	2.7	2.0	1.7	1.7	1.9	2.5	3.1	4.3	0.91
50.0	R	10.9	48.0	63.5	78.0	82.0	100.	200.	465.	1080.	1450.	15.0
	L	1.37	5.22	8.17	13.2	17.7	25.3	55.0	87.0	131.	150.	9.5
	R/L	7.9	9.1	7.8	5.9	4.6	4.0	3.6	5.3	8.2	9.7	1.6
75.0	R	14.1	73.0	94.0	155.	133.	172.					19.0
	L	1.37	5.23	8.22	13.6	18.6	27.5					9.7
	R/L	10.3	14.0	11.4	11.4	7.2	6.2					2.0
100.0	R	16.6	84.0	142.	267.	268.	415.					37.0
	L	1.38	5.27	8.55	14.3	20.3	30.6					10.4
	R/L	12.0	15.9	16.6	18.7	13.2	13.6					3.6
125.0	R	18.8	104.	190.	362.	462.						61.0
	L	1.37	5.32	9.07	15.1	21.0						10.8
	R/L	13.7	19.5	21.0	24.0	22.0						5.6
150.0	R	20.5	124.	260.	550.	840.						115.
	L	1.38	5.56	9.30	15.7	22.0						11.4
	R/L	14.9	32.3	28.0	35.0	38.0						10.0

Size of wire to the nearest B. & S. gage number. Inductance given in millihenries.
Resistance given in ohms and ohms per millihenry.

resistance increase is 13 times greater when used in a 10-layer coil than when used in a single-layer coil.

It must be noticed also that this great increase in resistance is not due to the internal capacity of the coil. These multi-layer coils were built on wooden reels in a special way first described by the author; the construction was such that a considerable air space (in this case 0.16 cm.) was used between consecutive layers, this construction giving such a low internal capacity that, up to the highest frequency used the inductance of the coil showed but slight increase. These multi-layer coils were octagonal in form and had 10 layers each, wound back and forth. Each layer was 2.6 cm. long, separated from the next layer by 0.16 cm. air. The inner diameter was approximately 10.5 cm. and the outside diameter varied with the size of wire used, being greater for the larger wires.

The great increase in resistance of coil No. 16 for example is really due to a redistribution of current throughout the cross-section of the conductor. Although the resistance increases 172 times in the frequency range used the inductance is lower at the highest frequency than at the lowest. There are cases shown in which the increase in apparent resistance increases very rapidly for the higher frequencies, even the small-sized wire. Thus the 10-layer coil wound with No. 26 wire increased its resistance from 21.3 ohms to 415 ohms, but at the same time the inductance increased from 23.7×10^{-3} henry to 30.6×10^{-3} henry. Hence for this coil the internal capacity was making itself felt so that the actual increase in apparent resistance must be regarded as due to the combined effect of redistributed current and internal capacity.

Three of the coils were wound with radio cable; in two of them there were used 48 No. 38 enameled strands in the cable—three twisted cables, each having 16 strands, were twisted together to make the cable. The solid wire most nearly approaching this cable in cross-section was No. 22. In the single-layer coils the solid wire increased its resistance by 220 per cent, and this radio cable coil increased by 72 per cent, only one-third as much increase as for the solid wire over the same range of frequency. In the multi-layer coil the solid wire increased its resistance from 6.7 ohms to 267 ohms, an increase of 40 times as the frequency was varied from zero to 100,000 cycles whereas the multi-layer coil wound with the radio cable increased (in the same range of frequency) from 5.1 ohms to 37 ohms, an increase of only 7.2 times, that is, the stranded wire coil showed a resistance increase due to skin effect only one-sixth as great as the nearest size solid wire. In this resistance increase there is some effect due to the internal capacity of the coil, and if this effect (which is approximately the same in amount for both coils) were taken into consideration the superiority of the radio cable over the solid wire would be even more striking.

From the results presented in Tables II and III there was calculated for each coil the "ohms resistance per millihenry" and the results are presented in the form of curves in Figs. 18 and 19. The most interesting conclusion to be drawn from these curves is the idea that the higher the frequency the smaller the wire should be to keep the ratio of resistance to reactance low. Thus in the single-layer coil it is evident that below 40 kilocyles No. 16 wire is better than No. 20 (such factors as cost, bulk, current-carrying capacity, etc., not considered) but above this frequency No. 20 wire is better than No. 16. The size wire which gives the best results depends greatly upon the form of the coil and its inductance. Thus recent tests of coils having a few microhenries inductance, intended for use in circuits employing about 10 megacycles (10^7 cycles), indicate that solid wire, size about No. 12 or No. 14, is the proper wire to give the lowest resistance.

For the multi-layer coils this effect of wire size is shown to a much

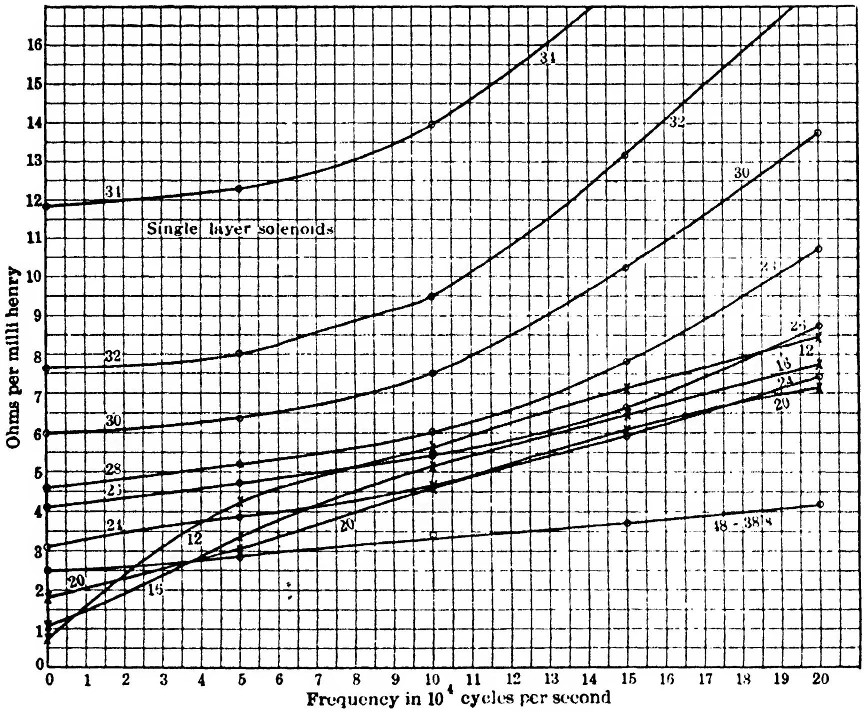


FIG. 18.—Variation of resistance with frequency in single-layer solenoids of various wires. Numbers on curves indicate size of wire, B. & S. gage.

greater degree and at lower frequencies; thus at 1200 cycles No. 34 wire has about 8 times as many ohms per millihenry as No. 12, but at 15 kilocycles the No. 12 wire has more resistance per millihenry than has the No. 34 wire.

The multi-layer coil of radio cable is indicative of what a good coil should be; its reactance at 75 kilocycles is 220 times as much as its resistance and this ratio holds over a wide range of frequency. Other coils have been built by the author, using better stranded wire, more of it, keeping the radial depth of conductor low, which showed a reactance 450 times as great as the resistance at 50 kilocycles. The ideas to be kept in mind in building good radio coils are to use carefully stranded and insulated cable, keep the radial depth of conductor small, keep the coil as compact as possible and at the same time to keep the internal capacity low, and avoid dielectric losses.

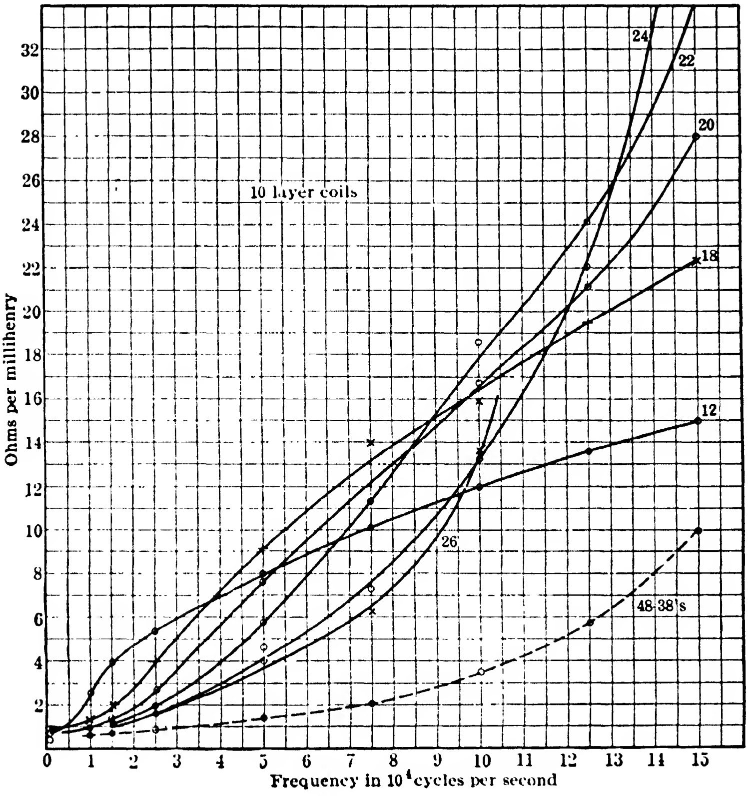


FIG. 19.—Variation of resistance with frequency of multi-layer coils; short coils of ten layers each, air space between layers. Numbers on curves give size of wire, B. & S. gage.

In Fig. 20 is shown the construction of a coil which has about 10 millihenries inductance and 7 ohms resistance at 50,000 cycles; sufficient air space was used between layers to keep the internal capacity to about 120 micro-microfarads. This coil operated satisfactorily when carrying 4 amperes with 12,000 volts across its terminals. If it were used in a good insulating oil it would probably be satisfactory when carrying 200 to 300 kilovolt amperes, although in this case its internal capacity would be more than doubled.

To illustrate in as striking a manner as possible the skin effect which may be present in poorly designed coils tests were made on two coils made of copper strip, wound edgewise. The first coil had $29\frac{1}{2}$ turns of strip $1\frac{1}{2}$ in. wide and $\frac{1}{16}$ in. thick; the spacing between successive turns was nearly $\frac{1}{16}$ in.; its inside diameter was $14\frac{3}{4}$ in. and outside diameter was $17\frac{3}{4}$ in. Its resistance for continuous current was 0.013 ohm and at 150 kc.

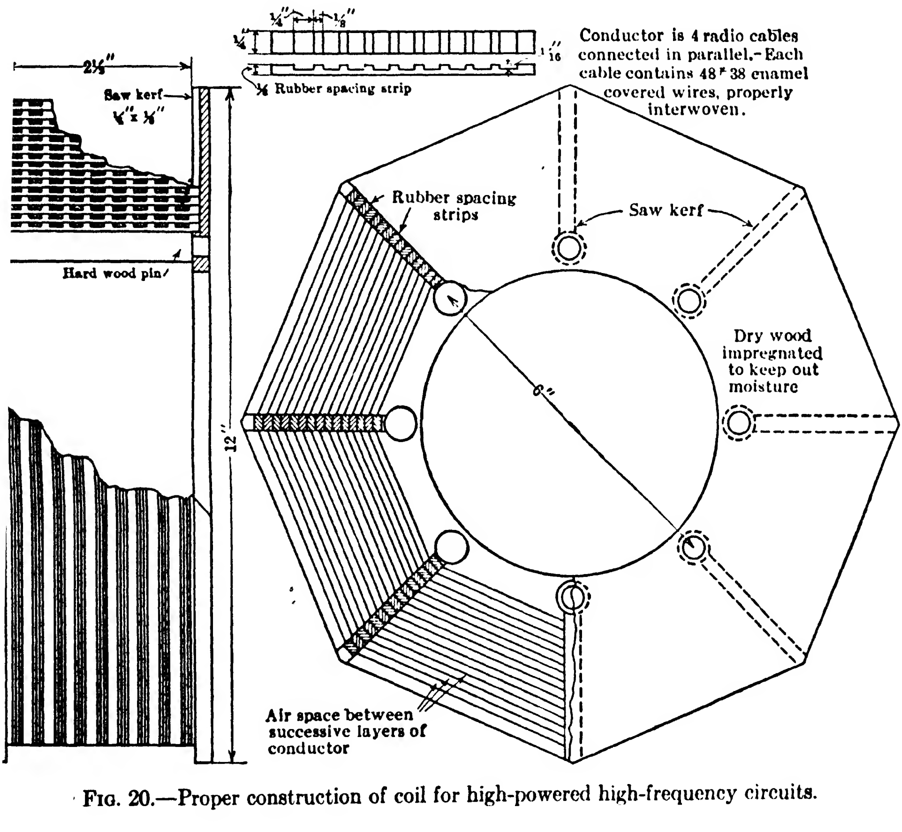


Fig. 20.—Proper construction of coil for high-powered high-frequency circuits.

it had 3.44 ohms resistance, an increase of 285 times. The second coil had $34\frac{1}{2}$ turns of strip 1 in. wide and $\frac{1}{16}$ in. thick, spacing between turns was nearly $\frac{1}{16}$ in. The c.c. resistance was 0.020 ohm and at 150 kc. it was 7.86 ohms, an increase of 393 times. These two examples bring out forcibly the fact that the radial depth of the conductor in a coil intended for radio circuits must be kept small. Further data on these two coils are given in Table VI, p. 225.

Resistance of Coils at Higher Radio Frequencies.—The foregoing values of coil resistances were obtained by bridge measurements; above about 100 kc. bridge measurements become unreliable unless added precautions are taken, so to get the coil resistances in the higher range of radio frequencies a resonance method was used.

In this scheme the coil to be measured is put in series with a variable condenser (of negligible series resistance) of suitable capacity to bring the circuit into resonance at the frequency being used. A suitable thermocouple ammeter serves to measure the current in the circuit and so indicate when resonance is obtained. At resonance there is no net reactance in

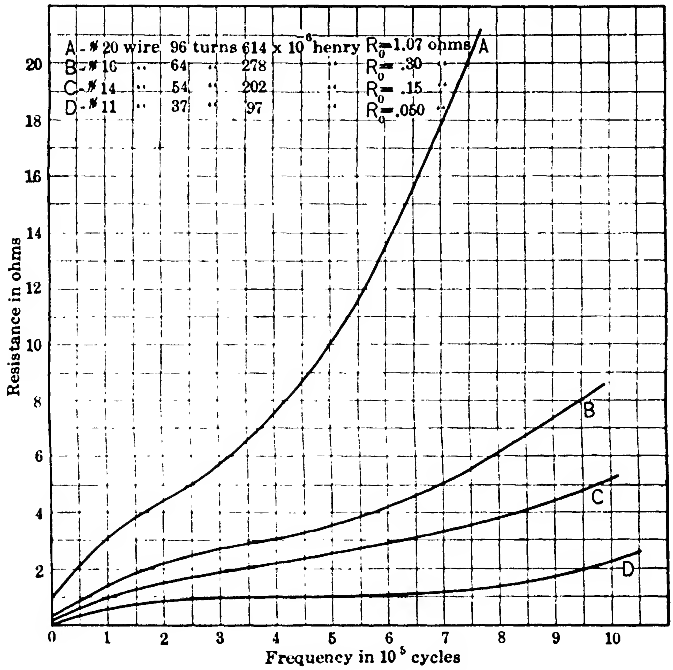


FIG. 21.—Comparison of single-layer solid wire coils at the lower radio frequencies.

the circuit, so that all of the impressed voltage is used up in resistance drop. If then the current (at resonant frequency) is reduced to one-half its original value by adding in series with the circuit a known, non-inductive resistance, the amount of added resistance must be equal to the original value of circuit resistance. If then the resistance of the thermocouple, condenser, and rest of the circuit (exclusive of the coil) is known, the resistance of the coil is readily obtained.

Using the resonance method some of the coils which had previously been measured on the bridge up to 200 kc. were measured at this frequency and the results checked within a few per cent. These coils were then measured throughout the radio frequency band up to frequencies as high as those for which the coils were suitable. The first coils tested gave results as shown in Fig. 21; these four coils can be recognized (by data on the curve sheet) as coils 1, 3, 4 and 6 of Table II, which gives their resistance up to 200 kc. The variation of resistance with frequency is evidently not a simple relation; it increases rapidly for the lower frequencies (lower as the wire is larger) and there occurs a point of inflexion in the curve. In the higher-frequency range the resistance curve is a

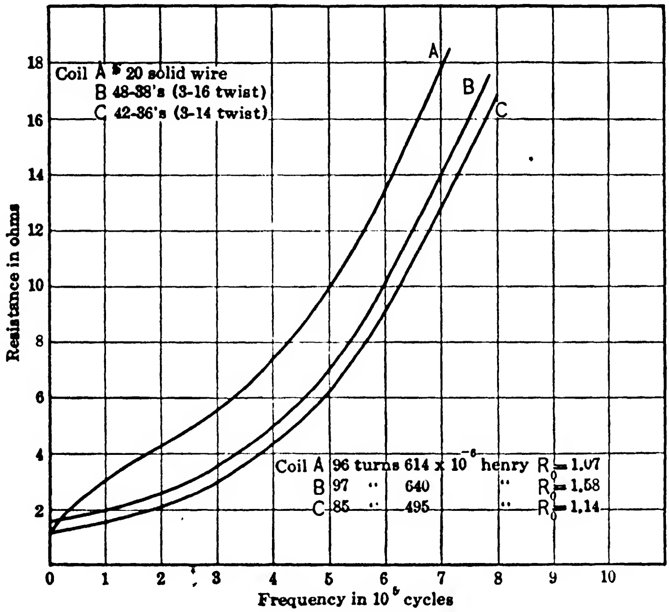


FIG. 22.—Comparison of cable coils and solid wire coils.

smooth one, the increase in resistance being greater (for a given increment in frequency) as the frequency is raised.

In Fig. 22 are shown the comparative merits of cable coils and those using solid wire. The coils were single-layer solenoids (coils 6, 14, and 15 of Table II, p. 180) and the curves show conclusively that within the frequency range from 100 to 1000 kc. the cable coils are better for radio purposes, which practically always demand coils of the lowest possible resistance. It can be seen, however, that the superiority of the cable coil over the solid wire is rapidly decreasing at the higher frequencies. Thus at 200 kc. coil *A* (solid) has about 75 per cent more resistance than the 48–38's cable coil, and this in spite of the fact that for zero frequency the cable coil has 50 per cent greater resistance because of the smaller cross-section of copper. These two coils are just the same size and have practically the same number of turns, as shown on the curve sheet. At 700 kc. the cable coil is still superior to the No. 20 solid wire but it is only about 25 per cent better, suggesting that at still higher frequencies possibly the solid wire may be the better of the two.

The coils of Fig. 22 are not suitable for the higher frequencies, as they have too much inductance. It will be found from experience that a coil suitable for radio circuits should not have more than about 1000 ohms reactance at the frequency considered. Thus coils suitable for a 1000-kc. circuit will have about 150 microhenries of inductance, while those suitable for 500 kc. may have 300 to 400 microhenries.

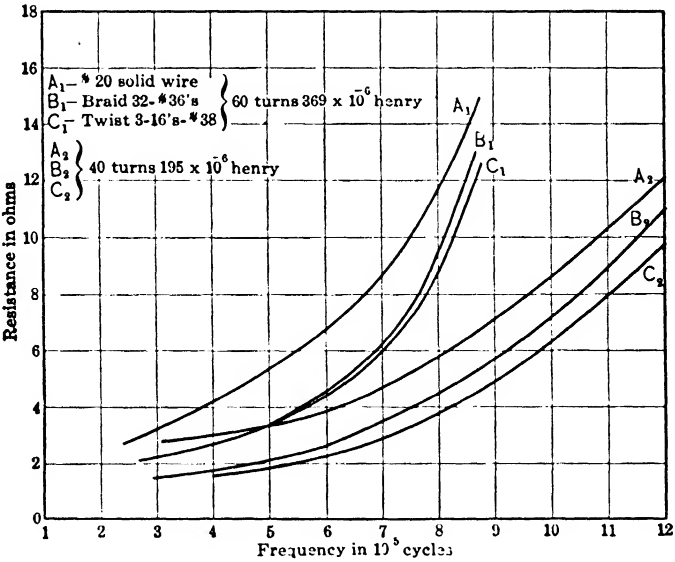


FIG. 23.—At these frequencies cable is still better than solid wire.

In Figs. 23 and 24 are shown the resistance of cable and solid wire coils suitable for frequencies from 300 kc. to 5000 kc., and it will be seen that at the extreme high frequencies solid wire coils gave the better results. It is not impossible that, if cable could be obtained using finer wire with a better type of insulation on the strands, the cable would be superior to the solid wire. Radio receivers for the highest frequencies are, however, built to-day with solid wire coils, as Fig. 24 indicates they should be.

At the highest radio frequencies even the coils of Fig. 24 have too much inductance. In a typical 20-meter receiver, for example, the coil had 8.8 microhenries inductance, and the resistance of the coil (including losses in condenser and input circuit of tube) was 28 ohms. The reactance of this coil is about 800 ohms, so its reactance is about 30 times as much as the circuit resistance.

Effect of Neighboring Coils.—It has been pointed out in Chapter I that the presence of another circuit will always increase the resistance of the circuit in question, and Eq. (108) (Chapter I) indicates that this resistance increase may reach large values if the second circuit happens to be in resonance with the frequency impressed on the first. The curves of Fig. 123, Chapter I, indicate how pronounced this effect may be.

Now coils, of and by themselves, constitute a closed circuit at the higher radio frequencies. Coils always have inherent capacity (one part of the coil acting toward another like one plate of a condenser), and this inherent capacity, in combination with the inductance of the coil, gives the coil a natural frequency of its own. This natural frequency is generally

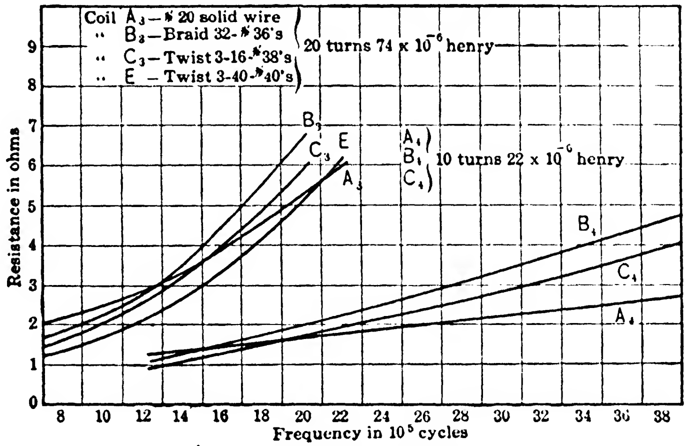


FIG. 24.—At these higher frequencies solid wire may be better than cable.

much above any frequency for which the coil is suitable, so that the natural frequency of a coil assumes importance only in special cases. We can say that the coils generally used in radio work have inherent capacity of about 5 micro-microfarads, as will be shown later in this chapter. Now we have said that a coil suitable for use on a 1000-kc. circuit will have about 150 microhenries inductance. But this amount of inductance in combination with 5×10^{-12} farad capacity will show resonance at about 6000 kc., far outside the range for which the coil is suitable.

It sometimes happens, however, that a receiver is made to cover a wide frequency range; in such a receiver several coils are incorporated, any one of which may be used by a selector switch. These coils may be wound on the same cylindrical form, and, if so, do have some mutual induction. Such a case is shown in Fig. 25(a). Coil A was designed for the highest frequencies, coil *B* for medium frequencies, and coil *C* for the lowest frequencies the set was to receive.

In listening to signals in the neighborhood of 2300 kc. it was found that the set was "dead." No signal of that frequency could be heard unless it was excessively strong. Measurement of the resistance of coil *A* showed the effect given in Fig. 25(*b*). At the critical frequency coil *A* showed a resistance of 90 ohms, whereas on either side of that frequency its resistance was about 3 ohms. Further tests showed that coil *C* had its natural frequency at 2300 kc.

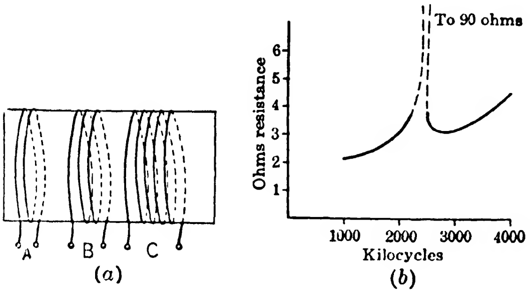


Fig. 25.—Showing the effect of an idle coil on the resistance of a smaller coil in proximity to it.

By short-circuiting coil *C*, the dead spot on coil *A* was completely eliminated; connecting the two terminals of coil *C* together practically nullifies the effect of the inherent capacity of the coil. The dead spot can also be practically eliminated by short-circuiting coil *B*, as such a condition results in almost no voltage being induced in coil *C* from coil *A*. This effect will be taken up under the topic of Shielding, later in this chapter.

Effect of Coil Mounting upon Its Resistance.—Any material on which a coil is wound, or to which its terminals are connected, will necessarily be in the magnetic and electric fields which surround the coil. Unless the material is conductive the magnetic field will have but little effect, but the electric field will produce dielectric loss (discussed later) in the material, which loss will show up as an increase in the apparent resistance of the coil. This effect is generally not large, a typical case being given in Fig. 26, showing the resistance of a coil wound on a bitumen spool. For the first

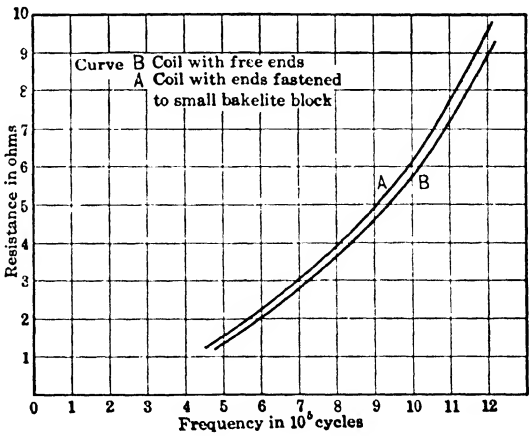


Fig. 26.—Even a terminal block adds somewhat to the effective resistance of a coil.

measurements the coil ends were left free and in the second case they were connected to binding posts mounted in a small block of Bakelite. If fiber had been used for the terminal block, the increase in resistance would have been much greater.

Effect of Using a Tapped Portion of a Coil.—It is generally bad design to use a tapped portion of a large coil when a low inductance is wanted. The unused portion of the coil will act like an autotransformer having comparatively high voltages set up in it. Capacity currents (due to inherent capacities) will flow, giving losses in the unused portion and the insulating material surrounding the unused portion of the coil will have high dielectric loss. This effect is shown in Fig. 27, in which the resistance of a 32-turn coil is given, and also the resistance of a tapped portion, including only 22 turns. And on the same curve sheet is shown the resistance of the 22 turns by themselves, after the extra 10 turns had been removed. Actually there is plotted in Fig. 27, not resistance directly, but the ratio of reactance to resistance. It is this ratio (which is practically the reciprocal of the power factor) which really gives the merit of the coil.

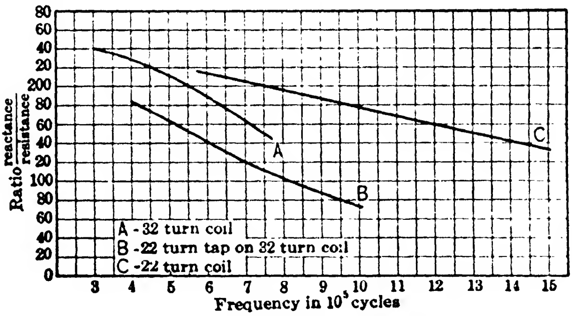


FIG. 27.—These curves show that tapped portions of coils should not be used.

If a tapped portion of a coil must be used, it is frequently preferable to short-circuit that part of the coil not being used, strange as it may seem; the resistance of that part of the coil being used is often decreased by putting the short-circuit on the unused portion.

Resistances of Typical Good Radio Coils.—In Fig. 28 there is shown the resistance of a whole set of radio coils suitable for covering the band of frequencies from 150 kc. to 3000 kc. They were made of cable (48–38's) on spools 4 in. in diameter. They were wound in either two or three layers, according to their inductance. The layers were arranged in the so-called " banked winding," in which the various turns, for a 3-layer coil, are put on in the sequence shown in Fig. 29. The first and second turns are wound side by side on the spool, which must have a rough surface. The wire is then turned up over the second turn and the third turn put on. The wire then jumps down and winds the fourth turn on the spool surface and then jumps on top of the fourth and winds the fifth, etc. The numbers in the cross-section of the wires show the sequence of winding.

The coils used to get the data of Fig. 28 were on bitumen spools and were thoroughly impregnated with shellac and dried before being measured.

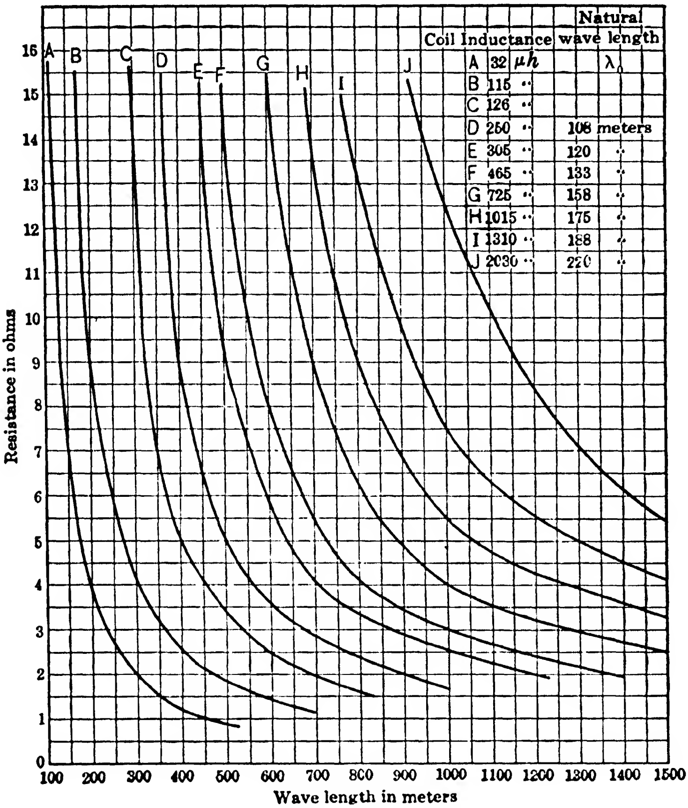


Fig. 28.—Resistance curves for a set of bank-wound coils suitable for radio circuits; they had either two or three layers of finely stranded cable.

When shellac is used on coils, it must be *very thoroughly dried* (preferably by heating the wires of the coil itself), or a large loss due to leakage currents will result.

Reasonable Power Factors for Radio Coils.—From the data given in Fig. 28, and other data the author has, it is possible to plot a curve as shown in Fig. 30, giving reasonable ratios of reactance to resistance for typical radio coils. To get ratios as high as those given in this figure the coil must be suitably chosen for the

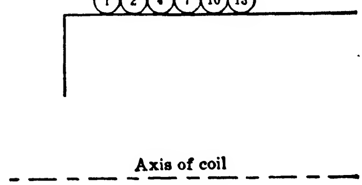


Fig. 29.—Showing how bank-wound coils are made.

frequency of the circuit. From this figure it is seen that a reasonable power factor for radio coils varies from possibly $\frac{1}{3}$ per cent at 200 kc. to about 1 per cent at 3000 kc. These figures refer to coils used in radio receiving apparatus. For the large coils used in transmitting stations, weighing hundreds of pounds, a power factor as low as 0.002, or even lower, may be reached.[1]

Theory of Increase in Coil Resistance.—In general, the exact theoretical investigation of the increase in coil resistance, as frequency is increased, involves considerable effort. Any such treatment must take into account the gradient of the magnetic field in which the various turns lie, the effect of the mutual induction of all the other turns on the turn being examined, effect of the space necessary between adjacent turns for insulation, etc. Even after many simplifications have been made, the solutions generally involve coefficients expressible only by infinite series sometimes only slowly convergent.

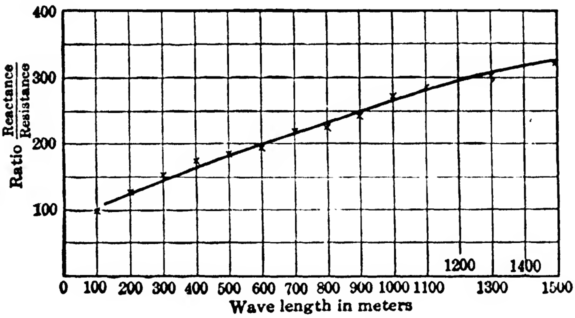


FIG. 30.—Reasonable ratios of reactance to resistance in good radio coils.

For single-layer solenoids of solid wire Lindeman has proposed the relation

$$R = R_0(1 + \alpha\sqrt{f} + \beta f^2), \quad \ldots \ldots \quad (7)$$

where R = a.c. resistance;
R_0 = c.c. resistance;
f = frequency.

In the formula α and β are constants which must be determined experimentally; they must be obtained from measurements at frequencies much lower than the natural frequency of the coil.

For single-layer solenoids of cable he suggests another formula

$$R = R_0 \frac{\pi^2\omega^2N^2r^2}{2\sigma^2y^2}, \quad \ldots \ldots \quad (8)$$

[1] See "Design and Efficiencies of Large Air Core Inductances," by Brown and Love, in Proc. I.R.E., Dec., 1925. Reasonable coils have about 800 ohms reactance at the middle of their useful frequency range.

where N = number of strands in the cable;
r = radius of one strand;
σ = resistivity of material in electro magnetic units;
y = radius of cable;
$\omega = 2\pi f$.

Effect of Broken Strands in Cable Coils.—It has often been stated that a serious drawback in the use of finely stranded cable for radio coils was due to excessive resistance in the coil if one or two of the strands were broken, or not properly soldered to the others at the terminals of the coil. Analysis shows no logical reason for such an effect, and the experimental results given below show that if there is such an effect it is negligibly small.

A single layer solenoid $4\frac{1}{2}$ in. in diameter 1.3 in. long was wound with 26 turns of cable made up of 3–16–38's enamel-covered wire. In making the cable the manufacturer twists together 16 of the No. 38 enamel-covered wires and then twists together three of these cables, covering the whole with silk wrap. The calculated resistance for the length of the cable used was 0.467 ohm, and the measured c.c. resistance was 0.483 ohm. One of the strands was undoubtedly broken inside the cable somewhere. The inductance of the coil was 109 microhenries.

Its resistance was measured throughout the range of frequencies for which such a coil is suitable. One strand was then broken, at one end, and the resistance was again measured. Then one more strand was broken (at the same end), making two broken strands, and its resistance was again measured. Similar tests were made as detailed in Table IV; whereas

TABLE IV

Frequency in Kilocycles	Resistance in Ohms						
	Complete	One Strand Broken	Two Strands Broken	Four Strands Broken	Ten Strands Broken	Twenty Strands Broken	Forty-seven Strands Broken
560	1 73	1.85	1.78	1.90	1.95	2.30	19.6
650	1.69	1.72	1.51	1.93	1.97	2.17	26.1
790	2 10	1 98	2.19	2.25	2.50	2.75	25.3
970	2 32	2 32	2.31	2.35	3.05	3.29	30.0
1250	5 46	5.50	5.53	5.50	5.45	6.60	30.4
1500	8 68	7.47	7.33	7.69	7.42	8.90	46.6

the results seem to be somewhat erratic, to a small degree, they do not show any great harm resulting from a few broken strands in such a cable.

Cable coils do, of course, have a marked disadvantage over solid-wire coils from the viewpoint of the manufacturer; it is a tedious task to clean the enamel from each of the strands, which of course must be done before they can be soldered together.

Neighboring Circuits Heated by Eddy Currents.—As an instance of the losses occurring in neighboring circuits it is interesting to note that one of the terminal posts of the coil pictured in Fig. 20 was fastened on a piece of hard rubber, and this rubber block was fastened to the wood end-piece of the coil with small iron screws. When operating this coil with 4 amperes at 50 kc. flowing in the winding the heat generated in those screws was such that they burned themselves free from the wood after the coil had been in the circuit but a short time. Nowadays the heating of masses of metals by the eddy currents set up by high-frequency currents is assuming real commercial significance. Furthermore, such eddy currents can be thus set up in metal masses which are inside an evacuated glass vessel, a very useful procedure in the manufacture of vacuum tubes.

Resistance of Iron-core Coils.—It was at first thought impossible to use iron-core coils for the high frequencies employed in radio circuits, but such is not actually the case (at least for the lower radio frequencies), although the gain in using iron is not so great for radio frequencies as it is for ordinary low frequencies. The difficulty in making efficient iron-core coils for high-frequency circuits is a two-fold one—the apparent permeability of the iron is much less than it should be, and the losses in the iron core greatly increase the effective resistance of the coil. Both of these undesirable effects are due to the same cause; the increase in resistance due to iron loss is mostly caused by eddy currents in the iron laminas, and these same eddy currents serve to keep the magnetic flux from penetrating and so make only the outer parts of the laminas useful as flux carriers. There is also some iron loss due to hysteresis, but this is small compared to the eddy-current loss.

The paths of the eddy currents in the laminas of an iron core are indicated in Fig. 31, the laminas being shown much thicker, of course, than they really are. The direction of the eddy currents, it is to be noticed, is opposite to that of the current in the winding, hence at any point A in the center of a lamina, the magnetomotive force acting is really that produced by the winding diminished by a certain amount due to these eddy currents. At low frequencies the back m.m.f. of these eddy currents is negligible compared to that of the main magnetizing coil, so that the flux density in the lamina is nearly constant throughout its cross-section and is about the same as it would be were no eddy current present. At higher frequencies, however, the eddy-current effect becomes increasingly greater, so that at radio frequencies the full value of magnetic flux exists only on the outer surface of the iron; in the inner parts the flux density decreases and it may be practically zero at a depth only a small fraction of a millimeter from the surface of the iron.

The strength of the eddy currents decreases with the thickness of the laminations; the plates used for the cores of radio coils should be only a few hundredths of a millimeter thick. To get the benefit of lamination it is essential that the plates be well insulated from one another, either by a fine quality of varnish or thin paper, or both. The burred edges of the plates, caused by imperfect fit of the punch and die used in making the plates, is especially bad in causing eddy currents.

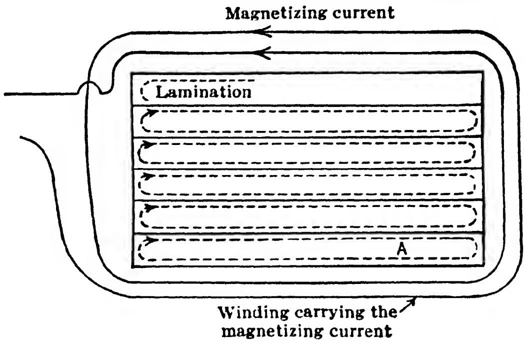


Fig. 31.—Eddy currents occurring in a laminated iron core.

The flux density in the steel plates has about the distribution shown by Fig. 32; the penetration of magnetic flux into an iron sheet decreases as frequency increases, increases with the specific resistance of the iron, etc., in fact follows the same distribution law as the penetration of current itself into a conducting medium given in Eq. (3). Because of this

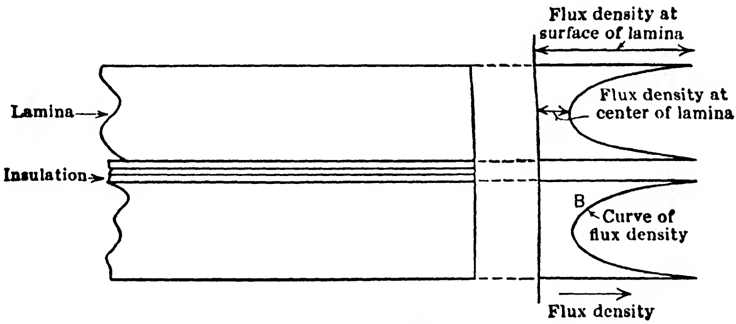


Fig. 32.—Flux density variation throughout the section of a lamination of an iron core; the higher the frequency the greater is the variation in flux density.

lack of penetration and because of the phase lag of the flux in the inner part of the core the apparent permeability of the iron decreases as the frequency increases, resulting in a decrease in the self-induction as the frequency increases. The curves given in Fig. 33 show how the resistance and inductance of a laminated iron-core coil change as frequency changes.

It will be noted that the increase in resistance is practically all due to eddy-current losses; the hysteresis loss is nearly negligible. It is evident that iron of the quality used in this coil must be used in laminations less than 0.0075 cm. (the thickness of those in the test coil) to maintain a low resistance at high frequency. The decrease in inductance of this coil is comparatively small in the range of frequencies used. The effective permeability of the iron in this core proves to be just 100.

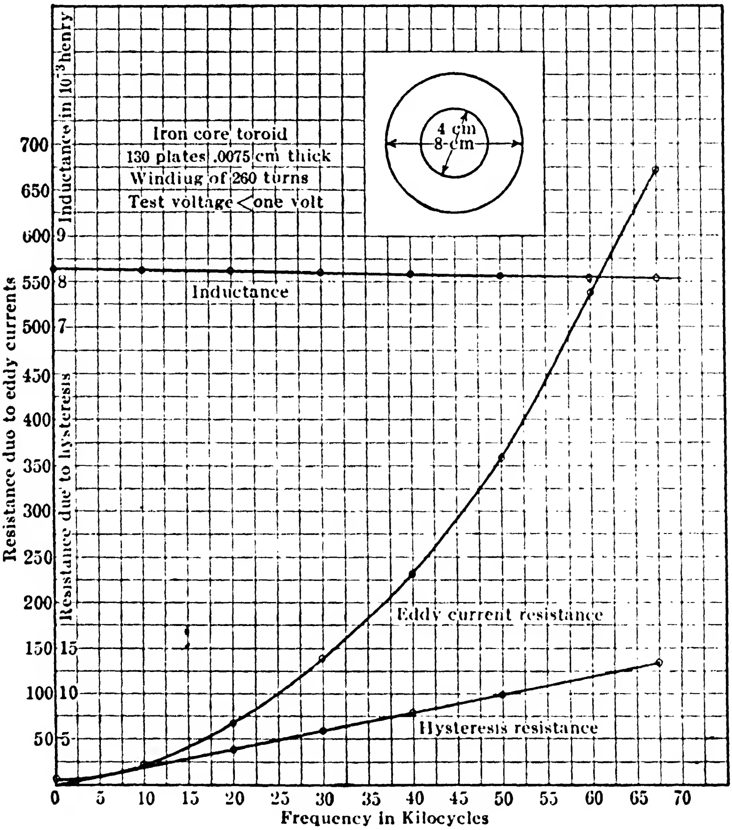


FIG. 33.—Characteristics of a toroidal-shaped coil having laminated iron core.

In Fig. 34 are shown the variations in L and R of another toroidal coil, using thicker laminations. Even for the comparatively low frequencies used in this test the decrease in inductance is very pronounced.

The effective permeability of the iron in this core varies from 100 at the lower frequency to about 60 at 25 kc. In the section of this chapter dealing with inductance more material in permeability of iron cores will be given.

Action of Iron at High Frequency.—It is to be remembered that the effective permeability is obtained by the ratio of the total flux produced in the iron and the magnetizing ampere terms of the coil. Certain experiments seem to indicate that the actual permeability of iron is independent of frequency, and this is probably the fact. The ratio of flux to ampere

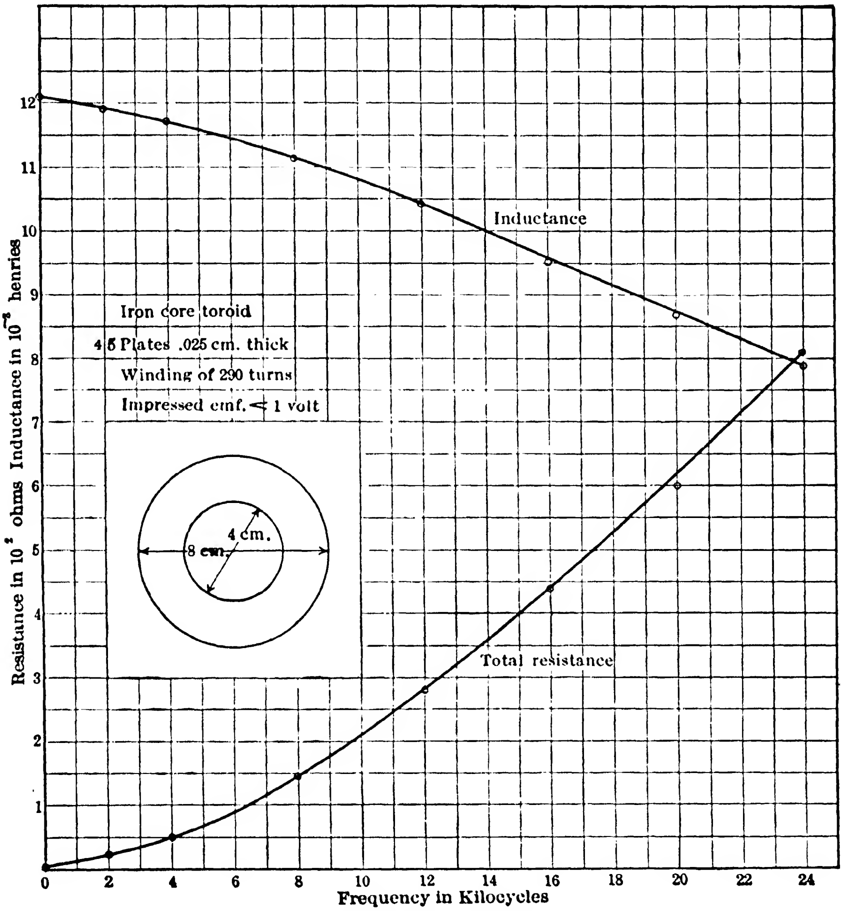


FIG. 34.—Characteristics of a toroidal-shaped coil having laminated iron core, laminations being much thicker than those of Fig. 33.

turns of the magnetizing winding goes down as the frequency increases, and this results in the decreasing " effective " permeability; however, if the ratio of flux density to the **net magnetizing force** is obtained a constant value will probably be found. That is, if we consider the ratio of the flux density to the (magnetizing force of the coil—the demagnetizing force of the eddy currents), probably no change takes place as frequency increases.

The flux density throughout the cross-section of a plate not only changes as indicated in Fig. 32, but its phase changes also. The flux at a certain instant may actually be in one direction at the surface of the iron plate and in the opposite direction in the center of the plate. The density of magnetic flux, with increasing distance from the surface of the plate, follows the same law of diminution in magnitude, and falling behind in phase, as does the current flowing down a long transmission line.

It is sometimes convenient to speak of the " depth of penetration of the flux into a piece of iron. Now actually the flux always penetrates to the center of the plate, but the meaning of the term depth of penetration is taken as that depth which multiplied by the density of the flux at the surface, gives a total flux equal to the actual total flux.

In Fig. 35 the flux density distribution in a plate is shown by the curve *ABCDE*, the distance to this curve from the edge of the plate, *IKHG*, being the flux density, the total flux in the upper half of the plate is shown in the cross-hatched area, *ABCKI*. Now in the lower half of the diagram the line *HF* has been taken in such a position that the area *FEGH* is equal to that of *ABCKI*. Then the distance *FE* is taken as the " depth of penetration" of the flux. MacLachlan has reported[1] some experiments in the action of iron at high frequencies; a few of his results are given here. Using silicon steel plates and " pure " iron plates, each 0.25 mm. thick, he found the depth of penetration as shown in the table here given. The specific resistance of the silicon steel was taken as 54.5 microhms per centimeter

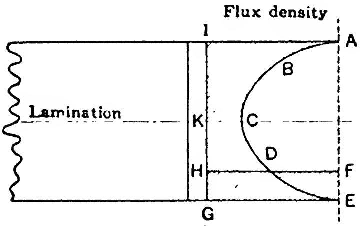


FIG. 35.—Diagram to illustrate the meaning of the term "depth of penetration."

	By Theory	By Test	Frequency
Silicon steel......	0.011 cm.	0.0075	2×10^5
Silicon steel......	0.0072	0.0045	5×10^5
Pure iron.........	0.0054	0.0025	2×10^6
Pure iron.........	0.0035	0.0017	5×10^6

cube, and that of the iron as 12.5 microhms. The higher the resistance the smaller are the eddy currents and hence the greater is the penetration; this is shown in the above table.

[1] Journal I.E.E., Vol. 54, p. 480.

The ratio of flux density to magnetizing force is of course much lower at high frequencies than at low ones, primarily because of the lack of penetration as explained above. MacLachlan gives some B–H curves for the silicon and iron plates he used; some of them are reproduced here. The flux density (B) plotted is that calculated on the assumption that the flux density throughout the cross-section of the plates was uniform; as we know this is not the case it is evident that the flux density at the surface of the plates was considerably higher than the value shown by the curves.

Fig. 36 shows the action of plates of silicon steel (0.25 mm. thick) with variations in frequency and magnetizing force. It is seen at once that the flux does not penetrate as well in the pure iron as it does in the silicon plates. At 2×10^5 cycles the total flux through the silicon plates is about twice that through the iron plates, but it is to be

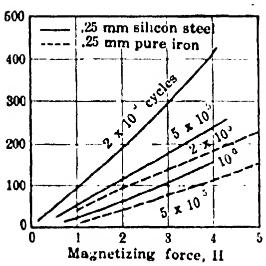


FIG. 36.—Average flux density vs. magnetizing force at radio frequency.

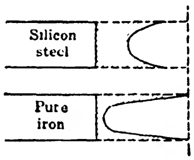


FIG. 37.—Possible flux density distributions in two different irons.

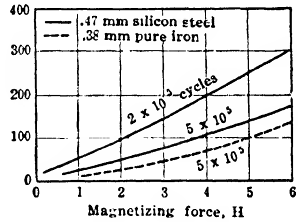


FIG. 38.—Curves similar to those of Fig. 36, for somewhat thicker plates.

remembered that possibly the flux density at the surface of the pure iron was as large as, or even larger than, that of the silicon plates, but owing to the lesser penetration the total flux, from which the B of Fig. 36 is calculated, is less.

This idea is indicated in Fig. 37, which shows a possible flux density distribution for the two plates. The silicon plate has much less surface density of flux than the iron one, but the total flux (shaded area) is greater for the former because of the better penetration. In Fig. 38 other B–H curves are given for plates of the same composition as those of Fig. 36, but different thickness. Even though there may be considerable flux penetration in the iron at radio frequencies its application is of but little value at present, because of the high losses occurring in the iron, due to eddy currents and hysteresis. To be of service for tuning radio receivers the reactance of a coil must be much greater than its effective resistance; the iron losses act to give a high effective resistance as shown by the accompanying table of power factors which MacLachlan found for his iron-core coils, using plates 0.25 mm. thick.

	Frequency	Power Factor
Silicon steel.......	2×10^5	0.90
Pure iron.........	2×10^5	0.84
Silicon steel.......	5×10^5	0.92
Pure iron.........	5×10^5	0.86

With power factors as high as these the reactance cannot be high compared with the resistance; in fact, it is actually much less than the resistance.

Suitable Iron for Cores.—In general, iron-core coils are used in radio circuits only for currents of audio frequencies. These vary from about 50 cycles per second to several thousand.

Iron which has been developed for use at the commercial frequencies of 25 cycles and 60 cycles is not of the right characteristics for radio apparatus. Ordinary " electric steel " has too high an eddy-current loss at the higher audio frequencies, whereas at engineering frequencies the eddy-current loss is small compared to the hysteresis loss. Steel with a high percentage of silicon (3.5 per cent) seems suitable for many purposes. At a flux density (maximum) of 15 gausses (15 lines per square centimeter) at 1000 cycles plates of this steel 0.03 cm. thick show an eddy-current loss of about 1 erg per gram per second and hysteresis loss of about 5 ergs per gram per second.

Permalloy.—It has recently been discovered that a certain nickel-steel alloy shows remarkable magnetic properties for very low magnetizing forces. If the alloy has about 79 per cent nickel its permeability may go to 100,000 or more; just what permeability the material shows depends principally upon what magnetizing force is employed and the treatment given to the alloy in the process of manufacture.

As will be explained later in this chapter iron cores used in radio circuits generally are subjected to a continuous magnetomotive force as well as the alternating one; in this case the permeability for the alternating current depends largely upon the continuous magnetomotive force present. This effect is especially pronounced with permalloy cores.

In case two alternating currents of different frequencies are acting on the core, each may have a marked effect on the permeability for the other, as is explained later under the topic of Self-induction.

Iron Dust for Cores.—Many experimenters have had the idea of using very finely pulverized iron dust for cores of coils to be used at high frequencies. By having the dust particles fine enough and putting them together, to make the core, in some way that will preserve a high resistance between the iron particles, it should be possible materially to reduce the eddy-current loss and at the same time retain a reasonable permeability.

The idea seems simple enough, but it took a deal of experimenting and money to accomplish the desired result; the dust which the early experimenters put together with shellac or other binder seldom had a permeability of more than five. The telephone engineers had great need for iron-dust cores (for telephone loading coils) and have now produced an iron-dust core of remarkable properties.

Electrolytic iron, saturated with hydrogen and very brittle, is pulverized, finely coated (sometimes with zinc) and mixed with an insulating cement; it is then put into ring-shaped molds and subjected to extreme pressure. By using a pressure in excess of the elastic limit of the iron particles, the dust grains are made actually to flow, thus binding themselves together. This dust has about the same density as ordinary iron, thus showing that practically all the cement is forced out. While by no means as mechanically strong as ordinary iron, these molded cores can be handled with impunity without breaking; on filing the core the exposed surface has, to the naked eye, exactly the appearance of ordinary iron.

Iron dust, thus suitably prepared, makes very excellent material for the cores of coils intended for high-frequency use, having very low eddy-current loss. It has, in common with all iron cores, the disadvantage of a permeability varying with magnetic density. A dust-core coil (toroid) was tested to show this effect and gave the results shown in Fig. 39; the measurements were carried out at a frequency of 2000 cycles. In Fig. 40 are shown the characteristics of this same coil measured at various frequencies. The ratio of reactance to resistance brings out very well the fact that a coil is always most efficient at some certain frequency. The permeability of this iron dust is about 50, holding this value to above 100 kc.

It is to be noticed for all these iron-core coils that, at radio frequency the resistance of the copper wire is negligible compared to the resistance caused by iron losses, hence it is of little use to employ for the winding a wire as large as those used in winding the coils whose characteristics are shown in the foregoing figures. A very fine wire may be profitably used for the windings, then a large number of turns (hence high L) may be put on a very small core—using an iron-dust toroidal core, the outer diameter of the toroid being 5 cm. and the core itself being slightly over 1 cm. in diameter, winding with fine wire, it is possible to make a coil with an inductance of about 0.25 henry, and having low enough internal capacity to be efficient at 60,000 cycles. Such coils are very convenient for the plate circuit of amplifying tubes, as they are very compact, and are not subject to outside disturbances, a toroid having practically zero mutual induction with any other circuit.

The losses in iron cores have been measured by a calorimetric method [1] and they indicate the same effect of eddy currents as that given by the decrease in self-induction shown above.

Resistance of Spark and Arc.—In radio circuits there is sometimes used an arc, or a spark gap, the resistance of which affects the operation of the set and must be considered when getting decrement, losses, etc. The resistance of a spark gap varies with many factors, principally the length

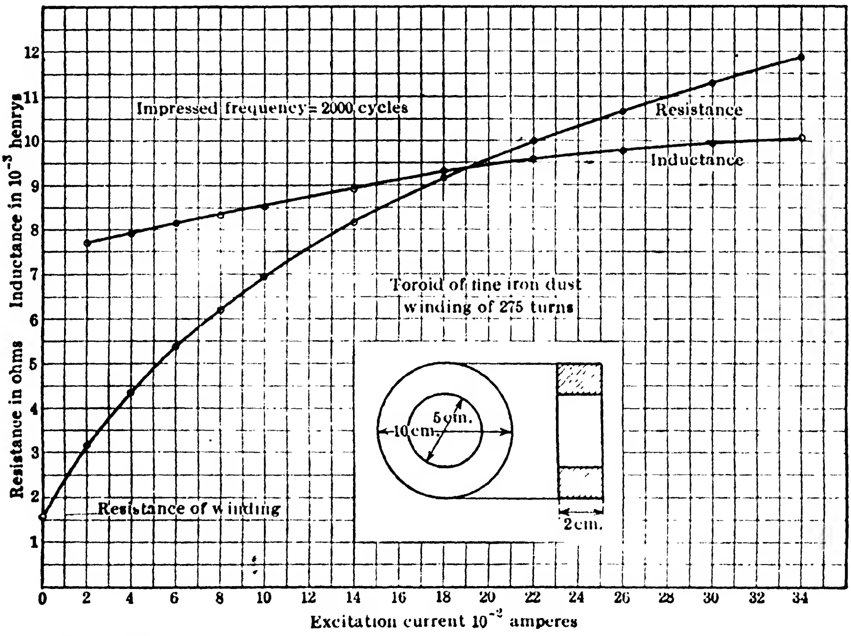


FIG. 39.—Variation of inductance and resistance of a coil having iron-dust core, showing increasing permeability with increasing current.

of gap and magnitude of current through the gap. Within the ordinary range of currents used in radio circuits the resistance of an arc or spark, for constant length of gap, varies inversely with the current to some power higher than the first, in such a way that the IR drop actually decreases with an increase of current. Other factors affecting the resistance of the gap are the nature of the gas through which the arc or spark is passing and the material of which the gap terminals are made. Silver and copper electrodes give a higher resistance gap than such metals as

[1] "Hysteresis and Eddy-Current Losses in Iron at Radio Frequencies," Nussbaum in Proc. I.R.E., Vol. 7, No. 1, Feb., 1919.

zinc, magnesium, etc.; hydrogen and illuminating gas give a higher resistance than air.

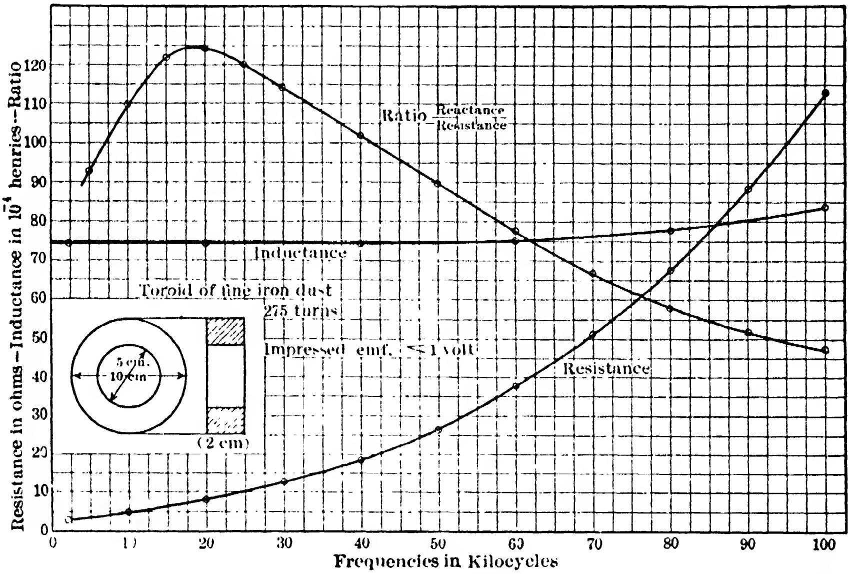


FIG. 40.—Effect of frequency upon the characteristics of a toroidal coil having a core of fine iron dust.

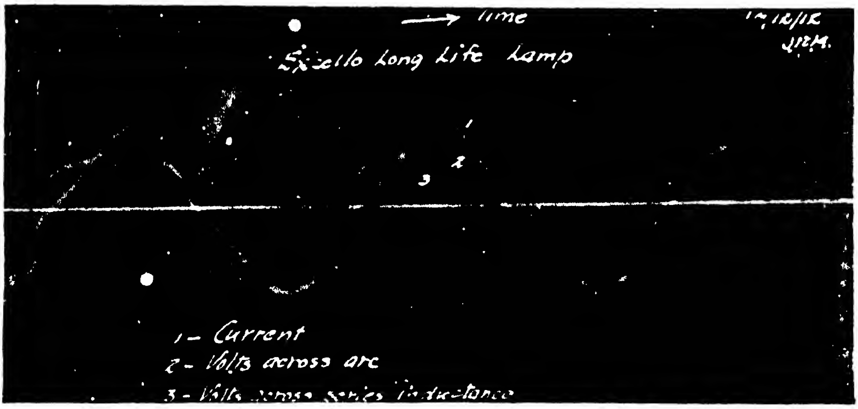


FIG. 41.—A sine wave of e.m.f. impressed across an arc in series with an iron-core inductance gave voltage and current forms as shown. There was much metallic vapor used in this arc.

Experiments indicate that for such currents and gaps as are used in radio sets the resistance (effective) of a spark gap is not more than 1 ohm, and is generally only a few hundredths of 1 ohm. This value of resistance is obtained from the heating effect, and so gives a kind of average value of the resistance during the cycle. For low frequencies the resistance

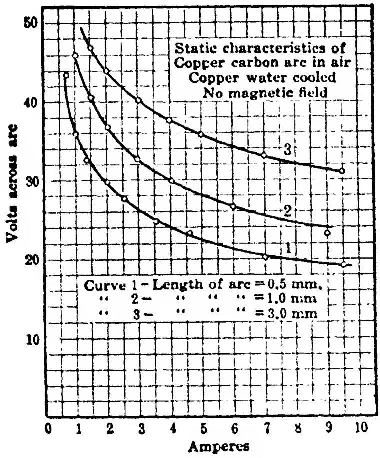


FIG. 42.—Resistance of small arc, in air.

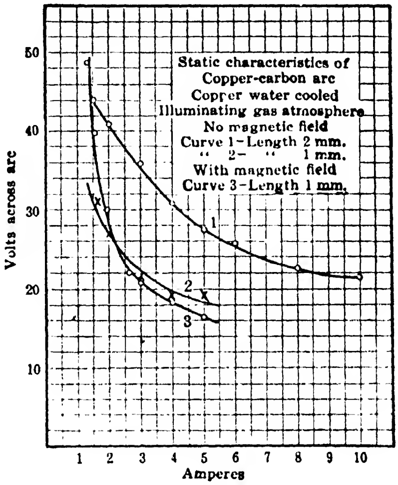


FIG. 43.—Resistance of small arc in illuminating gas, showing also the effect of a transverse magnetic field.

of an arc or spark *varies a great deal throughout a cycle of current*, and it probably does (even if to a less extent) at radio frequencies.

In Fig. 41 is shown the oscillogram giving the form of voltage across an arc and the current through the arc; if the resistance is defined as the ratio of volts to amperes it is evident that the resistance varies through widely differing values during the cycle.

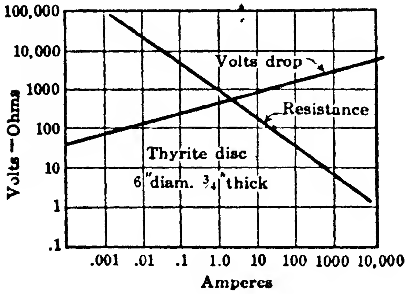


FIG. 44.—Characteristics of a disc of thyrite.

In most radio circuits the resistance which assumes importance is that offered to alternating currents, rather than the c.c. resistance. Nearly all circuits offer a greater resistance to the flow of alternating current than they do to continuous current, but the arc is one of the exceptions to this rule. The relation of current (c.c.) and potential difference in a certain arc is shown in Figs. 42 and 43; from the curves it is evident that increasing current requires less and less potential difference across the arc. The resistance of this arc for continuous current varies from 50 ohms to 2 ohms, according to conditions. *Now the a.c. resistance must be determined by the ratio of the voltage change to the current change,* i.e., $R = dv/di$, and from these curves it is evident that when dv is positive di is negative, so that the a.c. resistance of the arc is *negative.* This negative resistance for alternating currents is characteristic of nearly all circuits in which gas of some kind takes part in the conduction of the current; in some special cases even a pure electron stream may have a negative resistance for an alternating current, as will be explained when discussing vacuum tubes.

Alternating-current Resistance of Arc.—In general we may say that the a.c. resistance of an arc is negative; how large this negative resistance is depends upon many factors.

By the statement that the a.c. resistance of the arc is negative is not meant that an arc, operating on alternating current, does not absorb power and give off heat. Evidently such a statement would be absurd. The statement means that if an arc operating from a c.c. source is now made to carry *in addition* a small alternating current (thus making the arc current a pulsating one), for this *superimposed alternating current the resistance will be negative.*

If, for example, an arc is operating with 10 amperes of continuous current flowing through it, and a 500-cycle current of 1 ampere is sent through the arc, in addition to the steady current of 10 amperes, the 500-cycle current will encounter negative resistance. The arc will actually be cooler (give off less heat) with the 10 amperes c.c. and the 1 ampere a.c. flowing simultaneously than when the 10 amperes is flowing alone.

As the amplitude of the alternating current approaches that of the continuous current the negative resistance it encounters decreases in value; this is also true if the alternating current is kept low in magnitude, but the frequency is increased. Above about 500 kc. it is difficult to make an arc develop appreciable negative resistance.

Materials with Peculiar Resistance Characteristics.—Carbon is a material which acts differently from most other conductors in that its resistance decreases as its temperature rises; this means that as the voltage across a carbon filament is raised the current rises in even greater proportion. Some synthetic substances show this characteristic to a remarkable degree. One of them, with the trade name *thyrite,* shows a tremendous change in resistance values, as the voltage impressed on it is increased. In Fig. 44 are shown the variations in impressed voltage and resistance for various values of current; the piece of thyrite from which these results were obtained was in disc form, 6 in. in diameter and $\frac{3}{4}$ in. thick.[1]

Resistance of an Antenna.—An antenna is a circuit consisting of a capacity, inductance, and resistance in series, the resistance being fixed in value by many effects, among them the radiation of power from the

[1] Jour. A.I.E.E., May, 1930, p. 350.

antenna in the form of electromagnetic waves. The surface of the earth generally forms one plate of the condenser and the overhead wire system the other, Fig. 45. When the current circulates in the antenna, losses occur

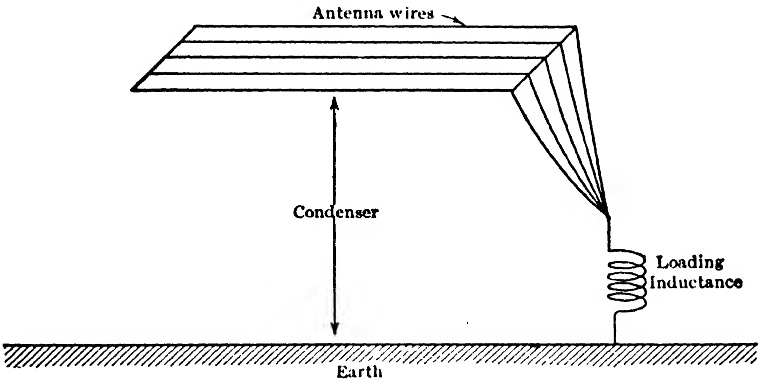


FIG. 45.—Antenna with loading inductance.

in the network of wires and in the earth due to actual heat loss produced by the conduction currents; losses occur due to induced currents in guy wires, etc.; losses occur in the earth's surface and any other dielectrics in

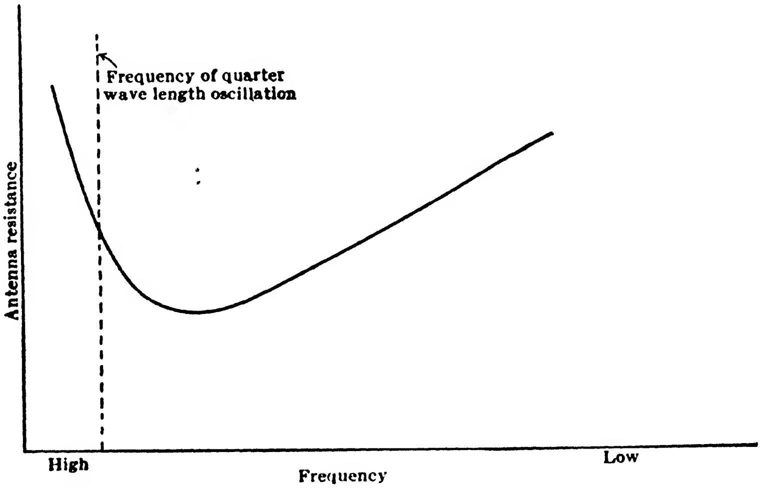


FIG. 46.—Typical resistance curve for an antenna, showing variation with frequency.

the field of the condenser such as trees, etc., and power is radiated. That resistance caused by radiation is the only useful resistance; the other factors increase the resistance of the antenna, but perform no useful function.

The resistance of an antenna varies with the frequency about as indicated in Fig. 46. With very high frequencies the resistance is high; it decreases to a minimum at a frequency about twice that of the natural oscillation of the antenna (without any added inductance) and then rises gradually, the amount of this rise being determined principally by dielectric losses in objects located in the electrostatic field of the antenna.

For small land antennas the minimum on the curve may be 20–30 ohms, for aeroplane antennas perhaps 5–10 ohms; for ships' antennas 3–6 ohms, and for the large antennas used for long-distance communication the minimum value may be 1–2 ohms. The more complete discussion of antennas and their characteristics will be given in Chapter IX.

Resistance due to dielectric loss and corona loss will be treated in the section dealing with capacity.

INDUCTANCE

Self-induction

Coefficient of Self-induction. Units.—The ordinary unit of self-induction, the henry, is much too great to serve for radio work; instead there are used the millihenry (10^{-3} henry) the microhenry (10^{-6}) and infrequently the centimeter (10^{-9} henry). The microhenry is most commonly used.

The fundamental viewpoint from which to consider the self-induction of a circuit is that of the amount of energy stored in the magnetic field of the circuit. This energy is given by the well-known formula, Energy $=LI^2/2$, where the energy is measured in joules, the current in amperes and L is the coefficient of self-induction in henries. From this equation we can get the definition that the coefficient of self-induction of a circuit in henries, is numerically equal to twice the number of joules stored in the magnetic field when the current in the circuit is 1 ampere. Hence any conditions which affect the magnetic energy in a circuit, the current staying fixed, must affect the coefficient of self-induction.

A piece of iron introduced into the magnetic field decreases the reluctance of the magnetic path, increases the flux and hence the magnetic energy, and therefore increases the L of the circuit. If a neighboring closed circuit is so placed that current is caused to flow in it by an alternating current in the coil in question, this induced current may be nearly in phase opposition to the current in the first circuit. Flux which is produced by A (Fig. 47) and which threads circuit B is opposed by the m.m.f. of the current induced in B, and hence the reluctance of this part of the magnetic path of coil A is increased. This decreases the flux produced by a given current in A and so proportionately decreased the L of coil A. It is evident that the closer circuit B is placed to circuit A, the more effect will its counter m.m.f. have on the amount of flux produced by A, and hence the more effect it will have in decreasing the L of coil A.

If by any means the current in B is made to lead the voltage induced in B by 90° (as may be nearly done by putting a suitable condenser in series with the circuit) then the m.m.f. produced by B occurs in such phase as to help the m.m.f. of coil A and hence produce an increase in the magnetic energy of A for a given current, thus the presence of coil B actually increases the effective self-induction of coil A. These effects were analyzed in the previous chapter and are calculable from Eqs. (109), etc., pages 125 et seq.

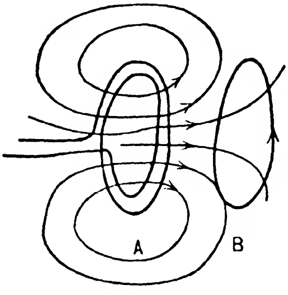


FIG. 47.—When the current in circuit A is varied, current will flow in circuit B, affecting thereby the resistance and inductance of circuit A.

Self-induction of Iron-core Coils.—As has been pointed out earlier in this chapter it is customary to use iron-core coils in radio apparatus except for those coils through which the radio-frequency currents flow. For circuits of 100 kc. perhaps, or possibly somewhat higher, iron-core coils offer certain advantages, cheapness, compactness, and freedom from stray magnetic fields when compared to air-core coils. To get the most advantage from iron, however, it is necessary to appreciate some of the peculiarities of iron when subjected to various magnetomotive forces.

We must distinguish three cases when discussing the behavior of iron-core inductances; coils in which simple harmonic currents are flowing, coils in which a continuous current as well as a sine wave of current is flowing, and coils in which two or more sine-wave currents of different frequencies are flowing at the same time. These cases are more complicated in the order of the above tabulation.

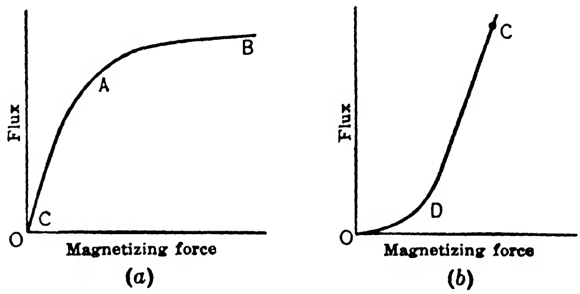


FIG. 48.—The ordinary form of B-H curve frequently leads to error; the bottom of the curve is shaped as in (b).

Permeability of Iron Cores.—The permeability of iron is indicated by the relation between flux and magnetizing force, as shown by the ordinary B-H curve. For ordinary iron this relation is given by such a curve as that of Fig. 48(a), of the well-known form. Up to a certain flux density (A in the curve) the flux is approximately proportional to the magnetizing force and we say the permeability is constant. On that part of the curve above A the increase in flux for a given increment in magnetizing force becomes continually smaller and we say the iron is becoming saturated.

In radio circuits we frequently deal with very small currents, involving small magnetizing forces. The only part of the B–H curve of interest then is the region of C (Fig. 48a). But if this part of the B–H curve is obtained carefully it will be found to have the form given in Fig. 48b, in which the point marked C corresponds to the point C of the (a) diagram. From the (b) diagram it is seen that for very weak magnetizing forces the iron acts much differently than it does for those ordinarily used in electrical machinery. As permeability is determined from the ratio of flux to magnetizing force it is apparent that the permeability of the iron at point C (Fig. 48b) is much greater than it is at point D. Whereas the electrical engineer, dealing with iron well up towards point A of Fig. 48, thinks of the permeability of iron as 1000 to 2000, the telephone engineer, using iron at a density of only a few lines per square centimeter or less (point D), thinks of iron as having a permeability between 50 and 100.

Hysteresis Loop.—If the iron is carried through a magnetic cycle, as it is in every a.c. circuit, the B–H relation does not go up and down the curve OAB of Fig. 48, but executes a loop as shown in Fig. 49. The area of this loop represents the work required to carry the iron through its magnetic cycle; it appears as heat in the iron core, and is said to be due to molecular friction.

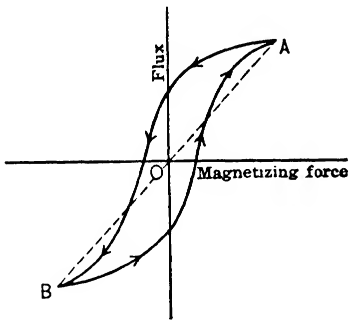


FIG. 49.—The permeability of iron is a variable throughout the a.c. cycle; for alternating currents it is generally considered constant of value determined by the line A–B.

The ratio between flux and m.m.f., as the iron goes through its hysteresis loop, is evidently a variable quantity. At two points in the loop the flux is zero for finite values of the magnetizing force, thus making the permeability (as ordinarily conceived) zero. At two other points the flux has definite values when the magnetizing force is zero, thus suggesting the idea of infinite permeability. Actually the permeability (for the alternating current) is thought of, and ordinarily measured, as the slope of the line AOB of Fig. 49. The slope of this line gives an average permeability and is thus the value used by the transformer designer, for example, in his calculations.

So we say that for alternating currents the permeability is determined by the slope of the hysteresis loop. If, now, we use this idea for various amplitudes of alternating current, we arrive at the conclusion indicated by Fig. 50. For weak m.m.f.s (small currents) the loop A is executed and the slope of this loop is comparatively small. So for weak alternating currents this iron would have low permeability. For larger alternating currents loop B would be executed and the permeability would

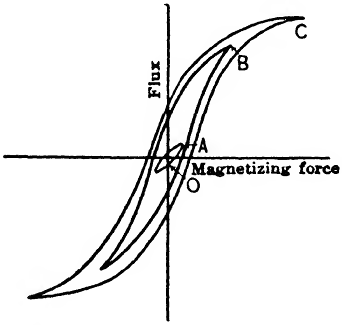


FIG. 50.—The slope of the hysteresis loop varies with the density of magnetization.

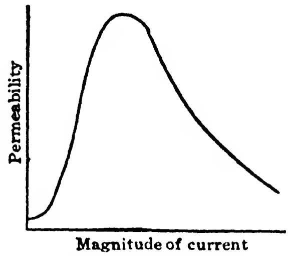


FIG. 51.—The average permeability, determined as in Fig. 49, varies with flux density as shown here.

be higher, while for still larger currents loop C is executed and the permeability is again lower.

Fig. 51 shows the variation in permeability suggested by the three loops of Fig. 50; the range of permeability in the curve of Fig. 51 is from about 100 (for weak currents) to a maximum of 2000, this representing ordinary "electric steel," while for permalloy the maximum value is perhaps 100,000. This high permeability is frequently not available to the radio engineer, however, for reasons given in the next section.

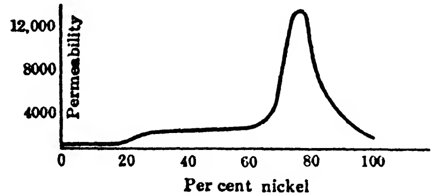


FIG. 52.—Permeability values for nickel-iron alloy, as the proportion of nickel is altered.

Magnetic Action of Alloys.—In recent years much work has been done on the magnetic properties of various alloys, principally those of iron, nickel, and cobalt. Not only does the addition of nickel and cobalt to iron change its permeability tremendously, but it also changes to an equally great degree its hysteresis loss and retentivity. Thus cobalt steel permanent magnets are only a fraction as large as the older tungsten steel magnets, for a given amount of flux, and yet they retain this flux for a much longer time and are more permanent.

The nickel iron alloys show remarkable changes in permeability as the percentage of nickel is changed. The so-called initial permeability (permeability for very small magnetizing force) follows the curve given in Fig. 52.

This diagram shows the remarkable change in permeability which occurs with about $78\frac{1}{2}$ per cent nickel; this alloy is so permeable that even the weak m.m.f. of the earth's field may magnetize it to saturation. It must be remembered that a material may have a very high permeability and yet not have much flux-carrying capacity. Thus in the $78\frac{1}{2}$ per cent nickel alloy (generally called permalloy) the density of flux at saturation is somewhat lower than the maximum flux density possible in ordinary electric steel. But what density the alloy will support is produced by very weak magnetomotive forces. This makes it valuable for loading ocean cables,

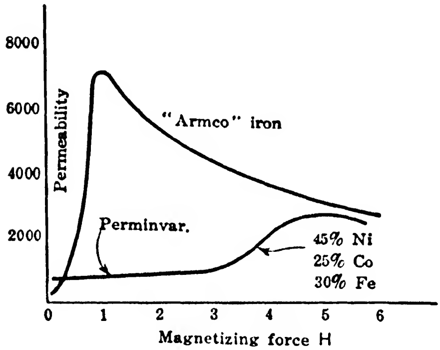


Fig. 53.—Comparison of perminvar and ordinary electric steel, showing constancy of permeability for the nickel-cobalt-iron alloy.

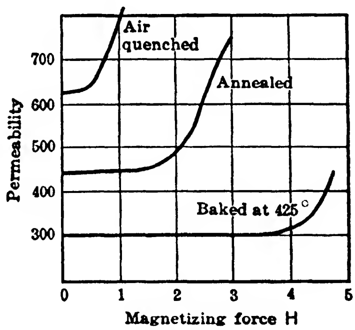


Fig. 54.—Showing the effect of heat treatment on the permeability of perminvar.

where it finds application in the form of thin tape or wire wound spirally over the copper conductor carrying the cable current. These currents are very weak, and the m.m.f. in the magnetic path around the cable is exceedingly small. Yet for just these small m.m.f.s, the permalloy shows its high permeability, hence its application in this service. This special alloy has the further desirable quality of practically no hysteresis loss. Furthermore it shows the interesting quality of no expansion or contraction when magnetized, that is, it shows no magnetostriction. Most magnetic materials change size or shape when they are being magnetized.

Another interesting alloy goes by the name of " perminvar," signifying that its permeability is invariable, or constant, under widely different conditions. In Fig. 53 are shown permeability curves for ordinary electric steel and for perminvar. It is seen that up to about 4 ampere turns per centimeter the ordinary electric steel goes through tremendous permeability changes, whereas the perminvar is practically constant. In Fig. 54 are shown the curves of perminvar permeability for different heat treatments; it is evident that this treatment is vitally important in determining the magnetic characteristics of the alloy.

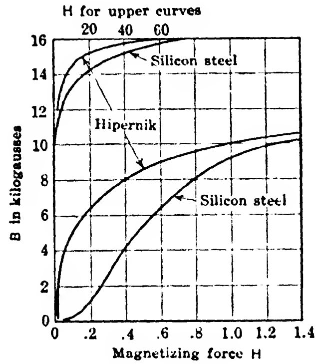


FIG. 55.—Comparison of Hipernik with ordinary electric steel.

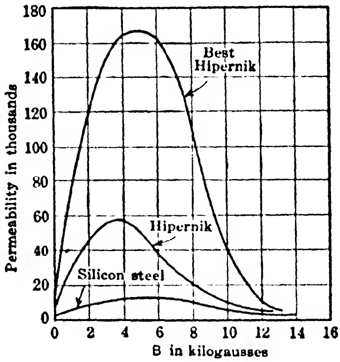


FIG. 56.—Possible values of permeability with all the oxygen removed from the nickel-iron alloy.

The group of nickel iron alloys having between 40 and 60 per cent nickel show remarkable properties after having been subjected to a certain heat treatment. Results obtained up to the present indicate that it is the very small amount of occluded oxygen, which remains after the ordinary purifying processes, that must be removed before the high values of permeability can be obtained. This refining process involves a final annealing of the alloy in an atmosphere of dry hydrogen, at a temperature of 1000°–1300° C. for many hours. This treatment apparently removes the very small residuals of sulphur, carbon, and oxygen. A few hundredths of 1 per cent of these substances acts very unfavorably upon the magnetic properties of the alloy.

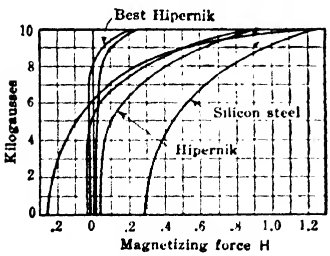


FIG. 57.—Comparison of hysteresis loops of Hipernik and electric steel.

Using a 50 per cent nickel alloy, with the heat treatment above referred to, a B–H curve was obtained as shown in Fig. 55; on the same curve sheet is shown the B–H curve for ordinary silicon sheet steel, the kind used in electrical machinery. These curves bring out the tremendous gain in permeability which this alloy has made possible, at no sacrifice in the obtainable flux density. The alloy, with only 50 per cent iron, will carry practically as much flux as the silicon steel containing 96 per cent iron; of course at the very high densities shown on the curve sheet the permeability for the alloy is no better than it is for the silicon steel.

In Fig. 56 are shown permeability curves for two grades of this special alloy (sometimes styled "Hipernik") as well as that for silicon steel. It is likely that great improvements in these new magnetic materials are still possible.

The shapes of the hysteresis loop of magnetic materials is extremely important to the radio engineer, even more so than the normal permeability curves, because he is generally interested in the a.c. permeability, when a continuous m.m.f. is also acting on the iron. The hysteresis loop gives information on this point better than does the ordinary μ–B curve. In Fig. 57 are shown the shapes of hysteresis loops for silicon steel, and two grades of Hipernik; their use will be discussed in a following section.

Continuous Current and Alternating Current Combined.—In many of its uses the iron-core coil used in radio is subjected to a comparatively large continuous m.m.f. and superimposed on this is a small alternating m.m.f. This is true, for example, in the plate circuit of a vacuum tube and in practically all of the filter circuits used in radio apparatus. The plate current of a vacuum tube amplifier might look as shown in Fig. 58; the continuous component has the amplitude A and the alternating component the amplitude B. How will the iron core act for the alternating component of the plate current?

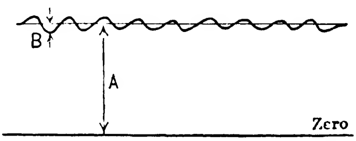


FIG. 58.—The plate current of a vacuum-tube amplifier has the form shown here.

A common misunderstanding of this point is indicated in Fig. 59. Let us suppose the continuous current magnetizes the iron up to the point A. The alternating component will then evidently make the iron execute a small hysteresis loop about point A. Instead of executing the loop C–B, however, the iron actually follows the loop C'–B', the slope of which is only a small fraction of the slope of the loop C–B. Thus the loop C–B might give the iron a permeability of 1500, whereas the actual loop C'–B' gives a permeability of only 500 or less. Furthermore, the slope of the loop C'–B' continually decreases as the point A moves up the B–H curve.

In Fig. 60 there are shown three hysteresis loops for ordinary electric steel as well as the normal B–H curve. Now if a continuous m.m.f. of magnitude Oa is put on the iron, this is presumably brought to point d of its B–H curve; this might be the result of the c.c. component of current in Fig. 58. Now the effect of the ripple in the current of Fig. 58 (that is, the a.c. component) is not to make the flux go up and down the B–H curve at point d, but rather the flux changes around a small hysteresis loop, the slope of which is practically the same as the slope of the top of the hysteresis loop due to an alternating current of m.m.f. equal to Oa. The slope of the small hysteresis loop, for point d on the B–H curve, is that of line A–C, and this determines the effective permeability for the ripple component of the current in Fig. 58.

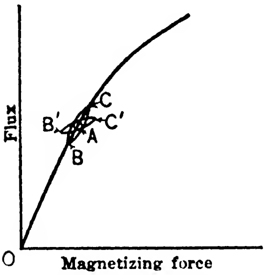


FIG. 59.—The average permeability for the ripple current of Fig. 47 is determined by the slope of the line B'–C', not by the line B–C.

If the c.c. component of Fig. 58 is increased to Ob the ripple will cause the flux to increase and decrease around a small hysteresis loop at point e; the effective permeability is now determined by the slope of line DE. And if the continuous m.m.f. is increased to Oe the ripple current will make the flux go around a small hysteresis loop through point f, having the slope FG. Now these slopes are continually decreasing, as they go through points higher up the B–H curve, and none of them have nearly the value of the slope of the B–H curve itself; they are only a small fraction of this value. Inspection of Fig. 57 shows that even hipernik will not show very high effective permeability if an appreciable continuous m.m.f. is present; the slopes of the tops of the hysteresis loops are by no means as steep as the B–H curves themselves.

Insofar as the author knows, a continuous m.m.f. always decreases the permeability for the alternating current; the one exception occurs when the iron has some remnant flux from a previous magnetization. This case is taken up in the next section.

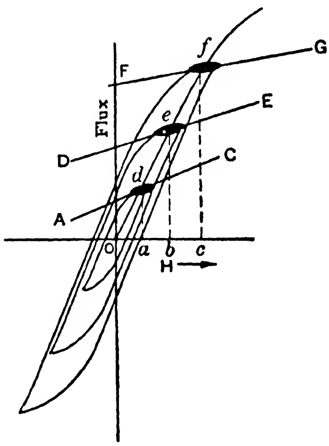


FIG. 60.—This diagram shows how the incremental permeability is affected by a continuous magnetizing force.

Permeability Affected by Previous Magnetization.—It may well be that a test shows an increasing a.c. permeability when a small continuous m.m.f. is impressed; such a case is shown possible in Fig. 61. A transformer core of nickel steel (composition unknown) was tested for a.c. permeability as the continuous current was carried through a whole cycle of values. After being thoroughly demagnetized, measurements of permeability were made for increasing continuous m.m.f. up to 20 ampere turns; this was decreased to zero and then raised in small steps to 20 ampere

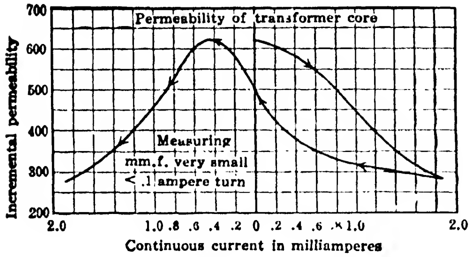


FIG. 61.—Experimental values of incremental permeability; they depend upon the continuous magnetomotive force acting, as shown by this cycle of values.

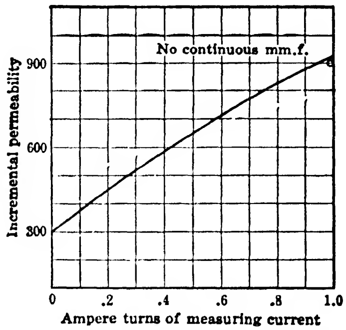


FIG. 62.—A.C. permeability depends upon the strength of the magnetomotive force of the measuring current.

turns in the reversed direction. The measuring current superimposed an alternating m.m.f. of 0.1 ampere turn. It will be seen from the curve that under certain conditions an impressed continuous m.m.f. would raise the a.c. permeability; this probably occurs only when the core has residual magnetization, from previous use.

The same iron core was tested for permeability with increasing a.c. m.m.f. used in the measuring circuit; the results are given in Fig. 62. From this pair of curves (Figs. 61 and 62) it is evident that care must be exercised in measuring permeability; it varies greatly with differing conditions of the test. The amount of the variation depends upon the iron used. Fig. 63 shows the experimental results from two commercial cores used for audio-frequency transformers, undoubtedly of entirely different materials. One had a high percentage of nickel, and the other was an older iron having probably several per cent of silicon. The effect of superimposed continuous m.m.f. in a ring sample of soft iron

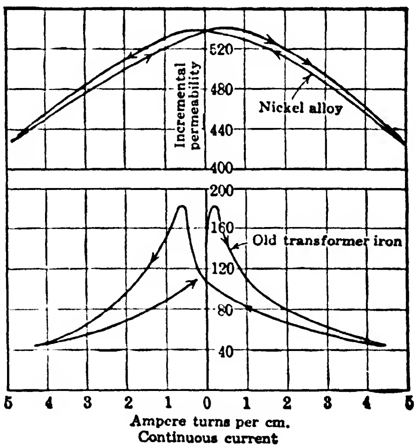


FIG. 63.—Cyclical values of incremental permeability for two kinds of steel.

(no silicon) is shown in Fig. 64; the coil used in this test had such dimensions that the curves for self-induction and permeability are plotted to the same scale.

It is to be noticed that the permeability calculated from results of tests such as this gives only an average value. The inner part of the iron ring offers a considerably shorter path than the outer part, and hence will have much greater flux density. And as permeability depends upon flux density it is evident that the permeability of the inner part of the iron core will be different from that of the outer part. Thus in Fig. 64, the magnetizing

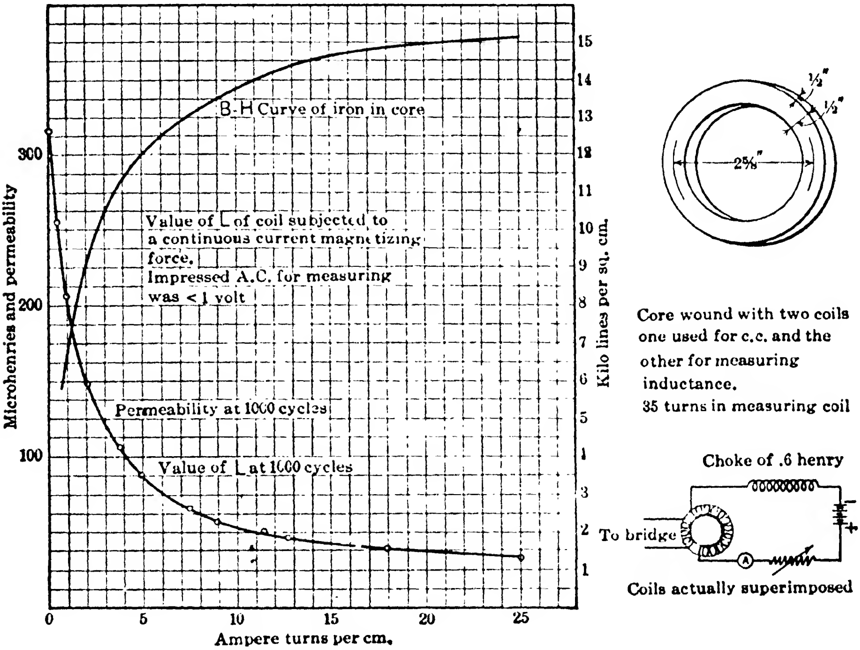


Fig. 64.—Showing how the permeability, for alternating current, changes as the continuous magnetizing force is increased.

force is given in ampere turns per centimeter, using the average diameter of $2\frac{5}{8}$ in. in calculating the length of the magnetic path. But the inner path (diameter $=2\frac{1}{8}$ in.) is 50 per cent less than the outer diameter of $3\frac{1}{8}$ in. Hence the ampere turns per centimeter for the inner path in the iron is 50 per cent greater than that for the outer, and this will result in great difference in flux density, and hence greatly different permeability. To reduce the effect of this path difference the test specimen should have a radial depth which is small compared to the mean radius of the ring.

Turner has reported[1] the results of commercial transformer tests; the

[1] I.R.E., Oct., 1929, p. 1822.

values of self-induction of the primary of an audio-frequency transformer are shown in Figs. 65, 66, and 67. The variations in self-induction which his tests show are in accordance with the analysis previously given.

Effect of an Air Gap in the Magnetic Circuit.—If an air gap is introduced into the magnetic circuit of an iron-core coil we know that the self-

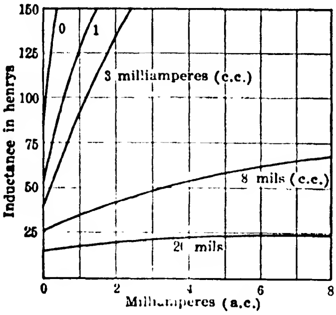


Fig. 65.—Inductance of a transformer core, for different values of a.c. magnetizing force, with various continuous currents flowing in the same winding.

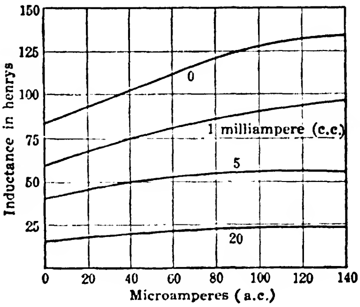


Fig. 66.—Action of the transformer core of Fig. 65, for weak alternating currents.

induction of the coil is diminished. How much it is diminished depends upon the comparative lengths of the air and iron parts of the magnetic circuit, and upon the degree of saturation of the iron. The self-induction which is diminished is that defined by "interlinkages per ampere," and the reason for the decrease in this quantity is the increase in the reluctance of the magnetic circuit.

Now it may well be that the introduction of the air gap actually increases the self-induction of the coil for superimposed alternating current, due to the lower degree of saturation brought about by the continuous current. Putting in the air gap corresponds to diminishing the ampere-turns per centimeter of iron of the c.c. magnetizing force and this, as shown by inspection of Fig. 64, actually raises the effective self-induction for the alternating current.

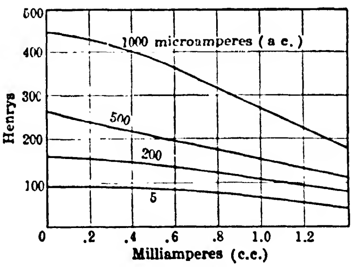


Fig. 67.—Results of Fig. 65 plotted in a different manner.

In Fig. 68 we have shown the effect of air gaps of different lengths on the effective self-induction of a transformer core. It is seen that the effect of the air gap may be very beneficial in some circumstances.

Alternating Currents of Different Frequencies Combined.—It may happen that the iron of a magnetic circuit is subjected to magnetizing forces from two sine waves of current at the same time. In general the frequencies might have any ratio and the two currents be of any relative amplitudes, but the case likely to be of interest in radio is that involving a comparatively large current of low frequency with small currents of frequency perhaps ten times as much. If the low-frequency current is

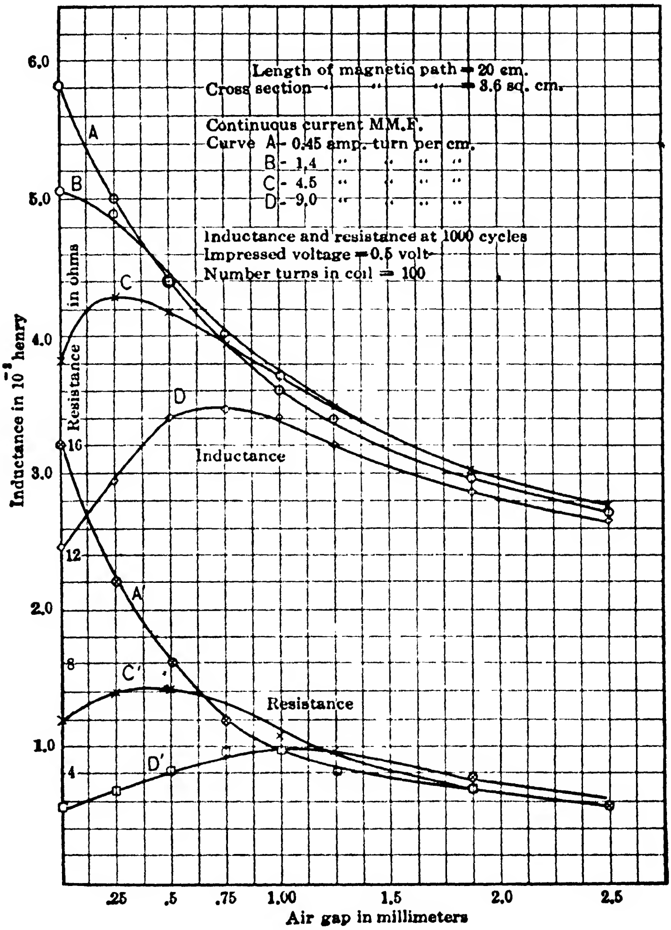


FIG. 68.—Putting an air gap in the magnetic circuit may actually increase the permeability for the alternating component of the current of Fig. 58.

not sufficient to saturate the core, the presence of the small high-frequency m.m.f. acts to increase the effective permeability for the lower frequency, in certain cases as much as four or five times. For larger low-frequency currents the effect of the superimposed high-frequency current may be to decrease the low-frequency permeability. The energy loss per cycle for the low-frequency current follows changes similar to those mentioned for permeability, thus tending to keep the power factor of the coil at the low frequency, independent of the high-frequency current.

The characteristics of the iron for the high-frequency current show very peculiar changes, due to the action of the low-frequency currents. The permeability of the iron for the high-frequency varies periodically with the low-frequency, showing that the iron offers different reluctance and hysteresis for the high-frequency, as it is in different portions of its low-frequency hysteresis loop.

Turner [1] has investigated this effect and a set of results from his paper is given in Fig. 69. The bridge in which his high-frequency measurements

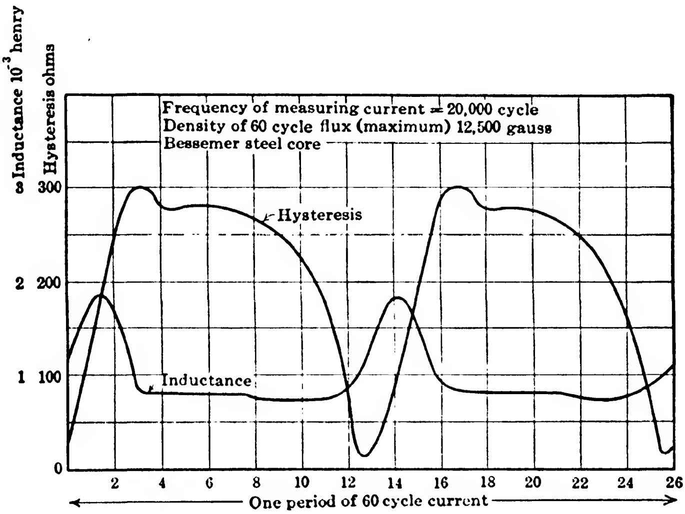


FIG. 69.—The high-frequency constants vary throughout the low-frequency cycle.

were made was connected to the test coil for only a very small fraction of the period of the 60-cycle wave which kept the iron going through its hysteresis cycle. To show how the maxima and minima of hysteresis loss and inductance correspond with the parts of the low-frequency hysteresis loop a polar diagram is given in Fig. 70. On radii from the center of coordinates to the various parts of the hysteresis cycle lengths are taken proportional to hysteresis loss and inductance. The envelopes of these magnitudes give the curves properly marked in Fig. 70. It can be seen from these curves, for example, that minimum hysteresis occurs just before the decreasing 60-cycle m.m.f. gets to zero, and that maximum inductance occurs just a little later in the cycle. The maximum hysteresis

[1] The Physical Review, Jan., 1923.

loss occurs when the 60-cycle flux is changing most rapidly, that is, on the nearly vertical parts of its hysteresis loop.

The alloy known as perminvar, described above, is valuable in service where two or more frequencies flow at the same time because it reduces the likelihood of one channel affecting the performance of another. If, for example, the loading coils of a telephone line were made of iron giving the performance shown in Figs. 69 and 70, it is evident that the carrier telephone channel would be modulated by the currents of voice frequency going over the same wires and these in turn would be modulated by the lower-frequency telegraph currents going over the same wires; these extra modulations are eliminated if the iron has a constant permeability.

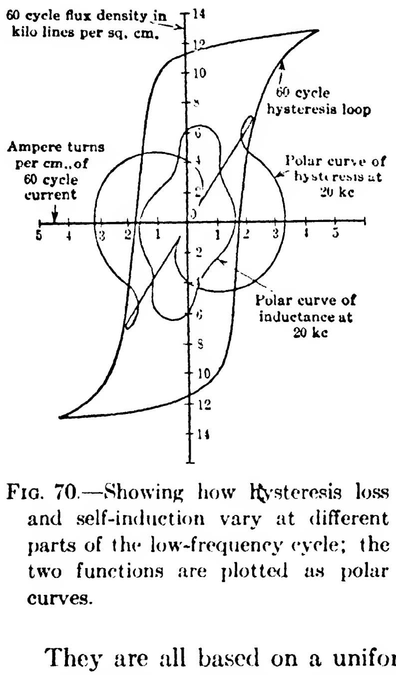


FIG. 70.—Showing how hysteresis loss and self-induction vary at different parts of the low-frequency cycle; the two functions are plotted as polar curves.

Formulas for Inductance only Approximate.—It is evident, therefore, that the L of a circuit may vary with frequency, current amplitude, current distribution, etc., and that its changes can best be predicted by examining in each case the distribution and density of the magnetic field produced by 1 ampere of current in the circuit; the value of L (in henries) of the circuit is equal to twice the number of joules of energy stored in this field. The derivation of the amount of the magnetic energy is difficult and tedious except in the most simple cases; it will not be attempted here, but the formulas themselves for the circuits most generally used in radio work will be given in comparatively simple form, the accuracy being for most cases better than 1 per cent.

They are all based on a uniform distribution of current in the cross-section of the conductor; in case the current is unevenly distributed, due to skin effect, etc., hyperbolic functions are generally required for an accurate formula.

For exact formulas the student should consult the various publications on the subject, notably those of the Bureau of Standards.

Self-induction of a Single Straight Vertical Wire Distant from All Other Conductors.—

$$L = 2l\left(\log\frac{2l}{r} - \frac{3}{4}\right) \text{ centimeters,} \quad . \quad . \quad . \quad . \quad (9)$$

where l = length of wire in centimeters;
r = radius of wire in centimeters;
L = coefficient of self-induction in centimeters;
log = logarithm to the base ϵ, as it is for all the succeeding formulas.

An interesting concept necessarily follows from the form of this equation, namely, that two short pieces of wire placed parallel to each other and end to end have mutual induction; such an idea would probably not be reached from the ordinary viewpoint of mutual induction. Eq. (9) makes the self-induction *per unit length* depend upon the total length of the wire, which means that as the total wire is made longer the self-induction per unit length is increased. This can be true only if there is mutual induction between any one piece of the wire and all others.

Eq. (9) assumes the material of the wire to have a permeability of unity. For uniform current distribution, and permeability differing from unity,

$$L=2l\left(\log\frac{2l}{r}-1+\frac{\mu}{4}\right)\text{ centimeters,} \qquad (10)$$

where μ = the value of the permeability.

For a Single Horizontal Wire.—

$$L=2l\left(\log\frac{2h}{r}+\frac{1}{4}\right)\text{ centimeters,} \qquad (11)$$

where l = length in centimeters;
r = radius of wire in centimeters;
h = height of wire, above earth, in centimeters.

For a Single Circular Turn of Round Wire.—

$$L=4\pi R\left[\left(1+\frac{r^2}{8R^2}\right)\log\frac{8R}{r}+\frac{r^2}{24R^2}-1.75\right]\text{ centimeters,} \qquad (12)$$

where R = radius of turn, to center of conductor, in centimeters;
r = radius of cross-section of conductor.

For a Single Layer Solenoid, Closely Wound.—

$$L=4\pi^2R^2n_1^2lK\text{ centimeter,} \qquad (13)$$

where R = radius of coil, to center of wire, in centimeters;
n_1 = number of turns of wire per centimeter length;
l = length of winding in centimeters;
K = summation of a certain series, which series depends upon the form of the coil. These series have been summed by

H. Nagaoka and are given in Table V. The value of K is given in terms of $\frac{2R}{l}$, i.e., the ratio of the coil diameter to the coil length.

An approximate empirical formula due to Wheeler[1] gives for a single layer solenoid, with turns wound close together,

$$L=\frac{a^2n^2}{9a+10b} \text{ microhenries}, \quad \quad (14)$$

in which a = mean radius of coil, in inches;
b = length of winding, in inches;
n = total number of turns.

Wheeler states that for coils in which b does not exceed 80 per cent of a the formula gives results correct to better than 1 per cent.

TABLE V

Diameter / Length	K	Diameter / Length	K
0.00	1.000	0.95	0.700
.05	0.979	1.00	.688
.10	.959	1.10	.667
.15	.939	1 20	.648
.20	.920	1.40	.611
.25	.902	1.60	.580
.30	.884	1.80	.551
.35	.867	2.00	.526
.40	.850	2 50	.472
.45	.834	3.00	.429
.50	.818	3.50	.394
.55	.803	4.00	.365
.60	.789	4.50	.341
.65	.775	5.00	.320
.70	.761	6 00	.285
.75	.748	7.00	.258
.80	.735	8.00	.237
.85	.723	9.00	.219
.90	.711	10.00	.203

The values of K given in Table V assume a current distribution uniform throughout the conductor, and so give too large a value of L, if, due to skin effect, the current concentrates in the inner side of the winding. The decrease in self induction due to this effect is shown in Table

[1] I.R.E., Oct., 1928, p. 1398.

VI, in which are tabulated the experimentally determined inductances and resistances of the two edgewise-wound ribbon coils referred to on p. 186, and pictured in Fig. 71. It may be found by calculation from the figures given that at high frequencies the current is practically concentrated in the inner side of the coil.

Best Form of Solenoid.—It may be seen from a few calculations, using Eq. (13), that a given amount of wire, to be wound into a single-layer solenoid, should have a certain form if the maximum inductance is to be obtained. This occurs when the diameter is 2.45 times the coil length. The variation of L with departure from this form is not great, however; thus if the ratio is made as low as 1.5 or as high as 4.5 the decrease in L (for fixed length of wire), is only 3 per cent.

TABLE VI

Resistance and Inductance of Edgewise-wound Ribbon Coils

Coil No. 1

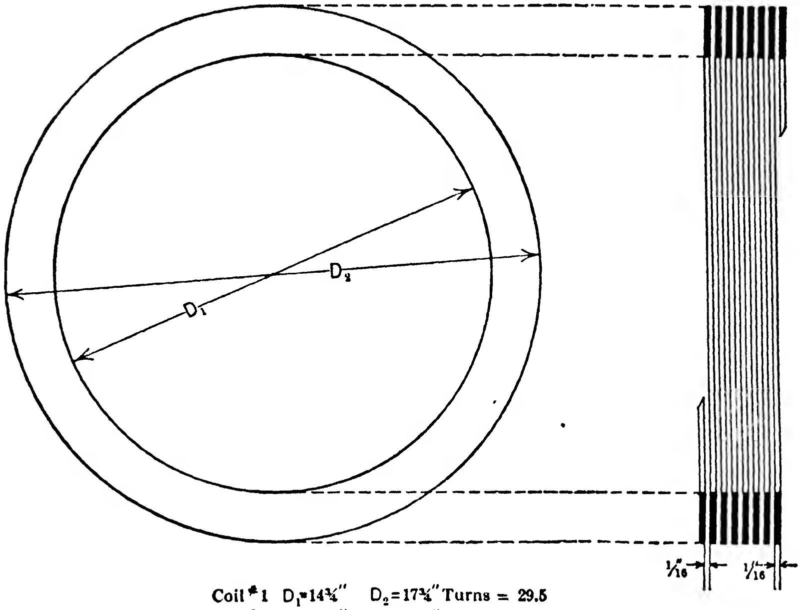
Frequency in 10^{3} cycles	0.043	0.088	0.128	0.248	0.338	0.450	0.730	1.250	3.50
L in 10^{-6} henry	489	485	482	476	472	470	466	464	460
R in 10^{-3} ohm	13	15	19	26	31	36	46	72	176

Frequency in 10^{3} cycles	7.00	16.4	25.2	50.0	75.0	100	125	150
L in 10^{-6} henry	458	455	452	451	454	457	456	460
R in 10^{-3} ohm	295	725	945	1345	1775	2205	2745	3440

Coil No. 2

Frequency in 10^{3} cycles	0.043	0.100	0.150	0.200	0.300	0.400	0.600	1.000	1.60	2.44
L in 10^{-6} henry	613	608	604	602	598	595	592	585	585	583
R in 10^{-3} ohm	23	25	26	30	40	45	49	70	100	145

Frequency in 10^{3} cycles	3.50	6.46	15.3	21.5	50	75	100	125	150
L in 10^{-6} henry	581	578	574	572	570	568	568	570	572
R in 10^{-3} ohm	245	495	1095	1345	2640	2940	3730	5280	7860



Coil #1 $D_1 = 14\frac{3}{4}''$ $D_2 = 17\frac{3}{4}''$ Turns = 29.5
Coil #2 $D_1 = 14\frac{3}{4}''$ $D_2 = 16\frac{1}{4}''$ Turns = 34.5

FIG. 71.—Short solenoid made of edgewise-wound copper ribbon.

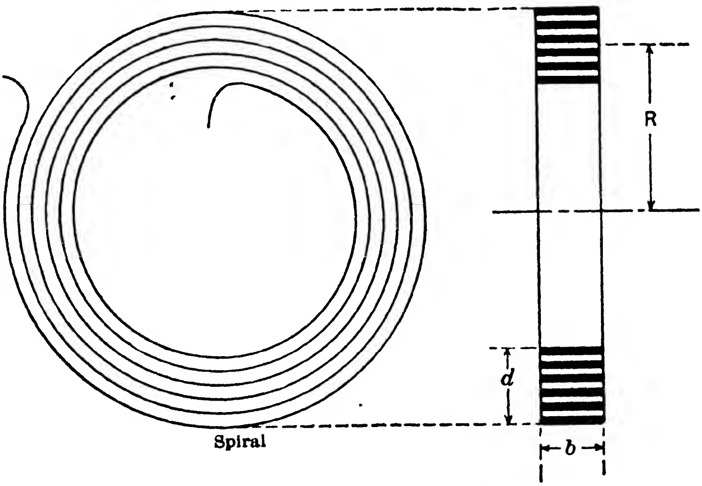


FIG. 72.—Spiral coil of ribbon wound flatwise.

Two-layer solenoids, one layer wound directly on the other, are not feasible for radio work, as the internal capacity is so high. Two-layer coils are sometimes used, however, the turns being arranged in a so-called " banked " winding. Multi-layer coils are, however, preferable, but they must be built in such a way as to keep the internal capacity low, as shown in Fig. 20, p. 186.

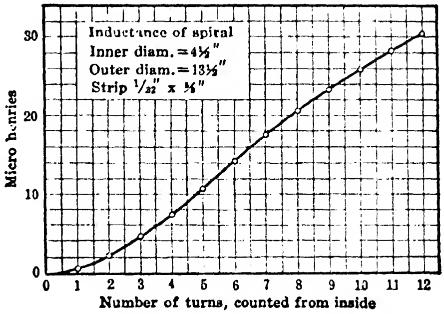


FIG. 73.—Inductance of a spiral similar to that shown in Fig. 72.

Inductances of a Flat Spiral, of Ribbon Conductors, Wound Flatwise, Turns Close Together.—

$$L=4\pi Rn^2\left[\left(1+\frac{3b^2+d^2}{96R^2}\right)\log\frac{8R}{\sqrt{b^2+d^2}}-C_1+\frac{b^2}{16R^2}C_2\right] \text{ centimeters,} \quad (15)$$

where R = mean radius of coil (see Fig. 72);
n = total number of turns;
b = width of strip = axial length of coil;
d = radial depth of coil = outside radius − inside radius.

C_1 and C_2 are constants depending on the shape of the spiral for their values. They are given in Table VII.

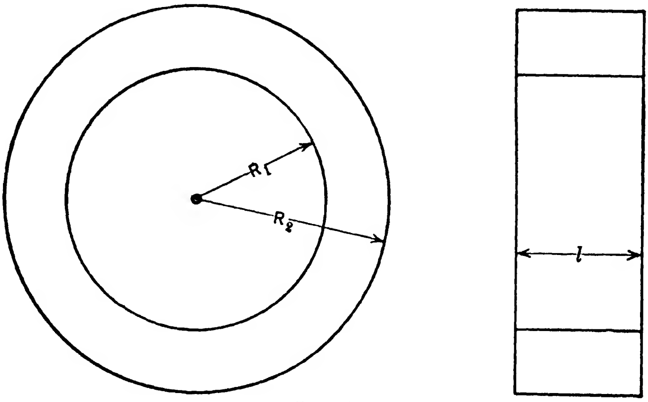
TABLE VII

CONSTANTS C_1 AND C_2 FOR EQ. (15)

Ratio $\frac{b}{d}$	C_1	C_2
0.00	0.500	0.125
.05	.549	.127
.10	.592	.133
.15	.631	.142
.20	.665	.155

Eq. (15) gives incorrect values if the turns are not close together; the values obtained from the equation must be decreased as much as 5 per cent for the spacing used in ordinary transmitting coils in spiral form. Fig. 73 shows how the value of L for a given spiral varies with the number of turns used

It is interesting to note that the same length of wire will give about the same inductance whether wound into a flat spiral or a single-layer solenoid, provided that the mean radius of the spiral has the same value as the radius of the solenoid.

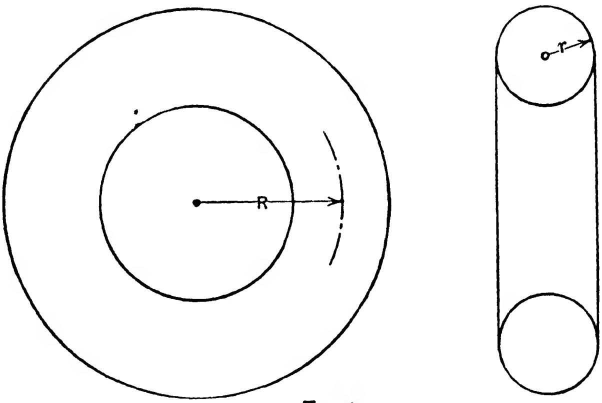


Toroid

FIG. 74.—Toroidal coil of rectangular cross-section.

Toroidal Coil of Rectangular Cross-section. (Fig. 74.)—

$$L = 2n^2 l \log \frac{R_2}{R_1} \text{ centimeter}, \quad \dots\dots \quad (16)$$



Torus

FIG. 75.—Toroidal coil of circular cross-section.

where n = total number of turns;
l = axial length of coil;
R_2 = outer radius;
R_1 = inner radius.

Toroidal Coil of Circular Cross-section (Torus). (Fig. 75.)—

$$L=4\pi n^2(R-\sqrt{R^2-r^2}) \text{ centimeter}, \quad . \quad . \quad . \quad . \quad (17)$$

where n = total number of turns;
R = mean radius of ring;
r = radius of cross-section of winding.

The great advantage of a toroidal coil is that it has practically no external magnetic field and so gives but little mutual induction with other circuits. Also a toroidal coil will, for similar reasons, not be affected by mutual induction from other circuits or sources. Used as a tuning coil in a receiving set it will not pick up any strays or other disturbing fields unless they be of excessively short wave length.

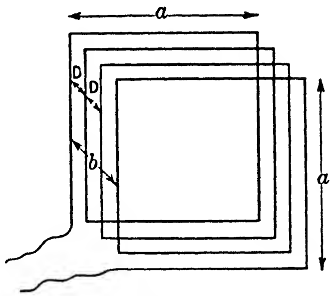


FIG. 76.—Single-layer square coil, such as is used for a coil antenna.

Single-layer Square Coil. (Fig. 76.)—

$$L=8an^2\left[\log\frac{a}{b}+0.726+0.223\frac{b}{a}\right]-8an[A+B] \text{ centimeters}, \quad . \quad (18)$$

in which a = side of square, measured to center of wire;
n = number of turns;
b = axial length of coil = $(n-1)D$;
D = pitch of winding, center to center (to be used in getting A and B from the following tables).

A and B are constants depending upon number of turns, pitch, etc., and are given in Tables VIII and IX, d being the diameter of the wire

TABLE VIII

$\frac{d}{D}$	A	$\frac{d}{D}$	A
1.00	+0.557	0.18	1.16
0.90	.452	.16	1.28
.80	.334		
.70	.200	.14	1.41
.60	+ .046	.12	1.56
.50	− .136	.10	1.75
.40	.356	.08	1.97
.35	.443	.06	2.26
.30	.647	.04	2.66
.25	.830	.02	3.36
.20	1.05		

TABLE IX

Number of Turns, n	B
1	0.000
2	.114
3	.166
4	.197
6	.233
8	.253
10	.266
20	.297
40	.315
60	.322
100	.328

used. Coils wound with rectangular conductor have slightly different constants than those given in the Tables on p. 229.

Flat Square Coil. (Fig. 77.)—

For this case Eq. (18) is applicable providing a is taken as $a_0-(n-1)D$

where a_0 = side of square, outside wire;
n = number of turns;
D = distance between turns, center to center.

The value of b is obtained from the depth of the winding, i.e., it is equal to $(n-1)D$.

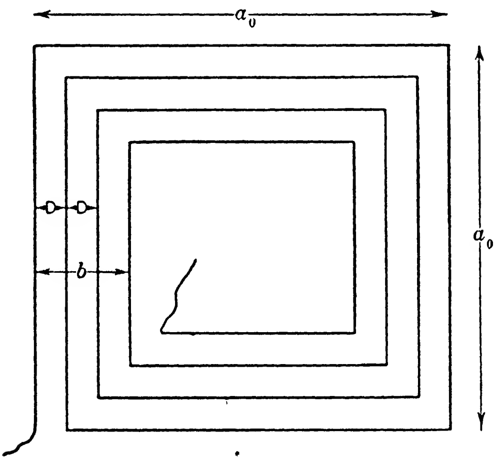


FIG. 77.—Flat square coil, used as coil antenna for short wave-lengths.

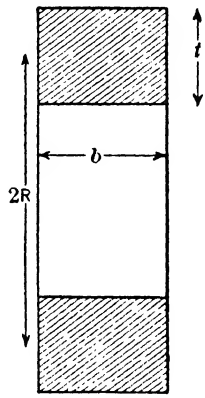


FIG. 78.—Multi-layer coil of rectangular cross-section; the cross-hatched area shows the cross-section of the winding.

Multi-layer Coils of Rectangular Cross-section.[1] (Fig. 78.)—

$$L=\frac{(2\pi Rn)^2}{b+1.5t+R}F'F'' \text{ centimeters}, \quad . \quad . \quad . \quad . \quad . \quad (19)$$

where R = mean radius of coil in centimeters;
n = total number of turns in coil;
b = axial length of coil;
t = radial depth of winding.

[1] An excellent article on the design of multi-layer coils was published in Univ. of California Publications in Engineering, Vol. 147, by F. E. Pernot.

F' and F'' are correction factors

$$F' = \frac{10b + 13t + 2R}{10b + 10.7t + 1.4R}$$

$$F'' = 0.5 \log_{10}\left(100 + \frac{14R + 7t}{2b + 3t}\right)$$

For accurate results with this formula the distance between wires must be small compared to the diameter of the wire.

Hazeltine has devised an empirical formula for multi-layer coils of rectangular cross-section, somewhat simpler than that given in Eq. (19). It is

$$L = \frac{0.8R^2n^2}{6R + 9b + 10t} \text{ microhenries} \quad . \quad . \quad . \quad . \quad . \quad (20)$$

in which n = total number of turns;
R = mean radius, in inches;
b = axial length of coil, in inches;
t = radial depth of coil, in inches.

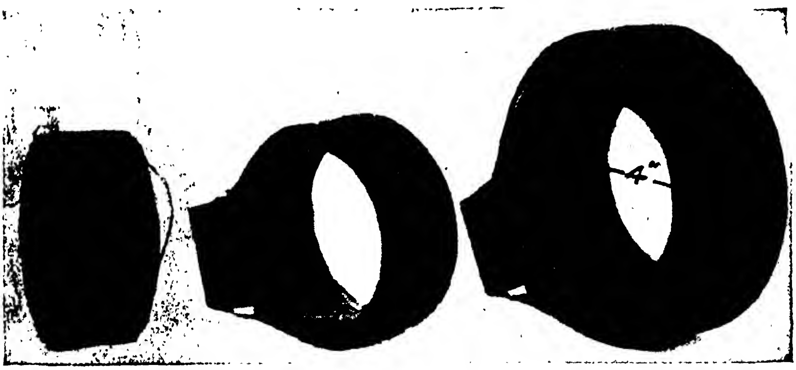


FIG. 79.—"Honeycomb" construction of multi-layer coil.

For coil forms giving approximately equal values for each of the terms of the denominator this formula is said to be good to better than 1 per cent.

A multi-layer coil of very ingenious construction is being made at present, using a so-called honeycomb construction. A picture of such a coil is shown in Fig. 79. The coil is self-supporting, in this respect being superior to the multi-layer coils described on p. 183, and although its internal capacity is greater than that of the type shown in Fig. 20, it is still sufficiently low to make it an excellent coil for radio circuits, especially those calling for many millihenries of inductance. The constants of one of these coils are shown in Table X.

TABLE X

CONSTANTS OF A HONEYCOMB COIL

Frequency in 10^3 Cycles	R in Ohms	L in 10^{-3} Henries	Reactance Divided by Resistance
0	9.39		
26.5	12 4	17.75	238
53	23 8	17.85	250
79.5	53.0	18.70	176
106	102.0	20.25	120

The dimensions of this coil were internal diameter = 5 cm., external diameter = 10 cm., cross-section of winding 2.5 cm. by 2.5 cm.

Most Efficient Form of Coil.—For a given length of wire it will be found that a certain form of coil gives a maximum inductance. Brooks and Turner have compiled a table to bring out the idea and some of their data is given in Table XI. In this table are shown constants of practically any form of coil into which the given length of wire could be wound.

TABLE XI

INDUCTANCE IN MILLIHENRIES OF 52.4 FEET OF MAGNET WIRE 0.1 IN. IN OUTSIDE DIAMETER (SMALL NO. 11 D.C.C.) WOUND INTO CYLINDRICAL COILS OF VARIOUS FORMS AS INDICATED.

(Dimensions are in Inches)

Description	Number of Layers	Total Turns	Mean Radius	Length of Coil	Thickness	Outside Radius	Inductance in Millihenries	Per Cent of Maximum
Spaced solenoid..........	1	80	1.25	16 0	0 1	1.3	0 058	17
Solenoid..................	1	80	1.25	8 0	0 1	1.3	.108	33
Double-layer solenoid......	2	80	1.25	4 0	0 2	1.35	.186	56
Thick tube..............	4	80	1.25	2.0	0.4	1.45	.279	84
Compact................	8	80	1.25	1.0	0 8	1.65	.331	100
Thick disc..............	16	80	1.25	0.5	1.6	2 05	.289	87
Thick disc..............	10	50	2.0	0 5	1.0	2.5	.301	91
Square section ring........	5	25	4.0	0.5	0.5	4.25	.244	74
Flat ring.................	4	8	12.5	0.2	0.4	12.7	.119	36
Thin disc.................	2	2	50.0	0.1	0.2	50.1	.042	13
Single turn..............	1	1	100.0	0.1	0.1	100.0	.025	8

As the same length of wire is used in all these coils their c.c. resistances are all the same. It would seem then that the single-layer solenoid is only 33 per cent as good as a banked coil having eight layers.

However, it must be remembered that the a.c. resistance is generally much greater (at radio frequencies) than the c.c. resistance, and that this increase in resistance is greater as the coil is more compact. This effect is well brought out in Tables II and III, which indicate that a coil of No. 12 wire, single layer, increases its resistance 12 times in the frequency range from zero to 150 kc., whereas a square section coil (10 layers) of the same wire, in the same frequency range, increases its resistance 170 times.

With this idea in mind we must conclude that the ratio of reactance to resistance (which is really the figure of merit for a radio coil) would probably be better for the single-layer solenoid than for the compact coil. This serves well to illustrate the fact that simple apparatus may behave quite differently at radio frequency than it does when continuous current is flowing through it.

Variable Inductances.—It is many times desirable to have a continuously variable inductance for tuning a circuit; two such types have been used, one a long solenoid with a sliding contact and the other a pair of coils connected in series, one rotatable inside the other, an inductance of this type being generally styled a variometer.

The solenoid with sliding contact is not good, because the sliding contact frequently lies on two turns at the same time, thus producing a short-circuited turn, decreasing very appreciably the self-induction from its proper value for the position of the contact, and, due to the current in the short-circuited turn, increasing the effective resistance of the coil. Also there is not much useful variation of inductance obtainable by this method; for long solenoids the value of L increases with the first power of the length only and the coil cannot be used effectively with the contact set to connect in only a small portion of the coil because of the losses occurring in the long unused portion. This part of the coil (generally called a " dead end ") is excited like the secondary of a step-up auto transformer; the charging current circulating in the dead end produces losses and so increases the effective resistance of that part of the coil which is used. Long solenoids intended to be used in steps should be divided up into a number of completely insulated sections, these being connected in series as required.

The variometer type of inductance is very convenient and useful, it being continously variable; the calibration curve of such an inductance is shown in Fig. 80, from which the probable range in inductance can be seen. If the ends of the stationary coil and rotating coil are brought out to separate terminals, the combination forms a very convenient scheme of magnetically coupling two independent circuits, a so-called "coupler."

A rather convenient scheme (even though somewhat inefficient) for making a continuously variable inductance out of a short solenoid, is to fix a copper disc on a shaft inside the solenoid, so that the axis of the disc may be made parallel or not to that of the coil. The eddy currents in the disc, with parallel axes, will very materially reduce the inductance of the solenoid.

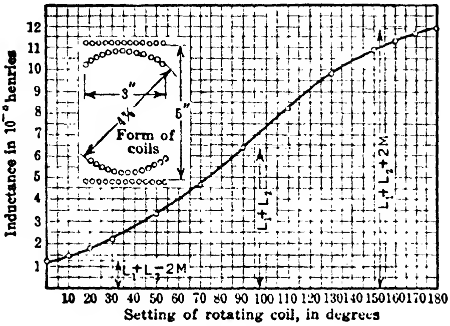


FIG. 80.—Calibration curve of a variable inductance commonly known as a variometer.

The effect of such a solid disc placed inside a short solenoid is given in Table XII. The coil was a single-layer solenoid 12 cm. in diameterandof 2 cm. axial length; the various discs were 11 cm. diameter and were placed inside the coil, centrally, with the plane of the disc perpendicular to the axis of the coil. It is evident that a copper disc very materially affects the inductance without prohibitive increase in resistance. By having the plane of the disc rotatable this scheme of varying the inductance of a coil may be useful, e.g., in heterodyne reception, where but slight changes in inductance are desired (to change signal note), and an increase in resistance of the coil is not of serious consequence.

TABLE XII

EFFECT OF METAL DISC INSIDE SOLENOID

Frequency in Kilocycles	Coil Alone		1/16-In. Copper Disc		1/16-In. Brass		1/64-In. Brass		1/64-In. Tinned Iron	
	L 10^{-3} Henry	R Ohms	L	R	L	R	L	R	L	R
1	1 060	3.03	0.807	3 49	0.895	3 77	1.025	3.53	1.055	3.25
5.35	1.052	3.20	0.762	4.18	0.785	5.13	0.870	6.40	0.970	6.40
50	1.058	3.35	0.751	5.79	0 756	8.13	0.783	12.6	0.865	18.6
149	1.092	5.08	0.760	9.17	0.765	13.5	0.792	17.8	0.861	44.3

This method of varying a self-induction is frequently used in tuning the antenna and other circuits of a short-wave transmitting set.

The best adjustable inductance is a multi-layer coil, with each layer (or every other one after the first four perhaps) separate from the others, equipped with the proper switch to connect in the circuit as many layers as desired.

MUTUAL INDUCTION

Mutual Induction.—The coefficient of mutual induction of two coils may be expressed in terms of energy in the same way as is self-induction. If two coils, so situated with respect to one another that part of the magnetic field of each is linked with the other, are connected electrically in series in such a way that their m.m.f.s add, the total energy associated with the magnetic field of the circuit is $\frac{1}{2}I^2(L_1+L_2+2M)$ and if the electrical connection is reversed, it is $\frac{1}{2}I^2(L_1+L_2-2M)$. Any change in the circuit which changes that portion of the magnetic energy due to M has a corresponding effect on the value of M. The value of M may also be considered as fixed by the voltage induced in one coil by current in the other as given in Eq. (6), Chapter I.

The M of the two coils is determined by their relative position; it may be changed, however, even if the relative position of the two coils stays fixed, if a third circuit is brought into the mutual field of the two coils. Thus two equal coaxial coils, placed with their ends close together may have a value of M about 0.7 as large as L_1, but if a copper sheet is inserted between the two coils the value of M as defined by Eq. (6) of Chapter I may be brought nearly to zero at high frequencies.

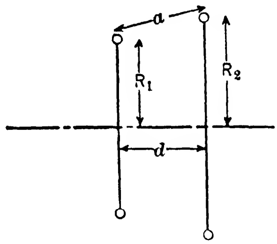


FIG. 81.—Cross-section through two single turns placed coaxially.

The values of M for a few ordinary arrangements are given below, the formulas being approximations as were those for L, the values obtained from the formulas being accurate to better than 1 per cent in most of the cases.

Two Single Turns, Coaxial. (Fig. 81.)—

$$M=4\pi\sqrt{R_1R_2}\left\{\log\frac{8\sqrt{R_1R_2}}{a}\left[1+\frac{3}{16}\alpha-\frac{15}{1024}\alpha^2+\frac{35}{128^2}\alpha^3\ldots\right]\right.$$

$$\left.-\left[2+\frac{1}{16}\alpha-\frac{31}{2048}\alpha^2+\frac{247}{6(128)^2}\alpha^3\ldots\right]\right\}\text{ centimeters.}\quad(21)$$

$$\alpha=a\sqrt{\frac{1}{R_1R_2}}.$$

When the circles have nearly the same radii and the distance between coils is small compared to the radius, the simpler form may be used,

$$M = 4\pi R_1 \left(\log \frac{8R_1}{d} - 2 \right) \text{ centimeters}, \quad . \quad . \quad . \quad . \quad (22)$$

in which R_1 = radius of smaller circle.

Experimental results showing how M varies for the case shown in Fig. 81 are shown in Fig. 82. The coils used were not actually single turns, but the cross-section of the winding was so small compared with the radius of the coil that they approximated single turns geometrically.

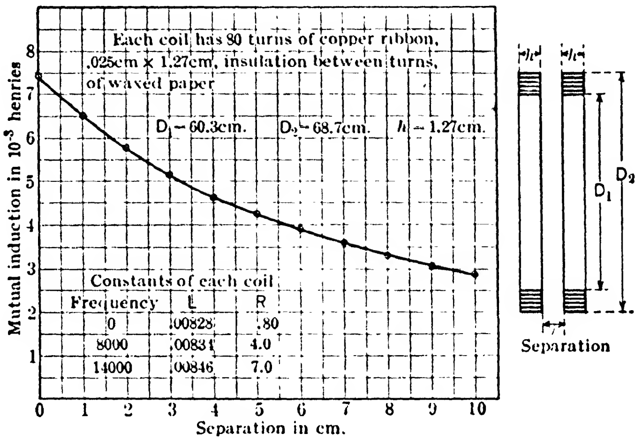


FIG. 82.—Variation in mutual inductance of two coaxial coils with separation; the two coils approximated single turns.

Mutual Induction of Two Coaxial, Circular Coils of Rectangular Cross-section (Fig. 83).—An approximate formula for this case (error for most practical cases less than 1 per cent) is

$$M = N_1 N_2 M_0 \text{ centimeters}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

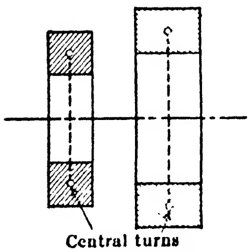


FIG. 83. — Two coaxial multi-layer coils.

where M_0 is the mutual induction between the central turns of the two coils (by Eq. (21)). The curves of Figs. 84 and 85 show the experimentally determined values of M for two typical cases.

Mutual Induction of Two Coaxial Solenoids.—The formulas to cover the various cases which may arise in this class are long; the reader is referred to the Bureau of Standards Bulletin No. 74 for discussion of the case. In Fig. 86 are shown, however, three curves for coils of different dimensions; from these curves M for other shaped coils can be approximated.

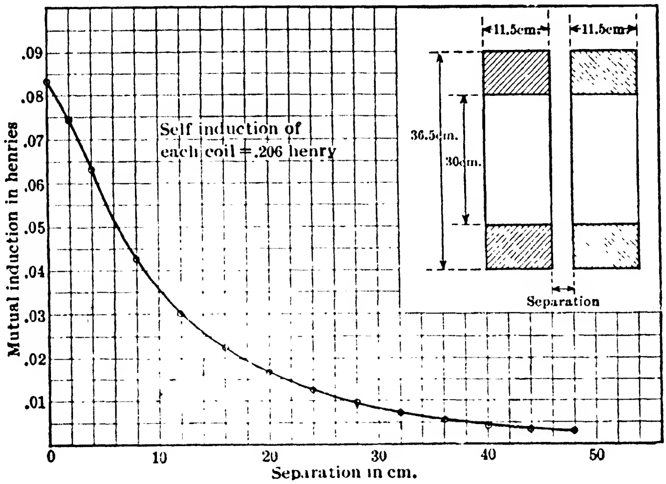


FIG. 84.—Variation of mutual inductance of two multi-layer circular, coaxial coils; separation measured between nearest sides.

Mutual Induction of Two Overhead Parallel Wires, Grounded, at Same Height from Ground.—

$$M = l \log\left(\frac{d^2+4h^2}{d^2}\right) \text{ centimeters}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

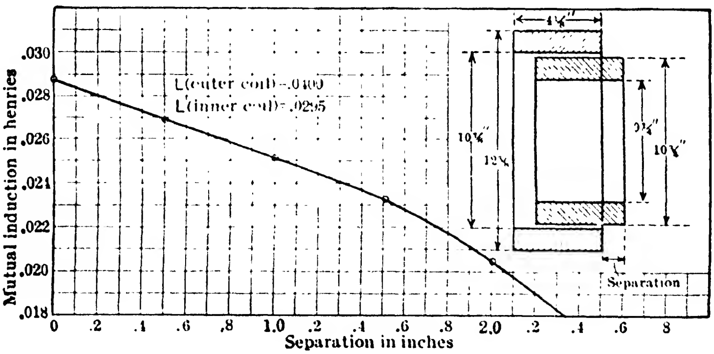


FIG. 85.—Variation in mutual induction of two multi-layer circular, coaxial telescoping coils.

where d = separation of the two wires;
h = height of wires above ground (same units as d);
l = length of one wire in centimeters.

Self-induction of a Two-wire Antenna, Made Up of Two Parallel Wires at Same Height from Ground.—

$$L'=\frac{L+M}{2}, \qquad (25)$$

L' = inductance of the antenna;
L = self-induction of one wire by Eq. (11);
M = mutual induction of the two wires by Eq. (24).

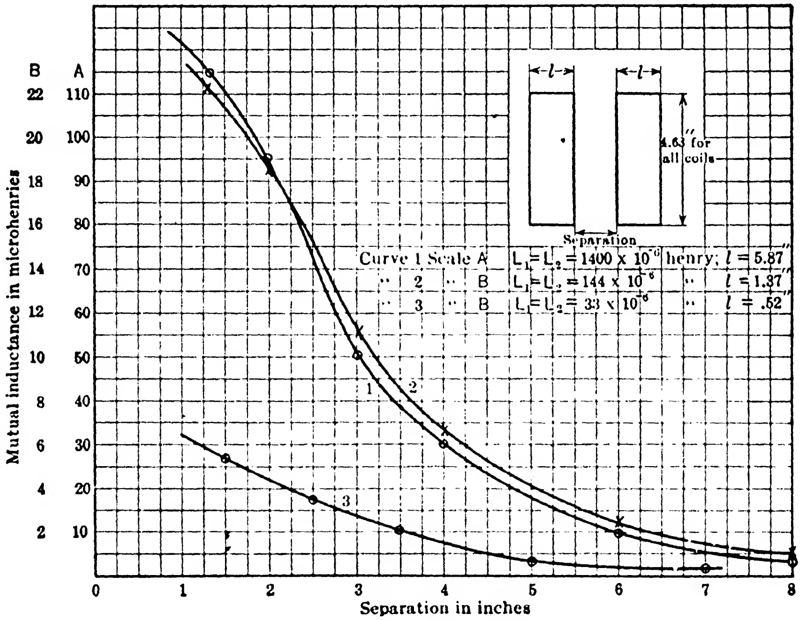


FIG. 86.—Variation in mutual induction of various single-layer solenoids placed coaxially.

Mutual Induction between Two Concentric Coils, as One Rotates (Fig. 87).—This combination of coils is frequently used in radio work, either to make a variable self-inductance or to couple two circuits together. The exact expression for M has not been calculated, but an experimentally determined value of M for a certain combination is shown in Fig. 87. In case the two coils are connected in series the self-induction of the combination is $L_1+L_2\pm2M$. In such variable inductances it is feasible to get a maximum value of L about 12 times as large as the minimum value of L. This range is determined by the manner in which the coils are fitted into one another. When both coils are wound in straight cylindrical form (short solenoids) the range in L will not be as great as when both coils are wound on spherical surfaces, making a closer fit possible. A typical calibration of such an inductance is given in Fig. 80, the form of the coils being shown on the curve sheet.

Mutual Induction between Two Coaxial Spirals.—A tedious calculation is necessary to calculate the value of M for two flat spirals arranged coaxially, but an idea of what may be expected is indicated in the experimentally determined curves of Fig. 88.

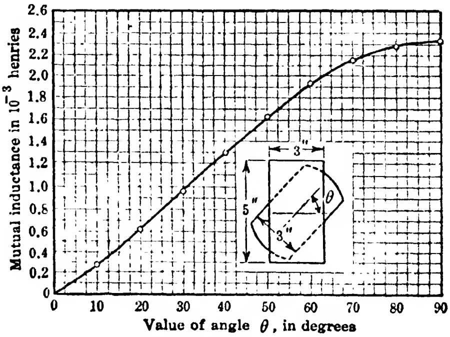


FIG. 87.—Mutual inductance of two coils, one rotating inside the other.

Two-ribbon-wound spirals of the dimensions given on the curve sheet were used; the number of turns in one spiral was fixed at 12, while the sliding contact on the other was used to vary its number of turns as indicated on the curve sheet—the value of M was measured for various separations of the two spirals. The results shown

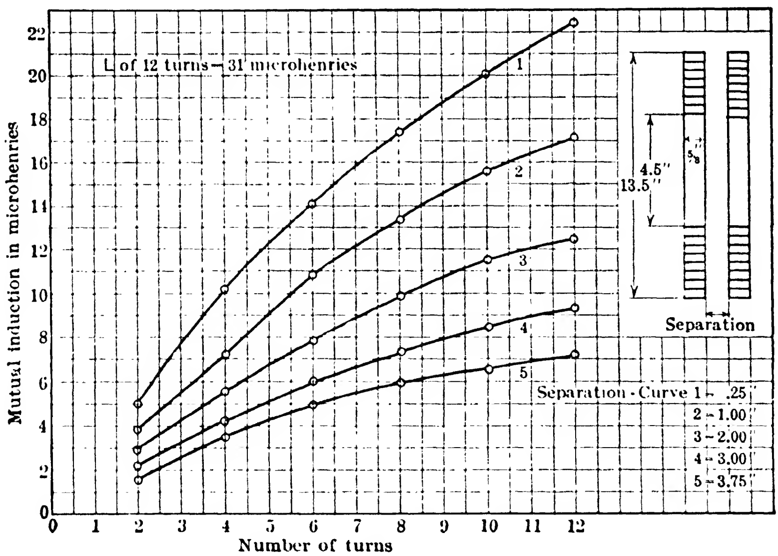


FIG. 88.—Mutual induction of the two flat spiral coils of an oscillation transformer.

in Fig. 82 indicate the higher values of coupling obtainable when the radial depth of the winding of the spiral is smaller.

A Peculiar Case of Mutual Induction.—The fundamental idea of mutual induction is that involving a rate of change of a magnetic field.

If a magnetic circuit set up by one circuit links with another, and the current in the first circuit varies, it follows that the magnetic field varies, and the variation is what gives the e.m.f. of mutual induction in the second. We would judge then that no voltage could be induced in a conductor if it was lying in a place where there is no magnetic field, but in certain cases this idea results in incorrect conclusions.

Let us consider the case of a wire lying axially in a tubular conductor, such as the wire of an ocean cable and its metallic sheath. Let us suppose the sheath is a continuous homogeneous tube, as indicated in Fig. 89. If an alternating current is sent through conductor B, will it induce a voltage in sheath A? Certainly the sheath A will lie in the changing field due to current in B, and this changing field will induce a voltage in A. If the resistance of B is small compared to its reactance, then the voltage induced in the sheath will be in the opposite direction to that impressed on conductor B. That is, if the voltage impressed on B is from left to right (Fig. 89), then the voltage set up in the sheath due to mutual induction, will be from right to left, 180° out of phase, as is the case with the secondary and primary of the ordinary transformer. If the resistance of B is high, the voltage in the sheath will lag only slightly more than 90° behind the voltage impressed on B.

Now if an alternating current is flowing in the sheath, will a voltage be set up in conductor B? If we remember that M is not a one-way characteristic, we must conclude that such a voltage will be set up. However, when we try to visualize the changing magnetic field which induces the voltage in B, we are in trouble.

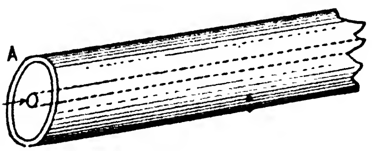


FIG. 89.—Representing a wire in a tubular conductor, such as an ocean cable.

A uniform current (uniform in distribution around the periphery) flowing in a tubular conductor *sets up no magnetic field inside the tube.* If we put a compass needle inside a tubular conductor carrying very large currents, it indicates no magnetic field whatsoever due to this current. So that conductor B is in a region of no magnetic field, yet when the current in A varies it does set up a voltage in this conductor B.

Mutual Induction from Electron Viewpoint.—The above anomaly can be cleared up by getting a more detailed picture of the phenomena of self- and mutual induction. We must endow each electron with a radially distributed electric field, this field being merely a condition in space which tends to make other electrons move in the direction of the lines by which we represent the field.

Now the only force which can make an electron move is that due to the fields of other electrons. A magnet exerts no force on a stationary electron, and it is easy to believe that if a magnetic field itself exerts no force on the electron, then a moving magnetic field can exert no force on an electron, unless this moving magnetic field sets up an electric field and this, it seems, is the fact. An electromagnetic field is an electric field in motion, and nothing but this. When the electrons in a conductor start to drift along a conductor (that is, the conductor is carrying current), then fields move along with them and it is this moving electric field which we denote as the magnetic field surrounding a conductor carrying current.

If an electron accelerates (current in the conductor changes), there is a " kink " sent out on the electron's field, which is depicted in elementary fashion in Fig. 90. In (*a*) the electron is at rest and in (*b*) it is being accelerated to the right. Now any change in the configuration in the field of the electron means a transfer of energy, and energy cannot travel at infinite velocity. Energy, when in the form of an electric field in free space, travels with the velocity of light, and so the kink in the electron's field.

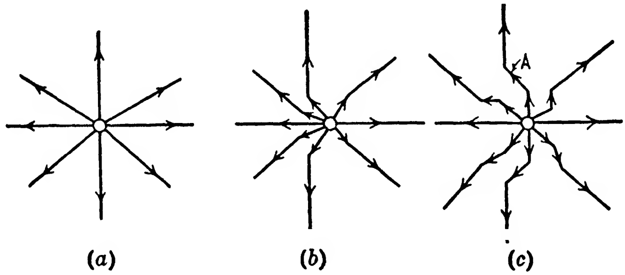


FIG. 90.—When an electron is accelerated a "kink" is sent out in its field.

Fig. 90(*b*), brought about by its acceleration, travels out with the velocity of light. In Fig. 90(*c*) the electron has again been brought to rest, and its field again becomes of uniform radial distribution, but the kink which was produced by the acceleration keeps moving outward with the velocity of light. This energy will travel outward forever as *radiated energy*.

Now we apply this idea of the behavior of an electron's field to the conductor of Fig. 89. In Fig. 91 we have shown a cross-section of the conductor and its sheath with three electrons, two in the sheath and one in the conductor. The two electrons *a* and *b* normally exert no net force on electron *c*, because these forces are equal and opposite.

But at the instant considered these two electrons have been accelerated, from left to right, and a kink has started to travel outwards over their respective fields. These two kinks are shown as having arrived at electron *c*, and it is now evident that at this instant electron *c* will be urged towards the left, and it is this effect, of the distorted electric fields, that we call the e.m.f. of mutual induction.

Granted that a magnetic field is an electric field in motion, it becomes evident from the picture of Fig. 91 why there is no magnetic field inside a tubular conductor. Suppose there is no acceleration, but a uniform motion of electrons a and b. Inside the tube their fields will have no kinks, but their radial fields will be moving slowly from left to right. Each electric field will set up its own magnetic field (we may say), but as the two electric fields are in opposite directions, and moving in the same direction, the two magnetic fields will be just opposite to each other and so neutralize.

Whereas we have considered only two electrons in the tubular conductor all the others may evidently be considered in pairs, so we reach the same conclusion for the whole conductor as we reached above for one pair of electrons.

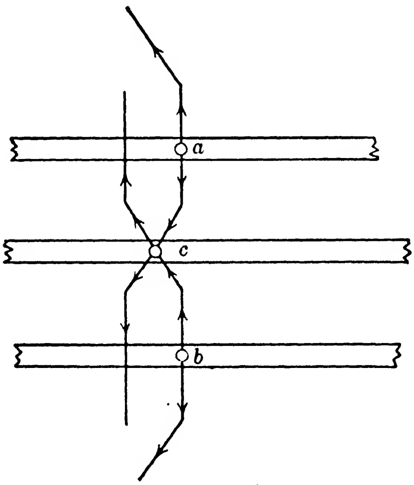


FIG. 91.—The mutual induction of the arrangement of Fig. 89 can be explained with the help of Fig. 90.

Electric and Magnetic Fields Acting Together.—If a student relies upon the foregoing ideas about mutual induction, when making measurements at radio frequencies, he will arrive at some results which seem to be absurd. Thus the arrangement of coils given in Fig. 87 is frequently used for coupling one circuit to another, and such an arrangement gives zero mutual induction when the coils are at right angles. This follows from elementary ideas regarding magnetic fields, as well as such curves as those given in Fig. 87.

Now it will be found in radio measurement that current will be set up in the second circuit when the coupler (as the arrangement of coils is called) *has its coils at right angles;* this unexpected current, however, can be brought to zero by actually turning the rotatable coil past its 90° position, thus giving some mutual induction according to Fig. 87.

This anomalous situation is brought about by the capacitive coupling, which exists between the two circuits because of the proximity of the two coils of the coupler. The capacitive coupling exists even though the two coils are at 90°; to offset the e.m.f. due to this capacitive coupling a reversed magnetic coupling must be introduced. The amount of the reversed mutual induction required varies with the actual form of the coupler coils, but is frequently as high as 10°.

CAPACITY

General Idea of Capacity.—The electrostatic capacity of a body may be thought of either in terms of the quantity of electricity stored for a given potential difference between the two surfaces constituting the condenser or in terms of the energy in the electrostatic field, the value of this capacity, in farads, being equal to twice the energy of the field, measured in joules, when the potential difference is 1 volt.

There may be still another idea of capacity when looking at a circuit from the standpoint of electrical reactions, just as there is for inductance and resistance. When a current flows in a circuit the circuit will generate counter forces called reacting forces or reactions. If the current flowing is 1 ampere the amount of reacting force set up in phase opposition to the current, in volts, is the resistance of the circuit in ohms—the reacting force set up in phase with the current is the negative resistance of the circuit, the reacting force set up 90° behind the current is the inductance reaction in volts, and the reacting force set up 90° ahead of the current is the capacity reaction. The capacity and inductance are calculated from their respective reactances, $\frac{1}{2\pi fC}$ and $2\pi fL$, f being known. The reactions may be caused by ordinary coils, condensers, and wires, but it must be remembered that in special cases a circuit may give capacity reaction when there are no condensers, and it may give inductance reaction when there are no coils present. Thus an overexcited synchronous motor is electrically equivalent to a condenser; a tuned electrostatic telephone when excited at certain frequencies develops (due to its motion) an inductance reaction, and there are no coils used in the telephone.

It must also be remembered that the capacity of a body in general changes with the frequency. Not only does the viscous action of the dielectric decrease the effective specific inductive capacity constant as the frequency is increased, but in many circuits, the capacity of which is under consideration, the potential distribution changes with frequency, and as the electrostatic energy (hence capacity) depends upon the potential distribution, the capacity may be expected to change with frequency.

The formulas given herewith are good only for stationary charges; if the circuit considered is electrically long, the values obtained from these formulas are not correct except at very low frequencies. The capacity calculated from these formulas is in centimeters; to change to micro-microfarads ($\mu\mu f$) the values obtained must be divided by 0.9 and to get milli-microfarads the values must be divided by 900. Where the abbreviation *log* is used the natural logarithm (to base ϵ) is intended.

Capacity of a Conducting, Isolated, Sphere in Air.—

$$C = r \text{ centimeters}, \quad \quad (26)$$

where r = radius of sphere in centimeters.

Capacity of Two Flat, Circular Parallel Plates in Air.—

$$C = \frac{r^2}{4d}\left\{1 + \frac{d}{\pi r}\left(\log \frac{16\pi(d+t)}{d^2} + \frac{t}{d} \log \frac{d+t}{t} + 1\right)\right\} \text{ centimeters}, \quad \quad (27)$$

where r = radius of plates in centimeters;
t = thickness of plates in centimeters;
d = separation of plates in centimeters.

Capacity of Two Flat Plates (Approximate Formula).—

$$C = \frac{KA}{4\pi d} \text{ centimeters}, \quad \quad (28)$$

where K = specific inductive capacity of dielectric;
A = area of one side of one plate in square centimeters;
d = separation of plates in centimeters.

Single Vertical Wire, Proximity to Earth Neglected.—

$$C = \frac{l}{2 \log \frac{l}{r}} \text{ centimeters}, \quad \quad (29)$$

where l = length in centimeters;
r = radius in centimeters.

In several experiments with the lower end of the wire close to the earth, the measured capacity exceeded that calculated from the formula by about 10 per cent.

Single Horizontal Wire, Earth for Other Plate.—

$$C = \frac{l}{2 \log \frac{2h}{r}} \text{ centimeters}, \quad \quad (30)$$

where l = length of wire in centimeters;
h = height of wire above earth;
r = radius of wire, same units as used for h.

This formula assumes that the charge in the wire distributes itself uniformly over the periphery. Actually the lower side of the wire has a slightly greater density of charge than the upper side, resulting in a formula in hyperbolic functions.

$$C=\frac{l}{2\cosh^{-1}\frac{h}{r}} \text{ centimeters. . . . . } \quad (31)$$

When $h/r=5$ Formula (31) gives a result 15 per cent greater than does (30). For greater values of h/r the discrepancy between the two is less. In general, whenever two wires are so close together that the separation is not more than 5 times their diameter, hyperbolic functions are required for precise results, rather than the ordinary logarithmic formulas, for either the inductance or capacity. In practice the ratio of h/r is much greater than 5 except for one or two cases, such as the wires of a telephone cable, etc.

Mutual Capacity of Two Horizontal Wires, Such as Two Wires of an Antenna.—

$$C=l\frac{\log\sqrt{\frac{d^2+4h^2}{d^2}}}{2\left[\left(\log\frac{2h}{r}\right)^2-\left(\log\frac{d^2+4h^2}{d^2}\right)^2\right]} \text{ centimeters, . . } \quad (32)$$

where l = length of one wire;
h = height of each wire;
r = radius of wire;
d = distance between wires.

The mutual capacity is not the same as the capacity of the two wires regarded as the two plates of a condenser, one charged positively while the other is charged negatively. It really represents a decrease in the capacity of one of the wires with respect to earth caused by the presence of the field of the other. In Fig. 92 this point is illustrated; the normal field of wire a to earth is shown by the full lines and that of wire b is shown by dotted lines, and it is evident that the two fields overlap. The total capacity of these two wires, to earth, is diminished to some extent by this overlapping of the two individual fields, and a measure of the decrease in capacity is given by the value of C from Eq. (32).

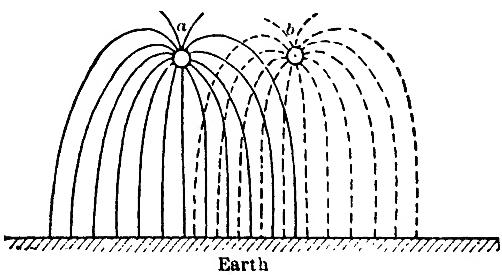


Fig. 92.—Diagram illustrating the "overlapping" of the electric fields of two antenna wires.

Capacity of Two Horizontal Overhead Wires with Respect to Each Other.—This is the case of using the two wires of Fig. 92, one as one side of a condenser and the other one for the other side of the condenser.

$$C=\frac{l}{4\ log\frac{d}{r}} \text{ centimeters, } \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (33)$$

where l = length of one wire in centimeter;
d = separation of the two wires;
r = radius of the wire in same units as d.

This formula supposes the distance between the wires is small compared to the height above the earth; for wires close to the earth, compared to their separation, this formula gives values of C too low.

Capacity of Two-wire Antenna.—This is the case of the two wires of Fig. 92 being connected together and their capacity with respect to earth being determined. It is equal to twice the capacity of one wire with respect to ground (Eq. (30)) diminished by the mutual capacity of the two wires (Eq. (32)).

In case the two wires are far apart the value of capacity is twice that of one wire, and as the wires approach each other the capacity decreases, until when the two wires touch, their combined capacity is not greatly in excess of that of a single wire.

It is interesting to note that the self-induction of a pair of wires (the two wires of an antenna, for example) increases as the wires approach, whereas the capacity of the pair diminishes. In fact the variation is nearly reciprocal, so that the product of L and C of the pair is independent of the spacing of the two wires.

The foregoing formulas for capacity of wires with respect to earth are not very accurate, not being corrected for end effects, etc. It does not seem worth while to use more elaborate formulas, however, because the presence of foreign bodies in the electrostatic fields of antennas, such as trees, masts, stays, etc., influences the value of capacity to a large extent. Also the height of a wire is ambiguous; this height is really to be measured to conducting earth (wet) and the height of the wires above wet earth may not be easy to determine.

Recently Austin [1] has given an empirical formula for the capacity of an antenna, the formula apparently being fairly accurate (say within 10 per cent) for any ordinary form of antenna. It is

$$C=\left(36\sqrt{A}+7.97\frac{A}{h}\right) \text{ centimeters, } \quad . \quad . \quad . \quad (34)$$

[1] Louis W. Austin, "Calculation of Antenna Capacity," Proc. I.R.E., Vol. 8, No. 2.

where A = area of the antenna in square meters;
h = mean height of the antenna, in meters.

In case the length of the antenna is more than eight times the breadth a slight additional correction is necessary, this increase being equal to $\frac{\text{length}}{\text{breadth}} \times 1.4$ per cent.

In calculating A, the length of the antenna is multiplied by its breadth, the area thus obtained being of course much greater than the actual surface of the antenna wires. With the ordinary antenna a spacing of one meter between wires will give a capacity about 90 per cent of that which would be obtained if sufficient wires were used to completely fill the space occupied by the antenna, so that neighboring wires touched each other.

Capacity of a Multiplate Condenser.—

$$C = \frac{KA(n-1)}{4\pi d} \text{ centimeters}, \quad . \quad . \quad . \quad . \quad . \quad (35)$$

where A = area of one side of one plate in square centimeters;
n = total number of plates;
d = separation of plates in centimeters;
K = specific inductive capacity of dielectric.

Various Forms of Variable Condenser.—It is in general more convenient to make a condenser continuously variable than to make an inductance of that kind, hence the tuning of a radio circuit is generally accomplished by using fixed steps of inductances and a continuously variable condenser. These variable condensers are made with either sliding plates, one set of plates moving in grooves in insulating blocks, or with rotating plates, one set of plates being mounted on a shaft.

If the sliding plates are rectangular (and move parallel to one side) or the rotating plates are circular (with shaft on which they rotate in the center), then the amount of capacity in the condenser will vary directly with the amount of movement (sliding or rotation) of the moving plates and the calibration curve will be a straight line. This straight line will not pass through the zero-zero point, because even with zero scale setting there is still an appreciable capacity in the condenser.

It is many times convenient to have the capacity vary with the setting to some other power than the first; thus if it is used in a wave meter it is convenient to have the capacity vary as the square of the setting and the wave-length scale will then be a straight line. For other purposes it is convenient to have a logarithmically varying capacity so that a scale division everywhere represents the same percentage change in capacity.

Both of these variations of capacity are obtainable in rotating plate condensers by properly shaping the rotating plates and suitably placing the shaft in which they turn.

Two typical calibration curves are shown in Fig. 93, for semicircular plates with central shaft, and for specially formed plates, with displaced shaft. In the first the capacity varies directly as the angle of the movable plates and in the second the scale reading is proportional to the logarithm of the capacity.

Special Forms of Variable Condensers.—A variable condenser in which the capacity increases linearly with the displacement of the movable plates, such as the curve marked 1 in Fig. 93, is not of the best form for use in radio circuits. As a laboratory standard it is suitable, but for most radio purposes the capacity of the condenser should vary with the displacement to some power other than the first.

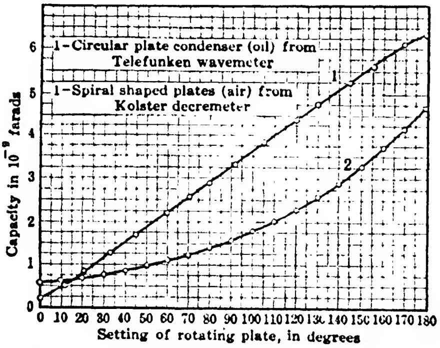


Fig. 93.—Calibration curves of typical condensers used in radio apparatus.

In a special form of wave meter, called a decremeter, it is desirable to have the capacity vary with the displacement in such a fashion that dC/C = constant. Each degree displacement should give the same percentage increase in capacity; this, it will be found, is the equation of the curve marked 2 in Fig. 93. Such a condenser has a logarithmic law for capacity variation.

It may be advisable to have a condenser such that, when associated with a coil, the circuit of which it is a part is resonant to wave lengths proportional to the setting of the condenser. As the wave length varies with the square root of the capacity, this means that the condenser plates must be so formed that the capacity of the circuit varies with the square of the setting. In designing the shape of these plates allowance must be made for the stray capacity of the circuit, capacity between connection wires, inherent capacity of coils, etc. Such a condenser is called a straight-line wave-length condenser.

In another type, of much more importance than those mentioned above, is the straight-line frequency type. The importance of this type follows from certain considerations in radio telephony, which require the spacing of broadcasting sections from each other by equal increments in frequency. This makes advisable a condenser so designed that equal increments in condenser setting advance the frequency, to which the associated circuit is tuned, by equal increments. Such a condenser will tune for the various broadcasting stations at equally spaced points in the condenser dial.

In making these special forms of variable condensers we may use standard rotor plates and specially cut stator plates, or vice versa; or both sets of plates may be especially formed. Certain other possibilities exist, as for example, peculiar shape of air gap between the plates, or a special cam drive between dial and rotor plates of the condenser.

On the basis of ordinary semicircular stator plates and specially formed rotor plates Forbes has shown [1] that the radius vector to the edge of the rotor plates must satisfy the relation

$$r=\sqrt{\frac{4D^2}{nkK^2\left[\frac{D}{K\sqrt{C_0}}-\theta\right]^3}+r_1^2}, \quad . \quad . \quad . \quad . \quad (36)$$

in which $D=\frac{1}{2\pi\sqrt{L}}$;

L = inductance of circuit;

n = number of dielectric spaces;

$k=10^{-11}/36\pi d$;

d = length of air gap between plates;

$K=\frac{f_{0^\circ}-f_{180^\circ}}{\pi}$ = cycles per radian of condenser scale;

f_{0° = frequency of circuit with condenser set at 0°;

f_{180° = frequency of circuit with condenser set at 180°;

C_0 = total capacity of circuit when condenser is set at 0°; this includes stray circuit capacity as well as that of the zero setting of the condenser;

θ = angle of rotation of rotor plates;

r_1 = radius of cut-out of stator plates, to accommodate the rotor shaft.

All dimensions are in centimeters, the inductance of L is taken in henries, the capacity in farads and angles in radians.

The capacity of such a condenser, for any angular position θ of the rotor is

$$C_\theta=\left[\frac{D}{\frac{D}{\sqrt{C_0}}-K\theta}\right]^2-C_0+\frac{nkr_1^2}{2}\theta. \quad . \quad . \quad . \quad . \quad (37)$$

[1] Proceedings I.R.E., Aug., 1925.

In Fig. 94 are shown the approximate forms of the movable plates of three types of variable condensers. *A* is formed properly to increase its capacity linearly with movement of the plate, straight-line capacity (*SLC*). *B* is designed to give such capacity changes that it, in cooperation with an inductance, so tunes its circuit that the relation between condenser setting and circuit wave length is a straight line, giving the straight-line wave

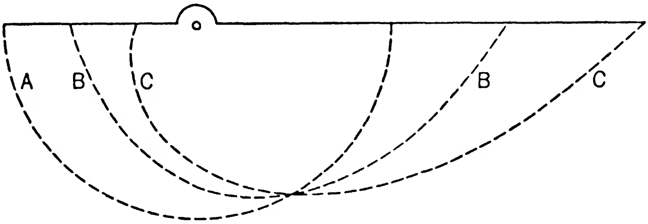


FIG. 94.—These three shapes for the rotating plates of a variable condenser (the stator plates being semicircular) result in such capacity changes as to give variations in capacity (*A*), in circuit wave length (*B*), or circuit frequency (*C*), which are linear with respect to rotation of the plates.

length (*SLW*); form *c*, in cooperation with a suitable coil, gives its circuit such resonance characteristics that the relation between condenser setting and circuit frequency gives a straight line; this is called the straight-line frequency condenser (*SLF*). In the latter two condensers the stray capacity of the circuit, connecting wires, etc., materially affect the form of the plate, especially at the low-capacity end.

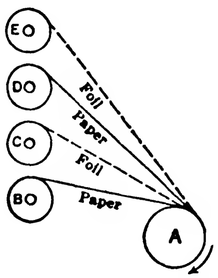


FIG. 95.—The general scheme of making paper condensers.

Condensers of Fixed Capacity.—The ordinary fixed condenser, such as is extensively used in telephone circuits, is made by rolling together two sheets of tinfoil and two sheets of paper, as indicated in Fig. 95. Rolls of tinfoil and paper, *B*, *C*, *D*, and *E*, are fed together onto roll *A*; when the proper length has been wound the paper and tinfoil are cut and the loosely rolled pack is slid off from roller *A*. Suitable provision for making connection to the tinfoil sheets is made while the sheets are rolling on to *A*. The loosely wound condenser is put into a tank which can be heated and evacuated; this process takes out most of the moisture and air from the condenser. Hot paraffin wax is admitted at the bottom of the hot, evacuated tank, submerging the condensers, and then pressure is let into the tank. The hot wax is thus made to completely permeate the condenser, which is then taken out, compressed into its proper shape, and allowed to cool. If the wax contracts, and permits the formation of minute cracks during the cooling process (as some waxes do), the condenser will be spoiled, as air and moisture will re-enter the paper from which they have just been expelled. The ends of the cooling condensers are sometimes covered with a layer of soft non-cracking wax to prevent this. Sometimes a good grade of non-hygroscopic oil is used for impregnation, instead of wax.

A paraffin-impregnated condenser, using a good grade of paper about one-half of one mil thick, will stand from 100 to 200 volts, for each layer of paper used. Thus to operate as a filter condenser for a 350-volt " power pack " a three- or four-paper condenser should be used. If there is not much fluctuation in the voltage, less thickness of paper is required. The condensers in a power pack close to the rectifier receive higher c.c. voltage, as well as a greater fluctuation, and therefore should have a greater dielectric strength.

A well-impregnated condenser will show an insulation resistance of 100 to 1000 megohms per microfarad (insulation resistance varies inversely as the capacity), and its current will lead the impressed voltage by more than 89°, showing a power factor of 1 per cent or less.

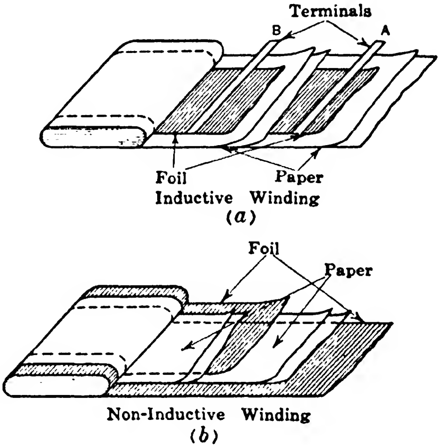


FIG. 96.—The connectors shown in (a) serve well at low frequencies but for radio frequencies the connections are made all along the edge of each tinfoil strip.

In case this type of condenser is to be used as a radio-frequency " by pass," its connections must be fashioned differently from those in condensers used only at low frequencies. Fig. 96, diagram *a* shows the normal connection scheme for a condenser to be used at telephone frequencies. The two tinned copper strips *A* and *B* are laid on the tinfoil sheets when the roll is being made; connection is made to their ends, which generally project through the end of the can in which the condenser is sealed. The charging current of such a condenser must run along the whole length of the tinfoil strips, and, at radio frequency, such a condenser may show an unexpectedly high resistance because of this effect.

By winding roll *A* (Fig. 95) with each tinfoil strip projecting past one edge of the paper, as shown in Fig. 96*b*, the complete length of each tinfoil strip is made available for connection. A low-melting-point solder serves to make a good electrical connection between the projecting tinfoil sheets and the condenser terminal. In such a condenser the maximum distance the current has to flow in the tinfoil sheet is its width, only 3 or 4 in. The result is that such a condenser shows but little effect of sheet resistance even at radio frequencies, and a condenser of such construction is always chosen for use as a by pass for radio-frequency currents.

Losses Occurring in Condensers.—When a condenser is used in a high-frequency circuit there are various losses which occur, all of which are detrimental and to be avoided if possible. The losses may be due to actual leakage from one plate to the other through or around the dielectric, to dielectric hysteresis, to I^2R loss in the conducting plates of the condensers, and due to corona losses from the edges of the plates. When a condenser is used in a receiving circuit (low voltage) only the first three sources of loss exist. All the data available go to prove that the losses in solid dielectrics vary with the first power of the frequency and with the square of the impressed voltage.

For an air condenser constructed with rugged plates of negligible resistance all of the losses in the condenser proper are negligible except at voltages high enough to give corona loss. However, the supports, terminals, etc., of the air condenser must be mounted on very good insulators, otherwise an appreciable resistance may be incurred due to the dielectric hysteresis and leakage at these points. Quartz or high-grade porcelain should be used at these points.

Condensers using glass, paper, rubber, or mica for the dielectric have some dielectric losses, although this loss in a well-constructed mica condenser (air and moisture excluded) seems to be very small; the dielectric losses in paper and some grades of glass are high. Dry oil is in general a very good dielectric with very low losses; the oil has an added advantage over a solid dielectric, in that a disruptive breakdown does not spoil the condenser, the oil repairing itself with no deleterious effects unless sufficient arcing occurs in the oil to produce considerable carbonization. A good grade of mineral oil is generally used, but the author has found castor oil to be excellent, having a high dielectric strength, low losses, and having such a high specific inductive capacity as to give about twice as much capacity as the same amount of mineral oil. The value of K for various dielectrics is shown in Table XIII.

It must be remembered when using solid dielectric condensers that practically all such materials as glass, rubber, paper, wax, etc., very rapidly lose their insulating properties as the temperature increases. In fact the operation of most solid dielectric condensers becomes unstable above a certain voltage; above this critical voltage the condenser will soon break down if left connected to the circuit. This is due to the cumulative effect of the losses in causing temperature rise, the higher the temperature the higher the losses become, thus again increasing the temperature. Some special paper condensers passed a puncture test of 4000 volts successfully,

TABLE XIII

SPECIFIC INDUCTIVE CAPACITY OF MATERIALS USED MORE GENERALLY IN RADIO CONDENSERS (MEASUREMENTS AT LOW FREQUENCY)

Material	Value of K	Material	Value of K
Ebonite	2.5–3.5	Resin	2.50
Ebonite	1.9 at about	Shellac	3.0–3.7
	4×10^7 cycles	Castor oil	4.7
Glass, density:		Olive oil	3.1
2.5–4.5	5.0–10	Petroleum oil	2.1
Glass, density:		Vaseline	2.2
2.5–4 5	2 7 at about	Isolantite	3.6
	5×10^7 cycles	Pure water:	
Gutta percha	3.3–4.9	1° C.	83.0
Mica	4 6–8.0	95° C.	57.0
Paraffin wax	2 0 2.5	Linear relation for	
Porcelain	4 38	intermediate tem-	
Quartz	4 50	peratures.	

Formica, Bakelite, Bakelite-dilecto, and such compounds have a value of K of about 5, generally lying between 5 and 6. See pages 266 and 267 for more detailed information on these materials.

but upon being connected to a 2000 volts 60-cycle line every one out of the twelve tested broke down in less than twenty minutes. The same experience was had to an even more marked degree with a lot of mica condensers.

It must also be remembered that a condenser passing a voltage test successfully when tested in a 60-cycle line may break down after a few minutes' operation on a high-frequency circuit with a voltage only a small fraction of the 60-cycle voltage which it withstood successfully. In certain parts of a vacuum tube the glass (as dielectric) is subjected to high-potential gradients at high frequency and it shows losses many times as great as might be expected from low-frequency tests.

Even quartz, which is one of the best dielectrics, shows this effect; a certain piece required 46,000 volts to puncture when the voltage was continuous, whereas it broke down at 18,000 volts (effective) after being connected to a 500-cycle line for a few minutes.

Volt-ampere Rating of a Condenser.—At low frequencies a condenser has a low volt-ampere rating because of its voltage limitation; not more than a certain voltage can be applied without danger of puncturing the insulation. At high frequencies the safe heating of the dielectric is the factor that limits the capacity of the condenser. These two limitations result in a variable K.V.A. rating of a condenser. A mica condenser (consisting of ten sections connected in series) of 0.001 μf capacity has a safe rating as follows: [1]

Frequency	Safe Current	Safe Voltage	Safe K.V.A.
0.1 kc.	1.9 amps.	3000 volts	5 7
0.3	5.65	3000	17 0
1.0	11.2	1780	20 0
3.0	11.4	605	8 9
10.0	11.2	178	2.0

Equivalent Series or Shunt Resistance of a Condenser.—All of the losses in a condenser can be grouped together and represented by a certain hypothetical series resistance; the value of this resistance will, in general, be different for different frequencies. Thus in Fig. 97 let a represent a

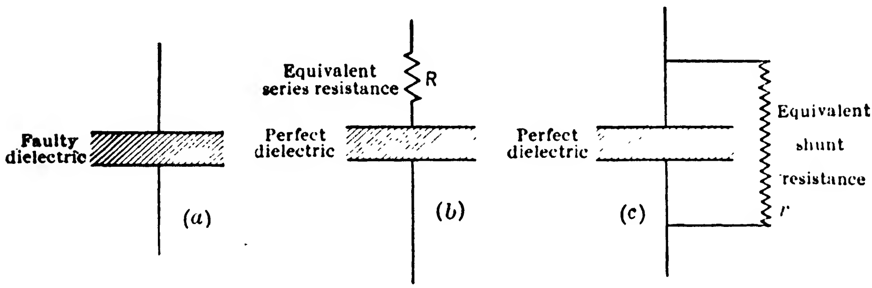


FIG. 97.—A condenser with imperfect dielectric may be represented by one having perfect dielectric in series with, or shunted by, a suitable resistance.

condenser which is drawing a current of 2 amperes at a certain frequency, and has a total power loss due to all causes of 7.5 watts. Then this faulty condenser can be well enough represented by a perfect condenser b (of the same capacity as a) having no losses, but having in series with itself a non-inductive resistance R, such that the charging current of condenser b, flowing through this resistance, will dissipate the same amount of power as is lost in the faulty condenser a. For the case cited above we shall have $R = \frac{7.5}{2^2} = 1.88$ ohms. The faulty condenser might also be replaced by a perfect condenser and a suitable leak, or shunt, resistance. If the voltage in the condenser is E, the loss in a shunt resistance is E^2/r, so in c, Fig. 97, is shown this arrangement and for the case cited, if the

[1] Mica condensers in high-frequency circuits. Maloff, I.R.E., April, 1932, p. 647.

voltage is 5000 volts the proper shunt resistance is obtained by putting $\frac{5000^2}{r}=7.5$ or $r=\frac{5000^2}{7.5}=3.34\times 10^6$ ohms.

These simple calculations hold accurately only for condensers of low power factor, say 0.02 or less, but as all good radio condensers have a lower power factor than this the method outlined above is accurate enough. The relation between the equivalent series resistance and equivalent shunt resistance is obtained from the relations (which, it must be remembered, hold good for low power factor condensers only)

$$I^2R=\omega^2C^2E^2R=\frac{E^2}{r},$$

or

$$r=\frac{1}{\omega^2C^2R}=\frac{X_c^2}{R}. \qquad (38)$$

R being the series resistance and r the shunt resistance.

A practical case where this change of shunt resistance to series resistance is of value is shown in Fig. 98. This gives part of the circuit of a radio-frequency amplifier, the frequency being fixed by the coil L and condenser C. Across the condenser C is connected the input circuit of the vacuum tube, which may have a resistance of 100,000 ohms. The question arises—will this leakage resistance of 100,000 ohms have a material effect on the selectivity of the circuit L-C? At the low-frequency setting C is equal to 0.0005 microfarad and at the high-frequency setting it is about 0.00005 microfarad. At the low frequency (600 kc.) the resistance of the coil L will be about 5 ohms and at the high frequency (1900 kc.) it will be about 20 ohms.

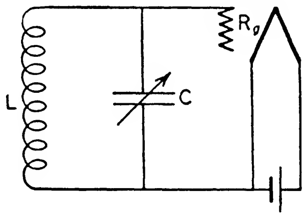


FIG. 98.—In the ordinary vacuum-tube amplifier the tuning condenser C has a leak resistance, R_g, in shunt with it.

Using Eq. (38) we find that for the 600-kc. frequency the 100,000 ohms shunt resistance is equivalent to 2.8 ohms series resistance and at the high frequency it is equivalent to 27 ohms series resistance. At the low frequency therefore the vacuum tube input circuit increases the width of the resonance curve by about 50 per cent and at the high frequency it makes the resonance curve more than twice its normal width. By using a " C " battery the resistance of the input circuit may be raised to about 500,000 ohms, which value is sufficiently high to produce but little effect on the resonance quality of the L-C circuit.

It will be noted that this interchange of series and shunt resistances involves certain premises. Thus the series resistance must be so low compared to the reactance of the condenser that its effect on the impedance of the condenser circuit is negligible—the shunt resistance must be so high that the current through it is so small as to make the line current (condenser current+leak current) essentially the same as the condenser current.

The equivalence of series and shunt resistance is based on equal power consumptions in the two arrangements, as evidenced by the derivation of Eq. (38); it may be that, for other considerations the two arrangements do not act alike. A case of this kind will be mentioned in Chapter III, in discussing the effect of condenser leakage on an oscillatory discharge.

For most dielectrics the equivalent series resistance varies inversely with the frequency, indicating a constant energy loss per cycle.

Electrolytic Condensers.—If two aluminum plates are placed in a suitable solution, and connected to a suitable source of c.c. power, an aluminum oxide forms on the positive plate which will gradually act to insulate this plate from the solution, thus reducing the current flow to practically zero. The thickness of the oxide film formed, and the time taken to form it, depend upon how high a voltage is impressed on the two plates; it may be only 0.00001 mm. for low formation voltages (say 20 volts), and it may reach a thickness of 0.001 mm. if the forming voltage is raised to 400. If a voltage much higher than this is used the film apparently cannot be formed sufficiently thick to prevent breakdown; the film will spark through and permit large current to flow, but if the voltage is reduced to about 400 the ruptured spot in the film will be repaired in a short time.

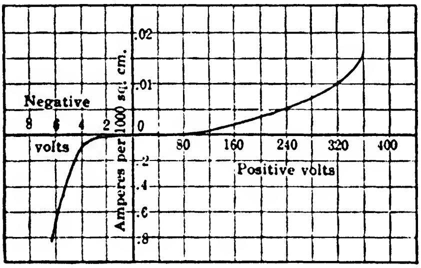


FIG. 99.—A pair of aluminum plates in ammonium borate solution, shows asymmetrical conductivity; notice the different scales for positive and negative voltage.

The electrical behavior of a pair of aluminum plates, in a solution of ammonium borate, is shown in Fig. 99. The oxide film on the positive plate had been formed by gradually increasing the impressed voltage up to 350 volts; the thickness of the film was about 0.001 mm. The film having been thus formed, the voltage was dropped to zero and gradually raised, and corresponding values of current read. Up to about 40 volts the current was not readable on the meter; above this voltage the current increased gradually until 350 volts was approached, and here it rose more rapidly. At about 460 volts the film began to break down and pass current measured in amperes. It will be noticed that for reversed voltage large currents flowed even for low voltage; the negative voltage scale is greatly magnified. From this record it appears that the film is also a rectifier, and it has been so used in many of the battery chargers sold with radio sets.

The capacity per unit area of film is very high, primarily because of the small thickness of the film. A film formed to 30 volts is probably about 0.00001 mm. thick, and it shows a capacity of 0.2 microfarad per square centimeter. If the film is thickened by being formed to higher voltages its capacity per unit area goes down proportionately; a film formed to 350 volts has about 0.018 microfarad per square centimeter of area.

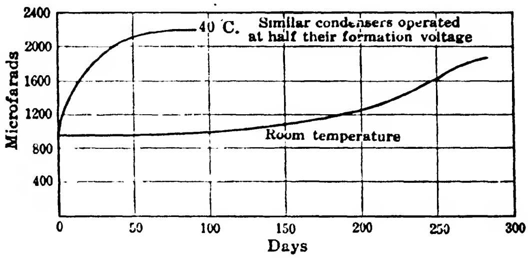


FIG. 100.—Change in capacity of electrolytic condensers, formed to 60 volts and then operated at 30 volts; with higher temperature the capacity increases more rapidly.

The oxide film gradually disintegrates if the cell is left idle; it should remain connected to a c.c. line of suitable voltage. There will be a small polarizing current flowing continually, maintaining the film in its proper condition. The amount of current varies somewhat with conditions but should not be more than a few microamperes per microfarad after the condenser has been connected to the line a week or so. At 300 volts impressed it might be 50 microamperes per microfarad, and at 400 volts perhaps 200 microamperes.

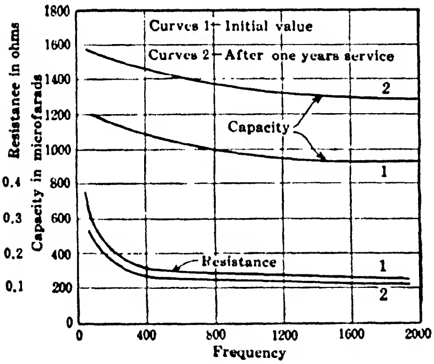


FIG. 101.—The capacity of an electrolytic condenser goes down as the frequency is raised, probably due to the resistance of the electrolyte.

If a condenser has been formed at a high voltage and is then operated at a lower one the capacity will increase with time, owing to the fact that not sufficient leakage current will flow to maintain the film. The film will decrease in thickness (and so leakage current increase) until the amount of leakage current that flows is just sufficient to compensate for the disintegration of the film. The change in film thickness, and hence in condenser capacity, takes place more rapidly if the electrolyte is warmed; in Fig. 100 is shown the change in capacity of two condensers formed to 60 volts and then operated at 30 volts. These results, as well as some of the others given here, were obtained in the Bell Telephone Laboratories.

These ammonium borate-aluminum condensers have a high capacity per unit volume, compared with those made of paraffin paper. The capacity *per unit area* is nearly 1000 times as great. However, the losses are much greater; whereas the power factor of the paper condenser may be 1 per cent, that of the electrolytic condenser may be from 10 to as much as 80 per cent! Fortunately, in their principal service, this high loss is of no great importance. This type of condenser is used practically always in filter circuits, to eliminate the ripples from a unidirectional pulsating power supply. Here the undesired high-frequency currents flow into the electrolytic condenser and their energy is used up as condenser loss; this is the effect desired of the filter.

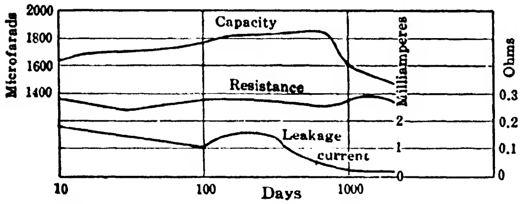


FIG. 102.—Performance curves of a normal electrolytic condenser.

The principal loss is due to the resistance of the electrolyte, which varies from 50 to 300 ohms per cubic centimeter according to concentration. The anode is made of deeply corrugated aluminum sheet, and at high frequencies (where the high power factor occurs) the bottom of the grooves is not very effective because the current has to flow through too long a path of solution. The decrease in capacity as frequency is increased is probably accounted for by this effect. In Fig. 101 are shown curves of effective series resistance and capacity of a condenser in which the film had been formed to 46 volts and the condenser operated at 28 volts. Of course the ripple voltage is always small

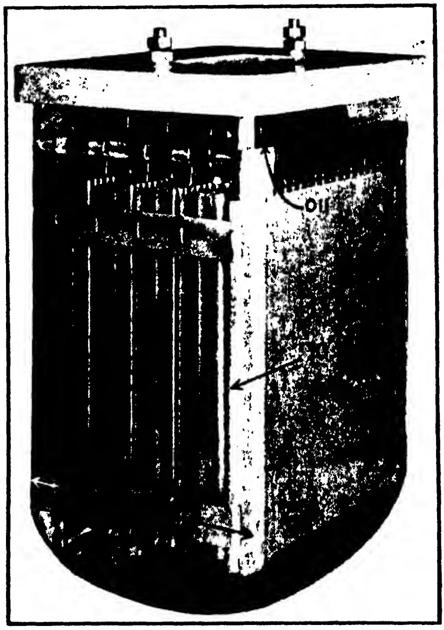


FIG. 103.—A 1500-microfarad electrolytic condenser; the electrolyte is covered with a thin layer of oil, to prevent evaporation.

compared with the polarizing voltages; the voltages given in the text are the continuous voltages impressed on the condenser. It will be appreciated that these electrolytic condensers are always operated on unidirectional voltage lines. They would not function as condensers if put on an ordinary a.c. line because the voltage across their terminals must never reverse; if it does, an excessive current will flow and the line will be practically short-circuited.

The performance of a normal liquid electrolytic condenser is shown in Fig. 102; its construction is shown in Figs. 103 and 104. The glass jar containing the condenser is 8 in. by 10 in. by 14 in., and the condenser

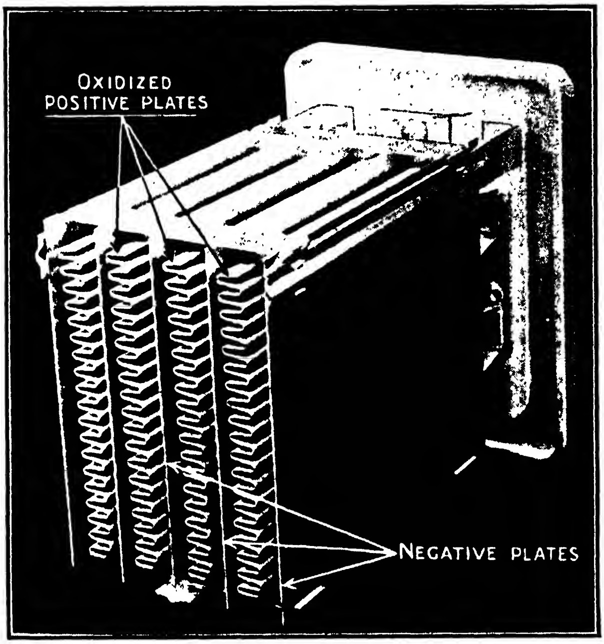


FIG. 104.—Construction of the condenser shown in Fig. 103.

weighs 42 lbs. complete. A layer of oil is poured on the electrolyte to keep it from evaporating, and in addition the cover is sealed tight to the jar. Corrosion of the aluminum plates takes place in time but apparently does but little harm providing the electrical contacts remain tight. An aluminum hydroxide gradually forms in the electrolyte, faster for a greater concentration of the electrolyte. One showing 75 ohms per cubic centimeter (a relatively concentrated solution) may need renewing before one year of use, whereas a solution showing 300 ohms per cubic centimeter lasts more than five years before precipitation of the hydroxide requires renewal of the electrolyte.

Recently, " dry " electrolytic condensers have appeared; in one form two thin aluminum sheets, one of them having a formed oxide film, are rolled up with a layer of electrolyte-impregnated gauze between them. In another it seems that a thin sheet of rubber or similar material is used in place of gauze. Some of them show very low leakage current and have a very large capacity per unit volume. In Figs. 105 and 106 are shown characteristics of one of these condensers, as published by the manufacturer. It is claimed that the condenser will operate on 350 to 450 volts without damage, even if this voltage has pulsations of 100 volts. These condensers are very compact, one rated to operate at 300 volts having nearly 2 microfarads per cubic inch; another designed for an operating voltage of 450 volts, with 575 volts peak permissible, has about $\frac{3}{4}$ microfarad per cubic inch. Figures on the probable life of these dry electrolytic condensers are not now available.

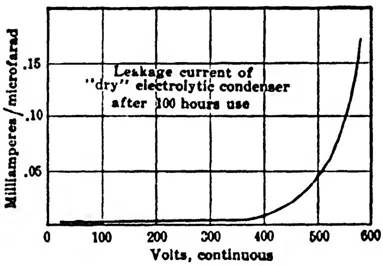


FIG. 105.—Leakage current of a dry electrolytic condenser.

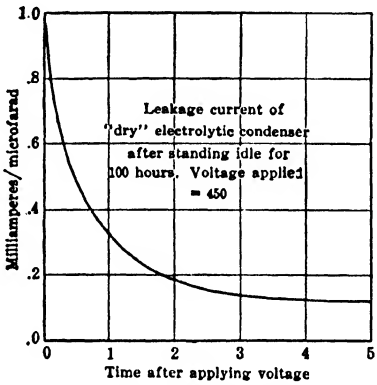


FIG. 106.—Decrease of leakage current with time the condenser is in use.

Characteristics of Ordinary Power Condensers.—Tests made by L. W. Austin on various condensers used for the transmitting circuits of radio sets gave results as shown in Table XIV. The tests were made at 14,500 volts and 300,000 cycles with damped wave excitation of 120 sparks per second.

In the case of the compressed-air condenser it is probable that practically all of the loss was attributable to dielectric losses in the insulated lead-in wire. At the voltages used it seems that the ordinary Leyden jar used in radio sets has considerable corona or leakage loss, because immersion in oil cut the losses to about 20 per cent of the value in air. Mica condensers were not tested at this time, but recent tests give them an efficiency rating nearly equal to the compressed air.

Tests on Pyrex glass give widely varying results, different tests giving power factors varying from 0.0028 to 0.0089. In general, annealed glass

TABLE XIV

Kind of Condenser	Power Factor	Capacity in 10^{-9} Farad	Equivalent Series Resistance in Ohms
Compressed air	0.001	5.8	0.14
Leyden jar in oil (glass)	.003	6.0	.28
Composition Murdock	.004	5.4	.41
Glass plates in oil	.005	4.2	.58
Moscicki condenser	.006	5.5	.57
Leyden jar in air	.016	6.1	1.4
Molded micanite	.023	4.1	2.9
Paper	.024	5.8	2.2

shows power factors about 50 per cent less than unannealed glass. Micalex (ground mica and lead borate fused together) is reported to have $K=8$ and $\cos \phi = 0.002$.

Test made by E. F. W. Alexanderson, using a high-frequency alternator for source of power, shows power factors greatly in excess of the values given by Austin's results. Some of the values obtained by Alexanderson are given in the accompanying table; the frequency used varied from 20 kc. to 90 kc., and the potential gradient from 5000 volts per cm. to 20,000 volts per centimeter. The power factor for most of the dielectrics tested increased somewhat with frequency increase, the amount of increase being small for the better dielectrics; thus the power factor for mica was constant within the range of frequency used, while glass increased from 0.013 to 0.016. All of the samples showed an increase of power factor with increased potential gradients, a slow increase at first, then more rapidly until rupture occurred; glass, e.g., increased its power factor from 0.013 to 0.015 with a change in potential gradient from 5000 to 12,000 volts per centimeter, and when the gradient was further increased to 19,000 volts per centimeter the power factor rose to 0.054.

For a gradient of 10,000 volts per centimeter and frequency of 50,000 cycles the results obtained were as follows:

TABLE XV

Material	Power Factor	Watts Loss per Cm. Cube
Built-up mica	0.019	0.15
Glass	.014	.25
Paper	.021	.26
Varnished cambric	.031	.35

The mica used was built up from small mica sheets and some binding cement; it seems likely that the losses in the cement and possible small air cavities caused more loss than did the mica itself; the small temperature rise in a good mica condenser built especially for radio work would indicate that a comparatively small part of the loss found by Alexanderson was due to losses in the mica itself, unless a poor grade has been used. He found some samples of built-up mica with a power factor as high as 0.07; it would seem likely that a lot of air was trapped in this sample. Some recent tests have indicated a power factor in mica as low as 0.0003.

It will be noticed that there is a considerable difference between Austin's results and those of Alexanderson; e.g., glass gave power factors of 0.005 and 0.014, in the different measurements. The difference is probably attributable to the different quality of glass used and also to the fact that different methods of experimentation were used; in one case the material was subjected to the loss continuously and in the other for only a small fraction of the time. Alexanderson used continuous-wave excitation and Austin a 120-cycle spark; the resulting temperature rise was undoubtedly different in the two tests.

Most of the solid dielectrics using synthetic resin for base have a power factor (at radio frequencies) of about 4 per cent.[1] Some show a power factor increasing with age of the material, the power factor of some of them increases with increase in frequency and in others a decrease of power factor occurs.

A dielectric which has had wide application in radio apparatus in recent years goes by the trade name of Isolantite; it is principally magnesium silicate suitably prepared and baked. After being ground and formed into a paste it is pressed into whatever shape is desired, and then fired in accurately controlled kilns. Its properties, as published by the manufacturer and compared with other insulating materials are given in Table XVI.

TABLE XVI

Material	Power Factor $\cos\phi$	Power Factor Multiplied by Specific Inductive Capacity $K \times \cos\phi$
Fused quartz	0.0007	0.0025
Isolantite	.002	.01
Electrical glass	.004	.022
Porcelain	.007–.15	.05–1.05
Good synthetic resin	.007	.045
Rubber compounds	.01	.03
Other synthetic resins	.03–.11	.12–.55

[1] See detailed data on pp. 266-267.

In Fig. 107 are shown the variations in specific inductive capacity, with change in temperatures, of isolantite compared with that of the average synthetic resin. In Fig. 108 is shown how the power factor of isolantite varies with frequency throughout the broadcast band, both dry and after it has been soaked in distilled water for 24 hours.

Phase Angle Difference of a Condenser.—In a good condenser the angle of current lead, ϕ, is very nearly 90°; the power factor of the condenser is the cosine of this angle, ϕ, or it may be put equal to sin ψ, where $\psi = 90° - \phi$. This angle ψ is called the phase difference of the condenser, and it is evident that if ψ is only 1° or 2° that the power factor of

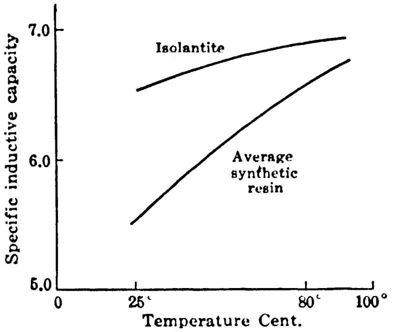


Fig. 107.—Specific inductive capacity of Isolantite (from data furnished by the manufacturer).

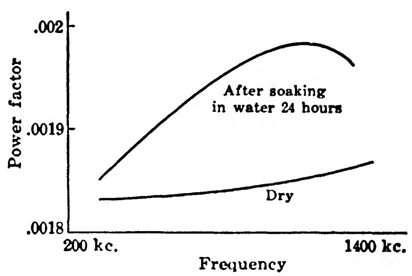


Fig. 108.—Effect of water upon the power factor of Isolantite (from data furnished by the manufacturer).

the condenser, cos $\phi = \sin \psi = \psi$, hence, the power used in a condenser is readily given in terms of ψ.

Power used

$$= EI \cos \phi = EI\psi = \omega CE^2\psi. \qquad (39)$$

The power factor of the condenser

$$= \frac{\text{Resistance}}{\text{Reactance}} = R\omega C. \qquad (40)$$

The phase difference of a condenser to be used for radio work should never exceed 0.3°; a greater value indicates excessive dielectric loss.

Variation of Dielectric Constant and Phase Angle with Frequency and Temperature.—Certain liquid dielectrics have been found to suffer large changes in both dielectric constant and power factor as the frequency and temperature are changed. According to theory widely accepted, these peculiar variations are due to dipoles (polarized molecules) which have reasonably definite periods of oscillation. If then the frequency of the impressed voltage happens to be reasonably close to this value anomalous changes in dielectric constant and power factor are to be expected. Fig. 109 gives typical results [1] obtained for an oil extracted from resin. Other oils show similar effects, and it is quite likely that practically all liquid dielectrics have this property in some temperature range.

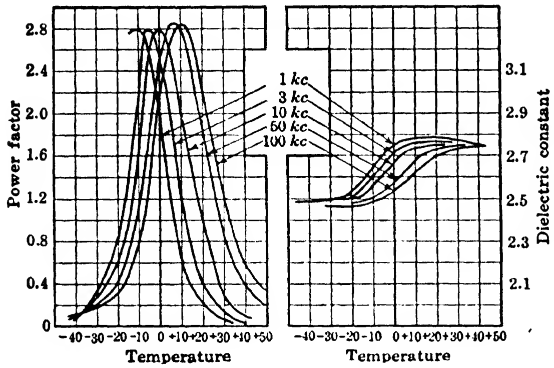


FIG. 109.—Certain substances show peculiar changes in power factor and specific inductive capacity with frequency and temperature; these results are for a certain resin oil.

Comparative Merits of Dielectrics. — When the phase angle difference of a dielectric is small, so that $\sin\psi = \psi$, we can get a measure of dielectric merit in terms of specific inductive capacity and ψ, as shown by Hoch.[2] If V is the volume of dielectric used, E is the voltage across it, f is the frequency, K the specific inductive capacity, ψ the phase difference, P the power used in dielectric hysteresis, and m a constant depending upon the various units employed, we have

$$P/V = mE^2fK\psi. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (41)$$

Thus the product, $K\psi$ gives the relative merits of two dielectrics insofar as power loss is concerned. For several commercial insulating materials, using continuous-wave excitation and resistance variation method of measurement Hoch found results as in Table XVII.

In general as the temperature of a dielectric is increased its losses increase at the same time. Hoch found temperature effects in the various commercial insulators as shown in Table XVIII.

Phase Difference Caused by Dielectric Loss is Constant for a Given Material.—The dielectric loss in most dielectrics varies with the square of the potential gradient in the dielectric, other quantities being fixed; this merely states the fact that the power factor of the condenser is independent of the voltage. Such has been found to be true for most materials, for voltages well below the rupturing strength of the dielectric. If then we have a condenser (of certain capacity), made with a certain dielectric, it will produce a certain loss, no matter how much or how little of the

[1] Morgan and White, Jour. of Franklin Institute, March, 1932, p. 313.

[2] Bell System Technical Journal, Nov., 1922. For other values, differing somewhat from those given here, see article by Guthrie in Proc. I.R.E., for Dec., 1924.

dielectric we use. Thus if we double the thickness of the dielectric (cutting the gradient in half) we must increase the area of the dielectric by two, thus using four times the volume of the dielectric as before. With the gradient cut in two the dielectric loss per unit volume is cut to one-fourth, but as we have four times the volume the total loss is the same. It seems then that the efficiency of a condenser cannot be improved by using more or less dielectric; a better dielectric must be substituted if the phase difference, ψ, is to be reduced.

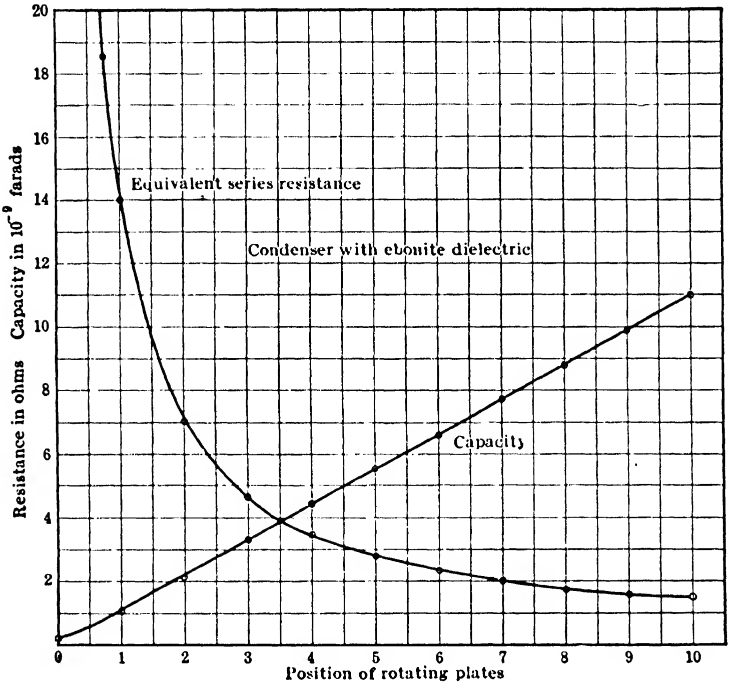


FIG. 110.—Variation of equivalent series resistance of a radio condenser having ebonite dielectric.

Fig. 110 gives curves showing the experimentally determined resistance and capacity of a variable condenser having ebonite for the dielectric. The condenser showed the product $RC = 14 \times 10^{-9}$ everywhere throughout the scale; the test was performed at 25,000 cycles, giving a value for $\psi = 0.0022$. In this condenser, therefore, the current leads the voltage by an angle of 89.874°.

Phase Difference of Actual Condensers.—There are of course two types of condensers used in radio, fixed ones using solid dielectric between their plates and variable ones using air (occasionally oil) for their dielectric.

TABLE XVII

DIELECTRIC CONSTANT, PHASE DIFFERENCE AND THEIR PRODUCT FOR SEVERAL COMMERCIAL INSULATING MATERIALS *

Material	Frequency Kilocycles	Dielectric Constant	Phase Difference, Degrees	Product
Phenol Fiber A	295	5.9	2.9	17.1
	500	5.8	2.9	16.8
	670	5.7	2.9	16.5
	1040	5.6	3.3	18.5
Phenol Fiber B	190	5.8	2.2	12.7
	500	5.6	2.5	14.0
	675	5.6	2.6	14.6
	975	5.6	2.8	15.7
Phenol Fiber C	200	5.4	2.1	11.3
	395	5.4	2.2	11.8
	685	5.3	2.3	12.2
	975	5.2	2.4	12.5
Phenol Fiber D	194	5.4	4.2	22.7
	500	5.2	3.9	20.3
	695	5.2	3.9	20.3
	1000	5.1	3.8	19.4
Wood (Oak)	300	3.2	2.1	6.7
	425	3.3	2.0	6.6
	635	3.3	2.2	7.3
	1060	3.3	2.4	7.9
Wood (Maple)	500	4.4	1.9	8.4
Wood (Birch)	500	5.2	3.7	19.2
Hard Rubber	210	3.0	0.5	1.5
	440	3.0	.5	1.5
	710	3.0	.5	1.5
	1126	3.0	.6	1.8
Flint Glass	500	7.0	.24	1.68
	720	7.0	.24	1.68
	890	7.0	.23	1.61
Plate Glass	500	6.8	.4	2.7
Cobalt Glass	500	7.3	.4	2.9
Pyrex Glass	500	4.9	.24	1.18

* All of the samples had been in the laboratory for some time during summer weather without artificial drying or other special preparation.

TABLE XVIII

VARIATION WITH TEMPERATURE OF DIELECTRIC CONSTANT, PHASE DIFFERENCE AND THEIR PRODUCT FOR SOME COMMERCIAL INSULATING MATERIALS (FREQUENCY 500 Kc.) *

Material	Temperature, Degrees C.	Dielectric Constant	Phase Difference, Degrees	Product
Molded Phenol Product A...	21	5 6	3.1	17.4
	71	6.9	6.5	45.0
	120	10.4	22.0	230.0
Molded Phenol Product B...	21	5 2	2.3	12.0
	71	6.1	3.7	22.5
	120	7.6	8.9	68.0
Molded Phenol Product C...	21	5 3	2.8	14.8
	71	6.1	3.6	22.0
	120	6.7	9.6	64.0
Phenol Fiber B............	21	5.6	2.5	14.0
	71	6 6	3.1	20.5
	120	6.5	4 6	30.0
Phenol Fiber C............	21	5 4	2.3	12.4
	71	6 0	3.9	23.5
	120	5 3	4.9	26.5
Phenol Fiber D............	21	5 2	3.9	20.3
	71	6 6	6.9	46.0
	120	6 3	13 5	85.5
Hard Rubber............	21	3 0	0 5	1.5
	71	3 1	1 2	3.7
	120	3 2	3.7	11.8
Pyrex Glass...............	20	4 9	0 24	1.18
	74	5 0	0.4	2.0
	125	5 0	0 7	3 5

* The measurements on each sample were made in the order in which they are given in the table.

The solid dielectric ones are generally used for by-pass, or blocking, condensers and the properties of their dielectric has but little effect on the performance of the radio set.

The variable air condensers are used in the tuning circuits and their phase difference may be of prime importance. If the equivalent series resistance of the tuning condenser approaches that of the coil with which it is associated its losses are important in that they affect the tuning of the circuit. The well-built condenser of to-day, however, has a resistance so small compared to the coil resistance that its effect in tuning is negligible. This is really because in a well-designed condenser the amount of dielectric used is small, what is used is of the best quality, and is placed on the condenser where the electric field intensity is low.

Measurement of many of the better types of variable condensers used in radio sets to-day show at frequencies of about 1000 kc. an equivalent series resistance varying from 0.5 to 1.5 ohms with the condenser set for maximum capacity. As the reactance of these condensers averages about 300 ohms, we can see that their power factors are considerably less than those of the coils with which they are necessarily associated. Typical results, for example, show

Capacity	Frequency	Resistance	Power Factor
0.0005	1500 kc.	0.8	0 0038
0.001	1000 kc.	0.4	0.0024

These measurements seem to indicate, when compared with the measured values at 1000 cycles, that the plates themselves of the condenser begin to offer appreciable resistance at the higher frequencies,[1] and for frequencies measured in many megacycles the condenser resistance may be equal to, or greater than, that of the coil.

Using a small tuning condenser of 0.00008 microfarad capacity and a calorimetric method of resistance measurement Ramsey found [2] resistances as given here

Frequency.........	10^6	$3\ 7\times10^6$	$7\ 5\times10^6$	1.4×10^7
Resistance.........	0.1 ohm	0.06	0 04	0 06

We ordinarily expect to find the phase angle difference of a given dielectric to be independent of frequency and for most substances this seems to be true. For liquids the effect seems to be more complex, measurements indicating a phase angle changing with frequency. For water Bryan [3] found that at room temperature

$$\psi=0.8^\circ+\frac{2.09\times10^6}{f}. \qquad (42)$$

[1] See article on Condensers, by Weyl and Harris, Proc. I.R.E., Feb., 1925.

[2] Phil. Mag., Dec., 1926.

[3] "Dielectric Losses at Radio Frequencies in Liquid Dielectrics," Bryan, Phys. Rev., Oct., 1923.

He found somewhat similar relations for other liquids. Fig. 109 shows that with liquid dielectrics peculiar changes in behavior may be expected as frequency and temperature are altered.

Internal Capacity of a Two-layer Solenoid.—A certain single-layer solenoid had an inductance of 1000 μh at 600 meters; another solenoid was at hand which had the same dimensions as the first, but it had two layers of wire instead of one, giving it twice as many turns. Tested at 1000 cycles it showed an inductance slightly more than four times as great as the single-layer solenoid, as it should, but when tested at 500,000 cycles it acted like a condenser, not like an inductance; in fact it ceased to act like an inductance for frequencies above 200,000 cycles. This peculiar behavior was caused by the internal capacity of the coil; this internal capacity may be represented to a certain degree of approximation, as a condenser connected in parallel with the terminals of the coil. It is then evident that above a certain frequency the current taken to charge the condenser will be greater than the current through the coil itself, making the com-

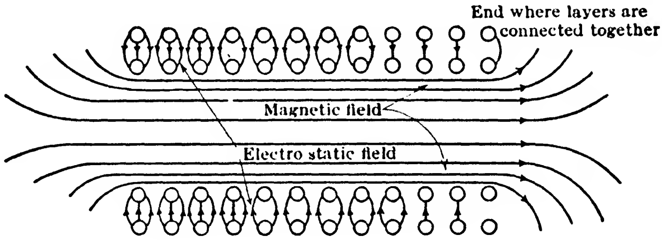


Fig. 111.—Magnetic and electric fields in an ordinary two-layer solenoid.

bination circuit act like a condenser, of capacity varying with the frequency.

The calculation of the internal capacity [1] of a coil is most conveniently carried out by calculating the electrostatic energy stored in the coil for a given voltage; the value of C is then at once obtained. The capacity of the two-layer solenoid referred to above will first be calculated. Fig. 111 depicts the arrangement of the electrostatic and magnetic fields of the coil when current is flowing through it; when the impressed e.m.f. is continuous, the difference in potential between the two layers varies directly as the distance from the end where the two layers connect together, being zero at this point. When the impressed e.m.f. is alternating, this potential difference between the two layers is no longer a straight-line variation but varies more rapidly in the center of the coil; the exact form

[1] In Phys. Rev., Aug., 1921, Breit publishes a note in which he says that for a short single-layer solenoid the internal capacity (in $\mu\mu f$) is nearly equal to 7 per cent of the perimeter of the coil, in centimeters.

of this potential difference curve varies with the frequency and with the shape of the coil.

It is noticed that the capacity of the coil is essentially that of two concentric cylinders, the separation of these two cylinders being determined by the average separation of the wire on the two layers of the coil, being perhaps four times the thickness of insulation of the wire, for wire and insulation with relative proportions as shown in Fig. 112; this is about the right scale for No. 26 double cotton-covered wire. With such wire then we can calculate the capacity by replacing the actual coil by two conducting cylinders having the same diameter and length as the coil, and separated by a distance equal to four times the thickness of insulations of the wire. As the separation of the cylinders is so small compared

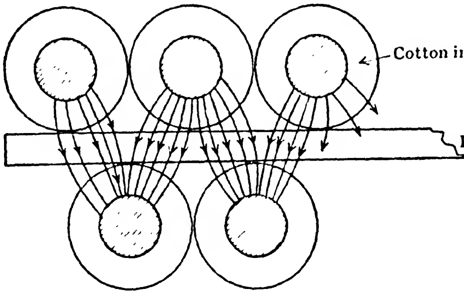


FIG. 112.—Electric field between neighboring conductors of a two-layer solenoid.

to the diameter, the capacity may be calculated by using the formula for flat plates. Hence we get

$$C=\frac{2\pi rlK}{4\pi 4t}=\frac{rlK}{8t} \text{ centimeters}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (43)$$

where r = mean radius of coil in centimeters;
l = length of coil in centimeters;
t = thickness of insulation in centimeters;
K = specific inductive capacity of dielectric. If the coil is impregnated with shellac, K is approximately equal to 3.

Eq. (43) gives the capacity for static charges, the two cylinders being insulated from one another; in the coil, however, the two cylindrical surfaces are actually connected together at one end. The problem then resolves itself in one of determining the equivalent capacity of two cylinders having a potential difference of zero at one end and E volts at the other. The diagram in Fig. 113 gives the elements of the problem; two plates, capacity per unit length $=\dfrac{rK}{8t}$ with potential difference as represented by the e curve in the upper part of Fig. 113. The energy stored in an element of length of the condenser is

$$dW=\frac{rK}{8t}\frac{e^2}{2}dx=\frac{rKE^2}{16t}\left(1-\frac{x}{l}\right)^2dx.$$

The total work stored in the electrostatic field,

$$W=\frac{rKE^2}{16t}\int_0^l\left(1-\frac{x}{l}\right)^2dx=\frac{rKE^2}{16t}\frac{l}{3}.$$

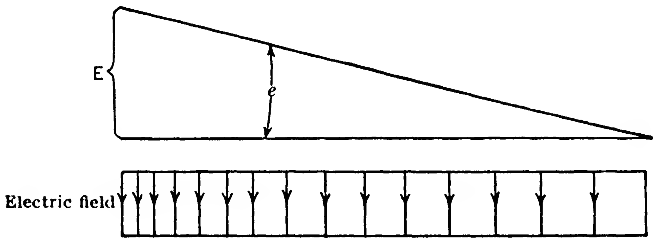


FIG. 113.—Variation in potential difference between the layers of a two-layer solenoid, at low frequency.

Now we define capacity in a problem like this by putting the total energy stored in the field equal to $\dfrac{C'E^2}{2}$ where C' is the capacity we desire to calculate.

So
$$\frac{C'E^2}{2}=\frac{rKE^2l}{48t}\quad\text{or}\quad C'=\frac{rKl}{24t}. \qquad (44)$$

By comparison with Eq. (43) we see that the equivalent capacity of such a coil, assuming uniform change in the potential gradient from one end to the other, is equal to one-third of the static capacity of the two surfaces.

The actual distribution of potential differences between the two layers is more nearly as shown in the curve of Fig. 114; such a distribution will result in C' being somewhat larger than the value obtained in Eq. (44).

Internal Capacity of a Multi-layer Coil.—A multi-layer coil constructed with an air space between each layer may have a comparatively small internal capacity in spite of the fact that it has 10 or 20 layers of winding; a short analysis shows that the internal capacity, as a matter of fact, decreases with an increase in the number of layers. If the capacity between two adjacent layers of the coil is C then the internal capacity of a coil having N layers is nearly C/N, as will be shown.

The electrostatic field in such a coil has a distribution about as shown in Fig. 115, which represents a cross-section through the winding of an air-spaced coil, having 8 layers. The cross-section is shown through one side of the coil only, the other side of the coil would be similar.

If a voltage of E volts is impressed across the terminals of the coil, the voltage between adjacent layers (at the ends where the two layers are not connected) is $E \div N/2$.

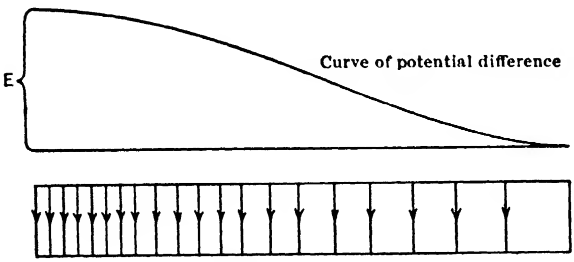


FIG. 114.—Potential difference between the two layers of a two-layer solenoid at high frequencies.

Let the normal geometrical capacity between two adjacent layers be C; this is the capacity between two cylinders, insulated from one another, of same dimensions as the layers of wire, and spaced by the space between the two layers of wire. If the potential gradient between two adjacent layers is assumed to vary uniformly from a maximum at one end to zero at the other (where the two layers connect together), as illustrated in Fig. 113, the energy between two adjacent layers is found by integration to be $C\frac{2E^2}{3N^2}$. As there are $N-1$ spaces between layers, where this much energy is stored the total energy stored is $(N-1)\frac{2E^2C}{3N^2}$.

Now if we consider a condenser made up of two cylinders having the same dimensions and spacing as the inner and outer layers of the coil, its capacity would be equal to $\frac{C}{N-1}$, the thickness of dielectric being $(N-1)$ times as thick between the innermost and outermost layers as it is between two adjacent layers. The stored energy in such a condenser would be

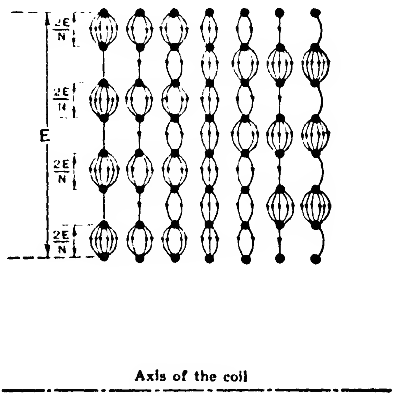


FIG. 115.—Electric field distribution in a multi-layer coil.

$$\left(\frac{C}{N-1}\right)\frac{E^2}{2}.$$

But from the previous paragraph the stored energy in the coil is actually

$$(N-1)\frac{2E^2C}{3N^2}=\left(\frac{N-1}{N}\right)^2\frac{2E^2}{3}\left(\frac{C}{N-1}\right),$$

from which it follows that the equivalent internal capacity of a multi-layer coil is equal to $\frac{4}{3}\left(\frac{N-1}{N}\right)^2 \times$ the capacity between the inner and outer layers, that is

$$C'=C_0\times\frac{4}{3}\left(\frac{N-1}{N}\right)^2, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (45)$$

where $C_0=$ capacity from inside to outside layer.

This capacity calculation neglects the "edge effect" which may increase the actual capacity over that given by Eq. (45) by as much as 100 per cent, depending upon the shape of coil.

In calculating the separation of the inner and outer layers the thickness of wire in the intermediate layers must not be included; thus if the space between the surfaces of the wires of two adjacent layers is 0.5 mm. and there are ten layers, the space between the inner and outer layers is 4.5 mm.

To keep the capacity low it is very necessary to keep the adjacent layers separated by an appreciable air space; if this space is too small the internal capacity is high, and the coil cannot be used at as high a frequency as it is possible with a coil of equal inductance having a low internal capacity. If paper, oiled cloth, wax, shellac, or similar dielectric is used to separate the different layers, there will be an appreciable dielectric loss in this material (thereby decreasing the efficiency of the coil) and the internal capacity will be increased by a factor equal to the specific inductive capacity of the material used. The bad effects of the dielectric used between layers increase as the amount of external tuning capacity is decreased.

It might seem from the previous reasoning that a large air space would be advisable, but such is not the case; as the air space is increased the value of inductance for a given amount of wire is decreased, the ratio of L to R being greater the more compact the coil. Just what air space is best the author has not determined, but with a coil made of well-stranded radio-cable an air space equal to the diameter of the cable has seemed suitable.

Natural Period of Multi-layer Coils.—It is possible to calculate the natural period of the air-spaced coils with a fair degree of precision, perhaps within 10 per cent. The capacity to be used in making the calculation is considerably greater than the value given in Eq. (45) because of the edge effects and a redistribution of potential in the coil. For some coils having a square cross-section of winding, outer radius about 30 per cent greater than the inner radius, the natural period could be determined by using for L its low-frequency value and for C just twice the value calculated from Eq. (45). As examples of how well the prediction may be made the data for two coils are given in Table XIX.

TABLE XIX

	Coil No. 1	Coil No. 2
Inner radius	4.7 cm.	4.7
Outer radius	5.5 cm.	6.3
Axial length of coil	2.5 cm.	2.5
Number of layers	10.0	18.0
Number of turns (total)	736.0	740.0
Air space between layers	0.060 cm.	0.033 cm.
Calculated capacity, Eq. (45)	14.2×10^{-12}	16.2×10^{-12}
Low frequency inductance	66.2×10^{-3}	72.2×10^{-3}
Calculated natural frequency	1.09×10^{5}	1.11×10^{5}
Measured natural frequency	1.16×10^{5}	1.05×10^{5}

The natural period of a multi-layer coil does not increase rapidly with increase in inductance. Thus if the number of layers in such a coil is doubled, the inductance is increased nearly four times, but, as the capacity has been decreased to only half its previous value, the natural period has been increased only about 40 per cent.

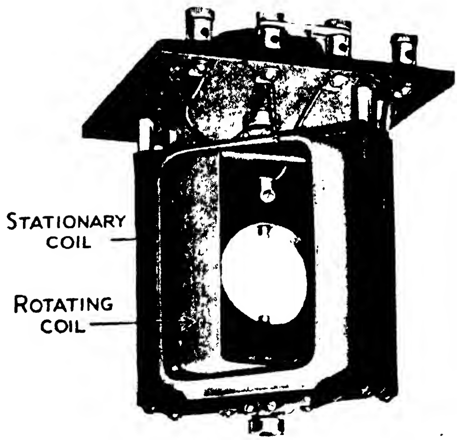


FIG. 116.—A laboratory type of variable inductance; one coil turns inside the other.

Natural Wave Lengths of Typical Coils.—For single-layer solenoids the natural wave length can be reasonably well calculated by using the low-frequency inductance of the coil and an internal capacity equal, in micro-microfarads, to the radius of the coil in centimeters. For multi-layer coils of the banked type the internal capacity must be taken considerably greater, especially when the coil is short. In Table XX are shown experimental results for a whole set of coils.

TABLE XX

Type of Coil	Diameter in Cm.	Length in Cm.	Number of Turns	Inductance in 10^{-6} Henry	Natural Wave Length in Meters	C_0-in Micro Micro-Farads
Single-layer solenoids on bitumen spools.	11.7	2.3	20	79	37	5.2
	11.7	4.8	40	200	60	4.7
	11.7	7.5	60	374	78	4.4
	11.7	17.8	141	1,140	122	3.6
	11.9	3.3	30	144	53	5.5
	12.0	42.3	931	26,000	625	4.2
	11.7*	7.4	126	214	213	6.0
	14.7	17.6	296	7,560	370	5.1
Single-layer solenoids on wood spools.	10.5	10.0	361	8,490	388	5.0
	10.5	10.0	290	5,540	323	5.3
	10.5	10.0	138	1,290	153	5.1
	10.5	10.0	96	614	104	5.0
	10.5	10.0	64	278	69	4.8
	10.5	10.0	41	120	49	5.6
	10.5	10.0	37	97	44	5.6
Three-layer banked coils on bitumen spools.	11.4	1.5	36	205	108	16.0
	11.4	2.1	52	465	133	13.3
	11.4	2.7	68	725	158	9.7
	11.4	3.4	84	1,015	175	8.5
	11.4	4.2	100	1,310	188	7.6
	11.4	5.7	132	2,030	225	7.0
	11.4	7.3	169	2,880	263	6.8
	11.4	9.8	232	4,220	297	5.9

* Fitted with taps and multipoint switch.

As another illustration indicating the effect of the internal capacity of coils in their performance we note the following interesting facts. A set of fixed inductances, used for laboratory secondary standards, showed natural frequencies as given here:

TABLE XXI

Value of inductance...	0.0001 henry	0.001	0.01	0.1	1.0
Natural frequency....	6×10^6 cycles	2×10^6	5×10^5	1.5×10^5	3.5×10^4
Resistance (c.c.)......	0.18 ohm	1.80	12.2	85.3	545

A set of variable inductances, built as shown in Fig. 116, one coil rotatable inside the other, gave natural frequencies, minimum power factors, etc., as follows:

TABLE XXII

Inductance Range	Natural Frequency at Maximum Setting	Minimum Power Factor	Frequency at Which Minimum Power Factor Occurred
0.00004 to 0.0004 henry	1.5×10^6	0.007	2.0×10^5
0.0004 to 0.004	5.0×10^5	0.008	7.0×10^4
0.0018 to 0.018	2.0×10^5	0.01	2.5×10^4

It will be appreciated that the internal capacity of a coil acts to seriously limit the amount of inductive reactance available. A coil cannot be used as a good reactance at a frequency much higher than about one quarter of its natural frequency, because of the rapid increase in effective resistance at higher frequencies; it will be found that none of the coils used in obtaining the above data make possible 100,000 ohms of inductive reactance.

General View of Capacitance and Inductance.—So far in this text we have always thought of coils and magnetic fields when considering inductance and, of course, that is the ordinary conception. But the true distinction of an inductance arises from the reactions set up in the circuit; if, due to the flow of alternating current, *the circuit sets up a counter electromotive force which lags 90° behind the current,* then the circuit is an inductive one. In the same way if, due to the flow of an alternating current, a circuit *sets up a reaction which leads the current by 90°*, the circuit is a capacitive one.

Now, in most cases, the student will recognize inductive and capacitive circuits by their physical make up; coils for inductances and conducting plates separated by a dielectric for a condenser form a reasonable basis for judgment. When both coils and condensers are used the circuit will be inductive or capacitive according to the dominant reaction in the equivalent series circuit.

There are, however, many devices which are entirely misleading as to their electrical character. Thus a bronze wire, stretched tightly across a magnetic field, has the appearance of a resistance only, plus a slight inductance calculated from Eq. (11). But when alternating current is put through such a wire it will vibrate and this vibration across the magnetic field will generate an e.m.f. in phase with the wire's velocity. But the stretched string is a mechanically resonant system and the phase of its velocity, with respect to the current flowing through the wire, will depend entirely upon the ratio of the electrical frequency (of the current) compared to the natural mechanical frequency of the wire. If the impressed frequency is less than the natural frequency the velocity leads the current and so sets up a c.e.m.f. which is 90° ahead of the current, but such a circuit must be classed as capacitive, and so *we must classify the vibrating wire as a condenser.*

If the impressed and natural frequencies are the same the velocity and current are in phase and so the c.e.m.f. is in phase with the current. Except for the slight inductance the wire naturally has (Eq. 11), *the vibrating wire acts like a resistance,* the value of this resistance, however, being much in excess of the ordinary resistance of the wire.

If the impressed frequency is higher than the natural frequency, the velocity of the wire lags 90° behind the current and so the c.e.m.f. due to the wire's motion across the magnetic field lags 90° behind the current *and so the vibrating wire acts like an inductance.*

A synchronously revolving alternator, over-excited, acts like a condenser, and under-excited acts like an inductance. A tuned electrostatic telephone may act like an inductance due to the phase of the diaphragm's motion with respect to the voltage actuating it.

A small piece of quartz, suitably cut from a good crystal, will vibrate vigorously if put in an alternating electric field (*piezo electric quartz*). This expansion and contraction of the quartz sets up variable surface charges on the quartz which act as c.e.m.f.s. in the circuit. Such a piece of quartz may act like a high resistance, a condenser, or an inductance, according to the relative values of the impressed frequency and the natural mechanical vibrational frequency of the piece of quartz. In this case we have the interesting spectacle of a slab of quartz with tinfoil coatings on opposite faces of the slab, acting like a coil.[1] This is perhaps the most striking illustration of the physical appearance of a device belying its electrical character.

SHIELDING

In most laboratory work the problem of shielding one circuit from another, for both electric and magnetic fields, is of great importance, and recently shielding of one sort or another has been incorporated in every good radio receiving set.

Electrostatic Shielding.—Ever since the very early days of electricity it has been known that no charge resides inside of a closed box of conducting material. "Faraday's cage" was the first useful embodiment of this idea; recently many research laboratories have constructed completely copper-inclosed rooms, where experiments can be carried out free from the disturbing effects of the electric fields always present in a laboratory or electrical shop. As long as the metal enclosing the room is a good

[1] Cady, Proc. I.R.E., April, 1922.

conductor it may be very thin and yet give practically complete shielding from external electric fields. In fact a thin copper mesh, or zig-zag arrangement of copper wires covering the walls, ceiling, and floor of the room, all connected together and grounded, will suffice reasonably well for electric fields of any but the highest frequencies.

In Fig. 117 is given the idea underlying electrostatic shielding. A positive charge at A is producing an electrostatic field at the apparatus B; to shield from this field the apparatus is surrounded by a metallic box C. Induced charges will be set up on this box as indicated and the electric field at B must now be calculated from the original charge A and the induced charges on the shield. Now the induced charges on this shield will always be such that *the net electric field inside the shield is zero.*

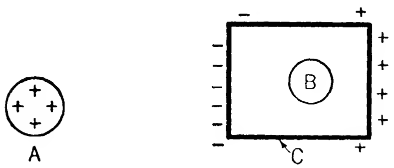


FIG. 117.—Induced charges on C prevent A having much effect at point B.

Generally, parts of the circuit of apparatus B are grounded, and in this case currents will run up and down the grounding wires of B, as charge A is moved around. To prevent this effect shield C should itself be grounded; thus its potential cannot change and apparatus B is shielded from external electric fields and, further, there will be no currents in its grounding wires. The above remarks hold good for all except the very highest frequencies.

A Cage around a Charge Does Not Shield.—It is to be noticed in Fig. 117 that the shield C is placed around the apparatus B and not around the disturbing charge A. If the shield were placed around the charge A the situation would be as shown in Fig. 118. Negative charges are induced on the inner surface of the metal box C and positive charges on its outer surface. All three charges now act at point B and the net field is just the same as if the shield had not been put around A. If, however, the shield C is grounded, the positive charge will disappear (electrons run up from the earth) and the resultant field at point B will now be zero.

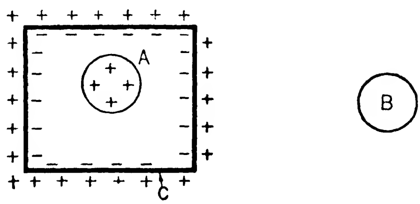


FIG. 118.—If the shield C is put around the charge A, no shielding action is obtained unless the shield C is grounded.

Effect of Frequency and Resistance upon Electric Shielding.—It will be noticed in Fig. 117 that shielding is accomplished by the combined action of the induced charges and the original charge. If the resistance of the shielding material is very high or the frequency is very high, the shielding against the electric field will be correspondingly less perfect. Thus if the shield is merely a wooden box covered with a moist cloth, the shielding against charge A will be perfect after sufficient time has elasped for the proper induced charges to take their proper places, but insofar as these induced charges are less than their proper values, or not in their proper places, the shielding is imperfect. The box covered with moist cloth, for example, will shield perfectly from charge A while this remains stationary; if, however, A is quickly moved toward B, larger induced charges must appear on the shield C, and while these extra charges are congregating the electric shielding at B is imperfect.

Shielding against Magnetic Fields.—A magnetic field may originate at a permanent magnet or at a circuit-carrying current. The two cases offer somewhat different characteristics, so we consider them separately and Fig. 119 takes up the case of a permanent magnet A, affecting apparatus B. By surrounding B with a box of highly permeable material (iron or permalloy) the magnetic field at B may be made to nearly disappear. If the magnetic reluctance of the box C were zero it would require no difference of magnetic potential to maintain the induced magnetization and the shielding would be perfect. The shielding is the more perfect as the reluctance of the shielding box C approaches zero.

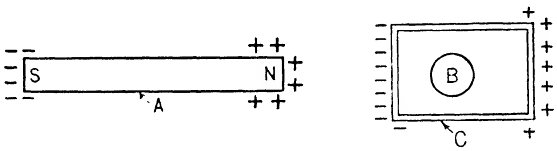


Fig. 119.—Space B may be shielded from magnet A by a highly permeable shield C.

It might be thought that surrounding the north pole of the magnet by the shielding box C would reduce the magnetic field at B to zero (Fig. 120), but such is not the case. The amount of shielding by such a procedure is practically nil, corresponding to the case of electric fields shown in Fig. 118. Induced magnetic charges (whatever those are) appear on shield C, negative charge on the inside and positive charge on the outside, and the combined effect of these charges and that of the magnet itself is nearly the same as if the shield were not present.

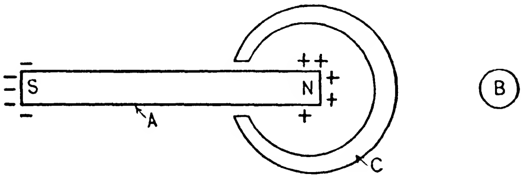


Fig. 120.—Putting an iron shield over one pole of the magnet does not shield point B.

If, however, the shielding box C is made to surround the whole magnet (Fig. 121) and the box C is of low magnetic reluctance, the shielding is practically perfect and there will be no magnetic field at B. This case corresponds to the case given in Fig. 118, when the shield is grounded.

Shielding against Electromagnetic Fields.—The amount of flux from a permanent magnet is fixed, independent of what changes are made in its magnetic circuit. Thus a horse-shoe magnet with the keeper, or armature, in place, has no more flux than when the keeper is removed. But if the permanent magnet is replaced by an electromagnet, the conditions are entirely different; with the keeper in place the amount of flux may be hundreds of times as great as when it is taken away. This is an essential difference in the two cases.

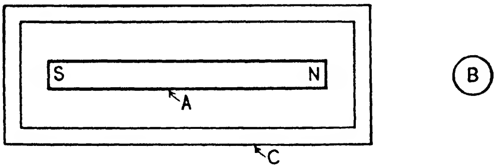


FIG. 121.—Surrounding the whole magnet by a permeable box C will successfully shield point B.

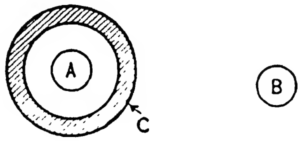


FIG. 122.—Putting a heavy iron pipe C around a wire A, carrying current, does not shield point B at all.

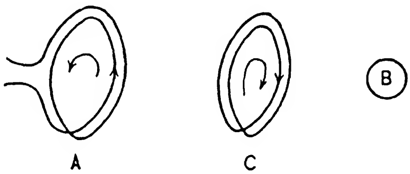


FIG. 123.—To shield a point from the effect of alternating currents other alternating currents in opposite phase must be used.

If a single conductor carrying current is surrounded by a heavy iron pipe, there is just as much flux density outside the pipe (and inside too for that matter) as if the iron pipe were not there. In Fig. 122, A represents the cross-section of a conductor carrying current, giving a certain undesired magnetic field at apparatus B. Putting a heavy iron pipe, C, around the conductor A does not change the magnetic field at point B in the least. The total flux around A has been greatly increased, but the density of the field at B is unchanged.

If an iron box, however, is placed around B, then this point is nearly shielded; how perfect the shielding is depends upon the magnetic reluctance of the box.

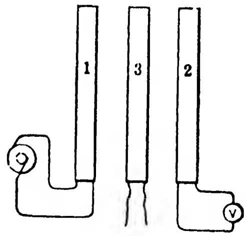


FIG. 124.—A coil arrangement for studying shielding action.

Shielding against Changing Magnetic Fields.—Changing magnetic fields are in general set up by coils carrying varying currents, in the general case, alternating currents. Now it will be noticed that to shield against electric charges we depended upon induced charges of opposite sign, to shield against magnetic poles we depended upon induced poles, and similarly to shield against changing electric currents (which set up both electric and magnetic fields) we must depend upon induced electric currents of opposite polarity, that is, opposite direction. To shield B (Fig. 123) against the combined electric and magnetic effects due to the alternating current in A we must induce in the shielding circuit, C, a suitable alternating current. In the case shown the shielding current can flow only in the two-turn coil C, and because of the confining of the current to this restricted path, the shielding will be correspondingly less than perfect. To completely shield space B we must surround it by a perfectly conducting box. The shape of such a box is of no importance, but it must be as nearly a perfect conductor as possible.

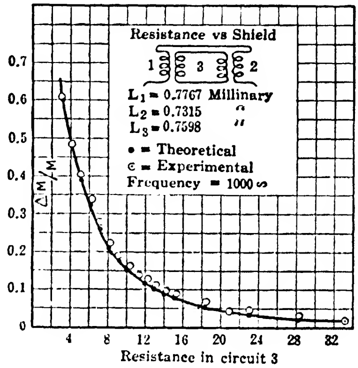


FIG. 125.—Effect of increasing resistance of coil 3 (Fig. 124).

The shielding action of the currents in such an enclosing box can well be studied by a coil arrangement as shown in Fig. 124, which shows in plan three short solenoids. If alternating current is sent through coil 1, its alternating magnetic field will reach in to Coil 2 and so set up an induced voltage in this coil. The magnitude of this induced voltage (for a given current in Coil 1) is a measure of the mutual induction between coils 1 and 2.

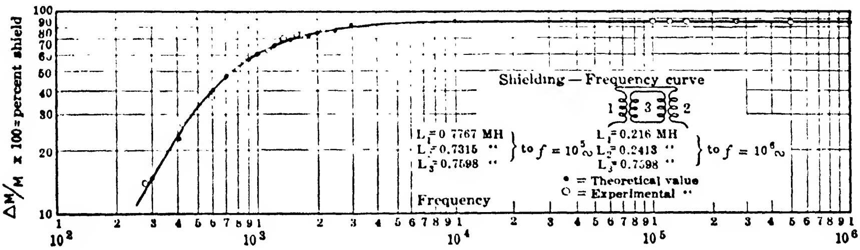


FIG. 126.—Showing how shielding action increases with frequency.

Now if coil 3 is closed, it will have currents set up in it by the alternating field of coil 1, and these currents will tend to diminish the amount of alternating flux reaching coil 2 from coil 1. The voltage induced in coil 2 will therefore be diminished. This decrease in induced voltage in coil 2, expressed as a ratio of the voltage induced in coil 2 with coil 3 not acting (open-circuited), is a measure of the shielding action of coil 3.

An analysis of the amount of this shielding shows that[1]

$$\Delta M_{1-2}/M_{1-2}=K\overline{\omega L_3}^2/Z_3^2. \qquad (46)$$

If then the resistance of coil 3 is increased, reactance remaining fixed, the shielding must diminish. In Fig. 125 this action is shown; the resistance

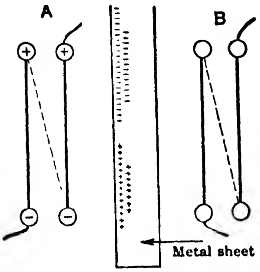


FIG. 127.—Metal plates shield better than coils; the shielding currents are distributed in bands in the plate.

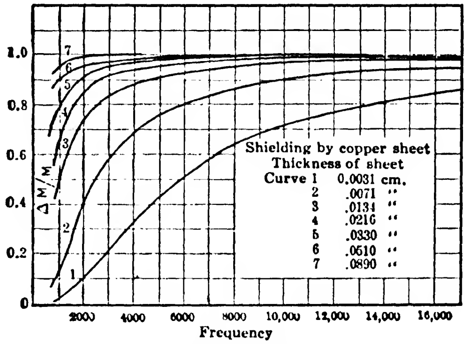


FIG. 128.—Shielding efficiency of copper.

of coil 3 was increased, and the shielding, $\Delta M_{1-2}/M_{1-2}$, measured. In Fig. 125 the experimentally determined points, as well as points calculated from Eq. (46) are plotted.

It is also evident from Eq. (46) that if the resistance of coil 3 stays fixed, the shielding will increase with frequency, becoming asymptotic to the value K which depends upon the relative placing of the three coils. Fig. 126 shows this change in shielding with frequency change.

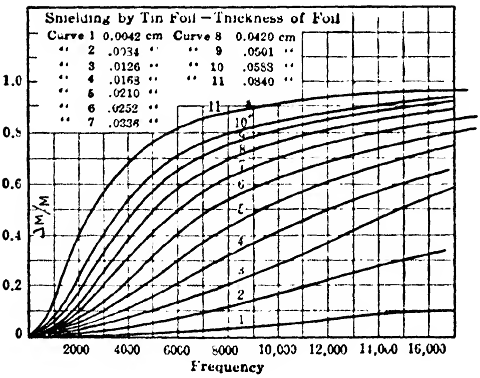


FIG. 129.—Shielding efficiency of tinfoil.

If a metal sheet is used in place of coil 3 it is evidently difficult to write an exact theory because the paths of the eddy currents in the metal sheet are difficult to evaluate. Fig. 127 gives an indication of the action; the eddy currents flow mostly in that face of the metal sheet closer to coil A and of course in phase relation nearly 180° behind the current in coil A. The plus and negative signs in Fig. 127 show somewhat the distribution of current in the cross-section of the metal sheet.

Quite evidently the shielding action of such a sheet will increase with thickness and with frequency. Fig. 128 brings out this idea very well.

[1] See article by Morecroft and Turner, Proc. I.R.E., August, 1925.

The copper sheets were 6 in. square, and the coils A and B were 4 in. in diameter, placed coaxially and quite close together.

The higher the resistance of the shielding metal the less will the shielding action be, as shown by Eq. (46) and by Fig. 125. This action is also shown by comparison of Figs. 128 and 129, the latter for tinfoil and the former for copper.

Shielding Increases Losses.—Of course the eddy currents in the shielding plates must increase the effective resistance of the coil setting up the changing magnetic field, which produces these eddy currents. Theory shows that the resistance added by the shielding currents will be a maximum for a certain thickness of shield, for any definite frequency and kind of metal, and this is brought out by the curves of Fig. 130.

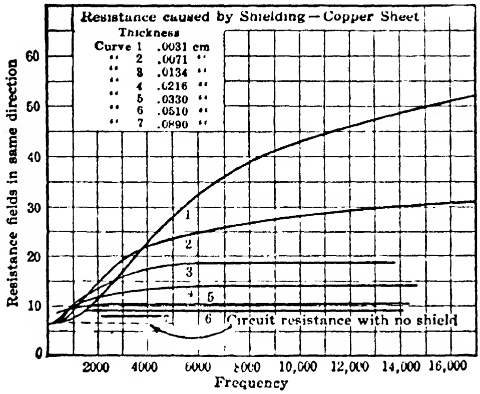


FIG. 130.—Resistance caused by shielding plates varies with frequency, material, and thickness of material. For one frequency a certain thickness gives maximum resistance.

Of course the shielding action continues to increase as the thickness of the shield is increased (see Fig. 128), but the resistance caused by the shielding action is a maximum for a definite thickness of sheet. For example, at a frequency of 2500 cycles, a copper shield 0.0071 cm. thick causes more increase in resistance than other sheets either thinner or thicker.

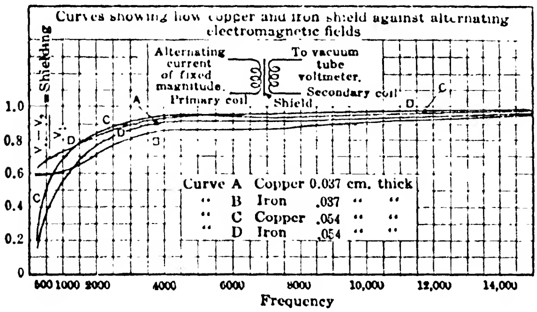


FIG. 131.—Comparison of iron and copper sheets as shielding material.

In comparing the relative merits of copper and iron sheets for shielding purposes, it is to be remembered that at very low frequencies low value eddy currents are set up, so shielding due to their action must necessarily be small. However, iron sheets will still give shielding because of the high permeability, so that we might expect iron to be better at low frequencies and copper at the higher frequencies. Fig. 131 shows this to be the fact; it happens that for the grade of iron tested, copper and iron are about equally effective for 1000 cycles per second. Below this frequency iron is better.

CHAPTER III

LAWS OF OSCILLATING CIRCUITS

Discharge of a Condenser through an Inductance and Resistance in Series.—Practically all radio sets which send out damped (or discontinuous) waves generate the high-frequency currents required by charging up a condenser from a suitable source of power, then letting this condenser discharge through an inductance in series with a spark gap. In general the oscillatory power so generated is transferred by coupling of some kind to another circuit from which it is radiated. The investigation of the form of oscillatory current in these coupled circuits will be taken up later in this chapter, we shall first investigate the discharge of a condenser in the single circuit.

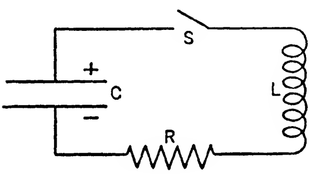


FIG. 1.—The charged condenser C will discharge through L and R when switch S is closed.

In Fig. 1 is shown the circuit; the condenser C charged to voltage E is to be connected to the circuit consisting of L and R in series when the switch S is closed. The switch is not used in the actual radio circuit, a spark gap performing its function, but the resistance of the spark gap somewhat complicates the analysis so that its action is deferred until a later paragraph.

It will be supposed at first that the condenser has no leakage; the equation of reactions of the circuit after the switch is closed is,

$$L\frac{di}{dt}+Ri+v=0,$$

v being the voltage across the condenser at any instant. Then

$$L\frac{d^2i}{dt^2}+R\frac{di}{dt}+\frac{dv}{dt}=0.$$

But we know that

$$i=C\frac{dv}{dt}.$$

so we have

$$L\frac{d^2i}{dt^2}+R\frac{di}{dt}+\frac{i}{C}=0,$$

or

$$\frac{d^2i}{dt^2}+\frac{R}{L}\frac{di}{dt}+\frac{i}{LC}=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The solution of a differential equation of this kind is obtained by an "intelligent guess." It is evident that i must be a function of t and furthermore that this function must be of such a form that the second derivative of the function plus the first derivative multiplied by R/L plus the function itself multiplied by $1/LC$ must add up to zero. By trial we find that if the current is of the form

$$i=A\,\epsilon^{mt}$$

Eq. (1) will probably be satisfied. Using this function we have

$$\frac{di}{dt}=mA\,\epsilon^{mt},$$

and

$$\frac{d^2i}{dt^2}=m^2A\,\epsilon^{mt}.$$

Substituting these values in Eq. (1), we get

$$A\,\epsilon^{mt}\left(m^2+\frac{R}{L}m+\frac{1}{LC}\right)=0. \qquad (1a)$$

As no useful solution is obtained by putting $A=0$, we use the condition

$$m^2+\frac{R}{L}m+\frac{1}{LC}=0.$$

There are two roots to this equation either of which will satisfy it. As Eq. (1) involves the second derivative of i we know there must be two independent solutions for i and these two values of m which we call m_1 and m_2, permit the two required solutions being written. The complete solution of Eq. (1) is the sum of the two particular solutions, so we write as the complete solution

$$i=A_1\epsilon^{m_1t}+A_2\epsilon^{m_2t}, \qquad (2)$$

where

$$m=-\frac{R}{2L}\pm\sqrt{\frac{R^2}{4L^2}-\frac{1}{LC}}=-\alpha\pm\beta$$

So we have

$$i=\epsilon^{-\alpha t}(A_1\epsilon^{\beta t}+A_2\epsilon^{-\beta t}) \qquad (3)$$

As initial conditions we have, at the instant the switch is closed ($t=0$), $i=0$, so from (3)

$$0=A_1+A_2. \qquad (4)$$

Also if

$$i=0,\ Ri=0,\ \text{so}\ \left(\frac{di}{dt}\right)_{t=0}=-\frac{E}{L},$$

Using (3) we get

$$\frac{di}{dt} = -\alpha\epsilon^{-\alpha t}(A_1\epsilon^{\beta t} + A_2\epsilon^{-\beta t}) + \epsilon^{-\alpha t}(\beta A_1\epsilon^{\beta t} - \beta A_2\epsilon^{-\beta t}),$$

and at $t=0$ this gives

$$\frac{di}{dt} = -\alpha(A_1 + A_2) + (\beta A_1 - \beta A_2) = (\beta - \alpha)A_1 - (\beta + \alpha)A_2,$$

so

$$-\frac{E}{L} = (\beta - \alpha)A_1 - (\beta + \alpha)A_2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Solving (4) and (5) for A_1 and A_2 we get

$$A_1 = -\frac{E}{2\beta L} \text{ and } A_2 = +\frac{E}{2\beta L},$$

which values substituted in Eq. (3) give

$$i = -\frac{E}{2\beta L}\epsilon^{-\alpha t}(\epsilon^{\beta t} - \epsilon^{-\beta t}), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

in which

$$\alpha = \frac{R}{2L} \text{ and } \beta = \sqrt{\alpha^2 - \frac{1}{LC}}.$$

The quantity α is always real, which means that the amplitude of the current continually decreases with increase of time. The quantity $(\epsilon^{\beta t} - \epsilon^{-\beta t})$, which determines the form of the current, while it is decaying, depends for its value on the quantity β; this may be either real or imaginary, according as α^2 is greater or less than $\frac{1}{LC}$. The form of the current will be analyzed for the three conditions—

$$\alpha^2 > \frac{1}{LC}, \qquad \alpha^2 = \frac{1}{LC}, \qquad \alpha^2 < \frac{1}{LC}.$$

Case 1. $\alpha^2 > \frac{1}{LC}$.

In this case β is a real quantity so we have

$$i = -\frac{E}{\beta L}\epsilon^{-\alpha t}\left(\frac{\epsilon^{\beta t} - \epsilon^{-\beta t}}{2}\right) = -\frac{E}{\beta L}\epsilon^{-\alpha t}\sinh \beta t. \quad . \quad . \quad . \quad . \quad (7)$$

The negative sign indicates that the effect of the current is to decrease E, i.e., to release the charge on the condenser—as to whether or not current is actually positive or negative depends upon the polarity of charge on the condenser assumed positive.

The form of current in this case is shown in Fig. 2; the lines properly marked give the two terms $\epsilon^{-\alpha t}$ and $\sinh \beta t$. The figure is drawn to scale for $E=100$ volts, $C=10\mu f$, $L=0.20$ henry, and $R=500$ ohms. By calculation we find $\alpha=1250$ and $\beta=1030$.

The maximum current is reached at a time calculated from putting the first derivative of Eq. (7) equal to zero.

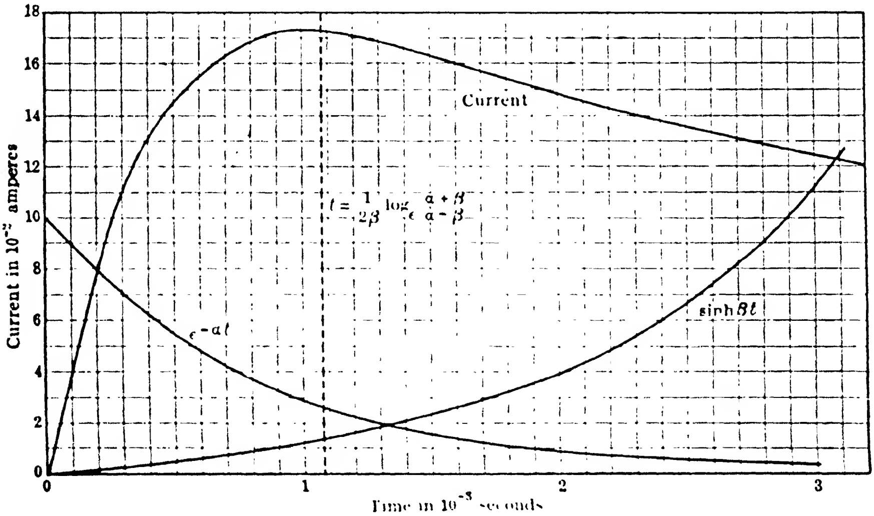


FIG. 2.—Calculated discharge current when the R of Fig. 1 is too high for oscillatory discharge.

This results in the equation

$$\epsilon^{2\beta t}=-\frac{\beta+\alpha}{\beta-\alpha}$$

or

$$t=\frac{1}{2\beta}\log_\epsilon\frac{\alpha+\beta}{\alpha-\beta}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Fig. 3 shows two oscillograms of a condenser discharge current for the case just analyzed.

CASE 2. $\alpha^2=\frac{1}{LC}$.

In this case we have $\beta=0$. We write the current in the form,

$$i=-\frac{E}{2L}\epsilon^{-\alpha t}\left(\frac{\epsilon^{\beta t}-\epsilon^{-\beta t}}{\beta}\right),$$

the value of the expression in the parenthesis being indeterminate. We evaluate it by differentiation and get,

$$\left(\frac{\dfrac{d}{d\beta}(\epsilon^{\beta t}-\epsilon^{-\beta t})}{\dfrac{d}{d\beta}(\beta)}\right)_{\beta=0}=\left(\frac{t(\epsilon^{\beta t}+\epsilon^{-\beta t})}{1}\right)_{\beta=0}=2t.$$

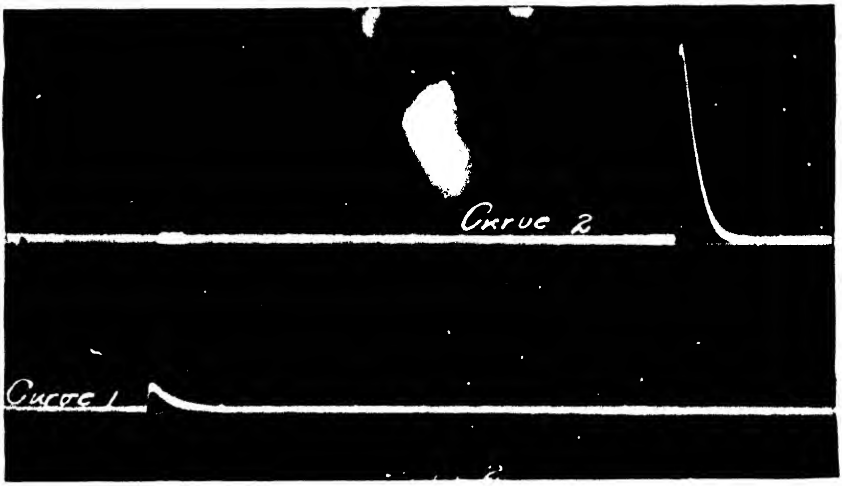


FIG. 3.—Oscillograms of discharges similar to Fig. 2.

We find the same result by expanding $\epsilon^{\beta t}$ and $\epsilon^{-\beta t}$ into series—thus

$$\epsilon^{\beta t}=1+\beta t+\frac{\overline{\beta t}^2}{2}+\frac{\overline{\beta t}^3}{3!}+\ldots$$

$$\epsilon^{-\beta t}=1-\beta t+\frac{\overline{\beta t}^2}{2}-\frac{\overline{\beta t}^3}{3!}+\ldots$$

so

$$\epsilon^{\beta t}-\epsilon^{-\beta t}=2\beta t+\frac{\overline{\beta t}^3}{3}+\ldots$$

Dividing by β we get

$$\frac{\epsilon^{\beta t}-\epsilon^{-\beta t}}{\beta}=2t+\frac{\beta^2t^3}{3}+\ldots$$

and when $\beta=0$ this reduces to $2t$.

Hence in this case the equation for the discharge current is,

$$i = -\frac{Et}{L}\epsilon^{-\alpha t}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (9)$$

The graph of such a discharge current is shown in Fig. 4 for $E=100$ volts, $C=10\ \mu f$, $L=$ 0.20 henry, and $R=282.3$ ohms.

The time at which maximum current occurs is obtained as outlined for the previous case and yields the condition that,

$$t=\frac{1}{\alpha}. \quad . \quad . \quad . \quad . \qquad (10)$$

For the conditions given this time is 0.001416 second after closing the switch.

Fig. 5 shows an oscillogram of such a critically damped circuit; the time scale on the lower part of the film permits the validity of Eq. (10) to be checked.

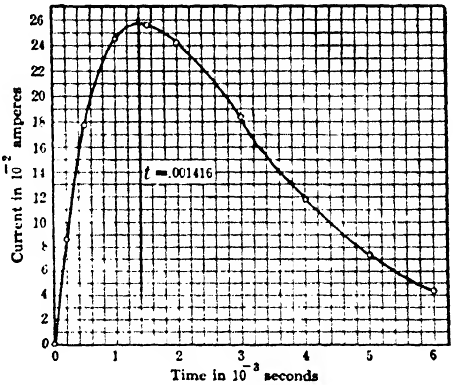


FIG. 4.—Discharge in a circuit in which R has the minimum possible value without permitting oscillatory discharge.

CASE 3. $\alpha^2 < \frac{1}{LC}$.

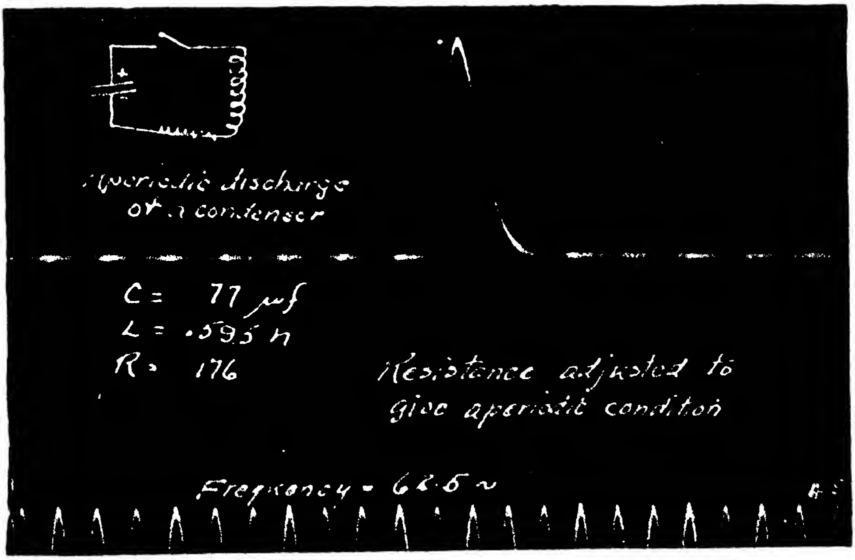


FIG. 5.—Oscillogram of current for conditions assumed in Fig. 4.

In this case β becomes the square root of a negative quantity, and we write it

$$\beta=\sqrt{\alpha^2-\frac{1}{LC}}=\sqrt{(-1)\left(\frac{1}{LC}-\alpha^2\right)}=j\sqrt{\frac{1}{LC}-\alpha^2}=j\omega.$$

Then from Eq. (6)

$$i=-\frac{E}{2j\omega L}\,\epsilon^{-\alpha t}(\epsilon^{j\omega t}-\epsilon^{-j\omega t})=-\frac{E}{\omega L}\,\epsilon^{-\alpha t}\left(\frac{\epsilon^{j\omega t}-\epsilon^{-j\omega t}}{2j}\right)$$

$$=-\frac{E}{\omega L}\,\epsilon^{-\alpha t}\sin\omega t \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

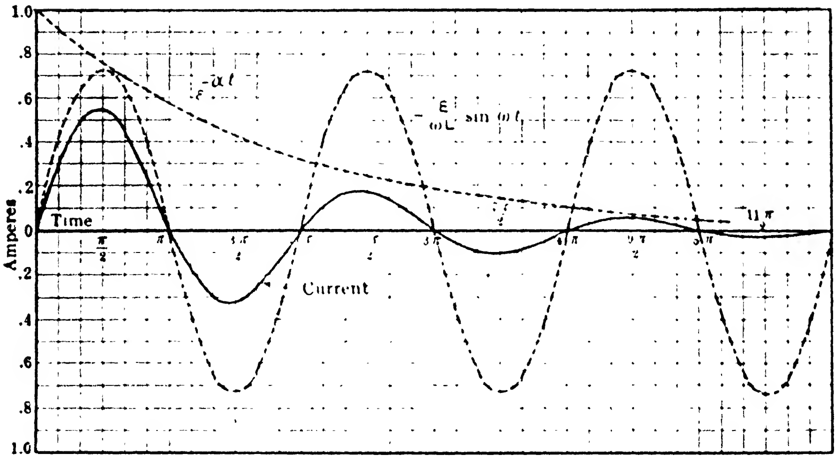


FIG. 6.—Value of R of Fig. 1 reduced sufficiently to permit the ordinary oscillatory discharge, giving a "damped sine wave."

The current in this case is oscillatory, its frequency being fixed by the value of ω; the term $\epsilon^{-\alpha t}$ represents the decay of the current, and the theoretical maximum value of the current is given by $\frac{E}{\omega L}$.

In Fig. 6 are plotted, in dotted lines, each of the terms of Eq. (11) for a circuit of $E=100$ volts, $C=10\ \mu f$, $L=0.20$ henry, and $R=50$ ohms.

We have

$$\alpha=\frac{R}{2L}=\frac{50}{2\times0.20}=125$$

$$\omega=\sqrt{\frac{1}{LC}-\alpha^2}=\sqrt{\frac{10^6}{0.2\times10}-125^2}=695.$$

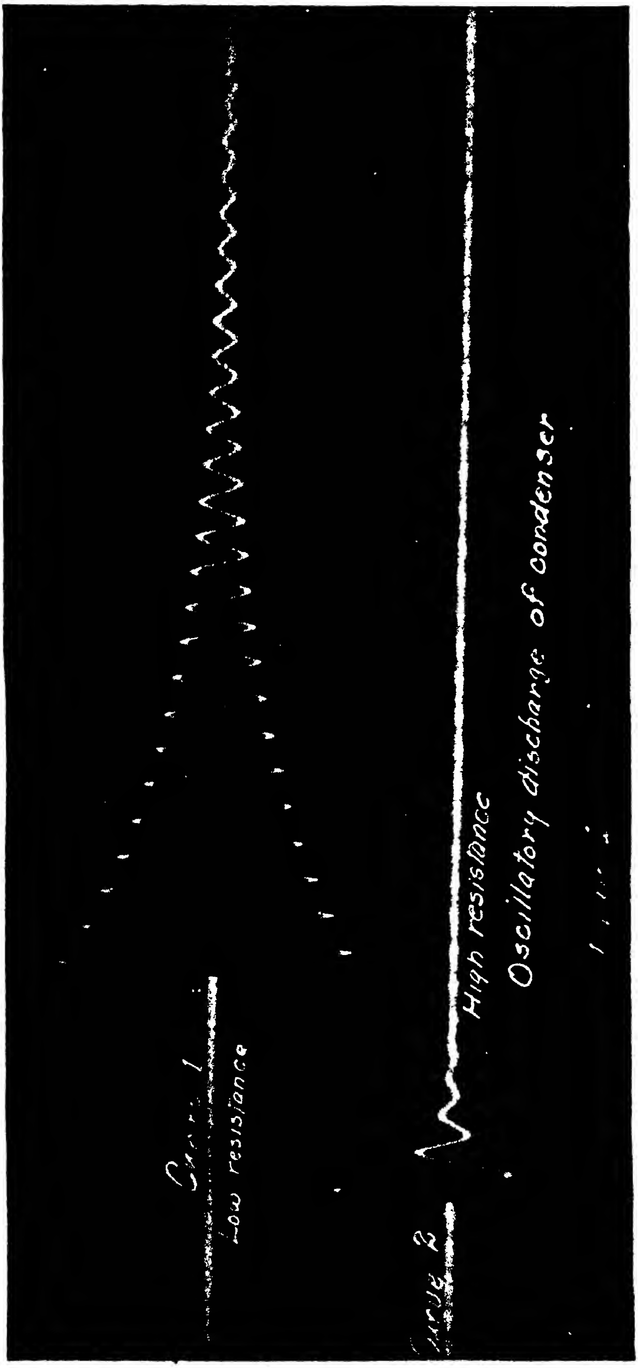


Fig. 7.—Oscillograms of oscillatory discharge for two values of resistance.

But $\omega=2\pi f$, hence

$$f=\frac{695}{2\pi}=110.5 \text{ cycles per second.}$$

The actual equation for the current is then,

$$i=-\frac{100}{695\times0.2}\epsilon^{-125t}\sin(2\pi\, 110.5\, t).$$

This curve is shown in the full lines of Fig. 6. It is generally called a "damped sine wave," the term $\epsilon^{-\alpha t}$ giving the damping. In Fig. 7 is shown the oscillogram of a damped sine wave showing how the actual current is of the form indicated by Eq. (11) for two values of resistance.

Effect of Condenser Leakage.—In case the condenser has appreciable leakage the solution takes a slightly different form. The circuit is now as shown in Fig. 8; the energy stored in the condenser when the switch is closed is partially consumed in the series resistance R, partially consumed in the leak resistance r, and the rest transformed into magnetic energy in the coil; then the magnetic energy in the coil is transformed back to electrostatic energy in the recharged condenser, but during the transformation more of the energy is wasted in R and r.

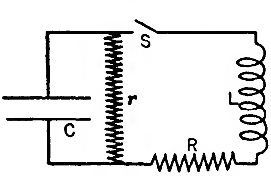


FIG. 8.—Oscillatory circuit in which the condenser is "leaky."

It was shown in Chapter II, Eq. (36), that a shunt resistance might be replaced by a resistance in series, and such might be done in Fig. 8. It would at once reduce the circuit to that of Fig. 1 and thus the problem would be solved.

So far as decrement is concerned this might be done; the power loss in the actual shunt resistance being the same as that in its equivalent series resistance we would expect the damping to be the same for either arrangement. However, there are some differences in behavior in a circuit having leakage and one having none, so we will solve the problem with the leak resistance left as such. The differential equation of the circuit becomes

$$L\frac{di}{dt}+Ri+v=0 \quad \text{or} \quad L\frac{d^2i}{dt^2}+R\frac{di}{dt}+\frac{dv}{dt}=0.$$

We have $i_c=C\frac{dv}{dt}$, and $i_c=i+i_g$ where $i_g=vg$, g being equal to $\frac{1}{r}$.

Now, in magnitude, $v=L\frac{di}{dt}+Ri$, so that $i_g=gL\frac{di}{dt}+gRi$.

Substituting then $\frac{dv}{dt}=\frac{i_c}{C}$ and using the value of i_g just obtained, we get

$$LC\frac{d^2i}{dt^2}+(RC+gL)\frac{di}{dt}+(1+gR)i=0,$$

which may be written

$$\frac{d^2i}{dt^2}+\left(\frac{R}{L}+\frac{g}{C}\right)\frac{di}{dt}+\left(\frac{1+gR}{LC}\right)i=0. \quad . \quad . \quad . \quad . \quad (12)$$

This equation is similar in form to (1) and its solution is of exactly the same form. For this case, however, we have

$$\alpha=\frac{R}{2L}+\frac{g}{2C}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

and as

$$\beta=\sqrt{\frac{1}{4}\left(\frac{R}{L}+\frac{g}{C}\right)^2-\frac{1+gR}{LC}},$$

we find by combining terms that

$$\beta=\sqrt{\frac{1}{4}\left(\frac{R}{L}-\frac{g}{C}\right)^2-\frac{1}{LC}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

The three cases considered in the previous section occur also for this circuit; the conclusions reached are the same, except where previously $\frac{R}{2L}$ determined the damping, we now have the quantity $\left(\frac{R}{2L}+\frac{g}{2C}\right)$.

The conditions for oscillations or no oscillation are affected by the condenser leakage in a manner not to be expected; with no leakage the non-oscillatory condition is reached when

$$\frac{R}{2L}=\frac{1}{\sqrt{LC}}$$

and for the leaky condenser the criterion is

$$\left(\frac{R}{2L}-\frac{g}{2C}\right)=\frac{1}{\sqrt{LC}}.$$

That is, a circuit which has sufficient series resistance to be critically damped may become oscillatory if sufficient leakage is introduced across the condenser.

For the circuit considered in the previous section the non-oscillatory condition was reached when R was adjusted for 282.3 ohms; we then had $\alpha=707$. If we now shunt the condenser by a leak resistance of 1000 ohms we have

$$\alpha=\frac{282.3}{2\times0.2}+\frac{10^5}{2\times10^3}=757;$$

that is, greater than before, but we now have an oscillatory circuit because $\left(\frac{R}{2L}-\frac{g}{2C}\right)$ is less than $\frac{1}{\sqrt{LC}}$. Thus we have the unexpected phenomenon of increased damping changing a non-oscillatory circuit to an oscillatory one. Fig. 9 shows the three currents for the circuit with leaky condenser, as in Fig. 8.

The frequency of the free oscillation is lowered by the series resistance of the circuit, but it is raised by the effect of shunt resistance until this

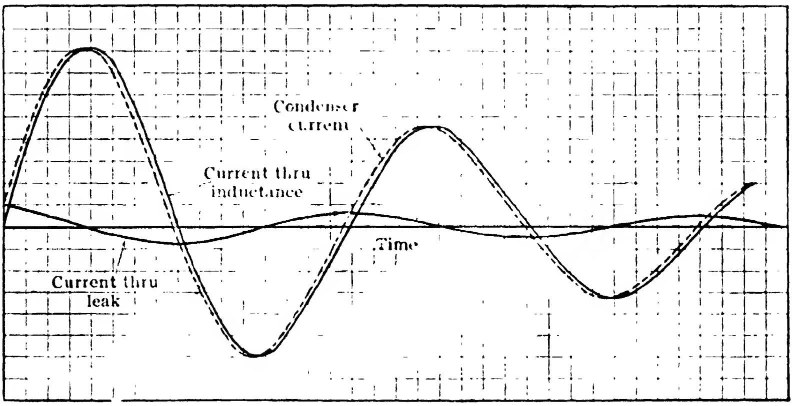


FIG. 9.—Calculated currents for circuit depicted in Fig. 8.

shunt resistance reaches the value such that $\frac{R}{L}=\frac{g}{C}$. If the shunt, or leak resistance, is made still less the frequency will again decrease; from this it is seen that the effect of a leak resistance (with no series resistance) is to increase the damping and increase the period, just as is the case for a series resistance above, but that when both are present the damping is increased by an amount depending on the sum of the series resistance and leak resistance, but that the effect of these two on the period is subtractive, and that a certain relation between them suffices for complete neutralization, so that the natural period is the same as it would be if the circuit had no dissipative reactions at all.

Frequency—Wave Length.—In the previous paragraph the frequency of an oscillatory discharge was shown to be fixed by the damping, inductance, and capacity. The effect of the damping constants in the frequency is, in ordinary radio circuits, so small that it can be neglected without appreciable error, so that this formula for frequency of an oscillatory circuit becomes

$$\omega = 2\pi f = \frac{1}{\sqrt{LC}},$$

or

$$f = \frac{1}{2\pi\sqrt{LC}}. \qquad (15)$$

In the formula f is in cycles per second, L in henries, and C in farads. Now in radio circuits the values of L and C are more generally measured in micro-units and the formula becomes,

$$f = \frac{10^6}{2\pi\sqrt{LC}}, \qquad (16)$$

where L is in microhenries and C in microfarads.

It has been customary to use the term *wave length* in radio literature, instead of frequency. When an antenna is excited by an oscillatory current of frequency f it sends out over the earth's surface electromagnetic waves which travel out from the antenna with the velocity of light, i.e., 3×10^8 meters per second. In any wave phenomenon the frequency and wave length (always designated in radio by the symbol λ) are connected by the formula

$$f\lambda = V, \qquad (17)$$

where V is the velocity of travel of the waves. We therefore find for the value of wave length of these electromagnetic radiations

$$\lambda = \frac{V}{f} = \frac{3 \times 10^8 \times 2\pi\sqrt{LC}}{10^6} = 1885\sqrt{LC}. \qquad (18)$$

In this formula λ is given in meters, L in microhenries, and C in microfarads.

Relation of Current and Voltage in Oscillatory Circuits.—The equation for the discharge of a condenser for the ordinary condition (Case 3, p. 290, Eq. 11) is

$$i = -\frac{E}{\omega L} \epsilon^{-\alpha t} \sin \omega t,$$

in which

$$\alpha = \frac{R}{2L},$$

and we have said that in the ordinary radio circuit ω is approximately equal to

$$\frac{1}{\sqrt{LC}}.$$

Eq. (11) therefore becomes,

$$i=-E\sqrt{\frac{C}{L}}\,\epsilon^{-\frac{Rt}{2L}}\sin\frac{t}{\sqrt{LC}}. \quad \ldots \ldots \quad (19)$$

The maximum current occurs one-quarter of a cycle after closing the switch, nearly; the effect of the damping term $\epsilon^{-\frac{Rt}{2L}}$ is to make the current a maximum shortly before the quarter cycle interval. The value of this current, neglecting the small effect of the damping for one quarter cycle, is equal to $E\sqrt{C/L}$.

Now this could have been predicted from the consideration of energy in the circuit; before the switch is closed all the energy is in the condenser and is equal to $\frac{CE^2}{2}$. One quarter cycle after closing the switch the voltage across the condenser is zero, so that all the energy must be in the coil, hence we may put

$$\frac{CE^2}{2}=\frac{LI^2}{2},$$

or

$$I=E\sqrt{\frac{C}{L}},$$

as we had before.

In an oscillatory circuit there is a certain amount of energy oscillating back and forth from coil to condenser, and being wasted during the transfer. The frequency of transfer will be the same no matter what the relative value of L and C, so long as their product is constant. It is sometimes desired to establish resonance in a circuit and keep the voltage low; in such a case a low value of L and correspondingly high value of C should be chosen. In radio receiving circuits, however, it is generally desired to obtain as high a voltage as possible; this is done by using as low a value of C as possible (sometimes as low as 100 micro-microfarads) and a correspondingly high value of L.

Damping and Decrement.—In Eq. (19) the factor $\epsilon^{-\frac{Rt}{2L}}$ represents a logarithmic decrease in the amplitude of the current; the value of $\frac{R}{2L}$ is called the *damping coefficient* of the circuit. For the average radio circuit this coefficient is of the order of 1000 to 10,000, being greater the shorter the wave length of the set. The damping coefficient multiplied by the time of one cycle is called the *logarithmic decrement* or merely the *decrement* of the circuit.

If we write the values of successive maxima of current (maxima in the same direction) we shall have from Eq. (19), calling T the period of oscillation,

$$I_1 = -E\sqrt{\frac{C}{L}}\epsilon^{-\frac{R}{2L}\frac{T}{4}} = I_0\epsilon^{-\frac{R}{2L}\frac{T}{4}}$$

$$I_2 = I_0\epsilon^{-\frac{R}{2L}\left(\frac{T}{4}+T\right)} = I_1\epsilon^{-\frac{R}{2L}T}$$

$$I_3 = I_0\epsilon^{-\frac{R}{2L}\left(\frac{T}{4}+2T\right)} = I_2\epsilon^{-\frac{R}{2L}T}, \text{ etc.}$$

From these we get

$$\frac{I_1}{I_2} = \frac{I_2}{I_3} = \frac{I_3}{I_4} \text{ etc.} = \epsilon^{\frac{R}{2L}T} = \epsilon^{\delta},$$

where δ is the decrement of the circuit. As $T = \frac{1}{f}$ it is evident that

$$\delta = \frac{R}{2fL}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

An ordinary radio set has a decrement about 0.1; the large stations may have a decrement as low as 0.02, while the upper limit for a transmitting station is 0.2, this being fixed legally as the maximum decrement a spark station is allowed. As is evident from Eq. (20), the decrement of a circuit depends directly upon the resistance of the circuit, this resistance being interpreted in the broadest sense as suggested on page 163. In transmitting stations the ground resistance of the antenna is likely to be very important in its effect on the decrement. The decrement measured from the upper curve of the film of Fig. 7 was 0.150, while that calculated from the constants of the circuit was 0.152.

In a continuous-wave transmitting station the source of high-frequency power maintains a constant amplitude to the successive cycles and the station is said to have a zero decrement; in certain receiving circuits using a vacuum tube for receiver, the effective resistance of the circuit is made to approach zero as nearly as desired, thus making the decrement of the receiving set approach zero. As explained in Chapter I, pp. 79–87, the decrement is the important factor in determining the selectivity of a receiving circuit, as it determines the sharpness of resonance.

Decrement Determined by Energy Waste per Cycle.—The decrement may be defined as the ratio of the energy dissipation per cycle to the energy transferred during the same interval of time. Neglecting the small change in value of maximum current during one cycle we have:

$$\text{Energy dissipated per cycle} = \frac{RI^2}{2f},$$

where I is the maximum value of current.

Suppose we consider the cycle to begin when I has maximum positive value and all the energy is in the coil, this energy being equal to $\frac{LI^2}{2}$. Now during one cycle this energy flows from the coil to the condenser, back to the coil (when I goes through its values of opposite polarity) back to the condenser and then back to the coil. The energy makes two complete transfers through the circuit so that the amount of energy transfer during one cycle is

$$2 \times \frac{LI^2}{2} = LI^2.$$

Hence,
$$\frac{\text{Energy dissipated}}{\text{Energy transferred}} = \frac{\frac{RI^2}{2f}}{LI^2} = \frac{R}{2fL} = \delta.$$

If the above analysis were carried through rigorously (taking account of decrease of I during the cycle), it would be found that the above relation for δ is correct.

Current, Voltage, and Energy in a Damped Wave.—During the decay of a wave train the corresponding maximum values of electrostatic energy of the condenser and electromagnetic energy of the coil remain practically equal; the voltage across the condenser goes through the same changes as does the current through the inductance. Using the relation between the voltage across the condenser and the current in the circuit

$$e_C = \int \frac{i\,dt}{C},$$

we get from Eq. (11) the approximate relation

$$e_C = E\epsilon^{-\alpha t} \cos \omega t.$$

At any instant, therefore, the energy in the condenser is

$$w_C = \frac{CE^2}{2}\epsilon^{-2\alpha t} \cos^2 \omega t.$$

And we have for the energy in the coil,

$$w_L = \frac{LI_0^2}{2}\epsilon^{-2\alpha t} \sin^2 \omega t.$$

If we substitute for I_0 its value determined above $\left(E\sqrt{\frac{C}{L}}\right)$ and then add we get

$$w_C + w_L = w = \frac{CE^2}{2}\epsilon^{-2\alpha t}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

From the equation we see that the original energy stored in the condenser, $\frac{CE^2}{2}$, undergoes a logarithmic decay, with damping coefficient twice that of the current.

The curves of current, voltage, and energy are plotted in Fig. 10; the total energy is obtained by adding the corresponding instantaneous values

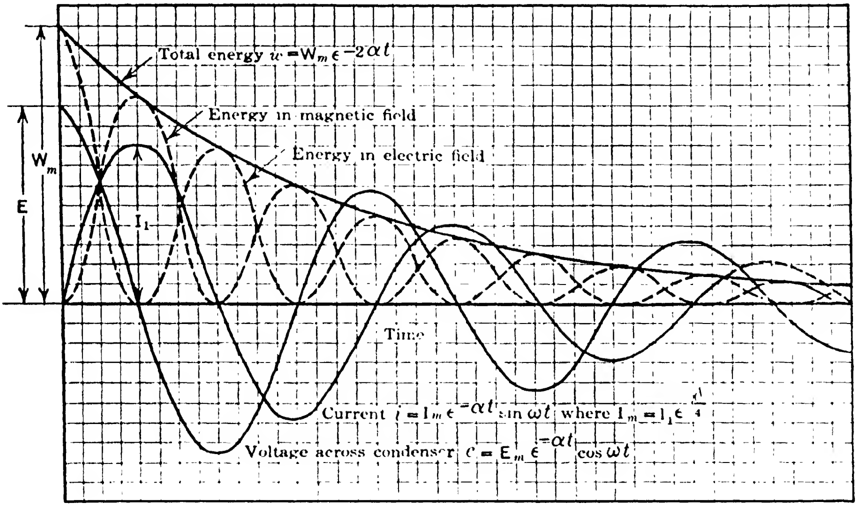


FIG. 10.—Conventional energy curve representation in an oscillatory circuit.

of the magnetic and electric energy. In reality the current and voltage are not exactly 90° out of phase, due to the effect of the resistance of the circuit so that the addition of the two components of energy does not give the smooth exponential curve shown in Fig. 10, but a wavy exponential curve as indicated in Fig. 11. Here a decrement of 0.3 has been assumed, giving a power factor of $\frac{0.3}{\pi} = 0.0955$; the phase difference of E and I is therefore 84.5°. The energy for the electric and magnetic fields no longer adds to give the smooth energy curve of Fig. 10, but indicates that the dissipation of energy from the system is more rapid at certain parts of the cycle than at others. When all the dissipated energy appears in the form of heat in the series resistance (as supposed for Fig. 11), the maximum rate of dissipation corresponds with the time of maximum current as it should; when the current is zero there is no energy being dissipated.

In case the condenser used in the oscillatory circuit is a leaky one, with leak conductance, g, the energy dissipated while an oscillatory current is flowing is used up partly in the series resistance of the circuit and partly in the conductance across the circuit. The rate of energy dissipation in the series resistance is i^2R and the rate of energy dissipation in the leak is e_c^2g. It will be noticed that these two power losses do not have their maxima at the same instant; when e_c is a maximum i is practically zero.

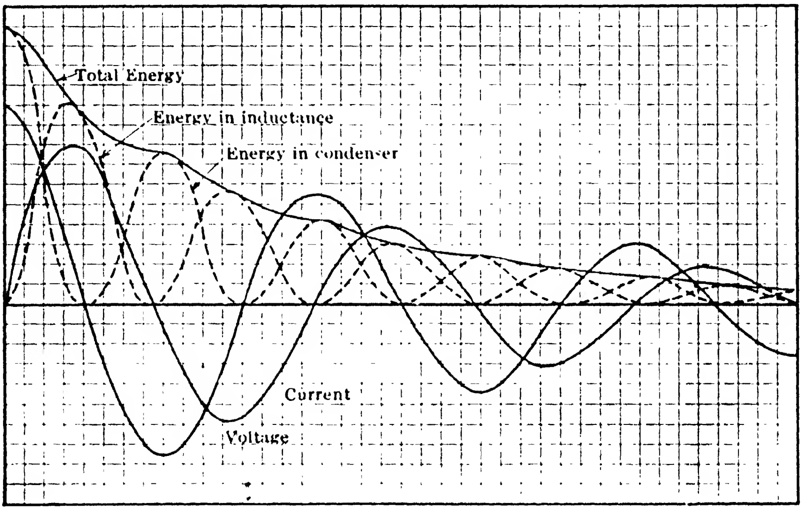


FIG. 11.—Actual energy curve for an ordinary oscillatory circuit.

If the series resistance and shunt resistance are properly proportioned the power dissipated in each will be the same; the proper ratio is obtained by putting

$$I^2R = E^2g = \frac{E^2}{r}.$$

Now we have,

$$I^2 = \omega^2C^2E^2,$$

and so,

$$\omega^2C^2E^2R = \frac{E^2}{r},$$

from which

$$R = \frac{1}{r}\,\frac{1}{\omega^2C^2}.$$

The relation may also be expressed

$$R=\frac{1}{r}\frac{L}{C},$$

or we may also put it in the form,

$$\frac{R}{L}=\frac{g}{C}.$$

Such a proportionality [1] in the series and shunt resistances of the circuit will result in a power consumption in the oscillating circuit which does not fluctuate throughout the cycle as it does when the relation is not maintained. Hence the energy decay in the circuit is not a wavy line as given in Fig. 11, but a smooth logarithmic curve as given in Fig. 10.

It is interesting to note that this proportionality of series and shunt resistances is the same as is required to make the natural period of oscillation the same as if no dissipative reactions were present in the circuit. The natural period of such a circuit was given in Eq. (14); it is seen that if

$$\frac{R}{L}=\frac{g}{C},$$

then

$$f=\frac{1}{2\pi\sqrt{LC}}.$$

Oscillatory Discharge through a Spark Gap.—If the oscillating circuit contains a spark gap the current is not of the form indicated by Eq. (11), because of the influence of the gap; as pointed out on page 204 the resistance of a given spark gap is not constant but depends upon the current flowing through it. The resistance of the gap is smaller the higher the amplitude of current through it, as is more or less evident from the appearance of a spark. The greater the current through the gap the larger is the cross-section of the hot, ionized gas conducting the current, and the more intensely is it ionized, both of these effects lowering the gap resistance.

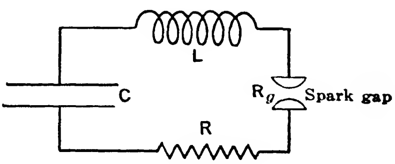


FIG. 12.—Oscillatory circuit in which a spark gap is used.

[1] The same proportionality has a peculiar significance when applied to long conductors, such as telephone lines. It was first pointed out by O. Heaviside that such relation between the various constants of a line give so-called "distortionless" transmission of speech waves.

When a charged condenser discharges through the circuit represented in Fig. 12, the equation of discharge is

$$L\frac{di}{dt}+(R+R_g)i+v=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

where R_g is the resistance of the spark gap. If R_g can be written as a simple function of the current this equation can be solved, but it is quite likely that the function is a complicated one, depending not only on the magnitude of current, but on the frequency of the oscillations.

The value of R_g undoubtedly varies a great deal throughout the cycle, but these variations can have much less effect on the magnitude and shape of the current than might be supposed. The resistance reaction is the only reaction limiting the value of the current of fundamental frequency $\left(\text{approximately } \frac{1}{2\pi\sqrt{LC}}\right)$ but any upper harmonics which the cyclically varying resistance might tend to produce in the circuit would be opposed by a reactance several hundred times as great as the resistance, because the inductance and capacity reactions balance only for the fundamental frequency. Hence the cyclical change in resistance may be neglected in so far as its affects the solution of the oscillatory current defined by Eq. (22).

By means of a Braun tube oscillograph photographs have been taken of the oscillations in such a circuit as that of Fig. 12, and the commonly accepted interpretation of these photographs is that the decay of current is linear with respect to time instead of exponential, as given in Eq. (11).

On the assumption that the resistance of the circuit did not affect the frequency and using the experimental data given by Zenneck, J. S. Stone has shown that if the gap resistance is written

$$R_g=\frac{2BL}{A-Bt},$$

A and B being constants, and the other resistance in the circuit is negligible the solution of Eq. (22) becomes,

$$i=-E\sqrt{\frac{C}{L}}\left(1-\frac{R_0t}{2L}\right)\sin\frac{t}{\sqrt{LC}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

where R_0=initial resistance of spark gap. Although not so stated in Stone's paper, this value R_0 must be approximately its resistance at the first current maximum. The other symbols have their ordinary meanings. The solution is really of little importance in radio work, because in no case is the spark gap the controlling factor in a radiating circuit. Such a circuit would use up practically all of its stored energy in heating the spark gap, so it is practice to remove the high-frequency power from the spark-gap circuit as quickly as possible and let it radiate from a circuit which has no gap. Even when the energy is in the spark-gap circuit the resistance of the gap is small compared to the *resistance introduced into this circuit by the coupled antenna circuit* as indicated by Eq. (108), p. 125.

The current of Eq. (23) is different from that of Eq. (11) in that the successive maxima have a constant *difference*, whereas those of Eq. (11) have a constant *ratio*. Thus the linear damping gives a wave train (group of oscillations) which has a definite end, whereas the logarithmic decrement never actually reduces the current to zero.

Number of Waves in a Train.—When an oscillatory current is expressible by Eq. (19) it is evident that there must be an infinite number of cycles per discharge of the condenser (or per wave train); the damping factor $\epsilon^{-\frac{Rt}{2L}}$ makes the current approach zero value, but theoretically it never reaches the zero value. It is customary in radio practice to say that a wave train has ended when the current amplitude has fallen to 1 per cent of its initial amplitude; this means of course that the energy remaining in the circuit is only (1 per cent)$^2=0.0001$ of its initial value.

The successive maxima of current are related by the equation

$$I_n=I_1\epsilon^{-(n-1)\delta},$$

and if

$$\frac{I_1}{I_n}=100,$$

we have

$$\log_\epsilon 100=(n-1)\delta,$$

or

$$n=\frac{4.6+\delta}{\delta}. \qquad (24)$$

Thus if the decrement of an antenna is 0.05 there will be $\frac{4.6+0.05}{0.05}=93$ complete cycles before the energy has been sufficiently dissipated to reduce the current to 1 per cent of its initial value.

In the case of a linearly damped wave train the number of waves is very few, principally because if the gap resistance is so large that the rest of the circuit resistance is negligible (a necessary assumption for linear decrement) the decrement is of such a high value that the wave train cannot have more than perhaps five to ten cycles before it is completely finished.

It is interesting to note that at the end of a logarithmically decaying wave train the condenser in the circuit is completely discharged, while the circuit with linear decrement may leave a considerable charge in the condenser at the end of a wave train. When the resistance of the spark gap becomes too high, towards the end of the train when the current is small, the gap opens (probably at a time when the current is zero), leaving the condenser charged to some appreciable voltage.

Effective Value of Current in a Damped Wave Train.—The effective value of a damped sine wave may be obtained by integration of the heating effect of the current

$$I^2=\frac{1}{T}\int_0^T (I_0\epsilon^{-\alpha t}\sin \omega t)^2 dt.$$

Evidently the value of this integral will vary with the length of time over which the integration is extended, and is to this extent indeterminate in value. As in practice one wave train follows another in rapid succession we are really interested in an integral of the form,

$$I^2=N\int_0^\infty (I_0\epsilon^{-\alpha t}\sin \omega t)^2 dt=NI_0^2\int_0^\infty \epsilon^{-2\alpha t}\sin^2 \omega t\, dt,$$

where N is the number of discharges per second.

This may be integrated by standard methods, after expressing the sine in the form of exponentials, and these yield the solution,

$$I^2=NI_0^2\frac{\omega^2}{4\alpha(\alpha^2+\omega^2)}.$$

Now we have

$$\omega=2\pi f$$

and

$$\alpha=f\delta,$$

so

$$\frac{\omega^2}{\alpha^2+\omega^2}=\frac{(2\pi f)^2}{(f\delta)^2+(2\pi f)^2}=\frac{1}{1+\left(\frac{\delta}{2\pi}\right)^2}.$$

So

$$I^2=\frac{NI_0^2}{4f\delta}\,\frac{1}{1+\left(\frac{\delta}{2\pi}\right)^2}.$$

Now, $\left(\frac{\delta}{2\pi}\right)^2$ is, for the most radio circuits, negligible compared to unity. If we write in place of the theoretical value of current I_0^2, its equal, $E^2\frac{C}{L}$, and put $\frac{1}{1+\left(\frac{\delta}{2\pi}\right)^2}=1$, we get the expression,

$$I^2=\frac{NE^2C}{4f\delta L}=\frac{NE^2C}{2R},$$

or

$$I=E\sqrt{\frac{NC}{2R}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

We could have obtained the same result by noticing that all the energy stored in the condenser is transformed into heat or radiation by the oscillatory current. So we can put

$$N\frac{CE^2}{2}=I^2R,$$

or

$$I=E\sqrt{\frac{NC}{2R}}$$

as before.

The value of I is what a hot-wire meter in the circuit would indicate; the maximum instantaneous value of the current directly after the discharge starts may be a hundred times as great as the value given by Eq. (25).

Effect of Neighboring Circuits on Frequency and Damping.—If another closed circuit of inductance and resistance is so situated that currents are induced in it by the oscillatory current of the first circuit, the damping of the first circuit is increased and the frequency is increased because of the decrease in inductance. The changes in L and R due to the extra circuit are calculable from Eqs. (95) and (96) of Chapter I; the effect on the decrement is increased not only by the increase in R, but also by the decrease in L.

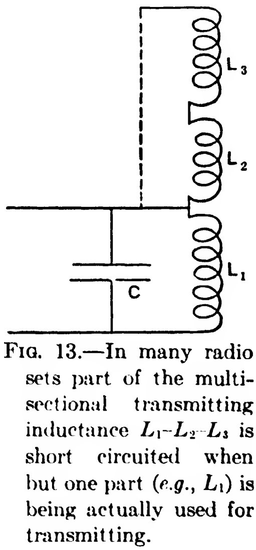


FIG. 13.—In many radio sets part of the multi-sectional transmitting inductance L_1-L_2-L_3 is short circuited when but one part (*e.g.*, L_1) is being actually used for transmitting.

In spite of these effects it is sometimes the practice to intentionally short-circuit part of the coils in a receiving or transmitting set. Thus in Fig. 13 is shown a diagram of such a scheme; the inductance is made in three sections, connected electrically, and also magnetically. When being used for short wave lengths (high frequency) only one section of the inductance, L_1, is connected in series with the condenser, the others being used when longer wave lengths are desired. Now with the connection as shown, the inductance acts like an autotransformer, generating very high voltages at the open end of the coil. This high voltage may cause excessive losses due to both corona and dielectric losses in the insulating supports. Also the voltage generated at the free end may be high enough to break down the coil insulation. To obviate these difficulties the parts L_2 and L_3 are short circuited, as shown by the dotted line, thus increasing the decrement of the L_1C circuit, as noted above. The decrement may in certain cases be even less with L_2 and L_3 short circuited than it would be if they were not short circuited.

Effect of a Neighboring Tuned Circuit on an Oscillatory Discharge.—When an oscillatory discharge takes place in a circuit to which is coupled, either by mutual capacity or mutual inductance, another circuit consisting of inductance and capacity, the form of the current is no longer a simple logarithmic decay, but is much more complicated, the exact form depending upon the coefficient of coupling between the two circuits, the relation between the natural frequencies of the two circuits, the resistances of each, and the type of spark gap used in the discharge circuit.

Before anyone takes up the study of the coupled circuits he should make himself a simple piece of apparatus, which offers the same peculiarities to motion as coupled circuits do to current. This apparatus is shown in Fig. 14; it consists of a board about 10 cm. by 30 cm. with two upright posts about 30 cm. high fastened at the ends of the board. Across the tops of the two posts is fastened a string and from this string are suspended two pendulums, A and B, the lengths of which are readily adjustable and which can be slid along the supporting string so that their points of support a and b, can be separated or brought together.

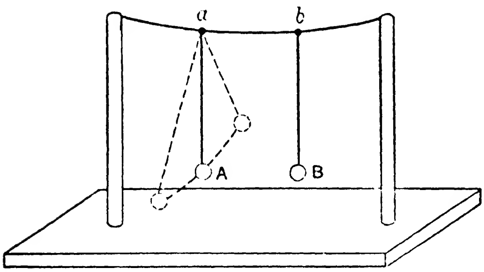


FIG. 14.—Coupled pendulum model.

This simple piece of apparatus is the mechanical analogue of two resonant electrical circuits; each of the pendulums has a natural period of its own and as it swings it tends to make the other pendulum oscillate also.

Suppose that bob A is pulled to one side, bob B being stationary; as A swings sidewise it, of course, pulls its point of support, a, sidewise and thus pulls point b sidewise with it. This motion of point b will gradually set bob B into motion, as the amplitude of motion of B increases that of A decreases and after perhaps twenty or thirty complete vibrations of A its motion will have been reduced to practically zero and that of B will have increased to a maximum, practically the same as the original amplitude of A. This remark holds good only if the lengths and weights of the two pendulums are the same.

Qualitative Analysis of the Pendulum Experiment.—The natural frequency of the pendulum is fixed by its length and the gravitational force; hence to change the natural period of a pendulum it is only necessary to change its length; the mass of the bob itself has no appreciable effect on the natural frequency. It must be noted, however, that the mass of the bob does have a considerable effect on the *amplitude of vibration* for a given energy in the oscillation, in fact for a given energy the amplitude of vibration varies inversely as the square root of the mass of the bob.

The damping, therefore the decrement, of a swinging pendulum is fixed by the ratio of the frictional forces (set up by the motion) to the mass of the bob; an aluminum bob will have considerable greater decrement than a lead bob, the two being the same diameter. Of two bobs of the same material the smaller will have the higher damping because the mass varies as the cube of the diameter and the air friction in the bob approximately as the square of the diameter; the air friction on the string will be the same for both. Hence a small bob, or one of less dense material, will have greater damping than a large heavy one.

The coupling of the two pendulums depends, for a given length of pendulum, on the distance apart of the points of support, a and b, and on the tightness of the supporting string. The farther apart the points of attachment a and b, and the tighter the string the less the coupling of the two pendulums.

The decrement of these pendulums is much less than the decrement of a radio circuit; if it is desired to give the pendulums a greater damping the bobs may be made to swing in a pan of water, or other liquid, or an air damping vane may be fastened to the pendulum string; the closer the vane is placed to the bob the greater will be its damping effect.

By watching the motion of the bobs under various conditions the following approximate deductions may be drawn:

1. For all conditions the motion of either bob is a complex harmonic motion, the amplitude varying periodically from a maximum to a minimum, the average value of the amplitude gradually decreasing.

2. The maximum variation in amplitude occurs in case the two pendulums have the same natural frequency, the minimum amplitude of each pendulum for this case being practically zero.

3. If the two pendulums have the same mass the maximum amplitude of each is nearly the same, if not of the same mass, the lighter bob has the greater maximum amplitude.

4. The period of oscillation for each pendulum (time between successive passages through zero displacement, in the same direction) is practically constant with similar pendulums, the same as the natural period of either pendulum at all times except when the amplitude is going through its minimum values; at this time a sudden reversal of phase takes place in the motion so that the motion gains (or loses) nearly one half a cycle at this time.

5. During the time a pendulum is gaining amplitude its motion lags

nearly 90° behind that of the other pendulum; when its amplitude is decreasing its motion is slightly more than 90° ahead of that of the other pendulum.

6. The amplitude of the first pendulum (the one originally displaced to start oscillations) varies from a maximum to a minimum, the value of this minimum depending upon the relative lengths of the pendulums; for equal lengths the minimum is practically zero, but the minimum increases in value as the ratio of lengths departs from unity value. For all conditions, however, the amplitude of the second pendulum varies from maximum to zero.

7. The time between successive maxima and minima of amplitude depends entirely on the coupling; the tighter the coupling the more rapidly the successive maxima follow one another.

8. Beats (periodic variations in amplitude) always occur unless the coupling is weak and damping is high. In fact practically the only way to prevent beats is to make the damping so high that for the coupling in question the time between beats (as determined for low value of damping) is sufficient to allow nearly complete dissipation of the energy originally put into the first pendulum.

9. If, after the first pendulum has given its energy to the second pendulum (first minimum amplitude of the first pendulum), it is in some way disconnected from the second by cutting its string (or merely by holding the first bob so that it cannot move), the second pendulum will oscillate at its own natural frequency, and with its own decrement, until all the energy originally in the first pendulum has been dissipated by the losses in the second. This condition illustrates the operation of a so-called *quenched spark* transmitting set.

Analysis of the Motion of Coupled Pendulums.—The peculiar motion of each of the oscillating pendulums discussed in the previous paragraph can be produced synthetically, more easily than would be supposed. If we let v_1 and v_2 be the actual velocities of the two bobs, both of changing phase and amplitude we may write,

$$v_1 = V_1\epsilon^{-\alpha_1 t}\sin 2\pi f't + V_2\epsilon^{-\alpha_2 t}\sin 2\pi f''t, \quad . \quad . \quad . \quad . \quad . \quad (26)$$

$$v_2 = V_1\epsilon^{-\alpha_1 t}\sin 2\pi f't + V_2\epsilon^{-\alpha_2 t}\sin (2\pi f''t+\pi), \quad . \quad . \quad . \quad (27)$$

where f' and f'' are lower and higher respectively than the natural period of each pendulum and V_1 and V_2 are maximum velocities of these two component velocities, and α_1 and α_2 are the damping factors of the coupled system for the two frequencies f' and f''. In Fig. 15 are shown the graphs of Eqs. (26) and (27), when applied to pendulums of equal length; the resultant v_1 and v_2 will be recognized at once as the form of the velocity of the two coupled pendulums, v_1 corresponding to the pendulum originally displaced to start vibrations; the reversal of phase at the times of minimum amplitude is well shown in the curves. In these curves the damping has been neglected; for the ordinary pendulum the damping is very small for so short a time as shown in Fig. 15.

After having clearly in mind the phenomena which occur in the pair of coupled pendulums we will analyze mathematically the currents in two coupled electrical circuits and shall find solutions exactly similar to Eqs. (26) and (27).

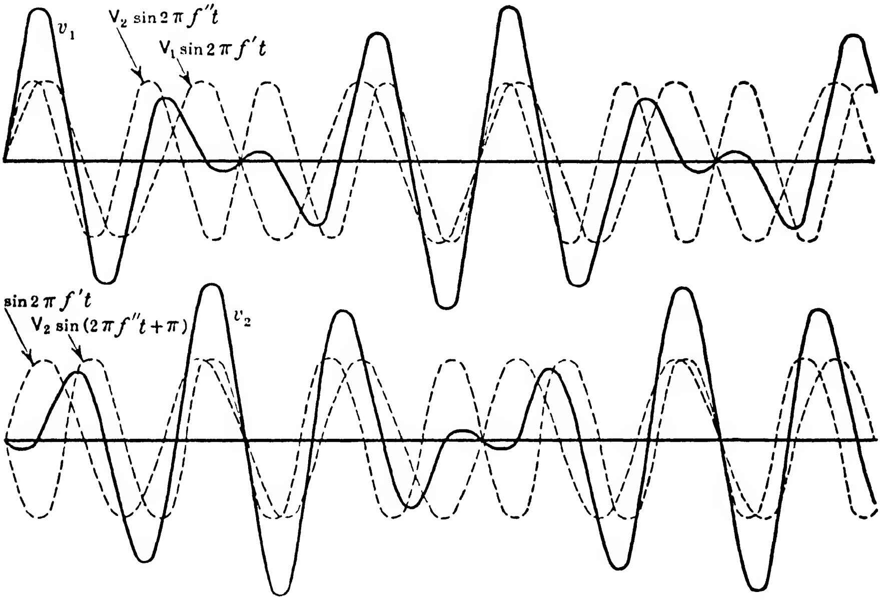


Fig. 15.—Full-line curves show actual motion of the two bobs of Fig. 14 for tight coupling; the dashed lines represent the two sinusoidal components of the actual complex motion.

Analysis of Oscillations in Coupled Circuits.—When the switch S (Fig. 16), is closed currents flow in each circuit and the equation of reactions for each circuit is given by

$$L_1D^2q_1+MD^2q_2+R_1Dq_1+\frac{q_1}{C_1}=0, \quad . \quad . \quad . \quad . \quad . \quad (28)$$

$$L_2D^2q_2+MD^2q_1+R_2Dq_2+\frac{q_2}{C_2}=0, \quad . \quad . \quad . \quad . \quad . \quad (29)$$

where the letters have the meaning shown in Fig. 16, M being the mutual induction between L_1 and L_2 and q_1 and q_2 being the charges on condensers C_1 and C_2 respectively. D stands for $\frac{d}{dt}$ and D^2 for $\frac{d^2}{dt^2}$, etc.

By differentiating (28) twice

$$C_1L_1D^3q_1+C_1MD^3q_2+C_1R_1D^2q_1+Dq_1=0, \quad . \quad . \quad . \quad (30)$$

$$C_1L_1D^4q_1+C_1MD^4q_2+C_1R_1D^3q_1+D_2q_1=0. \quad . \quad . \quad . \quad (31)$$

Similarly for circuit (2)

$$C_2L_2D^3q_2+C_2MD^3q_1+C_2R_2D^2q_2+Dq_2=0, \quad . \quad . \quad . \quad (32)$$

$$C_2L_2D^4q_2+C_2MD^4q_1+C_2R_2D^3q_2+D^2q_2=0. \quad . \quad . \quad . \quad (33)$$

Multiply (31) by C_2L_2 and (33) by C_1M and subtract

$$C_1C_2(L_1L_2-M^2)D^4q_1+C_1C_2L_2R_1D^3q_1$$
$$+C_2L_2D^2q_1-C_1C_2MR_2D^3q_2-C_1MD^2q_2=0. \quad . \quad (34)$$

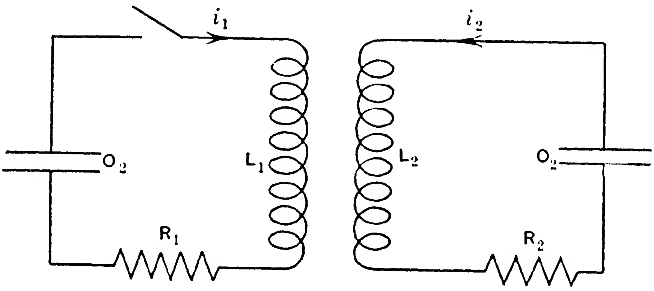


FIG. 16.—When the switch is closed C_1 will discharge through L_1 and R_1; current will also be set up in circuit 2, the actual current in the two circuits being similar to the motion of the pendulum bobs of Fig. 14.

Multiply (30) by C_2R_2 and add to (34).

$$C_1C_2(L_1L_2-M^2)D^4q_1+C_1C_2(L_2R_1+L_1R_2)D^3q_1$$
$$+(C_2L_2+C_1C_2R_1R_2)D^2q_1-C_1MD^2q_2+C_2R_2Dq_1=0. \quad . \quad (35)$$

Add (28) to (35) and get

$$C_1C_2(L_1L_2-M^2)D^4q_1+C_1C_2(L_2R_1+L_1R_2)D^3q_1$$
$$+(C_1L_1+C_2L_2+C_1C_2R_1R_2)D^2q_1+(C_1R_1+C_2R_2)Dq_1+q_1=0. \quad . \quad (36)$$

By a similar procedure an identical equation can be obtained for q_2.

The exact solution of Eq. (36) is not generally attempted in texts on radio; it is lengthy and the exact solution[1] differs but little from the approximate solution given below.

[1] For the exact solution the student is referred to an excellent article by F. E. Pernot in Vol. 1, No. 8, University of California Publications in Engineering.

Determination of the Two Frequencies of Oscillation.—In using Eq. (36) to obtain the frequencies of oscillation the solution is much simplified by making an assumption which is justified in all ordinary radio circuits, i.e., the resistance of the circuit has a negligible effect on the frequency of oscillation. We may therefore neglect the resistance terms in Eq. (36) in solving for the periods of oscillation; by doing this we get the comparatively simple equation,

$$\left(1-\frac{M^2}{L_1L_2}\right)D^4q_1+\left(\frac{1}{L_1C_1}+\frac{1}{L_2C_2}\right)D^2q_1+\frac{q_1}{C_1L_1C_2L_2}=0. \quad . \quad . \quad (37)$$

By substituting

$$k^2=\frac{M^2}{L_1L_2}, \quad \omega_1^2=\frac{1}{L_1C_1}, \quad \omega_2^2=\frac{1}{L_2C_2}$$

this becomes

$$(1-k^2)D^4q_1+(\omega_1^2+\omega_2^2)D^2q_1+\omega_1^2\omega_2^2q_1=0. \quad . \quad . \quad . \quad (38)$$

A similar analysis for q_2 would yield

$$(1-k^2)D^4q_2+(\omega_1^2+\omega_2^2)D^2q_2+\omega_1^2\omega_2^2q_2=0. \quad . \quad . \quad . \quad (39)$$

The solutions of (38) and (39) are, by inspection, of trigonometric form, so we put

$$q_1=A_1\cos(\omega t+\phi), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (40)$$

$$q_2=A_2\cos(\omega t+\phi'). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (41)$$

By differentiating these equations and inserting the values of the proper derivatives in Eqs. (38) and (39) we obtain the two values of ω.

$$\omega''=\sqrt{\frac{\omega_1^2+\omega_2^2+\sqrt{(\omega_1^2+\omega_2^2)^2-4\omega_1^2\omega_2^2(1-k^2)}}{2(1-k^2)}}, \quad . \quad . \quad (42)$$

and

$$\omega'=\sqrt{\frac{\omega_1^2+\omega_2^2-\sqrt{(\omega_1^2+\omega_2^2)^2-4\omega_1^2\omega_2^2(1-k^2)}}{2(1-k^2)}}. \quad . \quad . \quad (43)$$

If we now suppose the two circuits of Fig. 16 to be tuned alike, i.e., $L_1C_1=L_2C_2$, we can simplify Eqs. (42) and (43) very much. By introducing the condition that $\omega_1=\omega_2=\omega$ we get

$$\omega''=\frac{\omega}{\sqrt{1-k}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (44)$$

and

$$\omega'=\frac{\omega}{\sqrt{1+k}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (45)$$

From (44) and (45) we can get the value of the coupling coefficient as

$$k=\frac{(\omega'')^2-(\omega')^2}{(\omega'')^2+(\omega')^2}. \quad \dots\dots \quad (45a)$$

And it is to be noticed at this point of the analysis that these two frequencies are exactly the same as those given in Eqs. (130) and (131) of Chapter I for coupled circuits excited by an alternating e.m.f. of variable frequency. Indeed from the similarity of procedure we may conclude that a complex circuit having sufficiently low resistance, if left free to oscillate after being excited in some way or other, *will oscillate at those frequencies for which the system has zero reactance when excited by an alternating e.m.f.*

If, therefore, the natural periods of an electric circuit are desired it is only necessary to excite the circuit by an alternating e.m.f. of variable frequency and note those frequencies for which the power factor of the system is unity. When left free to vibrate, the circuit will, in general, oscillate at all these frequencies simultaneously, the energy dividing between the various frequencies.

This general theorem in resonance is a very useful one. Suppose three circuits as pictured in Fig. 17 all coupled together in any complex way possible; knowing all the constants of the circuits it would be possible to set up the differential equations and, after some laborious transformations, it would be possible to so combine them as to eliminate all but one variable. The resulting equation would, however, be difficult to solve, because schemes have not yet been evolved for solving an equation of the sixth degree.

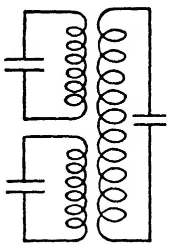


FIG. 17.—General case of three coupled circuits.

But if an alternating e.m.f. is introduced into the complex network, and the frequency of this e.m.f. be varied through as wide a range as necessary, there will be found three frequencies for which the network shows a comparatively low resistance only, the reactance being zero. These are the three free periods at which the network will oscillate if excited and left to itself.

The point where the power is introduced when determining the three resonant frequencies by impressing an e.m.f. is of no importance; it will be found that the same frequencies will result in unity power factor if the alternator is introduced in any line of the whole network.

It must be borne in mind that the previous remarks hold good only for circuits having low damping factors; the argument depends upon the assumption that this resistance does not appreciably affect the frequency of oscillation. The assumption is always warranted because the radio engineer is seldom interested in inefficient circuits, i.e., circuits having a low ratio of reactance to resistance.

We therefore write the solution of (38) and (39),

$$q_1 = A_1 \cos(\omega''t + \phi'') + B_1 \cos(\omega't + \phi'). \quad . \quad . \quad . \quad . \quad (46)$$

$$q_2 = A_2 \cos(\omega''t + \phi'') + B_2 \cos(\omega't + \phi'). \quad . \quad . \quad . \quad . \quad (47)$$

By differentiation of (46) and (47), we get the two currents,

$$i_1 = A_1\omega'' \sin(\omega''t + \phi'') + B_1\omega' \sin(\omega't + \phi')$$

$$= I''_1 \sin(\omega''t + \phi'') + I'_1 \sin(\omega't + \phi'), \quad . \quad . \quad . \quad . \quad (48)$$

and

$$i_2 = A_2\omega'' \sin(\omega''t + \phi'') + B_2\omega' \sin(\omega't + \phi')$$

$$= I''_2 \sin(\omega''t + \phi'') + I'_2 \sin(\omega't + \phi'). \quad . \quad . \quad . \quad . \quad (49)$$

The constants of Eqs. (48) and (49) must be chosen correctly to satisfy the initial conditions of the problem.

It will be noticed that these solutions give alternating currents of constant amplitude, evidently an impossible condition for the circuit of Fig. 16. The currents must rapidly die away as the energy originally stored in the condenser C_1 is used up in the resistances of the two circuits. The reason no damping term appears in the expressions for i_1 and i_2 is the neglect of the resistance terms of Eq. (36) in passing to Eq. (37). Of course, a circuit having no resistance has no damping.

Before proceeding to further analysis of the currents in the two circuits it is well to summarize the results so far obtained. *When the switch in circuit 1 is closed complex-shaped alternating currents begin to flow in both circuits 1 and 2; these complex currents are exactly represented by two currents of frequencies fixed by ω'' and ω', in each circuit.* We have therefore to determine the relative amplitude and phases of four currents I'_1 and I'_2 of frequency fixed by ω' (I'_1 in circuit 1 and I'_2 in circuit 2), and I''_1 and I''_2 of frequency fixed by ω'' (I''_1 in circuit 1 and I''_2 in circuit 2).

Relative Amplitude and Phases of Currents in Two Circuits.—An analysis of the phase and magnitude relations of the four currents I'_1, I''_1, I'_2, I''_2 was carried out by Chaffee and the deductions verified by an ingenious experiment; the results given below are taken from his paper.[1]

[1] "Amplitude Relations in Coupled Circuits," E. Leon Chaffee, Proc. I.R.E., Vol. 4, No. 3, June, 1916.

By using Eqs. (46) and (47) in combination with Eq. (28) (neglecting the resistance term in the latter) we get,

$$A_1\left(\frac{1}{L_1C_1}-\omega''^2\right)=\frac{A_2M\omega''^2}{L_1},$$

$$B_1\left(\frac{1}{L_1C_1}-\omega'^2\right)=\frac{B_2M\omega'^2}{L_1},$$

from which

$$\frac{A_2}{A_1}=-\frac{\omega''^2-\omega_1{}^2}{k\omega''^2}\sqrt{\frac{L_1}{L_2}},$$

$$\frac{B_2}{B_1}=\frac{\omega_1{}^2-\omega'^2}{k\omega'^2}\sqrt{\frac{L_1}{L_2}},$$

$$\frac{I''_2}{I''_1}=-\frac{\omega''^2-\omega_1{}^2}{k\omega''^2}\sqrt{\frac{L_1}{L_2}}, \quad \text{. . . . .} \quad (50)$$

$$\frac{I'_2}{I'_1}=\frac{\omega_1{}^2-\omega'^2}{k\omega'^2}\sqrt{\frac{L_1}{L_2}}. \quad \text{. . . . .} \quad (51)$$

Eq. (50) gives the ratio of amplitudes of the short waves in the two circuits and (51) that of the long waves. As ω'' is greater than ω_1 (see Eq. (42)), it is evident that I''_2 and I''_1 are in opposite phase; when one is positive the other is negative. The long waves I'_2 and I'_1 are in the same phase, however, as their ratio is positive, ω_1 being greater than ω' for all conditions of coupling.

In the oscillation transformer of a transmitting set, therefore, the effective flux for the short waves is much less than it is for the long waves; of course, this might be surmised, because the short wave in each circuit could only occur if the effective L_1 and L_2 were each diminished by the action of the other, which means currents in the two coils nearly 180° out of phase. The long waves, by a similar argument, must occur because the mutual action of L_1 and L_2 increases the effective inductance of each; this could only occur if the long wave currents in the two coils L_1 and L_2 are in phase, i.e., they magnetize their mutual field in the same direction.

To express the amplitude relation of the two currents in each circuit it is more convenient to express the relations of Eqs. (42) and (43) in terms of wave lengths. Using the relation $\lambda=\frac{2\pi V}{\omega}$ for each frequency involved (V being velocity of propagation of the electromagnetic waves), we get,

$$\lambda''=\sqrt{\frac{\lambda_1{}^2+\lambda_2{}^2-\sqrt{(\lambda_1{}^2-\lambda_2{}^2)^2+4k^2\lambda_1{}^2\lambda_2{}^2}}{2}}. \quad \text{. . .} \quad (52)$$

$$\lambda'=\sqrt{\frac{\lambda_1{}^2+\lambda_2{}^2+\sqrt{(\lambda_1{}^2-\lambda_2{}^2)^2+4k^2\lambda_1{}^2\lambda_2{}^2}}{2}}. \quad \text{. . .} \quad (53)$$

As Eqs. (42) and (43) were simplified by supposing the two circuits tuned alike, i.e., $\lambda_1=\lambda_2=\lambda$ we may write, for this condition, Eqs. (52) and (53) in the abbreviated forms

$$\lambda''=\lambda\sqrt{1-k}, \quad (54)$$

$$\lambda'=\lambda\sqrt{1+k}, \quad (55)$$

from which we get

$$k=\frac{(\lambda')^2-(\lambda'')^2}{(\lambda')^2+(\lambda'')^2}. \quad (56)$$

Eqs. (50) and (51) may also be written in terms of wave length and they become,

$$\frac{I''_2}{I''_1}=-\frac{1-\left(\frac{\lambda''}{\lambda_1}\right)^2}{k}\sqrt{\frac{L_1}{L_2}}. \quad (57)$$

$$\frac{I'_2}{I'_1}=\frac{\left(\frac{\lambda'}{\lambda_1}\right)^2-1}{k}\sqrt{\frac{L_1}{L_2}}. \quad (58)$$

To determine the ratio I'_1/I''_1 and I'_2/I''_2 we set down the initial conditions of the circuit. When $t=0$ (time of closing switch) $q_1=Q_0$, $q_2=0$, $i_1=0$, and $i_2=0$. Using these values in Eqs. (46), (47), (48), and (49), we get,

$$Q_0=A_1\cos\phi''+B_1\cos\phi'. \quad (59)$$

$$0=A_2\cos\phi''+B_2\cos\phi'. \quad (60)$$

$$0=A_1\omega''\sin\phi''+B_1\omega'\sin\phi'=I''_1\sin\phi''+I'_1\sin\phi'. \quad (61)$$

$$0=A_2\omega''\sin\phi''+B_2\omega'\sin\phi'=I''_2\sin\phi''+I'_2\sin\phi'. \quad (62)$$

To satisfy these conditions we must have $\phi'=\phi''=0$. Then we find as $A_2=-B_2$, and $I''_2=A_2\omega''$ and $I'_2=B_2\omega'$, that

$$\frac{I'_2}{I''_2}=-\frac{\omega'}{\omega''}=-\frac{\left(\frac{\lambda''}{\lambda_1}\right)}{\left(\frac{\lambda'}{\lambda_1}\right)}. \quad (63)$$

Dividing (50) by (51)

$$\frac{I''_2 I'_1}{I''_1 I'_2}=-\frac{\omega''^2-\omega_1^2}{\omega''^2}\,\frac{\omega'^2}{\omega_1^2-\omega'^2}=-\frac{\omega''^2-\omega_1^2}{\omega_1^2-\omega'^2}\,\frac{\omega'^2}{\omega''^2}$$

Multiplying by (63) we get,

$$\frac{I'_1}{I''_1}=\frac{\omega''^2-\omega_1^2}{\omega_1^2-\omega'^2}\,\frac{\omega'^3}{\omega''^3}=\frac{1-\left(\frac{\lambda''}{\lambda_1}\right)^2}{\left(\frac{\lambda'}{\lambda_1}\right)^2-1}\,\frac{\left(\frac{\lambda''}{\lambda_1}\right)}{\left(\frac{\lambda'}{\lambda_1}\right)}. \quad \ldots\ldots \quad (64)$$

For convenience in using the relations of Eqs. (63) and (64), the values of λ''/λ_1 and λ'/λ_1 have been calculated by Chaffee and are reproduced in Fig. 18. In this figure are shown the variations in λ''/λ_1 and λ'/λ_1 as λ_2/λ_1 is varied, this ratio being varied by varying λ_2 by a variation in condenser C_2. This keeps k constant as the ratio λ_2/λ_1 is varied.

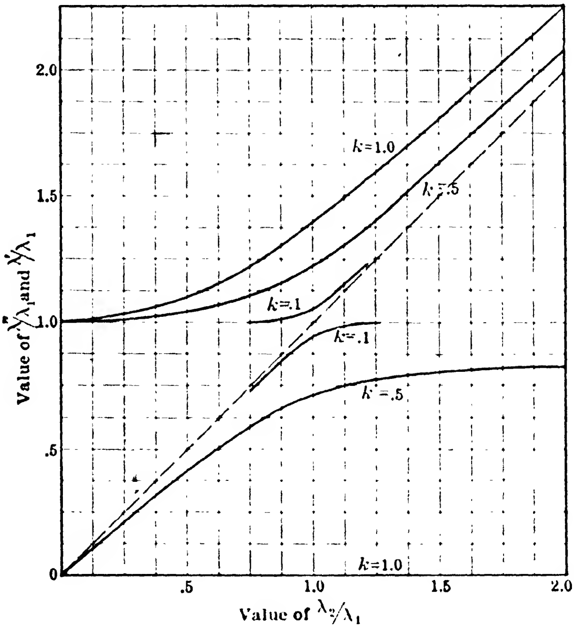


FIG. 18.—Variation in ratios of λ''/λ_1 and λ'/λ_1 as the ratio of λ_2/λ_1 is varied, for different values of coupling.

To get the magnitudes of the four currents, we solve for A_1, B_1, and A_2, B_2. From (59)

$$Q_0=A_1+B_1.$$

From (61)

$$I'_1=B_1\omega' \quad \text{and} \quad I''_1=A_1\omega'',$$

from which

$$\frac{B_1\omega'}{A_1\omega''}=\frac{I'_1}{I''_1}.$$

From this equation, by using (64),

$$\frac{B_1\omega'}{A_1\omega''}=\frac{\omega''^2-\omega_1{}^2}{\omega_1{}^2-\omega'^2}\,\frac{\omega'^3}{\omega''^3}$$

or

$$B_1=A_1\,\frac{\omega''^2-\omega_1{}^2}{\omega_1{}^2-\omega'^2}\,\frac{\omega'^2}{\omega''^2}.$$

Substituting this value of B_1 in above equation for Q_0 gives

$$A_1=Q_0\,\frac{\omega_1{}^2-\omega'^2}{\omega''^2-\omega'^2}\,\frac{\omega''^2}{\omega_1{}^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (65)$$

And as $B_1=Q_0-A_1$, we have

$$B_1=Q_0\,\frac{\omega''^2-\omega_1{}^2}{\omega''^2-\omega'^2}\,\frac{\omega'^2}{\omega_1{}^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (66)$$

From (62)

$$A_2=\frac{I''_2}{\omega''},$$

and by using (50) then substituting for I_1'' its equal $\omega''A_1$, we get

$$A_2=-A_1\left(\frac{\omega''^2-\omega_1{}^2}{k\omega''^2}\right)\sqrt{\frac{L_1}{L_2}},$$

then using (65)

$$A^2=-Q_0\,\frac{(\omega_1{}^2-\omega'^2)(\omega''^2-\omega_1{}^2)}{k\omega_1{}^2(\omega''^2-\omega'^2)}\sqrt{\frac{L_1}{L_2}}; \quad . \quad . \quad . \quad . \quad . \quad (67)$$

then from (60)

$$B_2=-A_2=Q_0\,\frac{(\omega_1{}^2-\omega'^2)(\omega''^2-\omega_1{}^2)}{k\omega_1{}^2(\omega''^2-\omega'^2)}\sqrt{\frac{L_1}{L_2}}. \quad . \quad . \quad . \quad (68)$$

Substituting the values of Eq. (65-68) in (48) and (49), we get,

$$i_1=\omega_1Q_0(F''_1\sin\omega''t+F'_1\sin\omega't) \quad . \quad . \quad . \quad . \quad . \quad . \quad (69)$$

and

$$i_2=\omega_1Q_0\sqrt{\frac{L_1}{L_2}}\,(F''_2\sin\omega''t+F'_2\sin\omega't). \quad . \quad . \quad . \quad . \quad . \quad (70)$$

In these equations

$$F''_1=\frac{\left(\dfrac{\lambda'}{\lambda_1}\right)^2-1}{\left(\dfrac{\lambda'}{\lambda_1}\right)^2-\left(\dfrac{\lambda''}{\lambda_1}\right)^2}\,\frac{1}{\left(\dfrac{\lambda''}{\lambda_1}\right)}.$$

$$F'_1=\frac{1-\left(\dfrac{\lambda''}{\lambda_1}\right)^2}{\left(\dfrac{\lambda'}{\lambda_1}\right)^2-\left(\dfrac{\lambda''}{\lambda_1}\right)^2}\,\frac{1}{\left(\dfrac{\lambda'}{\lambda_1}\right)}.$$

$$F''_2 = -\frac{\left[1-\left(\dfrac{\lambda''}{\lambda_1}\right)^2\right]\left[\left(\dfrac{\lambda'}{\lambda_1}\right)^2-1\right]}{k\left[\left(\dfrac{\lambda'}{\lambda_1}\right)^2-\left(\dfrac{\lambda''}{\lambda_1}\right)^2\right]\left(\dfrac{\lambda''}{\lambda_1}\right)}$$

$$F'_2 = \frac{\left[1-\left(\dfrac{\lambda''}{\lambda_1}\right)^2\right]\left[\left(\dfrac{\lambda'}{\lambda_1}\right)^2-1\right]}{k\left[\left(\dfrac{\lambda'}{\lambda_1}\right)^2-\left(\dfrac{\lambda''}{\lambda_1}\right)^2\right]\left(\dfrac{\lambda'}{\lambda_1}\right)}$$

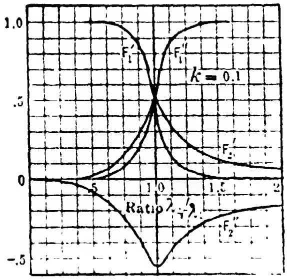


FIG. 19.—Values of the F coefficients for 10% coupling.

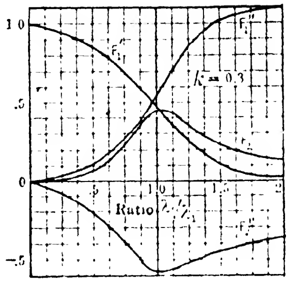


FIG. 20.—Values of the F coefficients for 30% coupling.

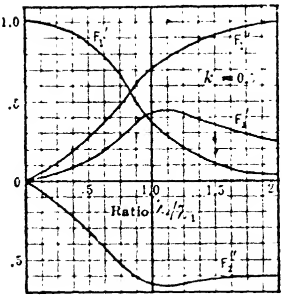


FIG. 21.—Values of the F coefficients for 50% coupling.

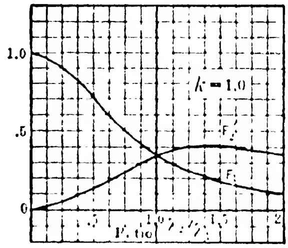


FIG. 22.—Values of the F coefficients for 100% coupling.

The values of these F factors are plotted in Figs. 19–22, which serve to show how the four different currents vary as C_2, the condenser in circuit 2, is varied, other things remaining constant. An examination of these curves shows that with weak coupling and tuned circuits the variation in amplitude (due to beats) is from maximum to zero as the values of F''_1 and F'_1 are equal in magnitude as are those of F''_2 and F'_2. For tighter couplings the ratio of λ_2/λ_1 must be different than unity to make $F''_1 = F'_1$

or $F''_2 = F'_2$. Furthermore, with other coupling than very loose no ratio of λ_2/λ_1 can be found which will make both $F''_1 = F'_1$ and $F''_2 = F'_2$ so that, except for very weak coupling the beats are not complete in both circuits, i.e., the minimum amplitude is not zero. It may be made zero in circuit 1 for any value of coupling by the proper amount of de-tuning, but the values of F''_2 and F'_2 are such as to preclude the possibility of zero amplitude beats for any except the weakest coupling, no matter how much λ_2 and λ_1 are made to differ.

A critical examination of the foregoing analysis shows that maximum current occurs in the second circuit when the ratio of λ_2/λ_1 is slightly greater than unity. This might have an important bearing on the use of a wave meter; this instrument is a coil and variable condenser which has an ammeter (or other device) for indicating resonance with the circuit, being tested. A precise analysis shows that maximum current will occur in this wave meter when its natural period is somewhat longer than that of the circuit being tested; as maximum current in the wave meter is ordinarily taken to signify resonance with the circuit tested it is evident that an appreciable error might be incurred.

It appears, however, that with a coupling between wave meter and the circuit tested as high as 10 per cent the error in wave meter reading is less than 1 per cent and as the wave meter coupling is, in practice, seldom more than 1 per cent or 2 per cent, the error is probably well within the precision of measurement.

The previous analysis of amplitudes, resulting in Eqs. (69) and (70) for the currents in the two circuits, was carried out without considering the resistance terms in the original equations, (28) and (29). The consideration of damping would have greatly complicated the derivations, and the damping factors can be introduced now without invalidating the previous work.

The damping factor of the high-frequency wave is the same for the high-frequency current in both circuits and similarly for the low-frequency wave. If we call the damping factors α'' and α', it is possible to derive the relations [1]

$$\alpha'' = \frac{1}{2(1-k)}\left(\frac{R_1}{2L_1} + \frac{R_2}{2L_2}\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (71)$$

$$\alpha' = \frac{1}{2(1+k)}\left(\frac{R_1}{2L_1} + \frac{R_2}{2L_2}\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (72)$$

α'' being for the high-frequency currents and α' for the low-frequency currents.

[1] A. Oberbeck, Wied. Ann. der Physik, 1895, Vol. 55, p. 623.

It is to be noticed that if Eqs. (71) and (72) are changed to give decrements (the two circuits being tuned), they assume the forms

$$\delta''=\frac{1}{\sqrt{1-k}}\left(\frac{\delta_1+\delta_2}{2}\right)$$

and

$$\delta'=\frac{1}{\sqrt{1+k}}\left(\frac{\delta_1+\delta_2}{2}\right),$$

where δ_1 and δ_2 are the decrements of the primary and secondary circuits, respectively. These solutions are approximate and good only when the decrements are low.

The complete solutions then become,

$$i_1=\omega_1 Q_0(F''_1\epsilon^{-\alpha''t}\sin\omega''t+F'_1\epsilon^{-\alpha't}\sin\omega't). \quad . \quad . \quad . \quad (73)$$

$$i_2=\omega_1 Q_0\sqrt{\frac{L_1}{L_2}}\,(F''_2\epsilon^{-\alpha''t}\sin\omega''t+F'_2\epsilon^{-\alpha't}\sin\omega't). \quad . \quad . \quad (74)$$

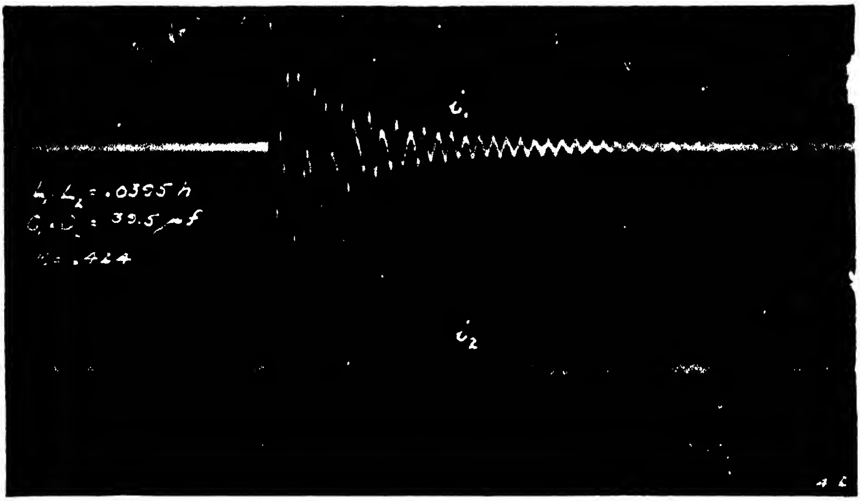


FIG. 23.—Currents in coupled tuned circuits with 42.4% coupling.

Actual Shapes of Currents in Coupled Circuits.—In Figs. 23–26 are shown oscillograms of currents in each of two coupled circuits, the circuits being practically identical. For each $L=0.0395$ henry and $C=39.5$ microfarads. The coefficients of coupling were 0.424, 0.282, 0.114, and 0.0707, respectively, for the several curves. The films do not quite bear out the preceding theory on amplitudes, as the values of F'_1, and F''_1

are evidently not near enough in amplitude to neutralize each other for even the minimum coupling, 7.07 per cent. It is quite likely that the

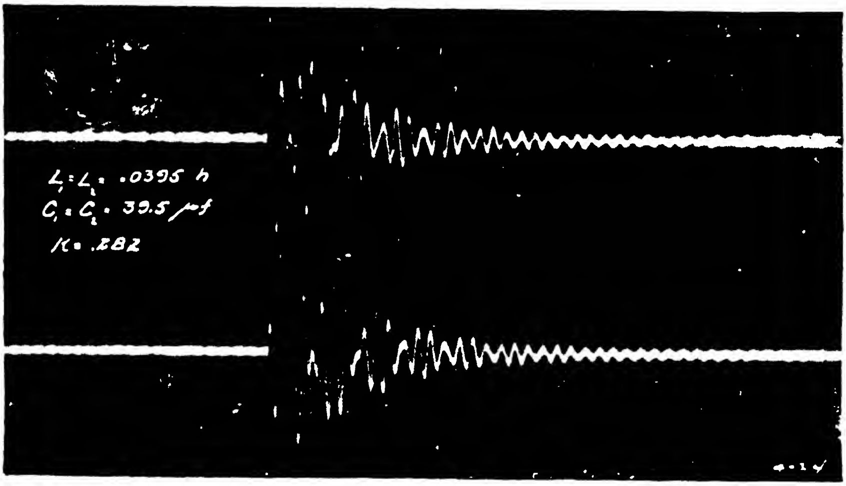


FIG. 24.—Currents in coupled tuned circuits with 28.2% coupling.

rather high decrement of the circuit had an appreciable effect on the various amplitude factors, not accounted for in the previous analysis.

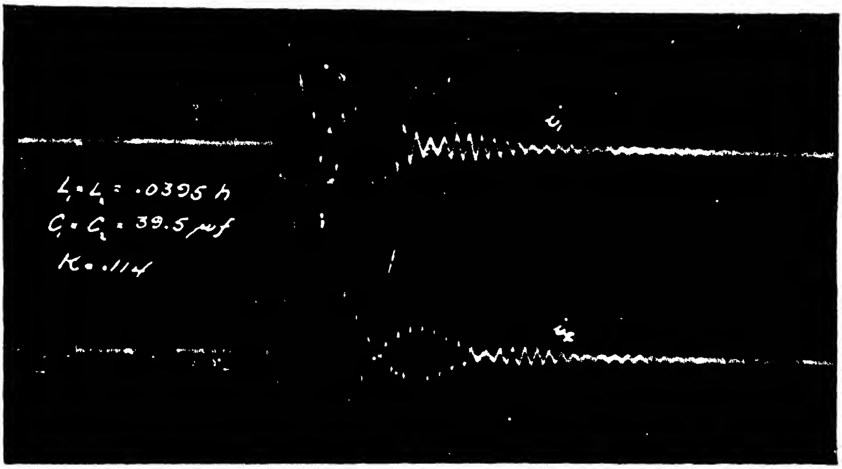


FIG. 25.—Currents in coupled tuned circuits with 11.4% coupling.

Frequency of the Actual Complex Current.—By inspection of the films shown in Figs. 23–26 it is seen that the time between successive zero points in the current wave is practically constant (indicating constant frequency);

in fact, careful measurement shows the frequency constant (for Fig. 26), within about 1 per cent, *except at the points of minimum amplitude,* where the time between successive zero points changes very much. Just what

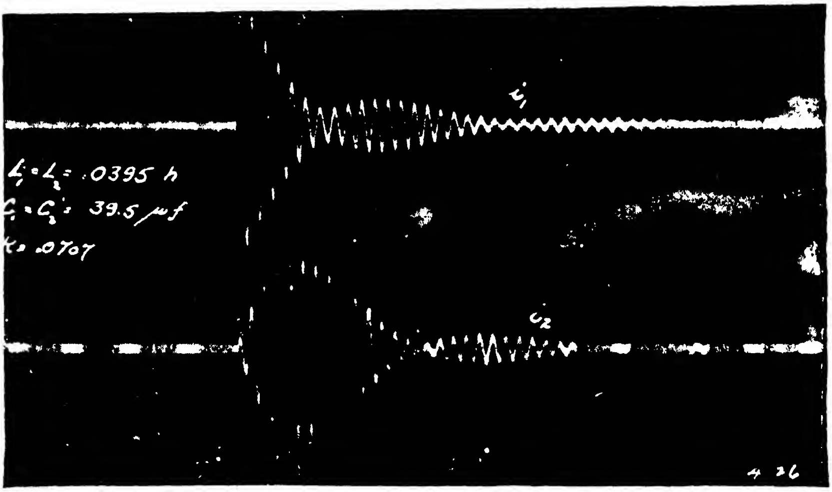


FIG. 26.—Currents in coupled tuned circuits with 7.07% coupling.

changes take place in the magnitude and phase of the current at this time depends altogether upon the relative amplitudes of the two component currents.

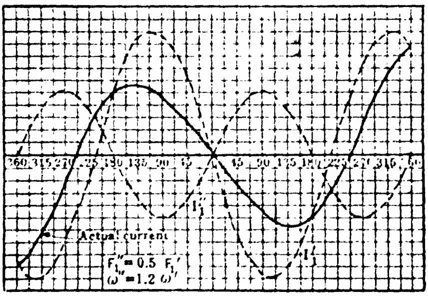


FIG. 27.—Form of current and minimum amplitude; high-frequency current much smaller than low-frequency.

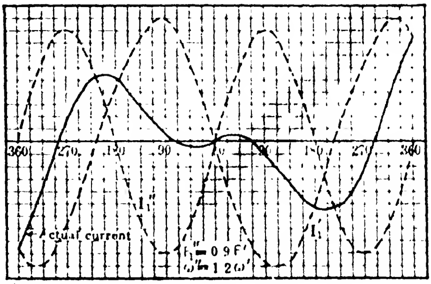


FIG. 28.—Form of current at minimum amplitude; high-frequency current of nearly the same amplitude as low-frequency.

In Figs. 27, 28, and 29 are shown three possible conditions at this time of minimum amplitude of the actual current. In Fig. 27 we have shown the condition for $F''_1 = 0.5\ F'_1$, in Fig. 28 for $F''_1 = 0.9\ F'_1$ and in Fig. 29 for $F''_1 = 1.25\ F'_1$. For all three figures we have $\omega'' = 1.20\ \omega'$, which means a value of coupling of the two circuits of about 20 per cent.

It might seem that as the frequency (time between successive zero points) of this "beating" current is constant, that a third circuit, coupled to the circuit carrying this complex current, would respond most strongly if tuned to this frequency. As a matter of fact but little response will be had in this third circuit if tuned to this actual frequency; if tuned to either of the component currents of this actual complex current, however, a strong response will be obtained.

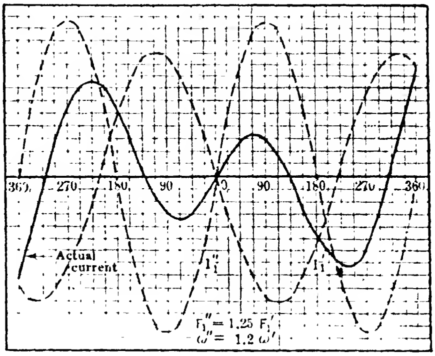


FIG. 29.—Form of current at minimum amplitude; high-frequency current greater than low-frequency.

Thus suppose the two circuits of Fig. 16 are each adjusted for a natural period of 100 cycles, and they are coupled 20 per cent. Then the two frequencies generated, when the condenser in circuit 1 discharges, will be (by Eqs. (44) and (45))

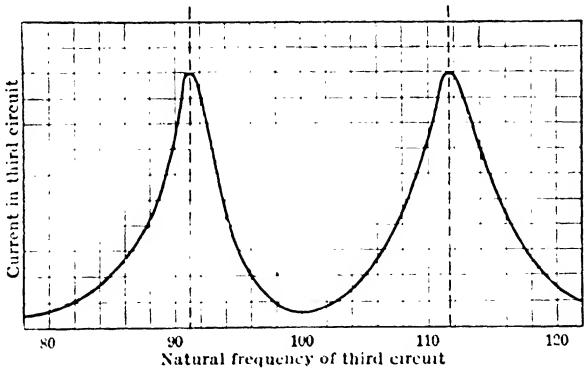


FIG. 30.—Amplitude of current in a third circuit coupled very loosely to either of the two coupled oscillating circuits.

$$f'' = \frac{100}{\sqrt{1-0.2}} = 111.7 \text{ cycles};$$

$$f' = \frac{100}{\sqrt{1+0.2}} = 91.2 \text{ cycles}.$$

The oscillatory current in each of the circuits will have a period of 0.01 second (except at the minimum amplitude points) but if a third circuit loosely coupled to either of the others is tuned to a natural period of 0.01 second the current induced in it will be much smaller than if it (the third circuit) is tuned to a natural frequency of either 111.7 or 91.2. The magnitude of current in this third circuit, as its natural frequency is changed by changing the value of its condenser, will be about as indicated in Fig. 30. The reason for this weak response to the 100-cycle tuning is the reversal of the phase in the exciting current at the minimum amplitude points; what current is built up in circuit 3 during time $t - t_1$, Fig. 31, is destroyed, or neutralized by the action during time $t_1 - t_2$, because of the phase reversal in the inducing current at time t_1.

Vector Representation of Current in Coupled Circuits.—Such a function as that given in Eq. (73) can be represented vectorially but, of course, the vector diagram is not of the ordinary type. If we let the instantaneous

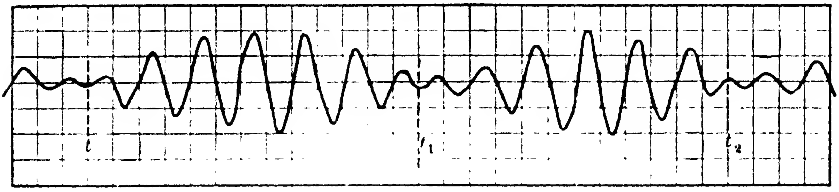


FIG. 31.—At minimum amplitude points the actual current in either of the two oscillating circuits reverses its phase.

value of the current be represented by the projection on the y axis of the resultant vector obtained by adding $F''_1\epsilon^{-\alpha''t}$ and $F'_1\epsilon^{-\alpha't}$ the construction will be as shown in Fig. 32. When $t=0$ the two vectors coincide in position; with increase in time the F''_1 vector advances its phase over that of F'_1, so after F'_1 has advanced 90°, as shown at OA, the vector F''_1 has moved to position OB. The resultant of these two vectors OR, gives, by its projection in the OY axis, OD, the instantaneous value of the actual current i_1. As the two vectors OA and OB rotate their magnitudes must continually diminish to keep them equal to $F''_1\epsilon^{-\alpha''t}$ and $F'_1\epsilon^{-\alpha't}$. The loci of the terminals of the vectors are logarithmic spirals about the point O. The logarithm of the ratio of the values of a vector, in two successive passages through the same phase gives the decrement of the current represented by that vector; thus we have $\log_\epsilon OM/ON = \delta''$ the logarithmic decrement of the current I''_1.

The unusual motion of this resultant vector as the two component vectors pass through phase opposition is indicated in Fig. 33. Vector OA, the one with less angular velocity, is shown stationary and the vector OB is shown in several successive positions around its phase opposition

position; OB is slightly greater in magnitude than OA. With OB in the position indicated by OB_1, the resultant of OA and OB_1 is shown by OR_1, etc. It may be seen that this resultant vector moves through the angle R_1OR_5, which is more than 180°, while the vector OB has moved about 45°.

The case of OB being smaller than OA is given in Fig. 34; in this case when OB goes through its opposition phase the resultant vector, instead of speeding up as it did in Fig. 33, slows down and goes through the successive values OR_1, OR_2, OR_3, etc., for the correspondingly marked positions of OB. If the two vectors OB and OA happen to have equal magnitudes as they go through phase opposition, the successive positions and values of the resultant vector are as shown in Fig. 35; for this con-

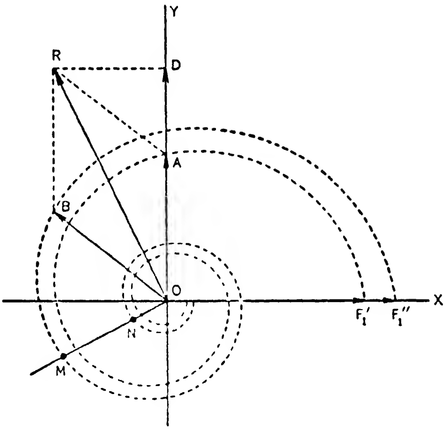


FIG. 32.—For damped sine waves the terminals of the e.m.f. vectors must lie on logarithmic spirals.

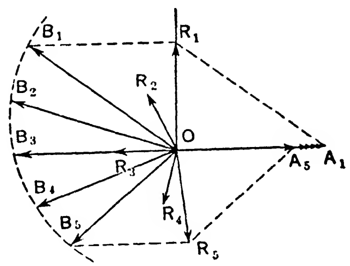


FIG. 33.—Resultant vector when high-frequency current has the greater amplitude.

dition when the two vectors are 180° out of phase the resultant vector is zero.

It is quite possible so to adjust the tuning of the two circuits that the vector OB is greater than OA at the start of the oscillations; then as the oscillations continue, OB, having greater damping than OA, will become equal to OA and then smaller. Hence in three successive beats it is possible to have the resultant vector OR go through phase changes as depicted in Figs. 33, 34, and 35, respectively, as the amplitudes of the actual current goes through its minimum values. The effects of these peculiar angular velocities of the resultant vector, in combination with its changes in magnitude, account for the peculiar form of the actual current during the one or two cycles of minimum amplitude. It is seen in Figs. 27 and

29 that the 180° phase shift which occurs at the point of minimum amplitude may be produced by either a gain of 180° or a loss of 180° at this time. Fig. 27 shows a loss and Fig. 29 a gain of nearly 180° during the time shown in those curves.

Frequency of Beats.—The beats are not well pronounced unless the two circuits are tuned to the same natural frequency; in this case all of the energy surges back and forth from one circuit to the other. With untuned circuits only a part of the energy is exchanged between the two, most of it remaining in the primary circuit; in this case the beats are not so pronounced as for the tuned circuits, because it is really the to-and-fro flow of the energy which gives the beats.

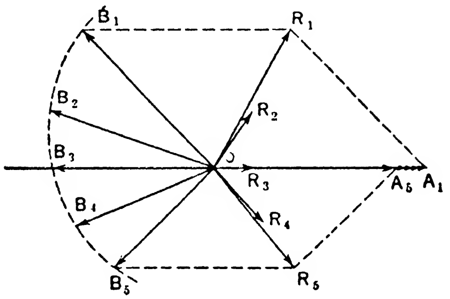


FIG. 34.—Resultant vector when low-frequency current has the greater amplitude.

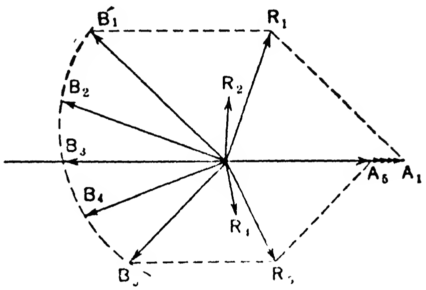


FIG. 35.—Resultant vector when both currents have the same amplitude.

In the case of tuned circuits the two frequencies are given by Eqs. (44) and (45),

$$\omega''=\omega\left(\frac{1}{\sqrt{1-k}}\right) \text{ and } \omega'=\omega\left(\frac{1}{\sqrt{1+k}}\right).$$

Hence

$$\omega''-\omega'=\omega\left(\frac{1}{\sqrt{1-k}}-\frac{1}{\sqrt{1+k}}\right)\cong\omega k. \quad . \quad . \quad . \quad . \quad . \quad . \quad (75)$$

This holds, of course, for low values of k only.

As the number of beats per second is equal to the difference in frequency per second of the two component frequencies, we must have the number of beats per second which are given by the relation $N=fk$.

We have shown previously that the frequency of the complex current for tuned circuits (except at the minimum amplitude point) is f.[1] The

[1] This is not a true frequency because of the varying amplitude; it is frequency determined by the time between successive zero values.

number of cycles of current per beat is therefore obtained from Eq. (75) by writing,

$$N=f''-f'=fk,$$

where N = beats per second. From this, we get

$$\frac{f}{N}=\frac{1}{k}. \qquad (76)$$

This equation is useful in determining the coupling of a pair of circuits from the shape of the complex wave of current. Thus if there is one beat for five cycles of current the coupling must be 20 per cent.

Form of Secondary Current if Primary Circuit is Opened at the Right Time.—The two circuits of Fig. 16 are used in practically every radio spark transmitting set; the condenser C_2 is, in the actual sets, the capacity of the antenna and part of the resistance, R_2 is the radiation resistance of the antenna. From the foregoing analysis of the current in circuit 2 it is evident that two wave lengths, λ'' and λ', would be radiated from the antenna; this is undesirable both from the standpoint of efficiency and interference, this latter factor being so important that government license will be granted only to those stations in which such precautions have been taken that practically all their power is radiated in one wave.

As previously stated, all the energy (to be transformed into oscillatory power) is originally stored in condenser C_1; when the switch S is closed this electric energy starts surging back and forth from L_1 to C_1 and also starts to flow over to circuit 2. If the two circuits are properly tuned all of the energy will have been transferred to circuit 2 in $1/2k$ cycles; unless prevented from doing so the energy then starts to flow back to circuit 1. Suppose, however, that circuit 1 is opened by some device or other, at that instant when all of its energy has been transferred to circuit 2; the re-transfer of the energy to circuit 1 is made impossible because no current can flow in circuit 1 if it is open.

Such an action is accomplished by a "quenched" spark gap to be described in detail in Chapter V. The forms of current in the primary and secondary circuit for this case are as indicated in Fig. 36; the curves are drawn for a coupling of 20 per cent. The number of cycles per beat for such a coupling is five; hence the time A-B during which energy is being transferred to the secondary (being one-half the time of a beat) will be $2\frac{1}{2}$ cycles. At time B the primary circuit is opened and from this time on the secondary circuit oscillates just as if the primary circuit was not present; in fact, electrically, circuit 1 is not present, it being open at the spark gap after time A–B.

The form of the current in the secondary circuit during time $A-B$ will approximate that given by Eq. (74); this equation is not strictly applicable because of the variable resistance (spark gap) in circuit 1. However, the resistance of the spark gap is probably negligible compared to the resistance due to the coupling of circuit 2 to circuit 1, so that Eq. (74) closely represents the form of current during time $A-B$. After time B the primary circuit is disconnected (by the opening of the spark gap) and the equation of secondary current is fixed by Eq. (19), the frequency and damping of the current being fixed by the secondary constants only.

The proper value of E to put in Eq. (19) is very nearly equal to

$$E\sqrt{\frac{L_2}{L_1}}\,\epsilon^{-\frac{\alpha_1+\alpha_2}{2}t'},$$

where E is the voltage of condenser C_1 when discharge began,

$$\alpha_1=\frac{R_1}{2L_1},\quad \alpha_2=\frac{R_2}{2L_2},\quad \text{and } t'=\text{time } A-B.$$

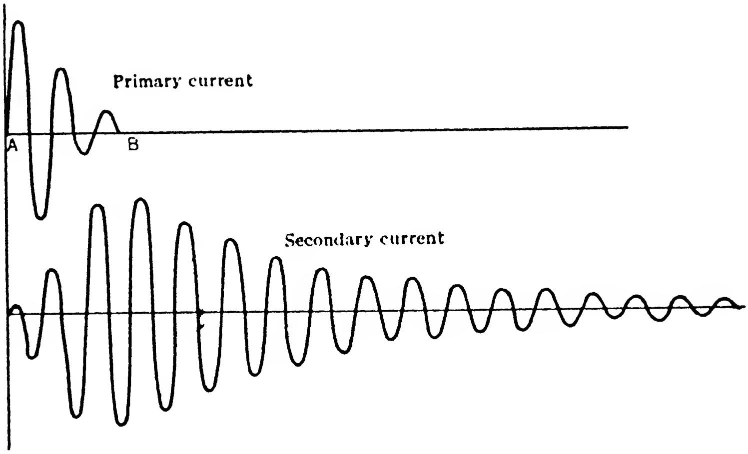


FIG. 36.—Forms of primary and secondary currents if primary circuit is opened at the first minimum.

With the conditions as represented in Fig. 36 the current in circuit 2 (except for the first one or two alternations) is of frequency f, the natural frequency of circuit 2, and the power is practically all radiated at this one frequency.

Possibility of No Beats without a Quenching Gap.—If the damping of the circuits is high and coupling is loose the beating phenomena will be absent, even if the spark gap in the primary does not offer any quenching action. This is illustrated by Fig. 37, in which oscillograms of primary and secondary current are shown for the circuit of Fig. 16, there being no spark gap at all in the primary circuit. The two circuits were tuned alike, the coupling was weak, $k=0.07$, and the decrement of each circuit was high, $\delta_1=\delta_2=0.30$.

This method of getting a current in the secondary of essentially single frequency is of no use to the radio engineer, because it really means that most of the energy originally stored in C_1 is dissipated as heat in the primary circuit; but little power is supplied to the secondary circuit where it is needed to carry out the useful function of radiation.

Oscillatory Discharge in One Circuit and Non-Oscillatory Discharge in the Other.—Under exceptional conditions it is possible with coupled circuits to have a non-oscillatory discharge in one circuit and an oscillatory discharge in the other. If the primary circuit has a high decrement

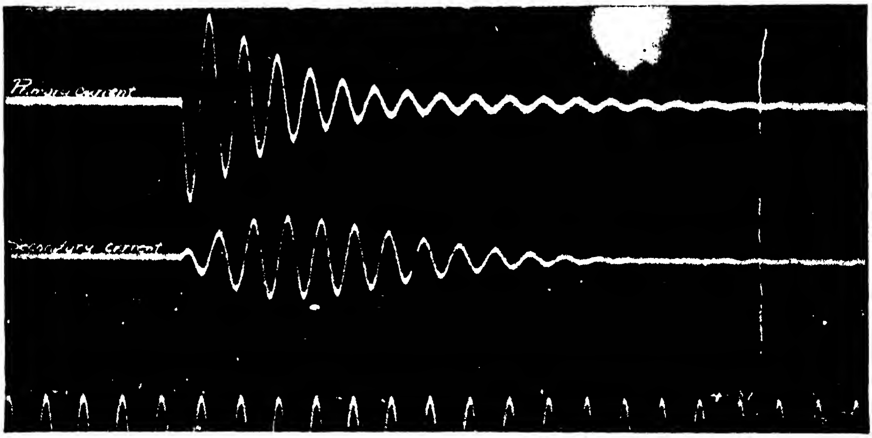


FIG. 37.—Forms of currents in coupled tuned circuits when the coupling is weak and damping is high.

and the secondary circuit a comparatively low decrement, then when the primary condenser discharges there may be a single, unidirectional pulse in the primary during which some of the primary energy is transferred to the secondary and some of it used as heat in the primary. Such a scheme is frequently used in small radio sets, and goes by the name of "impulse excitation." The primary circuit has generally a high decrement, having a large condenser and only one or two turns in its inductance. This gives a high value to $R/2fL$, especially when the resistance of the spark gap is taken into account. In addition to the high primary decrement, the gaps used in this method of generating oscillations are of the quenching type so that when they are functioning properly but one pulse exists in the primary and the secondary is left free to oscillate at its own period and its own decrement.

Oscillatory Circuit Excited by Continuous Voltage.—In case a circuit of L, C, and R, in series is connected to a source of continuous voltage E, Fig. 38, the equation of reactions is

$$E=L\frac{di}{dt}+Ri+\frac{q}{C}. \qquad (77)$$

By differentiating once this equation becomes the same as Eq. (1), and so its solution must be the same. The same three cases are to be considered here as they were for Eq. (1); the more important one of the solutions being that of Eq. (11). The initial and final conditions of the problems are different than those considered previously. Evidently at $t=0$, $v_c=0$ and at $t=\infty$, $v_c=E$; these conditions affect the equation of voltage across the condenser terminals, which becomes approximately,

$$v_c=E\left(1-\epsilon^{-\frac{Rt}{2L}}\cos\frac{t}{\sqrt{LC}}\right). \qquad (78)$$

This equation brings out the interesting fact that the maximum voltage across the condenser in such a circuit as that given in Fig. 38 is nearly double that of the source of e.m.f. to which the circuit is connected. This is illustrated by the film shown in Fig. 39; the voltage of the c.c. line to which the circuit was connected was 105 volts, whereas the maximum potential difference across the condenser was 190 volts. It is evident from this oscillogram that if the dielectric strength of a condenser is to be tested by connecting it to a source of continuous e.m.f., a resistance should be used in series with the condenser of sufficient magnitude to make the circuit aperiodic. If this is not done the maximum voltage across the condenser is not E, the voltage of the line used for testing, but is equal to $E(1+\epsilon^{-\frac{\delta}{2}})$, where δ is the decrement of the circuit.

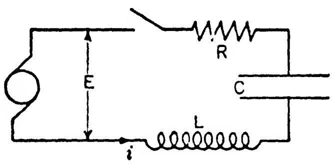


FIG. 38.—Oscillatory circuit connected to a source of continuous-current power.

This same phenomenon occurs frequently in everyday life. Suppose a springboard deflects 1 foot when a 200-lb. man is standing quietly on it at its free end. If the man starts to walk slowly out from the anchored end, the board will be deflecting 1 foot by the time he reaches the end. If, however, with the board undeflected the man steps on to the free end (say from a boat) the board will bend down *2 feet.* It will then spring back up and after a few oscillations will have a steady deflection of 1 foot. Thus the board has twice as much strain (maximum) if the man steps (not jumps) directly on to the free end as if he quietly walks out from the anchored end.

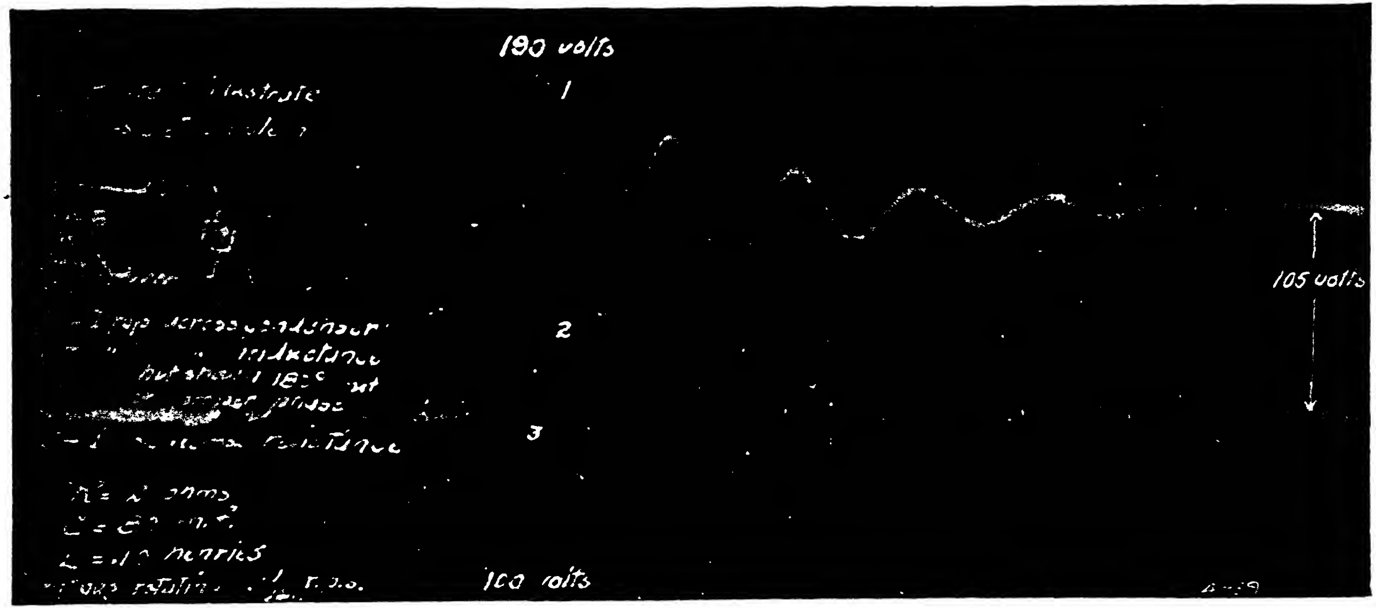


FIG. 39.—Oscillogram of current and voltages of the circuit shown in Fig. 38. The source of power for the 105-volt excitation was a battery having a resistance of 2.50 ohms.

Oscillatory Circuit Excited by Energy Stored in Inductance.—In certain radio-testing circuits oscillations are produced not by the energy stored in a condenser, but by the energy in the magnetic field of the inductance. The circuit is indicated in Fig. 40; in the actual testing set the battery circuit is made and broken many times a second, perhaps 1000, the function of the switch being performed by the contact points of a small buzzer. When the switch S is closed the condenser C charges at once to battery voltage and the current through L and R rises on a logarithmic curve (Eq. (12), p. 40), to a value E/R; the magnetic energy in the coil being $LE^2/2R^2$. When the switch is opened this magnetic energy is emptied into the condenser C, and then the energy surges back into L as described in the first section of this chapter.

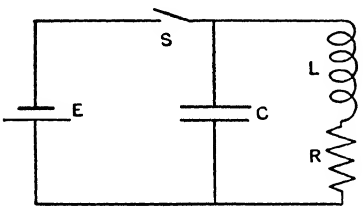


FIG. 40.—Oscillatory circuit to be excited by stored magnetic energy. This circuit is the same as used for "buzzer excitation" of radio circuits.

At the end of the first quarter of a cycle of the oscillation all the energy from the coil is in the condenser; it is then charged to such a potential difference E' that we have (if the decrement of the circuit is so low that the damping for one-quarter of a cycle may be neglected),

$$\frac{CE'^2}{2}=\frac{CE^2}{2}+\frac{LE^2}{2R^2}=E^2\left(\frac{C}{2}+\frac{L}{2R^2}\right),$$

or

$$E'=E\sqrt{1+\frac{L}{CR^2}}. \qquad (79)$$

The cycle of events in such a circuit as shown in Fig. 40 is shown in the film of Fig. 41; of course, all the constants of the circuit used in getting this film are much greater than those used in the so-called " buzzer wave generator " used in radio, but the form of voltages and currents are nearly the same as those occurring in the radio circuit.

Oscillating Circuits Excited by being Connected to a Line of Alternating E.M.F.—If a circuit of L, R, and C, in series, Fig. 42, is suddenly switched to an alternating-current line, the current must be zero, no matter at what point the e.m.f. wave the switch is closed; in general the condenser of the circuit will not be charged. Now in the steady state the current must have a certain value for any given value of e.m.f. as fixed by Eqs. (40) and (41) of Chapter I. Also the condenser must have a definite charge for this value of impressed e.m.f. It is evident, therefore,

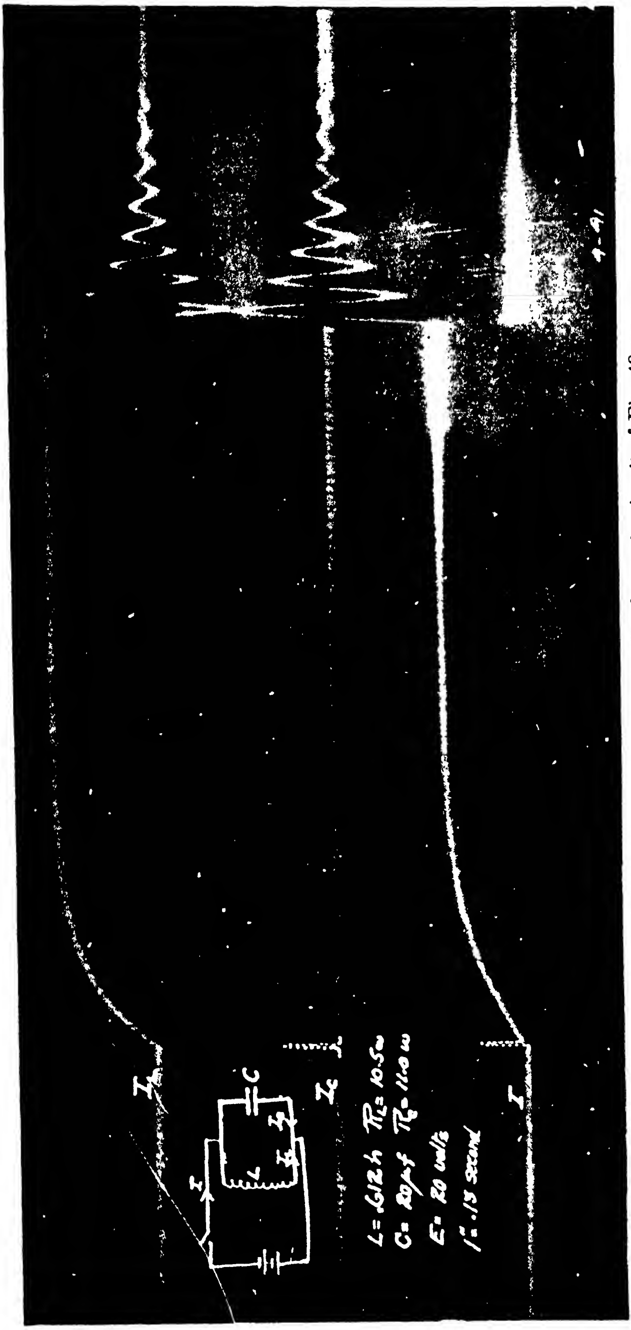


FIG. 41.—Oscillogram of current and voltages in circuit of Fig. 40.

that in general the initial conditions, when the switch is closed, will not satisfy the conditions required by the steady state.

For this reason the current for the first few cycles after switching the circuit to the line will be of irregular form; the circuit requires time to "settle down" to the steady state. Mathematically this is accomplished by adding to the equation for the steady current a suitable damped oscillation, the magnitude of which depends upon the time the switch is closed and the frequency of which is fixed by the L and C of the circuit.

The actual current after closing the switch is therefore the sum of the steady value of current and a damped oscillation at the natural period of the circuit, the two sufficing to satisfy the required initial conditions on closing the switch.

If the impressed voltage is $e=E\sin pt$, the circuit having constants, L, C and R and $\left(\frac{R}{2L}\right)^2<\frac{1}{LC}$ the solution is,

$$i=A\epsilon^{-\alpha t'}\sin(\omega t')+\frac{E}{\sqrt{R^2+\left(pL-\frac{1}{pC}\right)^2}}\sin(pt-\phi), \quad . \quad . \quad (80)$$

in which

$$\tan\phi=\left(pL-\frac{1}{pC}\right)/R,$$

$$\alpha=\frac{R}{2L} \text{ and } \omega=\sqrt{\frac{1}{LC}}, \text{ approximately,}$$

and t' is the time counted from the start of the imagined transient oscillatory current; it is sometimes written $(t+\Delta t)$ where Δt is the time between the start of the imagined transient term and the closing of the switch—this increment of time is indicated in Fig. 44.

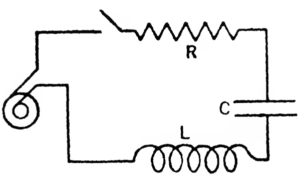


FIG. 42.—Oscillatory circuit to be connected to source of alternating-current power.

A and t' are to be suitably determined to satisfy the initial condition that $i=0$ and $v_c=0$. This condition $v_c=0$, supposes the condenser to be uncharged at the time of switching the circuit to the line; if it is charged to a certain potential difference V, then the initial conditions are $i=0$, $v_c=V$.

Let us suppose the steady state of the circuit is as represented in Fig. 43, and further let us suppose that the switch is closed at the phase indicated by θ. In the steady state the current should be I' and the voltage across the condenser should be V'. Actually the current at the time of closing the switch is zero, and we also suppose an uncharged condenser, so that $v_c=0$. We must then so determine t' and A of Eq. (80) that these initial conditions are satisfied.

The equation for voltage across the condenser, due to the transient term only, we write,

$$v_c=E_0\epsilon^{-\alpha t'}\cos\omega t',$$

and hence, the current due to the transient term is

$$i=C\frac{dv_c}{dt}=-\omega CE_0\epsilon^{-\alpha t'}\sin\omega t'-\alpha CE_0\epsilon^{-\alpha t'}\cos\omega t'.$$

We here make the same assumption we have previously made for similar circuits, that α is negligible compared to ω, and so we get,

$$i=-\omega CE_0\epsilon^{-\alpha t'}\sin\omega t'. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (81)$$

Using the condition that the voltage across the condenser must be zero at the time of closing the switch, we have

$$v_c+V'=0 \quad \text{or} \quad -V'=\epsilon^{-\alpha t'}E_0\cos\omega t'_0,$$

in which t'_0 is the value of t' when the switch is closed.

Then

$$E_0=-\frac{V'}{\epsilon^{-\alpha t'_0}\cos\omega t'_0}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (82)$$

Also

$i+I'=0$, so from (81) using also (82)

$$I'=-\frac{\omega CV'\epsilon^{-\alpha t'_0}\sin\omega t'_0}{\epsilon^{-\alpha t'_0}\cos\omega t'_0}$$

or

$$\tan\omega t'_0=-\frac{I'}{\omega CV'}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (83)$$

In case the damping of the circuit is small this equation may be written

$$\tan\omega t'_0=-\frac{I'}{V'}\sqrt{\frac{L}{C}}.$$

From this equation we get $\tan\omega t'_0$ and so may find the value of $\cos\omega t'_0$. Knowing $\omega t'_0$ and ω we get t'_0 and so can calculate $\epsilon^{-\alpha t'_0}$ and then substituting in (82), we get E_0; evidently $A=-\omega CE_0$, which can now be substituted in Eq. (80).

In Fig. 44 are reproduced in dotted lines the current and voltage curves of Fig. 43 and in dashed lines the transient current and condenser voltage determined from Eqs. (83) and (82) and (81). The addition of the steady value of current and the transient current gives the full line curve which is the actual current in the circuit after closing the switch.

In Fig. 45 is shown an oscillogram of the transient current after switching such a circuit as the one used in plotting the curves of Fig. 44. From Figs. 44 and 45 it may be seen that on switching a circuit, of L, R, and C, in series, to an alternating current line the condenser might be subjected to much higher voltages than occur in the steady state, nearly twice as much if the damping is low.

Periodic Disturbances in a Resonant Circuit.—In every radio spark set there is a circuit the equivalent of that shown in Fig. 46; in place of the switch shown there is a spark gap which performs the same function. An alternator supplies power to a condenser C, through a transformer T; once every cycle (or alternation) the condenser is short circuited for a very short time; in an actual set by the spark gap breaking down, in the set from which the following oscillograms were taken, by a suitable revolving switch. The time of switch closure was adjusted for maximum voltage in the secondary circuit and took place at the same phase of each cycle. The voltage across the condenser and current in the secondary of the transformer for this case have been worked out theoretically, but they are rather unwieldy, as one might suppose after an elementary consideration of the problem.[1]

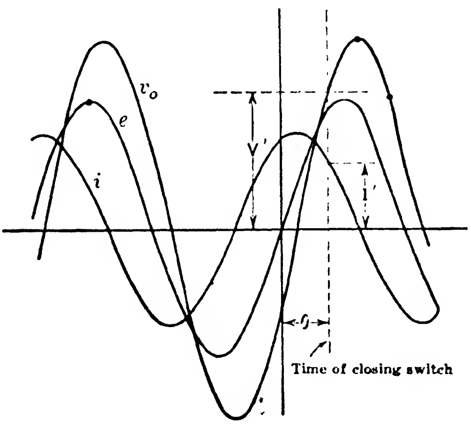


FIG. 43.—Proper "steady state" values of voltages and current of circuit of Fig. 42, at time of closing switch.

For each closure of the switch S, a transient term is introduced into the circuit, and as the damping is not sufficient to eliminate one transient before another is introduced, the actual current and voltage consist of the steady values with a whole series of transients superimposed. The form

[1] Fulton Cutting, "The Theory and Design of Radio Telegraphic Transformers," Proc. I.R.E., Vol. 4, No. 2, April, 1916. This article serves to show how complicated an exact treatment may become; in Chapter V, pages 413–414, are shown some curves which are calculated from simpler formulas, which curves represent quite accurately the form of disturbance in the ordinary spark transmitting set.

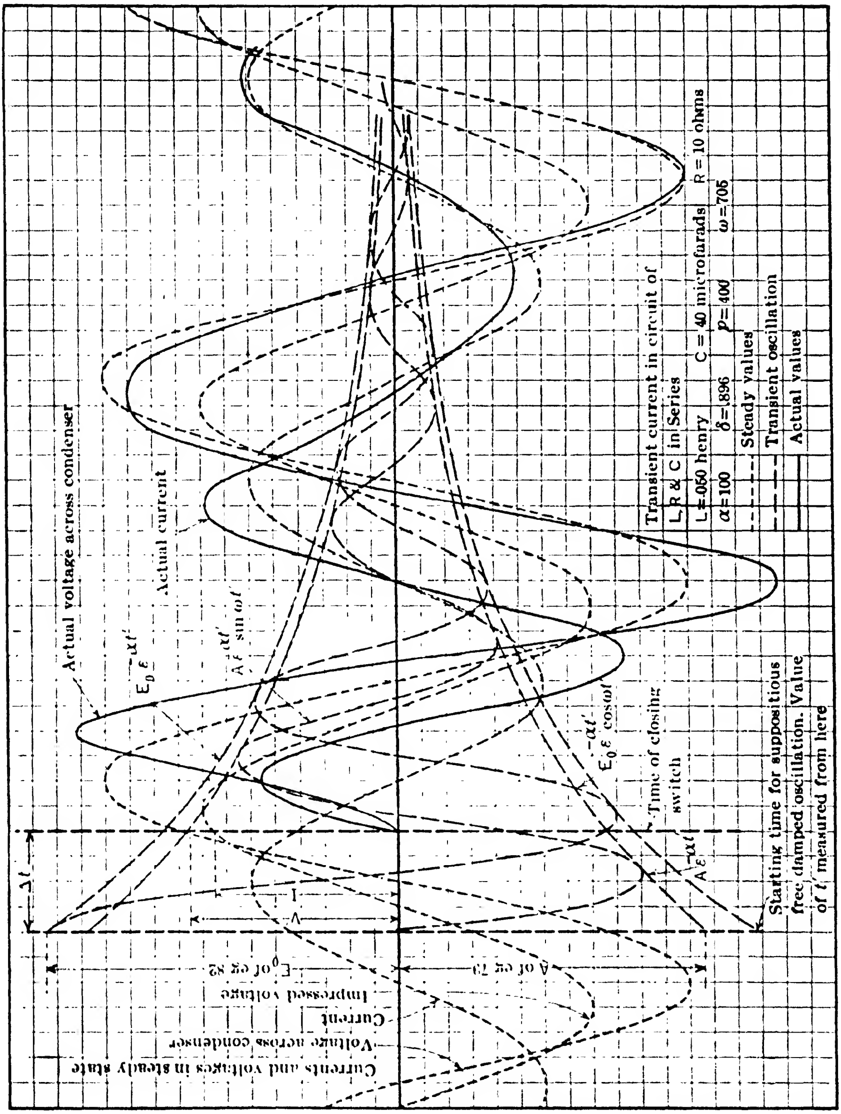


FIG. 44.—Analysis of the transient condition existing after closing the switch of circuit shown in Fig. 42.

of voltage and current depends largely upon the ratio of the frequency of the impressed e.m.f. to the natural frequency of the circuit.

In Figs. 47, 48, and 49 are shown the forms of voltage across the condenser and current in the secondary of the transformer for three values

FIG. 45.—Oscillogram of transient current corresponding to condition shown in Fig. 44.

of this ratio. In Fig. 47 the natural period was less than that of the alternator, in Fig. 48, the two were equal, and in Fig. 49, the natural period was greater than that of the alternator. As the films were taken at high speed they are not very distinct, so two cycles have been dotted in with ink.

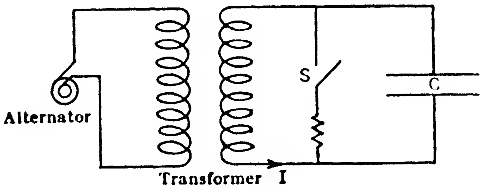


FIG. 46.—Elementary circuit illustrating the action occurring in every spark set when the spark gap breaks down.

It will be seen at once that any mathematical expression to represent these curves must be a complex one. With the switch adjusted to make one closure per cycle the circuit is a rectifying one; if the voltage across the condenser at the time of short circuit is E, it is evident that each cycle the secondary of the transformer carries more in one direction than it does in the other a quantity of electricity equal to CE.

Oscillating Circuit Excited by Pulse.—It often occurs in radio work that an oscillatory circuit is excited by a unidirectional pulse of some shape or other; thus it is quite likely that atmospheric disturbances in radio receiving circuits are due to some sort of highly damped oscillation or a series of short pulses.[1] The effect of a pulse on a resonant circuit will depend upon two factors, the ratio of the duration of the pulse to the natural period of the circuit, and the intensity or amplitude of the

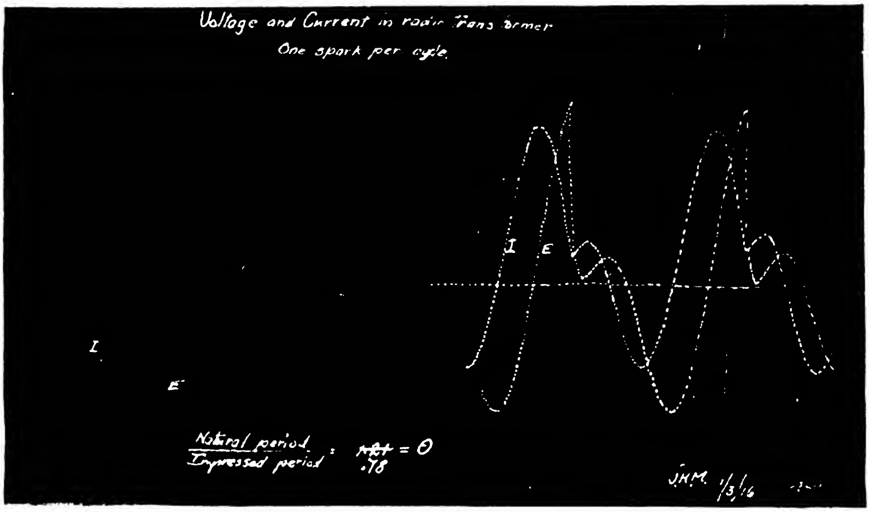


FIG. 47.—Voltage and current in such a circuit as that shown in Fig. 46, the switch S being closed synchronously. Natural frequency of secondary circuit greater than alternative frequency.

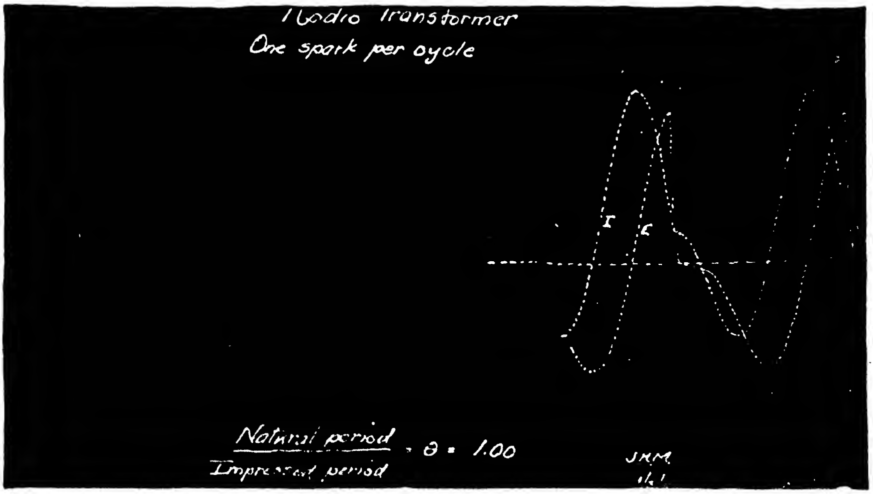


FIG. 48.—Similar to Fig. 47 but with secondary circuit having a natural frequency equal to that of the alternator.

[1] See p. 391.

pulse. Also to a minor extent the exact form of the pulse will determine the amount of disturbance produced.

The simplest kind of a pulse to consider mathematically is a " square " pulse, one in which the voltage rises suddenly from zero to a certain value, holds this value for a short time and then again drops suddenly to zero. If such a pulse of voltage is introduced into a circuit consisting of L, C, and R, in series the shape of the current can be obtained by properly combining the solutions of Eqs. (77) and (1). Eq. (77) gives the conditions when the voltage is applied to the circuit and Eq. (78) gives the voltage on the condenser at any time t after the voltage has been applied. When the pulse of voltage ends, Eq. (1) applies, the voltage on the discharging condenser being that determined from Eq. (78).

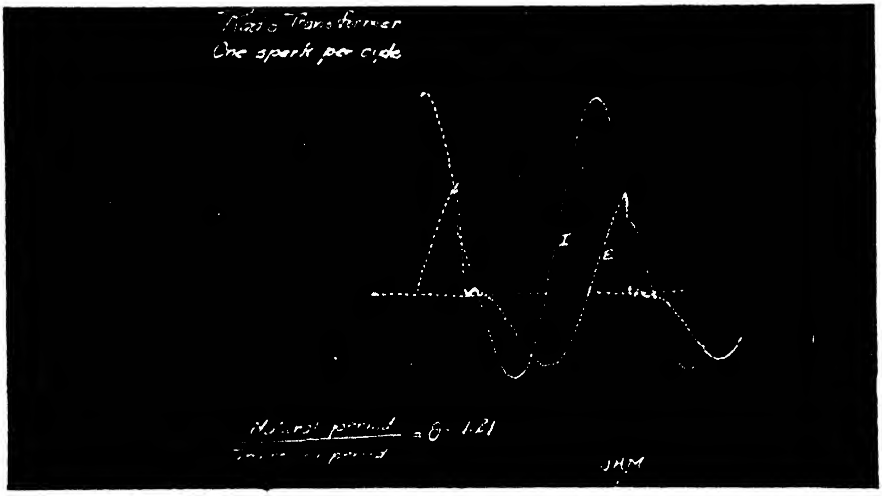


FIG. 49.—Similar to Fig. 47 but with secondary circuit having a natural frequency less than that of the alternator.

Thus in Fig. 50 is shown at a the pulse of e.m.f. introduced into the oscillating circuit, in b is shown in full lines the condenser voltage curve, determined from Eq. (78) and in the dotted line the current produced in the circuit by the introduction of voltage E.

Counting $t=0$ at the beginning of the pulse, we have

$$i=\frac{E}{\omega L}\epsilon^{-\alpha t}\sin \omega t, \quad \ldots \ldots \ldots \quad (84)$$

and

$$v_c=E(1-\epsilon^{-\alpha t}\cos \omega t), \quad \ldots \ldots \ldots \quad (85)$$

in which $\omega=\frac{1}{\sqrt{LC}}$ and $\alpha=\frac{R}{2L}$, the solution being approximate as α has been considered small compared to ω.

If the pulse has a duration, T, at the end of the pulse, the voltage in the condenser is

$$v_c = E(1 - \epsilon^{-\alpha T} \cos \omega T). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (86)$$

Now at the end of the pulse the solution of Eq. (1) is available if we substitute the proper initial conditions. The circuit solved in Eq. (1) was one in which the initial conditions were a charged condenser and the zero current. By inspection of Eqs. (86) and (84) it is evident that if

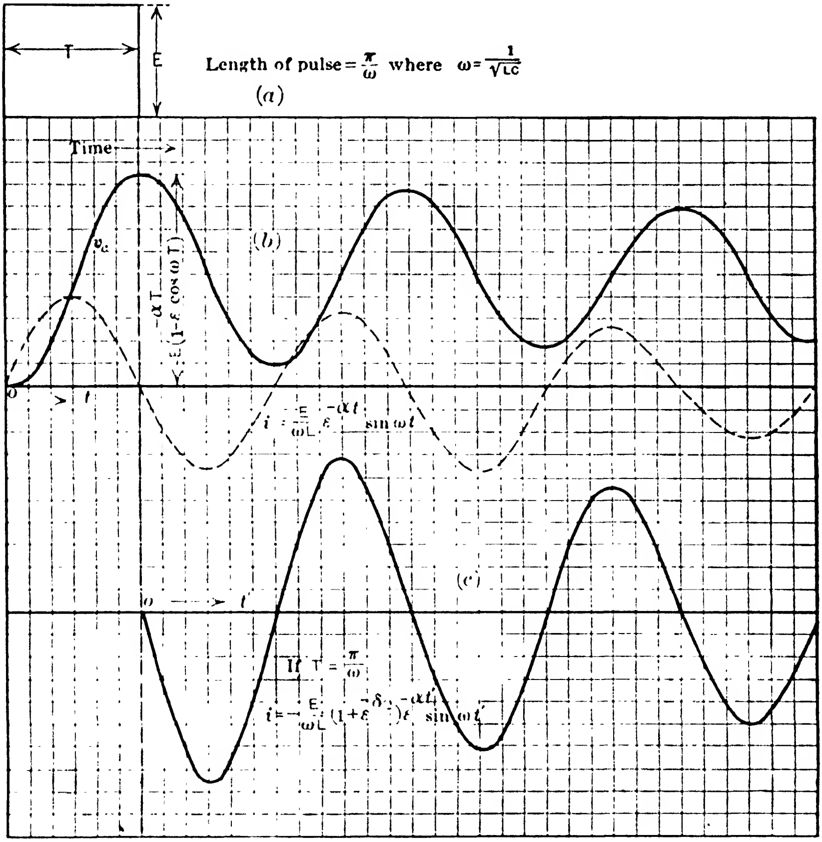


FIG. 50.—Effect of introducing a rectangular pulse of e.m.f. into an oscillatory circuit.

we make $T = \pi/\omega$ the current at the end of the pulse is zero ($\sin \omega T = 0$) and the voltage in the condenser is $v_c = E(1 + \epsilon^{-\alpha T})$, as $\cos \omega T = -1$.

The equation for current from the end of the pulse (for length of pulse $= \pi/\omega$) is therefore,

$$i = -\frac{E}{\omega L}(1 + \epsilon^{-\frac{\delta}{2}})\epsilon^{-\alpha t'} \sin \omega t'$$

where t' is reckoned after the end of the pulse. This current is shown in curve c, Fig. 50. At time $t' = \pi/2\omega$, $\sin \omega t' = 1$ and the value of current is

$$I_{\max.} = -\frac{E}{\omega L}(\epsilon^{-\frac{\delta}{4}} + \epsilon^{-\frac{3\delta}{4}}). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (87)$$

This is the maximum current obtainable from a rectangular pulse, of amplitude E, no matter what its duration may be. Any duration either more or less than π/ω will give less value to $I_{\max}$.

A more fundamental way of looking at the problem is to consider a voltage of $+E$ impressed on the circuit at the beginning of the pulse, and that this voltage is maintained; at the end of the pulse a voltage of $-E$ is impressed on the circuit and maintained. Each of the impressed voltages will produce a current, and the actual current at any time is the sum of the two.

The current after the second voltage $(-E)$ is impressed

$$i = +\frac{E}{\omega L}\,\epsilon^{-\alpha t} \sin \omega t - \frac{E}{\omega L}\,\epsilon^{-\alpha t'} \sin \omega t',$$

in which t is the time after the $(+E)$ voltage is impressed and t' is the time after the second voltage $(-E)$ is impressed. If the interval between the application of these two voltages is T_0, then the current after time T_0 has passed is

$$i = -\frac{E}{\omega L}\left\{\epsilon^{-\alpha t'} \sin \omega t' - \epsilon^{-\alpha(t'+T_0)} \sin \omega(t'+T_0)\right\}.$$

If the damping is comparatively small, the maximum current will occur when $\sin \omega t'$, and $\sin \omega(t'+T_0)$ are simultaneously equal to unity and of opposite sign. Moreover, it is evident that, because of the damping factors, this maximum current will have its greatest value when the above conditions occur for the smallest possible value of $\omega t'$. Inspection shows this to be when $\omega t' = \pi/2$ and $\omega T_0 = \pi$; this means that the length of the pulse (time between applying $+E$ and $-E$) should be equal to one-half the natural period of the circuit and the maximum current occurs one-quarter of a period after the end of the pulse. Putting $T_0 = T/2$ (T being the natural period of the circuit), the equation for current becomes

$$i = -\frac{E}{\omega L}\left(\epsilon^{-\alpha t'} + \epsilon^{-\alpha\left(t' + \frac{T}{2}\right)}\right) \sin \omega t'),$$

and if we now suppose $\omega t' = \pi/2$ and write the damping in terms of decrement, we get

$$I_{(\text{maximum})} = -\frac{E}{\omega L}(\epsilon^{-\frac{\delta}{4}} + \epsilon^{-\frac{3\delta}{4}}).$$

An oscillographic investigation of impulse excitation was carried out by the author [1] and some of the films obtained are shown in Figs. 51, 52,

[1] Proc. I.R.E., Vol. 8, No. 1, February, 1920.

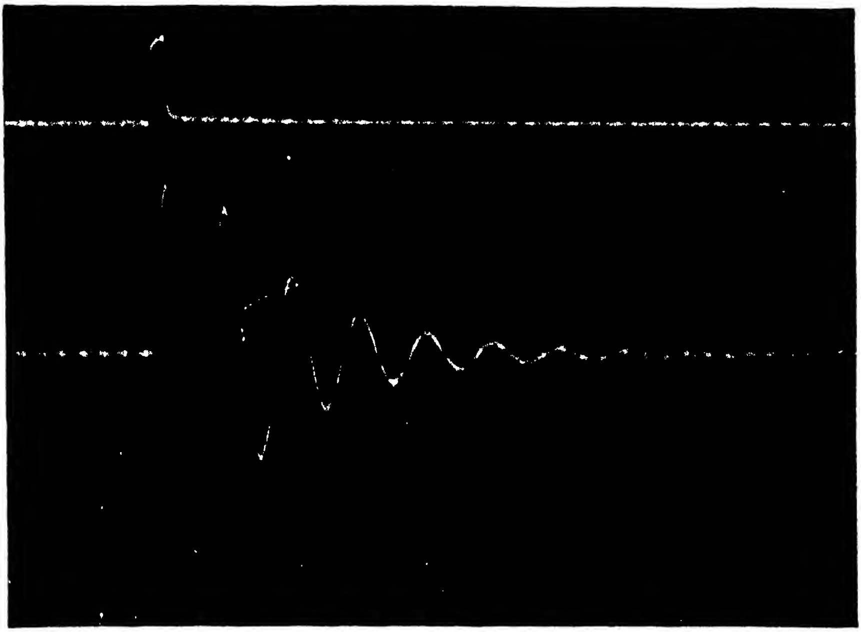


FIG. 51.—Oscillogram of oscillatory current produced by short pulse of e.m.f.

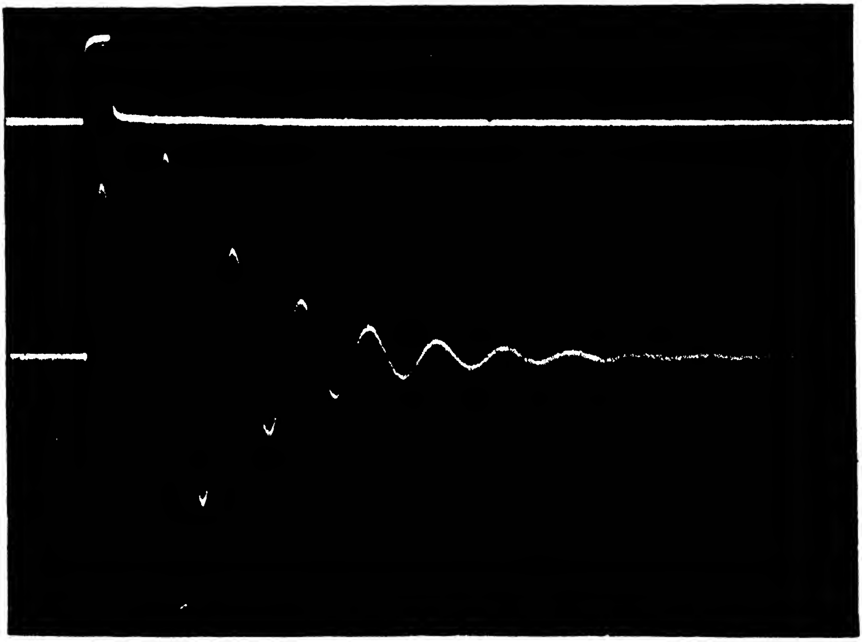


FIG. 52.—Similar to Fig. 51 but with longer pulse.

and 53. The upper record of the films shows the shape and length of pulse of e.m.f. introduced into the oscillatory circuit and the lower curve shows the current set up in the circuit by the pulse. A complete set of

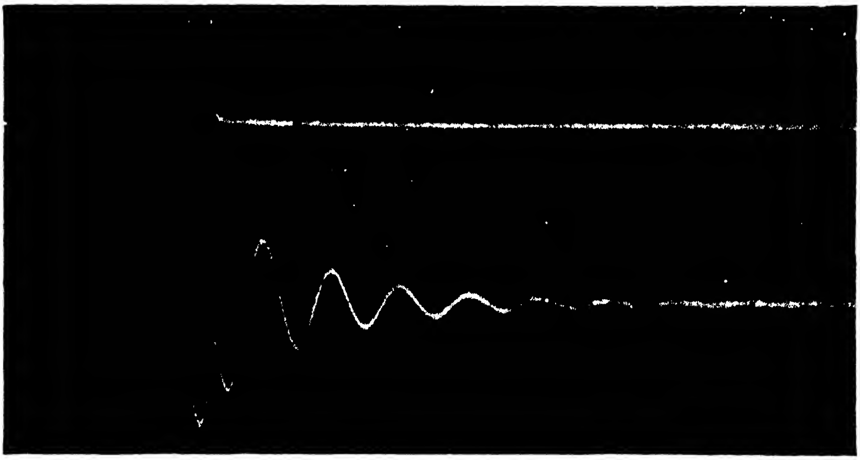


FIG. 53.—Similar to Fig. 52 but with longer pulse.

films was taken varying the length of pulse from less than 0.2 of the natural period of the circuit to several times the period. The amplitude of the first and second alternations were measured and their values plotted in

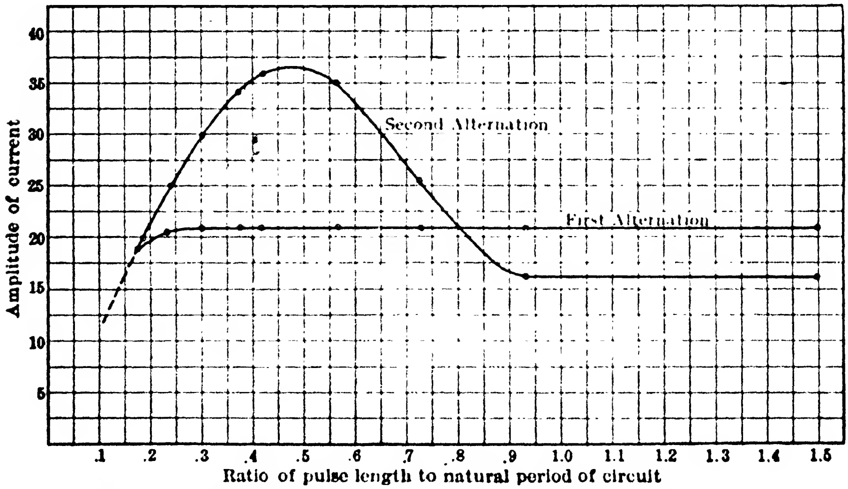


FIG. 54.—Variation in amplitude of oscillatory current with length of pulse.

terms of the ratio of pulse length to natural period of the circuit; the results are given in Fig. 54, and it is seen that they are in accord with the prediction of Eq. (87).

Fig. 55.—Irregular current produced by series of pulses.

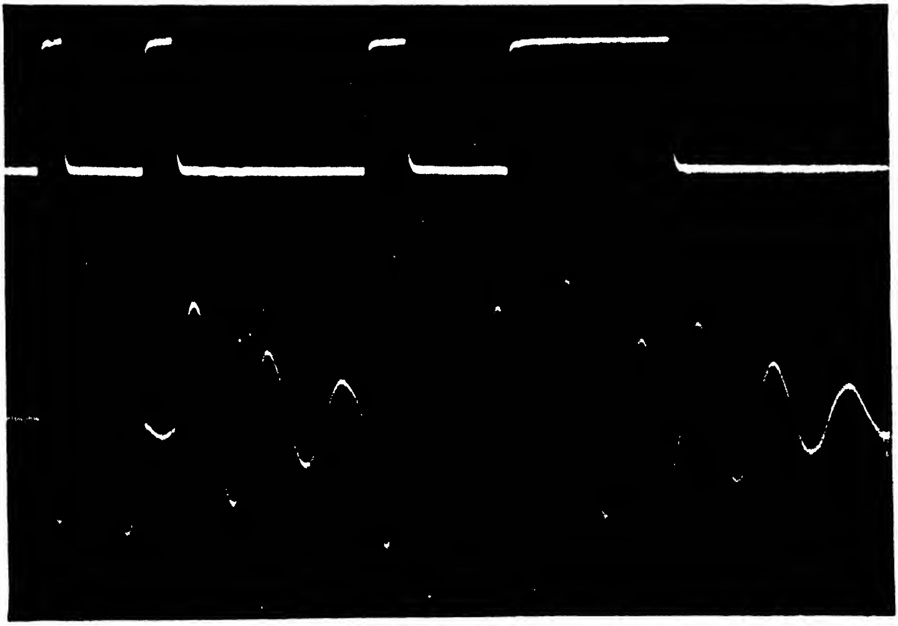


Fig. 56.—Irregular current produced by series of pulses.

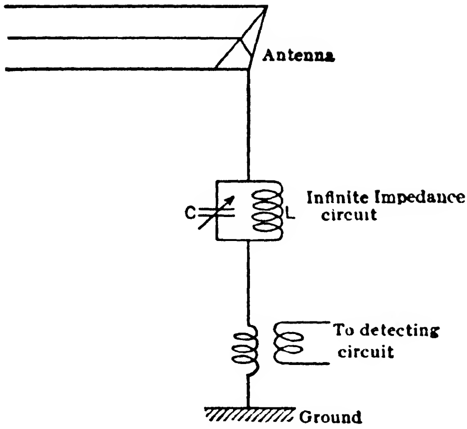


FIG. 57.—Showing the use of a parallel resonant circuit for weeding out undesired signals from an antenna. Such a parallel circuit is often called an "infinite impedance."

In case a series of irregularly timed pulses are impressed on an oscillatory circuit the resulting current will be of rather complex form; Figs. 55 and 56 show how such irregular pulses excite an oscillatory circuit. It is evident (Fig. 55) that pulses properly timed may practically neutralize one another as in this case the circuit was nearly dead after the last pulse.

Impulse Excitation of a Parallel Resonant Circuit.—A condenser and coil in parallel act like a circuit of very high resistance for an e.m.f. of the same frequency as that natural to the circuit. The value of the resistance is predicted by Eq. (56), Chapter I, and curves are shown in Chapter I, Figs. 88 and 89; because of this characteristic the circuit is often called an "infinite impedance" circuit. The Eq. (56) was derived from the steady state of the circuit, and so predicts nothing regarding the behavior of the circuit for other than steady alternating voltages; even in this case the equations are good only if the e.m.f. has been applied sufficient time for the transient terms to disappear.

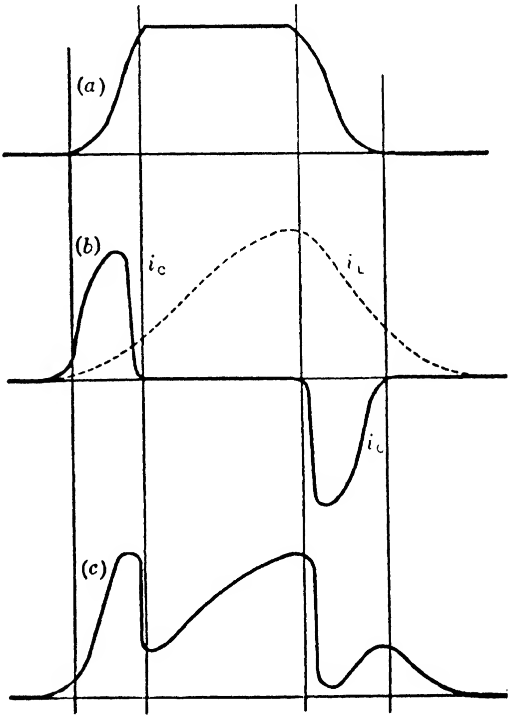


FIG. 58.—Action of an approximately rectangular pulse of e.m.f. (*a*) impressed across parallel circuit. Curves in (*b*) show the separate currents, the total current being that shown at (*c*).

Because of the "infinite impedance" characteristic, the circuit is often used to eliminate from a circuit some undesired frequency; thus in Fig. 57 the *L–C* parallel circuit is so adjusted that its natural period is the same as that of some undesired frequency which is impressed on the antenna. Now it is to be remembered that the *L–C* circuit offers "infinite impedance" only for the steady state and it is interesting to note the impedance offered by it to a pulse of e.m.f.

If the pulse is a square one, such as used in Figs. 50–53, the current flowing in the supply line (the impedance of the circuit other than that of the "infinite impedance" being neglected) will be about as shown in Fig. 58. The e.m.f. pulse form is shown in curve *a*; the full line curve of *b* shows the current flowing through the condenser and the dotted line that through the coil; in *c* is shown the actual current in the line, that is, the current which the "infinite impedance" circuit lets through. These curves, as mentioned before, are drawn on the assumption that the impedance of the rest of the circuit is negligible.

It would seem likely that if the circuit does have such a very high impedance for a certain frequency then it will offer a high impedance to a pulse, if this pulse is in the form of one alternation of a sine wave of the same frequency as that natural to the circuit. Fig. 59 shows the analysis of this case; *a* shows the form of pulse $e = E \sin \omega t$ (holding only between $\omega t = 0$ and $\omega t = \pi$); the curves of *b* indicate the currents through each branch, and curve *c* shows the current passed by the combined circuit. As the resistance in the parallel path is made to approach zero this line current approaches a true rectangular form, i.e., current of constant magnitude and equal to $2\pi fCE$, where f is the frequency of the e.m.f. of which the pulse is the alternation.

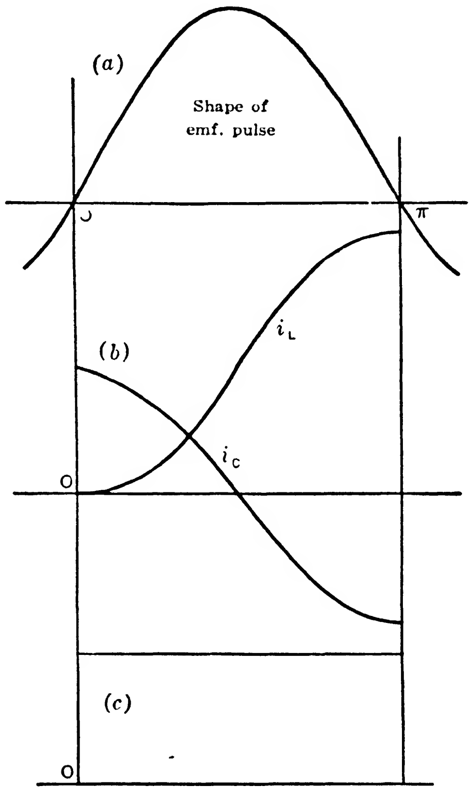


FIG. 59.—Effect of impressing a sinusoidal pulse of e.m.f. across a parallel circuit.

Analyzed mathematically, we have

$$i_C = \omega CE \cos \omega t$$

and

$$i_L = E\left(\frac{R}{R^2+(\omega L)^2}\sin \omega t - \frac{\omega L}{R^2+(\omega L)^2}\cos \omega t + \frac{\omega L}{R^2+(\omega L)^2}\epsilon^{-\frac{Rt}{L}}\right),$$

this being derived as a special case of Eq. (80).

For a good coil, i.e., one having a high value for $\omega L/R$, we have as an approximate value,

$$i_L = E\left(\frac{R}{(\omega L)^2}\sin \omega t - \frac{1}{\omega L}\cos \omega t + \frac{1}{\omega L}\epsilon^{-\frac{Rt}{L}}\right).$$

Adding to this the condenser current we get for the current passed by this parallel combination,

$$i = E\left(\frac{R}{(\omega L)^2}\sin \omega t + \left(\omega C - \frac{1}{\omega L}\right)\cos \omega t + \frac{1}{\omega L}\epsilon^{-\frac{Rt}{L}}\right).$$

If now the constants of the circuit are such that $\omega C = \dfrac{1}{\omega L}$, then

$$i = \frac{E}{\omega L}\left(\frac{R}{\omega L}\sin \omega t + \epsilon^{-\frac{Rt}{L}}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (88)$$

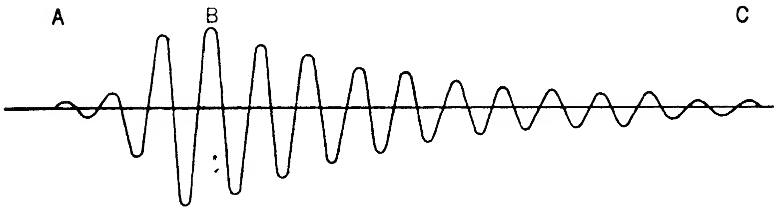


FIG. 60.—Form of voltage induced in a receiving antenna by the passage of one wave train as emitted from the ordinary spark transmitter.

Oscillatory Circuit Excited by a Damped Sine Wave.—Let us consider a circuit made of L, R, and C in series as e.g., the ordinary antenna, to be excited by a damped sine wave of voltage such as is induced many times each second in an antenna by a signal from a distant spark station. The voltage caused in an antenna by every discharge in the transmitting circuit is not exactly representable by a damped sine wave, because for the first few cycles its amplitude is increasing instead of decreasing; this is indicated in Fig. 60, which shows about the form of voltage actually induced in an antenna. The part A–B is relatively short ($2\frac{1}{2}$ cycles for 20 per cent coupling in the transmitting station) and the part B–C is really a damped sine wave represented by $E\epsilon^{-kt}\sin pt$.

The form of current induced in the antenna is determined in the usual way by putting the sum of its reactions equal to the impressed force.

$$L\frac{di}{dt}+Ri+v=E\epsilon^{-kt}\cos pt, \quad . \quad . \quad . \quad . \quad . \quad . \quad (89)$$

v being the voltage across the condenser.

Writing the impressed force as a cosine function indicates that we are going to consider the case of maximum voltage at $t=0$; such is the case when the circuit we are considering is acted upon by the oscillatory current set up in a neighboring circuit by the discharge of a condenser. Eq. (89) may be written in the form,

$$\frac{d^2v}{dt^2}+2\alpha\frac{dv}{dt}+(\omega^2+\alpha^2)v=\frac{E}{LC}\epsilon^{-kt}\cos pt, \quad . \quad . \quad . \quad . \quad . \quad (90)$$

by using the relations

$$i=C\frac{dv}{dt}, \quad \alpha=\frac{R}{2L}, \quad \text{and} \quad \omega^2+\alpha^2=\frac{1}{LC}.$$

Differentiating (89) twice with respect to time and eliminating the right-hand number by multiplying (90) by (p^2+k^2), its first derived equation by $2k$, and adding both the resulting equations to the second derived equation, we obtain

$$\frac{d^4v}{dt^4}+2(k+\alpha)\frac{d^3v}{dt^3}+[(p^2+k^2)+(\omega^2+\alpha^2)+4\alpha k]\frac{d^2v}{dt^2}$$
$$+[2\alpha(p^2+k^2)+2k(\omega^2+\alpha^2)]\frac{dv}{dt}+[(p^2+k^2)(\omega^2+\alpha^2)]v=0. \quad . \quad (91)$$

This equation is in standard form for integration, the value of v being of the form

$$v=V_1\epsilon^{-kt}\sin(pt+\theta)+V_2\epsilon^{-at}\sin(\omega t+\phi) \quad . \quad . \quad . \quad (92)$$

From this equation we get

$$i=I_1\epsilon^{-kt}\sin(pt+\theta')+I_2\epsilon^{-at}\sin(\omega t+\phi') \quad . \quad . \quad . \quad (93)$$

This solution shows that the current flowing in the circuit is made up of two components, one of the same frequency as the impressed force and one of the natural frequency of the circuit.

This might have been surmised from an elementary analysis of the problem. Suppose the damping of the impressed force is zero, then Eq. (80), page 334, would give the correct solution and this has two terms, one of the frequency of the impressed force (which is the solution for the steady state) and the transient term which dies away at a rate fixed by the decrement of the circuit.

The relative amplitudes of the two currents, I_1 and I_2, will evidently depend in some way upon the relative values of the damping factors, k and α, also upon the relative values of the frequencies fixed by p and ω. By getting the values of v, dv/dt, and d^2v/dt^2 from (92) and substituting them in (90) we find the value of V_1 to be

$$V_1 = E\frac{\omega^2+\alpha^2}{\sqrt{[\omega^2-p^2+(\alpha-k)^2]^2+4p^2(\alpha-k)^2}}, \quad . \quad . \quad . \quad (94)$$

and

$$\tan\theta = \frac{\omega^2-p^2+(\alpha-k)^2}{2p(\alpha-k)}.$$

Now by substituting in Eq. (92) the initial conditions that when $t=0$ $v=0$, then differentiating (92) and in this equation putting $dv/dt=0$ when $t=0$ we get the value of V_2 in terms of V_1. Substituting the value of V_1 from (94), we get,

$$V_2 = E\frac{(\omega^2+\alpha^2)\sqrt{\omega^2+(\alpha-k)^2}}{\omega\sqrt{[\omega^2-p^2+(\alpha-k)^2]^2+4p^2(\alpha-k)^2}}, \quad . \quad . \quad . \quad (95)$$

and

$$\tan\phi = \frac{\omega}{\alpha-k}\,\frac{\omega^2-p^2+(\alpha-k)^2}{\omega^2+p^2+(\alpha-k)^2}.$$

From the values of V_1 and V_2 we find at once I_1 and I_2 by using the relations $I_1 = pCV_1$ and $I_2 = \omega CV_2$:

$$I_1 = E\frac{p}{L\sqrt{[\omega^2-p^2+(\alpha-k)^2]^2+4p^2(\alpha-k)^2}}. \quad . \quad . \quad . \quad (96)$$

$$I_2 = E\frac{\sqrt{\omega^2+(\alpha-k)^2}}{L\sqrt{[\omega^2-p^2+(\alpha-k)^2]^2+4p^2(\alpha-k)^2}}. \quad . \quad . \quad (97)$$

The values of θ' and ϕ'—Eq. (93)—are determined from the values of θ and ϕ given above, by increasing each of them by $\pi/2$.

The exact form of current in the circuit is now fixed by the values of I_1, I_2, k, α, p and ω. It will be evident that if both damping factors are low and nearly equal, and the two frequencies, fixed by p and ω, are nearly equal, the conditions are the same as those for the secondary current in coupled circuits as illustrated in Fig. 26 of this chapter. If $p=\omega$ there can be no beats; for all values of damping, the current, with frequency $\omega/2\pi$ increases in value from zero to a certain maximum and then decreases again.

An analysis due to Bjerknes[1] shows that this current can be represented by the equation

$$i = mCM \cos (mt + \psi), \quad \ldots \ldots \quad (98)$$

in which $m = \frac{p+\omega}{2}$ and ψ is the phase of the impressed e.m.f. at time $t=0$.

If we let $n = \frac{p-\omega}{2}$, $a = \frac{k+\alpha}{2}$, $b = \frac{k-\alpha}{2}$, then

$$M = \left(\frac{E}{4mLC}\right)^2 \frac{1}{(n+b)} \left\{ P_1 + 2\left(\frac{1+\cos 2\psi}{m}\right) P_2 + 2\left(\frac{\sin 2\psi}{m}\right) P_3 \right\}, \quad . \quad (99)$$

in which

$$P_1 = \epsilon^{-2at}(\epsilon^{-2bt} + \epsilon^{2bt} - 2 \cos nt);$$

$$P_2 = \epsilon^{-2at}(n\epsilon^{2bt} - n \cos nt - b \sin 2nt);$$

$$P_3 = \epsilon^{-2at}(b\epsilon^{2bt} - b \cos 2nt + n \sin 2nt).$$

In certain cases the form of M is simpler than indicated in Eq. (99).

If $p=\omega$ and $k=\alpha$

$$M = \pm \frac{Et}{2mLC} \epsilon^{-at}. \quad \ldots \ldots \ldots \quad (100)$$

If $p=\omega$ and $k \neq \alpha$

$$M = \pm \frac{E}{4mbLC} \epsilon^{-at}(\epsilon^{-bt} - \epsilon^{bt}). \quad \ldots \ldots \quad (101)$$

If $p \neq \omega$ and $k=\alpha$

$$M = \pm \frac{E}{2mnLC} \epsilon^{-at} \sin nt. \quad \ldots \ldots \quad (102)$$

In case neither the damping factors nor frequencies are the same the general form given in Eq. (99) must be used.

In Fig. 61 are shown the forms of current in the oscillatory circuit for the cases given in Eqs. (100) and (102).

Resonance Curve of an Oscillatory Circuit Excited by Damped Sine Waves.—In Chapter I we analyzed the action of an oscillatory circuit (L, C, and R in series) when excited by an alternating e.m.f. of constant amplitude and showed that the form of the curve obtained when either C, L, or f was varied (E being held constant in amplitude), enables us to determine the decrement of the circuit. The form of the resonance curve for the steady state depends only upon the relation between the impedance of the circuit and the frequency of the impressed force. When

[1] V. Bjerknes, "Electrical Resonance," Wied. Ann., 1895, Vol. 55.

such a circuit is excited by a damped sine wave the reading of the indicating device for showing resonance will depend on both of the terms in Eq. (93). If a hot-wire ammeter is used to show resonance it is evident that its reading will depend upon the average integral of the square of each of the currents I_1 and I_2. Bjerknes and others have analyzed the value of this integral and by somewhat lengthy deductions have obtained the relation,

$$I^2=\frac{E^2}{16L^2}\frac{k+\alpha}{k\alpha}\frac{1}{(p-\omega)^2+(k+\alpha)^2}, \quad . \quad . \quad . \quad (103)$$

this holding good only for α much less than ω, k much less than p, and for frequencies close to the resonance point.

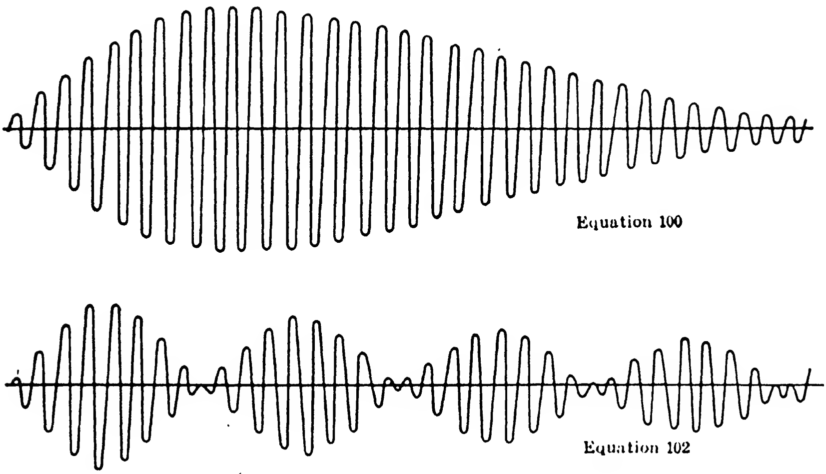


FIG. 61.—Two possible forms of current in the antenna excited by the wave trains from a spark transmitter.

At resonance ($\omega=p$) this reduces to

$$I_r^2=\frac{E^2}{16L^2}\frac{1}{k\alpha(k+\alpha)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (104)$$

From (103) and (104)

$$\frac{I^2}{I_r^2}=\frac{(k+\alpha)^2}{(p-\omega)^2+(k+\alpha)^2},$$

so

$$\frac{I_r^2-I^2}{I^2}=\frac{(p-\omega)^2}{(k+\alpha)^2}.$$

From this we get

$$k+\alpha=(p-\omega)\sqrt{\frac{I^2}{I_r^2-I^2}}.$$

If we now introduce the decrements, instead of damping factors, we have, putting $k=n\delta_1$ and $\alpha=f\delta_2$,

$$n\delta_1+f\delta_2=2\pi(n-f)\sqrt{\frac{I^2}{I_r^2-I^2}}.$$

It is to be remembered that n is the frequency of signal impressed on the wave meter and f is the frequency to which the wave meter is tuned. If now n is nearly equal to f, so we may put without much error $f/n=1$, we get,

$$\delta_1+\delta_2=2\pi\left(\frac{n-f}{n}\right)\sqrt{\frac{I^2}{I_r^2-I^2}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (105)$$

in which n is the frequency impressed on the wave meter and f is the frequency to which the wave meter is tuned.

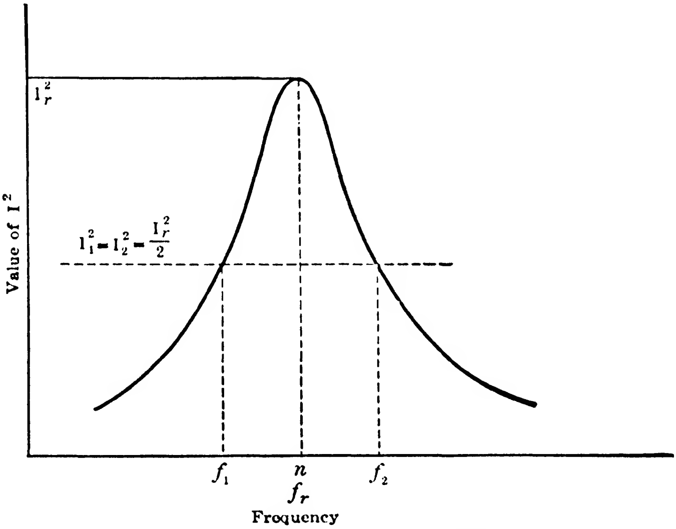


FIG. 62.—Resonance curve of a circuit excited by a series of damped sine waves.

If then we have plotted a curve showing the variation of the current in the oscillatory circuit, as its natural frequency is varied, we can calculate the sum of the decrements of the circuit and exciting voltage; if one of them is known the other may then be obtained. The curve between $(\text{current})^2$ and f will have the shape indicated in Fig. 62; when $f=n$, I has its maximum value I_r, and it decreases as f departs from n. The amount of decrease in I for a given difference between n and f is the same whether f is greater or less than n, provided the value of n/f does not differ materially from unity.

If we then read the two values of f (call them f_2 and f_1, f_2 being greater than n and f_1 being less than n), so chosen that the current is reduced to $1/\sqrt{2}$ of its resonance value we shall have

$$\delta_1+\delta_2=2\pi\frac{n-f_1}{n}\sqrt{\frac{0.5I_r^2}{I_r^2-0.5I_r^2}}=2\pi\frac{n-f_1}{n},$$

and also

$$\delta_1+\delta_2=2\pi\frac{f_2-n}{n}.$$

Adding these two values, we get

$$\delta_1+\delta_2=\pi\frac{f_2-f_1}{n}=\pi\frac{f_2-f_1}{f_r}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (106)$$

In this equation f_r is that frequency of the circuit which gives greatest current; this frequency we know to be practically the same as the frequency of the impressed force which we have been calling n.

As the frequency of a circuit varies with the square root of the capacity in the circuit, we may write Eq. (106) in terms of the amount of capacity used in getting the resonance curve. If C_r is the value used to get the maximum current and C_2 and C_1 correspond to f_2 and f_1 of Eq. (106) then we have, very nearly

$$\delta_1+\delta_2=\frac{\pi}{2}\frac{C_2-C_1}{C_r}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (107)$$

This is the equation generally used when using a wave meter for getting the decrement of a transmitting set; although approximations have been made in deducing it, the errors incurred are small if the sum of the two decrements is small (say less than 0.25), which is always the case in practical radio set.

CHAPTER IV

GENERAL VIEW OF RADIO COMMUNICATION

Wave-motion.—Since the transmission of intelligence by radio is brought about by sending out so-called electromagnetic waves, and since the transmission of these waves is somewhat similar to that of other kinds of waves, such as light, sound, heat, water, etc., we will first discuss wave-motion in a simple manner and then apply this to electromagnetic wave-motion.

In wave-motion a stress is transmitted from one point to another in an elastic medium without any permanent displacement of the medium itself in the direction in which the stress is transmitted. Thus, if a pebble be dropped in a still pond at *A*, an up-and-down motion of the water will be set up at *A*, which will be transmitted to a point at *B*, without any resultant motion of the water itself in the direction *A–B*, as evidenced by the fact that a float placed between *A* and *B* will not be displaced toward *B*.[1]

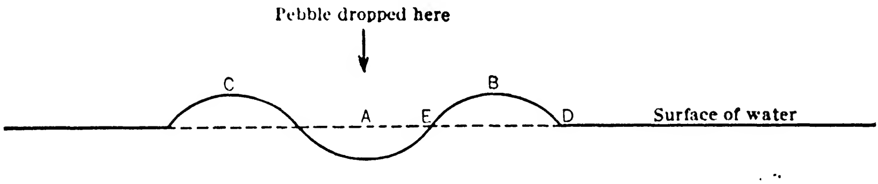


Fig. 1.—Cross-section through surface of water, immediately after pebble has been dropped at *A*.

In the case of water waves the pebble dropped at *A*, Fig. 1, will displace the water directly under *A* to the right and left (in fact in all directions) towards *B* and *C*, and will produce the bulges at *B* and *C* known as "crests." These bulges are due to the fact that the water displaced from *A* tends to raise the level all around *A*, but, on account of inertia, this cannot be done quickly enough, with the result that the level is raised most at *B* and *C*, and hence the "crests."

Considering the disturbances to the right of *A* alone, the particles of water in the space *EBD* will, because of gravitational forces, seek the average level of the water, and, in so doing, the bulge *EBD* will be made to

[1] If a wave is so high as to "break," i.e., an impure wave, this statement is not quite true.

disappear, and a depression will be created thereat due to the fact that, on account of inertia, the particles move beyond the average " level position." Not only is this the case, but the particles to the right of D will, one after the other, be urged, as if elastically tied together, to perform motions similar to those of the particles within EBD, so that a crest will presently appear to the right of D at the same time that a depression or " trough " is created in the region EBD. The result is that a " trough " and a " crest " of a wave will *appear to move* in all directions, away from the center of disturbance at A, and with a definite velocity.

It must be understood that the motion of the particles is limited to a small region around their positions of equilibrium, and that the wave is propagated by imparting this motion to one particle after the other, while each particle, after the disturbance has passed, remains in practically the same position as it originally occupied.

An exact analysis of the motions of the particles is extremely complex and will not be attempted here, but we will give certain well-known ele-

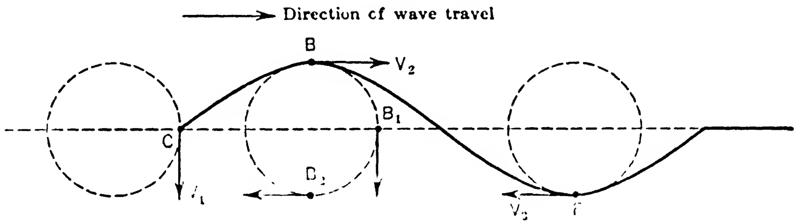


FIG. 2.—Motion of the water associated with a wave.

mentary facts regarding it because of analogy between certain points regarding water waves and electromagnetic waves.

Theory indicates and experiment corroborates that the water particles within the path of a wave execute motions which are in the simplest cases circular. Taking this case of circular motion of the particles and considering Fig. 2, at C a particle of water will just be passing through the undisturbed level and moving with a velocity V_1 in a downward direction, while at B another particle will be moving with a velocity V_2 in a horizontal direction. Thus, at every point within the volume of the water involved in the wave propagation each and every particle will be executing a circular motion in a clockwise direction, and the formation of a crest or trough is the result of various particles being, at any time, at different stages of their circular motions. Thus where the particles are moving horizontally in the same direction as that in which the wave is being propagated we obtain a crest, since a large number of particles are then at or near the top of the circle; where the particles are moving horizontally, but in a direction opposite to that of the propagation of the wave, a trough is obtained, since a number of particles are then at or near the bottom of the circles representing their respective motions.

Consider now the question of energy of a particle at B; it has velocity in the direction of the propagation of the wave, and, furthermore, it is displaced vertically upward with respect to the undisturbed level of the water; such a particle has kinetic energy in the direction of the wave propagation plus potential energy with respect to the average level. As the particle rotates the vertical displacement from the average level becomes less and less and so does the component of its velocity in the direction of propagation of the wave; so that when it occupies the position B_1, its potential energy has become zero and so has the kinetic energy in the direction of the wave propagation. As it moves still further it suffers a negative displacement and its velocity acquires a component parallel to direction of propagation of the wave but opposite thereto, so that by the time it has reached the point B_2 its potential and kinetic energy are equal and opposite in sense to those which it had while at the point B.

It may be shown that the potential energy of a particle at any point is exactly equal to the kinetic energy in the direction of propagation of the wave, provided that the wave is not changing shape. Because of this there must be a fixed relation between the displacement of a particle above or below the undisturbed level of the water and the component of the velocity of the same particle parallel to the direction of propagation of the wave.

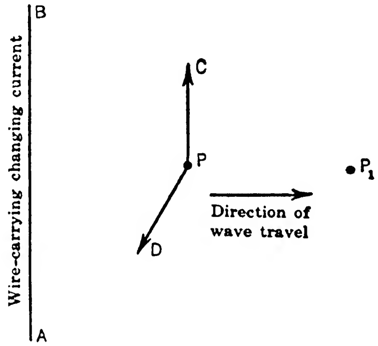


FIG. 3.—Electric and magnetic fields set up at P by wire A-B.

Again, the total energy of a particle of water (potential and kinetic in the direction of propagation of the wave) is continually changing as the wave progresses, the particle in question passing its energy along to the particle adjacent to it, in the direction of the wave propagation; this transfer of energy from one particle to another is the underlying principle of all wave-motion in water.

Electromagnetic Waves.—These are due to a disturbance of an electromagnetic nature and are such that they produce at points all around the center of disturbance a varying magnetic field and a varying electric field. Thus, if a wire, such as AB (Fig. 3) in space, has an alternating current flowing through it for a short interval of time, it will set up an alternating magnetic field and an alternating electric field all around itself, which fields, starting from the vicinity of the conductor, will travel away from it with the velocity of light. In so far as to set up a magnetic field or an electric field requires energy, it follows that a certain amount of energy must be detached from that available in the conductor in order that the electromagnetic disturbance be created at all. Thus energy is said to be "radiated," and the phenomenon itself is known as "electromagnetic radiation" or simply "radiation."

The electric and magnetic fields of radiation, at any one point, are in space quadrature, but *they are at all times in time phase with each other.* And not only is this the case, but there is a fixed relation between the values of the electric and magnetic fields at any instant; this relation is based upon the fact that, in order for the wave of electromagnetic disturbance to exist in space the energy, per unit volume of the medium, possessed by the electric field, must be equal to that possessed by the magnetic field. Thus the total amount of energy at any one point and instant is equal to twice that possessed by either field. The value of this energy is changing as the intensity of the two fields changes, and, as a matter of fact, the energy is being transferred from one point to the next by the elastic properties of the medium in which the disturbance travels. (The discussion of electromagnetic waves here given uses the idea of a medium as a carrier of the disturbance; to make this medium fill the role it is supposed to play in modern electron theory it must be considered as *the superimposed electric fields of all the electric charges in the universe.*) The reader will note that this is analogous to the case of water waves, where, at any point, the potential energy per unit volume is equal to the kinetic energy in the direction of propagation of the wave, and the total amount of energy is transferred from point to point within the medium, thus bringing about the conditions of wave motion.

In the case of water waves or electromagnetic waves if, at any point, some of the energy in one of the two forms (potential and kinetic for water, and electric and magnetic for electromagnetic waves) be withdrawn from the space wherein the wave exists, part of the energy in the other form will be immediately transformed into the former, with the result that equality of the two forms of energy will still apply, but the crests and troughs of the water waves will not be as high as before, nor will the amplitude of the electric and magnetic field intensities be as large as before.

It must be noted that this phenomenon is different from that of the creation of the ordinary magnetic or electric field around the conductor, which field never reaches far from the conductor (with appreciable intensity), and does not represent energy permanently *removed* from the conductor, since the variation of this field induces electromotive forces in the conductor, and thus an *exchange of energy* is kept up between the conductor and the field. This field is known as "induction field" to distinguish it from the "radiation field," wherein we are more vitally interested.

In the case of the "radiation field" at any point such as P, Fig. 3, the magnetic field would act along PD and the electric field along the line PC at right angles to PD, while the disturbance would travel in the direction of the arrow at right angles to both PC and PD. Both fields change in value and in direction in accordance with the variations of the current in the conductor producing the disturbance; if this is harmonic the two fields will change harmonically. At some other point such as P_1 the disturbance will arrive a little later than at P with the result that the fields at P and P_1 are out of phase, the phase difference depending upon the frequency and the velocity of propagation. Fig. 4 shows a

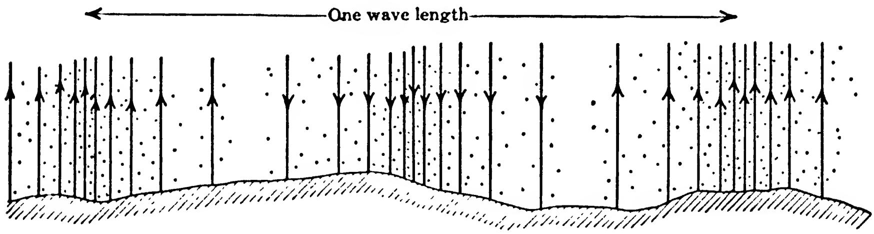


FIG. 4.—The electric field of the radiation wave is nearly perpendicular to the earth's surface; the magnetic field (shown by the dots) is parallel to the earth's surface. Magnetic and electric fields have their maximum intensity, at any instant, at the same place. The magnetic field reverses its direction when the electric field reverses. Magnetic fields are shown by dots in this figure irrespective of whether their direction is into, or out of, the page.

cross-sectional view of the magnetic and electric fields associated with the ordinary type of wave used in radio communication. The magnetic field has a direction parallel to the earth's surface and is shown by the dots of Fig. 4; the electric field is essentially perpendicular to the earth's surface. The two fields extend to some distance into the earth, becoming rapidly weaker with depth. The fields penetrate deeper into dry earth than into moist earth, penetrating a minimum depth into sea water.

The electric and magnetic fields always have their maxima and minima at the same points as the wave travels along with the speed of light away from the source of disturbance, i.e., away from the transmitting antenna.

Considering one of the two fields, say the magnetic, and plotting the instantaneous value of the field against distance from the conductor of Fig. 3 we would obtain a curve such as A in Fig. 5, which applies to any particular instant of time. A little later the plot of the field would be given by the curve B in so far as the intensity at every point in space will by then have varied so as to make the new curve possible. The wave has thus been shifted in the direction of the arrow by the amount C. If we plot a number of such curves we would find that the wave has shifted the distance λ, by the time the magnetic field has completed a cycle. We thus have that if:

λ = wave length in centimeters;
v = velocity of propagation of wave in centimeters per second;
f = frequency of the field in cycles per second.

Time for one cycle $= 1/f$

and
$$\lambda = \frac{1}{f} v;$$

or
$$v = \lambda f, \qquad (1)$$

which is the fundamental relation for any wave propagation.

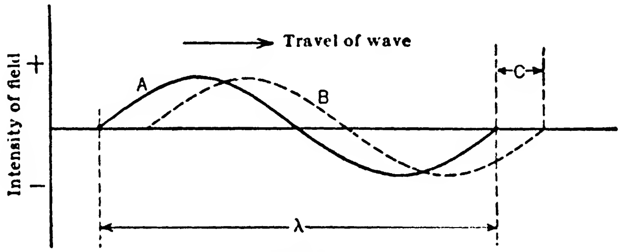


FIG. 5.—Illustrating one "wave length."

Velocity of Propagation.—For electromagnetic waves this is equal, as already stated, to the velocity of light when the wave is being propagated through air, but in general, for any medium, we have the following:

V = velocity of light in vacuum;
v = velocity of propagation of wave in any homogeneous medium;
μ = magnetic permeability of the medium;
k = specific inductive capacity (inductivity of medium).

$$v = \frac{V}{\sqrt{\mu k}}, \qquad (2)$$

since for air $\mu = 1$ and $k = 1$, $v = V$.

This formula neglects any possible effect on the velocity of propagation, of any losses occurring in the medium, due to conduction, hysteresis, or dielectric losses, etc.

As shown by Eq. (2) the velocity of propagation is dependent only upon the nature of the medium through which the wave is moving (its magnetic and electric constants), and is not affected by the wave length or by the frequency. Thus, in considering the conductor AB of Fig. 6, as the source of an electromagnetic wave, this wave would spread out in the direction of C with the velocity of light, but in passing through the mountain to the left the velocity would be lower. If instead of a mountain we should have a sheet of metal for which the value of k is ordinarily considered as infinitely large, then:

$$v = \frac{V}{\sqrt{\mu \infty}} = 0,$$

that is, the wave would stop completely;[1] the energy of the wave would be partly absorbed by the metal in the production of electric currents therein and partly reflected back. Not only would the wave travel in the direction C and D, but in every other direction, up into the air in

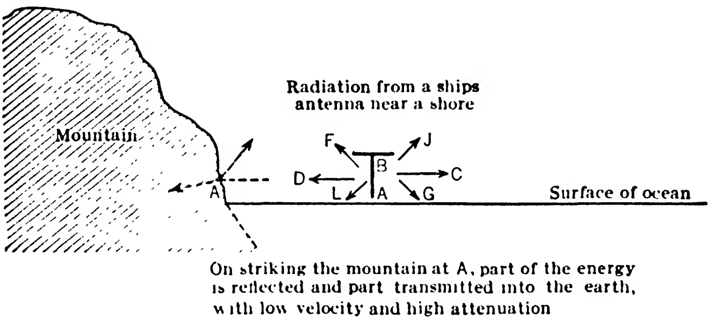


FIG. 6.—Energy radiated from an antenna is subject to the same laws of reflection, refraction, and absorption as ordinary light.

the direction of F and J, and into the water and earth in the direction of L and G.

As the wave travels outward some of the energy is absorbed by the medium if this be other than air: even in air there is some absorption of energy, especially in daylight, due to ionization making it partially conducting; and, in other materials, losses are produced by the varying electric and magnetic fields which absorb energy from the wave itself. The result of this is, of course, that a distance is soon reached where practically no more energy is available and the disturbance ceases to be communicated (in measurable intensity) any farther. The distance over which a certain amount of energy will travel through air, even in daylight, before being absorbed is, of course, much greater than through solid matter or even liquid, due to the fact that the losses (eddy currents, magnetic

[1] This conclusion is not strictly accurate; the velocity in such a case would be much less than the velocity of light, but would not be zero. The discrepancy arises from the very elementary viewpoint from which wave-motion is here considered.

hysteresis, dielectric hysteresis, etc.) are greater in these substances than in air.

Electric waves traveling over wires do not have as much velocity as the radio waves traveling through free space. The velocity of waves in wires depends to some extent upon the frequency, being greater for the higher frequencies. In the following table are given the velocity of travel of waves over pairs of wires in cables, with heavy loading, light loading, an unloaded pair, and open wires.

	Miles per Second
Cable pair 88 millihenry coils every 3000 ft..	10,000
Cable pair 44 millihenry coils every 6000 ft.	20,000
Cable pair No. 16 B. & S. wire unloaded...	130,000
Open wire pair	180,000
Radio.......	186,000

Waves Used in Communication.—If communication is to be carried on by radio waves, a fundamental requirement has to do with frequency. *The power radiated from an antenna varies as the square of the frequency,* other factors remaining the same; because of this fact, we find used in radio only frequencies above about 15,000 cycles per second. Generally the frequency is much higher than this, so we find the two terms kilocycle (1000 cycles) and megacycle (1,000,000 cycles) used. At present the lowest radio frequency is 15 kc. (kilocycles) and the highest about 30 mc. (megacycles). Experiments are being carried on with frequencies as high as about 300 mc., but such high frequencies have at present no commercial application.

In Fig. 7 [1] is shown the distribution of all the electromagnetic waves at present known. Below the chart are given the frequencies, and corresponding wave lengths, in meters or centimeters; above are given the wave length in Ångström units, the unit used in light measurements.

The use of radio waves is controlled by our Federal Government, in cooperation with those of other nations. Because of the world-wide character of radio communication, its control is necessarily one of international scope.

In so far as our government controls radio waves, they have been assigned roughly as shown in Fig. 8.[2] The broadcasting field embraces a band only about 1000 kc. wide, 550 kc. to 1500 kc. (547 meters to 200 meters). Below this range of frequencies, the 500-kc.-wide band is used mostly for point-to-point communication, the lower ones being used by the high-powered stations for transoceanic radio channels. Above the

[1] Electronics, April, 1930.
[2] Electronics, April, 1930.

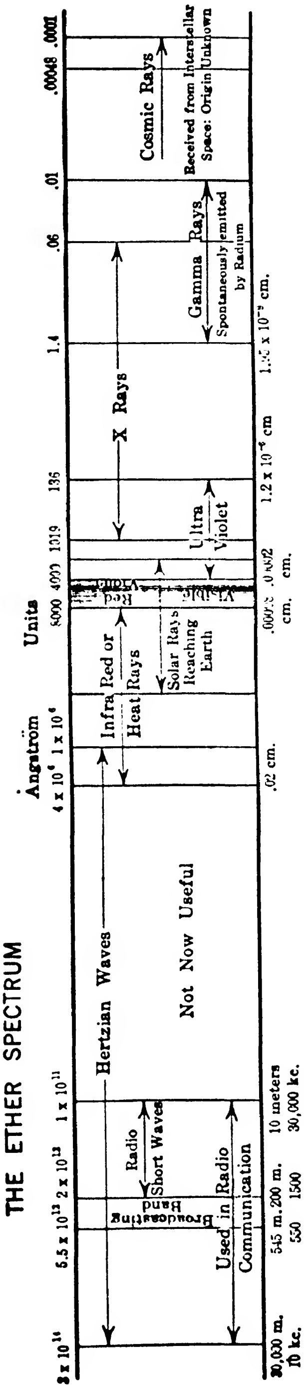


FIG. 7.—Wave lengths and frequencies of all the radiation waves at present known.

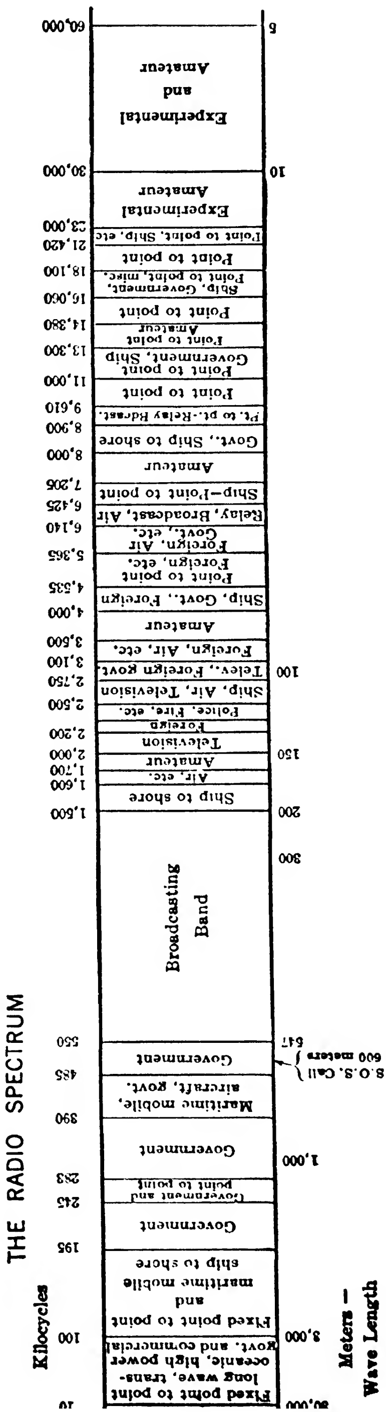


FIG. 8.—Assignment of radio waves, by the government, for different classes of radio communication.

broadcast range, there is a band 20,000 kc. wide; in spite of the large number of channels available in this wide band, they are all assigned by the government for some useful service.

Transmission on Different Waves.—The wave length which shall be used for a definite service is fixed by compromise of several viewpoints, the first of which is probably the reliability of the communication which can be obtained, with a reasonable amount of power at the transmitter. The availability of the desired wave length, the amount of disturbing effects at that frequency at the location of the receiver, the cost of the transmitter, possibility of directive transmission, etc., all must be considered.

The large transmitting stations, using hundreds of kilowatts of power, transmitting antennas several hundred feet high and receiving antennas several miles long, use the long waves of the radio spectrum (15 to 35 kc.). They furnish practically continuous transoceanic communication, being interrupted only during heavy electrical storms in the vicinity of the receiver. However, there are not many of these long-wave channels available.

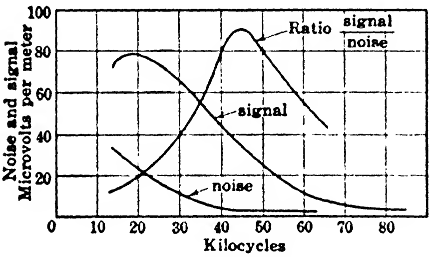


FIG. 9.—At a given location it will generally be found that there is a most favorable frequency, where the "signal to noise" ratio is a maximum.

With higher frequencies, the distance over which reliable communication can be obtained decreases up to frequencies of several megacycles. Thus a 5-kw. broadcasting station at 1 mc. gives reliable transmission over only a few hundred miles at night and much less than this in the daytime. For frequencies above 3 mc., however, the reliable distance again increases, and a wave of 15 meters (20 mc.) gives reasonably reliable communication during the day over many thousands of miles. Such a short wave, in combination with another about twice as long, will give nearly continuous communication up to 8000 to 10,000 miles. One is best for day communication and the other for night. They have what is called a " skip distance," which signifies that a short wave may be heard for a distance of perhaps 50 miles from the transmitter and then not heard for perhaps 2000 miles; after this distance is reached, the signals may be received strongly up to possibly 8000 miles. The distance between the 50 miles and 2000 miles is called the skip distance.

For a given transmitter it is frequently found that some wave length gives a better signal at the receiving point than either shorter or longer waves; such an effect is shown in Fig. 9. In this figure the signal received across the Atlantic is plotted as a function of frequency, and it is seen that about 20 kc. was the best. However, there was much " noise " at this frequency also, so that it would not be the best frequency to use, even if it had been available.

Different Types of Waves.—Radio communication is carried on by either dot-and-dash code (telegraph) or by the spoken word; different

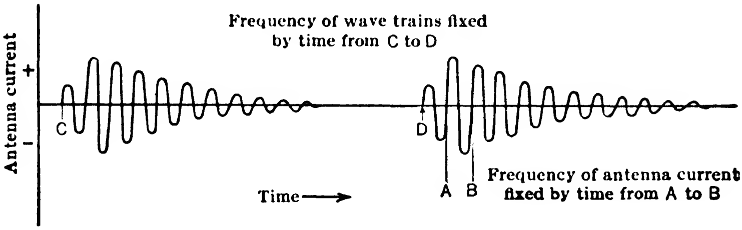


Fig. 10.—Type of antenna current in a spark station.

types of waves are used. In *spark telegraphy* a series of damped wave trains are sent out; the waves of a train follow each other at radio frequency and one train follows another at audio frequency—this type of signal is indicated in Fig. 10, which shows two wave trains. The frequency generally used in this type of communication is 500 kc. (600 meters), and the number of wave trains per second varies in different transmitters from 240 per second to 1000 per second. At the transmitter a spark initiates each wave train, so the number of wave trains is fixed by the number of sparks per second at the transmitter.

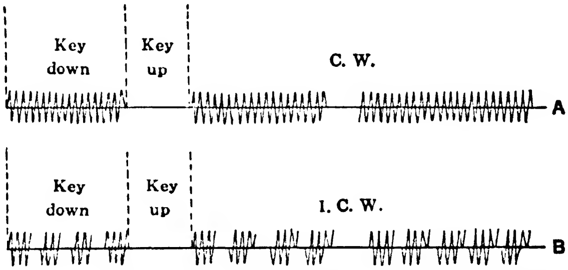


Fig. 11.—The type of wave used for a radio telegraph channel; in *A* an ordinary continuous wave signal is shown and in *B* the same signal is shown in "interrupted continuous wave" type of radiation.

In *continuous wave telegraphy* (C.W.) the antenna at the transmitter is excited continuously by a radio-frequency current of constant amplitude, as long as the key is held depressed. This is indicated in diagram *A* of Fig. 11, which shows the letter " u " sent by continuous wave telegraphy. The frequency of the current in the antenna might be 500 kc., the duration of a dot 0.05 second and of the dash 0.15 second. Then the dot would consist of 25,000 cycles of 500-kc. current, and the dash would be 75,000 cycles. This type of wave requires a special type of receiving set to be available; there must be an oscillating vacuum tube in the receiver or no audible signal will be obtained.

In *interrupted continuous wave telegraph* (I.C.W.), the radio-frequency current is fed into the antenna in groups of waves, as long as the key is held down. The group frequency depends upon the construction of the transmitter; it is generally about 500 per second. Diagram B of Fig. 11 shows an "u" sent by this method of communication. Assuming that the radio frequency is the same as for the C.W. signal, each group will consist of 500 cycles of current, and there will be 25 groups for a dot and 75 groups for a dash. This type of communication can be picked up by the same simple type of receiver as is required for spark telegraphy.

When the radio communication is carried on by *radio telephony*, there is no key for opening and closing the transmitter circuit; the antenna is continuously excited by a radio-frequency current. The amplitude of the radio-frequency current, however, goes up and down in accordance with the form of the sound wave to be transmitted. In case a musical note is being sent out, the radio-frequency current has a periodic amplitude fluctuation, the form of the fluctuation depending upon that of the musical tone. A radio-frequency current for a simple tone is shown in diagram *A* of Fig. 12, and in diagram *B* is shown the more complicated form possibly associated with part of a spoken word.

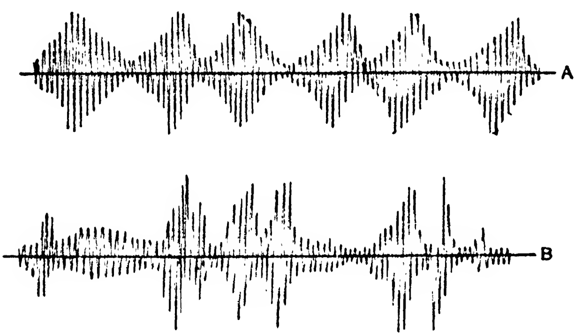


FIG. 12.—A radio telephone signal; in *A* is shown a musical note and in *B* part of a spoken word.

Frequency Standardization.—At present, there are in the United States over 600 transmitting stations operating in the broadcasting band of frequencies, i.e., from 550 to 1500 kc., and of course there are hundreds of other stations on the lower and higher frequencies. To prevent these numerous stations from interfering with one another, it is essential that they transmit on their assigned frequency with very little deviation therefrom. This problem has two aspects; how accurately can the government bureau maintain and measure frequency standards, and how well can a commercial station adhere to its assigned frequency?

By the use of small pieces of piezo electric quartz, suitably excited by a vacuum tube, it is found possible in the laborotory to maintain frequencies to about one part in a million. The engineers of the Bell Laboratories have done remarkable work along these lines; typical of their results are those given in Fig. 13.[1] This oscillator had a deviation from the mean value of only a few parts in 10^7, during 12 days of operation. This source of constant frequency was a quartz oscillator with accurate temperature control. As showing the degree of standardization now obtaining between different nations, the following tabulation gives the result of comparison of international standards by means of a quartz oscillator taken from our Bureau of Standards to several national laboratories and then back to the Bureau of Standards. Two different quartz oscillators, temperature controlled, were used as the medium of comparison. Dellinger reports [2] results as tabulated below:

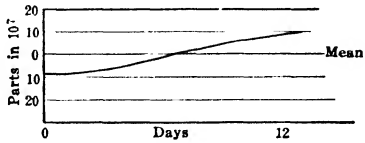


Fig. 13.—Showing the constancy of a frequency standard, an oscillating triode circuit controlled by a small piece of quartz crystal.

	Plate *A*, Cycles	Plate *B*, Cycles
Bureau Standards	200,122	200,142
England	200,118	200,128
France	200,134	200,149
Italy	200,119	200,137
Germany	200,131	200,152
Bureau Standards	200,115	200,138

It will be seen that during the months required for the trip one oscillator changed its frequency only 7 cycles per second in 200,000, and the other one, only 4. It will also be seen that the different national laboratories agreed with the mean frequency to about 10 cycles in 200,000. The agreement is undoubtedly closer today.

It is evident that very special schemes of measurement must be employed to determine frequency with the precision called for in the foregoing tests. In general these high radio frequencies are measured by the use of auxiliary frequencies which are adjusted to be sub-multiples of the frequencies to be measured, and these lower frequencies (or sub-multiples of them) are compared against a quartz-crystal-driven electric clock, which latter has been compared with the U. S. Naval Observatory standard clocks. Those skilled in this work attain remarkable facility and accuracy;

[1] Horton and Marrison, I.R.E., Feb., 1928, p. 137.
[2] Proc. I.R.E., May, 1928, p. 579.

using the scheme outlined above it is possible to measure frequency to one part in a million.[1]

Short-wave Transmission.—But a few years ago it was thought that wave lengths of less than 200 meters (1.5 mc.) were of no use for long-distance communication. When the wave-length band above 200 meters was assigned to broadcasting, the amateurs were assigned to shorter wave lengths, and they soon were establishing long-distance communication with wave lengths of less than 100 meters. This caused the commercial communication companies and government engineers to investigate these very high-frequency signals, and soon a tremendous amount of experimental data had been compiled from which the performance of short-wave channels could be ascertained.

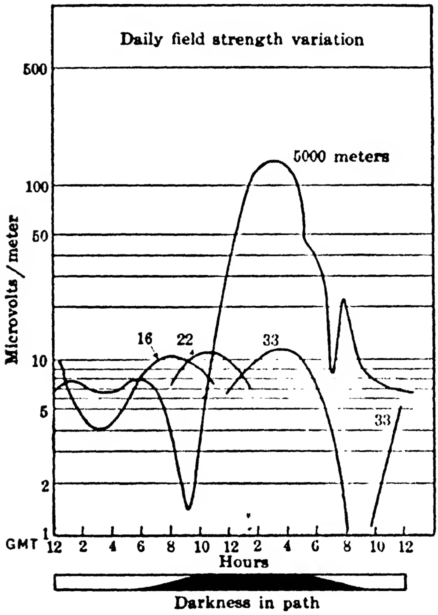


FIG. 14.—Showing how short-wave channels may supplement a long-wave channel.

The first transatlantic radio telephone channel used a wave length of 5000 meters; the transmission at this frequency varied a great deal throughout the 24 hours, so much so that it was regarded as advisable to supplement the channel with others of different frequencies. In Fig. 14 are shown the comparative transmission characteristics of a 5000-meter wave and three short waves 16, 22, and 33 meters. It can be seen that it is quite possible to supplement the long waves with short-wave channels and thus gain in reliability. The size of transmitting equipment required for the short-wave channels is much less than for the longer ones. The results of Fig. 14 are given by Blackwell[2] and represent the average of thousands of readings by engineers of the telephone company.

Espenschied[3] has classified the various available short-wave channels as to their day and night utility as in Fig. 15. Short waves are not good for less than about 1000 miles, so his curves start at this distance. For

[1] "Accurate methods of measuring transmitted wave frequencies at 5 and 20 megacycles per second."—Hall—I.R.E., January, 1931, p. 35.

[2] Blackwell, B.S.T.J., April, 1928.

[3] Espenschied, I.R.E., June, 1928, p. 773.

day transmission the best frequency is given on the graph by something below the " Day " line and above the dashed curve. For night transmission some frequency between the dashed line and the line marked " Night " should be used. For 24-hour transmission two frequencies should be chosen, one for day and one for night. A chart due to Baker and Rice is shown in Fig. 16. It has been assumed that 5 kw. of power are sent out by this transmitter and that the receiver requires a signal strength of at least 10 micróvolts per meter. Their chart then shows the distance over which communication can be established with the frequency which should be used.

Wilson and Espenschied [1] summarize all data available at that time in a curve sheet reproduced in Fig. 17. They take cognizance of the fact that the best wave for summer is not the best wave for winter conditions, especially for night transmission. The two sets of curves, Figs. 16 and 17, show reasonable agreement, although compiled largely from the results of different tests.

Skip Distance.—As the receiver is moved away from a short-wave transmitter, the signal falls off, in intensity, very rapidly, much more so than from a station sending out waves whose length is several hundred meters. It was undoubtedly because of this rapid attenuation that the usefulness of short waves for long-distance communication was questioned. Both Figs. 16 and 17 show a curve marked " Limit of ground wave"—and it will be seen that the possible distance over which communication can be established is small. Thus for 30 meters the possible distance is about 75 miles, yet this same 30-meter wave is good for communication up to 10,000 miles on a winter night! However, Fig. 16 shows that on a winter night this 30-meter wave would give too weak a signal for distances less than 4000 miles! Thus the 30-meter wave would permit communication up to 75 miles from the transmitter, would be too weak to be picked up from 75 miles to 4000 miles, and then give reliable communication from 4000 to 10,000 miles! From 75 to 4000 miles would be the skip distance of this wave for a winter night. For a summer day the skip distance would be from 75 miles to 400 miles.

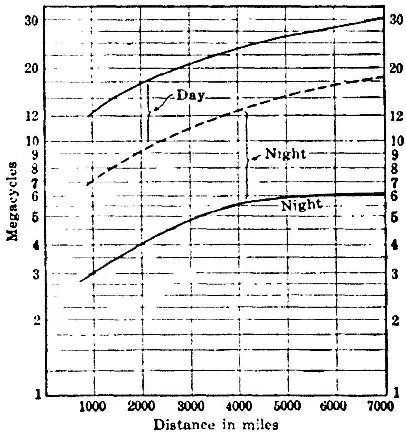


Fig. 15.—Day and night utility of short-wave channels; in general shorter waves are required for day transmission than for night transmission.

[1] B.S.T.J., July, 1930.

Taylor has recently reported on the skip distances of very short waves;[1] his experiments were made with frequencies from 20,000 kc. to 40,000 kc., and he found skip distances, in the day time, of from 800 to 1800 miles. At night time the results are too variable to reach any definite conclusion.

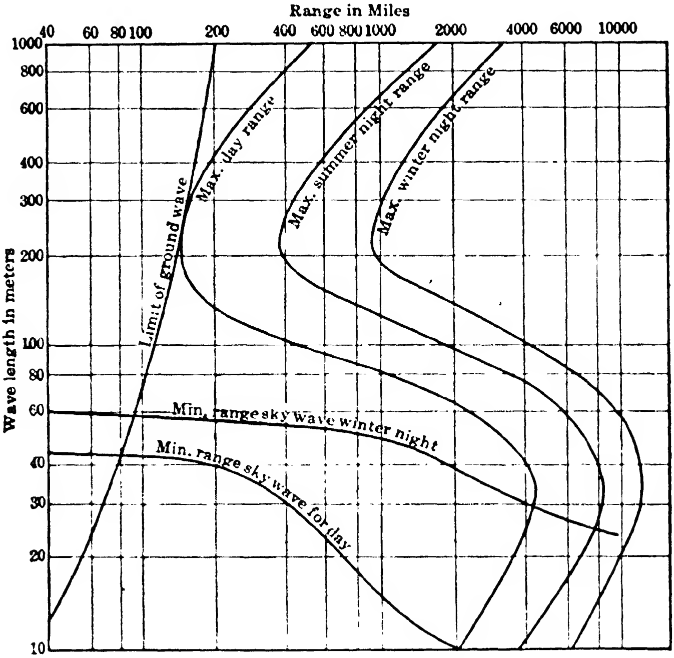


FIG. 16.—Relative propagation of various wave lengths, taking into account the reflection from the upper atmosphere. A 5-kw. transmitter, and required field strength at the receiver of 10-microvolts per meter, are assumed.

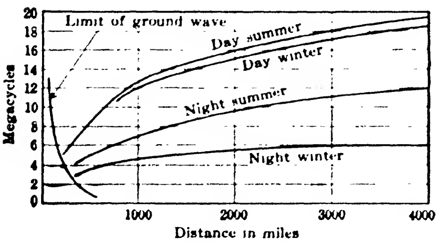


FIG. 17.—For night in summer a shorter wave is advisable than for a night in winter; to some extent the same is true of day transmission.

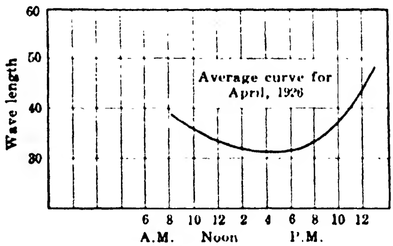


FIG. 18.—On one radio channel the shortest wave that could be used for communication varied throughout the day as shown here; the distance involved here was only 55 miles.

[1] I.R.E., January, 1931, p. 103.

In making some experiments with short-wave transmission, Clapp[1] found that, with a 50-watt transmitter, a receiving station 55 miles distant could get no signal unless a sufficiently long wave length were used. As the wave length was diminished, a definite value was found below which communication was impossible, because of the high attenuation. This "cut-off" wave length he found to vary with time of day, typical values being as shown in Fig. 18. At 6 P.M. he could transmit with a 32-meter wave, but at 12 o'clock at night a 32-meter wave was of no use; the shortest wave he could use was 45 meters.

Kennelly-Heaviside Layer.—With increasing height air becomes more and more rare; at sea level the pressure is 760 mm. mercury; at 5 miles it is about 300 mm.; at 10 miles about 90 mm.; at 20 miles about 10 mm., etc. Also with increasing height the intensity of the sun's radiation increases because a smaller part of it has been absorbed by the atmosphere. But the lower pressure and the increased radiation both tend toward ionization of the atmosphere; the number of free electrons per unit volume will therefore increase with increasing height. Of course, at a distance of thousands of miles there will be no sensible amount of air, so that here the number of free electrons per unit volume will be low in spite of the intense radiation. It will be apparent, therefore, that at some height above the earth's surface there will be a layer of ionized air; the layer will not be very definite in its limits and will probably be at different heights under different conditions, summer and winter, day and night, etc. This semi-conducting layer in the upper reaches of the atmosphere was apparently postulated by Kennelly and Heaviside independently; it is called the Kennelly-Heaviside Layer.

Now a radio wave is reflected by a conductor, so we might expect that this layer will reflect to a greater or less extent the radio waves impinging upon it. Furthermore, a knowledge of the laws of optics leads one to expect that it will act differently for waves of different frequencies. Such proves to be the fact; some wave lengths are reflected by it, while others go right through it.

An attempt at the mathematical analysis of the reflection of electromagnetic waves at media with varying conductivity and dielectric constant has been given,[2] but it can naturally lead to qualitative conclusions only, as we do not know either of these constants for the upper atmosphere; most of our knowledge of this reflecting layer must come from experiment.

Many clever experiments have been carried out with the idea of measuring the height of this layer; the time for a wave, generated on the earth, to travel up and back is measured by oscillograph record or similar scheme.

[1] I.R.E., March, 1929, p. 479.
[2] Elias, I.R.E., May, 1931, p. 891.

The height of this layer evidently varies, to a considerable extent, with weather conditions; it probably acts otherwise for long waves than for short ones. Various experimenters have reported it at heights of from 75 to about 700 km. Heising [1] finds heights from 155 to 400 miles, the average at night being 200 miles. His tests seem to show that the layer moved up and down with velocities of several miles a minute. Using a wave length of 40 meters, the yearly average height was between 120 and 240 km.; with 70 meters it was found to be 90 to 220 km. (day). With long waves heights of only 75 km. (day) and 90 km. (night) have been reported.

Merick and Hentschel [2] reported tests with an aeroplane transmitter and ground receiver. The interference between the ground wave and sky wave produced periodic fading at the receiver, the periodicity of the fading depending upon the speed of the plane and height of ionized layer, etc. They reported a height of 87 to 102 km. for daytime.

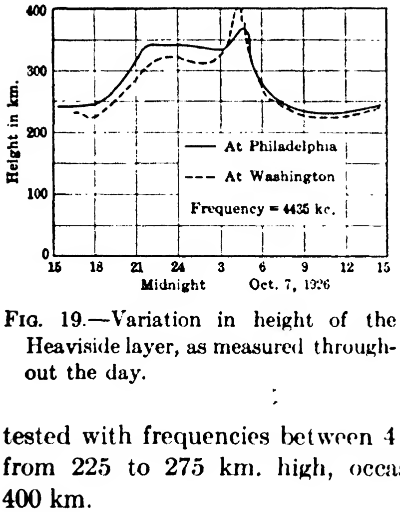


FIG. 19.—Variation in height of the Heaviside layer, as measured throughout the day.

Kenrick and Jen [3] report variations in the height of the layer, as measured with a 4435-kc. signal, to be variable throughout the day as shown in Fig. 19. They find as did Heising, that the layer rises more slowly than it falls. The different heights, as measured at Philadelphia and Washington, would indicate that the layer is cloud-like in its distribution, being much closer to the earth at one place than at another a few miles away.

Gilliland reports [4] that when tested with frequencies between 4 and 8 megacycles the layer is generally from 225 to 275 km. high, occasionally showing a height as great as 400 km.

Oscillograms frequently show several reflections of a single pulse, indicating that there are several layers, one above the other, acting as reflectors.[5]

Attenuation of Propagated Waves.—The electromagnetic waves set up by the transmitter are propagated in all directions through the ether at a velocity corresponding to that of light, as discussed in the earlier portions of this chapter. As the distance from the transmitter increases, their amplitude or intensity decreases, due to the waves spreading out in

[1] I.R.E., Jan., 1928, p. 75.
[2] I.R.E., June, 1929, p. 1034.
[3] I.R.E., April, 1929, p. 711.
[4] I.R.E., January, 1931, page 114.
[5] DeMars, Gilliland, and Kenrick, I.R.E., January, 1931, p. 106.

ever-widening circles and energy being absorbed by the different media through or over which the waves may be propagated. This decrease in intensity, expressed in terms of the initial intensity at the source, is called the attenuation of the wave.

Many investigations have been made to determine the attenuation of these waves, among the more important of which may be mentioned those carried out by L. W. Austin[1] in 1909–1910, using the station at Brant Rock, Mass., as the receiver and the transmitting sets on U. S. cruisers for sending. His results cover one special case only, namely, transmission during daylight over sea-water.

Through the points, obtained from these experiments, a curve was drawn, the equation of which as deduced by Austin[1] and Cohen, is as follows

$$I_r = AI_s \cdot \frac{h_s h_r}{d\lambda} \cdot \epsilon^{-0.0015\frac{d}{\sqrt{\lambda}}}$$

where A is a constant;

I_s is the effective current in the transmitter antenna;
I_r is the effective current in the receiver antenna;
h_s is the height of the transmitting antenna;
h_r is the height of the receiving antenna;
d is the distance between the two stations;
λ is the wave length of transmission.

All lengths are expressed in kilometers.

For the ranges covered by Austin's investigation, namely:

$$I_s = 7.0 \text{ to } 30.0 \text{ amperes}$$
$$h_r \text{ and } h_s = 12 \text{ to } 40 \text{ meters}$$
$$\lambda = 300 \text{ to } 3750 \text{ meters}$$
$$d = \text{up to } 1500 \text{ kilometers}$$

the constant A was found to be equal to 4.25.

The receiving antenna resistance was 25 ohms, effective.

The above formula for wave propagation is known as the Austin-Cohen formula; whereas it seemed to fit reasonably well for the conditions in the experiments from which the empirical constants were obtained, more recent tests, probably more accurate, and carried over a greater range of wave-length yield somewhat different constants.[2]

[1] L. W. Austin: Bull. Bureau of Standards, vol. 7, p. 315, 1911 and vol. 11, p. 69, 1914.

[2] More recent tests by Valauri on the strength of signals received at Leghorn from Annapolis indicate that the attenuation is much less than Austin's formula predicts. Vallauri's measurements gave a field strength at his receiving station about ten times as great as the value calculated from Austin's formula. See also "Long Distance Wireless Transmission," by C. F. Elwell, Inst. of Elec. Engineers, June, 1921.

Summarizing the results obtained to date we may write

$$I_r = AI_s \frac{h_s h_r}{d\lambda} \epsilon^{-\frac{\alpha d}{\lambda^\beta}}$$

and have for the empirical constants α and β,

	α	β
Sommerfeld.	0 0017	0.33
Austin-Cohen	0 0015	0.50
Fuller.	0 0045	1.4
Espenschied	0 005	1.25

The latter figures give the average of 5000 observations, for daylight transmission over the ocean through a 5000-km. distance. With an antenna input (transmitter) of 50 kw. the received signal varied during the day from 1 to 100 microvolts per meter, for a frequency range from 15 to 50 kc.

The signal from the 50-kw. station gave field strengths at different distances as shown in Fig. 20. The solid line shows the predicted field strength using Espenschied's constants, and the plotted points give the average of a year's observations. The signal was of 5000-meter wave length.

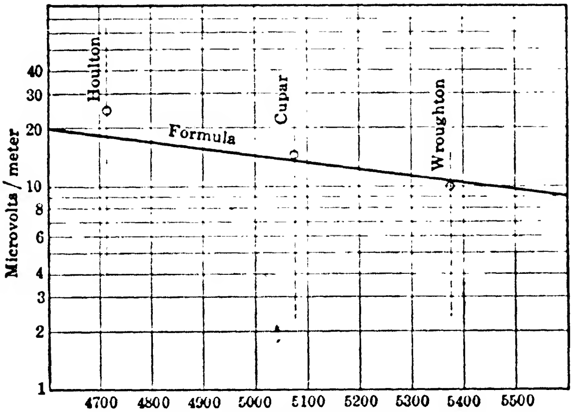


FIG. 20.—Measured signal strengths agree reasonably well with the calculated values, using Expenschied's constants.

Cherry [1] measured the attenuation of station 3LO (Melbourne) over land and over sea, and some of his results are given here. Fig. 21 shows the signal strength over sea, in the daytime, with exception of the last few miles. In Fig. 22 is shown the attenuation over land and sea; instead of plotting field strength there has been plotted, as ordinate, log (field strength times distance). The curvature of the earth affects the signal strength to

[1] Proc. Physical Society, April, 1930.

some extent; in the sea-water graph of Fig. 22 a correction has been applied to take care of this effect.

In Fig. 23 is shown a signal distribution curve over land of different characteristics; the shaded areas represent mountainous wooded regions with a high attenuation. In the accompanying table is shown the great

DROP IN SIGNAL STRENGTH THROUGH TIMBERED COUNTRY

Width of Wooded Strip, Miles	Intensity before and after Traversing Woods	
	$\lambda = 371$ Meters	$\lambda = 484$ Meters
6	12.4 to 1.6	14.5 to 5.7
5	8.6 to 2.0	8.0 to 3.6
6	7.5 to 3.7	6.6 to 4.1
13	10.5 to 2.0	10.5 to 3.6

Signal strength in millivolts per meter.

drop in signal strength as the waves travel through a short piece of woodland. By inspection of Fig. 23 it may be seen that the attenuation caused by a few miles of forest is as great as that taking place over 50 or 75 miles of flat land of good conductivity.

In Figs. 24 and 25 are shown the distances a given amount of power will send a signal of certain strength, both through city territory and through the country, with low attenuation. These results are from a paper by Edwards and Brown.[1]

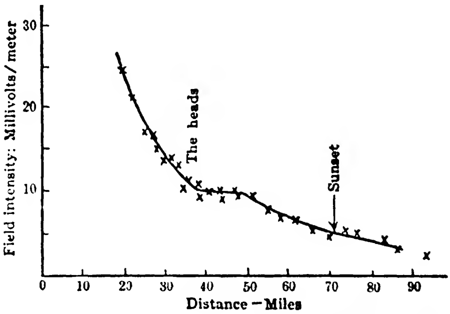


FIG. 21.—Variation in field strength with distance.

Refraction of Waves.—In Fig. 6, p. 361, is shown a conventional picture of an antenna, sending out radiated energy in all directions. Where the energy travels in the direction of C (Fig. 6) through the air, there are parts of the wave which are close to the water (or earth as the case maybe); it follows from Eq. (2) that the parts of the wave in the sea or ground cannot travel as fast as the rest and therefore lag behind, and the surface bounding the advancing wave in air is distorted, the " feet " of the wave being held back, thus reaching a given distance from the transmitting

[1] I.R.E., Sept., 1928, p. 1173.

station later than some other part of the wave which is more distant from the absorbing medium (earth or sea). This makes the wave front "lean over" as the wave advances. The table following lists some values of the wave-front angle as obtained by Austin[1] at the Bureau of Standards Laboratory.

Sending Station	Distance in Nautical Miles	Wave Length in Meters	Apparent Angle of Wave Front
New Brunswick	152	13,000	+3.1°
Annapolis	29	17,200	−2 2°
Nauen	3595	12,500	+3.4°
San Diego	1974	15,200	−0.8°

The angles listed are measured between the wave-front and the vertical; + indicates a bending forward; − indicates a bending backward. No marked change in angle was noted on measurements made at night: the foregoing represent daylight measurements.

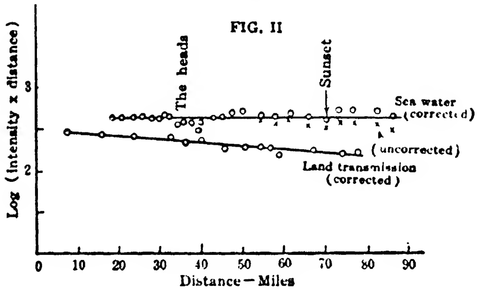


FIG. 22.—When corrected for the earth's curvature transmission over the ocean shows a nearly constant value for the function, log (intensity × distance).

The amount of this vertical tilt depends upon the conductivity of the earth or water over which the wave is traveling, being greater as the conductivity is higher. The effective resistance per centimeter cubed for different soils varies from perhaps 2500 ohms (bare swampy land) to as much as 200,000 ohms (mountainous, heavily wooded districts). For sea-water, it is about 100 ohms.

A 5000-meter wave, after traveling over the Atlantic and then over some hundreds of miles of land, showed (in Maine) an average vertical tilt of about 2.8°; in Cupar, Scotland, about 1°; and Wroughton, England, about 0.6°. As will be explained in the chapter dealing with antenna types, it is the forward tilt of the electric field of the advancing wave which permits the functioning of the long horizontal antennas used for transatlantic reception.

In addition to this change in vertical aspect, a wave front may also swing around in the horizontal direction, thus appearing to come from a

[1] "The Wave-front Angle in Radio Telegraphy," L. W. Austin, Journal Washington Academy of Sciences, March 4, 1921.

direction differing from the true direction of the transmitter. This effect seems to be greatest along shore lines; it is also noticed that at time of sunset a large change in apparent direction may take place if the wave travels approximately parallel to the direction of the sunset zone. Austin [1] reports some observations in long-wave European stations, observed in France. A 23,400-meter wave traveling 510 km. over land showed a weekly deviation as much as 20–25°, going even as far as 90° at times. The same transmission path, for an 18,900-meter wave, showed a variation of only 10–15°. Another long-wave station, distant 900 km., showed a weekly deviation of from 7° to 15° for a 13,000-meter wave and up to

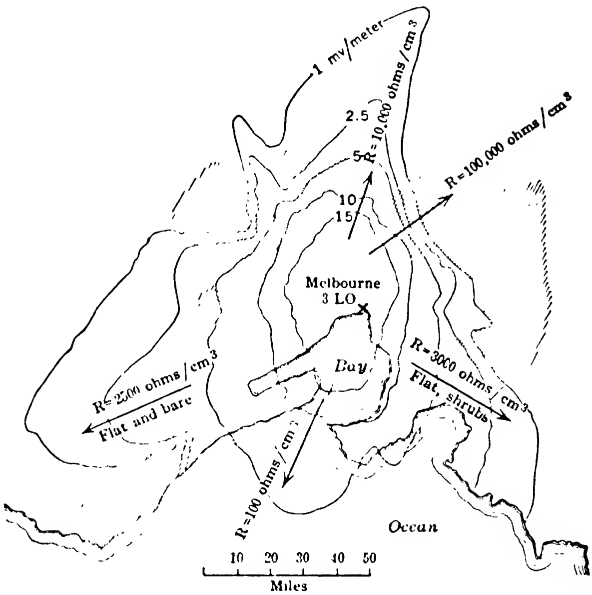


FIG. 23.—Showing how signal strength diminishes with increasing distance from station; the shaded areas are wooded mountains, and evidently cause much greater attenuation of signal than flat, bare country. The station was 5 kw., sending out a 371-meter wave.

30° for an 18,000-meter wave. Tests in the United States of long-wave European stations show night variations as much as 120° from true direction and only 2–3° by day. Up to about 20 miles, it seems there is no distinct deviation from true bearing, unless there are unusual conditions (shore lines, etc.).[2]

Pratt reports [3] tests on an aeroplane beacon transmitter (transmitting

[1] I.R.E., March, 1928, p. 348.
[2] For explanation of this direction change, see pp. 957 et seq.
[3] I.R.E., May, 1928, p. 652.

290 kc.) showing no serious departures from true bearing (as observed in an aeroplane) up to 25 miles. At 50 miles a deviation as great as 25 degrees from true bearing was found, and he reports the beacon as of no value (as a direction finder) at 100 miles. This same beacon tested from a ground receiver showed great variations in direction. One receiver

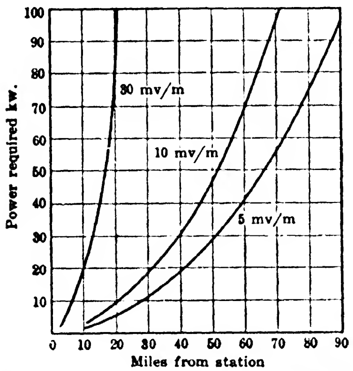


FIG. 24.—For given signal strengths these curves show powers required, with attenuations similar to those found in the average city.

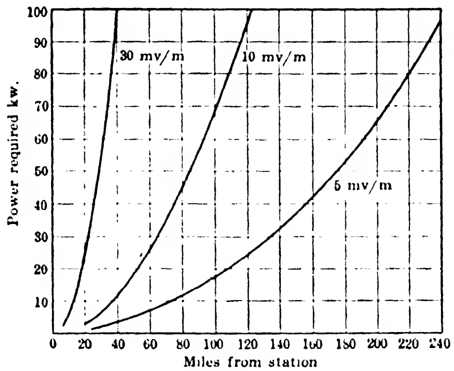


FIG. 25.—If the station is situated in the country, with not much woods or mountains, this set of curves shows the power required to give certain signal strengths at specified distances.

location, only 53 miles from the transmitter, showed a swing in direction of the wave travel of 62°, in only a few minutes' time. This was in a mountainous wooded district. Other points in the same kind of territory showed similar results.[1]

Taminura observed the variation in apparent direction of a long-wave station,[2] located 148 km. distant, in a true direction 13° east of north. The frequency of the station was 19.7 kc. Bearings were obtained by the loop antenna, and he noticed that during most of the time the zero setting was very broad, indicating either that the wave was coming in from several directions or else was rapidly variable in direction. He supposes, as do many others, that the changes in apparent direction are due

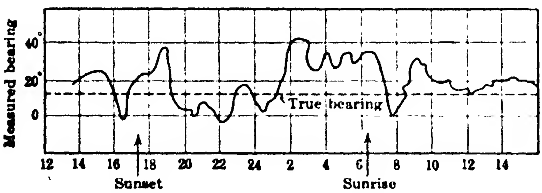


FIG. 26.—Variation in apparent direction of a 19.7 kc. transmitter, 128 km. distant from the direction-finding receiver.

[1] For later data on reliability of radio beacons, see p. 955.

[2] I.R.E., April, 1930, p. 718.

to some kind of interference phenomena between the ground and sky waves. In Fig. 26 is shown the measured direction throughout a typical day; it varied from the true direction by 7–8° west deviation to 25° east deviation. Variations are greater during the night than during the day.

Taminura noticed that except at the time of maximum deviation the direction finder gave broad minima, indicating a variability in the direction. This correlation is shown in Fig. 27, the upper curve of which indicates the deviation from the true direction and the lower curve the "broadness" of the zero setting of the direction finder.

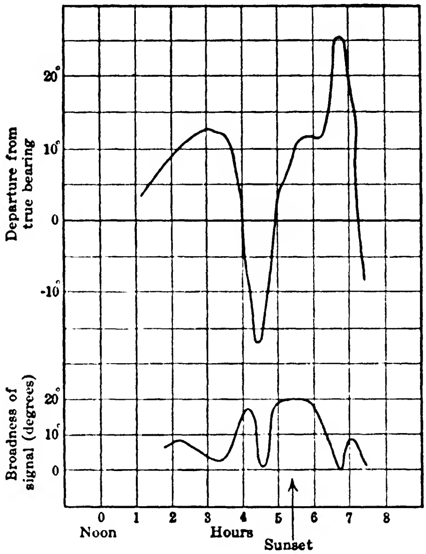


Fig. 27.—Showing correlation between the change in apparent direction of a distant station, and the broadness of setting of the direction-finding loop.

An interesting case of refraction is explained by help of Fig. 28. A transmitting station at A gave a nearly inaudible signal at C, only a few miles distant, while at D and more remote points in the same direction loud signals were observed. A district of high steel buildings at B absorbed practically all the radiation of station A going in the direc-

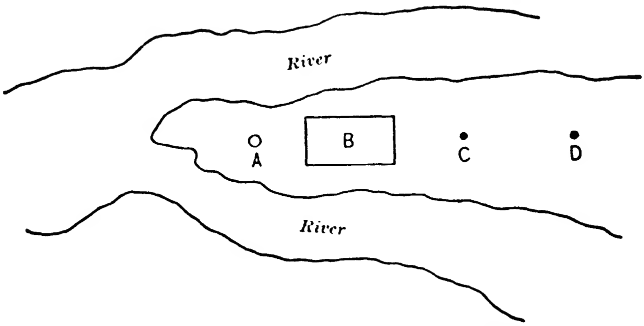


Fig. 28.—Radiation waves emanating from station A are absorbed by region B giving at C scarcely any energy. Point D, in the same direction as point C, has more energy than point C, energy flowing into point D from the directions of the rivers.

tion of C, making an "energy hole" between B and C. Radiation going up the two rivers, however, was not absorbed by district B, and this energy fed in towards point *D* to fill in this energy hole. Tests showed that the energy at *D* was coming from the directions of the two rivers, and not directly from station *A*.

Reflection of Waves.—In the chapter on antennas, it will be shown that suitable metallic reflectors will serve to send back a radio wave which strikes them; the reflector must be of sufficiently large dimensions and of the right electrical characteristics to act as a complete reflector.

The Kennelly-Heaviside layer of ionized atmosphere acts as a very efficient reflector, for waves sufficiently short, and it is undoubtedly the reflection from this layer which accounts for many of the puzzling facts met in connection with long-distance transmission, fading, etc.

As shown by Figs. 16 and 17, short-wave radio communication may be carried on by a wave sent directly from the antenna, called the *ground wave,* or by a wave which has gone up in the sky and been returned to earth (probably by reflection), called the *sky wave.* The echo tests referred to on p. 371 depend upon the radio pulse sent upwards, being returned to the earth by reflection from this ionized gas layer.

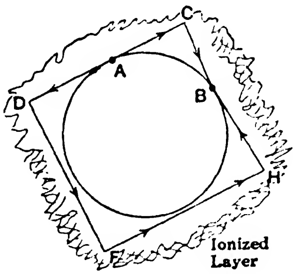
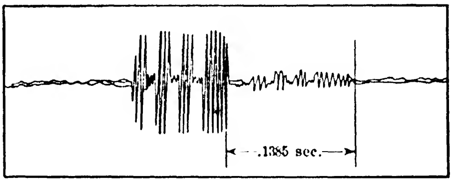
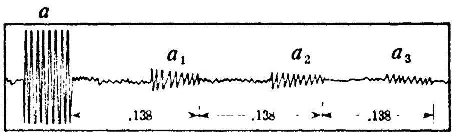


FIG. 29.—The ionized layer of air, by multiple reflections, or refraction, permits signals to travel from one point to another, showing both ways around the earth.



Letter V being sent from Nauen at 150 words/min. Wave length = 18.22 meters



Signal from Rio received at Geltow $\lambda = 15.6$ meters, showing main signal and three successive direct signals. a_3 had traveled 130,000 km.

FIG. 30.—Copies of oscillograms, showing the letter V received direct and after it had traveled once around the earth; the lower diagram is copied from an oscillogram showing distinctly the transit of a short wave signal three times around the earth.

A signal originating at *A* (Fig. 29) may reach *B* as a weak ground wave by traveling over the earth's surface, or it may arrive at *B* by reflection from the ionized layer over path *ACB*. This latter is called the sky wave. Also it is found that the signal may reach *B* by the longer path, going around the earth the longer way, possibly by two or more reflections from the ionized layer as indicated by *ADFHB*. Of course, the wave traveling this route would arrive at *B* somewhat later than the one going the more direct route; many oscillograph films of received signals do show both the direct and indirect signals.

Still more remarkable effects due to the Kennelly-Heaviside layer have been recorded; signals have been received after going around the earth three times! These undoubtedly require many reflections to carry them on such a path. Quäck and Mögel report [1] that their oscillograph records show a direct signal from a nearby station, and another signal which had traversed the complete earth's circumference. The upper diagram of Fig. 30 shows the letter *V* (three dots and one dash) received directly and then repeated just 0.138 second later. The second diagram is a reproduction of one of their oscillograms of a signal received from Rio de Janeiro in Germany; this is shown by the wave train *a* of the diagram. At a_1, a_2, and a_3 are shown with perfect definition three consecutive signals which are 0.138 second apart. Signal a_1 had gone around the earth once, a_2 twice, and a_3 had gone around the earth three times!

Some European observers (Hals, Störmer, and van der Pol) report echoes which occur many seconds after the signal was sent. Conjectures regarding the explanation of these echoes sometimes assume that the waves return to the earth after leaving it completely. Some clouds of electrons or similar effect, in the moon's orbit or beyond, serve as a reflector to send the escaped waves back to the earth.

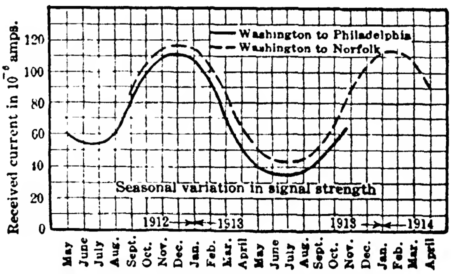


FIG. 31.—Variation of radio transmission occurring with seasonal frequency.

Effect of Season.—As has been discussed under the topic of attenuation, vegetation exerts a marked influence upon the propagation of radio waves over land; the table on p. 375 shows the relatively high absorption of waves traversing wooded regions. Austin has reported measurements of received signal strength throughout the year,[2] and the curves of Fig. 31 are plotted from his results. It is a fact well known to radio listeners that the signals from distant stations come in much stronger during the winter season than in summer.

In Fig. 32 are shown the results of signal strength measurements of a long-wave European station over a series of years; the full curve gives the microvolts per meter received in the United States and the dotted curve

[1] I.R.E., May, 1929, p. 791.
[2] I.R.E., June, 1915.

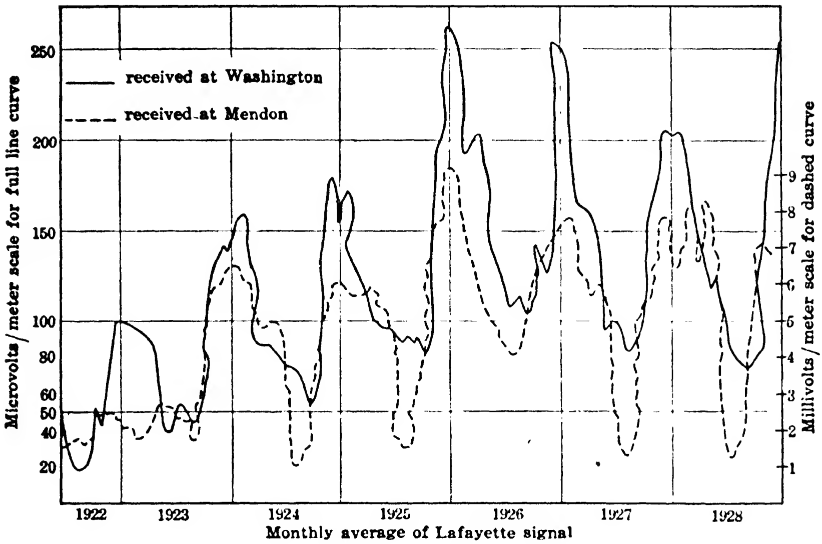


FIG. 32.—Seasonal variation in transmission over a long-wave channel; the full-line curve is for several thousand mile transmission and the dashed curve is for a receiving station near the transmitter.

that received a few miles away. These curves are plotted with the average monthly signal strength, so that only the seasonal variation is shown. It is seen that the same general type of signal change takes place over a short channel as over a long channel, but that in this case the duration of the season of minimum signal is much less for the short channel.

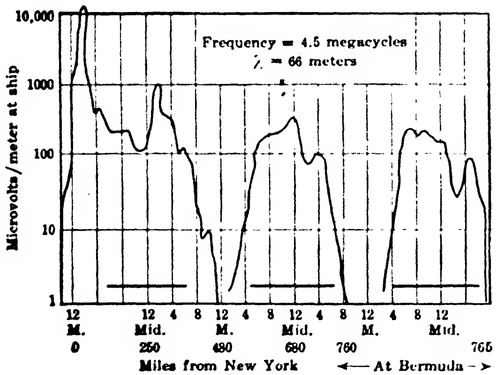


FIG. 33.—Signal strength received by ship on trip from New York to Bermuda; transmitter located close to New York.

Day and Night Variation.—Apparently daylight is not conducive to low attenuation; signals always travel much farther over the dark side of the earth than over the side receiving sunlight. In Fig. 33 are shown the received signal strengths on a boat voyaging from New York to Bermuda; the transmitter was near New York. It is seen that there is a general diminution in signal strength, from 10 millivolts per meter when the boat was near the transmitter to one hundredth as much when the boat was 700 miles distant. During the three days of the test the signals all through the night were 100 microvolts per meter or more, whereas during the intervening days they were so weak as to be unreadable. This was a 66-meter wave channel.

The same effect is observed over long-wave channels. Fig. 34 shows the diurnal signal variation over a 5000-meter transatlantic channel. The test lasted for 54 days during the winter; the two dotted curves are for the "worst" and "best" days, and the full-line curve shows the average signal strength during the entire 54 days.

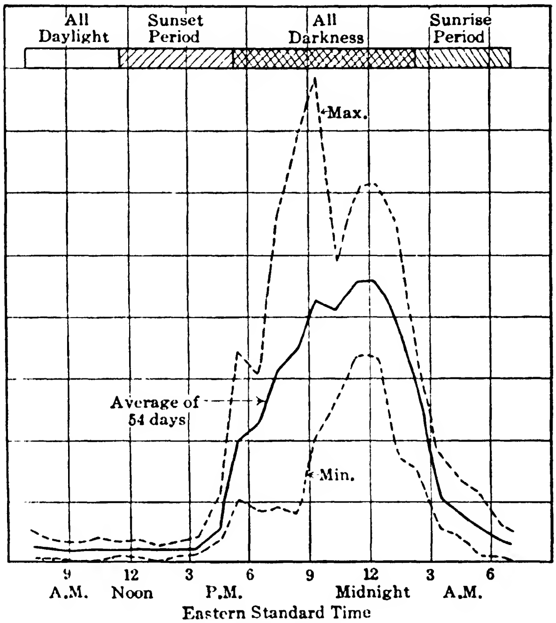


Fig. 34.—Showing variation in field strength for a trans-Atlantic channel. Darkness between transmitter and receiver increases tremendously the transmission efficiency.

Effect of Sun upon Radio Transmission.—As has already been pointed out, sunlight has a most important effect on radio transmission; it would seem likely therefore that variations in the sun's activity, as indicated by the size and number of sunspots, might be reflected in a corresponding variation in the functioning of radio transmission. Such proves to be the fact.

Fig. 35 is taken from a paper by Anderson[1] showing a distinct correlation between the number of sunspots, activity of the earth's magnetic field, and field strength of a 5000-meter signal traversing the Atlantic. It is seen that, whereas there is no evident correlation between the individual magnetic storm and the signal strength variations, there is an evident correspondence between number of sunspots and average field strength. The average field strength was nearly twice as great during a period of marked solar activity as when the sun was quiescent. With short-wave transmission there is a very marked relationship between magnetic storm and attenuation of transmission. In Fig. 36 are shown the variations in field

[1] I.R.E., Sept., 1929, p. 1528.

strength of two transatlantic channels, one 5000 meters and one 16 meters, as well as the horizontal component of the earth's magnetic field, all plotted against days before and after a magnetic storm. It can be seen that the short-wave transmission is seriously affected by the sunspots, whereas the 5000 meters show no observable effect. The horizontal component of the earth's field and the short-wave transmission show a linear relationship.

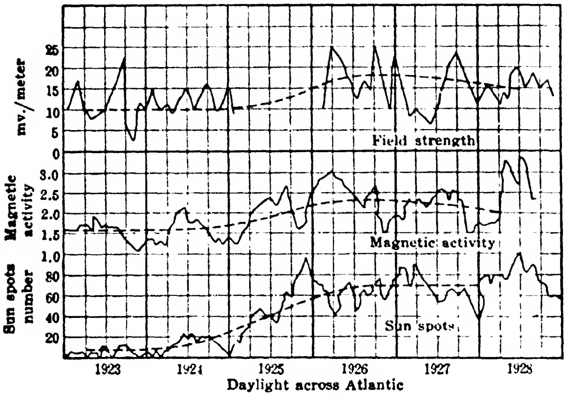


FIG. 35.—A rough correlation is shown between radio transmission and solar activity.

In Fig. 37 are shown variations in transmission of the 16-meter channel for five different magnetic storms. From curves such as these one can well imagine the complex nature of radio wave transmission through the ionized upper atmosphere and varying magnetic fields of the earth.

Anderson summarizes the results of 5 years' observation:

1. The higher daylight signal field strength on 60 kc. during periods of solar activity is associated more with general magnetic activity than with individual storm.

2. Individual storms do tend to increase 60-kc. day signals to some extent. Average was 30 per cent increase on day storm began, to about 75 per cent for 4 or 5 days following. Effects of individual storms vary greatly.

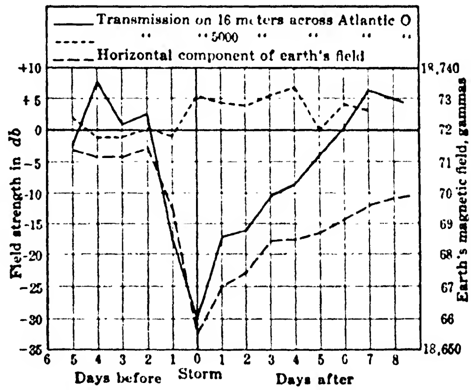


FIG. 36.—A long-wave channel seems to be independent of magnetic storms, but a short-wave channel is strongly affected by whatever causes the magnetic storm.

3. Day-to-day signal fluctuations on 60 kc. are much greater during periods of magnetic activity and are greater in winter than in summer.

4. Magnetic storm accompanied by great decrease in short-wave field strength on day of maximum activity. Even mild magnetic storms may

reduce signals below measurement value. Takes 1 to 8 days to recover, depending on storm severity.

5. Roughly a linear relation holds between millivolts per meter of short-wave field (in decibels above or below average millivolts per meter) and daily average of earth's horizontal component.

Temperature.—Pickard [1] reports a series of measurements to test for correlation between strength of received signal and temperature. He concludes "night reception and temperature are directly related, maximum reception accompanying maximum temperature and vice versa." This is the reverse of the relation previously found by Austin for day reception. Temperature at receiver only was found of importance. Received signal had no correlation with transmitter temperature.

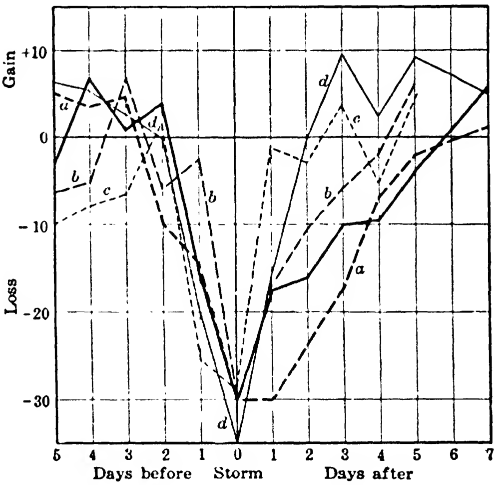


FIG. 37.—Correlation between transmission on a 16-meter channel, and five different magnetic storms.

Fading.—In addition to the signal strength variations so far discussed, varying with season, day and night, with sunspot activity, etc., there is another effect which is due to an apparently different phenomenon. A change in field strength takes place from one minute to the next, and sometimes shows complete cycles of waxing and waning several times a minute, or even second! These rapid changes in field strength are designated by the term "fading," something with which every broadcast listener is familiar.

It seems almost certain today, from the work of many investigators, that this type of signal variation is due to interference between waves arriving at the receiver over different paths and combining to accentuate, or neutralize, one another. Fig. 38 illustrates this idea; a signal arriving at B may come over the direct path or over the path that goes up to the reflecting layer and back. These two waves will assist one another if they arrive in the same phase and of course will tend to neutralize if they have opposite phases. This condition will depend upon the difference in path length; if these differ by a whole number of wave lengths, the waves will

[1] I.R.E., June, 1928, p. 765.

assist one another, and if they differ by an odd number of half wave lengths they will tend to neutralize one another.

Suppose point B is 100 miles from A. Then a signal getting to B over the sky path would have to travel about 500 miles, assuming that the ionized layer is 250 miles high. It is evident then that the sky wave will be very weak compared to the direct ray, so that no appreciable effect on signal strength will occur no matter what the relative phases may be. Thus in this case there would be no appreciable fading due to this type of interference.

However, if the direct ray had to travel through a region of high absorption, such as a city with steel buildings, it might well be that the direct wave has about the same strength as the sky wave. In this case the fading due to interference of the low waves will be a maximum, the signal changing from twice normal strength to zero.

In Fig. 39 such a situation is shown possible. The transmitting sta-

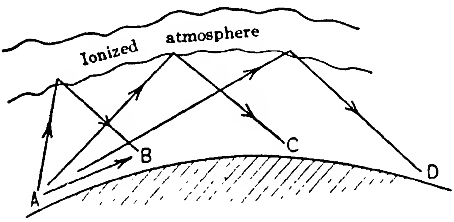


FIG. 38.—Signals arriving at B over the ground path and the sky path may seriously interfere with each other, especially if the attenuation of the ground wave is rather high.

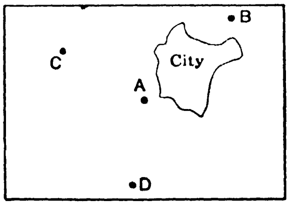


FIG. 39.—Signals received at B from transmitter A, are subject to much greater interference effects than are those received at C and D. The city buildings in the path of the ground wave may reduce its intensity at B to about the same value as that of the sky wave received at B.

tion is shown at A, on the outskirts of a large city. Receiver B receives a comparatively weak direct wave because of the high attenuation caused by a few miles of steel structure buildings. The direct and sky waves at B may therefore show marked interference; moreover, as the difference between paths of the direct and sky waves changes rapidly as the ionizing layer moves up and down (see Fig. 19), it is evident that sounds due to first one frequency and then another may be completely neutralized. Thus a peculiar sort of flutter may travel through the audio-frequency scale, making the different tones increase and decrease in intensity as the wave lengths of the two radio paths change. This is the most disagreeable type of *fading* and cannot be helped except by moving the location of either A or B. It is to be noticed that receivers C and D, situated the same distance from A as is B, will experience no fading at all, because the direct wave arrives with an amplitude much greater than that of the sky wave.

It will be appreciated that as the receiver is moved farther from the transmitter the distances of the direct path and sky path approach each other so that the attenuation over the two paths becomes more nearly equal. In this case, we should expect the interference between the two waves to be more complete, resulting in more complete extinction of the signal, and this corresponds to the well-known fact that stations 500 or more miles distant give signals that wax and wane to a great extent. This general conclusion regarding distance and fading is often contrary to the fact, however, because of special local conditions.

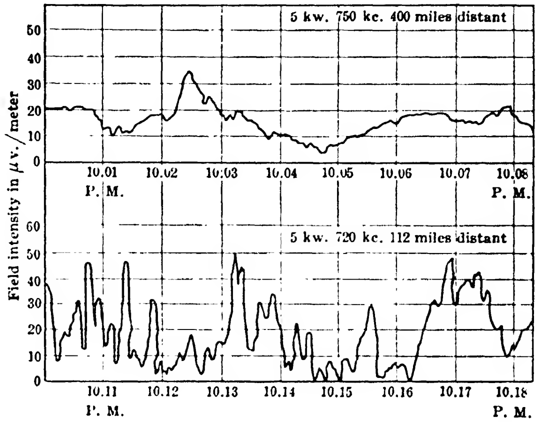


FIG. 40.—Field strength measurements, showing fading of signals from two stations at different distances.

Jansky has reported[1] on field strength conditions in Minnesota; Fig. 40 shows two curves from his paper. The upper one shows conditions 400 miles away from a station and the lower one shows corresponding signals for a station only 112 miles away. The fading is quite evidently much worse for the nearer station.

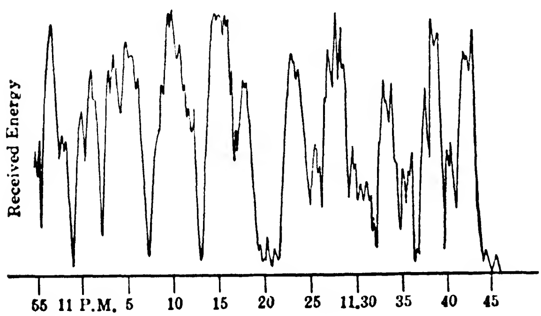


FIG. 41.—A typical record showing fading of the received signal in broadcast reception.

Pickard has reported on many cases of fading, and Fig. 41 shows one of his records. Here the fading presents quite a marked periodicity, the signal dropping to nearly zero intensity about once in 5 minutes. Many cases of fading are reported in which the frequency of the waxing and waning was so rapid that the recording instrument could not follow it.

Friis has carried out some very ingenious experiments to determine the reason for fading in short-wave channels. He investigated the propagation

[1] I.R.E., Oct., 1928, p. 1365.

of both the vertical and horizontal components of the electric field, using the cathode ray oscillograph to show rapid phase changes. His results indicated that the plane of polarization of the horizontal component of the electric field shifted as much as 30°, in a transatlantic 16-meter channel.[1] With the sunset line between receiver and transmitter the angle of polarization was especially variable.

Parkinson [2] also ascribes fading to change in the angle of polarization of the interfering waves. Potter [3] gives an exhaustive report on the propagation of 13 and 18 mc. across the Atlantic, using many oscillograph records to explain the fading of signals on these channels.

Static, Strays, or X's.—In addition to the often experienced interference which may be produced by the operation of other transmitting stations natural electrical disturbances occurring in the atmosphere also cause serious interference in the reception of signals. These disturbances set up electromagnetic waves, which, upon striking the receiving antennas, set up troublesome, interfering sounds in the phones, to which the names static, strays, or X's have variously applied. These strays have been arbitrarily placed by De Groot [4] in the following classes:

(*a*) Loud and sudden clicks. These do not interfere seriously when no other interference effects are present, and have been shown to originate in nearby or distant lightning discharges.

(*b*) A constant hissing noise in the receivers, giving the impression of a softly falling rain, or the noise of water running through tubes. This type occurs occasionally when there are dark, low-lying electrically charged clouds near the antennas, and is apparently caused by intermittent, unidirectional currents in the antennas. Charges of electricity come upon the antennas from the atmosphere through direct physical contact, and thence discharge to ground, producing a current.

(*c*) This type produces a continuous rattling noise in the telephone, something like the tumbling down of a brick wall, and is usually present to a greater or less extent. In the tropics, where interference from static is especially severe, this type predominates, and is always present. It is frequently so severe as to prevent entirely the reception of signals.

Strays are most prominent at night, but are not so troublesome at that time, due to the great increase of signal strength. Their intensity and character is a function of the time of day, the season of the

[1] I.R.E., May, 1928, p. 658.
[2] I.R.E., June, 1929, p. 1042.
[3] I.R.E., April, 1930, p. 581.
[4] "On the Nature and Elimination of Strays," Proc. I.R.E., April, 1917.

year, and the location of the station; thus, in the tropics, De Groot found the most unfavorable time to be that of the trade wind. In general, the worst trouble is experienced when the sun's altitude is highest. Their intensity is probably dependent somewhat on the dryness of the air and wind conditions, increasing dryness and high wind velocities increasing interference due to this cause.

Measurements [1] in the southeastern part of England showed that March was the month of least disturbance, June having the most, with twice the number of disturbances of March. Also, the apparent direction of arrival of the static disturbance (type (*c*)) showed simple diurnal and seasonal variations, following a general law of counter-clockwise swing in direction accompanying increase in solar altitude.

An investigation [2] of the British Meteorological Office appears to establish conclusively that areas in which rain is falling are a definite source of atmospheric disturbance, particularly the advancing edge of such areas. Austin [3] concludes that static disturbances generally come from inland, rather than from the sea. By analogy from signal characteristics, he reasons that static which increases with wave length comes from a great distance, while static which varies little with wave length originates at points less remote. Austin's measurements show that as the wave length for which the receiver is tuned is increased, the static audibility increases; the increase being roughly proportional to the wave length.[4]

Potter [5] reports the results of many measurements of static on short-wave channels (15 to 60 meters) and shows the strength of this disturbance to be much less than on the longer wave channels. He comes to the conclusion that the intensity of the atmospheric disturbances varies inversely with the frequency, for frequencies not too close to the lower end of the radio spectrum.

By means of continuous records [6] of static disturbances, extending over long periods of time, a better knowledge of the directional and intensity

[1] "Directional Observations of Atmospherics," R. A. W. Watt, Royal Society Proceedings, Jan. 1, 1923.

[2] "The Origin of Atmospherics," R. A. W. Watt, Nature, Nov., 1922.

[3] "Determination of the Direction of Atmospheric Disturbances, or Static, in Radio," L. W. Austin, Journal Franklin Institute, May, 1921.

[4] "The Relation between Atmospheric Disturbances and Wave Length in Radio Reception," L. W. Austin, I.R.E., Feb., 1921.

[5] I.R.E. Oct., 1931, p. 1731.

[6] An automatic static recorder consists of a loop antenna connected to a radio-frequency amplifier, the output of the latter supplying a sensitive thermocouple and galvanometer. The deflections of the galvanometer needle are recorded on a moving chart; the arrangement is calibrated by observing the galvanometer deflection when a signal of known strength is impressed on the antenna; this may be done automatically and at definite intervals.

variations experienced will undoubtedly be obtained. Such a survey is in progress and has already indicated that:

(a) Static is a minimum in the early morning hours, reaching a minimum shortly after sunrise. This change occurs regularly and is independent of the weather; it is a maximum with the receiving loop placed in the east and west direction, due probably to the "shadow line" effect discussed under the daily variation of signal strength. (Page 345.)

(b) Most of the static received on the Eastern Coast of the United States originates in the Southwest.

Austin has made many measurements of static on long-wave channels (reported in many volumes of the Proc. I.R.E.) and from them it is evident that the worst atmospheric disturbance occurs in the hot summer months. One set of his results is given in Fig. 42; this shows that the seasonal variation in static repeats itself from year to year. The measurements of Fig. 42 were taken on a 20,000-meter channel (15 kc.).

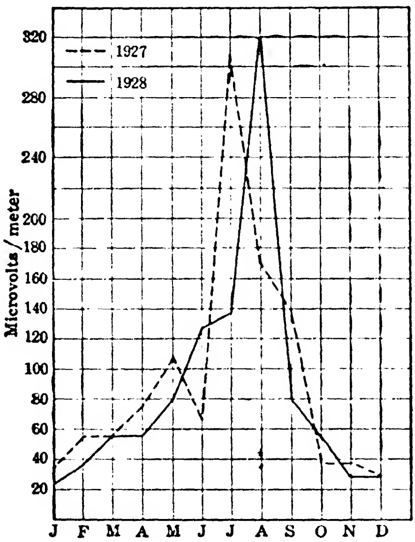


FIG. 42.—Curve showing variation in intensity of static (atmospheric disturbance) throughout the year; measurements made or a 15 kc. channel.

Measurements made at Southgate, England (by engineers of the American Telephone and Telegraph Co.), showed static to be more severe on the lower-frequency channels, as shown in Fig. 9, p. 364. They found also that signals transmitted across the Atlantic on different frequencies came through with different absorption giving the field strength marked "signal" on the curve sheet. This was for a fixed power at the transmitter. The ratio of signal to noise (static) is greatest for a frequency of 45 kc., so if this channel were available it would be the best one to use.

The fact that static has greater intensity on the lower-frequency channels and the fact that it is apparently non-periodic lead one to believe it is merely pulses, having a pulse length greater than one-half the period of 15 kc. because the noise was strongest at this frequency. This would make the pulses somewhat longer than 0.0001 second.

Using a cathode ray oscillograph, and a simple circuit layout which should permit the oscillograph to trace the pulse shapes, Cairns [1] reports on the form of static occurring in Australia. In Fig. 43 are shown a few of the conventionalized pulse forms he observed. Opposite each form is given the percentage of pulses falling in that class, the total pulses observed being over 1000. The right-hand figure, triangular in form, generally showed noticeable serrations on its sides. This condition occurred for practically all the other pulses also; the frequency of these ripples was about 10^5 cycles per second, and their amplitude about 8 millivolts per meter. This ripple was present to some extent on all the pulses whenever the radio receiver gave the continuous "roar" of summer static. Cairns estimated that the most prevalent type of pulse was a 3 half cycle one (*a* of Fig. 43), having a total duration of 0.003 second and a maximum field intensity of 0.14 volt per meter.

Most of the pulses showed that polarity which indicates that electrons were accumulating on the antenna.

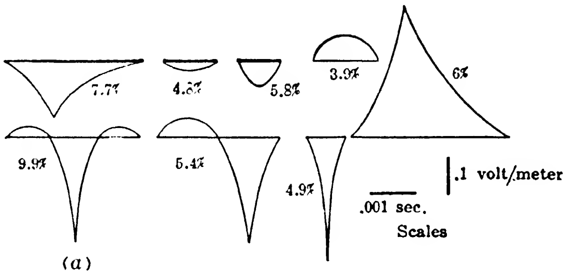


FIG. 43.—Typical shapes of "static," as reproduced from oscillograph records; scales of time and field strength given.

Yokoyama and Nakai have reported [2] that in Japan most of the "grinders," as they call the continuous roar, originate in the tropics; as heard on a 10,000-meter channel, their direction was reasonably constant during the daytime, but during the night their direction was indefinite. They seemed to come from an entirely different direction, generally the local mountainous country where electric storms were prevalent.

Of course, static pulses and roars interfere seriously with the reliability of radio communication. Various attempts have been made to eliminate static, but it is still with us. Non-periodic pulses are difficult signals to weed out. Progress in the problem of obtaining a high ratio of signal to static has been made principally by the use of directive receiving antennas, and selection of a receiver site where static is reasonably low. Of course, in the case of broadcast listeners, neither of these expedients is feasible; here the problem is solved only by increasing signal strength, that is, using more power at the transmitter.

[1] I.R.E., Dec., 1927, p. 985.
[2] I.R.E., Feb., 1929, p. 337.

Directional Radio.—Radio being a type of electromagnetic wave, it should follow the same laws as do ordinary light waves, and it does. Now it is possible to reflect a light wave, or send it in a narrow beam in any desired direction. Suitably formed mirrors will accomplish both these tasks. However, a mirror, to perform its task, must have a sufficiently smooth surface, and must have dimensions (length and breadth) measured in hundreds of wave lengths of the beam it is desired to reflect. This is simple in the case of light because of the shortness of the light wave, but when a wave is, say, 100 meters long, it is not feasible to build a mirror which is plane and, say, 10 wave lengths in dimension. The directing of radio beams therefore, even if possible, will not be as definite a phenomenon as is the reflection of light.

Hertz, the father of radio, using very short waves in his laboratory, actually did reflect waves and also focused them with a big convex lens, but

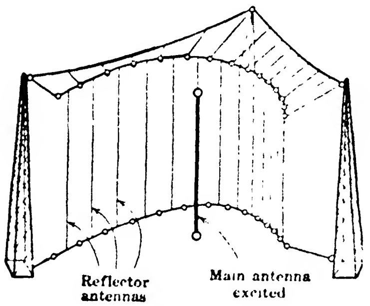


FIG. 44.—A short wave transmitting antenna, with a series of reflectors arranged behind it on a parabolic curve.

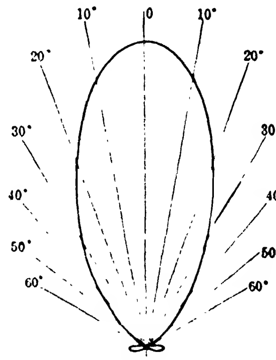


FIG. 45.—The radiation distribution from such an antenna as shown in Fig. 44.

it is not easy to duplicate his laboratory results on a large scale. Marconi did a lot of early work in directive transmitting systems, copying exactly the laboratory apparatus of Hertz, of course on a larger scale. He installed directive transmitters for the British government for communication with fixed receivers in various countries of the commonwealth, Canada, Australia, and South Africa. A conventional sketch of one is given in Fig. 44; steel towers support a steel cable network from which are hung a series of vertical tuned antennas, these being held on a parabolic alignment. The vertical transmitting antenna is held at the focus of this parabolic collection of reflecting antennas, and the emitted radiation is reasonably well focused, as shown by Fig. 45; this represents the electric field strength around the transmitter, as a polar diagram. Most of the radiated energy is included in an angle of about 60°.

More complicated directive transmitting antenna will be shown and discussed in Chapter IX. These directive antennas are only useful for short-wave transmission, because of the great size otherwise required.

The receiving antenna may be made directive by using a loop for receiver. (See p. 885.) This type of antenna receives maximum energy from a transmitter when the plane of the loop passes through the transmitter and practically none when it is placed at right angles to this position.

A long horizontal wire, perhaps 15 ft. above ground, serves on many radio channels as a directive receiving antenna. It may be made several wave lengths long, some of those used on transoceanic telegraphic channels being several miles in length. In more complicated receivers, several of these directive antennas are combined to feed the received signal energy into a common receiving circuit. A more detailed description of this system is given in Chapter IX, p. 960.

Transmitting Apparatus.—From what has gone before, it is evident that the transmitting station must have essentially three things: a generator for producing high-frequency power; a scheme for controlling the amount of this energy in accordance with the signal it is desired to send, dot-and-dash code, voice waves, music, etc.; and a radiating system which will efficiently throw off into space a large share of the generator's output.

The generator for telegraph code messages of a few years ago was a 500-cycle alternator which periodically charged a bank of Leyden jars, with suitable switching arrangement (a so-called quenched spark gap) for letting them discharge into a high-frequency oscillating circuit. This oscillating circuit was magnetically coupled to the radiating system (the antenna), and so there was sent off from the antenna a damped group of waves for every discharge of the Leyden jars. The jars were arranged to charge from the alternator, and discharge into the oscillating circuit, about 1000 times a second. The radiated wave had the form shown in Fig. 10, p. 365. At present, this scheme is used only in the poorer sets of the merchant marine.

For continuous-wave transmission on long waves, an induction type of generator, or an oscillating arc, is frequently used, but both of these methods are being superseded by an oscillating vacuum tube, the functioning of which will be taken up in Chapter VI. For frequencies above about 100 kc. the vacuum tube is the only available source of high-frequency power.

To send code messages, a key must be provided which starts and stops the flow of high-frequency current in the antenna. For small transmitters the key may accomplish this directly, but for powers measured in kilowatts the hand key controls large ones operated electromagnetically, which start and stop the antenna power.

For a voice-modulated signal, an amplifier is required which, after the voice power (of about 10^{-5} watt) is changed into electric current, will raise its energy level about 100 million times and then use this magnified voice power to control the flow of high-frequency power into the antenna.

Signal Distribution around a Station.—The antenna of the transmitting station shakes off energy in all directions, and this travels away from the antenna with nearly the velocity of light. The electric field of the wave

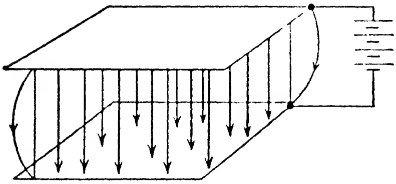


FIG. 46.—Diagram to illustrate the meaning of "volts per meter."

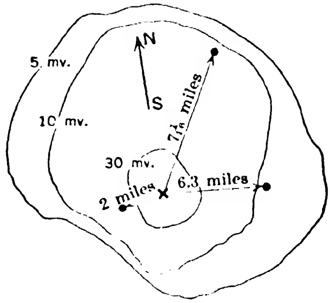


FIG. 47.—Field-distribution around a 1 kw. station with wooded country in the southwest direction.

is nearly vertical (from a vertical antenna) as indicated in Fig. 4, p. 359. The strength of an electric field is measured in *volts* (or fractions thereof) *per meter.* Suppose we have in Fig. 46 two large, flat, metal plates, connected to a battery. The plates will be charged like the two plates of a condenser, and there will be an electric field between them as indicated in the diagram. If the plates are 1 meter apart, and the battery has an e.m.f. of 10 volts, this electric field will have a strength of 10 volts per meter. If the battery is cut down to 1 volt, the field strength will be 1 volt per meter, etc. The electric fields associated with radiated waves are measured in volts per meter very close to the antenna, in millivolts per meter for distances up to perhaps 100 miles, and beyond this distance in microvolts per meter.

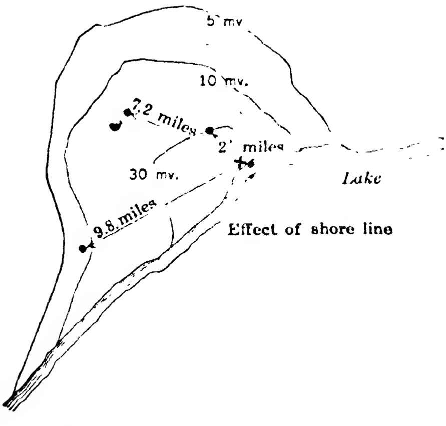


FIG. 48.—Another 1 kw. station located on a lake shore; the low flat country along the water's edge gives comparatively low attenuation.

The signal strength does not decrease, with increasing distance, to the same degree in all directions. Two typical distributions from 1-kw. stations are shown in Figs. 47 and 48. Fig. 47 shows the field distribution around a station having a country of high absorption in the southwest direction. The radiated wave falls off to only 5 mv. per meter at a distance of only 5 miles from the antenna, whereas in the northwest direction the signal travels about 12 miles before it has fallen to 5 mv. per meter. In Fig. 48 is shown the distribution around a 1-kw. station located on a lake shore. In this case, the weakening of the signal in the southwest direction takes place slowly, the country being flat, damp, and of low absorption. Along this shore the signal can travel 18 miles before it has dropped to 5 mv. per meter. In the northwest direction the country is wooded and so absorbs the signal more rapidly. In both Figs. 47 and 48 are marked the average distances to which the signal travels before it falls to 30 mv. per meter, 10 mv. per meter, and 5 mv. per meter respectively.

In Fig. 49 are plotted two curves showing the average distribution of signal intensity around the same transmitter set up in two different locations within about 25 miles of each other. It is quite evident that location B is several times as valuable a site as A (attenuation only being considered); the attenuation with increasing distance from the station is more than twice as great in one location as in the other. Before selecting a site for a transmitter it will quite evidently pay to set up a temporary transmitter and measure the attenuations at the various possible locations; that one giving a high field strength to the greatest number of listeners is the most valuable.

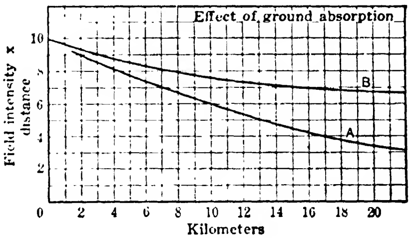


FIG. 49.—Attenuation of field strength from the same transmitter, tested at two different locations within 25 miles of each other; one location gives much greater attenuation than the other.

Edwards and Brown report the distribution around five typical broadcasting stations,[1] giving the data on the transmitter and the average distance at which they found signals of different strengths. Some of their results are tabulated in the following table, all stations having been prorated as 1000-watt stations.

Station	Antenna Height, Meters	Antenna Current, Amperes	Frequency, Kilocycles	Average Distance of Signals, Miles		
				30 Mv. per Meter	10 Mv. per Meter	5 Mv. per Meter
A	21.7	8.0	850	$1\frac{1}{4}$	$6\frac{1}{16}$	$10\frac{3}{8}$
B	39.5	6.0	700	$2\frac{1}{16}$	$6\frac{3}{16}$	$7\frac{1}{16}$
C	23.5	7.8	640	$1\frac{1}{4}$	$6\frac{9}{16}$	$8\frac{1}{2}$
D	32.3	8.0	770	$2\frac{5}{8}$	$7\frac{1}{4}$	$9\frac{3}{16}$
E	18.8	6.8	1110	$2\frac{7}{8}$	$8\frac{1}{4}$	$10\frac{3}{4}$

[1] I.R.E., Sept., 1928, p. 1173.

In Fig. 50 is shown the signal distribution around a 5-kw. station in Minnesota. The 100-microvolt per meter and 50-microvolt per meter lines only are given. Here, it will be seen, there are two directions, south and west, in which the signal travels with much lower attenuation than in the other directions. Fig. 23 gives an interesting signal strength map, and the frontispiece itself shows the signal distribution from a station located in the lower part of New York City.

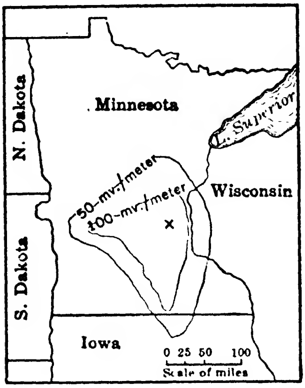


FIG. 50.—Signal distribution around a 5 kw. station in flat country; the two field strength lines are for 100 microvolts per meter and 50 microvolts per meter.

Effect of Shore Line on Wave Propagation.—When using short waves (and possibly longer ones) there is apparently a great attenuation caused by the first mile or so of land between transmitter and receiver, whether the signal is leaving the land to go over water or vice versa.

In Fig. 51 is shown the strength of signal at a distant receiving set located on the ocean front, as the transmitter was moved away from the shore line inland. The transmitter was in New Jersey and the receiver on the south shore of Long Island. It is seen that if the signal has to travel over only one mile of land the field strength at the receiver drops to about one-quarter the value it has when the transmitter is on the ocean front. In other words, the signal should be arranged to start from one ocean front and end at another; no land should intervene between transmitter and receiver.

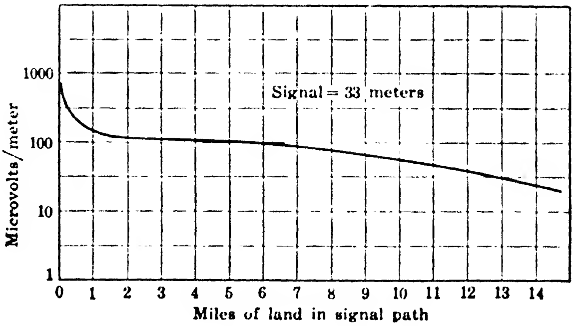


FIG. 51.—Short waves attenuate greatly (at ground level) in crossing a shore line; if the receiving station is located one mile inland the signal strength is only one quarter as much as it is at the shore line.

The phenomenon is evidently one of refraction, the shore line acting towards the radio wave as the edge of an opaque body acts toward a ray of light. If the receiver is moved up (in elevation) as it is moved inland, the rapid drop in signal strength does not occur; the same is true of the transmitter if this is elevated according to the distance inland.

Heising[1] has investigated this effect, and as a result of his analysis recommends that as the receiver, or transmitter, is moved inland its elevation be increased in accordance with the curves of Fig. 52. A 30-meter station, if 1000 ft. inland, should have an elevation of 270 ft.

Ship to Shore Radio Telephone.—It is now possible to talk with any of the larger steamships in the America-England service; a short-wave transmitter at Ocean Gate, N. J., can be connected with any subscriber in the United States. About 50 kw. of power can be put in any one of four channels (17,120, 12,840, 8560, or 4752.5 kc.), and the ship sends back to the receiving station at Forked River in the most suitable one of four other frequencies: 17,640, 13,210, 8830, or 4177.5 kc.

Two of the transmitting antennas are of the diamond form (p. 959) called Bruce antennas, and the other two are of the sawtooth type (p. 958). These antennas are not as directive as the arrays used for transatlantic telephone channels (p. 846) because the ships with which it is desired to communicate may be anywhere in quite a wide zone. An antenna with highly directive radiation might miss the ship altogether.

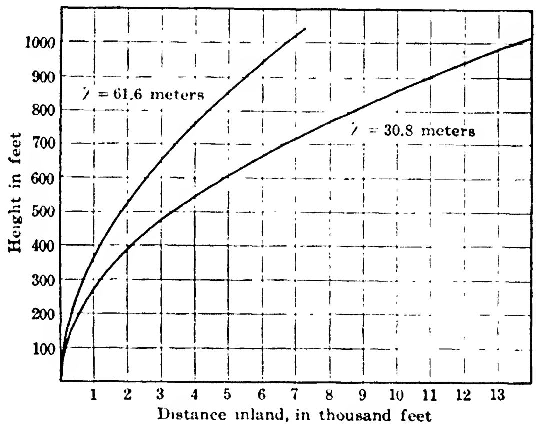


FIG. 52.—As the receiving station is moved inland, its height should be increased as shown here.

Communication with Aeroplanes.—As explained under the section in radio beacons, p. 966, our aeroplanes are continually guided in their course by signals from suitable ground stations, on reasonably long-wave channels, about 1000 meters. Attempts have been made, especially in Germany, to keep in telephone communication with the moving planes by the use of very short waves. Esau and Hahnemann,[2] using waves 3 to 4 meters long, came to the conclusion that if the plane could be seen from the transmitting station, radio communication on this short-wave channel was reasonably sure. They experienced no fading or atmospheric disturbances on this short-wave channel.

Neither the transmitter nor receiver should be close to the ground. As an example, with the receiver on the ground, in a certain instance they got no signal whatever, but by elevating the receiver 10 or 15 meters communi-

[1] I.R.E., Jan., 1932, p. 77.
[2] I.R.E., March, 1930, p. 471.

cation was perfectly satisfactory. A long range depends upon both receiver and transmitter being at high elevation. These short-wave beams evidently do not bend around the earth to any appreciable extent; when earth intervened between low receiver and transmitter no signal at all was received no matter how much power was used; but if one or both stations were sufficiently raised satisfactory signals were received with only a fraction of the power used before.

Using a transmitter of 70 watts output, on a high mountain peak near Jena, radio phone communication was held with a plane flying the course shown in Fig. 53. Communication was not established until the plane had passed the point in its route closest to Jena, but was then maintained over the shaded part of the route. When the plane was flying south, at 1000 meters elevation, the signal was held for a longer distance than when it was flying north at 350 meters.

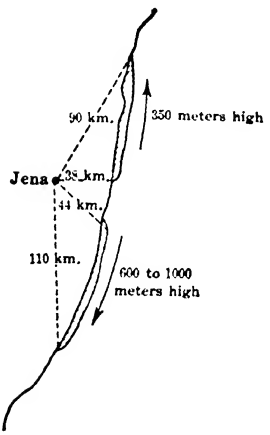


FIG. 53.—Using a 70-watt transmitter, with waves a few meters long, communication with an aeroplane was had at the distances shown here.

The fact that communication was not established until the plane had passed the transmitter was probably due to the fact that the metal frame of the plane effectively shielded the radio receiver from waves coming from the front.

Transmitters on the aeroplanes are necessarily limited in power output because of weight considerations. They vary from about 20 watts to 100 watts or over, and nearly always supply their output to a trailing wire antenna (see p. 898). A small transmitter is described by Bock,[1] and somewhat larger ones by Gross,[2] together with European practice in aviation radio.

Receiving Apparatus.—For fixed channel communication (point to point), with skilled engineers in charge of the receiving apparatus, very complicated receiving arrangements may be used. Complicated antenna systems, all very carefully adjusted and maintained, are used to obtain the highest possible ratio of desired signal to noise, and to discard all signals from any but the one station with which communication is desired.

For ordinary broadcast reception, the antenna is a short piece of wire perhaps 50 ft. long and 25 ft. high, feeding into a compact receiving set with generally only two control dials, one to select the desired station and the other to regulate the strength of the loud speaker response. Such a receiver generally has a rectifying tube and circuit arrangement for getting continuous current power for the receiver circuit; this is called the "power

[1] I.R.E. Sept., 1931, p. 1569. [2] I.R.E. March, 1931, p. 341.

pack." Then there are from two to five low-power vacuum tubes with associated tunable circuits feeding one into the other, called the radio-frequency amplifier. It is this part of the receiver that enables the user to separate one station from another.

The radio-frequency amplifier feeds into the detector tube (sometimes called the demodulator), and this feeds into the audio-frequency amplifier. The detector tube changes the modulated high-frequency current into an audio-frequency current, the shape of the audio-frequency current being the same as the form of modulation of the radio-frequency current supplied to the detector.

The audio-frequency amplifier consists generally of two vacuum tubes and associated transformers; the second tube, called the output tube, is of considerably greater capacity than the other tubes of the set. It supplies to the loud speaker generally about 50 milliwatts of power, but at times (for low organ notes, for example) its output may reach several watts.

The volume control is ordinarily a scheme for varying the bias voltage on the grid of one or more of the tubes used in the radio-frequency amplifier. Occasionally a "local-distance" switch is fitted into a set; by it a resistance may be introduced into one of the tuned circuits of the radio-frequency amplifier. This is done when listening to a local station; by its use the volume is greatly cut down and the fidelity of reproduction somewhat improved.

A receiving set should be selective, as well as sensitive, and should reproduce faithfully the signal it receives from the antenna. These various features of a receiver, and the method of obtaining them, will be taken up in Chapter VIII, dealing with radio telephony, and Chapter X, dealing with amplifiers.

Amounts of Power Used.—The size of the large long-wave telegraph station used for transoceanic communication can well be appreciated from the values given in the accompanying table. The antenna heights are

		Frequency, Kilocycles	Wave Length, Kilometers	Antenna Current, I_a	Effective Height, h
LY	Bordeaux	15.9	18.9	535 amp.	180 meters
FK	St. Assise, Paris	15.0	20.0	475	180
FT	St. Assise, Paris	20.8	14.4	344	180
AGW	Nauen, Berlin	16.5	18.1	414	170
AGS	Nauen, Berlin	23.4	12.8	389	130
GBR	Rugby	16.1	18.6	685	185
GBL	Leafield	24.4	12.3	210	75
MAK	Cayey	33.8	8.87	104	120
KET	Bolinas, San Francisco	22.9	13.1	600	51
NPL	San Diego	30.0	10.0	89	120

from 50 to 180 meters (effective). The actual antenna height may be 30 per cent greater than this. The currents used in the antenna are from 100 to 700 amperes—representing antenna power measured in the hundreds of kilowatts.

A radio telephone channel for spanning the Atlantic uses about 50 kw. of power with a frequency of 60 kc. and considerably less on the short-wave channels; here of course directive transmitting and receiving antennas are used.

The larger broadcasting stations use an antenna power of 50 kw., and the ordinary ones about 5 kw. of power.

The power picked up by the receiver antenna is extremely small. A small antenna might have an effective height of 5 meters and an effective resistance of perhaps 30 ohms. This antenna with a sensitive receiver would give a good output with a received signal strength of 5 microvolts per meter.

Now a 5-meter antenna on a 5-microvolt-per-meter signal picks up 25 microvolts, and this, acting on an antenna of 30 ohms, gives a current of 0.8 microampere, which represents a power of $(0.8 \times 10^{-6})^2 \times 30 = 19.2 \times 10^{-12}$ watt, that is, about 20 micro-microwatts.

If the signal field strength is 10 millivolts per meter (30 to 40 miles from a 5-kw. station) the energy picked up will be about 3 microwatts.

If we take the average of these two figures, we reach the conclusion that the average broadcast receiver abstracts from the advancing wave about one-hundredth of 1 microwatt. That is, if 10 million receivers (all there are in the United States, probably) are being used simultaneously, they receive a total of one-tenth of 1 watt of power. The sum total of power of the transmitters supplying the signals for these receivers would be in the neighborhood of 1000 kw.

These figures show the gross inefficiency of this scheme of transmitting power by radio. Only the minutest fraction of the power sent out is ever picked up.

Very Short Waves.—The curves of both Figs. 15 and 17 indicate that waves of less than 10-meter length are of questionable utility in commercial long-distance communication; such waves are, however, being used today by amateurs, and it will probably be found that under certain conditions they are very valuable. As the wave length decreases, the possibility of directive transmission and reception increases, static disturbance diminishes, and the number of channels available increases.

During the last few years, S. Uda has reported a series of very ingenious experiments on short-wave transmission.[1] He worked with waves varying in length from 4.4 meters to about 0.5 meter, showing how they can be

[1] His reports are abstracted in the English supplementary issues of the Journal of the Institute of Electrical Engineers of Japan.

directed, distances covered, etc. With a few watts of 4.4-meter wave he was able to establish communication up to 60 km., provided the transmitter and receiver were on hilltops. He found that the earth, or even the ocean,

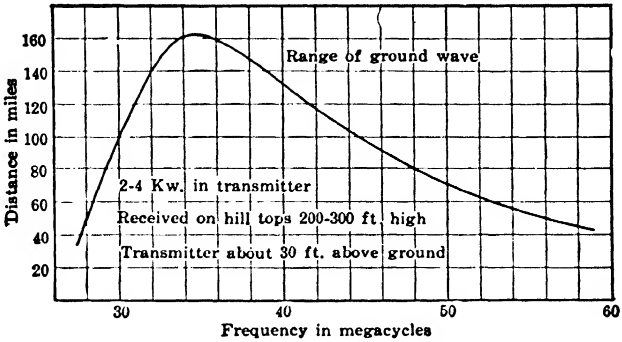


FIG. 54.—Range of ground wave as frequency is varied.

produced very high attenuation on these short waves, so that the wave close to the earth was useful over very short distances.

Okabi reports [1] that he has been able to establish telegraph communication up to 30 km. and telephone up to 10 km. with very small power on a 45-cm. wave.

A fine paper on the use of short waves (5 to 10 meters) is given by Beverage, Petersen and Hansell[2] in which they summarize the experimental results available to date. In the curve of Fig. 54 they give their conclusion from much experimental data obtained by themselves and others. Using elevated positions for transmitter and receiver, with 2 to 4 kw. in the transmitting antenna, they found a useful range somewhat more than 100 miles, the best frequency being about 35 mc.

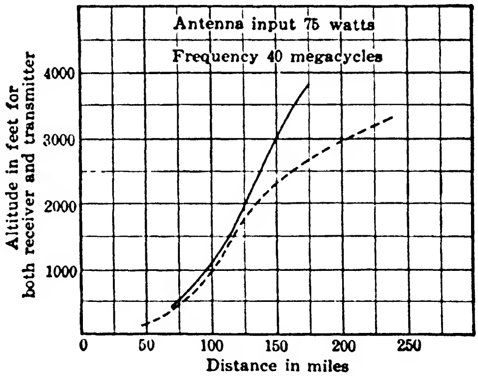


FIG. 55.—Possible distance of reliable communication with 40-megacycle radiation, both transmitter and receiver must be raised as the distance of transmission is increased.

These short waves apparently do not bend around the earth's surface, neither do they reflect from the Heaviside layer; increasing the power at the transmitter does not appreciably increase the range after this reaches

[1] I.R.E., June, 1930, p. 1028.
[2] I.R.E., Aug., 1931, p. 1913.

about 100 miles. The only way the range can be increased is by elevating the position of both transmitter and receiver. In Fig. 55 are shown the conclusions they reached as a result of tests between the high mountain peaks of the Hawaiian Islands. A good signal can be received up to the distance shown by the full-line curve, and it seems that by the use of directive antennas the commercially useful range can be increased to that shown by the dashed line.

Karplus [1] has analyzed the various types of receivers available for short wave detection. Of course the ordinary radio receiver is of no use in detecting these very short waves.

Transmission with Ultra Radio Frequencies.—It is also of interest to note that signals may be, and have been, transmitted by the invisible radiations from a hot body.[2] The wave lengths of these radiations are beyond the visible range of light waves, i.e., greater and are known as infra-red radiation; their frequency is 6×10^8 to 4×10^{11} kc. The transmitter consists of an arc projector or similar device, while the receiver consists of a sensitive thermopile located at the focus of a parabolic mirror. The thermopile is connected to an amplifier and interrupting device, the latter being required to obtain audibility. It is stated that communication has been established over distances of more than 20 km. by this system.

It is interesting to note that workers in the fields of light and those in the field of radio waves are gradually bringing together the range of frequencies in their respective fields. Nichols and Tear have extended the field of damped electric waves to those only 0.22 mm. long; this is a frequency of about 10^6 mc. The same experiments have succeeded in isolating and measuring heat waves as long as 0.42 mm., thus bridging the gap between the radio and light spectrums.

[1] Communication with quasi optical waves—I.R.E. Oct., 1931, p., 1715.

[2] "Radio Telegraphy by Infra-red Radiation," J. Jerbert-Stevens and A. Larigaldi, Comptes Rendus, 169, July 21, 1919. See also Karplus, "Communication with Quasi-optical Waves," I.R.E., Oct., 1931, p. 1715.

CHAPTER V

SPARK TELEGRAPHY

Spark Transmission and Equipment.—The transmission of intelligence by means of electromagnetic and electrostatic energy radiation from an open oscillator, produced by the high-frequency oscillatory discharge of a condenser in an associated circuit, is called spark telegraphy.

In the diagram of connections (Fig. 1) and description of the transmitter, a certain conventional and more or less standard arrangement of equipment has been assumed. Commercial transmitters may differ from this arrangement in several details, for instance, in the method of energy supply, form of spark gap used, and the kind of coupling employed between the closed and open (radiating) oscillatory circuits.

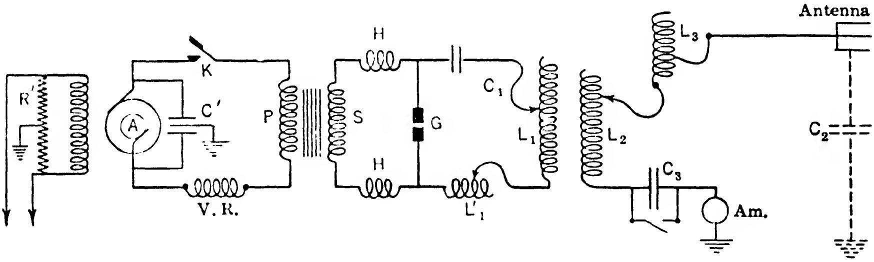


FIG. 1.—Circuit diagram of the ordinary spark transmitter.

An examination of the set shows it to consist of three main circuits: (1) a low-voltage, low-frequency circuit which includes the alternator (A), the key (K), the low-tension winding of the step-up power transformer (P), and a variable reactance ($V.R.$), which is also called a reactance regulator or choke coil; (2) a high-voltage, high- and low-frequency circuit, including the high-tension winding of the step-up power transformer (S), the capacity (C_1), the inductances (L_1) and (L'_1) and the radio-frequency choke coils (H), the spark gap G being shunted across the circuit as shown in the diagram; that part of this circuit comprising L_1, L'_1, C_1, and the gap, in series is called the *closed oscillating circuit*; (3) a third circuit, known as the *open* (*radiating*) *oscillatory circuit*, of high frequency only, containing the following equipment: Inductance L_2, coupled inductively with L_1, and forming with L_1 the *oscillation transformer*; a tuning inductance L_3, the antenna or aerial (represented in the diagram by a fictitious lumped capacity C_2), the hot-wire ammeter A_1, and the short-wave condenser C_3 equipped with short-circuiting switch.

Protective Equipment.—In addition to the above apparatus, the set is equipped with certain protective devices. High resistances, or condensers, are connected across the field and armature terminals of the alternator and its driving motor, to prevent the flow of high-frequency currents in these highly inductive circuits. These high-frequency currents may be caused to flow by direct inductive effects from the closed and open oscillating circuits, particularly if the space is restricted, as on shipboard, and the several circuits are close together. High-frequency current flowing or tending to flow through a high inductance means a high-potential drop across the winding, this potential usually being concentrated ("piled up") at the end turns of the winding. This potential may be sufficient to puncture the winding insulation, and to prevent this the coil is shunted by resistance or capacity, through which the high-frequency current can easily flow, without any abnormal potentials being produced. Also this resistance or capacity usually has its neutral point connected to ground, to prevent excessive potential stresses with respect to ground. The connection of a protective resistance (R') and condenser (C') is indicated in the diagram (Fig. 1).

Since the breakdown of the gap virtually short-circuits the high tension side of the step-up transformer, some means must be introduced to prevent the abnormal current flow which would otherwise occur under this condition. If this is not done, the transformer and alternator may be damaged, and the arc across the gap become a sustained condition, preventing the recharging of the capacity C_1, with resultant decrease of the high-frequency energy. Three means may be used, all three involving the insertion of reactance in the low-voltage supply circuit: (*a*) High reactance (high impedance) in the alternator; (*b*) an iron-cored inductance in the supply leads to the transformer; (*c*) high leakage reactance in the step-up transformer.

The action of this added reactance is to rapidly decrease the secondary terminal voltage as the current flow increases, when the gap breaks down. In addition to the above, a more or less resonant adjustment of the circuit constants may be used to secure an equivalent result. The action of this arrangement is discussed in detail on pp. 409 et seq.

To prevent any appreciable high-frequency current from flowing through the high-tension winding of the transformer, thus setting up high potentials with liability of puncture to the winding, high-frequency reactance coils H are inserted between the gap and the transformer. These coils have a very low impedance to the flow of current of alternator frequency, but present a very high impedance to the high-frequency discharge current, which is thus forced to follow the gap circuit. These coils may be simply a helix of copper wire wound on a porcelain or Bakelite spool, and are designed to possess a high "turn insulation." They can thus safely withstand potential strains which would cause puncture to the transformer winding if permitted to occur at this point. Another advantage is that the high-frequency current is forced to flow in the low-resistance gap circuit, instead of through the higher-resistance by-path presented by the transformer winding. The damping is thus decreased, and the operating efficiency of the set increased. (The damping is decreased by the decrease of closed circuit I^2R losses.) With modern transmitters, where the end turns of the high-tension winding have been specially insulated to withstand the high potentials, these high-frequency choke coils are usually omitted.

Condensers.—The Audio-frequency Circuit of the Transmitting Set.—The condensers used in a transmitting set are known as "power condensers," to distinguish them from those used in a receiving set, which are known as "receiving condensers." A power condenser must, as its very name implies, be capable of handling large amounts of power without serious deterioration or breaking down.

The requisites of a power condenser are:

First.—That the insulation between plates shall be such as to prevent its being punctured by the high voltage used.

Second.—That the losses shall be small. (See p. 252, Chapter II.) The dielectrics generally used in power condensers are: air, glass, oil and mica. Of these air has the minimum specific inductive capacity, and it causes practically no losses, while the other dielectrics have a much higher specific inductive capacity but suffer more or less energy loss. As regards breakdown voltage air is at a disadvantage as compared with the other dielectrics, but at pressures higher than atmospheric the breakdown voltage for air is very high and it increases in nearly direct proportion to the absolute pressure. A comparison of the characteristics of these dielectrics is given in Chapter II, pp. 253 and 261–262.

It will be seen from the characteristics of the various dielectrics that if a condenser of a certain capacity is to be designed, the air condenser would have the largest dimensions and the mica condenser the smallest. However, the losses in the air condenser would be very small, while those in a poorly constructed mica condenser might be so high as to make its use prohibitive. Glass condensers in the form of Leyden jars have met with much favor in the radio field and they have been extensively used. Each jar has a capacity of about 0.002 μf, and is capable of withstanding a voltage of about 15,000; for any particular desired voltage and capacity the jars are grouped in series multiple, so that the combination will have the required capacity and breakdown voltage. Condensers with glass as the dielectric are also made with flat pieces of glass covered with tin foil, the space requirements of such condensers being much smaller than for the Leyden jars. They do not stand continued use, however, as well as the Leyden jars, because of the greater amount of heating due to smaller cooling surface.

Oil condensers are not very much used in general practice, but their use is very commendable in places where there is no possibility of spilling the oil. It must be borne in mind, that although the dielectric properties of oil are unfavorably affected by a flash through it so that oil condensers cannot be expected to give as good service after the oil has once been flashed, they are still serviceable after a breakdown, whereas a solid dielectric condenser, such as mica or glass, is completely spoiled.

The mica condenser is a very desirable one, and has been used to largely supplant the Leyden jar for ship sets and similar installations. It is compact, and, if properly constructed, has a loss so small as to be hardly measurable. The impregnation of the condenser with suitable wax must be done sufficiently well to drive out all air completely, as the trapped air bubbles, suffering corona loss, are the source of local heating and thus weaken the dielectric strength of the mica. It must be noted that these condensers are made to be used at the rated voltage and frequency for *intermittent service only* and that even a good mica condenser if used continuously at the rated voltage and frequency (for spark operation) will have its wax melted after an hour or so.

Compressed-air condensers are very suitable where very high voltages and low losses are required; the structure of the condenser, i.e., the metal plates and their insulating supports, is placed in a steel container capable of safely withstanding a pressure up to a dozen atmospheres or more and dry compressed air is pumped in until the required pressure is obtained. It may be easily seen that an air compressor and gauge are necessary auxiliaries of such condensers, and that their use cannot be considered, except for very large land installations, or for laboratories.

On the whole the Leyden jar with its simplicity of construction and large heat-radiating surface affording cool operation is a favorite type of transmitting condenser and would be even more widely used were it not for its large space requirements and liability of breakage.

Transmitting condensers are very seldom constructed so that their capacity may be continuously varied in view of the insulation difficulties resulting from the high voltages dealt with.

The value of the capacity of the condenser used in the closed circuit of a spark transmitter is fixed by the voltage, the spark frequency, and the power of the set. The energy stored in a condenser is $CV^2/2$, and if this energy is discharged N times per second we have $\text{Power} = NCV^2/2$.

We immediately note that the power varies directly with N, C and V^2. The value of N is more or less fixed, because it represents the " tone " of the set and the best tone is supposed to be that due to $N=1000$ per second. Therefore, if a certain amount of power must be imparted to the condenser a suitable choice must be made of C and V. With a very high voltage the dielectric and leakage losses are likely to be high, and the difficulties of insulating the various parts of the set are such as to make it impractical, and a limit in this direction is soon reached after which, if more power is required, the condenser capacity must be increased. Voltages of 100,000 might be used in large land installations, but in small land and in ship installations the range is 10,000 to 20,000 volts.

It may be easily seen that in large power installations the condenser must have a very large capacity, even though a high voltage is used. For instance, assume:

$W=50{,}000$ watts;
$V=100{,}000$;
$N=1000$ per second.

Then
$$C=\frac{2\times 50{,}000}{1000\times 100{,}000^2}=0.01\ \mu f.$$

Since this capacity affects the wave length it is plain that even though a small inductance be used in the closed circuit, the wave length will be large; and this is one reason why the wave length of high-power installations is large; there are other reasons which are taken up on p. 372. In the example given, even if the inductance in the closed circuit were 200 μh (which is comparatively small) the wave length would be:

$$1885\sqrt{200\times 0.01}=2660 \text{ meters.}$$

Again, from the formula:

$$\frac{CV^2}{2}N=W$$

we see that

$$\frac{I^2R}{\eta}=W=N\frac{CV^2}{2}, \quad \quad (1)$$

where I = the current in the antenna;
R = effective resistance of antenna;
η = efficiency of transformation from condenser to antenna.

Hence,

$$I=V\sqrt{\frac{\eta NC}{2R}}, \quad \quad (2)$$

or the antenna current varies with the square root of the capacity.

To show this a test was made on a transmitter with the apparatus connected as shown in curve sheet, Fig. 2, where the ammeter measures the high-frequency current in the closed circuit. Of course the current in the antenna, which is here not shown, would be directly proportional to the current in the closed circuit. In this test the gap length was kept constant, the value of the capacity was varied and the voltage of the alternator was regulated until, for every case, a spark occurred for each

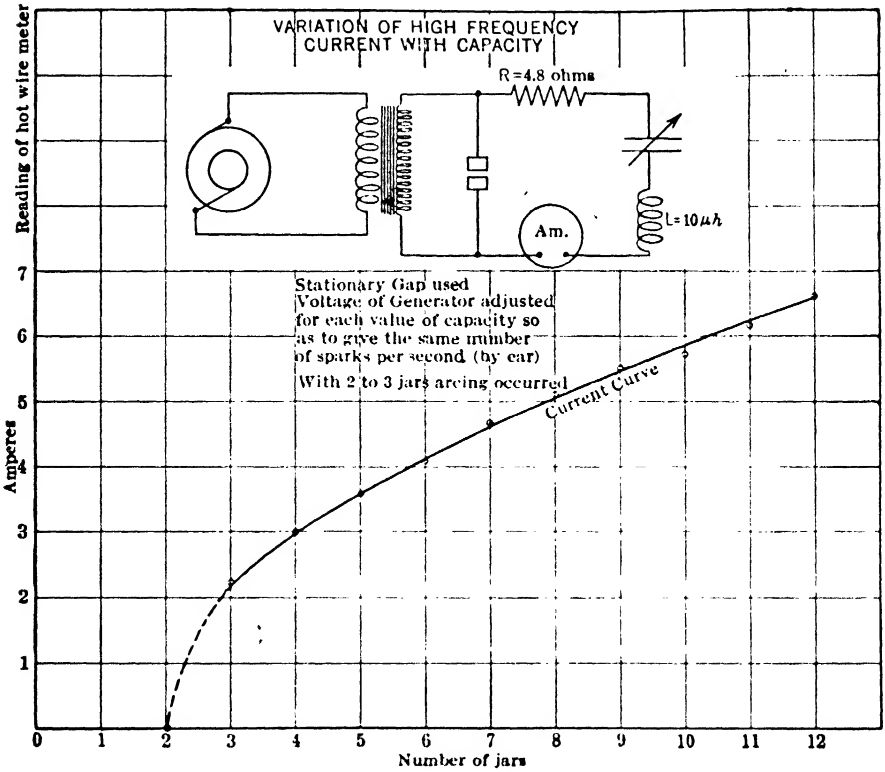


FIG. 2.—Variation of the high-frequency oscillatory current with the amount of capacity used.

alternation (as could be approximately determined by the pitch of the spark note); this meant that the voltage to which the condenser was being charged was the same, and, furthermore, that the condenser was being charged and discharged once for every alternation. Under these conditions the high-frequency current should be proportional to $\sqrt{C}$, and the square of the current proportional to C. The curve obtained shows this to be approximately the case, except that an intercept is noted at the point corresponding to two jars due to the fact that for such a low capacity the gap could not be kept from arcing, which, in turn prevented the periodic and regular charging and discharging of the condenser.

Design of Audio Circuit.—We may now discuss more fully the choice of the various parts of the so-called " audio circuit," which comprises (see Fig. 1) the alternator, the variable reactance, the step-up transformer, the choke coils, and the condenser. This circuit may be simplified by noting that a transformer may be treated approximately as a simple circuit consisting of an inductance and a resistance entirely transferred to the high- or to the low-tension side; furthermore, any impedance in the secondary circuit may be transferred to the primary by multiplying by a suitable factor. On this basis the audio circuit may be simplified to that of Fig. 3.

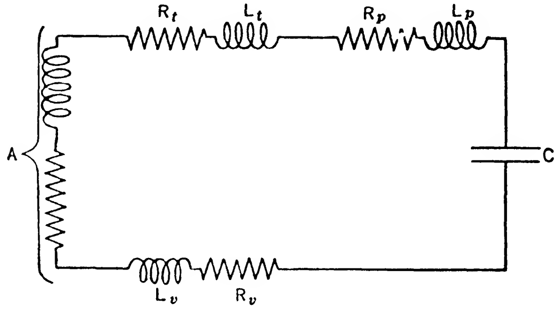


FIG. 3.—An approximate simplification of the low-frequency circuit of a radio transmitter.

where A = alternator armature, having both resistance and inductance;
R_t = resistance of both transformer coils transferred to low-tension side;
L_t = leakage inductance of both transformer coils transferred to low-tension side;
R_p = resistance of protective choke coils transferred to low-tension side;
L_p = inductance of protective choke coils transferred to low-tension side;
C = condenser capacity transferred to low-tension side;
R_v = resistance of variable reactance coil in low-tension side;
L_v = inductance of variable reactance coil in low-tension side.

The circuit of Fig. 3 may be still further simplified to that of Fig. 4 where

A = alternator *without* inductance or resistance;
R = resistance of entire circuit, including alternator armature, transformer coils, protective choke coils, variable reactance coils, on basis of low-tension side;
L = inductance of entire circuit (ditto);
C = capacity of condenser transferred to low-tension side.

It has already been stated that this circuit should be adjusted so that it will have a natural frequency about equal to that of the alternator;

it has also been shown how the capacity of the condenser may be calculated if the voltage to be used and the power required are known; it follows then that, knowing the capacity, the value of the total inductance L in the circuit of Fig. 4 may be easily calculated from formula:

$$f=\frac{1}{2\pi\sqrt{LC}},$$

and this inductance may then be apportioned between the alternator, the variable reactance, the transformer, and the choke coils. Calculation of a typical case is carried out as follows:

Assume that the transformer ratio is 1 : 80; then, any inductance or resistance in the high-tension side may be transferred to the low-tension side by dividing by 80^2, or 6400, while a capacity in the high-tension side may be transferred to the low-tension side by multiplying by 6400.

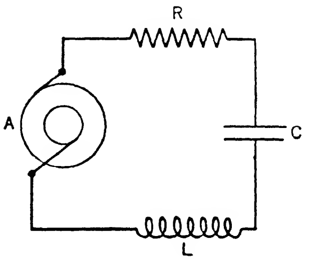


FIG. 4.—Simplest possible representation of the low-frequency circuit, not quite equivalent to the actual circuit. It neglects the magnetizing current of the transformer, which may be comparatively large in an open core transformer, as is generally used.

In our case:

Capacity of condenser in high-tension side $=0.012$ μf. Hence, equivalent low-tension capacity $=0.012\times6400=77.0$ μf.

If the audio circuit must resonate at 500 cycles per second,

Total equivalent low-tension inductance

$$=\frac{1}{500^2\times4\pi^2\times77.0\times10^{-6}}=0.00133 \text{ henry}.$$

Of this inductance probably the largest part is, in a modern set, found in the alternator, while the transformer has comparatively little inductance, and the balance is made up by the choke coils in the high-tension side and the variable reactance *V.R.* in the low-tension side. Most alternators used in spark transmitters are of the inductor type, and this type always has a comparatively high internal reactance; if it were not for this fact the transformer, which is generally of the open-core type, would have more reactance than the alternator.

In order to show the manner in which the whole audio circuit may be made to resonate the curves of Fig. 5 are here given as being representative of an actual set. In obtaining these curves the field current and speed of the alternator were kept constant, while the capacity in the high-tension side of the transformer was changed, with the circuit connections as shown in Fig. 5. Under these conditions, as the capacity, and, therefore, the natural frequency of the circuit, was varied, the current in the primary of the power transformer as well as the voltage across it varied and reached a maximum at the point corresponding to resonance conditions. A capacity of about 5.5 Leyden jars is seen to have produced resonance. In an actual set the adjustment of the capacity, or of the

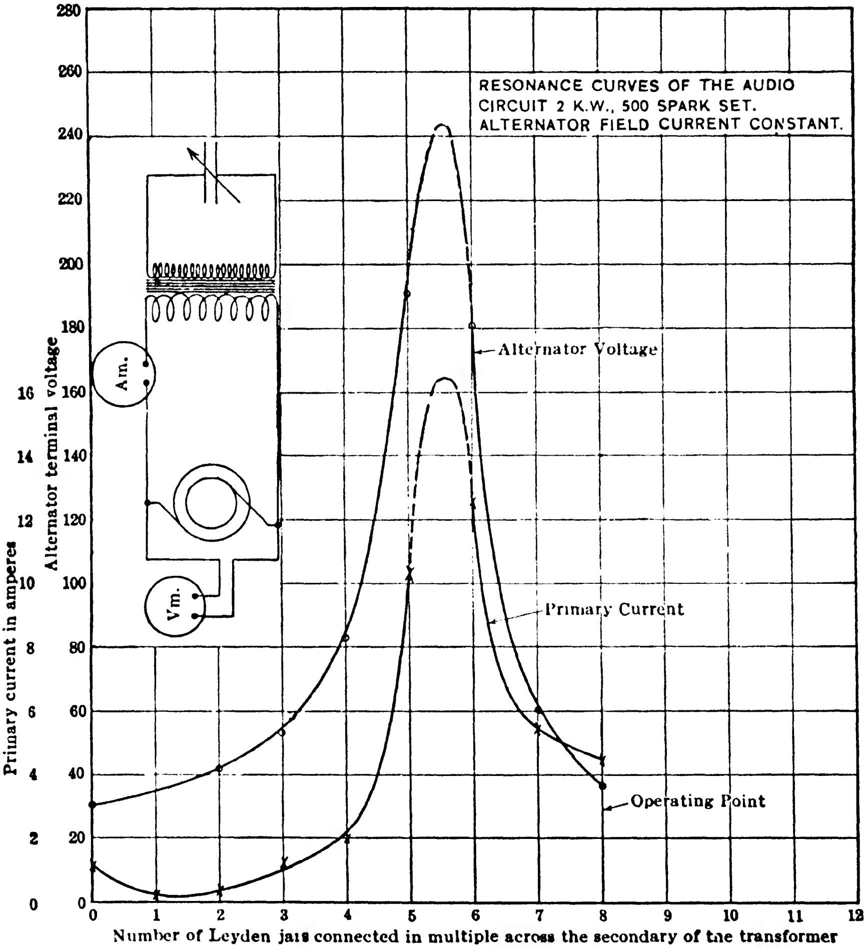


FIG. 5.—Variation of alternator voltage and primary current of a 2-kw. spark transmitter as the capacity in the secondary of the transformer was varied; field current of alternator and speed held constant. Gap set too long to permit sparking at voltage of test.

inductance, is made about 20 per cent to 30 per cent larger than necessary to give resonance at audio frequency, thus making the natural frequency of the circuit somewhat lower than the alternator frequency. This point will be more fully emphasized further on. In the case represented by the curves of Fig. 5 the set was actually operated with 8 Leyden jars across the secondary.

It now remains to investigate, as far as the conditions will allow, the transient phenomena taking place in the audio circuit as the condenser is charged and discharged; we especially mean to refer to the variation of the condenser current and voltage as the alternator e.m.f. is impressed upon the audio circuit, and thereafter, as the gap breaks down. As shown in Fig. 4, we are dealing with an oscillatory circuit, having the resistance, inductance, and capacity R, L, and C, respectively, upon which there is impressed a harmonic e.m.f. The equation for the instantaneous value of current for such a circuit was derived in Chapter III, p. 332, and is

$$i=\frac{E}{\sqrt{R^2+\left(pL-\frac{1}{pC}\right)^2}}\sin(pt-\phi)+A\epsilon^{-\frac{Rt'}{2L}}\sin\omega t', \quad . \quad . \quad . \quad (3)$$

in which p = angular velocity of impressed force;

ω = angular velocity of natural oscillations of the circuit;

ϕ = phase difference of E and I in the steady state;

A = a constant to be determined;

t' = time of duration of the transient term.

In deriving this equation (p. 335) it was shown how to solve for A and t', these depending for their value on the time the voltage is introduced into the circuit. In a radio set no switch is actually used, but the equivalent effect is caused by the operation of the spark gap; when the gap is sparking its resistance is so low that the secondary of the power transformer is short-circuited and this is the condition for comparatively small current in the armature circuit. When the gap opens (ceases to carry current) the effect of the condenser of the closed oscillating circuit is to so neutralize the inductance of the transformer and armature that the current rises to comparatively large values. We may get a fair idea of the behavior of the actual radio circuit, therefore, by supposing that Eq. (3) holds good, the voltage, $E \sin pt$, being introduced into the circuit at *the instant when the gap opens.* As mentioned when analyzing the action of this circuit before (page 336) the general solution is difficult, but we can get fairly easy solutions if we assume that the condenser is completely discharged at every oscillation and that the gap opens when the voltage of the generator is zero; this latter condition may be approximately satisfied by suitable adjustment of the set.

Assuming that the low-frequency circuit is resonant to the alternator voltage, that the gap opens when generator voltage is zero, and that the condenser is discharged, Eq. (3) becomes

$$i=\frac{E}{Z}\sin pt-\frac{E}{Z}\epsilon^{-\frac{Rt}{2L}}\sin\omega t=\frac{E}{R}\sin pt(1-\epsilon^{-\frac{Rt}{2L}}). \quad . \quad . \quad . \quad (4)$$

From this we find the voltage across the condenser; it is

$$v_c=\frac{E}{CR}\left\{-\frac{1}{p}\cos pt+\frac{\epsilon^{-\frac{Rt}{2L}}}{p^2+\left(\frac{R}{2L}\right)^2}\left(\frac{R}{2L}\sin pt+p\cos pt\right)\right\},$$

which when $(R/2L)^2$ is small compared to $(p)^2$ gives

$$v_c=\frac{E}{pCR}\left\{\frac{R}{2pL}\epsilon^{-\frac{Rt}{2L}}\sin pt-\cos pt(1-\epsilon^{-\frac{Rt}{2L}})\right\}. \quad . \quad . \quad . \quad . \quad (5)$$

This equation shows that the condenser voltage rapidly changes its phase during the first few alternations, the sin pt term predominating at first, the cos pt term being zero; as soon as $\epsilon^{-\frac{Rt}{2L}}$ departs appreciably from unity the cos pt begins to predominate and this continually increases

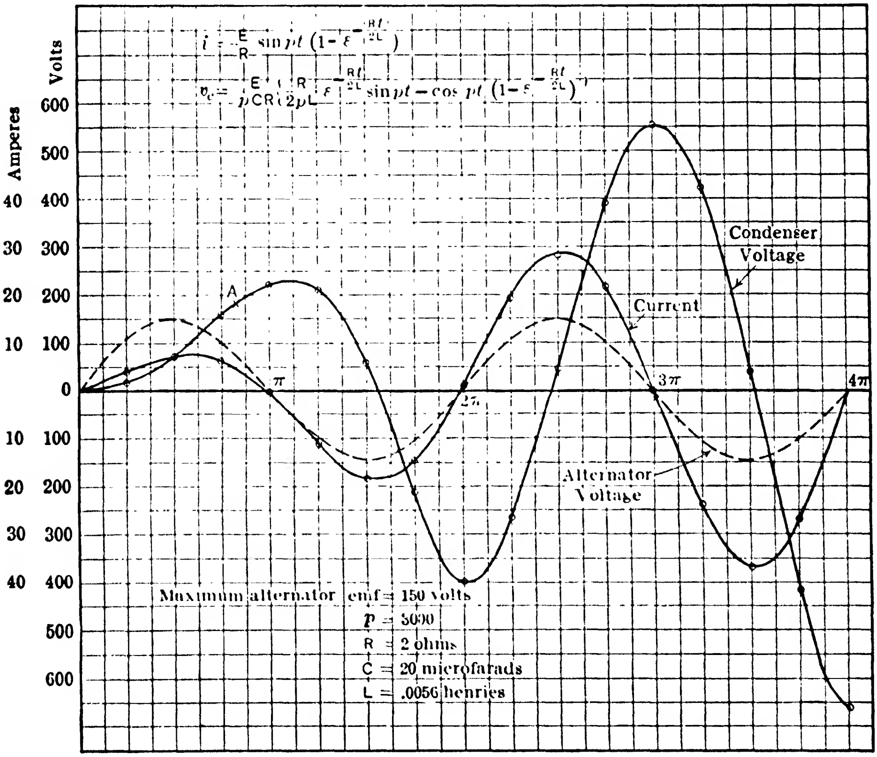


Fig. 6.—Transient current in audio circuit of a spark transmitter, circuit tuned to alternator frequency.

with increasing time due to the increasing value of $(1-\epsilon^{-\frac{Rt}{2L}})$. Thus in the steady state $(\epsilon^{-\frac{Rt}{2L}} \cong 0)$ Eq. (5) reduces to the familiar form $v_c = -(E/pCR) \cos pt$.

The curves of Fig. 6 show the form of current and voltage across the condenser for a typical circuit, the values of the various constants being noted on the curve sheet. It will be noticed that the condenser voltage reaches its maximum values at approximately the times when

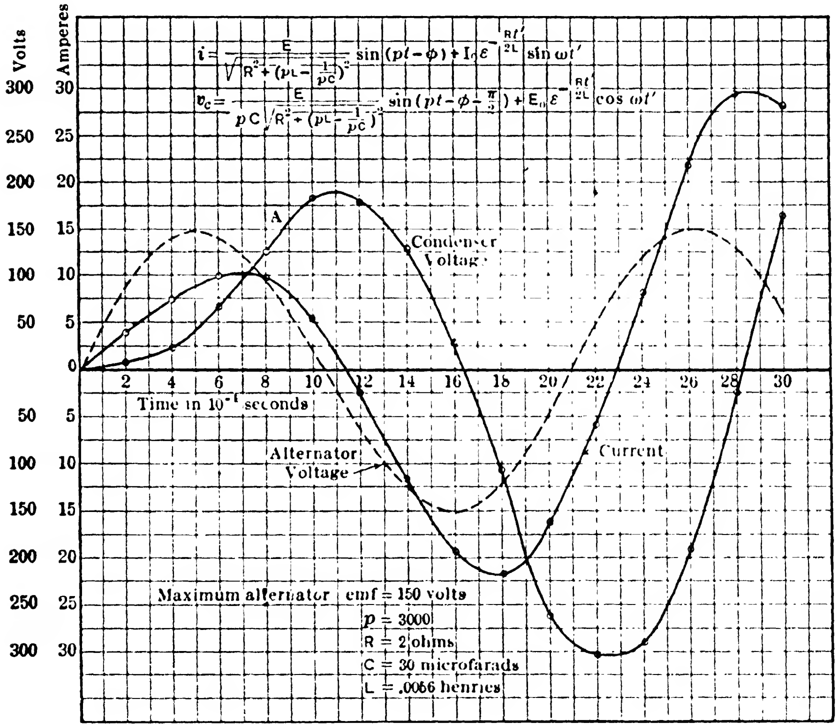


FIG. 7.—Transient current in audio circuit of a spark transmitter, circuit frequency being about 20 per cent lower than alternator frequency.

the impressed voltage is zero, and hence the spark gap will break down at about this time; the resulting oscillatory current in the closed oscillating circuit at once discharges the condenser, the spark gap opens and the voltage of the alternator, passing through its zero value, is again impressed on the circuit to produce the next transient. It is to be seen that if events follow the order given here the assumed condition ($e=0$ when gap opens) is satisfied.

It is found in practice that the condition of resonance assumed in this analysis tends to produce irregular sparking, giving the signal a ragged note, so actually the natural frequency of the circuit is made about 20 per cent lower than the frequency of the alternator. On the assumption that the spark gap again opens when the generator voltage is passing through zero the curves of Fig. 7 have been constructed for the same circuit as used for Fig. 6 with the exception that the capacity has been increased from 20 μf to 30 μf, this giving about the same amount of detuning as is used in practice.

For this case the form of current and condenser voltage are obtained by the use of Formulas (80)–(83) of Chapter III. Supposing that the gap opens the circuit at the instant the circuit voltage (that produced by the alternator) is zero and increasing, it is found that for the steady state the current should be -23.5 amperes and the voltage across the condenser should be -92 volts. To satisfy the condition that the actual current must be zero as well as the drop across the condenser a transient term must be added to the steady state solution; by the process outlined in Chapter III this transient is found to be satisfied by charging the condenser to -385 volts and starting this transient term 0.000756 second before the alternator voltage goes through its zero value—this transient term has the natural frequency of the circuit (given practically correct by putting $\omega=(1/\sqrt{LC})$) and a damping fixed by the R and L of the circuit. The actual current is obtained by taking the sum of the steady term and transient term and is

$$i=\frac{150}{\sqrt{2^2+\left(3000\times 0.0056-\dfrac{10^6}{3000\times 30}\right)^2}}\sin(3000t-70.7^\circ)$$
$$+28.2\epsilon^{-\frac{2(t+0.000756)}{2\times 0.0056}}\sin\{2440(t+0.000756)\}$$

Actually the effective resistance of the circuit would be much greater than 2 ohms (the value given in this problem), owing to iron losses in the circuit, but as they are neglected in this problem R is given the value due to the copper wire.

Similarly the equation for voltage drop across the condenser is found to be represented by:

$$V_c=\frac{150\times 10^6}{3000\times 30\sqrt{2^2+\left(3000\times 0.0056-\dfrac{10^6}{3000\times 30}\right)^2}}\sin\left(3000t-70.7^\circ-\frac{\pi}{2}\right)$$
$$-385\epsilon^{-\frac{2(t+0.000756)}{2\times 0.0056}}\cos\{2440(t+0.000756)\}$$

It may be seen from Fig. 7 that the voltage across the condenser is rising more rapidly, at times $t=\pi$, than was the case for the resonant condition depicted in Fig. 6; it is quite likely that this more rapid rise in condenser voltage, by causing the spark to take place at a more definite time, accounts for the more regular behavior of the spark when the circuit is detuned as supposed in Fig. 7, than when the circuit is resonant.

In Figs. 6 and 7 the forms of current and condenser voltage have been shown for nearly two cycles; actually if a spark occurs at the time indicated by the letter A on the condenser voltage curve (which is the time the spark should actually occur) the condenser voltage drops to zero and it, as well as the current, goes through the same changes from π to 2π as it did from $0°$ to π. The actual forms of the condenser voltage for the circuits analyzed in Figs. 6 and 7 are shown in Figs. 48 and 49 of Chapter III, p. 339. It will be seen that the above analysis does give fairly accurate results.

Types of Spark Gaps.—The construction of the different commercial types of spark gaps in use at the present time may be conveniently subdivided into the following classes:

(*a*) Synchronous rotating gap.

(*b*) Quenched gap $\begin{cases}\text{self-cooled}\\ \text{fan-cooled.}\end{cases}$

Synchronous Rotating Gap—Construction and Operation.—The construction of the synchronous rotating gap is indicated in Fig. 8, where the rotating electrode shown is simply a toothed wheel, rigidly fastened

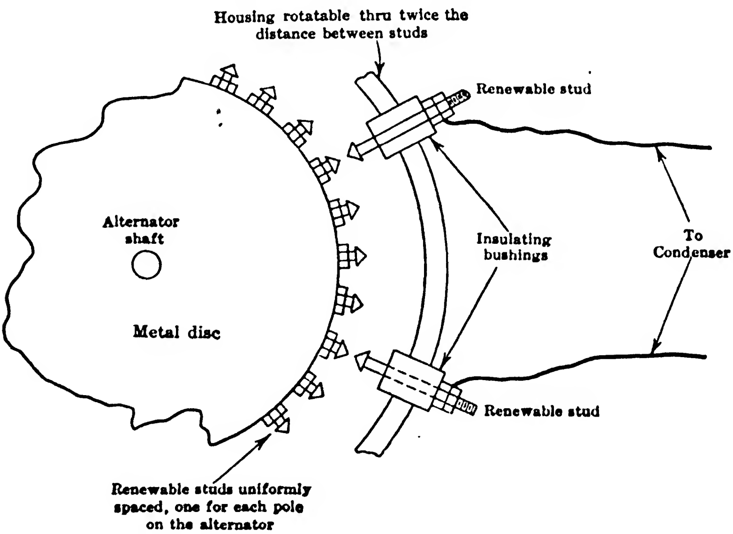


FIG. 8.—Arrangement of parts of a synchronous rotating gap; instead of using a metal disk for the rotating member this is sometimes made of a disc of Bakelite or similar material, the rotating studs being then all connected together by a metal strip.

to the alternator shaft. The spark jumps from one fixed electrode to the disc, through the disc and thence back through the second gap to the other electrode.

The position of the disc on the shaft is adjusted permanently so that the teeth line up with the fixed electrodes at the time of maximum values (positive and negative) of the voltage wave, and the gap separation adjusted so that the breakdown voltage is slightly below the maximum voltage. Under these conditions the gap breaks down once during each half cycle, and assuming a 500-cycle supply, the group frequency is evidently 1000. The number of teeth on the disc is determined by the number of alternator field poles. For instance, if the alternator be equipped with 24 poles, the disc would have 24 teeth, and 24 breakdowns would occur per revolution. This would correspond to an alternator speed of 2500 r.p.m. if a group frequency of 1000 were desired.

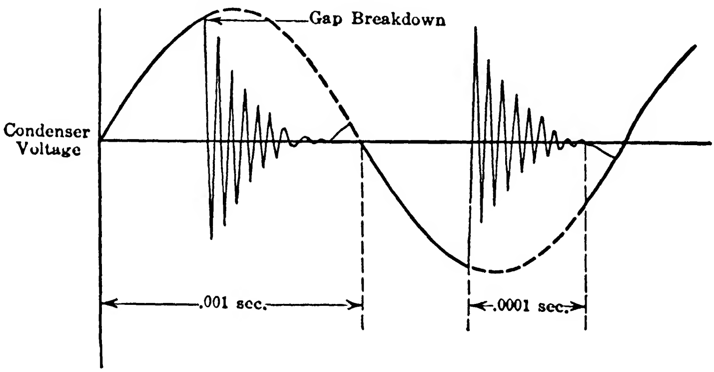


FIG. 9.—Conventional representation of audio- and radio-frequency currents; actually the voltage across the condenser does not have the sinusoidal shape given here but has the form given in Figs. 6 and 7.

Clearly, the number of breakdowns per revolution may be controlled by substituting discs with different tooth spacing. Thus, we could omit alternate teeth, and cut the group frequency in half, etc. The tone of a signal may thus be altered easily and quickly, in case this is found desirable because of interference effects present. The quality of the note may be made quite distinctive by introducing regular irregularities in the arrangement of teeth as, e.g., omitting every third tooth.

The action of the gap, assuming one breakdown to occur every half cycle, is indicated conventionally in Fig. 9. Actually the condenser voltage is not a sine wave, but has the peculiar form shown in Figs. 6 and 7.

Synchronous Gap Application.—The synchronous gap possesses a low operating resistance, due to the electrodes being close together at the time of discharge, and automatically recovers its insulating properties between discharges due to the electrodes being widely separated during this interval. Arcing is prevented by the separation of the electrodes increasing as the wave train passes, and also by the fanning and cooling action of the rapidly moving electrodes. Partial discharges cannot occur, as the gap separation may be adjusted for breakdown near the voltage maximum. This form of gap will successfully handle large amounts of power and high spark frequencies, and has been widely used on commercial spark transmitters of large capacity.

Quenched Gap.—The property which a gap possesses of returning very quickly to its un-ionized condition is termed "quenching." In the rotating gaps, quenching is obtained principally by the air blast which

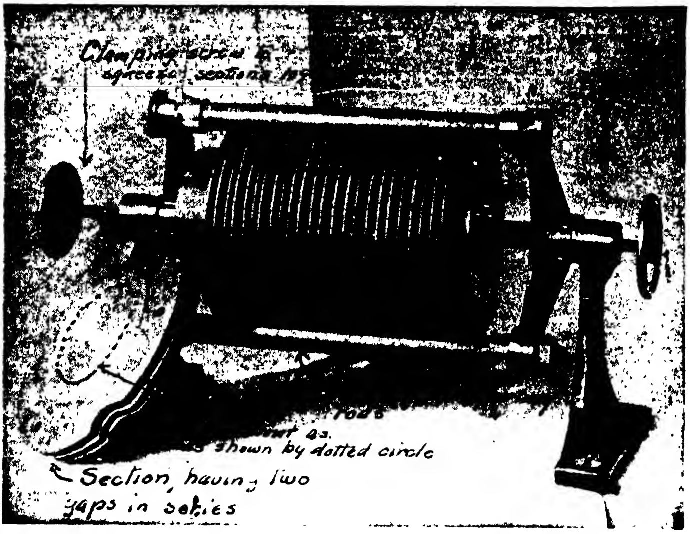


FIG. 10.—Photograph of a commercial type of quenched gap; three discs clamped together by insulating screws make up a unit, there being two gaps in series per unit.

occurs at the sparking contacts, and also to some extent perhaps by the high velocity of the moving electrodes, thus preventing arcing and permitting the voltage to build up again across the condenser, as already noted.

In place of mechanical quenching action, as illustrated by the rotary gaps, an electrical quenching type is also widely used, which is known as the "quenched gap." In this type, the return of the gap to a condition of high dielectric strength is obtained through very rapid de-ionization of the gap between the electrodes. The construction of a typical gap is shown in Fig. 10. The following description of its action will explain the peculiar cellular form of construction illustrated.

Quenched Gap—Requirements of Rapid De-ionization.—For the gap to operate satisfactorily, that is, return to its un-ionized condition in an extremely short time, the following conditions must be fullfilled:

1. The spark must take place in a space in which no oxide is formed. This is because the oxide will deposit on the sparking surface of the gap and soon short-circuit it.

2. The metal surfaces must be kept cool and the electrodes must therefore be good heat conductors. Silver or copper are the metals which best fulfill this requirement. Usually silver-plated copper electrodes are employed.

3. No part of the gas which forms the gap dielectric must be far from a cool metal surface, that is, a very short gap only may be used.

Quenched Gap—Construction.—The above requirements are satisfied in the commercial form of gap as follows:

1. The spark takes place in a practically air-tight chamber. The several elements or sections of the gap are separated from one another by the insulating gaskets as shown (Fig. 11) and the whole clamped tightly together. When the gap is first operated, the air, which is initially between the gap faces, becomes separated into its elements, mainly oxygen and nitrogen, the oxygen combining with the copper electrodes to form copper oxide, thus leaving an atmosphere of essentially pure nitrogen between the gap faces. The black oxide of copper disappears after the gap has been in operation a short while, the gap faces being found bright and clean if the gap is disassembled for inspection. (The exact reason for the disappearance of this oxide is not apparent—it is probably absorbed into the material of the separating gasket, under conditions present when the gap is in operation.)

2. In addition to using good heat-conducting materials, such as silver and copper, for the electrodes, the efficient cooling of the gap is assisted by means of cooling vanes or fins, which radiate the heat produced during the operation of the gap. These fins are clearly indicated in the photograph (Fig. 10). There has been recently developed a staggered form of gap construction, which permits air circulation on both sides of each element.

This construction, whereby cooling is accomplished by increased radiating surface, represents what is known as the self-cooled type. It is sometimes necessary, with the higher-powered sets, to supply a small motor-driven fan to cool the gap satisfactorily.[1]

3. The requirement that no particle of gas in the gap shall be remote from a metal surface is satisfied by subdividing the gap into sections, the number of sections increasing as the " breakdown " voltage value is increased. Each gap provides somewhat less than 0.01 inch separation,

[1] Most modern gaps receive their supply of cooling air from a fan mounted on the alternator shaft, thus dispensing with the extra motor required for blower.

with a breakdown voltage of approximately 1200 volts. Thus no particle of gas in the gap is more than 0.005 inch away from the metal, and the gap is rapidly de-ionized. This rapid de-ionization is due principally to the loss of electrons by diffusion, although recombination of electrons and positive ions is also a factor. By loss of electrons by diffusion is meant the removal of electrons from the gas to the face of the gap, due to the attraction of the induced positive charges on the gap faces. As the most distant electron has only a short distance (0.005 inch) to go before arriving at the gap face and the attracting charge, the time required is extremely small.

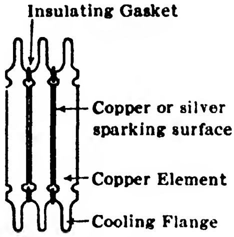


FIG. 11.—Cross-sectsonal sketch of part of a quenching gap.

Instead of having an initial dielectric of air between the gap faces, one of several gases may be used. Using air, carbon dioxide, or nitrogen, Pieck[1] has shown that quenching is about the same for any of these gases and greater than for hydrogen. The gas pressure adversely affected the quenching and, at high pressures, this quenching is also affected by the wave length, being best at the longer wave lengths. These results may be deduced by consideration of the theory of the gap as outlined above.

In case the quenched gap does not function as it should, there may be more than one spark per alternation; this is shown in Fig. 12. When this

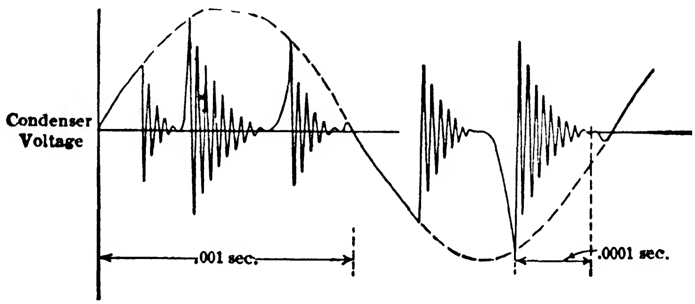


FIG. 12.—If the quenched spark gap does not function properly, there may be two or more sparks per alternation.

occurs the musical quality of the note changes and also the gap is very likely to overheat and spoil. Such an effect may be produced by improper coupling between the closed oscillating circuit, where the gap is used, and the antenna circuit.

[1] "On the Phenomenon of Quenched Sparks," V. Pieck, Jahrbuch d. Drahtl. Tele., Jan., 1920.

The Radio-frequency Circuits.—These consist of the closed and open oscillatory circuits, coupled together through the oscillation transformer. The whole of the radio-frequency circuit for a two-coil oscillation transformer is shown in Fig. 13. The closed and open circuits are tuned to the same frequency.

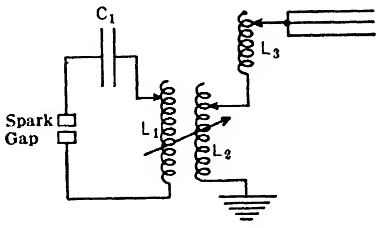


FIG. 13.—The two coupled radio-frequency circuits of a spark transmitter.

The theory applying to the above is that which has been discussed in connection with two inductively coupled oscillatory circuits (see Chapter III, pp. 309, et seq.). The main point to be considered is that when the two circuits are closely coupled there are produced in each two currents of frequencies differing from the natural frequency of the two circuits; when the natural frequencies of the circuits are the same, then the frequency and wave length of the component currents are given by (see Chapter III, pp. 311, et seq.).

$$f'' = \frac{f}{\sqrt{1-k}}, \qquad \lambda'' = \lambda\sqrt{1-k},$$

$$f' = \frac{f}{\sqrt{1+k}}, \qquad \lambda' = \lambda\sqrt{1+k},$$

where f and λ = natural frequency and wave length of either circuit;
f' and λ' = frequency and wave length of one of the component currents;
f'' and λ'' = frequency and wave length of the other component currents;
k = coefficient of coupling.

The relative amplitudes of the two currents have been discussed in Chapter III, pp. 313, et seq.; generally the higher-frequency current has the greater amplitude. Furthermore, the higher-frequency currents of the primary and secondary are about 180° apart, while the lower-frequency currents of the primary and secondary are about in phase. The effect of all this is to produce current " beats " in the primary and secondary with a frequency equal to the difference of the frequencies of the component currents; again, while the resultant current in the primary is passing through the small amplitude values of the " beat cycle," the secondary current is passing through the high amplitude values of the " beat cycle," and vice versa. This is illustrated by the curves of Fig. 14, where the dotted-line curves represent the resultant primary and secondary currents;

it will be noted that the primary resultant current starts with a high amplitude at Q and decreases to a low amplitude at R, while the secondary resultant current does just the opposite. In plotting the curves it has been assumed that neither circuit suffers any losses, and the result is that the decrement of the component currents is zero, while the resultant currents would also periodically repeat themselves through the " beat cycle" without any decay. This of course is not true of an actual case, where, on account of the losses in both circuits, the decrement would have a definite value, and the resultant currents would " decay " somewhat as shown in Fig. 15, which represents the component and the resultant primary and secondary currents for circuits with decrements. Another assumption made is that the gap used is such (open spark gap) that it

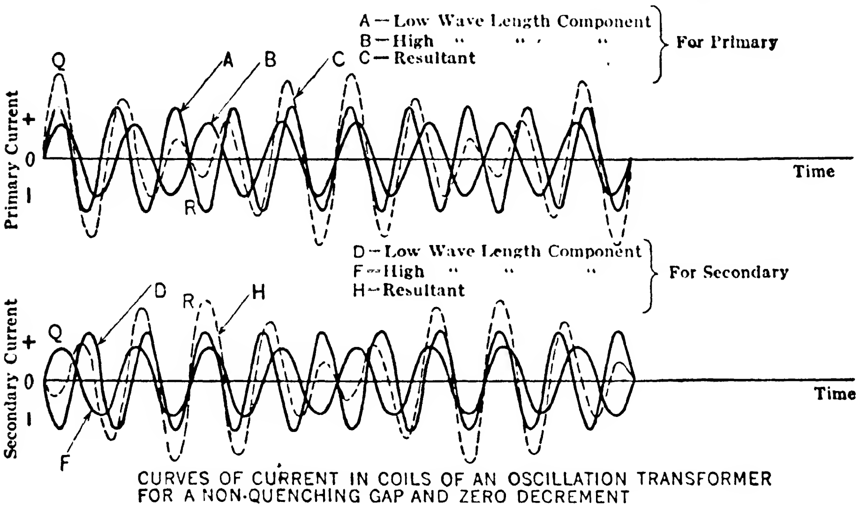


FIG. 14.—Currents in the two circuits of Fig. 13, no damping assumed.

remains closed for considerable time after its breaking down, so that the currents may flow through the closed circuit.

The phenomenon of the " beats " takes place most pronouncedly when the coupling between the primary and secondary of the oscillation transformer is closest. For loose coupling the two circuits oscillate at very nearly a single frequency equal to their natural frequency, but when this is the case the secondary current is generally low. On the other hand, when the coupling is very close, although the current in the antenna is large, yet since it is made up of two component currents of two widely different frequencies the antenna will radiate energy at these two different frequencies; this is very objectionable because the total available energy is subdivided, and hence the range of transmission diminished, and also because it would interfere with other stations. As a matter of fact, the law in the United States requires that the energy of no other frequency shall exceed 10 per cent of that of the frequency on which the station is transmitting.

As outlined above, we find that when an open spark gap is used, which remains closed for some time after its breaking down and thus permits a current to be maintained in the closed circuit, we are confronted by either one of two evils, i.e., *low current* in antenna at a single frequency for loose coupling, and large antenna current *of two frequencies* for close

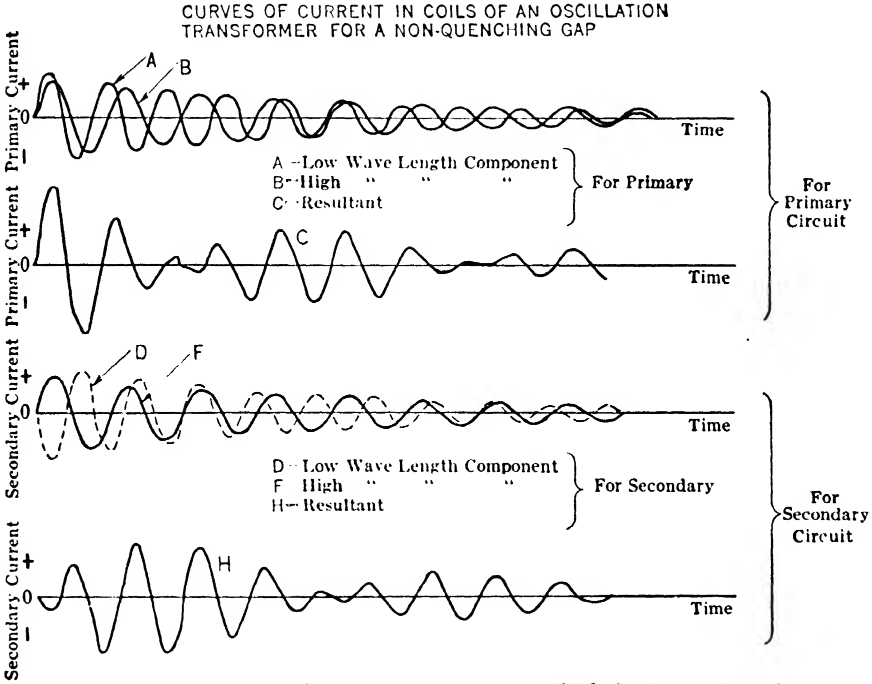


FIG. 15.—Currents in the two circuits of Fig. 13, high damping assumed.

coupling; besides, for both loose and close coupling, energy is wasted in the primary, since the latter has a current flowing in it for a longer time than necessary, which produces unnecessary losses, and subtracts from the energy which might otherwise be given to the antenna.

In order to overcome these difficulties advantage is taken of the fact that, as has already been pointed out, and as shown in Fig. 15, the primary and secondary currents (for close coupling and an open spark gap) pass through beat cycles, and that the amplitude of the primary current has minimum values at the same time that the amplitude of the secondary current has maximum values. It is plain that if the primary current were automatically interrupted when passing through its minimum amplitude values, the secondary circuit would then go on oscillating at its own frequency and damping. This, of course, would be made possible by the fact that the primary current would be interrupted when the secondary current amplitude values are a maximum and hence when almost the entire energy is in the secondary. According to this plan the current would be interrupted in the primary at the completion of the first one-half of a " beat-cycle " as shown at *A* in Fig. 16. To interrupt the primary current several methods may be used, the simplest of which is by replacing the ordinary open gap by the so-called " quenched gap." The construction of this already has been described on p. 418. Its characteristic is that it " opens " when the current in the primary of the oscilla-

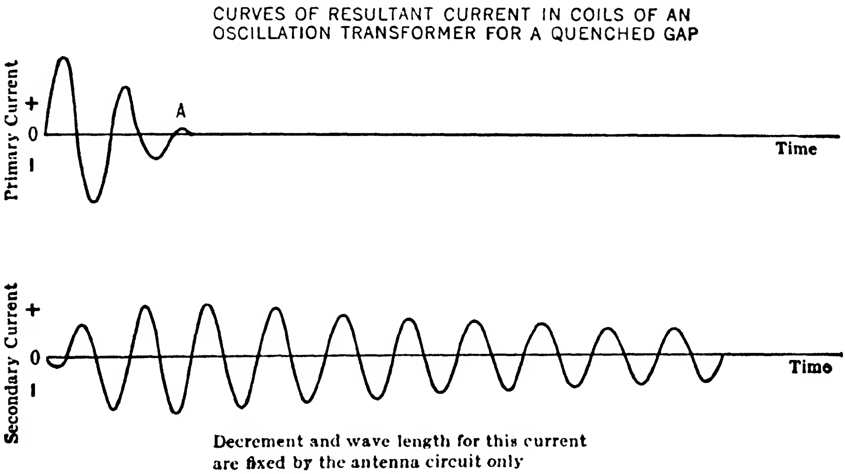


FIG. 16.—Currents in the two circuits of Fig. 13 if the gap used in the closed circuit is of the quenching type.

tion transformer passes through its low values, probably because the comparatively few ions, which are formed between the sparking surfaces during the time of low-current amplitude, disappear very quickly, thus making it impossible to maintain low currents through the gas between the sparking surfaces. Such a gap is said to " quench " the spark formed upon the discharge of the condenser. A quenched gap may be and is generally operated with a close coupling of the oscillation transformer, because the closer the coupling the greater the amount of energy transferred to the antenna circuit; but if the coupling should be made extremely close then it is possible that the gap may refuse to quench, because of the very short time during which the closed circuit has its low amplitude current; this may not be sufficient to permit the gap to quench. A critical coupling, therefore, exists at which the gap quenches best; this coupling is quite close and far closer than could be used with an ordinary open gap; the secondary current, as indicated by an ammeter, is a maximum for the critical coupling. Of course if the gap is quenching properly the secondary current should have a frequency equal to its natural frequency, and, since no current flows in the primary, the efficiency is higher and the decrement lower than for the "open gap."

The adjustment of a transmitting set as regards the coupling of the closed and open circuits, the gap, and the tuning of the two circuits is best determined by obtaining the "energy distribution curve." Such a curve is obtained in the following manner: a search coil of one or two turns is introduced in the antenna circuit as shown at S, Fig. 17, and a wave-meter circuit, consisting of L_4, C_4, and a hot-wire meter A, is loosely coupled to S. With the transmitter in operation the capacity C_4 is set at different values, and reading of A is obtained; thus, as the natural wave length of the circuit of $C_4 - L_4 - A$ is varied, the ammeter reading varies. A curve plotted with values of the natural wave lengths of circuit $C_4 - L_4 - A$ against squares of ammeter readings is known as "energy-distribution curve," and shows the relative amounts of energy radiated by the antenna at each wave length. Another way to look at it is that, since the circuit $C_4 - L_4 - A$ is nothing but a receiving circuit loosely coupled to the transmitting antenna, it follows that the energy distribution curve also represents the energy reaching the receiving circuit when it is adjusted to different natural wave lengths. Whichever way one chooses to look upon the "energy distribution curve," it is plain that it is of great importance in the study and adjustment of a transmitting set. Two typical sets of such curves are given in Figs. 18 and 19, and a study of these will bear out some of the points brought out in the previous discussion. In these curves the ordinates represent squares of currents, and they were in one case read on a so-called "wattmeter"[1] and in the other on a thermo-galvanometer.

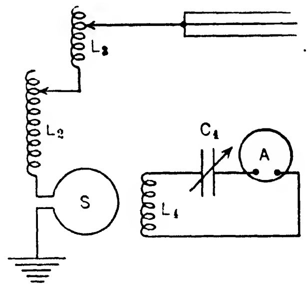
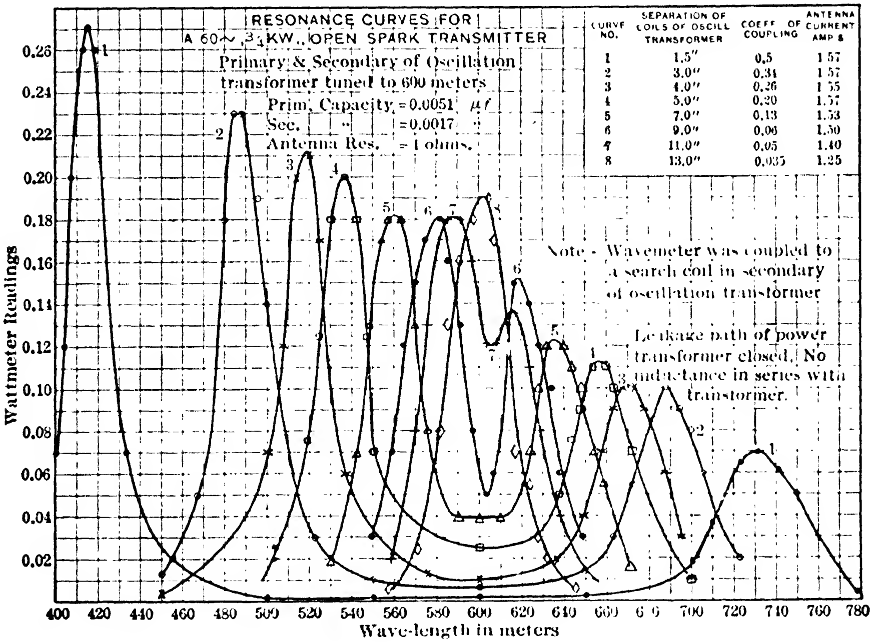


FIG. 17.—Use of wave meter for getting the energy distribution curve of antenna circuit.

Fig. 18 shows curves for an open gap and for different amounts of coupling, curve 1 being for the closest and curve 8 for the loosest coupling. It will be seen that for any but the loosest coupling there are

[1] Wattmeter is the name often given, in radio measurements, to a hot-wire ammeter the scale of which is calibrated to indicate the power expended in the resistance of the instrument itself.

two maxima in the radiation of the antenna at two different wave lengths more or less separated from each other; thus, for curve 1, the two wave lengths are 732 and 415 meters, while for curve 7 they are 616 and 588 meters. On the other hand, for curve 8 maximum energy is radiated at the one wave length of 602 meters, i.e., the natural wave length of the closed and open circuits. Again, by referring to the table inserted in Fig. 18, we note that the antenna current was a minimum for curve 8 (1.25 amperes) and a maximum for curve 1 (1.57 amperes). Or, as already pointed out, the loose coupling produces an antenna current which, though



CURVE NO.	SEPARATION OF COILS OF OSCILL TRANSFORMER	COEFF OF COUPLING	ANTENNA CURRENT AMPS
1	1.5"	0.5	1.57
2	3.0"	0.34	1.57
3	4.0"	0.26	1.55
4	5.0"	0.20	1.57
5	7.0"	0.13	1.53
6	9.0"	0.08	1.50
7	11.0"	0.05	1.40
8	13.0"	0.035	1.25

FIG. 18.—A set of resonance curves for a spark transmitter having a non-quenching gap; even when the coupling is as low as 5 per cent two distinct waves are emitted from the antenna.

smaller than for close coupling, radiates maximum energy at a single frequency or wave length.

In Fig. 19 are shown some energy distribution curves for a set having a quenched gap, the values of coupling used being noted on the curve sheet. It will be seen that the radiation for any but the weakest coupling was impure, i.e., took place at more than one frequency. As the coupling is increased, in a quenched gap transmitting set, from very low values, the antenna current, as read on the ammeter, will increase with the increasing coupling; for a certain coupling the antenna current reaches a maximum and then decreases sharply for a further small increase in coupling. The value of coupling just less than that at which the antenna current decreases is the proper one to use; it is the maximum value which can be used and still maintain the quenching action of the set.

Adjusting the Spark Transmitter.—In adjusting the transmitter shown in Fig. 1, to radiate at a certain wave length and energy output, the following schedule of procedure is followed:

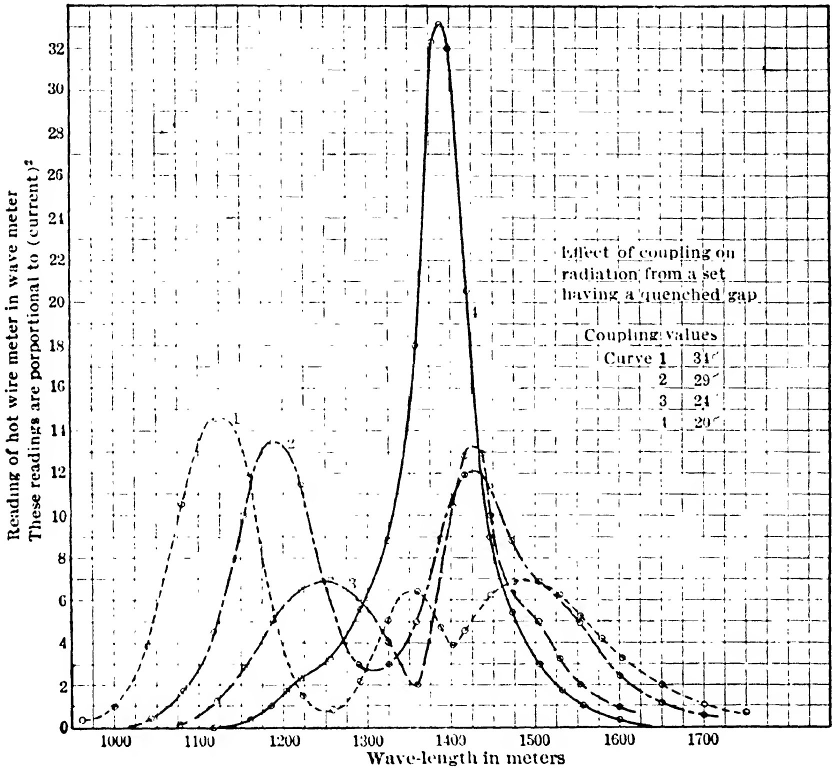


FIG. 19.—Resonance curves of a spark transmitter using a quenching gap; the gap would not properly quench if the coupling exceeded 20 per cent. For curves 1, 2, and 3 partial quenching is indicated by the presence of three "humps" on the resonance curve.

1. With the antenna circuit open, the closed oscillating circuit is adjusted to the wave length desired, by varying the value of inductance L_1.[1] The primary capacity is usually fixed in value and is not readily

[1] In the discussion L_1 stands for the total inductance in the closed oscillating circuit, i.e., the sum of the inductances of L_1 and L'_1 of the diagram; as previously noted, the extra inductance in the closed oscillating circuit, L'_1, is very seldom used.

changed, whereas the inductance L_1, forming also the primary of the oscillation transformer, is always of the variable type. The wave length at which the circuit will oscillate may be marked on the inductance L_1, different values of L_1 corresponding to different wave lengths, since C_1 is fixed and L_1 and C_1 being given in micro-units.

$$\lambda_{\text{meters}} = 1885\sqrt{L_1C_1}.$$

This calibration is usually made by the manufacturer before the set is delivered. In certain emergency or special conditions, however, this may not have been done, in which case a wave meter is loosely coupled to L_1, and L_1 adjusted, until the wave meter indicates a maximum deflection for the wave length, at which the set is to transmit.

2. After the closed circuit has been adjusted to the desired wave length, the antenna circuit is closed and *loosely* coupled to the closed circuit. The antenna inductance L_2 (or L_3 if in circuit) is then varied until maximum current is indicated on the antenna ammeter, under which condition the two circuits are in resonance. This adjustment may be checked, by coupling the wavemeter loosely to the loading coil, if in circuit, and noting the wave length at which maximum deflection of the wave meter ammeter is obtained. This should be the same as the wave length for which the closed circuit was adjusted.

It is important to note that the wave meter should not be coupled to the oscillation transformer secondary when making this check, but to some coil remote from L_2. If no loading coil is used, a small search coil, consisting of a turn or two of wire, should be inserted in the circuit, remote from the oscillation transformer, and the wave meter coupled to this coil. This procedure is required because of the relations of the flux, which surrounds both windings of the oscillation transformer, when both windings are carrying current, and under which condition, double-frequency current flows in each circuit. It was shown (see p. 313) that the lower-frequency currents in each winding are practically in phase, while the higher-frequency currents are practically 180° out of phase. The flux relations of the oscillation transformer, assuming a flat spiral construction, are thus as illustrated in Fig. 20.

It is apparent that a wave meter placed between the two coils, as indicated in Fig. 20, will indicate reasonance at the higher-frequency value, while if placed in the axial position will show resonance at the lower-frequency value. Intermediate positions will result in a combination of effects of the two fluxes, and the indications would therefore be inaccurate and confusing. It is thus always advisable to couple the wave meter to a single remote coil in the antenna circuit. The disturbing effects of the oscillation transformer fluxes in the wave-meter indication, exist to some extent even with loose coupling and both circuits correctly tuned. However, a wave meter coupled to the loading coil or search coil would give true indications under any condition. When the antenna is carrying much current, no search coil at all is required; if the coil of the wave meter is placed near the earth lead sufficient coupling will be obtained.

3. The set, after adjustment of the closed and open circuits as outlined above, is in condition for sending at the given wave length. The coupling should then be adjusted so that the energy radiated at this wave length will be a maximum. This will not be at the highest value of coupling obtainable, nor will the antenna current be a maximum necessarily for this condition. Maximum antenna current indicates a maximum energy radiation, but the distribution of this energy, as discussed on p. 425, is of more importance than the total radiation, and if the coupling is

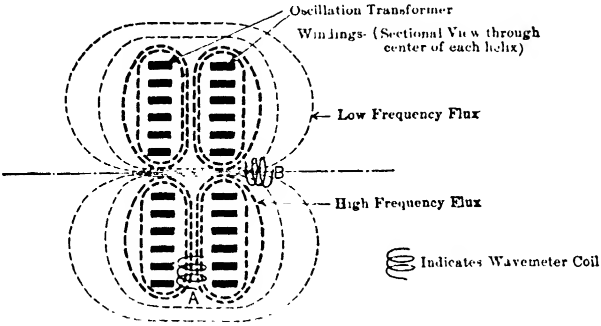


FIG. 20.—Cross-section through an oscillation transformer showing the distribution of flux due to the high- and low-frequency currents. With wave-meter coil on the axis low-frequency reasonance is obtained, whereas with coil between the two parts of the transformer high-frequency resonance is obtained.

too close, the radiated energy may be distributed over a large range of wave lengths. Thus the efficiency of the set, as measured by the energy reaching the receiving station, which is tuned to the wave length for which the transmitter is adjusted, may be very much reduced. This is a very important point and one often overlooked; namely that the criterion for best operation is *maximum energy radiation at the wave length for which the set is adjusted and not maximum antenna current.*

Characteristics of the Spark Transmitter — Energy Distribution Curves.—The energy distribution curves of a transmitter under different coupling values are shown in Fig. 18, the form of these curves evidently being determined by the coupling used, which in turn is governed by the following factors:

First.—The total amount of energy to be radiated. If the receiving station is at a considerable distance, more energy will be required and vice versa. This, however, is a minor factor, as the energy control is primarily obtained by spark gap adjustment.

Second.—The desired distribution of the energy radiated over the different wave lengths as illustrated by the curves (Fig. 18).

Under certain conditions, as, for instance, the sending out of distress signals, etc., a broad distribution of the energy radiated is of prime importance and close coupling would be used. A large number of stations, all of which may be tuned to different wave lengths, would thus be reached. This condition is shown by curve No. 4.

Under normal operating conditions, however, the distribution of the radiated energy is of greatest importance, and the coupling is adjusted so as to cause a minimum of interference with other stations, within range, for whom the message is not intended. Under this condition the maximum energy is radiated at the wave length for which the receiving set is tuned (as indicated by curve 8) and thus a maximum strength of signal would be obtained at the receiving station, although the coupling used is considerably less than that used in curve 4.

Adjustment of Power Input to the Transmitter.—The above description has considered the adjustment of the set for desired power output conditions. Power input adjustments will now be considered.

Since the high-frequency power input is equal to $\frac{1}{2}C_1E^2N$, this power may be controlled by varying the quantities C_1, E, and N.

Normally, the group frequency (N) will not be varied, as the operating efficiency of the set will probably be considerably decreased for speeds other than noted. Also, as mentioned previously, the characteristics of the phones at the receiving station are usually such as to make them most sensitive to a group frequency of about 1000 cycles per second, and it is therefore undesirable to deviate from this value to any considerable extent.

In practical installations, the closed-circuit capacity (C_1) is usually fixed in value, and could not be varied to secure a change in the power input.

The voltage to which C_1 is charged, E, is readily controlled, however, by adjusting the separation of the spark gap in the proper manner. This, therefore, forms the means whereby the power input may be controlled, and although limited in range, as discussed below, is widely used in practice.

In case a quenched gap is used the power input is controlled by using the proper number of gaps in series, many for high power and perhaps only one or two for short-range sending. It must be remembered that as the gap length is changed, or the number of sections of a quenched gap varied, the voltage of the alternator must be correspondingly altered to prevent arcing and irregular discharges.

Capacity and Inductance of the Closed and Open Circuits.—The proper capacity to be connected into the closed circuit will depend on the amount of the high-frequency power which it is intended to generate, the charging voltage to be employed, and the group frequency, since

$$W_{\text{watts}} = \tfrac{1}{2} C_1 E^2 N,$$

when $C_1 =$ closed circuit capacity in farads;
$E =$ potential in volts to which condenser is charged;
$N =$ group frequency in wave trains per second.

Thus, if we assume

$W = 2\tfrac{1}{2}$ kw. $= 2500$ watts;
$E = 15{,}000$ volts;
$N = 1000$ (alternator frequency $= 500$)

we have,

$$2500 = \tfrac{1}{2} \times C_1 (1.5 \times 10^4)^2 \times 1000$$

$$C_1 = 0.022 \text{ microfarad.}$$

The closed-circuit inductance, which also acts as the primary of the oscillation transformer, will be determined by the above value of C_1 and the maximum wave length at which the set is expected to radiate, Since:

$$\lambda_{\text{meters}} = 1885\sqrt{L_1 C_1}.$$

where $L_1 =$ closed circuit inductance in microhenries;
$C_1 =$ closed circuit capacity in microfarads,

we have, assuming $\lambda_{\max} = 1000$ meters

$$1000 = 1885\sqrt{L_1 \times 0.022}$$

$$L_1 = 12.8 \text{ microhenries.}$$

For proper operation the open (antenna) circuit constants L_2 and C_2 must satisfy the relation

$$L_1 C_1 = L_2 C_2,$$

where L_1 and C_1 are the same as indicated above;
$C_2 =$ total effective capacity of antenna circuit;
$L_2 =$ total effective inductance of antenna circuit.

Usually, C_2 will be considerably less than C_1 due to the difficulty and expense of building large-capacity antennas, and thus L_2 usually exceeds L_1 in value. If we assume an antenna capacity of 0.0024 microfarad,[1]

[1] Represents approximately an "L" antenna, length of top = 200 ft., height = 98 ft., number of wires = 6.

then

$$L_2 = \frac{0.022}{0.0024} \times 12.8 = 117 \text{ microhenries.}$$

All of this inductance would not be contained in the secondary winding of the oscillation transformer; a large part of it would be supplied by the loading inductance, while a relatively small portion would be found in the antenna itself. In the antenna referred to above, the inductance would be, perhaps, 40 microhenries.

Thus, the closed and open circuits are tuned to the same wave length, and if the coupling between them has been properly adjusted (see above), a maximum amount of energy will be radiated at 1000 meters and the efficiency of operation will be a maximum for the given conditions. The procedure to be followed in adjusting for a different wave length or changing the energy radiated by the set has already been described.

Elements of the Receiving Station.—The general connections and action of a receiving set have already been discussed. (See p. 398.) Primarily, it constitutes a circuit which absorbs a portion of the electrostatic and electromagnetic energy which reaches it from the transmitter, combined with certain devices to make this absorbed energy produce maximum visible or audible effects, so that its reception may be evidenced and intelligence thus transmitted.

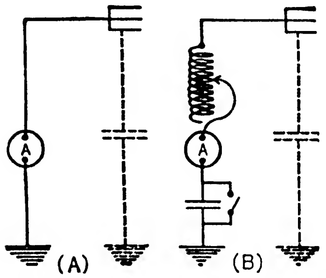


FIG. 21.—Simple schemes for receiving radio signals.

The antenna circuit represents the energy-absorbing element of the receiving set. The waves of electromagnetic and electrostatic energy sent out by the transmitter induce an e.m.f. in the antenna circuit, the natural frequency of which is the same as that of the transmitter. If the antenna were connected directly to ground as indicated in Fig. 21A, the circuit would be complete and a current would be caused to flow as long as energy is radiated by the transmitter. If a sensitive ammeter (thermogalvanometer) were inserted as shown, then this energy reception would be made visible, and thus a message might be transmitted between the two stations, if the radiated energy is interrupted in accordance with a prearranged code.

This arrangement represents the simplest possible form of receiver, but is never used, except for experimental purposes, owing to the impracticability of the sensitive ammeter which is required. The antenna would probably not be tuned to the frequency of the e.m.f. induced in it, and the resultant current would be extremely small, requiring a very sensitive instrument for its detection.

It is obvious that this current could be materially increased by so adjusting this circuit that its natural frequency is made the same as the frequency of the e.m.f. induced in it, i.e., the radio frequency of the received signal. This would be most easily accomplished by inserting a variable inductance in the circuit, provided the natural frequency of the antenna is above that of the incoming energy (a variable condenser would be inserted if the received energy has a frequency above that of the antenna). However, the current is very small even with the circuit adjusted to resonance, and the ammeter would of necessity be of a very delicate construction, making it impractical to use. Such an instrument would possess considerable resistance, which would still further limit the current flowing when the circuit is adjusted to resonance. In addition, its indications are inherently sluggish, and would require such a slow speed in sending, as to make its application for receiving purposes completely impractical. The addition of a variable inductance or capacity for tuning the antenna circuit is indicated in Fig. 21*B*.

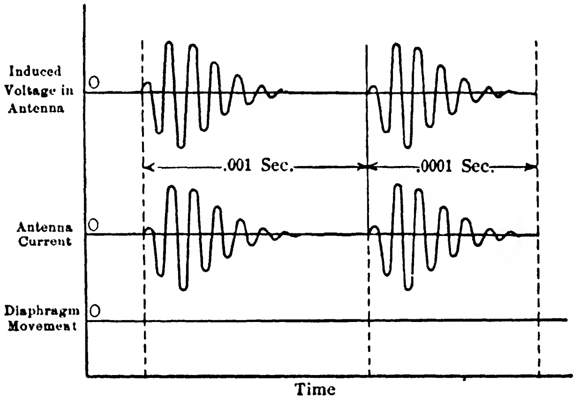


FIG. 22.—Conventional diagram of current in antenna of receiving station; even if such high-frequency currents could flow through the telephone the diaphragm could not move so rapidly.

Audible Detection.—In place of the above scheme of detection, which may be termed the visual method, a detector is used which causes the incoming energy to produce audible effects. Thus we might substitute a telephone receiver in the antenna circuit in place of the ammeter. The receiver is described in detail below. (See p. 437.) Briefly, it consists of a soft iron diaphragm actuated by current flowing through a winding placed on a permanent " U " magnet, the poles of which are placed closely adjacent to the diaphragm. An alternation of current of a certain polarity thus increases the pull on the diaphragm, whereas the reverse alternation will decrease the pull. The diaphragm is thus moved inward and outward, setting up vibrations in the air which are heard as sound by the observer.

The placing of such a receiver in the antenna circuit would, however, produce no sound in the phones even though high-frequency current were flowing in the antenna. This is due to the fact that the mechanical inertia of the diaphragm will not permit it to follow the extremely rapid reversals of the radio-frequency current. This reversal would occur at the rate of 1,000,000 times per second for a 300-meter wave signal. Also, even though it were possible for the diaphragm to respond to this current, no sound would be heard, as the frequency would be far above the limit of audible frequencies (about 20,000 cycles for young people and 10,000–12,000 for adults). The conditions are indicated in the above curves (Fig. 22).

The receiver would also add thousands of ohms impedance to the antenna circuit, so that only negligible high-frequency current could flow.

Application of the Rectifier.[1]—If we place in the antenna circuit, in addition to the phones, some device possessing unilateral conductivity, that is, a greater resistance to current flow in one direction than in the other, we would obtain a net or cumulative effect for each wave train, since the effect on the diaphragm in the one direction would then exceed the effect in the reverse direction. Thus the diaphragm would be given a resultant deflection, springing back to its initial position only after the wave train had passed. Thus, if 1000 wave trains strike the antenna per second (group frequency of transmitter = 1000), the diaphragm would be impulsed 1000 times per second, and the observer would hear a 1000-cycle note whenever the transmitter radiated energy. These conditions are graphically shown in Fig. 23. The amplitude variation of the e.m.f. for this figure should be carefully noted. The first cycle of a spark telegraph signal is of maximum amplitude in only one circuit of both sending and receiving stations, namely, the closed circuit of the transmitter. In all other circuits, time is required to build up the oscillations to their full amplitude, due to the electrical storage of energy which takes place during this period, just as in setting a mechanical system into oscillation, maximum amplitude is not obtained on the first impulse. (Unless the system starts with original distortion, as e.g., a pendulum held to one side and then released, which condition corresponds to that existing in the closed circuit of the transmitting set.)

The complete receiving circuit with the asymmetrical resistance (commonly known as a "detector"), is indicated in Fig. 24. The term "detector" is not strictly applicable, for it does not detect, but enables the receivers to detect, or make evident to the senses, the energy that is supplied to the telephone receivers. Owing to the high resistance of the phones, and the asymmetrical character of the rectifier resistance, this circuit is not selective, and it would be difficult to receive except under the unusual condition that only energy from the sending station desired

[1] For complete discussion of rectifiers the student should consult "Electric Rectifiers and Valves" by Güntherschulze and De Bruyne, John Wiley & Sons, Inc.

is reaching the antenna. Also the magnitude of the current flowing even under the best conditions is very small. For these reasons it is desirable and advantageous to connect the detecting apparatus in a separate circuit coupled inductively to the antenna by means of coils L_1, L_2, Fig. 25, the two forming what is known as the receiving coupler.

Inductively Coupled Receiver.—With this connection, the primary or antenna circuit may be tuned accurately to the frequency of the incoming energy, and since all high resistances have been removed, the antenna current will attain a maximum value very much greater than possible with the preceding arrangements. Therefore the e.m.f. induced in the secondary and the resulting current flow will be maximum and the signal strongest, although it will still be relatively weak on account of the high

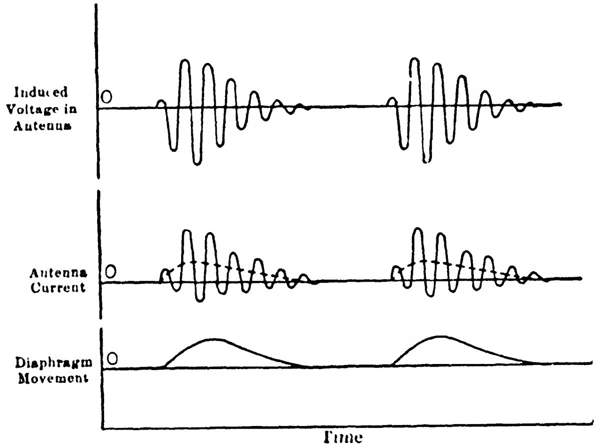


FIG. 23.—When a rectifier is used the antenna current is asymmetrical; more flowing in one direction than in the other; such a current will give the telephone diaphragm one impulse per wave train.

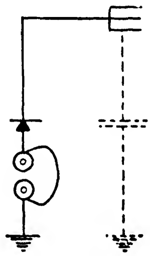


FIG. 24.—A possible scheme for using telephone and rectifying crystal.

resistances in the second circuit, which diminishes the resultant current. This combination of circuits possesses some selectivity due to the adjustment of natural frequency possible in the low-resistance antenna circuit. To enable the circuit to be tuned over wide ranges of wave length, an additional inductance L', known as a "loading" inductance, is inserted as shown for very long wave lengths, while the condenser C_1 may be cut into circuit if it becomes necessary to tune for very short wave lengths. This condenser is therefore known as a "shortening" or "short-wave" condenser.

To increase the selectivity and sensitivity of the set, a tuning condenser C_2 is placed across the coil L_2, giving the final circuit illustrated in

Fig. 26, which represents the arrangement most generally used in very simple receiving sets.

Neglecting for the moment the detector and phones connected in shunt across the condenser, it is readily seen that we may tune the secondary circuit to resonance with the primary circuit, and thus secure a maximum current flow in the circuit L_2C_2. L_2 is generally used for

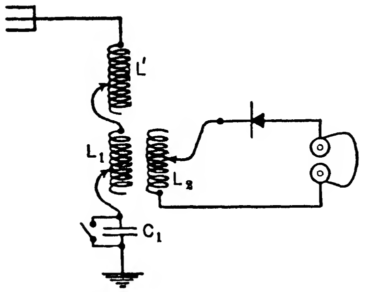


FIG. 25.—A receiving scheme using two circuits, the secondary being untuned.

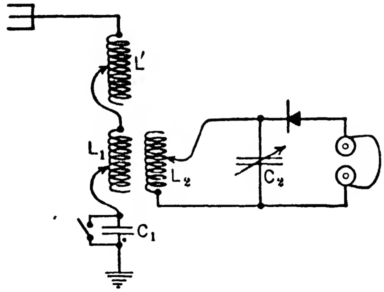


FIG. 26.—The ordinary receiving circuit using two tuned circuits, telephone in series with rectifier being shunted across the tuning condenser of the secondary circuit.

rough tuning, while a finer adjustment may be secured by adjusting C_2. Since the e.m.f. set up in the circuit L_2C_2 increases with the number of turns in L_2, it is desirable to have this inductance as high as possible without involving so much internal capacity action as to diminish seriously the selectivity of the set. Similarly, the condenser required for any wave-length adjustment should be relatively small. Under these

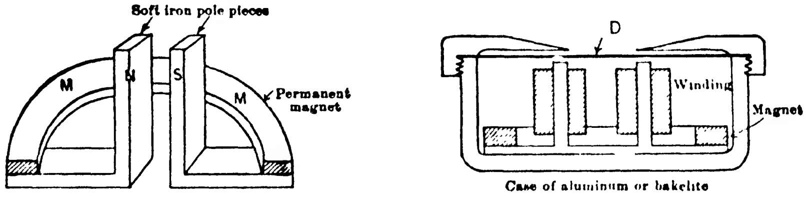


FIG. 27.—Essential elements of the ordinary watch-case telephone receiver.

conditions the radio-frequency voltage across the terminals of L_2 and C_2 (for a given amount of received energy) will be a maximum, and this radio-frequency voltage will, in turn, cause a maximum unsymmetrical current to flow in the detector-telephone circuit, and therefore maximum signal strength will be obtained. The action of the phones and detector in this circuit is exactly similar to their action illustrated in Fig. 23.

If the resistances involved in the primary and secondary circuits of the set are small, then this receiving circuit possesses considerable selectivity. Undesired signals may be tuned out and the efficiency and operating characteristics of the set are very much better than those of the previous circuits discussed.

The Telephone Receiver.—The construction of the telephone receiver usually employed for the reception of radio signals, known as the " watch-case " type, is shown in the accompanying sketch. (Fig. 27.)

It consists essentially, of a permanent magnet M, with pole pieces N and S, upon which are wound coils consisting of many turns of fine wire, through which the audio-frequency pulses of current pass. A diaphragm D is placed closely adjacent to the faces of the pole pieces as shown. When no signal is being received, this diaphragm is under a constant pull or attraction exerted on it by the permanent magnet M. This steady pull, which we may call P, is proportional to the square of the flux flowing through it and the permanent magnet, i.e.,

$$P=K\phi_s^2, \quad \ldots \ldots \ldots \quad (6)$$

where ϕ_s is the steady flux value.

When the current of proper polarity flows through the winding, the flux will be increased proportionately (neglecting saturation effects) or

$$\Delta\phi=K'i, \quad \ldots \ldots \ldots \quad (7)$$

wherein i represents current in the winding.

Therefore the total flux is:

$$\begin{aligned}\phi_t&=\phi_s+\Delta\phi\\&=\phi_s+K'i, \quad \ldots \ldots \ldots \quad (8)\end{aligned}$$

and the total pull under this condition becomes

$$\begin{aligned}P&=K\phi_t^2=K(\phi_s+K'i)^2\\&=K\phi_s^2+2KK'\phi_si+KK'^2i^2. \quad \ldots \quad (9)\end{aligned}$$

The total pull thus consists of three components, one of which is constant ($K\phi_s^2$) and thus has no effect on the diaphragm vibrational amplitude, while another is proportional to the current variation squared (KK'^2i^2). This term represents a distortional component of double frequency and it is therefore designedly made relatively small. The remaining component of pull is proportional to the current variation ($2KK'\phi_si$) and this component is therefore made as large as possible; the amplitude of the diaphragm vibrations will then be proportional to the amplitude of the current variation (i) and it will also be directly proportional to the flux due to the permanent magnet (ϕ_s). Thus, to make the vibrations of the diaphragm a maximum for a given current variation, $K\phi_s$ is designedly made large compared to $KK'i$, which means that the flux (ϕ_s) produced by the permanent magnet is much greater than the flux produced by the current in the winding ($\Delta\phi$). Under these conditions, distortional effects are minimized and maximum amplitude of diaphragm vibration and signal strength (sound) for a given signal current (i) secured.

It is, of course, very important that the deflections of the diaphragm follow accurately the value of the force ($2KK'\phi_s i$) as, otherwise, the sound waves set up by the diaphragm movement will be distorted. This is an extremely important feature in the design of loud-speakers, where the difficulties are even greater because of the larger displacements required.

The d.c. resistance of a receiver such as described above would be about 2000 ohms, as many as 10,000 or more turns of fine wire (about No. 40 A. W. G. or smaller) being employed to make up the winding. The impedance to an alternating current will, of course, be greater than this, depending on the frequency of the current and the effective resistance of the circuit. At 400 cycles a certain receiver of this type had an impedance of 2900 ohms; at 800 cycles an impedance of 3900 ohms; and at 1000 cycles an impedance of 4400 ohms.

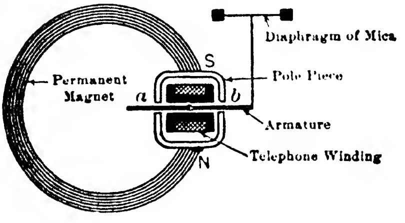


FIG. 28.—Essential elements of the Baldwin balanced-armature telephone receiver.

A well-built receiver of 11,000 turns for currents of the order of 1 microampere shows a resistance (a.c.) of about 3000 ohms and a reactance of 10,000 ohms. These values hold good for currents from 10^{-9} ampere to 10^{-5} ampere and increase somewhat for currents greater than this. In a quiet room this headphone gives an audible signal (at 1000 cycles) for a current somewhat less than 10^{-9} ampere. This means an input of 10^{-15} watt.

The Balanced Armature Receiver.—Another type of receiver more recently developed, known sometimes as the Baldwin receiver, possesses the advantage that the diaphragm is not initially stressed, and thus may be more responsive and sensitive to the pull exerted on it by the flux ($\Delta\phi$) caused by the signal current (i). The construction is indicated below (Fig. 28).

It is evident that, when no signal is being received, the armature being balanced in its neutral position (the flux traversing the gaps a and b in the same direction and being equal in value), no pull is exerted on the mica diaphragm. If a signal pulsation of current passes through the receiver winding, however, it produces a flux, which, combining with the permanent flux results in an asymmetrical distribution of the flux, causing a force to be exerted on the armature and thus on the diaphragm.[1]

Other advantages claimed for this type of receiver in addition to the one mentioned above, are:

(1) The magnetic circuit is of low reluctance and thus small signal currents will produce relatively greater fluxes and greater forces.

(2) The armature is similar in its mounting to a lever, with a force acting at each end. The diaphragm may be attached to the oscillating armature at any desired point near the outer end, as shown in Fig. 28, or near the center. When the diaphragm is large the latter arrangement is generally best.

It is to be noted that this device is not truly balanced (when no signal is being received) in the case of detectors where the initial current is not zero, as in the vacuum tube, and crystal equipped with polarizing battery. The pull due to this current, however, is extremely light, compared to

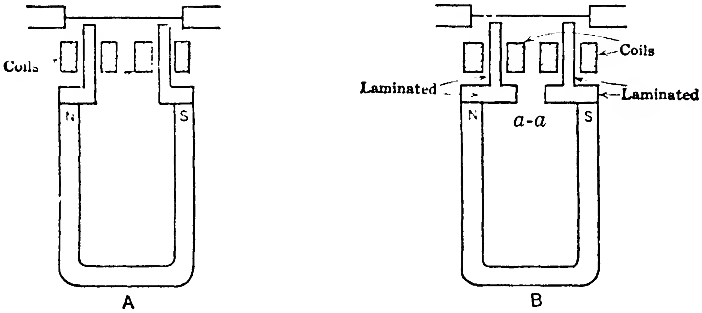


FIG. 29.—The magnetic circuit of an improved telephone receiver; the pole pieces are of laminated iron.

the heavy pull exerted by the permanent magnet in the usual type of construction, and the diaphragm may be considered as essentially unstressed.

The Seibt Receiver.—The sensitiveness of receivers of the electromagnetic type may be improved by decreasing the reluctance of the path which the a.c. flux will traverse, and also by laminating the material used in that path, thereby reducing eddy currents and counter m.m.f.s. This is accomplished in the Seibt[2] phone as shown in Fig. 29*B*. Fig. 29*A* shows the usual design for comparison.

The reluctance of the a.c. flux path is much reduced by the extensions *a-a* on the pole pieces. These may not be extended too far, however, as the permanent flux would then be by-passed and not traverse the

[1] The student is referred to E. I. Bucher, "Practical Wireless Telegraphy," p. 168, for more detailed description.

[2] "Telephone Receiver of Increased Sensitivity," G. Seibt, E.T.Z., March 2, 1922.

diaphragm. This construction, it is stated, more than doubles the diaphragm vibration for a given strength of signal. It also acts to make the magnet more permanent.

In consideration of the above, it is obvious that no closed conducting paths should be linked with the path of the a.c. flux. Thus metallic coil spools or collars would act as closed circuit secondary windings, setting up a counter m.m.f. to the primary m.m.f. and diminishing the efficiency of the telephones.

By the addition of auxiliary equipment, the telephone receiver may be converted to a visio instead of an audio detector.[1] Thus the diaphragm of a sensitive telephone receiver, carrying a few microamperes of 1000 cycles alternating current, may be optically magnified by means of a mechanical device, and so measured and recorded on a moving film. The visible indications offer a decided advantage in certain measurements, viz., comparing the relative strength of signals.

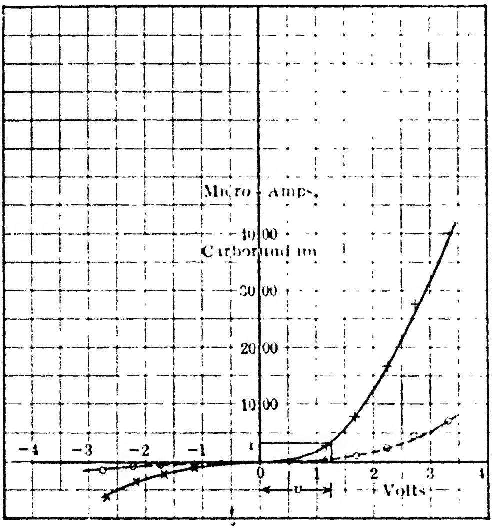


FIG. 30.—Characteristic curves of a carborundum rectifying crystal, using a fine steel point for making contact. The maximum rectifying action occurs when a polarizing voltage of about 1.2 volts is used.

Thermophone.—By utilizing a number of very fine wires, 0.001 to 0.0015 mm. in diameter, through which the receiver current is caused to flow, the temperature of the wires will follow almost instantaneously the variations in receiver current. This temperature, in turn, causes the motion of a small column of air which is heated by the wires, and the motion of this air column sets up sound waves reproducing the received speech.

This device, which has had only laboratory application as yet, is stated to have no inherent distorting characteristics, i.e., the sound waves truly follow the receiver current variations. This is said to be due to the absence of the diaphragm required on other types. The column of air used in the device possesses a natural frequency of 8000 to 10,000 cycles per second, which is well beyond the normal range of speech vibrations.

[1] "Photographic Recording and Measurement of Radio-Telegraph Signals," King, Royal Society, Canada, Trans., p. 143, 1921.

Characteristics of Crystal Rectifiers.—The unilateral conductivity possessed by various crystals is shown by the following curves (Figs. 30 to 33, inclusive). These curves indicate the relatively large currents obtained when e.m.f. of various values and of a given polarity are impressed across the rectifier circuit and the comparatively small (sometimes negligible) currents obtained when the e.m.f.s are reversed. These curves represent the "d.c. characteristics" of the crystals in contradistinction to the "a.c. characteristics" discussed below, and are obtained by means of the experimental circuit indicated in Fig. 34 (Insert A).

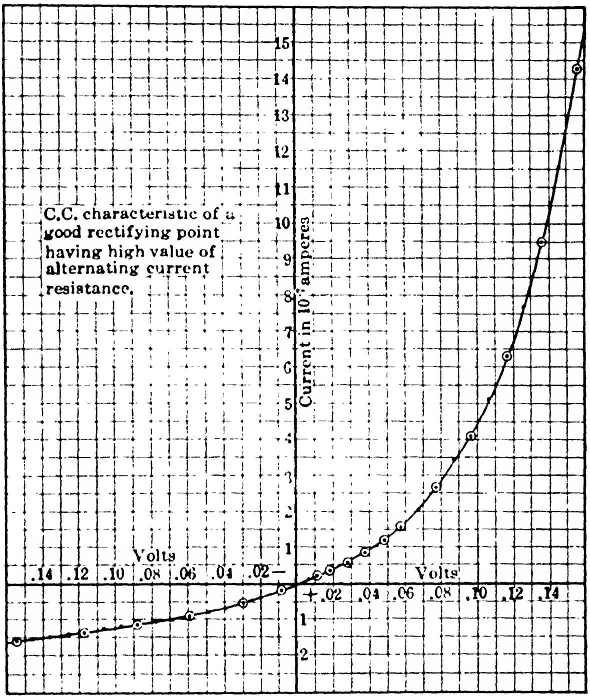


Fig. 31.—Characteristic of a good rectifying point on a galena crystal; the a.c. resistance is high, a desirable characteristic.

Fig. 30 illustrates the characteristics obtained for a carborundum (silicon carbide) crystal. The curve is interesting as it illustrates the function of the local battery sometimes used in series with the detector and phones, and known as a "polarizing" battery. The connection of this battery in the detector circuit is illustrated below (Fig. 33).

It is evident that with any detector the greatest asymmetrical effect (and thus maximum signal strength) will be obtained, if we adjust the crystal to operate at the point of maximum curvature. In the case of the curve considered, this does not occur at the zero voltage value but at +1.25 volts approximately. Therefore, the local battery potentiometer would be adjusted to impress an initial voltage of +1.25 volts on the crystal. Under these conditions the signal voltage impressed on the potentiometer, detector, and phones in series, would vary the current above and below the initial value (indicated by *i* in Fig. 30) and a maximum asymmetrical current would thus be secured.

The "d.c. characteristic" is also indicated on the figure for a poor rectifying point (dotted curve), which indicates that with this adjustment, the asymmetrical effect of the crystal is practically *nil*, and the resistance of a uniformly high value.

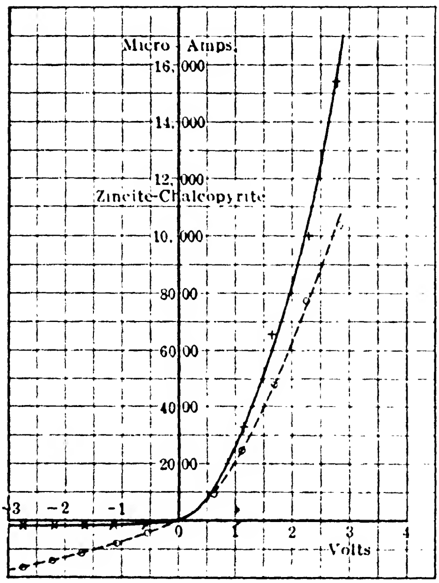


FIG. 32.—Characteristic curves of a "Perikon" rectifier, utilizing the contact between zincite and chalcopyrite.

Fig. 31 shows the characteristic curve of a carefully selected point on a galena crystal. This point indicates very good rectification for very low voltages; it is well known that galena is a good detector for weak signals. Not only does it rectify well, but the point has a high a.c. resistance, so that the selectivity of the receiving circuit is not spoiled when the crystal is connected in shunt with the tuning condenser. The a.c. resistance of a crystal contact depends upon the signal strength, being lower for more intense signals. For weak signals this galena crystal shows a resistance of about 500,000 ohms.

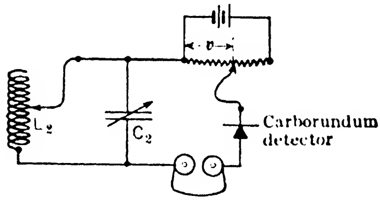


FIG. 33.—Scheme of using such a crystal as carborundum in a receiving circuit. The best rectifying action is obtained by suitable adjustment of the potentiometer on the polarizing battery.

Fig. 32 illustrates characteristics obtained with a combination of two crystals, zincite (red oxide of zinc) and chalcopyrite (iron-copper sulphide), which differs from the preceding cases in which a sharp metallic point is placed in light contact with the crystal. The curves indicate that considerable rectification will be secured even if operating on a poor point. This would make the combination desirable for use where adjustments may be frequent, due to vibrations or similar disturbances, which

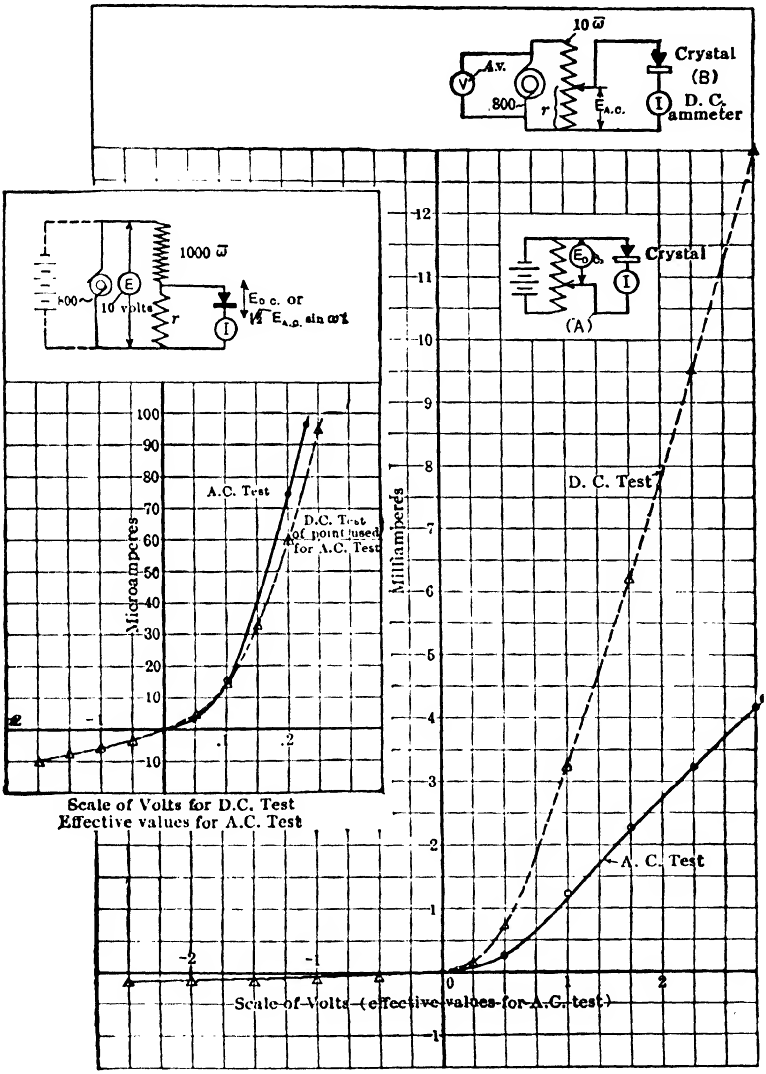


FIG. 34.—Comparison of d.c. and a.c. characteristics of a Perikon detector. The a.c. characteristic was obtained by measuring the current through the rectifier with d.c. ammeter, when an alternating e.m.f. was impressed. The insert shows the really important action, as it is seldom that more than a small fraction of a volt is set up across the rectifier when it is used in a receiving circuit.

may be present, as in the case of portable receiving equipment and stations on shipboard.

The asymmetrical effect of the detectors described above, when an alternating potential is impressed across the circuit in which they are connected is indicated by the curves (using a zincite-chalcopyrite detector) shown in Fig. 34. The experimental circuit used is shown in Fig. 34 (Insert B). Both the d.c. and a.c. characteristics are indicated, the former being plotted between the d.c. voltage and corresponding current as heretofore, while the latter is plotted between the effective a.c. voltage and the d.c. component of the rectified alternating current, as read by the same d.c. ammeter used in obtaining the "d.c. characteristic." The instantaneous values of the current flowing in the detector circuit when a sine wave e.m.f. is impressed have been plotted in Fig. 35 wherein the corresponding voltage values are also indicated. It should be noticed that the negative alternations of the current are practically negligible in amplitude, while the positive alternations are not of sine-wave form but considerably more peaked, owing to the variation in detector resistance, which decreases as the current increases.

This current may be graphically resolved into its d.c. and a.c. components as shown in the figure. The latter component will not affect the d.c. ammeter, the deflection of which is proportional to the magnitude of the d.c. component only, the value of which it indicates. Thus, for an effective a.c. voltage of 1.41 volts, Fig. 34 (maximum value equal to 2 volts as shown in Fig. 35), the reading of the ammeter is 2 milliamperes, which is the magnitude of the d.c. component as indicated in Fig. 35. Thus the curve obtained from the a.c. test indicates the d.c. component plotted to various corresponding a.c. voltages as indicated by the curve in Fig. 34.

The insert curve in Fig. 34 illustrates the a.c. and d.c. characteristics to a magnified scale in the region of the zero voltage point. It is interesting and important to observe that this d.c. characteristic indicates satisfactory rectification, for very small voltage values, such as would exist across the detector-phone circuit under normal conditions, although the more extended curve (main curve of Fig. 34) would seem to indicate that for small voltage variations, a polarizing e.m.f. of about +0.25 volt would be desirable. These data demonstrate the fact that if the characteristic curves are to be considered reliable, and truly indicative of what the rectifier will do in its application to radio signal reception, it is desirable to investigate them for low values of impressed voltages and not carry them out to voltage values as large as 2 volts, a magnitude practically never encountered in normal radio reception using head phones.

Desirable Characteristics of Crystal Rectifiers.—Crystal detectors or rectifiers should possess the following qualities:

1. They should be mechanically rugged and well constructed. This means that they should be able to hold their adjustment and not be easily disturbed. These are especially desirable characteristics in field or marine sets, where jarring or vibration is likely to be present.

2. The crystals should be sensitive, that is, should possess good rectifying properties, if their setting is properly adjusted. Too great a sensibility is not desirable as satisfactory adjustment is usually obtained with difficulty. Also it may be difficult to retain the sensitive adjustment.

3. The crystal should be easily adjusted. It is a distinct disadvantage if any marked difficulty is found in adjusting the setting for good reception,

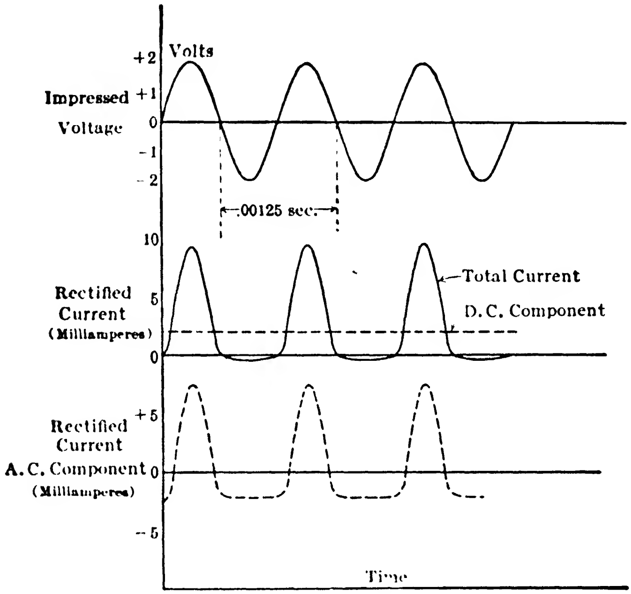


FIG. 35. Analysis of the current through a rectifying crystal.

as valuable time may be lost in this way if a signal is coming in and the detector not operating properly.

4. The crystal should possess self-protecting characteristics to prevent itself from being "burned out" and the setting destroyed, if abnormally powerful energy radiations are received, such as static and other atmospheric disturbances.

A complete explanation for the asymmetrical conductivity of two dissimilar crystals in contact, or a crystal in contact with a metal point has not yet been advanced. It may be due, in part, to the thermoelectric effects produced by the heating of the junction when the detector is carrying current.[1] The rectifying properties have also been considered

[1] W. H. Eccles, Proc. Phys. Soc., London, Vol. 25, p. 273, June, 1913.

as being due to electrolytic action, which occurs at the surface of the crystal.[1] A more likely explanation, however, is based on the "surface work" for electron evaporation from the two crystals; if the "surface works" are different, the contact point must offer asymmetrical resistance.[2]

Application of the Bridging Condenser.—It will be recalled that the telephone receiver possesses considerable impedance for a 1000-cycle alternating current. It might seem that it would be impossible to send a radio-frequency current through the phone circuit, and this would be the case were it not for the distributed capacity which is associated with the windings in the phone, and the phone cords. This capacity is in shunt with the telephone windings, and is represented by the fictitious condenser, C', in Fig. 36.

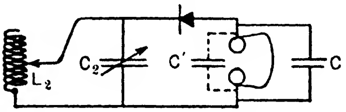


FIG. 36.—Use of a "bridging" condenser in parallel with receiver.

The current may then be considered to divide in the circuit shown in Fig. 37, the radio-frequency component passing through the distributed capacity, which has a relatively low impedance to high-frequency current, while the audio-frequency component flows through the phone windings.

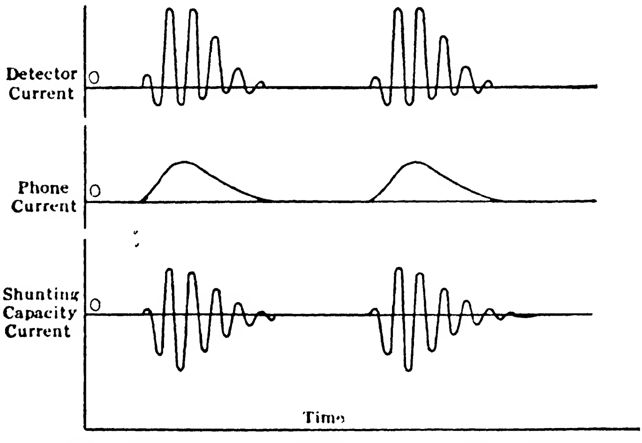


FIG. 37.—Currents in the branched circuit shown in Fig. 49.

Normally, however, the distributed capacity of the phone cords, etc., is only a few micro-microfarads in value and is not large enough to supply a sufficiently low impedance by-path for the high-frequency component. It is therefore usual to connect additional capacity across the phones,

[1] R. H. Goddard, Physical Review, Vol. 34, 1912.

[2] It is to be noted that any explanation must be able to take care of the fact that certain crystals rectify in one direction for low voltages, and in the opposite direction for higher voltages, not rectifying at all for some intermediate voltage.

as shown in Fig. 36, condenser C. This additional capacity is known as a bridging condenser, or by-pass condenser, and may have a value of approximately 5000 $\mu\mu f$. A low impedance path for the high-frequency component is thus supplied, thus permitting a larger fraction of the high-frequency signal voltage to act across the rectifying crystal and hence increasing its rectifying efficiency. This, in turn, increases the amplitude of the audio-frequency component flowing through the phones and the strength of the received signal is thereby increased.

The bridging condenser is sometimes considered as a capacity which receives a cumulative charge during the passage of a wave train, due to the asymmetrical conductivity of the rectifier, comparatively little current passing through the phones. When the wave train has passed this condenser discharges through the phones, since it cannot discharge back through the detector, due to high resistance in this circuit. This unidirectional discharge passes through the phone winding as a current pulse equivalent to the d.c. component previously described. Thus one click is produced in the phones per wave train, and the observer hears a note of audio-frequency pitch as previously described.

An additional effect of the bridging condenser is that it lowers the resonant frequency of the windings and associated capacity, which represent a multiple-resonance circuit. The average of forty pairs of phones (watch-case type) showed

$$L = 1.73 \text{ henries}$$

$$C = 10^{-10} \text{ farad (cord capacity).}$$

These constants give a resonant frequency of about 10^4 cycles per second. The addition of a 5000 $\mu\mu f$ bridging condenser lowers this to 1.5×10^3 or much nearer the value 10^3 cycles per second, which represents the usual audio frequency used in radio telegraphy. This latter feature is of no value in connection with radiophone reception; in fact it would constitute a disadvantage to have the circuit sharply selective to any one audio frequency, as distortion would result. The selectivity in this case is not sharp, however, because of the high resistance involved.

The above relations were indicated in a general way in Fig. 93, Chapter I. The curves there given indicate the receiver characteristics with no bridging condenser connected.

Vacuum-tube Detector.—The crystal rectifier has been practically superseded by the three-electrode vacuum tube because of the latter's greater sensitivity, reliability, and ease of adjustment. The action of the tube is discussed in detail in Chapter VI, to which the reader is referred for a full explanation of its rectifying action. One advantage of the tube over the crystal lies in the fact that its rectifying ability is measured by the accuracy of its design and construction, while that of the crystal is an inherent property of the substance, and cannot be altered or improved. With the latter, the "best" point of operation is determined experimentally and may or may not represent the best performance of which the crystal is capable. On the other hand, the tube may be accurately and definitely adjusted for best operation, which as already stated, exceeds that of the best crystal rectifiers.

Audibility and Selectivity.—The receiving circuit for a spark set has already been studied on p. 436, and it now remains to discuss the characteristics of such a circuit and the adjustments necessary to secure the best results. Before doing this, however, we must define two quantities, upon which the comparison of receiving systems is based, i.e.: "audibility" and "selectivity."

"Audibility" may be defined as the ratio of the audio current flowing through the telephone receivers to that which is necessary to make

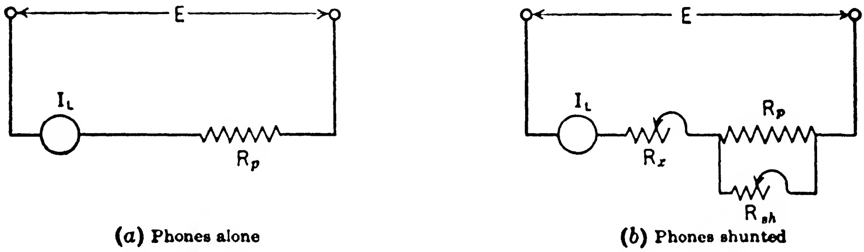


FIG. 38. Circuit for making audibility test.

the signals just audible. To speak of a receiving circuit having an audibility of, say, 20, means that the current in the receiving circuit is twenty times that which is just necessary to produce a just audible signal.

The relative strength of signal or "audibility" as it is defined above, is most conveniently determined by means of an audibility meter, which is merely a calibrated resistance placed in shunt with the telephone receiver. The signal is first listened to with no resistance shunted across the phones. Resistance is then connected in and decreased until the signal is just audible. The audibility is then,

$$\text{Aud.} = \frac{I_L}{I_{\min}} = 1 + \frac{R_p}{R_{sh}} = \frac{R_p + R_{sh}}{R_{sh}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where

R_p = phone resistance in ohms;
R_{sh} = shunt resistance in ohms for just audible signal;
I_L = current through phones without shunting resistance;
$I_{\min}$ = current through phones with shunting resistance connected.

The above expression assumes that the total current, and resistance in the detector circuit, remain constant as the shunt resistance is varied. We must thus insert a series resistance (R_x) as the shunt resistance (R_{sh}) decreases. The equivalent circuits are shown in Fig. 38.

Then

$$I_L = I_p + I_{sh} = I_p + \frac{I_p R_p}{R_{sh}}$$

or

$$\frac{I_L}{I_p} = 1 + \frac{R_p}{R_{sh}}$$

as given above.

For R total to remain constant, we must have:

$$R_x + \frac{R_{sh} R_p}{R_p + R_{sh}} = R_p$$

and

$$R_x R_p + R_x R_{sh} + R_{sh} R_p = R_p^2 + R_p R_{sh},$$

or

$$R_x = \frac{R_p^2}{R_p + R_{sh}}. \qquad \qquad (11)$$

Thus, for a given phone resistance, a definite relation exists between R_x and R_{sh}. An audibility meter is, therefore, designed for a certain phone resistance and only phones of this resistance should be used with it. Fig. 39 shows the internal connections of an audibility meter.

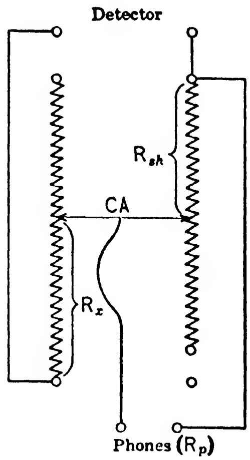


FIG. 39.—Arrangement of resistances in the audibility meter.

R_x and R_{sh} are tapped to a double row of dial buttons traversed by a single contact arm. Each position of the arm thus corresponds to a definite " audibility " value. The instrument is not very accurate, its setting depending largely on the observer's acuteness of hearing. This varies greatly with conditions such as room noise, condition of health, length of time operator has been listening, etc. Visual measurements, using sensitive thermo-galvanometers, possess greater accuracy but require more skill and time than is usually justified.

The audibility as defined above is directly proportional to the current in the receiving antenna and, for weak couplings, say less than 5 per cent, inversely to the coupling coefficient between the receiving antenna circuit and the receiving closed circuit. Again, for short distances the receiving antenna current may be shown to vary as follows:

$$I_r \propto \frac{I_s h_s h_r}{R\lambda d}, \qquad (12)$$

where I_r = receiving antenna current;
I_s = transmitting antenna current;
h_r and h_s = height of receiving and transmitting antenna, respectively;
R = effective resistance of the receiving antenna, including the resistance due to the closed circuit being coupled to it;
d = distance between the two antennas.

From the above it follows that, if the coupling between antenna and closed tuned circuit is very loose (generally the case in practice)

$$a \propto \frac{I_r}{k} \propto \frac{I_s h_s h_r}{R\lambda dk}, \qquad (13)$$

where
a = audibility;[1]
k = coupling coefficient between receiving antenna circuit and the receiving closed circuit.

"Selectivity" of a receiving system may be defined as the ratio of the natural wave length of the transmitting and receiving antenna circuits to the difference between this wave length and the length of some other wave which (of same field intensity as signal wave) will give a response just audible. Thus, if:
λ_n = natural wave length of the two antenna circuits;
λ_a = length of wave (of same field intensity as signal wave) which will give a just audible response in the telephone receivers.
S = selectivity.

Then:

$$S = \frac{\lambda_n}{\lambda_n - \lambda_a}. \qquad (14)$$

It will be seen that selectivity is a measure of how little the reception of signals from a certain transmitting station will be interfered with by the presence of electromagnetic waves of a different wave length emanating from other stations. Thus if $\lambda_a = \lambda_n$, a condition impossible to realize, then,

$$S = \frac{\lambda_n}{0} = \infty$$

[1] This formula is approximate only; actually the audibility does not vary inversely with R because the detector efficiency is involved in the magnitude of R. For a theoretical discussion of the best coupling for detectors of different types the reader is referred to Chapter XV of Pierce's "Electric Oscillations and Electric Waves."

or the selectivity is infinitely large, and no interference will be registered at the receiving station.[1]

It may be shown that when the transmitting and receiving systems are tuned, selectivity is affected by the sum of the decrements of the transmitting and receiving systems [2] and also by the audibility of the receiving

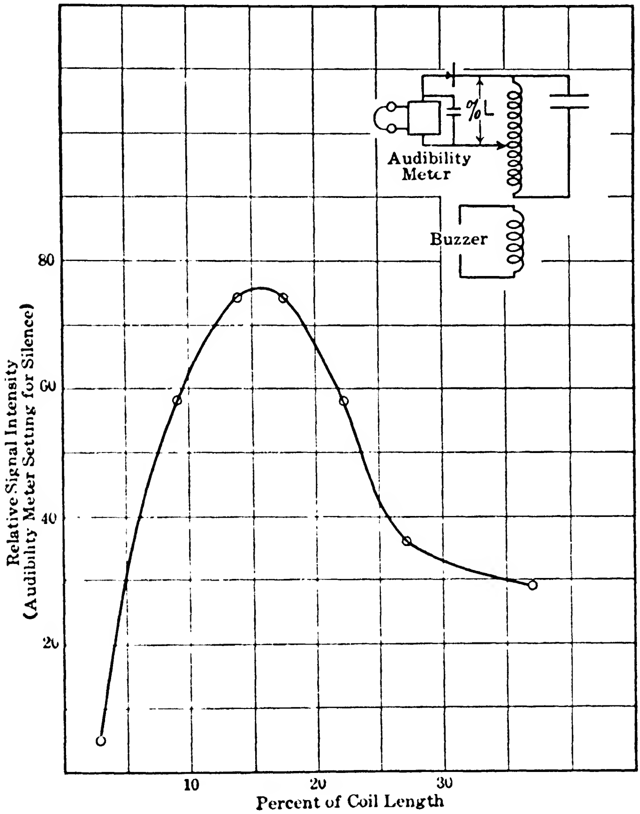


FIG. 40.—Showing that maximum signal may be obtained when connecting the rectifier-phone circuit across only part of the tuning inductance.

circuit (in so far as this is affected by k), approximately as shown by the following formula:

$$S = q \times \frac{1}{\delta_t + \delta_r} \times \frac{1}{a}, \quad . \; . \; . \; . \; . \; . \; . \; . \quad (15)$$

[1] For a comprehensive treatment of the selectivity of a receiving system, see "Behaviors of Radio Receiving Systems to Signals and Interference," by Peters, Journal A.I.E.E.

[2] See Chapter III, pp. 351–354.

where

q = a constant;

δ_t and δ_r = decrements of transmitting and receiving circuits, respectively;

a = audibility.

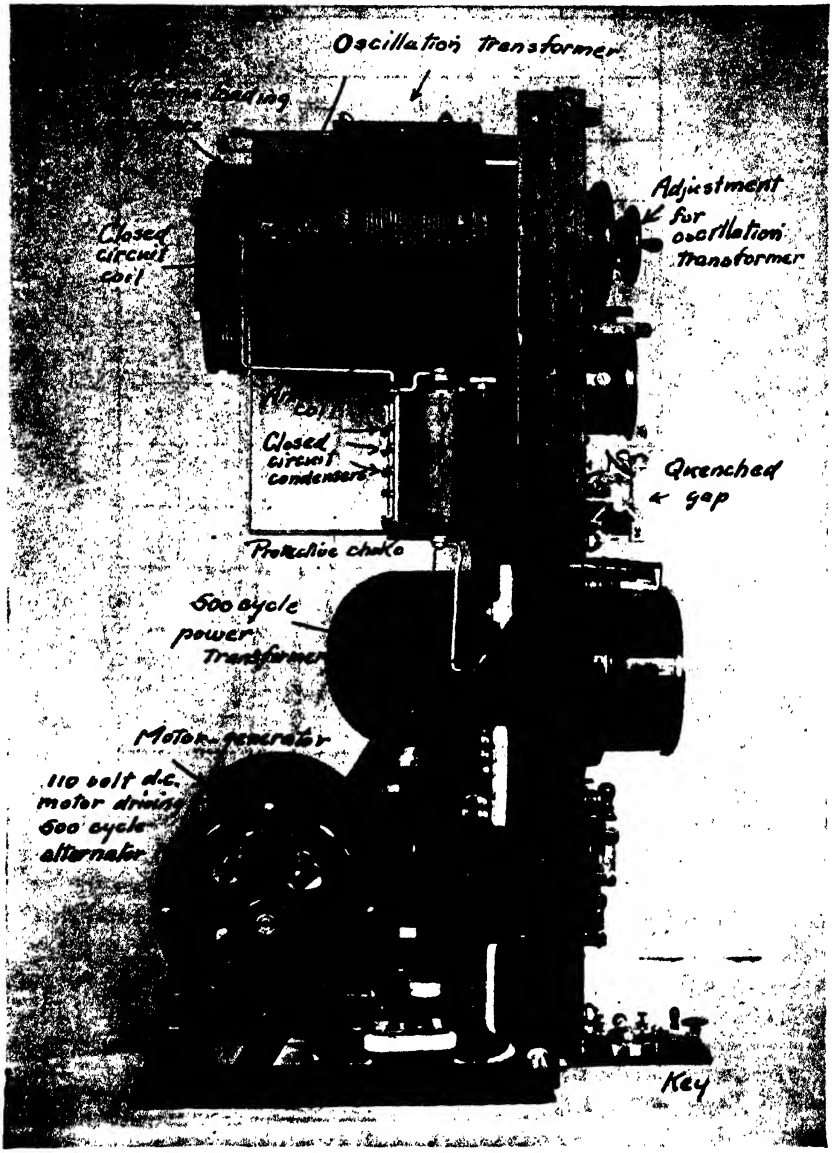


Fig. 41.—Side view of a small spark set.

Practically no selectivity can be obtained with the transmitting and receiving systems out of tune.

When making the adjustments of a receiving set the aim should be to obtain the maximum selectivity compatible with a reasonable audibility; but it must be borne in mind that these two quantities are inversely proportional to each other and that, in general, a high audibility means a low selectivity, and vice versa, as shown by the formula above.

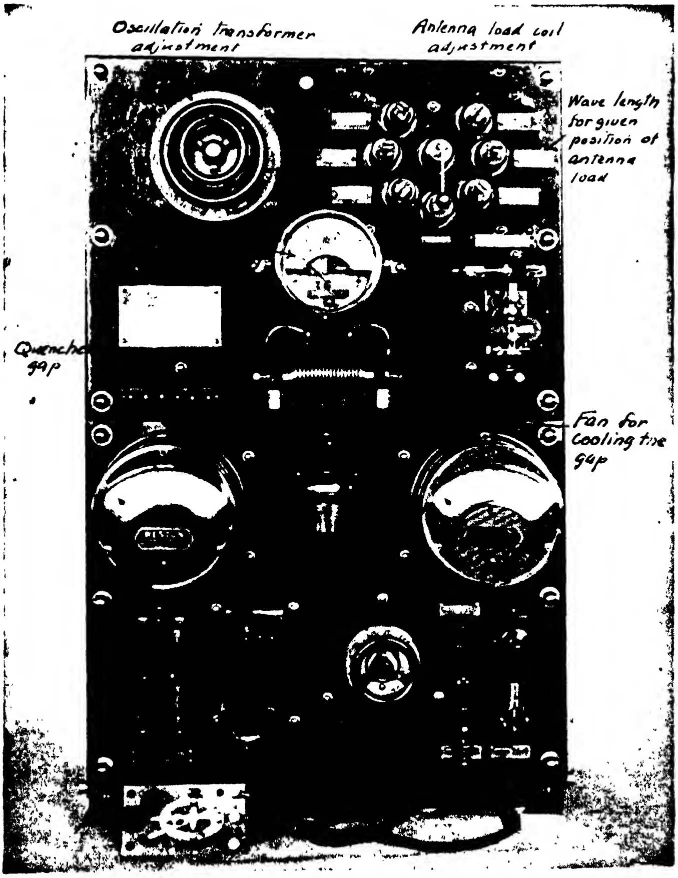


FIG. 42.—Front view of the set shown in Fig. 41.

Proper Connection of Rectifier.—In such a circuit as that of Fig. 26 it will frequently be found that the signal has greatest intensity when the rectifier is not connected across the entire tuning coil of the secondary circuit, but across only a piece of it. The resistance (a.c.) of the rectifier decreases as the impressed voltage is increased, resulting in comparatively poor selectivity if the crystal is connected across the whole coil.

The proper proportion of the coil to use for the rectifier connection depends upon its resistance, being greater as the resistance is greater. A typical illustration of this effect is shown in Fig. 40, in which case best signal was had when the rectifier was connected across only 15 per cent of the coil.

Arrangement of Apparatus of a Spark Set.—The various parts required for a spark transmitting set are generally assembled in compact form; in the case of a low-powered outfit practically all the apparatus, with the exception of the hand key for sending, may be mounted directly on a panel. A neat design for a 500-watt, quenched-spark transmitter is shown in Figs. 41 and 42; the legends on the cuts make them self-explanatory. The larger land stations of course require large switch boards and auxiliary apparatus; in fact the outfit really comprises a complete isolated power plant equipment.

CHAPTER VI

VACUUM TUBES AND THEIR OPERATION IN TYPICAL CIRCUITS

Constitution of a Conductor. Possibility of Electron Emission.—As outlined in Chapter I, a conductor is made of atoms (or molecules) with some of the electrons free from atoms, moving back and forth, from one atom to another. Unless the conductor is at absolute zero temperature its atoms are constantly in a state of agitation, having non-coordinated motions in all directions. The free electrons share the motion of the atoms, and due to their comparatively small mass (about 1/200,000 that of the tungsten atom) their average velocity is very much greater than that of the atoms.

Now the atoms of a metal tend to separate from each other at high temperatures or, we may say, the metal tends to evaporate just as water evaporates at ordinary temperatures. We must imagine the surface tension of a metal great enough to prevent appreciable evaporation at ordinary temperature; the velocity of motion of the atoms is not sufficient to carry them through this surface tension. With very high temperatures, however, those atoms having the highest velocity break through the surface tension and so start the process of vaporization, which becomes more and more rapid as the temperature rises. To accomplish actually the vaporization of the ordinary metal requires that the heating be done in vacuum, otherwise oxidation generally occurs instead. The number of atoms evaporated from a given surface depends upon the temperature and the latent heat of evaporation of the metal being tested.

Now it seems quite likely that if, when the atoms acquire a high velocity they are able to break through the surface tension of the metal the electrons can do the same thing, hence we get the idea of *electrons evaporating.*[1] This evaporation of the electrons will take place at lower temperature than that of the atoms of the metal itself because of the higher average velocity of the electrons.

It is possible to abstract electrons from metals by other means than heat, but not so effectively. Certain of the alkali metals give off electrons when light strikes them. The surface must be chemically clean, and such a surface is generally obtained by distilling the metal into a highly evacuated bulb. When light waves strike such a surface electrons are

[1] See O. W. Richardson's book, "The Emission of Electricity from Hot Bodies."

liberated, their number being proportional to the intensity of the light. This action makes possible the *photoelectric cell*, which has great importance in picture transmission, by wire or radio, and in talking films.

If a metal surface is bombarded by high-speed electrons, other electrons are "splashed" out of the surface where the high-speed one strikes; this phenomenon is called *secondary emission.* The number of electrons splashed out depends upon the speed of the bombarding electron, as many as five being ejected from the metal surface by the impact of only one electron, if this is traveling at high speed.

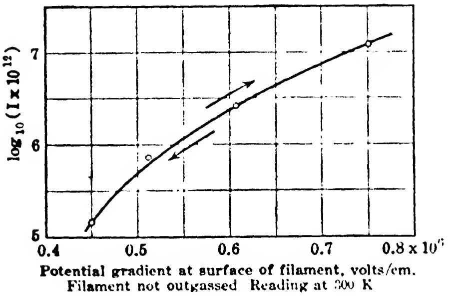


FIG. 1.—Electron current from a filament at room temperature. The filament had not been outgassed, and the potential gradients required were less than 10^6 volts per cm.

If extremely high-voltage gradients are impressed upon an electrode (1,000,000 volts or more per cm.) it seems that electrons are pulled out of the metal even at room temperature. Just how much of this effect is due to ionization which must be present when such very high potential gradients are used is not known. Some gas must always be present in a tube no matter how highly it is exhausted, and this gas is of course ionized. In that case the bombardment idea of the preceding paragraph might be considered as an explanation.

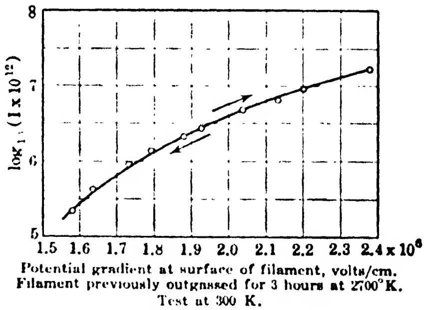


FIG. 2.—When the filament of Fig. 1 had been outgassed it required more than twice as high a potential gradient to get the same current.

For many years the possibility of pulling electrons out of cold metals was in dispute; some experimenters reported positive results and some negative. When it is remembered that the highest attainable vacuum there are still many millions of gas molecules per cubic centimeter still present, it will be evident that even if current is obtained, its origin may well be questioned. Apparently reliable results are reported by Piersol,[1] and two of his curves are reproduced in Figs. 1 and 2. Apparently, with a gradient at the surface of the tungsten filament of about 10^6 volts per centimeter, he obtained reproducible currents of the order of microamperes. Such

[1] Phys. Rev., March, 1928, p. 441.

currents are of great theoretical interest but have no practical application as yet, because much larger currents are readily attainable when the emitting surface is sufficiently heated.

Theoretical Prediction of the Number of Electrons Emitted from a Hot Body.—The number of atoms (or molecules) of an ordinary liquid which evaporate was known to vary with the latent heat of evaporation of the substance and temperature according to the equation

$$N = A\sqrt{T}\epsilon^{-\frac{a}{2T}}$$

where N = number of atoms evaporating per second per square centimeter of surface;
T = absolute temperature, on the Centigrade scale ordinarily called degrees Kelvin;
a = latent heat of evaporation;
A = a constant.

Richardson was the first to draw an analogy between the evaporation of atoms and possible evaporation of electrons from a hot metal. Reasoning from the above equation he came to the conclusion that the number of electrons evaporating per second (current) could be expressed by the equation

$$i = a\sqrt{T}\epsilon^{-\frac{b}{2T}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

in which i = current of emission per square centimeter of hot surface;
T = absolute temperature of hot metal;
b = latent heat of evaporation of electrons = 105,000;
a = a constant.

More recent workers give a formula of somewhat different form—thus Dushman[1] regards it more reasonable to write the relation in another form (also due to Richardson), namely,

$$i = AT^2\epsilon^{-b_0/T}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

in which i = amperes of current per square centimeter of hot surface;
T = degrees Kelvin (absolute Centigrade);
A = a universal constant; it varies somewhat with the method of derivation, but is between 50 and 60;
$b_0 = \phi_0 e/k$ in which
e = electron charge;
k = Boltzmann's constant;
ϕ_0 = Richardson's work function.

[1] "Electron Emission from Metals as a Function of Temperature," Phys. Rev., Vol. 21, No. 6, June, 1923.

The value of ϕ_0, in volts has been determined accurately for some of the metals, notably: platinum (6.27), tungsten (4.52), thorium (3.35), molybdenum (4.41), tantalum (4.07), calcium (2.24), and caesium (1.81). In general, the values are lower the larger the atomic number.

It might seem that Richardson's equation and Eq. (2) could not possibly represent similar curves, one involving $T^{1/2}$ and the other T^2. However, the two equations plot to nearly the same form, because the exponential term really controls the form of the curve.

As this predicted current was due to the thermal activity of the emitting surface Richardson suggested the term *thermionic current*, a name which is at present used to some extent; the term *electron current* is also used, but this is really not distinctive, because all currents, arising from whatsoever cause, are due to the flow of electrons.

Dubridge reports[1] that he finds the photo-electric work function and the thermionic work function for platinum the same, that is, it requires the same amount of work to pull out an electron whether this is done by light energy or heat energy. He finds that Dushman's A (Eq. 2) is by no means a universal constant; on several carefully outgassed specimens of platinum he found A to be 200 times as much as Dushman assumes.

The emission of electrons predicted by Eq. (2) would give currents from a heated tungsten filament about as shown in Fig. 3; it is evident that very large currents might be expected from a tungsten filament at temperatures well within the safe operating region.[2] Of course, ordinarily there is no current of such magnitude due to emitted electrons; although the number of electrons indicated in Fig. 3 is really emitted, they at once re-enter the surface so that on the whole there are no electrons leaving the hot surface. As soon as an electron leaves the filament it (the filament) is left charged positively and so attracts the emitted electron; thus there are as many electrons falling back into the filament as are expelled by the internal thermal agitation.

Suppose, however, that there is, close to the heated filament, a positively charged metal plate; an expelled electron will have two forces acting on it, one tending to make it fall back into the filament, and the other pulling it toward the positively charged plate. Which force has the preponderating effect will depend, of course, upon the value of the positive plate potential; if this is made sufficiently high, all the electrons emitted from the hot surface will be drawn to the plate, none of them re-entering the hot emitting surface.

[1] Phys. Rev., Feb., 1928, p. 236.

[2] The melting-point for tungsten is 3270° C.; reckoning the safe operating temperature as that which gives the filament 2000 hours' life, the safe temperature increases somewhat with the diameter of the filament, being perhaps 2200° C. for a filament 0.01 cm. diameter and 2300° C. for one 0.04 cm. diameter.

The value of the current under this condition is called the *saturation current*; this value of current measured for different filament temperatures should satisfy Richardson's equation because all the electrons emitted go over to the plate.

As early as 1902 Richardson published experimental results confirming his theory. Many other experimenters published results seemingly contradicting the relations given in Eq. (1), and for several years Richardson's theory was the subject of dispute.

It seems that very minor changes in the amount of gas[1] in the tube used, or the condition of the surface of the hot metal, completely nullified the results obtained, and such has been found to be the case. H. A. Wilson found, e.g., that the emission from a hot platinum filament might be reduced to 1/250,000 of its normal amount by first heating the filament in oxygen or boiling it in nitric acid; also he found that the presence of a slight amount of hydrogen around the heated filament completely destroyed the effects of the oxygen and nitric acid. On the other hand, it is found that the electron emission from tungsten is very much increased by such an impurity as thorium; if a small percentage of thorium is present in a tungsten filament the emission is many times as great as though pure tungsten were used.

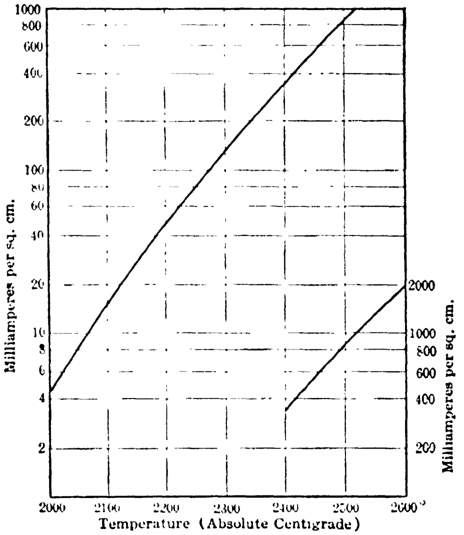


FIG. 3.—Theoretical values of current due to electron emission from a pure tungsten filament.

As a result of Wilson's experiment it was evident that the condition of the hot surface was of utmost importance in determining the emission; the layer of oxygen-filled platinum on the surface practically prevented emission. Yet a year afterward Wehnelt showed that if a platinum filament was coated with lime (calcium oxide) the emission of electrons at a given temperature was vastly greater than from the platinum itself.

Langmuir's experiments, performed with tungsten filaments in extremely high vacuum, proved without doubt the truth of Richardson's

[1] See article by Lockrow, Phys. Rev., Feb., 1922, for effect of gas on emission.

prediction and indicated that the various experimenters whose tests had showed the opposite had not been careful enough in the manipulation of their experiments and in the interpretation of the results. He found that the higher the vacuum the more consistently did experiment and theory agree, whereas others had concluded that gas was absolutely essential if the thermionic current was to be obtained. In one of Langmuir's tests he showed that the presence of only 0.000001 mm. pressure of oxygen was sufficient practically to stop the emission of electrons from a hot tungsten filament. It seems then that the condition of the surface of the hot electrode affects the emission of electrons much as the evaporation of water is prevented by covering the surface with a thin layer of oil or similar substance.

Distribution of Electrons near the Surface of a Hot Metal.—In Fig. 4 is shown, in rather crude fashion, the manner in which the electrons are concentrated near the surface of a hot body, the three figures being for

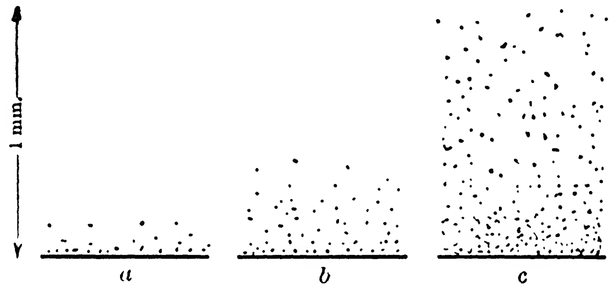


FIG. 4.—Conventional diagram to represent the distribution of electrons near the surface of a hot metal, for increasing temperature.

temperatures of perhaps 2100°, 2300°, and 2500° absolute temperature. In (*a*) but few electrons are coming off and these have such a low velocity that they are pulled back into the tungsten before they have moved out from the tungsten perhaps 0.001 cm. In (*b*) more electrons are emitted and their mean velocity has increased so that more of them move farther away from the surface before falling back. In (*c*) is shown a much denser electron atmosphere near the surface and also extending to considerable distance from the tungsten surface. In one tungsten filament tube tested by the author it was found that at normal operating temperature only 1/8000 of the electrons emitted reached a distance 0.15 cm. from the hot filament, most of them never going very far (perhaps 0.001 cm) from the surface.

In Fig. 5 is shown a set of curves corresponding to the conditions given for Fig. 4, the area under each curve gives the number of electrons emitted from the filament and the form of the curve illustrates how the number of electrons having a given velocity changes as the temperature is increased. At temperature T_1, but few electrons are emitted and they have on the average a low velocity, practically none having a velocity greater than V_1; at temperature T_2, many more electrons come off and on the average they have a higher velocity; the same effect, but more of it, is shown for the highest temperature T_3.

Some idea of the velocity with which these electrons leave the surface of the tungsten can be easily obtained. From certain experiments we know that an electron must fall freely through a potential difference of about 4 volts, before it gains sufficient energy to break through the "surface tension" or "surface constraint" of the average metal such as

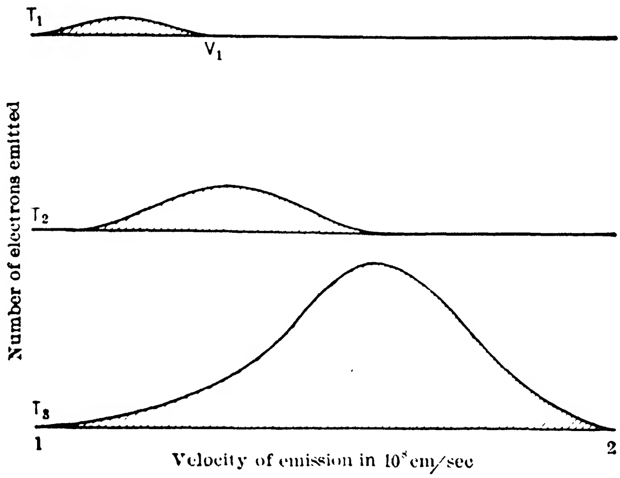


FIG. 5.—Velocity distribution for electrons emitted from hot tungsten, for three different temperatures.

tungsten, tantalum, or molybdenum. If we use the relation that, in any accelerating system,

Potential energy lost = kinetic energy gained

we put

$$Ve = \tfrac{1}{2}mv^2, \qquad \ldots \ldots \ldots \quad (3)$$

in which V = potential difference through which electron has fallen (e.s. units);

e = charge of electricity on one electron;

m = mass of electron;

v = final velocity of electron, assuming it to start from rest.

Transposing, we get $v^2 = 2V\,e/m$, and e/m has been determined many times, its value being 5.3×10^{17}, in electrostatic units.

Hence if an electron falls through 1 volt difference of potential (1 volt $= \frac{1}{300}$ e.s. unit) the above relation gives v approximately 5×10^7 cm. per sec. As the surface constraint of tungsten is about 4 volts we see that an electron, to break through, must have a velocity of about 1×10^8 cm. per sec.

As noted above an electron must be moving towards the surface with a velocity of about 1×10^8 cm. per sec. to break through the surface tension of tungsten. If they approach the surface (from the inside) with just this velocity when they do get through the surface their velocity is gone and they at once fall back into the hot metal. From the principles of thermodynamics it has been calculated that if the tungsten is at 2400° Kelvin (absolute Centigrade scale) 90 per cent have (after breaking through the tungsten surface) velocity greater than 0.4×10^7 cm. per sec.; 70 per cent have greater than 0.6×10^7 cm. per sec.; 25 per cent have greater than 1.3×10^7 cm. per sec.; 1 per cent have velocity greater than 2.5×10^7 cm per sec. and only 1×10^{-6} per cent have a velocity as high as 4×10^7 cm. per sec. A temperature of 2400° Kelvin is about that of the ordinary tungsten incandescent lamp.

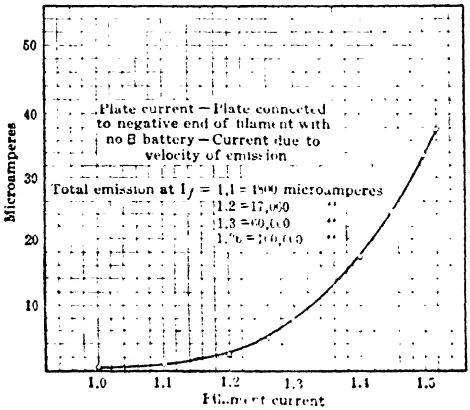


FIG. 6.—Electron current from a hot tungsten filament to an adjacent cold plate, at the same potential as the lowest potential of the filament. Current due to velocity of emission of the electrons.

If a cold metal plate, electrically connected to the filament outside the tube, is in the same vacuum tube as the hot filament, and close to it, some of the high-speed electrons may have sufficient velocity to carry them from the hot filament to the cold plate; they then flow along in the circuit connecting the plate to the filament. This thermionic plate current can exist even though the plate is at the same potential as the lowest potential point in the filament. Such an effect is shown in Fig. 6; the amount of the plate current recorded was due to electrons emitted from the filament with such a high velocity that their inertia carried them across the 0.2-cm. space separating the plate from the filament.

It will be noticed how the number reaching the plate increases rapidly with the value of filament current, due to the two effects mentioned above, greater emission and higher velocity of emission. The total emission of electrons from the filament for various filament currents is noted on the curve sheet of Fig. 6; above a filament current of 1.36 amperes this total emission could not be accurately measured, for reasons to be taken up later. The filament used in getting the curve of Fig. 6 was only about 3 cm. long and of approximately the same diameter as that of a 100-watt tungsten lamp, yet it will be found by calculation from the values given on the curve sheet that at 1.3 amperes in the filament the emission was about 4×10^{17} electrons per second, and of this number there were 4×10^{13} which had sufficient velocity to carry them away from the filament an appreciable fraction of a centimeter.

From the previous analysis of electron emission from a hot body it will be realized that the condition close to the surface of a hot filament resembles very much the atmosphere surrounding the earth, a depth of earth atmosphere of 1 kilometer corresponding to a depth of "electron atmosphere" of about 0.01 mm. Just as the earth's atmosphere becomes less dense with increase of distance from the surface, so does the density of electrons decrease with increase of distance from the filament; the upper part of the earth's atmosphere contains the more rapidly moving atoms of gas just as is the case of the high-speed electrons getting farther away from the filament than those emitted with lower velocity.

Effect of Oxide Coating on Electron Emission.—As noted on a previous page very slight changes in the surface conditions of a metal greatly affect the ease with which electrons can evaporate. The presence of the slightest amount of water vapor will reduce the emission to almost zero—it is said to "poison" the surface.

Very early work showed, however, the value of coating a cathode with certain oxides, and many of the tubes in use today employ that type of emitting surface, credited to Wehnelt. Thus certain oxides poison a metallic surface and others are very beneficial so far as electron emission is concerned.

The oxide-coated filament was originally made by coating a thin platinum-iridium (Pt, 95 per cent; Ir, 5 per cent) ribbon with successive applications of barium and strontium oxides in paraffin wax, and fixing this coat by heating the filament. The process was repeated sufficient times to give an oxide coating of 1 to 2 mg. per square centimeters of filament surface. Experiments have shown that other materials, of less precious metals are equally, or more, effective as filaments.[1] Pure nickel is not sufficiently strong when heated, so an alloy of nickel, cobalt, iron, and titanium (Konel metal) has been adopted. It radiates much more effectively than does platinum, being about 120° C. cooler than platinum when radiating 4 watts per square centimeter (Pt at 900°, Konel metal, 780°). Nevertheless, the oxide-coated Konel filament seemed to give somewhat more emission than the oxide-coated platinum filament under equal power

[1] Lowry, Phys. Rev., June 1, 1930, p. 1367.

input. This seemed to indicate that the core metal had an important effect on the electron emission from the oxide surface.

Becker [1] has given a very complete analysis of the emission phenomena occurring with oxide-coated filaments, especially applicable to barium and strontium oxides. He comes to the conclusion that some of the barium oxide is reduced and that the emission really takes place from an irregular surface (probably one molecule deep) of metallic barium on the surface of the oxide.

When first applied, the coating is in the form of carbonate and is white but after being " broken down " by continued heating of the filament, the carbonate is reduced to an oxide which sometimes turns a dark gray color; in this transformation it seems probable that the metallic coating of barium comes to the surface. Most recently, Becker and Sears [2] have concluded (*a*) that the core material does not directly affect the emission but it may greatly affect the ease with which barium is produced by heat or electrolysis; (*b*) that the thermionic-electrons originate in the oxide just underneath the surface layer of barium; and that (*c*) most of the current through the oxide is conducted by electrons, a small portion being carried by barium and oxygen ions.

In Fig. 7 are shown the comparative merits of pure tungsten and oxide-coated filaments, the coordinates being suitably proportioned to give a straight line between emission and filament power used.[3] The emission shown is for 1 sq. cm. of filament surface, and the tungsten wire is supposed to be 3 mils in diameter.

The heavy vertical lines give the approximate useful life of the two types of filament when operated at the temperature (degrees Kelvin) noted on the curve sheet. These curves indicate that pure tungsten requires almost ten times as much power as the oxide-coated filament, for the same rate of electron emission.

Thomson [4] reports that under certain conditions he finds emissions greatly in excess of those heretofore published. Some of his surprising results are given in Fig. 7-A; at what he styles "normal operating temperatures" he finds an emission efficiency as high as from 1 to 4 amperes per watt of filament power. The thermionic work function he found to be between 0.75 and 0.905 volt.

Of course if a cathode (filament) is properly heat-insulated much higher than ordinary emission is to be expected. Thus Hull [5] shows how to heat-

[1] Phys. Rev., Nov. 15, 1929, p. 1323.

[2] Phys. Rev., Dec. 15, 1931, p. 2193.

[3] This form of coordinate paper has come into universal use for plotting emission characteristics; it is due to Davisson of the Bell Laboratories.

[4] Phys. Rev., Oct. 15, 1930, p. 1415.

[5] Journal A.I.E.E., Nov., 1928.

insulate a cathode and get an emission of 0.6 ampere per watt of filament power. This scheme is of no avail, however, in high-vacuum tubes, as the actual amount of emission current which is obtainable from such a protected filament is small.

Thoriated Tungsten Filament.—At the present time many of the commercial tubes use a thoriated filament, this really being a tungsten

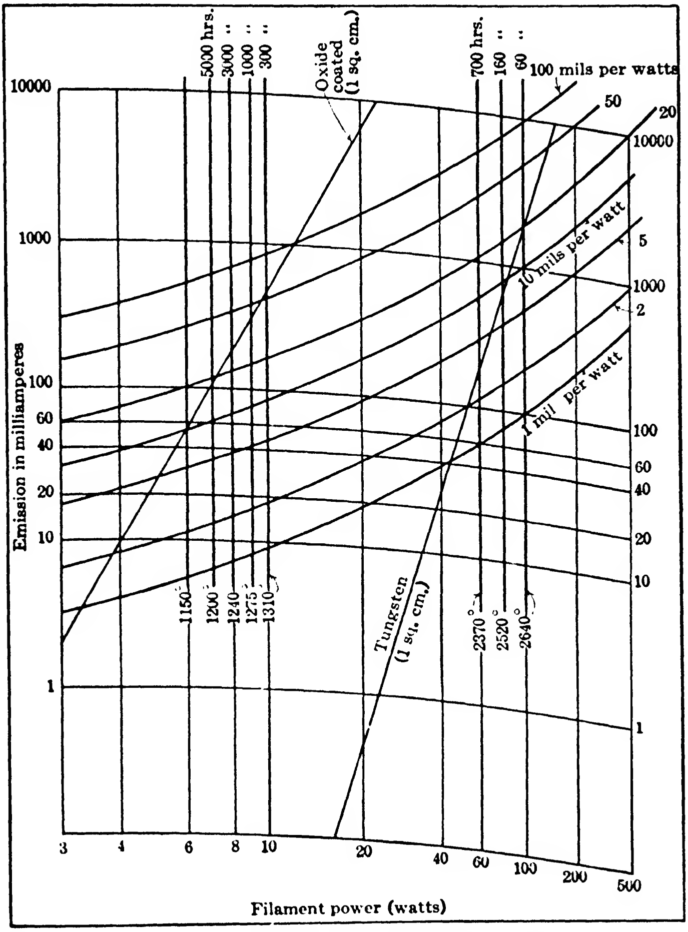


FIG. 7.—Showing the comparative emissive power of tungsten and oxide-coated cathodes at different temperatures.

filament having a very thin layer of the metal thorium on its surface. The electron emission seems to be from the thorium, the tungsten merely serving to heat the thorium and to renew this layer as it disintegrates. The development of this special cathode is apparently due to Langmuir and his co-workers.

The filament is made of tungsten in which about ½ per cent of thorium oxide and some carbon are dissolved. As the temperature of such a filament is gradually raised two important phenomena come into play. At first some of the thorium oxide is reduced to metallic thorium and then this gradually works its way to the surface of the filament. If the temperature is kept below a certain limit this metallic thorium stays on the surface constituting a layer one atom deep. Where more thorium atoms work their way to the surface and come up under other thorium atoms already there, the latter at once evaporate, thus maintaining the layer only one atom thick.

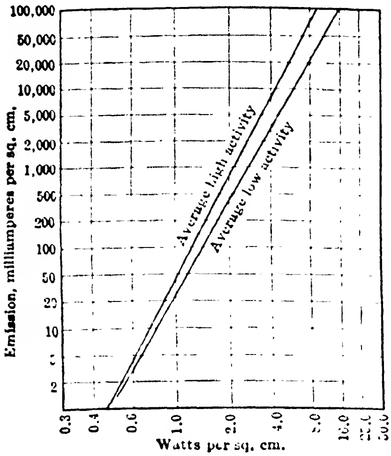


FIG. 7A.—Under certain conditions unusually high emission has been found, much greater than is obtained in ordinary commercial tubes.

If the temperature is raised a few hundred degrees the thin thorium layer completely evaporates, leaving a tungsten surface. To be sure, at this higher temperature, the metallic thorium is formed (from the oxide) more rapidly and comes to the surface more abundantly but it does not stay on the surface; it evaporates at once. This peculiar behavior of the thoriated cathode has been completely analyzed by Langmuir[1] as the outcome of a series of brilliant experiments. Certain other metals act somewhat like thorium, such as caesium, etc.[2]

Activation of Thoriated Tungsten.—The formation of the layer of thorium upon the tungsten filament has been styled its *activation*; the removal of the layer is styled the *deactivation* of the filament; these two actions take place according to certain complicated logarithmic laws which Langmuir has derived. Fig 8 gives a typical pair of curves, obtained as

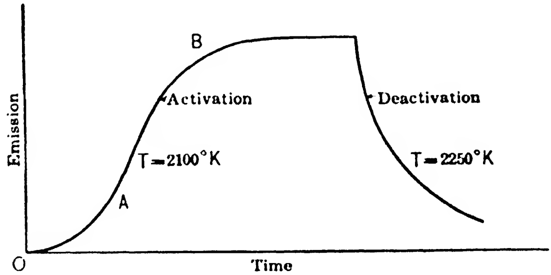


FIG. 8.—Showing the accumulation and dissipation of the thorium coating on a thoriated filament.

[1] "The Electron Emission from Thoriated Tungsten Filament," Phys. Rev., Oct., 1923.

[2] "Electron Emission from Absorbed Films on Tungsten," by Kingdon, Phys. Rev., Nov., 1924.

follows: The thoria-impregnated filament is heated to 2500 K. in a thoroughly evacuated bulb, this being immersed in liquid air to condense any gas which might appear in the bulb during the test. A fresh tungsten deposit on the interior walls of the bulb also helps to clean up (adsorb) any gas which might appear. At 2500° K. most of the metallic thorium which might be on the surface evaporates but to ensure this the filament is " flashed " for a few seconds at 2800° K.; this procedure leaving a chemically pure surface.

The filament is now held at its activating temporature (about 2100° K.) and periodically its emission is measured. The emission increases with time, as more thorium appears on the surface as shown in the first part (*OAB*) of Fig. 8. During this time the thorium oxide is being reduced to thorium and this is gradually migrating to the surface of the tungsten. The amount of emission does not vary directly with the percentage of the tungsten surface covered with thorium, as one might expect. If we take the emission from a pure tungsten filament at 2100° K. as unity the emission from a completely thorium-covered filament is about 100,000. But if half the tungsten surface is covered with thorium this emission is not 50,000; it is only 316! The rate at which the activation takes place depends of course on the temperature, as well as the final value of current when the emission has become constant. Thus at 2000° K. a certain filament gave a final emission of 0.001 ampere and reached 90 per cent of this in about 3 hours; at 2050° K. the final emission was 740 micro-amperes and 90 per cent of this was reached in about 1 hour; at 2150° K. the final emission was 250 micro-amperes and 90 per cent of this was reached in 30 minutes.

Deactivation.—If the layer of thorium is in some way driven from the surface the emission drops to that of tungsten, which is practically zero at the operating temperature of a thoriated filament (about 1500° K.). The thorium may be dissipated by too high a temperature or by bombardment by positive gas ions. In Fig. 8 is shown a deactivation curve brought about by operating the filament at too high a temperature. The excessive rate of thorium evaporation at this temperature rapidly dissipates the layer of thorium atoms in spite of the increased supply of these from the inside of the filament.

All tubes have some gas left in them and under certain conditions this gas becomes ionized and the positive ions are projected into the filament. The thorium is thereby knocked off, or so contaminated that it no longer emits an appreciable number of electrons.

An interesting curve sheet showing the deactivation of thoriated tungsten is shown in Fig. 9; it is taken from a paper by A. W. Hull.[1] A pure tungsten filament and a thoriated tungsten filament were both operated in a

[1] Journal A.I.E.E., Nov., 1928, p. 798.

pressure of 0.03 mm. of argon, the tungsten at 2450° K. and the thoriated filament at 1900° K. The current from the tungsten filament rose gradually to 3 amperes and held at that value as the voltage on the anode was increased to 130 volts. The current from the thoriated filament rose much faster, reaching a value of 7.6 amperes at 26 volts on the anode. Just above this voltage the ionized argon molecules were able to bombard the filament with sufficient energy to disintegrate the atom-deep thorium layer on the cathode and so spoil the emission. The current therefore fell off rapidly, being practically zero at 100 volts anode pressure.

Reactivation.—If the thorium layer has been either evaporated or poisoned the filament may generally be restored to normalcy by a simple treatment. The surface is first cleaned by operating for a short time at excessive temperature; it is then operated for a longer period at about 2100° K. and so follows the curve OAB of Fig. 8. A 3-volt filament for example should be operated at 10 volts for 30 seconds to rid the surface of impurities and then at 4.5 volts for 10 minutes. During this treatment there must be no voltage on the grid or plate of the tube. Correspondingly a 6-volt filament should be operated at 15 volts for 1 minute and then at 7.5 volts for 10 minutes. This reactivation process may be carried out several times before the supply of thoria in the filament is used up.

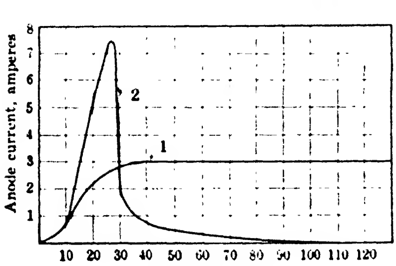


FIG. 9.—When surrounded by an atmosphere of 0.03 mm. of argon the electron emission from pure tungsten and thoriated tungsten varied as shown by these curves; above the ionizing voltage of argon, the positive ion bombardment knocked all the thorium off the thoriated filament, reducing its emission to practically zero.

Comparison of Oxide-coated and Thoriated Filaments.—The thoriated filament evidently has the distinct advantage of permitting recuperation after spoilage, which as yet is not possible with the oxide-coated filament, but it may well be that improvements in the character of the oxide coating may give it a life as long as that of the thoriated filament with several reactivations.

The gas can apparently be more completely eliminated from the thoriated filament during manufacture, which is a distinct advantage in high power, high-voltage tubes. Apparently also the thoriated filaments permit of uniform product, due to simpler process of manufacture.

According to data given by King,[1] however, the present type of

[1] "Thermionic Vacuum Tubes and their Application," Bell System Technical Journal, Oct., 1923.

oxide-coated filament is considerably more efficient than the thoriated type. In Fig. 10 are reproduced some of the data given by King; it appears from this chart that the oxide-coated, platinum-nickel filament uses only about half as much power for the same emission. The operating temperature of the oxide filament, for the same emission as the thoriated one, is

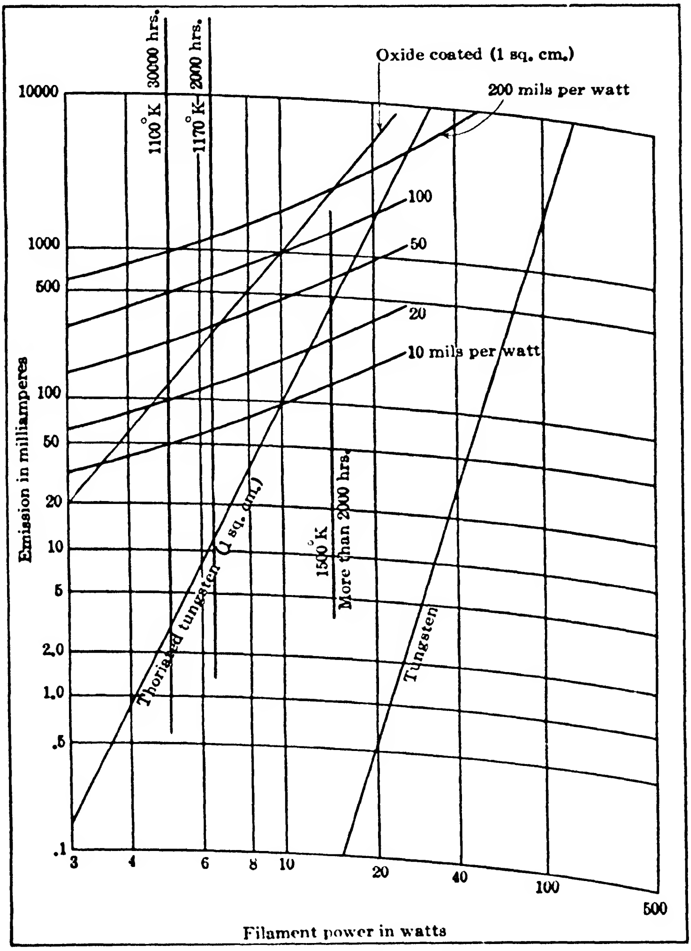


FIG. 10.—Comparative merits of pure tungsten, thoriated tungsten, and modern oxide-coated cathodes.

about 400° lower. For reasonably efficient emission the oxide-coated filament is a dull red, the thoriated filament is yellow, and the tungsten filament is dazzlingly bright.

Two-electrode Vacuum Tube.—The property of hot bodies *in vacuo* permitting passage of electrons to a cold electrode in the same vessel was originally called the Edison effect; it was noticed in incandescent lamps as early as 1884. In 1896 Fleming gave the results of a series of experiments in thermionic currents through *vacuo*, but in the light of our present knowledge it seems that a large part of the current measured by him was due to conduction by the ionized gas in the tube he was using. He found some characteristics which were really due to the electron emission, notably the unilateral (one direction only) conductivity of the apparatus, the non-linear relation between the plate potential (with respect to the filament) and the plate current, and the fact that a large separation of plate and filament tended to reduce the amount of plate current obtainable. He found, however, that the plate current was unstable and that the better the vacuum the less the plate current became; both of these effects show that ionized gas was largely responsible for carrying the plate

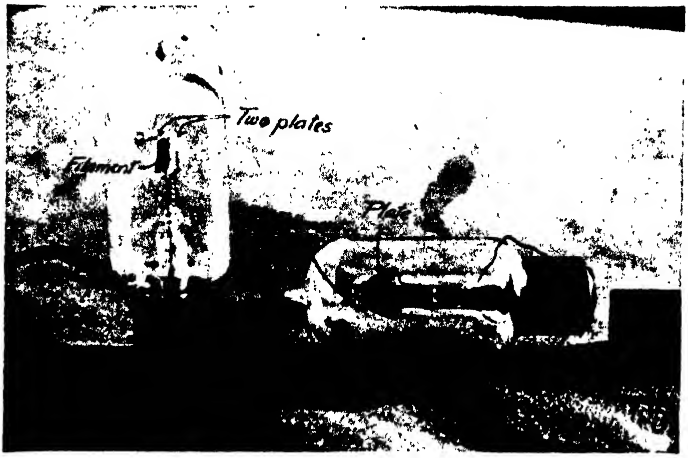


FIG. 11.—One type of Fleming valve, used on early Marconi receiving sets as detector.

current. The unilateral conductivity of a vacuum tube having two electrodes, one hot and the other cold, was utilized by Fleming for the detection of damped high-frequency waves and was patented by him in 1905. This patent was a very important one in the field of radio telegraphy; it goes by the name of the "Fleming valve" patent. A cut showing a Fleming valve is given in Fig. 11. More recent devices which function because of the unilateral conductivity between hot and cold electrodes in a partial vacuum are the mercury rectifier, the tungar rectifier, and the kenotron.

The *mercury rectifier* uses a hot spot on a pool of mercury as the source of its electrons, the necessary temperature of the hot spot being maintained by heat caused by the plate current itself; ionized mercury vapor serves as the carrier of the current which can pass one way only.

The "*tungar*" *rectifier* operates in a manner different from that of the mercury rectifier, in that a hot tungsten filament serves as the source of electrons, this filament requiring an auxiliary source of power for maintaining its requisite temperature. The tube is filled with an inert gas (generally about 2 lb. absolute of argon), and this gas is ionized by the electrons from the hot filament; the carrier of the plate current is thus in this case ionized gas for the main part, the number of electrons emitted from the hot filament being sufficient to carry perhaps 1/500 of the current to the plate.

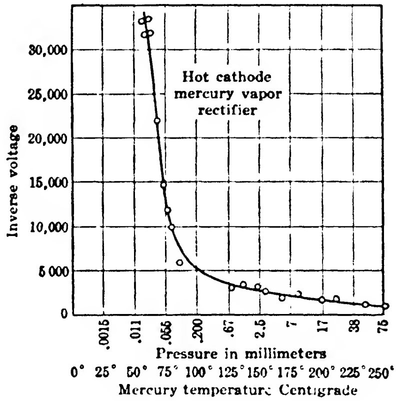


FIG. 12.—Variation of the "break-down" voltage of a mercury vapor tube, for various pressures of mercury vapor.

The hot cathode mercury vapor rectifier has recently been developed, largely due to the work of A. W. Hull, and is now widely used. A two-electrode tube, having an oxide-coated filament, has sufficient mercury in it so that when in normal operation the mercury vapor pressure is a few hundredths of 1 mm. This is sufficient pressure to permit the tube to carry its full emission current with a drop of only about 18 volts, but the tube will not break down and carry reversed current unless the inverse voltage is raised to many thousands. In Fig. 12 are shown the relations between mercury vapor pressure, temperature of the mercury vapor, and "break-down" voltage required before the tube will carry current in the reverse direction. If one of these tubes, having a few drops of mercury, is heated to, say, 75° C., it will carry many amperes (depending upon the size of the hot filament) in one direction with a drop of from 15 to 20 volts, but as the pressure of the vapor, at this temperature, is only 0.055 mm. (see Fig. 12), it will require 12,000 volts to make the tube carry reversed current. If the tube is cooled off so that the temperature of the mercury

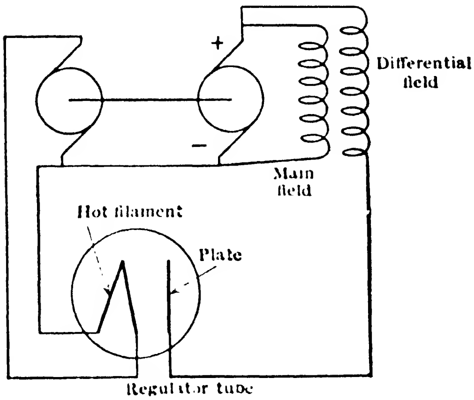


FIG. 13. Use of a two-electrode tube as a voltage-regulator for a variable speed generator.

vapor is only 60° C., its pressure will decrease to about 0.02 mm., and the required inverse voltage, for breakdown, is about 30,000 volts.

The *diode* or *kenotron* is a rectifying tube which really operates as a thermionic valve; the tube is exhausted as thoroughly as possible, so much so that whatever gas may be present plays an unimportant rôle in the functioning of the device. The plate current is never greater than that actually emitted by the hot filament. These rectifying tubes are made in large sizes, sufficient to rectify several kilowatts of power; the vacuum in these is so high that no appreciable current is carried in the reversed direction (electrons from plate to filament) even if 100,000 volts is impressed.

In small sizes they have been used as voltage regulators for self-excited

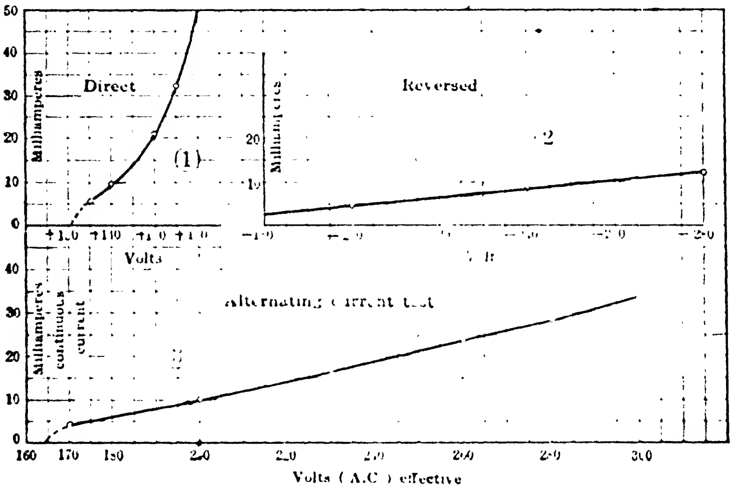


FIG. 14.—Characteristics of a low gas pressure, cold electrode rectifying tube. Its rectifying action depends upon the asymmetry of its two electrodes.

generators, the speed of which is variable. By having a differential winding on the field poles, which is supplied with current through a regulator tube, and connecting the filament of this regulator tube across a low-voltage winding on the armature, a small generator may be made to maintain practically constant voltage over a wide range of speed variation. The scheme of connection is shown in Fig. 13, and the reasons for the tube maintaining such constant voltage over such a wide speed range will appear from an examination of the characteristics curves of such a tube.

There is another type of rectifying tube which has had quite some application in radio work; it is used for changing an a.c. supply into a continuous current for the various circuits of the thermionic tube used in radio receivers. It depends entirely upon ionization of the gas with which the tube is filled at low pressure. The electrodes are so shaped that electrons leaving one of them travel only a short distance before being stopped, whereas those starting from the other electrode travel much farther through the gas before meeting the other electrode. Now an electron has to travel through a rarefied gas a considerable distance before setting up much "progressive" ionization and so making the path conductive. The result is that whereas the tube "breaks down" at about 150 volts, and conducts readily in one direction, it takes about 700 volts to make it conduct in the opposite direction.

The characteristics of one of these tubes (styled the S tube) are shown in Fig. 14. Like any device depending upon ionization for conduction of current, it is unstable and must be used with sufficient series resistance; this was done in getting the curves of Fig. 14. The voltages shown were measured at the tube terminals. In the a.c. test (curve 3) an alternating voltage, with effective values as given in the curve sheet, was impressed, and the current that passed was used on a c.c. ammeter; this curve then shows the true rectifying action of the tube.

This is not a thermionic tube at all, as neither of the electrodes is heated; its rectifying properties arise from the shape of the electrodes.[1]

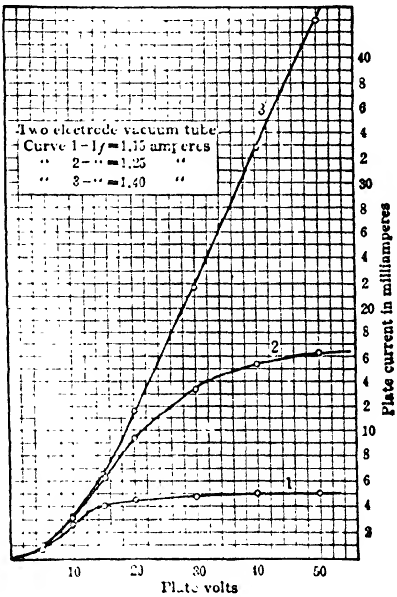


FIG. 15.—Variation of plate current with plate voltage (for various filament currents) in a small kenotron.

Characteristic Curves of a Two-electrode Vacuum Tube—Value of Saturation Current.—If the filament current of a diode is maintained constant and plate voltage varied, readings being taken of plate voltage (with respect to the filament) and plate current, curves will be obtained having the shape shown in Fig. 15; here three curves are shown for three different filament currents as noted on the curve sheet.

Curve 1, Fig. 15, shows the variation of plate current for a filament current of 1.15 amperes; it is evident that as the plate voltage is increased from zero the plate current rises more rapidly than the first power of the voltage until about 10 volts is impressed; for higher voltage a smaller increase in plate current is obtained and above 30 volts no further increase

[1] For more detailed explanation of the action of this type of rectifier, see Journal A.I.E.E., Sept., 1922.

in plate current is obtained, even if the plate voltage is increased to 300 volts. It is evident that a plate voltage of 30 is sufficiently high to attract to the plate *all the electrons which the filament emits*, at the temperature reached with a filament current, I_f, of 1.15 amperes. This value of plate current, which is limited only by the emitting power of the filament, is called *saturation current of the tube.*

Saturation current evidently will be determined in magnitude by the temperature and area of the filament surface; also for higher filament temperatures (higher emission) it will require higher plate voltage to obtain saturation current. Thus when I_f is raised to 1.25 amperes, saturation current is increased from 5 milliamperes, its value for $I_f = 1.15$ amperes, to about 16.5 milliamperes, and whereas in the first case 30 volts on the plate was sufficient to obtain saturation current, in the second case even 50 volts was not quite sufficient to reach saturation.

When I_f was increased to 1.40 amperes, the emission was so great that a plate voltage of 50 was not nearly enough to obtain saturation current and the value of saturation current is going to be very high, judging from the shape of the curve. Its value was actually determined in another test and found to be 140 milliamperes.

Considering curve 3 of Fig. 15, it is apparent that for any plate voltage shown on the curve sheet, the number of electrons arriving at the plate is only a small fraction of the number emitted by the hot filament; e.g., with a plate potential of 20 volts the current to the plate was only 10.8 milliamperes, whereas the total emission of electrons from the filament is sufficient to give a plate current of 140 milliamperes. It is therefore evident that to obtain at the plate all the electrons emitted from the filament a certain minimum voltage must be impressed on the plate. The reason for this is given by an analysis of the electron distribution between the hot filament and cold plate.

Form of Saturation Current.—It will be noticed that the plateau of current (Fig. 15), where it is assumed that all the evaporated electrons are being drawn over to the plate, is not quite level, i.e , the current continually increases to a small degree as the anode voltage is raised. With oxide-coated filaments, and thoriated filaments, this lack of saturation is much more marked than in the case of pure tungsten, which was used in getting the results of Fig. 15. It seems probable that this lack of saturation is entirely a surface effect; the emitting surface has small irregularities on it which result in different electric gradients at the surface and hence lack in uniform behavior of the emitting surface. Thus if there is a small crevice in the surface of the filament it is evident that very high voltages will be required to " reach in " to the bottom of the hole and take away all the evaporated electrons. As the chance of such crevices in the filament surface is greater with coated metals than with pure tungsten we should expect the slope of the "saturation plateau" to be greater in the case of the coated filament, and such proves to be the fact. An analysis of this surface action has been given by Becker and Mueller [1] and by Reynolds.[2]

Space Charge.—In Fig. 16 is shown in very elementary fashion the distribution of electrons between the plate and filament; we will consider the electric forces acting on two of the electrons a and b. Electron a is urged to the plate by two forces, the attraction from the plate and the repulsion from all the electrons between it and the filament; it will undoubtedly go to the plate. But electron b, although attracted by the plate, is repelled by all the electrons between the plate and itself; whether it will move toward the plate or re-enter the filament depends upon the relation between these two forces. It is evident that close to the surface of the filament the effect of all the electrons in the space between the filament and the plate (constituting the *space charge*) will practically neutralize any effect of the plate, *unless the plate voltage is high enough to give a force of attraction greater than the repulsive force exerted by the space charge.*

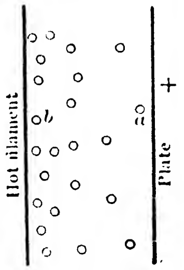


FIG. 16.—Elementary representation of the distribution of electrons between the hot filament and cold plate of a diode.

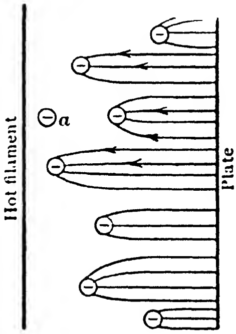


FIG. 17.—If there are sufficient electrons between the plate and filament, the lines of force from the plate do not penetrate as far as the filament, thus leaving some electrons near the filament free from attraction to the plate.

There is another way of looking at the problem; to bring the plate to a certain potential with respect to the filament requires a certain quantity of electricity, determined by the electrostatic capacity of the condenser formed by the plate and filament. Suppose this quantity of "positive" electricity is q, there will be then $4\pi q$ lines of electrostatic force leaving the plate, in the direction of the filament. These lines of force must end on q charges of negative electricity; but if the space charge is sufficiently large to furnish the requisite q the lines of force from the plate never penetrate to the filament.

An attempt to picture this is made in Fig. 17, for the picture as drawn electron a experiences no force at all from the plate, and so does just the same as it would if the plate were not there, i.e., goes back into the filament. But if the plate is brought to a higher positive potential, by putting more charge on it, more lines of force will emanate from the plate

[1] Phys. Rev., March, 1928, p. 431.
[2] Phys. Rev., Jan. 15, 1930, p. 158.

and so some may end on electron a and so attract it to the plate. It must be remembered that the above picture of what happens is very crude and artificial; the lines of force really have no entity and a is attracted, to some extent, by the plate for the condition shown in Fig. 17, but the attraction is negligibly small.

Of course it is to be remembered that the voltage gradient is not uniform between plate and filament. If the separation between plate and filament is 1 cm. and the voltage difference is 100 volts, it might be assumed that the voltage gradient between plate and filament was 100 volts per centimeter, but such is far from the fact. The potential gradient is always greatest at the filament surface and is greater for a small filament than for a larger one. Thus with a filament 0.00146 cm. in radius, axially placed in a cylinder of 1.06 inside radius, the potential gradient at the filament surface is 103 times the voltage between plate and filament.

It has been shown by Child[1] that when the emission of the electrons from the filament is much greater than that required by the plate current, the plate current may be expected to vary according to the relation

$$i=K\frac{E^{3/2}}{x^2}, \quad \ldots\ldots\ldots \quad (4)$$

in which E = potential of plate with respect to filament;
x = distance between filament and plate.

When E is measured in volts and x in centimeters this becomes

$$i=2.33\times10^{-6}\times\frac{E^{3/2}}{x^2} \text{ amperes per square centimeter of plate} \quad . \; . \quad (5)$$

If the plate is cylindrical in form with the hot filament placed in its axis this relation becomes

$$i=14.65\times10^{-6}\times\frac{E^{3/2}}{r} \text{ amperes per centimeter length of cylinder,} \quad . \quad (6)$$

where r = internal radius of cylinder.

The diameter of the filament is supposed small compared to the diameter of the cylinder in getting this formula.

If we have an equation in the form $x=y^a$ we have also $\log x=a \log y$, so that if the data for the curve $x=y^a$ are plotted on logarithmic coordinate paper the curve should become a straight line, the slope of which gives the value of the exponent a. Curve 3 of Fig. 15 was transposed to logarithmic paper and is shown in Fig. 18; it is seen that the exponent itself is variable, having a value about 2 for low plate voltages and rapidly decreasing for the higher values.

[1] Phys. Rev., 1911, Vol. 32, p. 498.

It could not be expected that the experimental results would agree with theory, because *the voltage between the plate and filament is different in different parts of the filament.*

This point must be borne in mind in interpreting all experimental results on vacuum tubes; practically all theoretical conclusions are reached from the premise of *uniform potential gradient* between the plate and all parts of the surface emitting the electrons. For the lower plate voltages this assumption is not even approximately true. The tube used in getting the results shown in Fig. 15 had an IR drop in the filament of 6 volts so that the potential relations in the tube may be about as shown in Fig. 19. The voltage difference is 20 at the negative end of the filament and only 14 at the positive end, having values between 14 and 20 at the intermediate points.

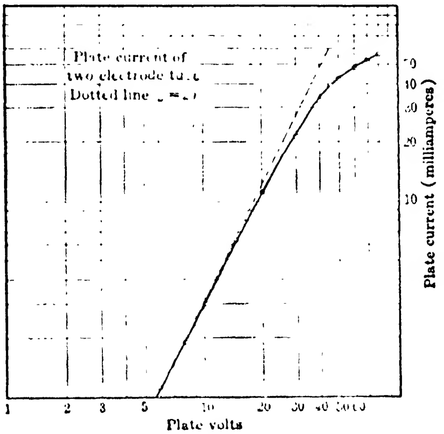


FIG. 18.—Curve 3 of Fig. 15 transposed to logarithmic coordinates.

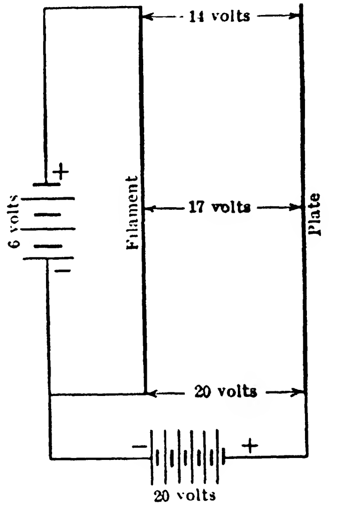


FIG. 19.—Variation of potential drop between the plate and different parts of the filament; the drop at one end of the filament is 14 volts and at the other end is 20 volts.

For such tubes we cannot expect to get theoretically correct results for the performance under any conditions; especially when the characteristic varies with the plate voltage to a power higher than the first (as the 3/2 or square) the departure of experiment from theory must be expected. The author built a tube in 1915 as shown in Fig. 20 in which the spiral tungsten filament A used for heating was entirely enclosed in a tungsten thimble B; this thimble constituted the hot surface from which the electrons were emitted. Such a construction gives a uniform potential gradient between the emitting surface B and the cylindrical plate C and so permits experimentation under the conditions assumed in theory. With this construction it is not possible to get the tungsten thimble as hot as the filament and so the emission is rather low, unless an oxide coating is used on the thimble. Such a construction permits the use of a high voltage filament and, a much more important point, *the electron current from the hot surface is not directly limited by the carrying capacity of the filament.* As will be explained later this feature becomes important in high-power tubes; in these tubes the electron current to the plate may be as high as 12 to 15 per cent of the filament current, so that the filament current *is 12 to 15 per cent greater at one end of the filament than it is at the other end.*

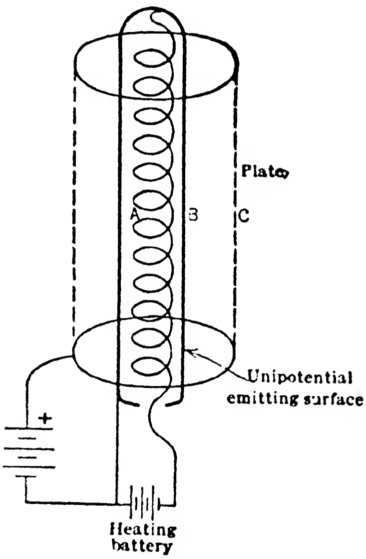


FIG. 20.—Showing the author's original scheme for getting a unipotential surface emitting electrons; the electrons come off from the thimble B, which is heated by the filament A.

With the construction of Fig. 20 it is possible to use alternating current in the filament proper and still not have the alternating potential drop in the filament affect the electron stream leaving the cathode. It is preferable in such a case, however, to completely insulate the filament from the thimble cathode.[1] Such an insulated construction makes possible the heating of the thimble cathode by the energy of electron bombardment. If the filament is maintained highly negative with respect to the thimble (which is generally best kept at zero potential), electrons leaving the filament will strike the inside surface of the thimble and there give up their energy, thus heating the thimble. There is practically no limit to the temperature producible by this method, except the volatilization of the thimble.

The idea of using indirectly heated thimble cathodes is used in very many of the modern tubes, although in a slightly different form. As shown in Fig. 21, a small cylinder of fused magnesia, A, has two small holes running through it lengthwise. Through these two holes is threaded a tungsten filament F, which is heated from the low voltage winding of a transformer giving about 2 volts. Tightly enveloping the magnesia cylinder is a cathode, in the form of a metal cylinder B, coated on the outside with electron-emitting oxides. A connection to this cathode, C, corresponds to the filament connection of the ordinary tube.

[1] See "A Combined Kenotron Rectifier and Pliotron Receiver Capable of Operation by Alternating Current Power," by Hull. Proc. I.R.E., April, 1923.

The filament heats the magnesia cylinder and this in turn heats the thimble cathode which emits the required electrons. By making the magnesia cylinder of small heat capacity the time required to make these tubes function has been reduced to about 10 seconds. The two holes in the magnesia cylinder are put very close together so that the *alternating magnetic field* set up by the one turn filament may be kept negligibly small. The filament is designed for a small *IR* drop so that the *alternating electric field* surrounding the filament may be small.[1]

Effect of Plate Current on Filament Current.—It is shown in Fig. 7 that a tungsten filament does not give appreciable emission until it is very hot, so that we may have conditions as shown in Fig. 22; the arrows indicate the direction of electron flow. With a plate current of 0.5 ampere the tube in question has a current of 3.3 amperes at one end and 3.8 amperes at the other end, as indicated in the diagram. End *B* of the filament is at a much lower temperature than end *A*, and is contributing but little of the plate current, as the emission is too low. End *A*, on the other hand, is furnishing most of the plate current and is also being operated at much too high a temperature. With a tube as shown in Figs. 20 and 21, the filament proper suffers no loss of electrons, so has the same current throughout its length.

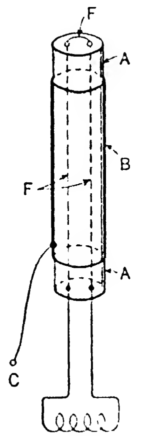


FIG. 21.—The modern form of indirectly heated cathode.

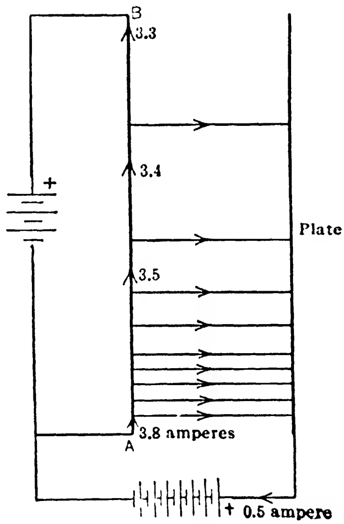


FIG. 22.—The emission of electrons from various parts of the same filament differs very much; because of the current to the plate, the filament current (hence filament temperature) is much greater at one end of the filament than at the other.

When using alternating current for heating the filaments it is possible to use a connection scheme which largely obviates the trouble just referred to; Fig. 23 shows how this is done. As filaments generally are com-

[1] For an analysis of the disturbances introduced into a radio receiver by the alternating current filament see "Analysis and Reduction of Output Disturbances Resulting from A.C. Operation of Heaters of Indirectly Heated Cathode Triodes," McNally, I.R.E., Aug., 1932, p. 1263. He discusses the relative intensities of disturbances due to electric and magnetic fields and also means for minimizing them.

paratively low voltage the a.c. power supply ordinarily requires a step-down transformer and the plate circuit is connected to the filament circuit at the midpoint of the transformer secondary. With the polarity reversal of this secondary each end of the filament in turn carries the larger current (lower end of the filament in Fig. 22) thus bringing about a much more uniform heating of the filament.

Variation of Emission with Filament Current.—Curves Showing Space Charge Effects.—The variation of emission with filament temperature is indicated in Eqs. (1) or (2), but as the experimenter generally has no means of measuring the temperature of the filament, he measures the filament current instead; the curves given in Fig. 24 show how the emission varies with filament current. In these curves is also shown the effect of space charge limiting the plate current. It is evident that the filament used in this tube gives practically no emission with currents less than 1.0 ampere. With a plate voltage of 100 the plate current rose rapidly with increase in filament current reaching 135 milliamperes at a filament current of 1.40 amperes.

When the plate voltage was dropped to 50 and the same variation of filament current carried out the plate current reached a value of only 48 milliamperes at $I_f = 1.40$ amperes. With plate voltages of 20 and 5 the maximum plate currents were 10.6 milliamperes and 1 milliampere respectively. Now with $I_f = 1.40$ the emission is 135 milliamperes as shown in curve 1; with the plate at a positive potential of 5 volts (with respect to negative end of filament) only 1 milliampere was obtained, that is, only 1/135 of the electrons emitted by the filament reached the plate, the rest re-entering this filament *due to the space charge overcoming the comparatively weak field from the plate.*

Speaking in terms of the idea depicted in Fig. 17 we can say that the lines of force from the plate penetrated but a short way into the electron atmosphere; the great mass of the emitted electrons which, it must be remembered, stay very close to the filament, never feel the tractive effect of the positive plate. Those few having exceptionally high outward velocity (due to their velocity of emission and suitable collisions with the other electrons in the electron atmosphere) reach the outer regions of the atmosphere and so get attracted to the plate.

Even for the lower values of filament current (Fig. 24) the four values of plate voltage do not give the same plate current as might be expected. This is due to the fact that the IR drop in the filament is appreciable; in the special tube pictured in Fig. 20 all curves coincide in the lower parts.

In comparing the curves of Fig. 24 with those of Fig. 15 it is to be noticed that although they have the same general shape they have entirely different meanings. In Fig. 15 the flat parts of the curves indicate that saturation current has been obtained and in the lower curved portions the space charge is limiting the current; in Fig. 24 the lower curved portions indicate that saturation current is flowing and the upper flat parts indicate that space charge is limiting the plate current.

The Three-electrode Tube or Triode.—The three-electrode tube differs from the two-electrode tube just analyzed in that a third electrode (called the *grid*, because of its ordinary form) is employed to control the plate current. In its normal form the grid is a metal mesh of some kind interposed between the plate and filament; the electrons passing from the filament to the plate have to go through the holes in the grid mesh and their passage to the plate is controlled to any desired extent by the potential of the grid with respect to the filament. In this form of tube therefore the plate current may be controlled by three factors, the filament current, the grid potential and the plate potential.

The control electrode, or grid, in the ordinary form of tube as invented by De Forest, is inside the tube, directly in the path of the electrons traveling from filament to plate. It is possible, however, to use a control electrode outside the tube, although it seems as though this kind of control offers more difficulties than the internal grid. In one type of outside grid tube tried by the author the control worked for a few seconds and then the accumulation of electrons on the inside walls of the tube made it stop functioning. The inner walls of the glass must be made partially conducting to prevent this accumulation of charge.

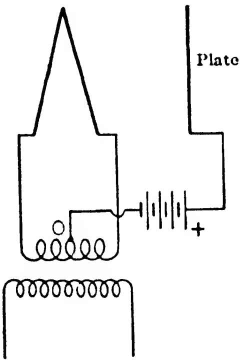


FIG. 23.—When using alternating current for filament excitation the plate circuit is generally connected to a middle tap on the filament winding of the transformer.

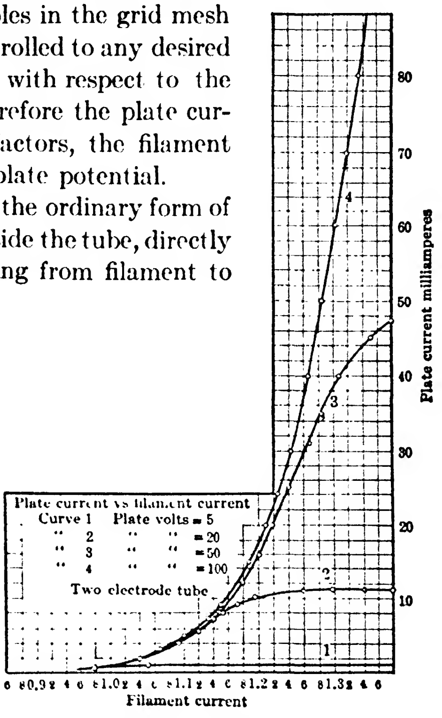


FIG. 24.—Variation of plate current in a diode as filament current is varied; for any one curve the plate voltage was held constant.

The inside control electrode need not be placed between the filament and plate; it will work to some extent even if it is on the side of the filament opposite to that on which the plate is situated. Its action in such a tube is not as efficient in controlling the plate current as in the normal placement; in the analyses to follow it will be supposed that the grid is

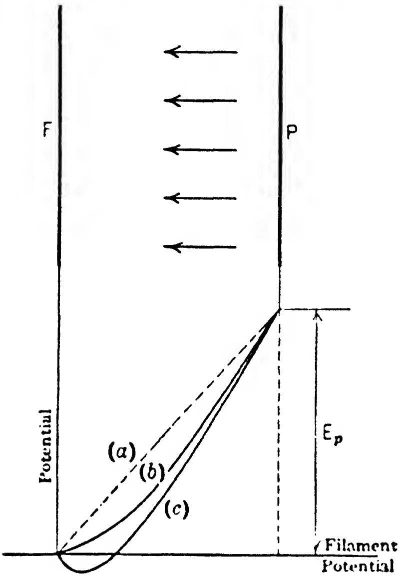


FIG. 25.—Two metal plates, one E_p volts higher potential than the other, have a uniform potential gradient between them, the potential being about as shown by dotted line a; if plate F is covered with an electron atmosphere the potential is changed to the form shown by line b, or even that shown by line c. In case the plate F is a filament, thus having a very small surface compared to the plate P, nearly all the voltage drop will occur close to the filament, instead of occurring uniformly as shown by dashed line a. The slope of the potential gradient close to the filament may be 50 or 100 times the average slope.

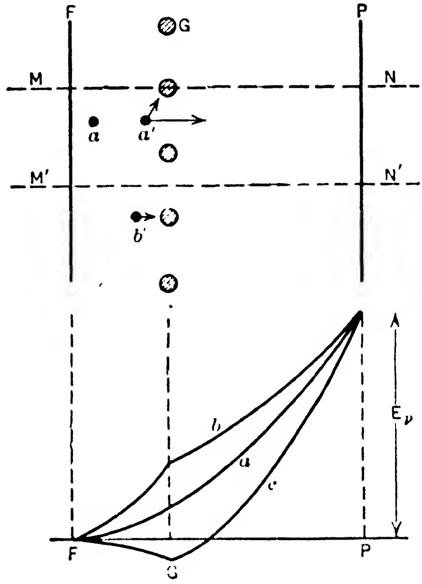


FIG. 26.—In a three-electrode tube the potential distribution between the filament and plate may be as shown by either b, a, or c, according to the potential of the grid.

inside the tube between the filament and plate and the curves given to illustrate the text will be records obtained from such tubes.

Potential Distribution in the Three-electrode Tube.—The three-electrode tube functions because of the effect of the grid on the potential distribution between the filament and plate; it is therefore necessary to have a clear idea of this potential. In Fig. 25 is shown by the dotted line (a) this potential distribution between two metal plates, one marked F to represent the filament, the other marked P to represent the plate. The filament is supposed at zero potential and the plate at positive potential E_p. With a uniform field distribution as shown in the upper part of the figure the potential between plate and filament falls off uniformly. In the actual tube such a uniform potential gradient does not obtain; owing to the comparatively small surface of the filament the potential falls more rapidly near the filament than near the plate.

If we now suppose an electron atmosphere to cover the surface of F the potential distribution is changed to some such form as indicated by the full line (b) in Fig. 25. The potential gradient becomes much lower in the vicinity of F because most of the field of P ends on electrons in the vicinity of F and so never reaches F; in fact if the emission is much greater than the plate current (practically always the case with three-electrode tubes in normal operation) the potential very close to the surface of F may even become negative, the potential distribution being then as indicated by curve c of Fig. 25. In Fig. 26 is represented a filament F, plate P, and grid G (shown in cross-section by the small circles) in the lower part of the figure is shown by the line marked (a) the potential distribution between P and F without any action from G; the curved form of this line is caused by the electron atmosphere around F. It must be remembered that most of the electrons emitted are very close to F and re-enter F without having moved very far toward G. The potential gradient in which the great majority of the electrons lie (close to F) is very small, hence they experience but little tractive effort from P.

If now G is made positive the potential distribution is changed to the line marked b in Fig. 26. The potential gradient between G and F has been much increased so that many of the electrons which previously fell back into the filament will now move toward G. Referring to the upper part of Fig. 26, electron a, which, without positive grid, would have fallen back into the filament, now moves toward G, and so is found in some such position as a'. In this position it experiences two attractions, one from G and one from P. Because of the relatively higher potential and larger surface of P most of the electrons which arrive at this position will move to the plate, instead of going to G as might be supposed. There may arrive at position a', however, an electron which has some velocity in the direction of G; the result of this velocity and the two attractions from G and P may result in its going to G instead of P. Other electrons moving from F toward G may find themselves in such a position (with respect to the grid wires) as shown by b'; these electrons will almost surely go to the grid instead of to the plate.

We may therefore conclude that the interposition of a positively charged grid between the filament and plate will partially neutralize the effect of the space charge; more of the electrons emitted from the filament will move away from it, some of them going to the grid and some going to the plate. A positive grid then increases the plate current, plate potential remaining fixed.

A negatively charged grid will result in a potential distribution somewhat as shown by curve *c* of Fig. 26; if the grid is negative to the extent shown by the curve the plate current will be reduced to practically zero, because none of the electrons (except a very few which are emitted with exceptionally high velocity) can move against the negative potential gradient between *F* and *G*. It must be noticed of course that the potential curve on such a line as indicated by *M*–*N* (Fig. 26) will be different from that on such a line as *M′*–*N′*; the grid potential will not be so effective on the line *M′*–*N′* as on a line lying closer to one of the grid wires.

It will be appreciated at once that this effect of the grid in controlling the flow of electrons to the plate will depend on various features of construction of the tube. The grid will exercise the most control when its wires are very fine and close together, and when it completely surrounds the filament. Unless the grid is considerably larger (in length and breadth) than the space occupied by the filament many of the electrons will go from the filament *around* the grid and thus arrive at the plate without having been subjected completely to the controlling action of the grid.

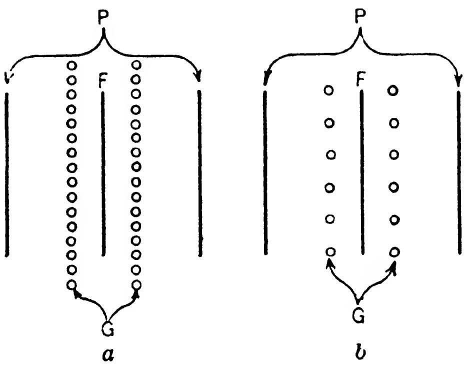


FIG. 27.—The construction shown in *a* will give the grid *G* much greater controlling action than that shown in *b*; the more completely the grid encloses the filament and the finer its structure the greater will be its controlling action.

This idea is illustrated in Fig. 27; the construction shown in *a* will permit the grid to exert a much greater control over the electron stream than will the construction shown in *b*.

Model of Triode.—It is possible to build a hydraulic model of a three-electrode tube which illustrates very well the general ideas involved in the tube action. A jar (Fig. 28) (such as glass storage-battery container) has placed in the lower part a pipe *A*, closed at its two ends, which is full of small holes on its lower side and is connected to an air supply of very low pressure. A rubber sheet (such as the rubber used by dentists) is fastened to the side of this pipe *A* and also to a rod *C* in the upper part of the jar, horizontal and parallel to *A*. To make the model simple only one-half of the three-electrode tube is represented; a metal sheet *E* fastened to *A* makes all the air bubbles which escape move to the left (in Fig. 28) and so run up on the under side of the rubber sheet and escape past *C*.

This stream of bubbles represents the electron stream from a filament, A being the filament and C the plate, C being at higher level than A, as must be the potential of the plate with respect to that of the filament.

A stick D has several parallel wooden pins fastened to it, the lower ends of which are fastened (by tacks) to the rubber sheet close to pipe A, as shown. When D is moved up and down, the lower ends of its pins lift up and down those parts of the rubber sheets to which they are attached; in Fig. 29 is shown a sketch of the rubber sheet with the bar D lifted, and in Fig. 30 is shown the form of the rubber sheet when the bar D is depressed. If the pressure of the air in the pipe A is properly adjusted, the flow of air bubbles up the under side of the rubber sheet resembles (more closely than any analogy the author has seen) the flow of electrons in a three-electrode tube.

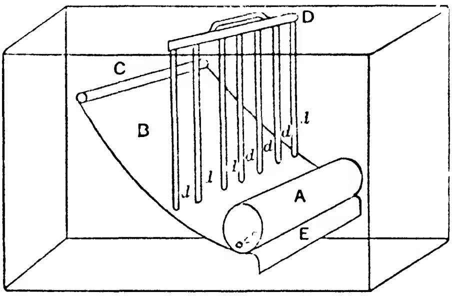


Fig. 28.—Hydraulic model of the three-electrode tube.

The action of the bar D with its attached pins, producing small hills and valleys in the rubber sheet, illustrates well the action of the grid. The depression of the pins, making it more difficult for the air to pass up along the sheet, illustrates a negative grid, and when the pins are

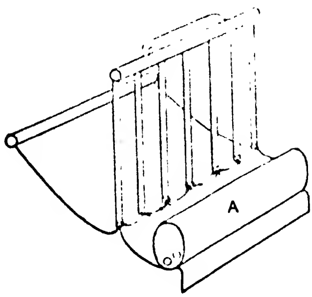


Fig. 29.—Hydraulic model of the three-electrode tube.

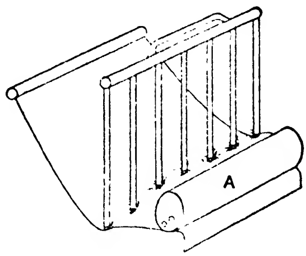


Fig. 30.—Hydraulic model of the three-electrode tube.

lifted up the increased flow of air corresponds to the increased plate current with positive grid. By having the pins, d, d, etc., in the form of tubes open at their lower ends and having corresponding holes in the rubber sheet, some of the air bubbles will run up these tubes when handle D is lifted, thus imitating the action of the positive grid attracting some of the electrons, causing grid current.

The effect of the space charge is not simulated very well by the model; the accumulation of air between the " grid " (row of pins, d, d, d) acts to prevent other bubbles of air coming through the small holes in the " filament " (pipe A) but this action is not strictly analogous to the mutual repulsion of the electrons in the actual space charge effect.

Fields of Use of Three-electrode Tube.—Detector, Amplifier, Generator or Converter, Modulator.—The three-electrode tube was first used as a detector of radio signals from spark stations; it was much more sensitive than its competitors, the magnetic detector, Fleming valve, etc., and so rapidly displaced these as a detector. In its original form as manufactured by De Forest a potential of about 30 volts was used on the plate; the normal plate current was a few hundred microamperes. Although these original tubes were rather erratic in their behavior, and not uniform in their characteristics (one tube not being like another) by careful adjustment of filament current and plate voltage, they were nearly as good detectors as the present types.

As the grid potential of a three-electrode tube controls the plate current (the power for which is supplied by a local battery) it is evidently applicable as a relay, the signal voltage controlling the delivery from the local power supply. When properly adjusted the grid circuit takes an extremely small power to operate, so that compared to the amount of power used in the grid circuit the amount controlled in the plate circuit may be thousands of times as great.

If the grid circuit is adjusted to take no power itself the power amplification is infinite; it must be remembered, however, that to operate the grid circuit certain coils, condensers, and resistances are required; taking the losses in these necessary associated circuits into account the power amplification is not infinite, but it is very large even then. Thus a certain tube used in telephone circuits as an amplifying repeater has a power amplification of about one thousand times.

If an alternating potential difference is impressed on the grid of a tube the plate current periodically increases and decreases. Thus pulsating current in the plate circuit may be made to produce fluctuations in the grid potential by means of a suitable transformer, the primary of which is connected in the plate circuit and the secondary connected between the filament and grid. If a suitable condenser is connected across either the primary or secondary winding to give a natural period to the circuit, the fluctuations in the plate current will be maintained by their action on the grid potential.

With this arrangement then the plate current fluctuates between certain maximum and minimum values, the voltage of the grid alternates, and in the condenser (no matter which circuit it is connected with) an alternating current flows. The device thus becomes a generator of a.c.

power; it might perhaps be more properly called a converter for changing c.c power into a.c. power. The frequency of the alternating current is fixed by the L and C of the oscillatory circuit, and the amount of power available depends on the average value of the plate current and the voltage of the battery or generator supplying the current.

A triode may be operating as a converter, to produce high-frequency currents for radio signaling. Now it may be desired to vary the amount of r.f. power in accordance with voice waves, that is to carry in radio telephony. The minute power of the voice waves is amplified suitably and then by means of a second triode suitably connected in the plate or grid circuit of the first, this voice energy controls the output (perhaps many kilowatts) of r.f. power. A triode thus used, to control the output of another, is called a modulator.

Various Types of Tubes.—According to the purposes for which they are to be used many different types of tubes have been evolved. Tubes designed for detecting high-frequency currents need to have a power output of only a very small fraction of a watt; they are generally fitted with small filaments, because but little emission is required and the voltage used in the plate circuit is low. Typical tubes use a filament current of 0.25 ampere at 6 volts and use a plate battery of about 45 volts. Tubes used for amplifiers are more generally higher plate voltage, perhaps 100 or 200 volts; the size of filament is about the same as used for a detector tube. Tubes used for generating power are designed for higher plate voltage, from 300 to 20,000 volts; as the amount of power available depends upon the value of plate current and this in turn upon the emission, the filament is much larger than in the amplifier and detector tubes. A 5-watt tube (output) might require a filament current of 1.0 ampere at 6 volts; a 500-watt tube might require 6 amperes at 20 volts. Later a tabulated list of ratings for various tubes will be given.

The grids used vary from a very fine mesh of the finest tungsten wire obtainable (wound 40 per cm.) to a lattice work of comparatively coarse wire spaced about 3 per cm. The grid may be flat or cylindrical according to the form of tube.

The plates used are of various forms; they vary from a short zigzag-shaped tungsten wire perhaps 5 cm. long, or a small thimble about 0.5 cm. in diameter and 0.5 cm. long to two heavy plates about 5 cm. square. The material used for the grids and plates of air-cooled tubes is generally nickel or tungsten or molybdenum; the air-cooled tubes designed for generating much power are likely to have all of their metal parts, filament, grid, and plate of tungsten.

In Fig. 31 are shown some of the more common tubes: A and B are power tubes of 50- and 250-watt ratings, respectively; C and D are small-power tubes designed for an a.c. output of about 4 watts, E, F and G

serve as either detectors or amplifiers; *H* is a De Forest audion of the original type, *I* is a cylindrical tube made for amateur use; *J* is one type of Marconi tube; *K* and *L* are two amplifying bulbs, the latter having extremely fine grid and very smal lplate (a nearly invisible zigzag wire); *M* is a special tube with grid brought out at tip of the bulb. The former tubes are of American manufacture; at *N* is shown an English power tube, at *O* a special French amplifier bulb and at *P* a small English detector and amplifier tube. At *Q* is shown a special type of tube called a dynatron, explained on p. 655.[1]

Some tubes are designed with hot cathodes in the form of filaments, carrying continuous current or alternating current. Many are built with the electron-emitting surface in the form of an oxide-coated cylinder, which

FIG. 31.—Various types of air-cooled tubes.

is indirectly heated (Fig. 21). The tubes of a radio receiver for the most part have to handle very minute amounts of power, measured in microwatts, or micro-microwatts. It is evident that such tubes can be designed in any desired manner, with practically no regard for plate heating, etc. The last tube of the series, however, the so-called "output tube," must furnish sufficient power to operate a loud speaker, from a few milliwatts to a few watts depending upon the type of program. (An average loud speaker volume is obtained with 50 milliwatts input to the speaker.) This output tube may therefore use several watts of power, so that temperature rise of internal tube parts must be considered.

[1] For a historical sketch of the change in tube design during the past ten years see "Recent Trends in Receiving Tube Design," by Warner, Ritter, and Schmit, I.R.E., Aug., 1932, p. 1247.

Limits of Operation of a Tube.—There are in general two limiting factors in the use of a vacuum tube—overheating and consequent collapse of the parts or of the bulb itself, and ionization of the residual gas in the tube. It is impossible to completely evacuate a tube so that some gas is always present; if too high a plate potential is impressed or too high a filament current (with fairly high plate voltage) is used this residual gas will ionize sufficiently to change the operating characteristics of the tube by an amount depending upon the amount of gas present.

With a tungsten filament tube the evacuation process is carried out more thoroughly than with the oxide-coated filament so that destructive ionization is not likely. The limit of the tungsten filament tubes (aside from the prescribed limit for filament current) is the safe heating of the plate and grid, generally the plate, because the grid circuit is so adjusted that the grid takes but little current. This heating is due to the power, used in accelerating the electrons as they move from the filament to the plate, being given up when the electrons are stopped by hitting the plate; the phenomenon is called *electron bombardment*. The amount of power so used on the plate is equal to the product of the plate voltage and the plate current; if this product varies cyclically (as it actually does when the tube is being used for power converter), its average value must be taken in calculating the amount of power used in bombarding the plate.[1]

With many of the tubes in common use today a thoriated tungsten filament is used instead of pure tungsten. As was explained on p. 467 if this type of filament is too heavily bombarded by positive ions the layer of thorium is disintegrated and the filament ceases to give appreciable emission. The tube is thus made inoperative, not due to excessive heat of either filament or plate, but to change in character of the filament surface.

The safe power to be used in bombarding the grid is much less than that for the plate, for two reasons; the surface of the grid is generally much smaller than that of the plate, and the possibility of heat radiation from the grid is less than that of the plate.

The large tube shown at *B*, Fig. 31, has a rating, for example, of 250 watts plate and 25 watts grid. Thus a plate current of 0.25 ampere (steady value) would be permissible with a plate voltage of 1000 volts, and with this amount of power used in the tube the plate becomes quite a bright red color. The two tubes shown at *C* and *D* have a safe plate

[1] A bombardment equivalent to 10 watts per square centimeter of a smooth plate will bring its temperature to about 1300° C.; such a temperature gives the plates a fairly bright red color. By properly roughening and darkening the plates to increase the radiation coefficient, much more power will be sent off at the same temperature. To keep the plate of an air-cooled tube down to a safe operating temperature, however, not more than 5 watts per square centimeter of plate should be allowed; too high a plate temperature liberates gas from the plate metal and spoils the tube.

capacity of 12 watts; with a plate voltage of 300 (their rated value) the average plate current should not exceed 40 milliamperes.

In Fig. 32 is shown the largest commercial American air-cooled tube; the glass bulb is 6 in. in diameter. The plate of this tube is large enough to dissipate 1000 watts of power, in addition to the 210 watts of filament power, practically all of which has to radiate from the plate.

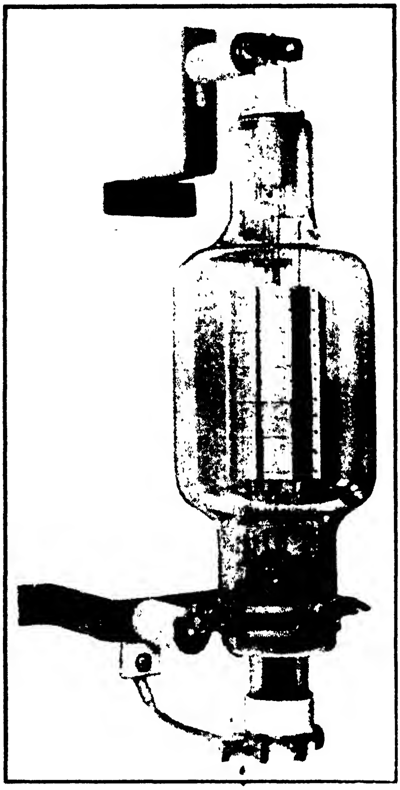


Fig. 32.—The largest air-cooled tube at present produced commercially in the United States.

Filament Voltage	10 volts
Normal Filament Current	21 amperes
Average characteristics on plate voltage of 3000 volts and grid bias of −200 volts	
Plate Current	.350 amperes
Average Plate Resistance	1800 ohms
Amplification Constant	10
Approximate Direct Interelectrode Capacities	
Plate to Grid	17.9 mmf.
Plate to Filament	7.9 "
Grid to Filament	15.4 "
Maximum Operating Plate Voltage	3000 volts
Negative Grid Bias for above Plate Voltage	275 volts
Continuous Plate Dissipation	1000 watts
Peak Plate Dissipation	1200 watts
Maximum Overall Length	21$\frac{11}{16}$ ins.
Diameter of Bulb	6 ins.

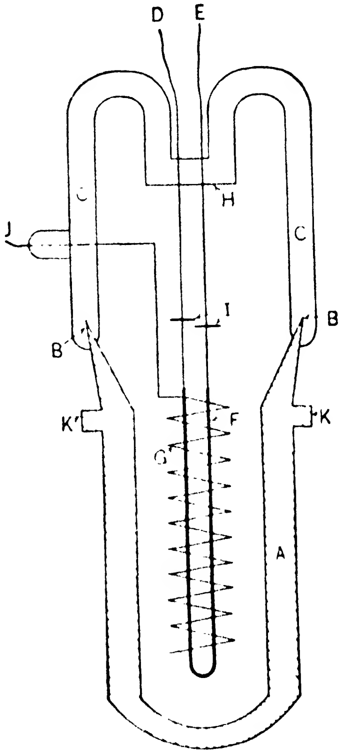


Fig. 33. Conventional cross-section of a modern water-cooled tube; the copper tube A (which constitutes the plate) is sealed to the glass part C C, at the feather edge B B. This construction permits an air-tight seal. The copper tube is about $\frac{1}{32}$ in. thick.

Water-cooled Tubes.—In attempts to increase the possible power output of a triode the heating of the plate constitutes the limiting factor.

Not only the plates themselves become prohibitively hot, but the glass bulb itself becomes so hot from the heat radiated from the plates that it is likely to collapse. Bulbs of fused quartz have been used, but even this expedient permitted an output of only 1 or 2 kilowatts.

The real step in advance was made when the plate was used as the wall of the vacuum tube itself; in this design it is possible to submerge the cylindrical plate in cooling water and where only a few hundred watts could be dissipated from the interior type of plates, the external plate can carry off a great many kilowatts, and the possible power rating of triodes was extended into hundreds of kilowatts.

Fig. 33 gives a conventional cross-section of the modern water-cooled tube. The plate of the triode is in the form of a heavy copper cylinder, A, open at the top and drawn out to a feather edge B–B. The heavy glass tube, C–C, is sealed to the tubular plate at B–B, the joint being air tight because of the low tensile strength of the thin edge.

This special metal-to-glass seal was the one factor that had previously held up the development of high-power tubes. Ordinarily a copper-glass seal will break when cooling due to the different expansion coefficients. But by making the copper mechanically weak at the sealing edge, B–B, it does stick to the glass in spite of the force tending to break it loose.

The filament, F, is generally a spiral of tungsten wire, led into the tube at the airtight seal H. As these filament wires are necessarily heavy, the metal-to-glass joint here also is made in the special manner, using a feather-edged cone of copper to seal to the glass. This small cone is welded to the filament wire (which passes through the apex of the cone) in an air-tight joint. To prevent the intense heat radiation from the filament, overheating the joint at H, little baffle plates, I, are fastened to the filament lead-in wires. The grid G, in the form of a spiral, surrounds the filament, and its connection (for the outside circuit) is brought out of the glass tube at J. In use the tube operates in a vertical position, the plate A fitting into a small tank through which cold water is rapidly circulated. The collar, KK, serves to support the tube when placed in its cooling tank.

The very high voltage between plate and filament would be difficult to control, it would seem, because of the water-circulating system. However, by making the water connections between the cooling tank around the plate and the water supply (which is ordinarily grounded) of long rubber hose, the amount of leakage current from plate to ground is only a few milliamperes, in spite of the fact that the plate is generally about 10,000 volts above ground. It might of course seem better practice to ground the plate, and so let the filament operate at a negative potential of about 10,000 volts, but apparently this scheme is not desirable.

The water-cooling feature makes it possible to use a very compact design for these tubes; a 10-kw. tube for example might have an overall length of 18 in. the plate A being $2\frac{1}{2}$ in. in diameter and 10 in. long, of $\frac{1}{8}$ in. copper. The filament might use 50 amperes at 20 volts and the normal voltage between plate and filament might be 15,000 volts.

The largest commercially produced triode of this type is shown in

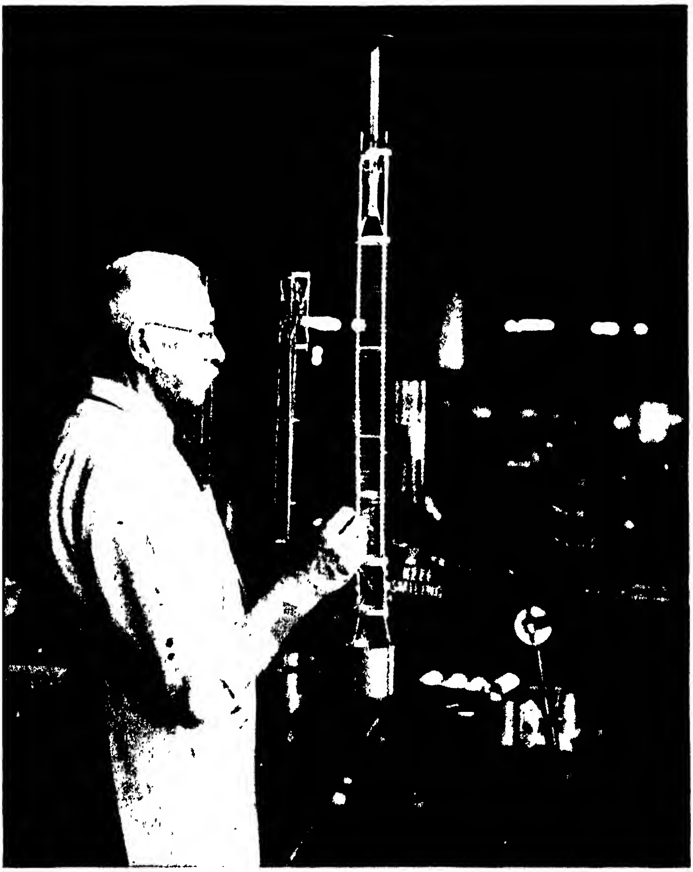


FIG. 34.—Grid assembly for a 100-kw. water-cooled triode.

Figs. 34 and 35.[1] The spiral grid, inside which the hairpin filament assembly is to be mounted, is shown in Fig. 34; rigidly fused to the grid support is a heavy quartz rod, projecting above the grid in Fig. 34. (When the triode is assembled this rod is *below* the grid and filament.) On top of this rod is a molybdenum disc which is slightly smaller than the inside

[1] Radio Engineering, Dec., 1931.

diameter of the tubular plate. When the grid-filament structure is inserted into the plate (Fig. 35) this disc serves to hold the lower end centrally in the copper tube.

The copper-to-glass special seal has already been made, so that the last step in the assembly is the fusing together of the two pieces of glass tubing, as indicated in Fig. 35. After this is done the tube is connected

FIG. 35.—Last step in the assembly of the 100-kw. triode the grid of which is shown in Fig. 34.

to the pump system and is properly evacuated. In the accompanying table are given all the mechanical and electrical constants of the 100-kw. tube and for comparison the corresponding contents of a 10-kw. water-cooled tube.

A somewhat later design of high powered tube uses a double-ended construction. The water-cooled plate is 4 in. in diameter and 26 in. long; both ends of this copper cylinder are machined to a thickness of only

0.003 in. for sealing to the glass ends. The grid of this triode is water cooled, as well as the plate, and it is capable of generating 200 kw.[1]

GENERAL INFORMATION	UV-862	UV-207
Filament voltage	33.0	22.0 volts
Current	207.0	52.0 amps.
Type—Tungsten		
Average characteristic values:		
Amplification factor	48	20
Plate resistance	2,800	3,500 ohms
Grid-plate transconductance	17,150	5,700 μ mhos.
Approximate direct interelectrode capacities:		
Plate to grid	80	27 $\mu\mu f$
Grid to filament	52	18 $\mu\mu f$
Plate to filament	2	2 $\mu\mu f$
Maximum overall dimensions:		
Length	$60\frac{3}{4}$	$20\frac{1}{4}$ inches
Diameter	$6\frac{1}{8}$	$4\frac{5}{32}$ inches
R.F. POWER AMPLIFIER—CLASS B:		
Maximum operating d.c. plate voltage	20,000	15,000 volts
Maximum unmodulated d.c. plate current	5	1 amp.
Maximum plate dissipation	100,000	10,000 watts
Maximum r.f. grid current	60	30 amps.
Typical operation:		
Plate supply voltage	18	12 kv.
Unmodulated d.c. plate current	4.20	0.90 amp.
Peak output	100,000	14,000 watts
Carrier output—Mod. factor 1	25,000	3,500 watts
OSCILLATOR AND R.F. POWER AMPLIFIER CLASS C:		
Maximum unmodulated d.c. plate voltage	20,000	15,000 volts
Maximum d.c. plate current	10*	2 amps.
Maximum d.c. grid current	1	0.2 amp.
Maximum plate dissipation	100,000	10,000 watts
Maximum r.f. grid current	60	30 amps.
Typical operation:		
Plate supply voltage	18	12 kv.
Output	100,000	15,000 watts

* If plate modulation is used, divide this figure by (1 + modulation factor) to get the maximum d.c. plate current.

It is reported that in England a 500-kw. water-cooled tube has been built.[2] This tube does not have an air-tight seal between its metal and insulation portions, but is always used with a diffusion pump in operation, to keep its vacuum as low as necessary. The diffusion pump uses

[1] Described by Mouromtseff in I.R.E., May, 1932, p. 783.

[2] Radio Engineering, April, 1932, p. 17.

oil instead of mercury and requires no liquid air trap to make it function properly. (See p. 499.)

The amount of water required for these tubes depends of course upon the temperature of the water and the power being used on the plate. From 12 to 25 gallons per minute for a 100-kw. tube is present practice. The water must not have much dissolved mineral matter; its resistance should be at least 4000 ohms per cubic centimeter.

With an active circulation of water the plates of a water-cooled tube can dissipate 70 to 80 watts per square centimeter of plate.

Special safety devices must be used with these tubes, due to the fact that if the water supply fails for a few seconds, the plate will be melted, so great is the energy of electron bombardment.

In these power tubes the grid connection, where it comes out through the glass, must be kept well away from the filament lead-in-wires; thus it would be a poor design in Fig. 33 to bring the grid lead, J, out through the same place, H, where the filament wires go through. The hot glass between grid and filament leads, for such construction, would electrolyze when the tube was in operation and this would result in leaky joints.

Effect of Gas in a Vacuum Tube.—Ionization.—The modern vacuum tube is a true electron relay; it functions entirely by means of the stream of electrons emitted from the filament, and these electrons in motion constitute the only current in the tube. This ideal is not quite realized by any vacuum tube, but it is so nearly approached that whatever other current may exist is so small as to make its effect negligible when considering the action of the tube.

The earlier types of vacuum tubes (Fleming valves and Deforest audions) were not at all well evacuated in the light of modern practice; there was a deal of gas left in the bulb at the completion of the evacuation process and this gas made the tubes very erratic and undependable in their behavior.[1] Not only would various bulbs, supposedly similar, have very different characteristics, but any one bulb would not act consistently, and many tricks had to be employed to make the bulbs perform to the best advantage.

An exact study of the effect of gas in a vacuum tube cannot be given here; only those points which bear directly on the operation of the tube in radio practice will be outlined. The student is referred to some such book as Thomson s "Conduction of Electricity through Gases" for a more thorough analysis than will be attempted here.

A cold electrode in a vacuum tube, unless subjected to considerable

[1] It is quite evident, however, that Fleming appreciated the necessity of a high vacuum to make the tubes constant in behavior; the superiority of present evacuation is due not so much to any conception of its importance, perhaps, as to the better pumps now available.

electron bombardment, will not give off electrons in appreciable quantities; thus in a two-electrode tube if the plate is made negative with respect to the filament no current will flow, because if the plate is made negative any current which flows from plate to filament must be caused by electrons leaving the cold plate. Experiment demonstrates the truth of this statement; if other possible carriers of current are eliminated (such as actual leaks inside or outside the tube, or gas inside the tube) the amount of current which will flow is too small to be measured unless excessive voltage gradients are employed. (See Figs. 1 and 2 of this chapter.) We may safely conclude that when a cold electrode (either grid or plate) of a tube shows current in such direction as to indicate electrons flowing from it, inside the tube, the tube has in it gas which is serving as a conductor of current.[1] This statement neglects the possibility of secondary emission of electrons due to excessive bombardment by electrons coming from the filament; this effect will be treated in a later paragraph.

Ordinarily a gas is a good insulator and will not carry current, but when under rather low pressure it may be made to carry very large current if by some means it becomes *ionized*. By this term is meant the breaking up of the normal gas atom into two parts, a free electron and positively charged nucleus; this breaking up of a gas atom corresponds to the "break-down" of any ordinary insulator when it is subjected to too high a potential gradient.

In a Geissler tube the gas becomes ionized (showing the well-known blue glow) only when rather high potentials are used, generally several thousand volts. Now in the vacuum tube used for radio receivers, high voltage is practically never used; ionization of the gas in the tube may occur with voltages as low as 30 or 40. This is due to the fact that the hot filament furnishes the electrons which by their motion (caused by the positive plate potential) serve to start the ionization of the gas atoms. In a Geissler tube no such means is at hand for starting the ionization, hence the comparatively high voltage required to show the effect.

The role played by the electrons emitted from the filament in producing ionization is easily shown by a simple test. If a tube which is known to be faulty is subjected to normal plate potential with cold filament, no plate current will flow and the tube will show no signs of ionization. Now if the filament current is gradually increased, emission of electrons will commence and a slight plate current will flow, at a certain filament temperature, depending upon how much gas there is in the tube,

[1] It must be remembered that even with the highest vacuum obtainable there is still a tremendous number of gas molecules in the evacuated space; it is likely that in highest vacuum tubes used to-day (10^{-8} mm. of mercury) there are of the order of 10^{8} gas molecules per cubic centimeter. In the ordinary vacuum tube used in a receiving set there are about 10^{11} gas molecules per cubic centimeter.

the familiar blue haze will appear in the bulb, accompanied generally by a very large increase in the plate current, thus showing that the filament must be emitting a certain minimum number of electrons before appreciable ionization of the gas occurs.

If but a small amount of gas is present the pale blue glow may be so weak as to be invisible, but the presence of appreciable quantity of gas is generally shown by erratic changes in the plate current.

Some oxide-coated power tubes show a bright fluorescence on the plate when being used, generally in the form of a pattern of the grid. It is easy to mistake this effect for ionization because of the blue color from the fluorescing plate; if the plate is hidden from the eye (by the hand or a piece of cardboard) it will be seen that there is no blue glow in the space inside the tube. The intensity of the effect of fluorescence depends upon the condition of the surface of the plate, which is generally covered with more or less oxide.

Danger to a Tube from Ionization.—When a tube ionizes the consequences resulting depend upon the type of tube being used and upon how quickly the condition is removed. In the case of a detecting tube, or amplifying tube, the state of ionization will generally stop the functioning of the tube, its characteristics being entirely different when the tube is filled with a semi-conductor (the ionized gas) than those of a normal electron tube. If either the plate voltage or filament current is reduced the ionization will disappear and the tube may operate as well (or possibly better) than it did before ionizing.

In the case of a power tube the situation is different; unless the plate potential is immediately reduced the tube may be completely spoiled. Ionization practically never occurs in a tungsten tube because of the high degree of vacuum ordinarily used; the oxide filament tube is much more likely to suffer from it. In these tubes there is always a lot of gas in the metal parts of the tube, filament, grid, and plate.

Now when ionization starts the electrons of the ionized gas travel to the plate, it being positive, but the positive nuclei travel to the filament and subject it to a bombardment. This bombardment results in extra heating of the filament, generally in one spot, which extra heating tends to aggravate itself and burn the filament out at this point. The hotter the filament the greater the electron emission, and also gas is likely to be emitted from the filament at this hot spot; where the gas and electron emission both increase the ionization increases, increasing the bombardment of the filament at this spot, and thus by the cumulative action burning it out. At the time the filament burns out it releases a lot of gas which, becoming ionized, may permit the passage of such a large current from the plate as to result in a miniature "explosion" inside the tube, completely wrecking the parts and breaking the bulb.

When a power bulb with oxide filament once ionizes it is practically valueless [1] until re-exhausted; the ionization itself will probably result in the emission of extra gas from the bombarded parts, so that the tube has much more gas in it after ionization than before.

Evacuation of a Vacuum Tube.[2]—Because of the deleterious effects of gas the electron tube must be very carefully freed from any appreciable quantity of it. With modern pumps the getting out of the gas from the *space* inside the bulb is very simple and rapid but this is not sufficient. Metals, oxides, and glass absorb a deal of gas which gradually comes out; so that a tube pumped " clean " will soon show gas because of its emission from the parts of the tube. This emission is very slow at ordinary temperatures, so that a tube might be pumped a long time without getting sufficient gas from the parts to prevent further emission. If, however, the glass and metal parts are heated, the gas is expelled from them very rapidly, and this is the scheme used in evacuating tubes; the whole tube is subjected to a " baking " process while connected to the pumps.

Even before assembly the metal parts of a tube are baked to a red heat in an atmosphere of dry hydrogen. Hereafter as the assembly proceeds the operators handle the metal only with clean gloves, to prevent contamination of the surface by grease, etc. After assembly, when the completed tube is on the vacuum machine, the metal parts are first heated by eddy currents induced in them by a solenoid carrying high-frequency currents being placed over the tube. The eddy currents get the grid and plate hot, thus rapidly expelling the absorbed gases, and these are immediately removed by the pumps. Finally the filament is heated, at a temperature somewhat greater than normal and the plate and grid again heated (this time by bombardment), thus again improving the vacuum.

This heating during the evacuation process should be carried much higher than any temperature at which the tube may operate; thus if in practice the plates and filament operate a dull-red heat they should be run for several minutes at a bright red heat during evacuation. This overheating of the parts is regularly done with tungsten tubes but it cannot be carried out to the same degree with the oxide-coated filament or thoriated filament tubes. The coated filament is easily spoiled if subjected to too high a temperature, and this limits the possibility of complete evacuation. For this reason, as previously mentioned, the oxide-coated power tubes are

[1] It may be used, however, for generating a small amount of power, providing the plate voltage is kept sufficiently low; thus a 300-volt tube which has ionized badly may sometimes be used by reducing the plate voltage to perhaps 250.

[2] For the modern pump used see article by Langmuir in G. E. Review, Dec., 1916. For general ideas on evacuation see a series of articles by Dushman in Vol 23 of the G. E. Review, 1920. In Vol. 24, pp. 58-68, he treats adsorption by metals; pp. 244-252, adsorption by glass; pp. 669, et seq., the use of phosphorus for "clean-up," and on pp. 810, et seq., he gives the general idea of "clean-up."

much more subject to destructive ionization during operation than are the tungsten tubes.

Recently the engineers of the Bell Telephone Laboratories [1] have found it possible to use oil diffusion pumps in place of mercury diffusion pumps, with greatly improved action. Using two of these special oil diffusion pumps in series, backed by a fore pump good for 10^{-3} mm. mercury, with a charcoal trap at room temperature, they have been able to get, and hold for days, a pressure as low as 10^{-8} mm. No liquid air, or even ice, was found to be necessary. The oil must be one which has a very low vapor pressure at room temperature; in England, Burch has used special oils known as Apiezon " A " and " B," and in America oils of the butyl phthalate class are used. It seems likely that such pumps will be used in the future, operating continuously on high-power tubes while they are in service. Tubes designed for this method of operation do not require special glass-to-metal seals, and may be designed so that burned-out filaments can be easily replaced.

Generally a tube intended for amplifying purposes in a receiving set is finally evacuated to a pressure of 10^{-5}–10^{-6} mm. of mercury (from 0.01 micron to 0.001 micron of pressure). After a tube is used gases are liberated from its parts and this would be deleterious, especially for thoriated or oxide-coated filaments. To care for this a small piece of one of the volatile alkali metals is put in the tube during the process of assembly and, after the tube has been evacuated and sealed up, this is volatilized and condenses on the inner surface of the glass bulb, giving the latter the well-known silver appearance. This freshly deposited metal acts effectively like a sponge for any appreciable gas which appears in the tube, adsorbing the gas as fast as it appears. This " getter," as it is called, thus protects the filament surface from becoming contaminated (poisoned) by gas ions, under which condition its emissive power would be much reduced.

The " getter " scheme for obtaining the final vacuum is employed on all small tubes used in radio receivers, and in general it is used also for power tubes which have oxide-coated, or thoriated tungsten, filaments. For the largest tubes, using pure tungsten filaments, final evacuation is obtained by the pumps, with all parts of the tube, including the glass, as hot as is permissible without danger of collapse; no getter is used in these large tubes.

Detection of Gas in a Three-electrode Tube.—In general gas shows itself by causing erratic changes in the plate current, as either the filament temperature or plate voltage is varied. Thus with a fixed plate voltage and low filament temperature there may be no appreciable ionization, but when the filament current exceeds a certain value, a sudden increase in plate current occurs, and a light blue haze may appear. With the

[1] Becker and Jaycox, the Review of Scientific Instruments, Dec., 1931, p. 773.

appearance of this blue haze the plate current may jump to ten times its proper value. This ionization phenomenon is not constant in its occurrence, as evidenced by Fig. 36, in which are shown three curves from the same tube, one after the other. Curve 1 was taken first; ionization set in with plate potential of 40 volts, causing a large increase in plate current, which value was maintained for one minute. The plate voltage was then reduced to zero and again increased, and with same filament current as before, ionization set in at 60 volts, indicating that during the maintenance of the ionization current previously, some of the gas has been occluded in the glass walls of the tube or elsewhere. This idea is substantiated by the fact that when ionization did set in (somewhat above 60 volts) the current jumped to only 1000 microamperes, whereas previously it had gone to 1200 microamperes. In a short time the ionization ceased, as indicated by disappearance of the blue haze and decrease in plate current to an even lower value at 66 volts than it had at 60 volts before ionizing. Upon again increasing the voltage the current followed the values shown. Upon dropping the plate voltage once more to zero and going through the same range as before the plate current varied as shown by curve 3, no ionization at all occurred. This action is quite typical of tungsten filament tubes; they tend to clean themselves of any gas present in the bulb.

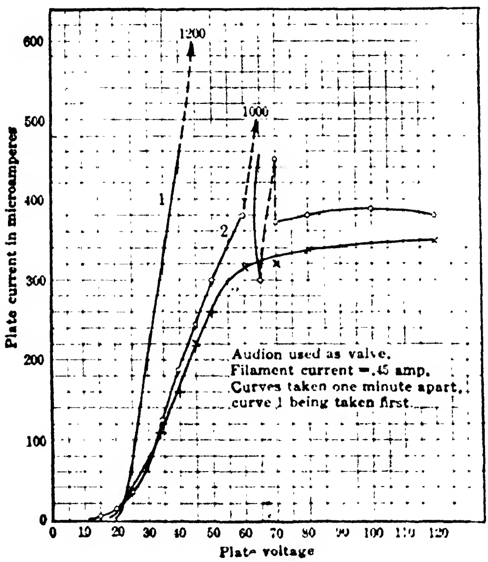


Fig. 36.—Disappearance of gas from a tube; curves were taken in the order 1-2-3; ionization showed on the first curve, to a lesser extent in the second and not at all in the third.

In Fig. 37 is shown a peculiarity of a tube having a small amount of gas present; a kind of "hysteresis" cycle occurs, the current not going through the same values for decreasing plate voltage as for increasing plate voltage. This was probably caused by a change in the surface of the filament. At voltages higher than 15 this tube showed a drooping current-voltage curve, which means that its a.c. resistance (for limited values of impressed alternating e.m.f.) is negative; as long as it held this characteristic, this tube might be used as a two-electrode tube for producing oscillations, its operation being the same as that of a Duddell singing arc.

The normal variation between plate current and grid voltage in a three-electrode tube gives smooth curves, but if gas is present abnormal shapes may be obtained. Fig. 38 shows an effect of this kind and for each of the plate voltages used a "hump" occurs in the plate-current curve. The position of this hump shifts to different grid voltage for the different plate voltages used in the test.

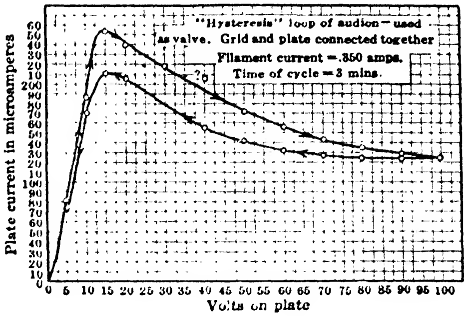


FIG. 37.—In this tube the effect of the gas present was to so alter the emitting properties of the filament that the saturation current was appreciably different with increasing and decreasing plate voltages, showng probably change in emissivity of the filament.

In Fig. 39 is shown a more striking example of this same peculiarity. The curve is for a well-pumped modern tube using tungsten filament; it is undoubtedly due to the presence of mercury vapor in the tube. If a mercury-vapor pump is used for evacuation, some of the mercury vapor will be left in the tube unless a proper freezing trap is used. If a tube showing this effect is used for a detector of radio signals it is remarkably sensitive if adjusted to just the right grid potential by a suitable potentiometer.[1]

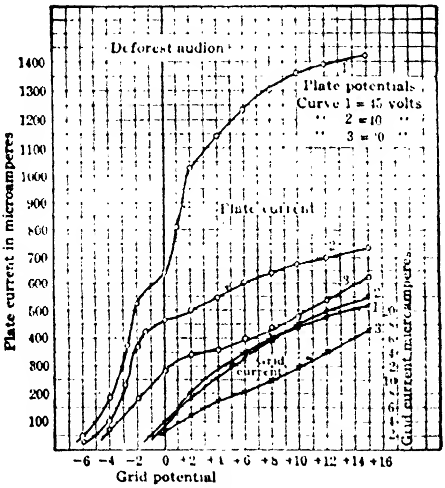


FIG. 38.—A small amount of gas in a three-electrode tube may produce more or less regular "humps" in the plate-current curve.

If a tube is completely freed from gas the current to the grid will not reverse when the potential of the grid is made negative. Even in a very well-pumped tube, however, there is a slight reversed current to the grid when the grid is negative, caused by the positive ions of gas in the tube. This grid current depends

[1] It must be remembered that when the tube is subjected to *very high-frequency variations* in its potential it is quite likely that the plate current does not vary in the manner indicated by the curve obtained in d.c. test, such as that given in Figs. 37 and 39.

upon the gas present being ionized by the electron flow to the plate and is zero if the electron flow is zero. The more plate current there is the more is the gas ionized and hence the greater is the grid current.

The effect is shown in Fig. 40, which shows the grid current in a well-pumped 250-watt tungsten tube; it is seen that for plate voltage of 100 the reversed grid current is much less than it is for a plate voltage of 200; this is due to the lower plate current at the lower plate voltage producing less intense ionization of the gas present. As the grid potential was increased (in the negative direction) the grid current decreased instead of increasing as might be expected. This is due to the decrease of plate current with the lower grid potentials.

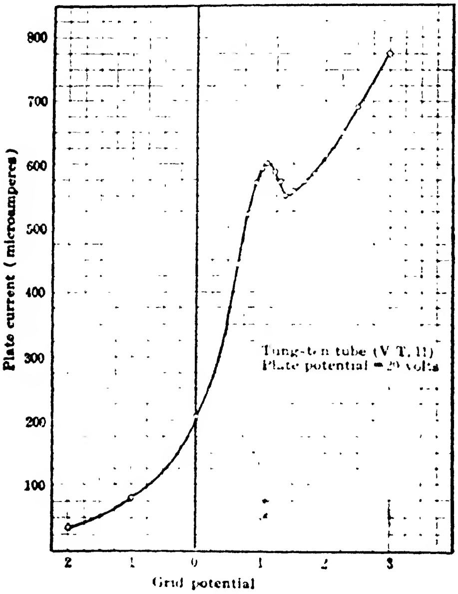


FIG. 39.—Showing the effect of a small amount of gas in producing a well-defined "hump" in the plate-current curve.

A tube having considerable gas in it may be made extremely sensitive as a detector if adjusted with a plate or grid voltage nearly sufficient to produce ionization; the slight increase in grid potential due to the incoming signal may then cause ionization to occur with a resultant great increase in the plate current. Such tubes are naturally more critical of adjustment than a true vacuum tube and for this reason will probably not be very popular in spite of their greater sensitivity. The modern high-vacuum tube used as a detector is much more dependable as a detector, even if less sensitive; its comparative lack of sensitivity can be overcome by using more amplification in the receiving set.

Several detailed studies of gas-content detector tubes have recently been made, notably those of Brown and Knipp.[1]

Tubes with Four Electrodes.—Although the three-electrode tube

[1] See "Alkali Vapor Detector Tubes," by Hugh A. Brown and Chas. T. Knipp, University of Illinois Bulletin, Vol. XXI, No. 11. For oscillographic records of the age of gas-content tubes see "Oscillographic Study of Electron Tube Characteristics, by E. Leon Chaffee, Proc. I.R.E., Dec., 1922.

(triode) was a great advance over the two-electrode tube (diode) it did not mark the limit of possibility for this type of apparatus. A four-electrode tube (tetrode) may be much better than the triode for certain purposes, and several such have been designed and are now extensively used. Furthermore, considerable advantage, so far as possible power output is concerned, is gained by using still another electrode, making a five-electrode tube, or pentode. Further mention of these multi-grid tubes will be made later on, after the characteristics of the triode have been analyzed.

Characteristic Curves for Three-electrode Tubes. The so-called "static" characteristic curves of a three-electrode tube show how the plate current and grid current vary as the grid potential is changed over a sufficient range to cause the plate current to vary from its maximum operating value to zero, the plate potential being constant while the series of points for the curve is being obtained. The same curves are taken for several values of plate potential.

Another set of curves is sometimes used showing the variation of plate and grid currents

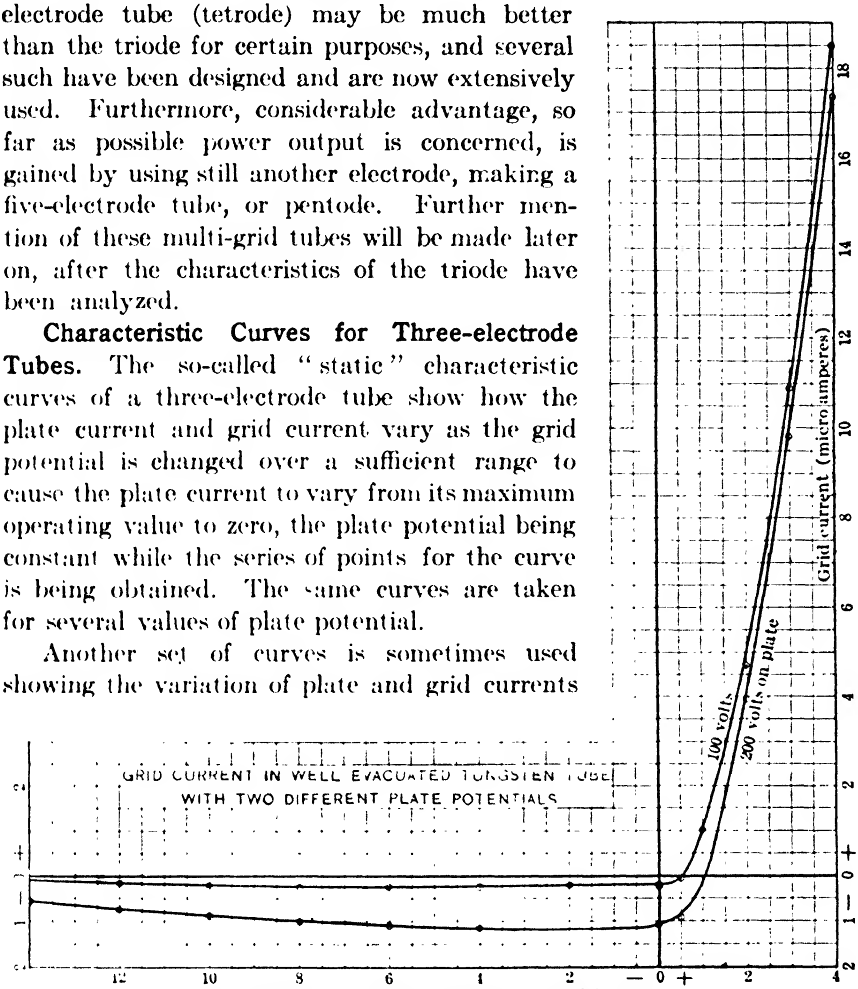


FIG. 40. Even in the very high vacuum tubes the grid shows a reversed current when its potential is negative; these curves are for a 250-watt power tube having a high degree of evacuation.

as the plate potential is varied from zero to its maximum safe value, the grid potential remaining constant, a series of such curves is obtained for various grid potentials.

Another, and probably more useful, set of curves show how the plate and grid currents vary as the grid potential is varied, the plate potential varying, during the process of getting the curve, in the same way it does when the tube is actually used in a detecting or generating circuit. When being used the three-electrode tube always has an impedance of some kind in series with the plate circuit. The value of the voltage used in the plate circuit is constant, not varying as the grid potential is varied, by signal or otherwise; it is therefore evident that as the grid potential varies, thus varying the current in the plate circuit, the plate potential

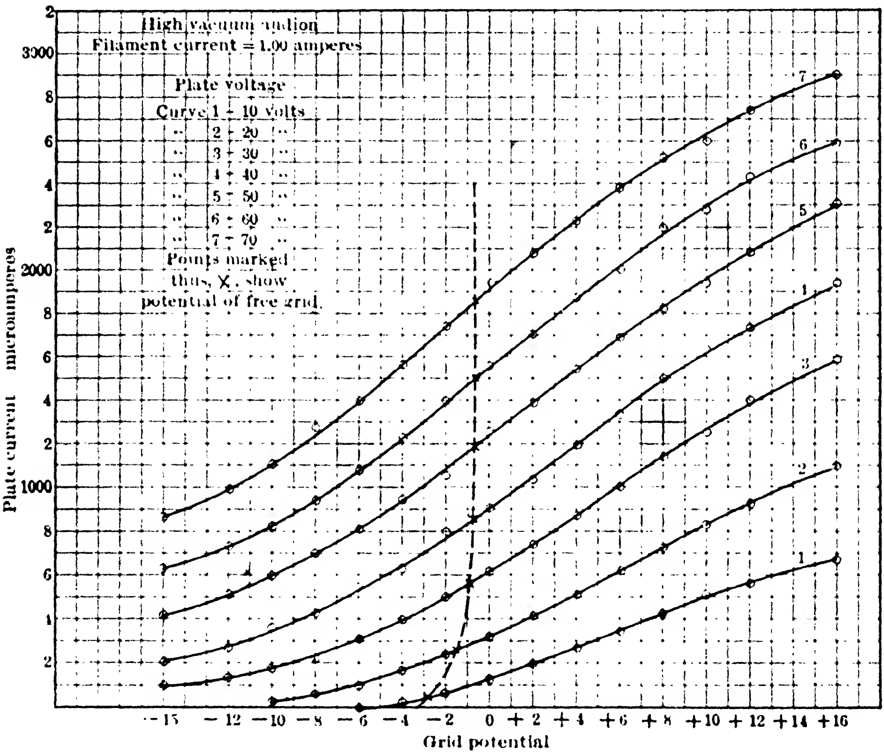


FIG. 41.—An old De Forest audion, after being well evacuated and baked, showed just as regular characteristics as the modern tube.

must vary because it is equal to the plate-circuit voltage minus the drop in the series impedance, and this drop varies with the grid potential.

This last set of curves is the one which most readily permits the prediction of the behavior of the tube. A resistance should be put in the plate circuit equal to that which is used when the tube is actually operating; a plate-circuit voltage should be used such that when the grid is set at the same potential as its average potential under operating conditions the plate current is the same as its average operating value. The plate-circuit voltage is frequently called the " B " battery voltage.

It has become customary in speaking of grid potential to refer the grid to the *negative end of the filament*; unless otherwise stated all the curves shown in this text are so given. In case the characteristics are desired when the grid is connected to the positive end of the filament it is only necessary to move the " zero grid potential " along to the right, on the curve sheets as given, by an amount equal to the *IR* drop in the filament.

In case the filament is heated by alternating current the zero potential point is taken as the mid-tap of the transformer winding supplying the filament power. (See Fig. 23.) In case a "heater type" of tube is used the electron-emitting cathode is taken as zero potential.

In Fig. 41 is shown a set of plate-current curves from an old De Forest audion, after it has been re-evacuated to take off all possible gas. The plate circuit had no added resistance except that of the *B* battery, which was so low that the variation in plate current did not appreciably affect the plate potential. On the curve sheet is shown the locus of the " free grid potential," i.e., the potential at which the grid set itself when its external terminal was completely insulated. This point will be taken up in more detail later.

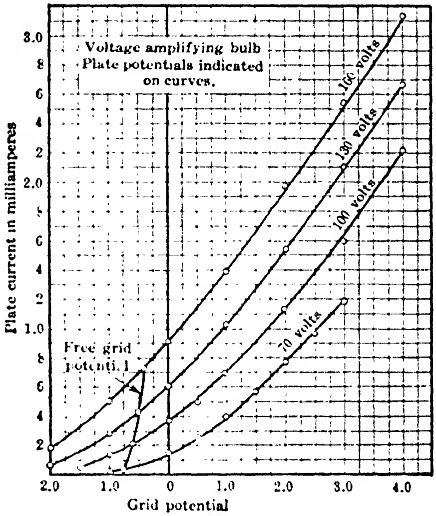


FIG. 42. -Plate-current curves for a tube intended as a voltage amplifier.

For the tube used in getting the curves of Fig. 41 it will be noticed that the grid voltage was more effective (in controlling the plate current) than the plate voltage in the ratio of about 2 to 1. Thus to get 1 milliampere of plate current it is necessary to use either ($E_p=20$, $E_g=13$), ($E_p=30$, $E_g=5.6$), ($E_p=40$, $E_g=1$), ($E_p=50$, $E_g=-3.4$), ($E_p=60$, $E_g=-7.2$) or ($E_p=70$, $E_g=-12$). Using the two extreme values, we see that a decrease in plate potential of $(70-20)=50$ volts is neutralized (in so far as it affects plate current) if the grid potential is increased from -12 volts to $+13$ volts, or a change of 25 volts.

In Fig. 42 is shown a set of curves from a tube designed for amplifying telephone line signals; free grid potentials in this tube follow about the same changes as for the tube used in Fig. 41. The much greater control of the grid of this tube is seen from the values of plate voltage and grid voltage for a current of 0.001 ampere. This is obtained with either ($E_p=160$, $E_g=0.2$), or ($E_p=70$, $E_g=2.6$) so that an increase in grid potential of 2.4 volts offsets a decrease in plate potential of 90 volts; the effectiveness of the grid is thus thirty-eight times as great as that of the plate.

In Fig. 43 is shown a set of curves for a tube having the plate and grid very close to the filament, the grid being comparatively coarse compared to that of the tube of Fig. 42. In Fig. 43 the grid potentials are referred *to the positive end of the filament*; as the filament IR drop was about 3 volts it is seen that if the grid were connected to the negative end of the filament the grid current would be practically zero. This tube is generally used as a detector with the grid normally somewhat positive.

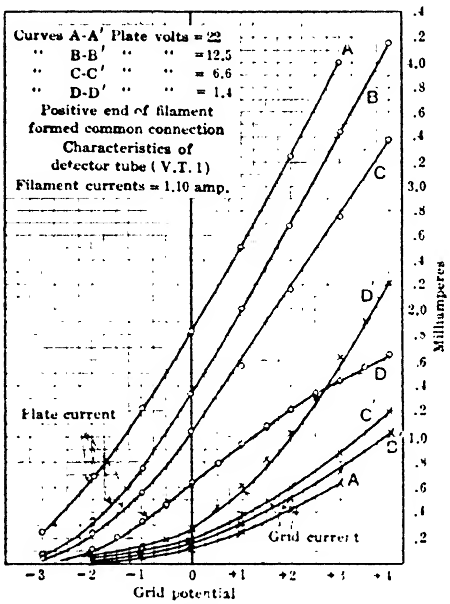


Fig. 43.—Characteristic curves for an ordinary detector tube, for a wide range of plate voltages. For the lowest plate voltage the grid current and plate current are about equal. In this test *the positive end* of the filament was used for the common junction, so that when the grid voltage was reported as zero the grid was actually about 3 volts positive, with respect *to the negative end* of the filament.

It will be noticed that the grid current (for a given grid potential) decreases as the plate potential is increased. When the grid and plate are positive by about the same amount (curves D and D' with grid 3 volts positive), each takes about the same thermionic current; the greater area of the plate compensates for the greater proximity of the grid and filament.

Fig. 44 shows the effect of filament current on the static characteristics of a 250-watt power tube. The grid currents with negative grid potentials are too small to be plotted on the curve sheet. The filament currents were measured at that end carrying the smaller current.

As was pointed out in discussing Fig. 22, the current in a filament varies throughout its length when it is delivering electrons to the plate, the amount of variation depending directly on the value of the plate current. In Fig. 45 is shown a set of curves to illustrate this point; a constant voltage of 32 was impressed on the filament, the grid was held at a positive potential of +100 volts and the plate voltage varied from zero to 200 volts.

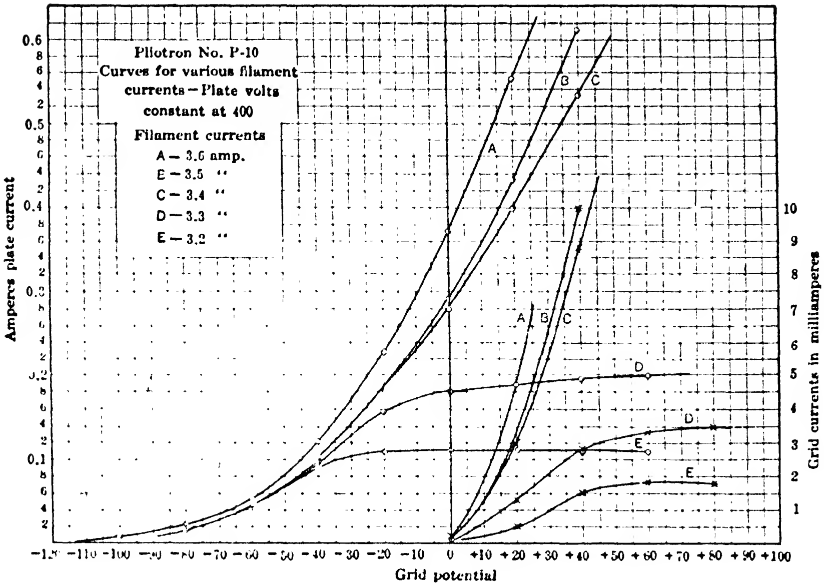


FIG. 44. Effect of filament current on the characteristics of a tube intended for a power generator.

This set of curves serves not only to show the peculiar changes in filament current, but also how, as the plate voltage increases, the grid current is reduced. The sum of the grid current and plate current gives, for all values of plate voltage, the difference between the two filament currents. The resistance of the filament of a vacuum tube under such conditions is not a simple function of volts and amperes; it involves all the theory of a long, leaky telegraph line.

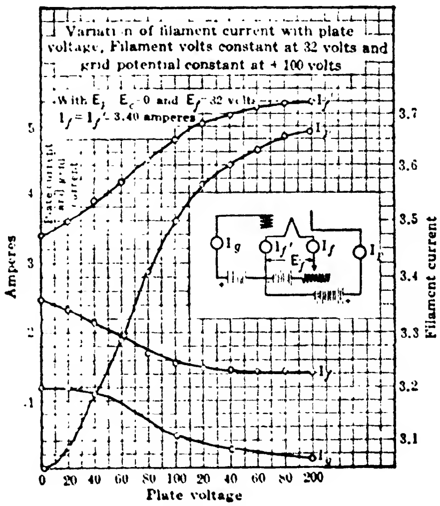


FIG. 45.—Showing the effect of the plate voltage upon the filament current of a power tube, the voltage impressed on the filament being constant. The change in grid current produced by increasing plate potential is also shown.

The safe filament current for these large power tubes is always rated in terms of the maximum current, that is, the end of the filament where the plate current and battery heating current combine to give a current greater than normal battery current.

In Fig. 46 are shown curves for the same tube as used for Fig. 44; the filament current (larger value) was held at 3.60 amperes and various voltages were impressed on the plate. With low plate voltage it is seen that when the grid becomes positive the plate current undergoes a rapid decrease. This combination of high positive grid voltage and low plate voltage occurs when the tube is used for generating power and results in peculiar-shaped plate current instead of a sinusoidal variation as is generally assumed.

In Fig. 47 are shown the characteristic curves for a C. E. P–20 (20 grid wires per inch) pliotron obtained by holding the grid potential constant while varying the plate voltage. For all these curves the grid currents were only a few microamperes. In this tube it is evident that 1 volt on the grid has the same effect on plate current as 11 volts on the plate. In Fig. 48 are shown the plate current-grid voltage curves of two of the modern receiving tubes, of which millions are in use today. In Fig. 49 are shown some typical curves for a fine mesh grid 250-watt power tube (P–30). In this tube the grid voltage is twenty-two times as effective as the plate voltage in determining plate current. It will be noticed how quickly the grid current rises as the plate potential decreases beyond a certain limit.

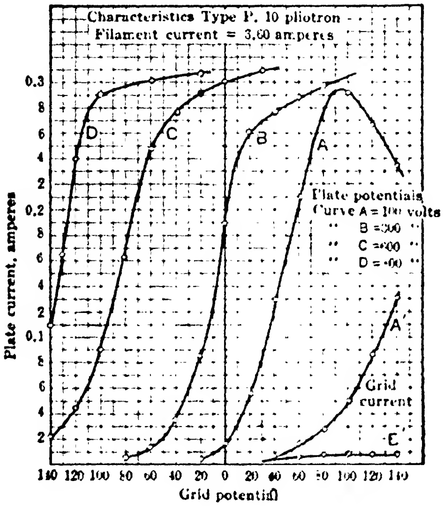


FIG. 46.—Static characteristics of a Type *P* pliotron for various plate voltages, filament current being constant.

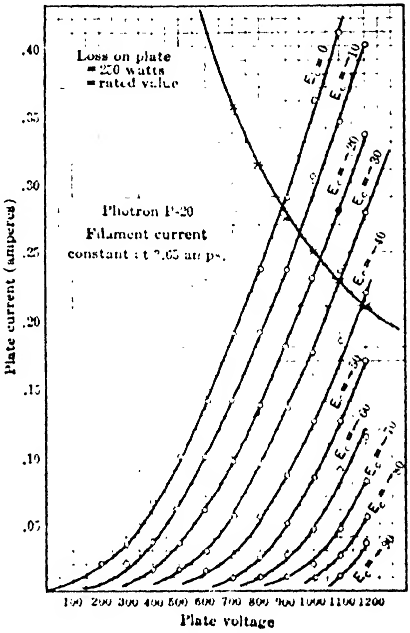


FIG. 47.—Static characteristics of a Type *P* tube for various fixed grid potentials and variable plate voltage. The curve in the upper part of the diagram shows the limit of operation of the tube.

In Fig. 50 are shown the curves of a type 27 amplifying and detecting triode, having an indirectly heated cathode, that is, the " heater " type of tube. In Fig. 51 are shown the characteristics of an " output " triode, designed to deliver to the loud speaker a power as high as 1.6 watts. In a later section of this chapter are shown the characteristic curves of four- and five-electrode tubes, that is, tetrodes and pentodes.

Potential of the Free Grid of a Three-electrode Tube.—When the grid of a vacuum tube is entirely disconnected from other circuits it is said to be " free," meaning that it is free to assume any potential circumstances may demand. Actually a grid is never really free, because there is always some leakage from the grid to the plate and filament even in tubes with extremely high vacuum. If the value of this leak resistance is perhaps 50 megohms the grid may be reckoned as free, although in many tubes a much greater resistance exists and the grids are correspondingly more "free."

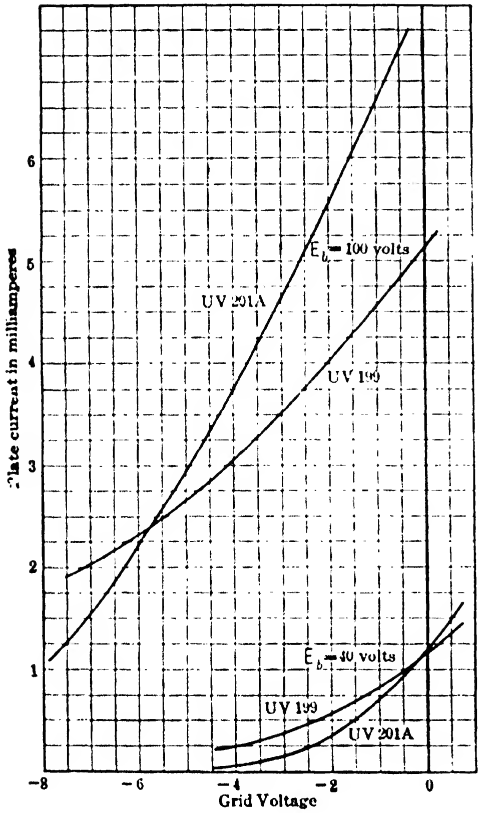


FIG. 48.—Characteristic curves for two of the modern receiving tubes; the type 199 tube is designed for operation from dry cells, taking only 0.06 ampere in its filament.

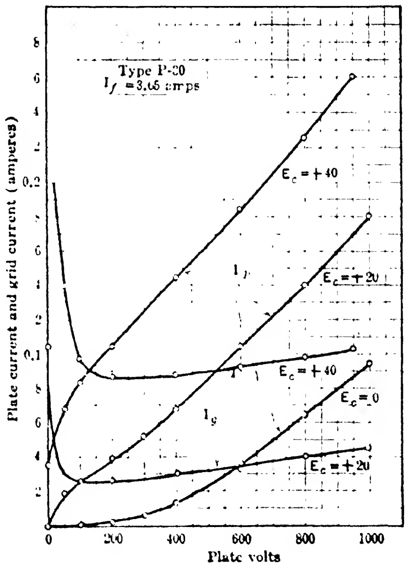


FIG. 49.—Similar to the curves of Fig. 47, this tube having a grid with finer mesh.

It is almost an axiom in vacuum-tube operation that a grid should never be left free. Consistent operation of the tube is almost impossible unless the resistance between the grid and filament is of definite value, and sufficiently low; it seldom exceeds 1 megohm in ordinary detecting or amplifying sets.

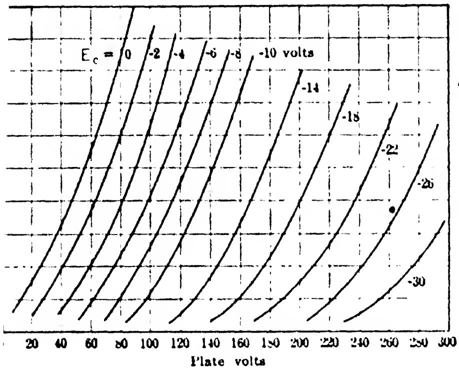


FIG. 50.—Characteristic curves of the type 27 triode.

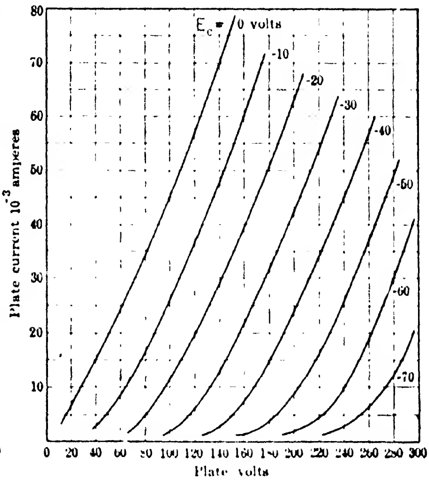


FIG. 51.—Characteristics of the type 45 triode.

TYPE 27 TRIODE

I_{heater} = 1.75 E = 2.5

E_p =	45	90	135	180 volts
E_c =	0	−4.5	−9	−13.5 volts
I_p =	4	5	5.5	6 mils
R_p =	10^4	10^4	9×10^3	9×10^3 ohms
μ =	9	9	9	9
Gm. =	900	900	1000	1000μ mhos
Output	. . . . .	30	78	164 milliwatts

TYPE 45 TRIODE

Filament volts   2.5 volts

Filament current   1.5 amps.

Plate volts	180	250 volts
Grid bias	33	50 volts
Plate current	26	32 milliamps.
Plate resistance	1650	1900 ohms
Amplification constant	3.5	3.5
Transconductance	1800	1850 micro mhos
Power output	780	1600 milliwatts

In Fig. 52 is shown a connection in which a free grid is used; tube 1 is repeating into tube 2, the fluctuations of plate voltage of 1 being

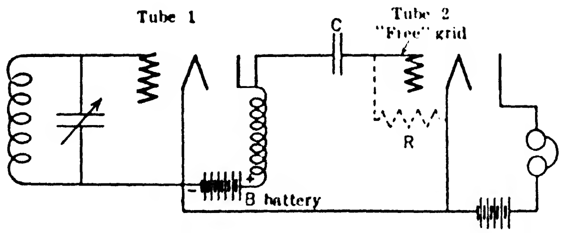


FIG. 52.—A circuit illustrating the meaning of the term "free grid," the grid of the second tube is electrically free to assume any potential that circumstances may demand.

impressed on the grid of 2. The grid of 2 cannot be connected directly to the plate of 1 because this plate is at comparatively high positive potential, due to its *B* battery. By putting an insulating condenser *C* of sufficiently large capacity between the plate of 1 and grid of 2 the fluctuations of plate voltage repeat through the condenser into the grid, but the grid is insulated from the high positive continuous e.m.f. of the plate of 1.

Now such a grid is said to be free; the insulation of condenser *C* may be hundreds of megohms, so that the grid is nearly free to assume any potential whatever. Because of the irregular action of a tube so connected a high resistance leak of one megohm or less (as indicated by the dotted

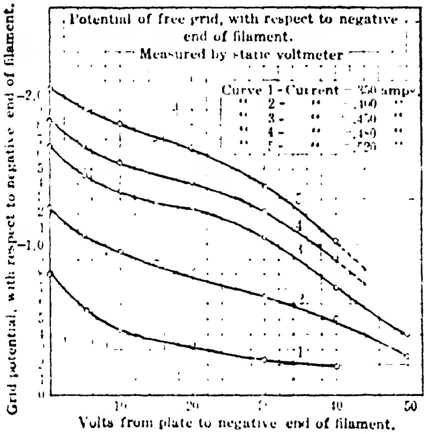


FIG. 53. Variations in free grid potential of a small highly evacuated tube for various plate voltages and filament currents; measurements by a highly insulated sensitive static voltmeter.

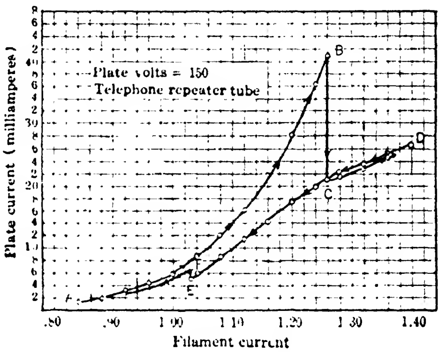


FIG. 54.—A peculiar cycle obtainable from a tube having a free grid; as the filament current was increased and then decreased the plate current went around the loop as indicated by the arrow heads; plate potential was kept constant.

line connection) is always advisable, to keep the grid, normally, at a suitable potential.

Some of the effects produced by a free grid will be indicated by the accompanying curves. In Fig. 53 is shown how the potential of a free grid may be expected to change as the plate voltage is increased from zero, for various filament temperatures. The higher the plate voltage the closer the grid potential approaches zero potential, i.e., that of the negative end of the filament. With zero plate voltage the grid goes negative as much as 2 volts, due undoubtedly to the accumulation of electrons which have left the filament with enough initial velocity to carry them as far as the grid.

The potential of the free grid depends entirely on the order in which the successive adjustments are carried out; thus, if the grid is left free, filament current brought to normal, and then plate potential brought to normal, a value for free grid potential may be obtained entirely different from that if the plate were first put at its proper potential and then the filament current brought to normal.

In Fig. 54 is shown the curve obtained (with free grid) by holding the plate at 150 volts, increasing the filament current from a low value to a high value and then decreasing the filament current through the same range. A peculiar loop is obtained, explained by the fact that as the plate potential was applied before there was a liberal supply of electrons in the vicinity of the grid the grid went positive. This positive grid gave comparatively large values of plate current from A up to the point B on the curve sheet; here the grid suddenly lost most of its positive charge due to bombardment by many electrons, and became nearly zero in potential with a consequent decrease in the plate current. From C to D and back to C the grid had nearly the same potential for increasing as for decreasing filament current, but from C to E the grid potential was much lower than it was for the corresponding values of filament current, when increasing values were being taken. At E the grid suddenly increases its potential a small amount and for the remainder of the cycle it has about the same potential as it had for increasing filament current; other tubes showed exactly the same effect.

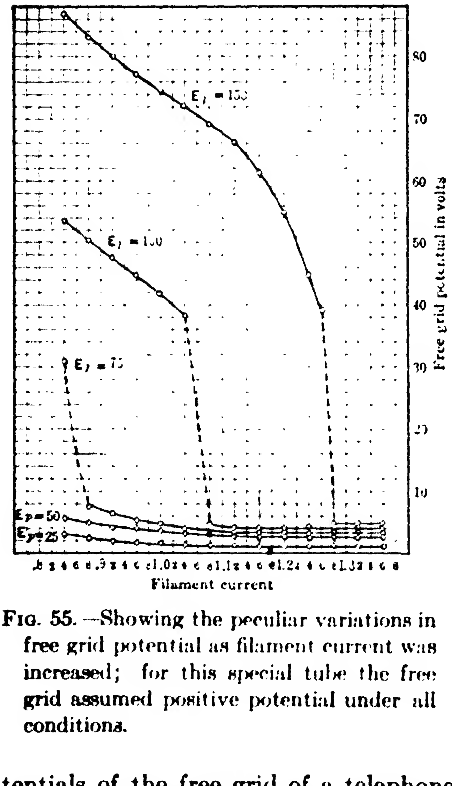


FIG. 55.—Showing the peculiar variations in free grid potential as filament current was increased; for this special tube the free grid assumed positive potential under all conditions.

In Fig. 55 are shown the potentials of the free grid of a telephone amplifying tube. For low values of filament current the free grid assumes a potential about half that of the plate, then as the filament current is increased the grid potential decreases gradually until a critical value of filament current is reached. At this critical filament current (i.e., critical supply of electrons) the grid potential suddenly falls to a comparatively low value, which value decreases somewhat as the filament current is still further increased. It will be noticed that for this tube the free grid is always positive, whereas for most high-vacuum tubes the grid potentials are negative.

While the free grid triode is not generally serviceable it may very well be used in certain circuits. Thus Ferrié [1] and his co-workers used a triode with free grid, in connection with a photoelectric cell, to measure the energy coming from the stars. In this case the grid was made more or less free according to the number of electróns passing across the light sensitive cell. It is reported that he obtained a current amplification in this way of 10,000 times, using only one triode.

Relations between Currents and Potentials in a Three-electrode Tube.—From experimental results already presented it is evident that the grid current and plate current of a three-electrode tube vary with either filament current, plate voltage, or grid voltage. It is also evident that the grid current is negligibly small compared to the plate current, and that the plate current is not affected directly by the grid current except under unusual conditions, as, e.g., curve A of Fig. 46. Unless we are specifically interested in distortion in an amplifier, or in the case of an oscillator, in the losses in the grid circuit, the grid current may be neglected. Furthermore, unless the conditions are such that saturation current is reached (plate current using all the electrons emitted from the filament) the filament current does not affect the plate current to a great extent. We shall therefore examine, in this section, the relations between plate current and grid and plate potentials, neglecting grid currents and the effect of too small a filament current.

We have seen that the plate current depends upon both plate voltage and grid voltage, to some power higher than the first, and that the grid potential is much more effective in controlling the current than is the plate potential. We may therefore write

$$I_p = A(E_p + \mu E_g)^x, \qquad (7)$$

where I_p = plate current in amperes;
A = a constant depending upon type of tube;
E_p = potential of plate to negative end of filament;
E_g = potential of grid to negative end of filament;
μ = relative effectiveness factor of E_g;
x = an unknown exponent, possibly variable.

Langmuir has given this equation with the value of x as 1.5; Van der Bijl has given the equation with the value of x as 2.0, having also an added quantity inside the parenthesis, a small constant in which such factors as velocity of emission of electrons, contact difference of potential of the electrodes, etc., are taken care of.

A little reflection will show that this simple relationship (Eq. 7) is

[1] "Amplification of Weak Currents and their Application to Photo Electric Cells," Ferrié, Jonaust, and Mesny, Proc. I.R.E., Aug., 1925.

approximate only. The factor A is probably not a constant, and the exponent x is not a constant, but depends as does A, upon the value of E_p and E_g. In a paper by Petersen and Evans [1] the general theory of the triode is worked out on the assumption that

$$I_p = f(E_p E_g) = \Sigma a_{mn} E_p^m E_g^n. \quad \quad (8)$$

This is a double power series in which the general coefficient has the form $a_{mn} = \frac{1}{m!\,n!} \frac{\partial^{m+n} f(00)}{\partial E_p^m \, \partial E_g^n}$. When investigating the detailed performance of a triode a complicated relationship of this nature must be assumed, but for the first simple analysis of the action of a triode, Eq. (7) will suffice, and is used in the subsequent analyses. Our solutions will not in general indicate distortion, harmonics, etc., because the relationship of Eq. (7) is too simple to truthfully represent the facts.

Amplification Factor of a Triode.—The quantity μ is the theoretical voltage amplifying power of the tube; it is ordinarily taken as a constant, its value depending solely upon the geometry of the tube. Many tests show this to be true for the ordinary use of the tube; it may be that with very low plate voltage and high grid voltage μ changes somewhat, but in the ordinary working range of E_g and E_p it is practically constant. As previously stated, it varies in different types of tubes from 2 to 200 or more.

When many determinations of μ are to be made, it is worth while to arrange some apparatus as shown in Fig. 56, a scheme due to J. M. Miller. The resistance R_2 is preferably 10 ohms and R_1 is a decade resistance box having units, 10-ohm, 100-ohm, and 1000-ohm units; the 1000-ohm units are used very seldom, but few tubes having high enough values of μ to require them.

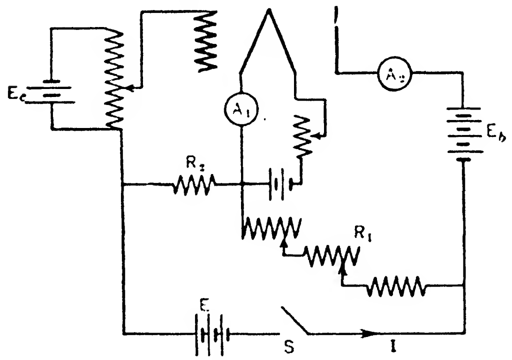


FIG. 56.—An arrangement of apparatus for rapidly determining the voltage amplification factor of a tube.

An ammeter A_1 serves to read the filament current, and milliammeter A_2 serves for plate current. This meter should have two or three scales, so that for various types of tubes to be tested the plate current will give indications well up on the scale. The filament battery should be perhaps 6 volts and E_b and E_c should have voltages suitable for the tubes to be tested.

[1] Bell System Technical Journal, July, 1927, p. 442.

With S open E_b, E_c, and I_f are put at their proper values and the reading of A_2 is noted. When S is closed, permitting current I to flow around the circuit E, R_1, R_2, E, the reading of A_2 will in general change; by properly adjusting R_1, however, it will be found that the reading of A_2 (which is the plate current) does not change when switch S is closed. The ratio of R_1 to R_2 for this adjustment gives μ.

By examination of Fig. 56 it will be seen that depressing key S decreases the voltage impressed on the plate by an amount IR_1, and raises the voltage of the grid by an amount IR_2. From inspection of Eq. (7) it is evident that if I_p does not change when S is closed,

$$(\Delta E_p+\mu\Delta E_g)=0, \qquad (9)$$

where ΔE_p and ΔE_g are the changes in E_p and E_g due to closing switch S. We therefore have the relation

$$IR_1-\mu IR_2=0,$$

or

$$\mu=\frac{R_1}{R_2}. \qquad (10)$$

With this scheme it is possible to investigate the dependence of μ on E_g, E_p, and I_f very quickly.

The value of R_1 should not be so high that the drop through it, due to the plate current, is an appreciable fraction of E_b, otherwise the plate potential will not be E_b, but something less, and must be calculated.

The value of the exponent x should theoretically be a constant, but in the actual tube it is constant only for a limited range of voltages. The voltage between the plate and filament is different for the different parts of the filament, and the velocity of emission of the electrons may not be negligible when the plate and grid voltages are low.

If the grid is held at zero voltage the relation between I_p and E_p is $I_p=AE_p^x$. The determination of the exponent x is most easily carried out by plotting the various values of I_p and E_p on logarithmic cross-section paper; if x is a constant the graph is a straight line with slope equal to x. If the graph is not a straight line the value of x varies, but it may be determined for any value of E_p by measuring the slope of the graph.

Figs. 57 and 58 give the variations between plate current and plate voltage for two amplifying and detecting bulbs designed for 40 volts on the plate and an IR drop in the filament of about 3.5 volts. The two curves were transposed to logarithmic paper, giving the graphs shown in Fig. 59, the straight dotted line shows the slope the two curves would have if the plate current varied with the square of the plate voltage. From

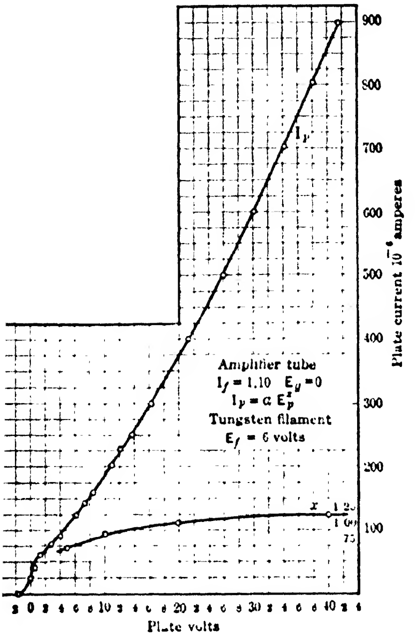


FIG. 57.—Variation of plate current in a tungsten filament tube as E_p is varied and grid potential held constant; values of the exponent of Eq. (7), p. 513.

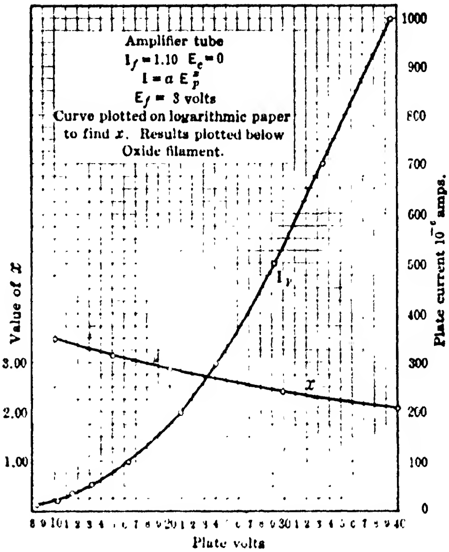


FIG. 58.—Curves similar to those of Fig. 57, the tube used having an oxide-coated filament.

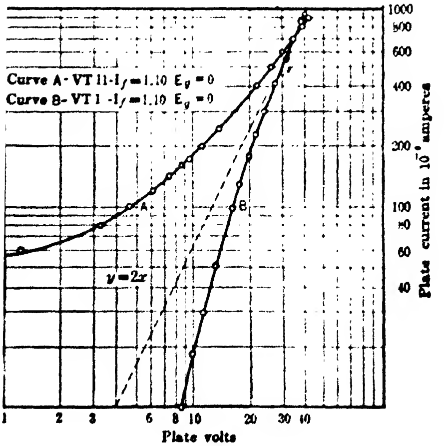


FIG. 59.—The curves of Figs. 57 and 58 transposed to logarithmic coordinates; this graph shows that the exponent for Eq. (7) is neither 1.5 nor 2, but is a variable for both tubes.

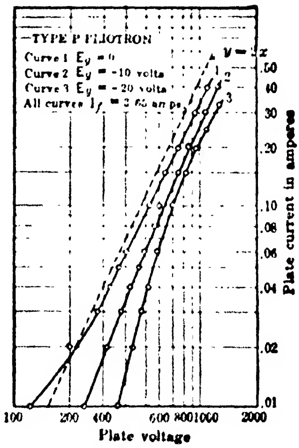


FIG. 60.—Logarithmic plot of the plate-current curve for a high-voltage power tube; with a certain constant negative grid potential the plate current of this tube varies with the square of the plate voltage.

the logarithmic graphs the values of x were measured and transferred to Figs. 57 and 58 to give the curves of x there shown.

In Fig. 60 is shown the logarithmic graph for a high-power pliotron, the rated plate voltage being 1000–2000 volts; it is seen that for high plate voltages the plate current varies as the square of the plate voltage. For the lower plate voltages the IR drop in the filament (about 20 volts) and the velocity of emission of the electrons tend to give an exponent other than 2; however, if the grid is held at -10 volts, the plate current follows the square law very closely throughout the range of the graph. A greater negative potential makes the plate current vary with higher power of plate voltage for the lower values of plate potential; this is to be expected from inspection of Eq. (7).

Factors Determining the Amplification Constant of a Triode.—The screening action of a grid was first discussed in Maxwell's "Electricity and Magnetism," Vol. I; here it is worked out for a flat grid and flat cathode, infinite in extent. Miller[1] has shown how closely Maxwell's derivation applies to the construction of an ordinary triode.

More recently King[2] has analyzed the dependence of the factor μ upon the parameters of the tube. The theoretical voltage amplification factor of the tube is shown to be expressible by the relation

$$\mu=\frac{2\pi a n}{\log_e \dfrac{1}{2\pi r n}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

in which a = distance between grid and plate;
n = number of grid wires per centimeter;
r = radius of grid wires.

In the derivation of this formula the grid, cathode, and plate have all been assumed as infinite parallel planes; although actual triodes depart very far from this requirement experimentally determined values of μ for several tubes of widely different construction check quite well with the value calculated from Eq. (11).

An interesting point in the theoretical analysis is the fact that the distance between the grid and filament plays no part in determining the value of μ; the closeness of the grid to the plate does, however, have a controlling influence.

In one tube, for example, $a=0.2$, $r=0.025$, $n=2.3$ and μ (calculated from Eq. (11)) calculates 2.85; the experimental value was 2.50. In another, $a=1.35$, $n=31.5$, $r=0.00127$ and μ (calculated) is 194; the experimental value was 217.

[1] Proc. I.R.E., Vol. 8. No. 1.
[2] Physical Review, Vol. 15, No. 4.

If more accurate theoretical results are desired than Eq. (11) yields it is necessary to consider the non-uniform distribution of the charge on the grid wires; due to their proximity the charge on the surface is not uniform, being a minimum where the wires are closest together. When this factor is taken into account, as in all other branches of electrical theory, the resulting formulas involve hyperbolic functions of the triode parameters.

Resistance of the Plate Circuit of a Triode and its Variations.—There are three circuits to be considered in getting the characteristics of three-electrode tubes, the filament, the grid to filament circuit, and the plate to filament circuit. The grid to filament is called the *input circuit* of the tube, and the plate to filament is called the *output circuit* of the tube.

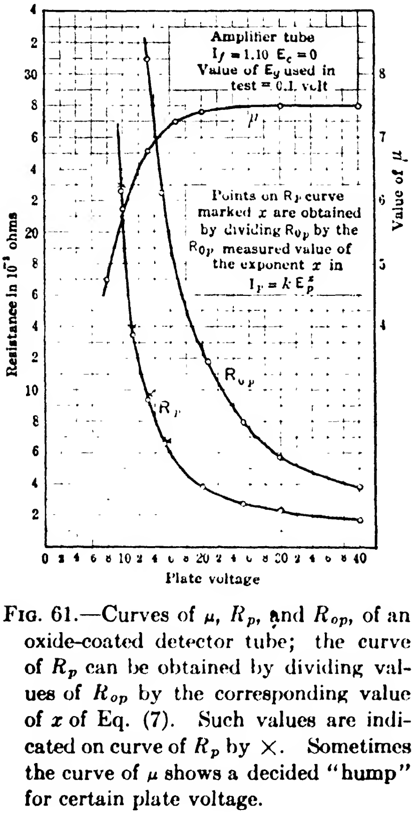


FIG. 61.—Curves of μ, R_p, and R_{op}, of an oxide-coated detector tube; the curve of R_p can be obtained by dividing values of R_{op} by the corresponding value of x of Eq. (7). Such values are indicated on curve of R_p by $\times$. Sometimes the curve of μ shows a decided "hump" for certain plate voltage.

In the ordinary small detecting and amplifying tube the filament current is practically independent of any changes in the grid and plate circuits. In large power tubes, however, this is not so, the resistance of the filament varying a good deal as either the grid or plate potential is varied, this variation being shown by impressing constant voltage on the filament and then impressing various potentials on the grid and plate. The accompanying changes in grid and plate current cause a non-uniform current to flow through the conductor, under which condition the filament has a resistance different from that when the current is the same everywhere through its length.

The resistances of the input and output circuits vary throughout extreme ranges, and they are different for alternating current than for continuous current; we shall first consider the output circuit. The ratio of plate voltage to plate current is generally called the *output impedance.* As there can be no appreciable lag in the motion of the electrons behind the impressed electric field, it might seem more appropriate to speak of output resistance instead of output impedance. But it is to be remembered that there is capacity between the plate and other electrodes, and at high radio frequencies this capacity may have an appreciable effect on the current flow from plate to filament. From this viewpoint the plate filament circuit has impedance, not merely resistance.

If we maintain the grid at zero potential (or any other fixed potential) and neglect the plate circuit capacity the plate current will vary with plate voltage only and we may speak of plate-circuit resistance. With constant grid potential and different plate potentials the values of plate current determine this resistance of the plate circuit, R_{op}, for continuous currents.

The value of the a.c. resistance, R_p, is determined by the ratio of dE_p/dI_p, the grid voltage being maintained constant; if we keep the grid at zero voltage we may write

$$I_p = AE_p^x$$

or

$$\frac{dE_p}{dI_p} = \frac{1}{AxE_p^{x-1}} = R_p.$$

But the c.c. resistance is

$$R_{op} = \frac{E_p}{I_p} = \frac{1}{AE_p^{x-1}}.$$

From these we get the relation

$$R_p = \frac{R_{op}}{x}. \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

In Fig. 61 are shown the curves of R_p and R_{op} for an oxide-filament amplifying tube; the points indicated by circles on the R_p curve were obtained by the a.c. measurement and those indicated by crosses were obtained by dividing the points on the R_{op} curve by the proper value of x. On the same curve sheet is shown the value of μ for this tube; it is nearly constant in the working range of the tube (E_p between 20 and 40 volts) and falls off with the lower plate voltages; in some other triodes the value of μ increased with the decrease in plate voltage.

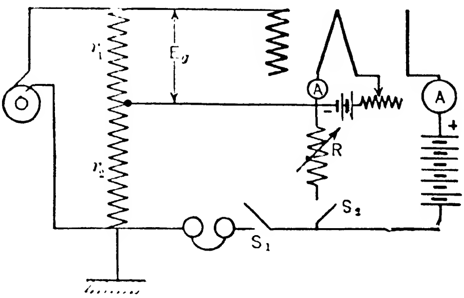


FIG. 62.—An arrangement of apparatus for conveniently measuring the μ of a tube as well as the a.c. resistance of the output circuit.

The value of R_p is found experimentally by the scheme outlined in Fig. 62, originated by J. M. Miller; the same arrangement serves to measure μ by a.c. test. Fig. 63 shows a curve of μ obtained by the method; it shows μ to be independent of filament current. With S_2 open the ratio of r_1 to r_2 is varied until no signal is heard in the telephone and we then have

$$\mu = \frac{r_2}{r_1}. \qquad (13)$$

In measuring R_p, r_1 and r_2 are fixed at some convenient value (say equal to each other; the ratio must be different than that given in Eq. (13)), and with S_2 closed R is varied until no signal is heard in the phone. With this adjustment we have

$$R_p = R\left(\frac{r_1}{r_2}\mu - 1\right). \qquad (14)$$

Flowing through the potentiometer there is a current i, which gives

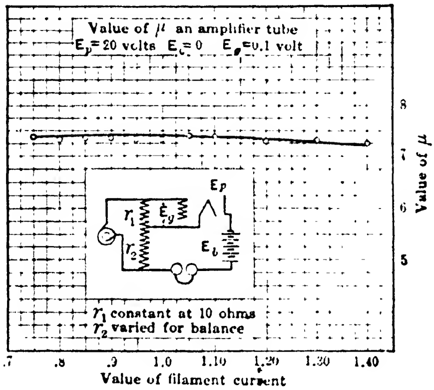


FIG. 63.—Value of μ of a small amplifying tube, obtained by the scheme outlined in Fig. 62; this shows μ to be nearly independent of the filament current.

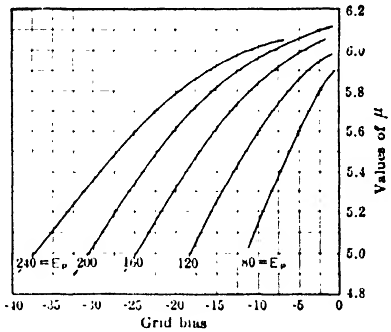


FIG. 64.—Variation in μ of a telephone repeater tube.

a drop between grid and filament $= Ir_1 = E_g$. If the alternating voltage E_g is impressed on the grid of a vacuum tube it will produce an alternating current in the plate circuit, I_p. Of course the actual plate-circuit current is not alternating, it is pulsating; this pulsating current may be resolved into a steady current I_{op} and an alternating current I_p. The current I_{op} is produced by the steady values of E_g and E_p and the alternating current I_p is caused by the variations in E_g.

The magnitude of this current I_p can be calculated by remembering that a voltage E_g in the grid circuit is equivalent to a voltage μE_g in the plate circuit. This voltage, μE_g, will cause an alternating current to flow in the plate circuit which is equal to $\dfrac{\mu E_g}{R_p+R}$. This current flowing through the resistance R must give a drop equal to $\dfrac{\mu E_g R}{R_p+R}$, and if there is no signal heard this drop must equal that across r_2 (which is equal to $E_g\dfrac{r_2}{r_1}$) when a balance is obtained. We therefore have

$$E_g\frac{r_2}{r_1}=\mu E_g\frac{R}{R_p+R}$$

Solving this equation for R_p we get the relation given in Eq. (14) above.

In case the resistance R does not permit a balance to be obtained, it being too small, the ratio of r_2/r_1 can be suitably altered.

The relation between I_p and E_g is not a linear one and it is therefore evident that R_p must vary throughout the cycle of change in E_g. The value of R_p is therefore represented correctly only by a constant (the

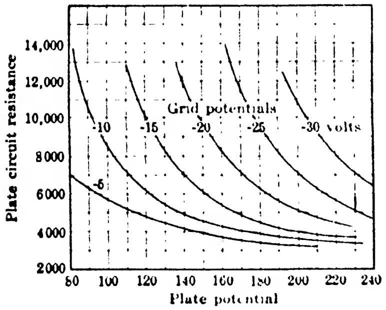


Fig. 65.—Variation in plate circuit resistance of a telephone repeater tube.

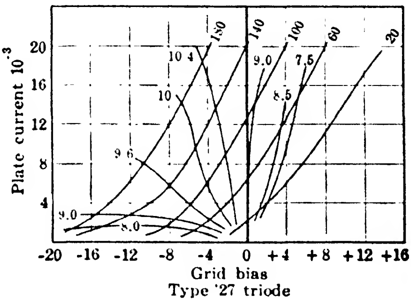


Fig. 66.—Variation of the amplification factor of the type '27 triode.

value of which we call R_p) and a series of harmonic terms; these harmonic terms become more pronounced as E_g is varied through wider ranges.

In the measurement of R_p by the method outlined above it will be found that complete silence cannot be obtained at the balance point; the note heard in the telephone is complex and only the fundamental note can be balanced. A balance will generally be most easily obtained if comparatively low values of E_g are used, say not more than 0.1 volt; moreover, it will be found that the value obtained for R_p varies with E_g, becoming greater for high values, as explained on p. 604.

Variations in the Amplification Factor of a Triode.—Simple theory leads to the conclusion that the amplification factor of a triode is a constant, depending only upon its geometry. But owing to the fact that the grid never completely encloses the filament, and the additional fact that different parts of the filament may be at different potentials, the amplification factor is found experimentally to be variable. In Figs. 64 and 65 are shown the values of μ and R_p for a telephone repeater triode.[1] It is seen that for the lower potential gradients in the triode R_p increases and μ decreases.

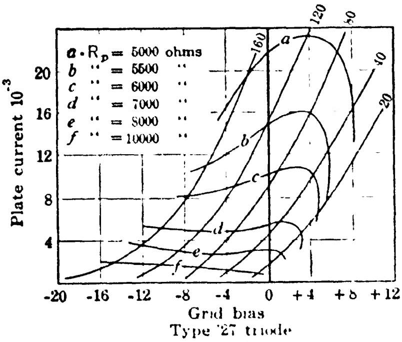


Fig. 67.—Variation in plate circuit resistance of the type '27 triode.

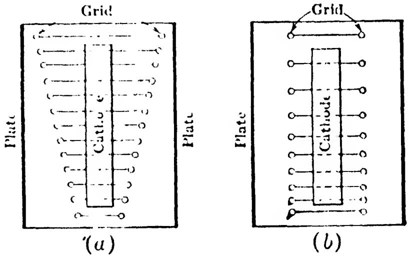


Fig. 68. Two constructions of grid which give triodes with an amplification factor decreasing as grid bias is increased.

In Figs. 66 and 67 are shown some results obtained by Terman and Cook[2] on a heater-type triode, the values being plotted in a somewhat different manner from those of the previous figures. Whereas all these figures show large variations in μ and R_p, it will be seen that, for reasonably small departures from specified grid and plate voltages, the changes in R_p and μ are relatively small. It is only for exceptionally low plate voltages and high positive or negative grid bias that the values change to any great extent.

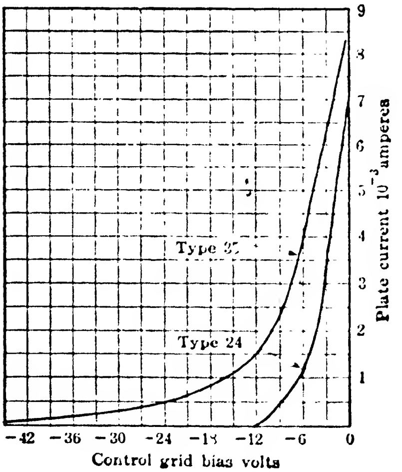


Fig. 69.—Comparison of grid control of plate current in a tube having constant μ (24) and one having a variable μ (35).

Tubes with Variable Amplification Factor.—For certain purposes a tube with a variable μ proves to be advantageous, notably when the amplification of a set is controlled by the grid bias of one of the radio-frequency tubes. The amplification control is smoother and the modulation distortion is less if the tube has a μ which decreases with increase in grid bias. Using a grid of helical construction, a variable μ can be

[1] Petersen and Evans, Bell System Technical Journal, July, 1927.
[2] I.R.E., June, 1930, p. 1044.

obtained by making the helix of tapering form or by varying the spacing of turns. These two constructions are indicated in Fig. 68; in *a* the spacing of the grid turns is uniform but they are of gradually increasing diameter, and in *b* they are of uniform diameter but have a gradually increasing spacing.

In Fig. 69 are shown the forms of plate current for two tubes of nearly the same construction, with the exception of the grid form. The type 24 tube has an ordinary uniform grid and nearly constant μ; the type 35 tube had a variable μ due to its grid construction. It is evident that for increasing negative bias on the grid of the 35 tube the μ of the tube becomes gradually smaller.

Input Circuit.—The resistance of the input circuit (grid-filament) for continuous current is practically infinite for all values of negative voltage; the current taken by the grid of the average tube when the grid is at lower potential than any part of the filament is of the order of 1 microampere or less. With a positive grid the current to the grid varies approximately as the square of the grid potential. When the grid and plate are at the same positive potential, the two currents are of the same order of magnitude (see Figs. 49, 46, and 43), so that the grid-filament resistance R_g is about the same as the plate-filament resistance R_p. It goes through the same kind of changes with respect to filament current, grid voltage, etc., as does R_p. It is to be noted from the curve sheets, however, that whereas an increase of E_g *decreases* R_p an increase in E_p causes an *increase* in R_g.

To measure the a.c. input resistance,[1] a scheme such as that illustrated in Fig. 62 is not directly applicable; for any ordinary scheme of measurement a transformer will be required to decrease the grid-circuit resistance to a value readily measured.

The author has used a bridge for measuring the characteristics of the input circuit of various tubes, the measurement being made at 50,000 cycles. The scheme used is illustrated in Fig. 70; the same setting of the bridge permitted the measurement of both capacity and resistance of the tube input circuit. The 50,000-cycle power was supplied to the bridge by wire A, the other side being grounded. Suitable high-resistance leaks are shunted across C_1 and C_2, these resistances being free from appreciable distributed capacity. Certain precautions have to be observed in using such a bridge as noted in an article by the author in the Proc. I.R.E.[2]

With suitable values of C_2 and R_4 the bridge is balanced with S open, the values of C_2 and R_4 being recorded. When S is closed the balance is destroyed, due to the capacity and conductance of the tube input circuit; by properly decreasing C_2 and increasing R_4 the balance may be

[1] See p. 536 for the effect of this input resistance on the tuning of the receiver circuit.

[2] "Some Notes on Vacuum Tubes," Proc. I.R.E., Vol. 8, No. 3, June, 1920.

again obtained. The total capacity and conductance in the (4) arm must now be the same as when S was open, so that the capacity of the input circuit is at once obtained as the difference in the two settings of C_2; from the two values of R_4 the conductance of the input circuit can be readily calculated.

In Figs. 57–60 are shown the variation in the input circuit of a small detecting tube rated at 1.1 amperes filament current and 20–40 volts in the plate. Unless the tube is defective the conductance is practically zero until about 0.8 ampere is used for heating the filament. It then rises rapidly until with normal filament current the conductance is about

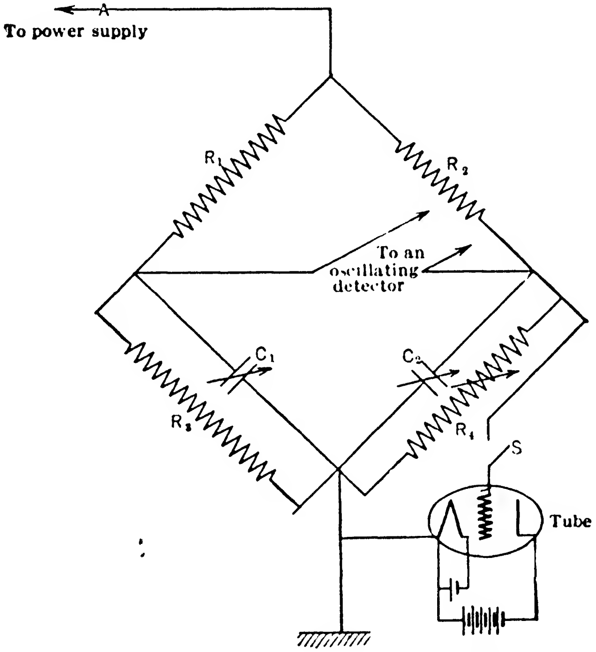


FIG. 70.—A suitable bridge arrangement for making high-frequency measurements.

12 micromhos, showing an input resistance of about 80,000 ohms. The values of E_p and E_g used are noted on the curve sheet.

In Fig. 72 is shown the variation of the input conductance as plate voltage was varied; this decrease in conductance with increasing plate potential could have been predicted from inspection of curves such as given in Fig. 43. In Fig. 73 is shown the variation in input conductance as the grid is made more negative, and in Fig. 74 is shown the effect of the magnitude of the alternating voltage impressed on the grid for testing.

The variation in grid conductance in Fig. 73 represents a variation in loss in the input circuit as the " bias " of the grid is changed. This has been used in one well-known scheme for controlling the amount of regeneration in tuned radio-frequency amplifier circuits.

It is evident from the four curves given above (which are all for the same tube) that if the input resistance of a tube is to be kept high the grid must at all times be negative (with respect to the negative end of the

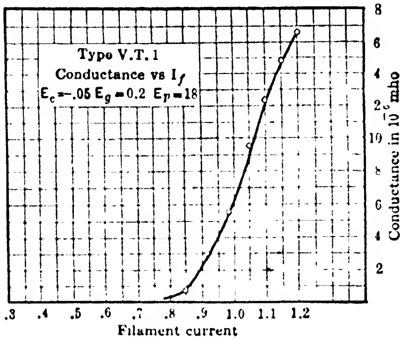


FIG. 71.—Variation of input circuit conductance with filament current.

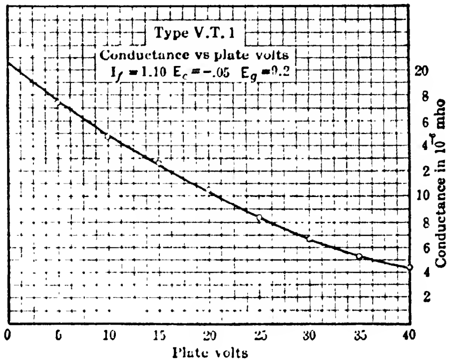


FIG. 72.—Variation of input circuit conductance with plate voltage.

filament). For the tube the characteristics of which are given above, the grid should normally be negative about 0.5 volt more than the maximum value of the voltage to be impressed on the input circuit. The resistance of the input circuit, as one component of the impedance of the

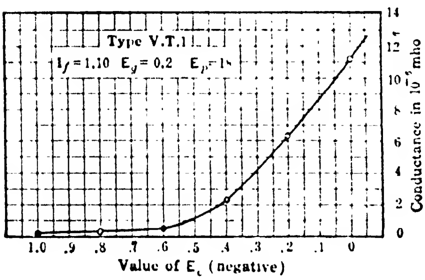


FIG. 73.—Variation of input circuit conductance with grid potential.

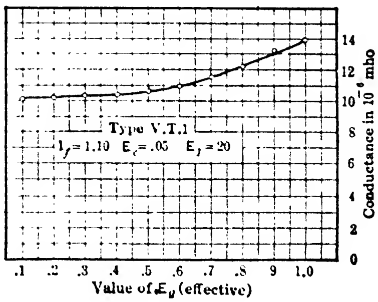


FIG. 74.—Variation of input circuit conductance with amplitude of voltage impressed on the input circuit.

input circuit, is of great importance if the tube is to be used as detector or amplifier; if the tube is to be used as an amplifier the input-circuit resistance may very seriously affect the selectivity of the receiving circuit, because of its damping effect on the signal.

In Fig. 75 is shown the variation of resistance of the input circuit of the type 201A triode; it is seen from the curve that there must be several volts of negative bias on the grid of this tube if the input circuit is to show a high resistance.

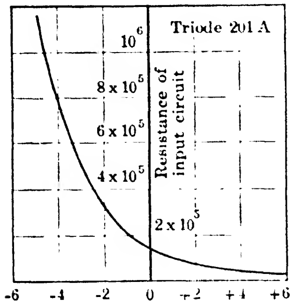


FIG. 75.—Variation of input circuit resistance of type 201A triode, as grid bias is altered.

Inter-electrode Capacity.—The capacity between the various electrodes of a triode is easily measured by means of the bridge shown in Fig. 76. The two resistors R_1 and R_2 are the ratio arms and each has about 500,000 ohms resistance. The two small condensers C_1 and C_2 are on the same shaft, so arranged that as the capacity of one increases that of the other decreases. C_3 is a fixed condenser of about 50 $\mu\mu f$ capacity. C_4 is a calibrated straight-line capacity condenser of about 40 $\mu\mu f$, and C' is a small uncalibrated vernier condenser to make the bridge balance, with nothing connected to posts 1, 2, and 3, and with C_4 set at its zero (maximum capacity). C_1, C_2, and C' are varied to get this balance.

The triode is then connected (by a special socket) to the bridge at 1, 2, and 3, say with the plate, grid, and cathode in the relative positions shown by p, g, and c. Condenser C_4 is then diminished to restore the bridge balance; it may be necessary to change the setting of C_1–C_2 to obtain a good balance. The amount of capacity by which C_4 is diminished to get the new balance is the capacity between those electrodes of the tube connected to posts 1 and 2. If the special connection from 3 to M is not used, the capacity so measured is C_{gp} increased by that of C_{cp} and C_{cg} in series. (C_{gp} = grid to plate capacity, C_{cp} = cathode to plate capacity, and C_{cg} = cathode to grid capacity.)

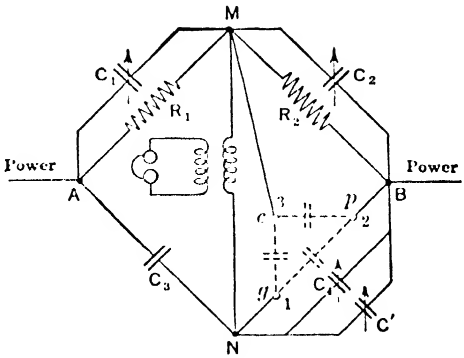


FIG. 76.—Special bridge arrangement for measuring inter-electrode capacities.

By reinserting the triode properly in the special socket connected to posts 1, 2, and 3, and again balancing the bridge, the other two interelectrode capacities can be measured.

In certain tetrodes the plate is shielded from the control grid by another grid structure which completely surrounds the plate; this screen grid has the effect of reducing the capacity between the plate and control grid to practically zero. To do this, of course, it must be properly connected to the plate-cathode circuit.

In the accompanying table are given the capacities of some of the commonly employed triodes, as well as a few of the special tetrodes. For information regarding power rating of these tubes, etc., see the table on p. 708.

Type	C_{gp}	C_{cg}	C_{cp}
171A	8.2 $\mu\mu f$	4.5 $\mu\mu f$	2.5 $\mu\mu f$
226	8.1	3.6	2.1
227	3.3	3.6	2.8
237	2.0	3.5	2.2
864	2.3	5.4	3.5
250	9.0	5.0	3.0
841	8.0	5.0	3.0
50 watt	12.0	9.0	8.0
250 watt	60.0	40.0	25.0
222	0.025	3.5	12.0
224	0.01	5.0	10.0
232	0.02	5.8	11.6
235	0.01	5.0	10.0

The last four tubes are of the screen-grid type; they have almost negligible capacity between control grid and plate, but this is offset to some extent by the greatly increased capacities C_{cp} and C_{cg}. The capacity C_{cg} is called the input capacity; the capacity C_{cp} is called the output capacity. The former acts in parallel to the tuning condenser of the input circuit (if used in a r.f. amplifier), and the latter acts to shunt the load circuit into which the tube is supplying its output.

Instead of measuring interelectrode capacity by the bridge method outlined above, Loughren and Parker recommend impressing a radio-frequency voltage and actually measuring the charging current.[1] As an alternative, they suggest the use of a standard variable condenser in parallel with the interelectrode capacity, and by obtaining the same current in the circuit with and without the triode connected, its interelectrode capacity is obtained from the two settings of the variable condenser.

Input Capacity Effected by Plate Circuit.—In practically all circuits involving the use of a vacuum tube it is required to have an impedance of some sort in the plate circuit; this impedance may be a resistance, a choke coil, or the primary winding of a transformer, and the value of this impe-

[1] I.R.E., June, 1929, p. 957.

dance is generally of the same magnitude as the a.c. resistance of the plate circuit of the tube, R_p, or somewhat greater.

When such an impedance is used in series with the B battery the voltage on the plate E_p varies when the grid voltage E_g is varied and the amount of fluctuation in E_p is generally much greater than E_g. If an impedance is used in the plate circuit, which is very high compared to the tube resistance, the fluctuation of E_p is nearly equal to μE_g. It is always somewhat less than this value, and we put it equal to αE_g where α lies between zero and μ, depending on the plate circuit impedance.

Let us suppose a resistance, R, used in the plate circuit; it is at once evident that as E_g increases, increasing thereby I_p, E_p must fall because of the increased value of I_pR. The forms of E_g, I_p and E_p are then as shown in Fig. 77; when the grid voltage rises (with respect to the filament) the plate voltage falls, and the actual plate voltage is represented by $(E_{op}-i_pR)=(E_{op}-\alpha E_g \sin pt)$, where $E_g \sin pt$ is the voltage impressed between the grid and filament, i_p is the instantaneous value of $I_p \sin pt$, the resulting fluctuation in plate current, and E_{op} is the value of the plate voltage when E_g is zero.

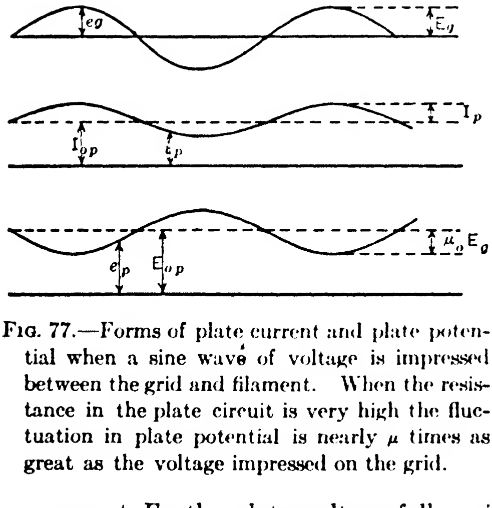


FIG. 77.—Forms of plate current and plate potential when a sine wave of voltage is impressed between the grid and filament. When the resistance in the plate circuit is very high the fluctuation in plate potential is nearly μ times as great as the voltage impressed on the grid.

We have then to consider the charging current taken by the grid when acted upon by an alternating voltage E_g, the condenser C_{G-F} being charged by voltage E_g and the condenser C_{G-P} in parallel, being charged to a voltage $(\alpha+1)E_g$, as shown in Fig. 78. The factor $(\alpha+1)$ occurs because when the grid voltage rises with respect to the filament, an amount E_g, the plate voltage falls, *with respect to the filament*, by an amount αE_g; it therefore falls *with respect to the grid*, an amount $(\alpha+1)E_g$.

The amount of charging current, therefore, which must be furnished by the input circuit is given by

$$I=2\pi f E_g(C_{G-F}+(\alpha+1)C_{G-P}),$$

from which the effective capacity of the input circuit is found to be

$$C_{\text{Input}}=C_{G-F}+(\alpha+1)C_{G-P}. \qquad (15)$$

Thus the effective capacity of the input circuit is not only much greater than the geometrical capacity, but it varies with any factors which affect α, the voltage amplification factor of the tube and circuit.

Due to the fact that the two voltages E_p and E_g are not exactly 180° apart, the capacity of the input circuit of a tube will actually be somewhat less than that predicted from Eq. (15).

This mutual capacity of the two condensers brings in another very interesting phenomenon: *the field of the grid-plate condenser may so react on the grid-filament condenser as to give a voltage in this condenser in phase with the impressed e.m.f. of this condenser (i.e., the e.m.f. impressed on the input circuit) so as to give the input circuit a negative conductance.* Such an effect would result in the plate circuit reacting on the input circuit to augment any voltage impressed on the input circuit.

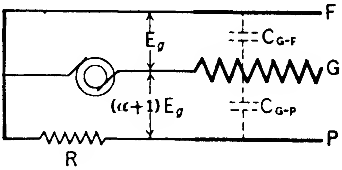


FIG. 78.—The circuit impressing the voltage E_g to the input circuit must furnish enough current to charge the condenser C_{G-F} to a voltage E_g and the condenser C_{G-P} to a voltage $(a+1)\ E_g$.

Using the bridge scheme illustrated in Fig. 70, the capacities and conductances of the input circuits of several of the typical commercial triodes were measured at 50,000 cycles. In Fig. 79 are shown the capacity and conductance of an amplifying tube with normal conditions of plate voltage, filament current, etc., as the external plate circuit resistance was varied; on the same curve sheet is shown the value of the voltage amplification factor of the tube for the various plate circuit resistances. It is seen that the capacity of the grid-to-ground circuit (same as input circuit, because the filament is generally grounded) increases from 17 $\mu\mu f$ (micro-micro-farads) to 71 $\mu\mu f$ as the plate circuit resistance was increased from zero to 80 kilohms.

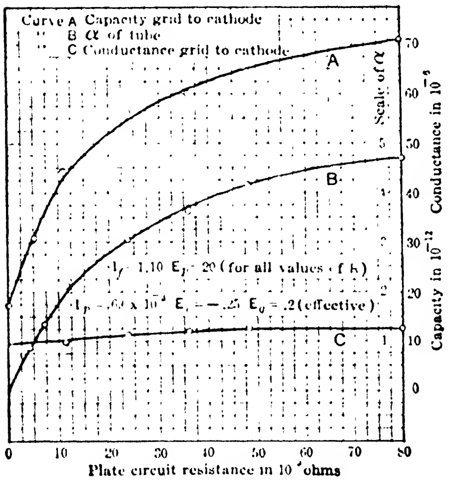


FIG. 79.—Capacity and conductance of an amplifying tube as the resistance in the plate circuit is varied; the α of the tube and circuit is shown also, so that the dependence of the effective input capacity upon α may be noted.

As the capacity C_{G-F} of this tube was 5.9 and the capacity C_{G-P} was 11.1 $\mu\mu f$, and the value of α is 4.65 for $R = 80$ kilohms, it might be expected that the input capacity would be equal to $(5.9+(4.65+1)\ 11.1)$

$= 68.6\ \mu\mu f$. The conductance of the input circuit was positive for all values of plate circuit resistance and gradually increased as R was increased.

In Fig. 80 are shown the capacity and conductance for the same tube, the plate circuit impedance being an inductance with a reactance resistance ratio between 25 and 50. In this case the increase in capacity is greater than when an equal amount of resistance was used in the plate circuit. Thus, with a reactance in the plate circuit of 50 kilohms the input circuit has a capacity of 82 $\mu\mu f$, whereas a resistance of 50 kilohms gave an input capacity of only 65 $\mu\mu f$. This difference in behavior of reactance and resistance is due to the fact that the α of the circuit is greater in one case than in the other, as will be explained later.

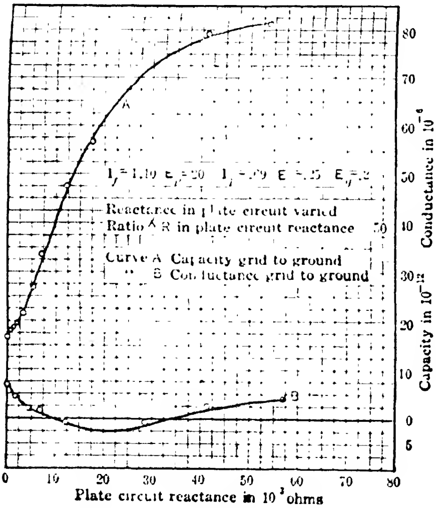


FIG. 80.—Capacity and conductance of the input circuit of an amplifying triode, as the plate circuit reactance is varied. Note that the input conductance is negative throughout a certain range of the reactance.

That any capacity present between the grid and plate, and which is not in the field of the grid-filament condenser, is increased by the factor $(\alpha+1)$ was proved by actually connecting a capacity of 20 $\mu\mu f$ across the plate-grid terminals of the tube and noting the increase in the effective capacity of the input circuit, the α of the circuit being 4.2. The capacity of the input circuit increased by 102 $\mu\mu f$, whereas calculation would make it increase by $(4.2+1)\times 20$, or 104 $\mu\mu f$.

The conductance of the input circuit of the tube was negative throughout a certain range of plate circuit reactance, thus indicating transfer of power from the plate circuit back to the grid circuit, with no other coupling between the grid and plate circuits than what existed in the tube itself. This curve shows that the three-electrode tube is not inherently a "one-way repeater," as has been commonly supposed; the output circuit does control the input circuit to an appreciable extent,[1] sufficient in fact to maintain the tube in operation as a generator of a.c. power when it is connected to the proper circuit. If the grid circuit and plate circuit are each tuned to the same frequency, as indicated in Fig. 81,

[1] It will further be appreciated that the grid-plate capacity serves as a coupling condenser for the input and output circuit, resulting in the double resonance peak tuning condition analyzed on pp. 133 et seq.

the tuning condensers are sufficiently small (and the coils fairly efficient), the coupling of the two circuits *inside the tube* may be sufficient to maintain the tube in the oscillating state, alternating currents flowing in circuits L_2C_2 and L_1C_1.

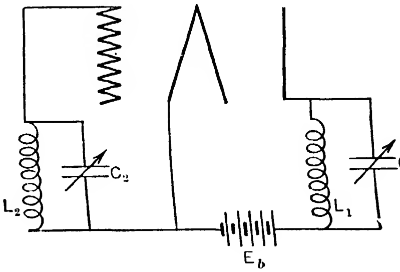


FIG. 81.—In such a circuit as this, with efficient coils used in both circuits, with suitable values of the capacities the tube will maintain itself in an oscillatory state, due to the negative conductance as shown in Fig. 80.

In Figs. 82 and 83 are shown the characteristic curves of two of the other tubes tested. It is seen that the same general shape holds for all three electrode tubes, the difference being one of degree only. The capacity of the grid-ground circuit of the ordinary tube, when it is operating with the normal amount of resistance (or reactance) in the plate circuit, is from five to ten times as much as the geometrical capacity of the circuit, and the amount of this increase *is controlled principally by the capacity between the grid and plate.*

As has been pointed out, the characteristics of the input circuit of a tube depend upon the relative phases of the input voltage and the voltage variation between the plate and filament. As this phase relation will evidently depend upon the kind and amount of reactance in the external portion of the plate circuit we may expect the input characteristics to vary with the input frequency because this will determine the reactance, other things being constant. This effect has been investigated theoretically by Ballantine,[1] who shows that for resistive plate circuit the effective input capacity decreases with an increase in frequency and the input conductance increases with an increase in frequency. For reactive plate circuit the effect of frequency may be to either decrease or increase the input circuit constants, depending upon the amount of the reactance used.

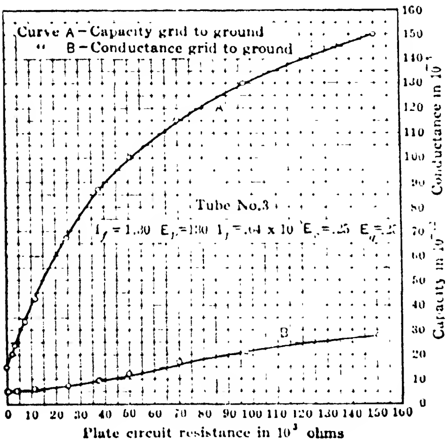


FIG. 82.—Variation in conductance and capacity of the input circuit of a telephone repeater tube as plate circuit resistance is varied.

[1] Stuart Ballantine, "The Thermionic Amplifier," Phys. Rev., Vol. XV, No. 5.

Reducing the Input Capacity.—The effect of the input capacity of a tube is frequently troublesome; it makes the tuning of the input circuit of the triode depend to a certain extent on the adjustment of its plate circuit and, more important, the negative resistance caused by the overlapping of the grid-filament condenser and plate-filament condenser frequently causes undesired oscillations.

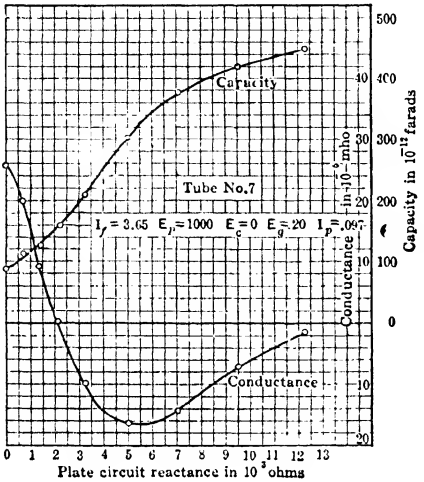


FIG. 83.—Variation of conductance and capacity of the input circuit of a large power tube as plate-circuit reactance is varied.

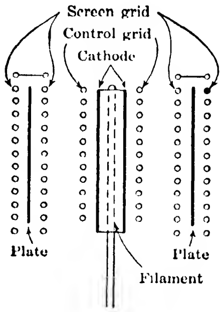


FIG. 84. The shielded grid tube; the shielding grid is made to completely surround the plate.

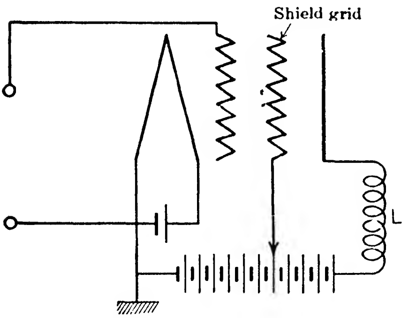


FIG. 85.—By this connection the shielding grid is prevented from changing its potential with respect to ground, thus preventing the varying electric field of the plate from reaching through to the control grid.

With the ordinary type of triode it is possible to use a circuit arrangement which neutralizes the effect of this mutual capacity effect; such a scheme has been much used in tuned radio-frequency amplifiers, where the undesired oscillations are very likely to occur. But it is possible to eliminate the trouble by using a specially designed four-electrode tube, in which a special shielding grid is used. The general idea of this form of tube is shown in Fig. 84. The control grid is the ordinary helical one of fine wire and the plate has the customary cylindrical form. The shield grid is a double helix, one inside and one outside the cylindrical plate. This shield grid is connected to ground (in so far as alternating fields are concerned) as indicated in Fig. 85. Actually it is connected directly to a suitably positioned point in the B battery, but its positive potential, so fixed, cannot vary with respect to ground. This construction therefore prevents the electric field of the plate reaching through and affecting the control grid. That is, no matter how much the plate voltage fluctuates, with respect to ground, due to the potential variations across coil L, these fluctuations cannot be felt at all in the grid-filament circuit. It is claimed by Hull that this construction so stabilizes the amplifying action of the triode that he finds it possible to build amplifiers having a voltage amplification of 100,000 times, and experience no trouble from undesired oscillations.

Transconductance of a Triode.—The real function of a triode used as an amplifier is to produce a large change in the plate-circuit power for small changes in grid voltage. By combining the plate circuit resistance with the amplification constant of the tube we get some idea as to how well the tube performs this function.

We have

$$\mu=\frac{de_p}{de_g}$$

with i_p constant.

$$r_p=\frac{de_p}{di_p}$$

with e_g constant.

Then

$$\frac{\mu}{r_p}=\frac{di_p}{de_g}=s_m. \qquad (16)$$

This quantity, s_m, has been called the *mutual conductance* of the triode but is now properly called the *grid-plate transconductance*, and this is generally abbreviated to merely *transconductance*. Being the reciprocal of a resistance (as μ has no dimensions) it is measured in mhos. The mho is, however, too large a unit for the ordinary triode, so that s_m is ordinarily measured in micromhos.

By multiplying the signal voltage, e_c, by the transconductance s_m, we find the amount of alternating current set up in the plate circuit by the signal, this on the assumption that there is no external impedance in the plate circuit.

In Fig. 86 are given the characteristics of the two standard receiving tubes being used today. The UV-199 tube uses about 0.06 ampere in its filament, at 3 volts, and the UV-201A uses 0.25 ampere at about 5 volts. With 100 volts on the plate of the 201A the data of Fig. 86 tell us that there will be set up in this plate circuit, by a signal of 1 volt on the grid, a fluctuation of plate current equivalent to an alternating current of 800 microamperes.

If a load is connected in the plate circuit, a loud speaker, for example, of the same resistance as the plate circuit of the triode itself (10,000 ohms) then the 1-volt signal will produce only one-half as much alternating current in the plate circuit, that is, 400 microamperes.

Tubes with Gradually Changing Transconductance.—In Fig. 68 there are shown two types of triode construction which result in variable μ, this decreasing as the negative bias of the grid is increased. Such a construction results in a value of transconductance which varies more

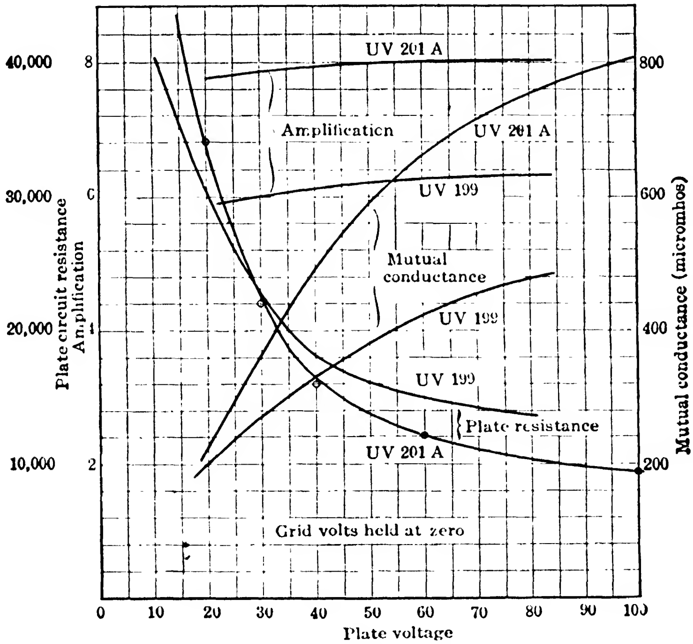


FIG. 86.—Transconductance, amplification factor, and plate-circuit resistance of two of the triodes much used in the past.

gradually as the grid bias is used in amplification control, than does the ordinary grid construction. This is a desirable characteristic, and variable μ tubes are used in this way in modern receivers. In Fig. 87 are shown two curves of transconductance, both for modern tubes; type 24 has a uniform grid construction and type 35 has a variable μ construction. This type of tube (gradually varying s_m) has other advantages which will be discussed in Chapter VIII under the topic of modulation distortion.

Operation of Three-electrode Tube as Detector of Damped-wave Signals. Grid Condenser. Leak Resistance. Normal Grid Potential.— Any detector of high-frequency currents must in some way cause low-frequency pulsations of current through the telephones when the device itself is actuated by high-frequency currents. The frequency of the low-frequency pulsations is fixed by the number of damped-wave trains arriving at the antenna per second in case of reception of a signal from a spark station, and is fixed by local conditions when receiving from a continuous station. The case we shall consider in this section is for spark signals only; damped-wave trains of the form shown in Fig. 88 are to be detected by the three-electrode tube. The time between wave trains A may be from 0.005 to 0.0005 second; the duration of a wave train B may be from 0.00001 to 0.001 second, and the time of one cycle, C, may be from 0.0000001 to 0.00003 second.

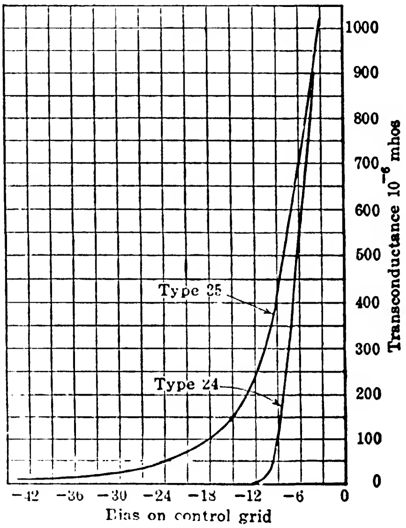


FIG. 87.—Showing the grid-plate transconductance of two modern receiving tubes. The type 35 is a variable μ tube.

Whereas the following discussion is carried out for spark signals, which have today comparatively little importance, it has direct applicability to radio telephone reception. Exactly the same ideas are involved in the detection of telephone signals as for the reception of spark signals.

The function of the detector is to produce, in the telephone, fluctuations of current, of frequency fixed by the time A, as large as possible with a given

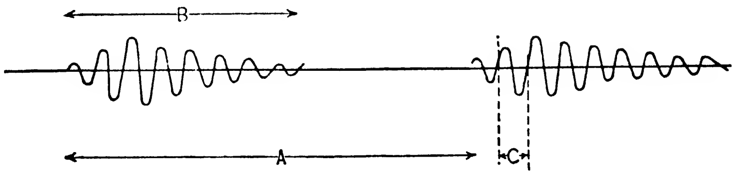


FIG. 88.—Conventional representation of part of a damped-wave signal.

amplitude of signal voltage. The scheme of connections used when no condenser is inserted in series with the grid of the tube is shown in Fig. 89; the ground terminal of the input circuit is generally connected to the negative end of the filament or to some point in the circuit at a lower potential than the negative end of the filament. This is possible by either of the two schemes sketched in Fig. 90; in (*a*) a resistance R is inserted in the negative filament wire and the potential of point A is thus lower in potential than the negative end of the filament by an amount I_fR, generally 1 volt or less, whereas in (*b*) a battery C is inserted in series with the input circuit to properly lower the grid potential. In case a careful adjustment of this potential is desired (generally not necessary) the grid may be connected to battery C through a suitable potentiometer connection.

In modern radio sets the power supply for plates and grids is obtained from a rectified alternating power supply. This power supply circuit always has a high resistance connected across the rectified supply, and the proper plate potentials and grid potentials are obtained by connecting the respective circuits to the proper points on this resistance.

The reason for maintaining the grid at a negative potential is evident

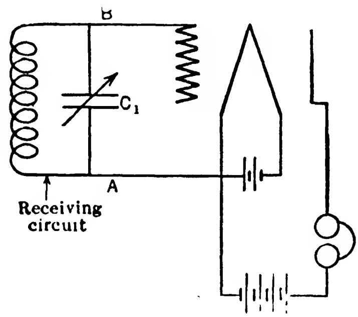


FIG. 89.—Connection scheme for using a three-electrode tube as detector, without use of a condenser in series with the grid.

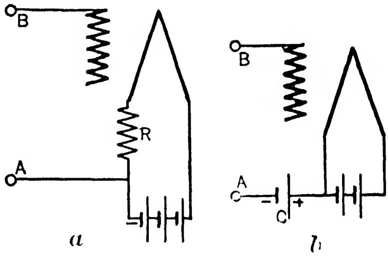


FIG. 90.—Two schemes for maintaining the average potential lower than the lowest potential point of the filament.

in looking at the input circuit conductance curves previously given; suppose the conductance of the grid circuit is 10^{-5} mho, and the signal being received is 600 meters, the tuning condenser C_1 (Fig. 89) being set at 200 $\mu\mu f$. A conductance of 10^{-5} mho is equivalent to a shunt resistance of 10^5 ohms around condenser C_1, and this is approximately equivalent (by Eq. 38, Chapter II) to a resistance of 25.4 ohms in series with C_1. But such a large resistance would materially interfere with the selectivity of the receiving circuit, in fact would make it practically useless if there was much interference; the resistance of the receiving circuit itself would be only a few ohms, perhaps five.

When discussing radio telephony there will be analyzed another reason for maintaining the grid at a suitable negative potential, namely the distortionless reproduction of speech and music.

The characteristics of the tube being as shown in Fig. 91, the normal grid potential being E_{og}, the question is how much will the telephone current (Fig. 89) change during the time one of the wave trains of Fig. 88 is actuating the grid. By actually plotting the values of plate current for each grid potential we get the curve of plate current shown by i_p in Fig. 91, while the grid potential is undergoing the changes indicated by the curve e_g. The increase in the average value of the plate current is indicated by the dotted line in Fig. 91, and this average increase, during the time the grid is being excited by a wave train, is what determines the response of the telephone diaphragm. Such a use of the static characteristic of the tube is permissible only if the receiving circuit is so arranged that, as the signal is received, the plate potential does not appreciably vary; this condition implies an external plate circuit of impedance which, compared to the internal plate resistance, is negligible for the frequency of the signal.

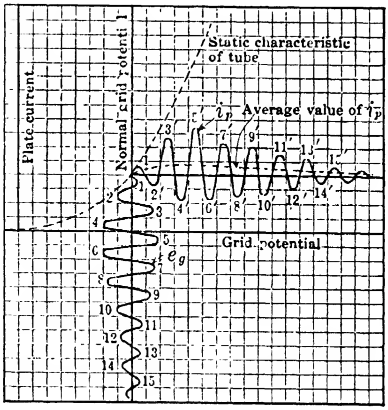


FIG. 91.—Analysis of the action of the three-electrode tube as detector of damped wave signals; assuming a certain variation in grid potential the resulting fluctuation in plate current can be plotted from the plate current, grid-potential curve of the tube.

The author arranged a tube circuit so that its input voltage and plate current could be recorded on an oscillogram, when a damped sine wave of about 100 cycles (having the general form of an actual wave train from a highly damped spark station) was impressed on the input circuit; some of the films obtained are presented herewith. In Fig. 92 are shown the input voltage, plate current and telephone current when the grid was made normally 2.5 volts negative with respect to the filament. A large capacity condenser was shunted around the coil representing the telephone of an ordinary receiving set so that the " high-frequency " current was not forced to flow through this coil. This condenser charged up during the first part of the wave train more rapidly than it discharged through the coil, so that its charge increased. Then as the wave train was reduced to zero by damping, the fluctuations in plate current ceased and the condenser continued to discharge through the coil; this action caused the current through the coil to lag somewhat behind the wave train impressed on the grid, as is evident from the film.

The signal used in getting this film, as well as those to follow, was much stronger than an actual radio signal; the change in " telephone " current in Fig. 92 is about 5 milliamperes, whereas actually a fairly strong radio

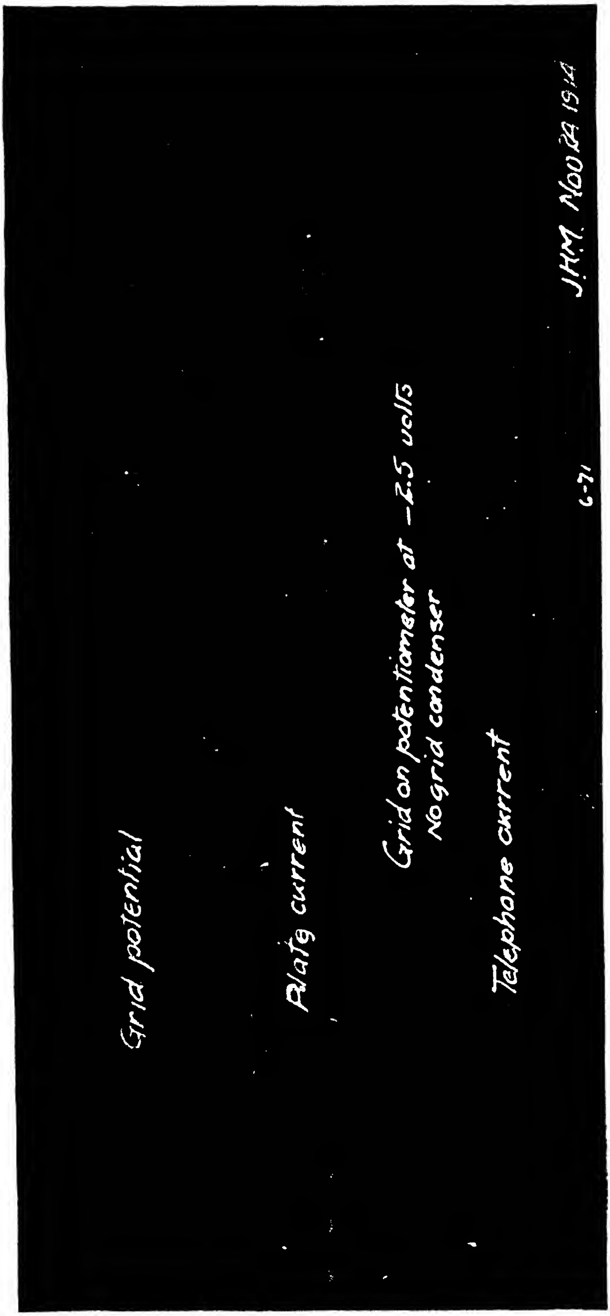


FIG. 92.—Oscillogram showing rectification without grid condenser; these actual curves agree exactly with curves predicted as in Fig. 91. The frequency of the artificial signal used in getting this film was about 100 cycles per second.

signal may not produce a change in the telephone current of more than a few microamperes. Figs. 91 and 92 show the rectifying action of a tube brought about by the increase in plate current being greater than the decrease; the grid was put as such a negative potential that the tube was operating well down on its characteristic about as indicated at A, Fig. 93.

The grid potential was then made positive sufficiently to rectify by giving a greater decrease than increase in plate current; Figs. 94, 95, and 96 show the forms of potentials and currents when putting sufficient positive potentials on the grid to bring it to points B, C and D (Fig. 93) respectively. It is to be noted in the film shown in Fig. 96 that at the highest positive grid potentials the plate current had actually decreased; the amount of current taken by the grid was sufficient to bring about a decrease in plate current. In each film the zero lines of potential and currents are shown.

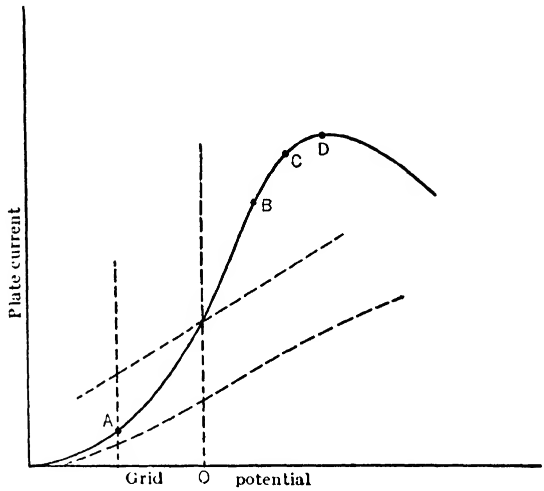


FIG. 93.—Form of the plate current, grid-potential curve of the tube used in getting the films of Figs. 92, 93, 94, 95.

An elementary analysis show that the efficiency of a tube for the purpose of detection (i.e., its rectifying power) depends largely upon the radius of curvature of the plate-current grid-potential characteristic. We put

$$I_p = f(E_g).$$

With no signal $I_{op} = f(E_{og}')$ and when the signal voltage ΔE_g is impressed on the grid

$$I_{op} + \Delta I_p = f(E_{og} + \Delta E_g)$$

$$= f(E_{og}) + \Delta E_g \left(\frac{dI_p}{dE_g}\right)_{\substack{at\\E_{og}}} + \frac{\Delta E_g^2}{2}\left(\frac{d^2 I_p}{dE_g^2}\right)_{\substack{at\\E_{og}}} + \ldots$$

Then we have as an approximation

$$\Delta I_p \doteq \Delta E_g \frac{dI_p}{dE_g} + \frac{\Delta E_g^2}{2}\,\frac{d^2 I_p}{dE_g^2}.$$

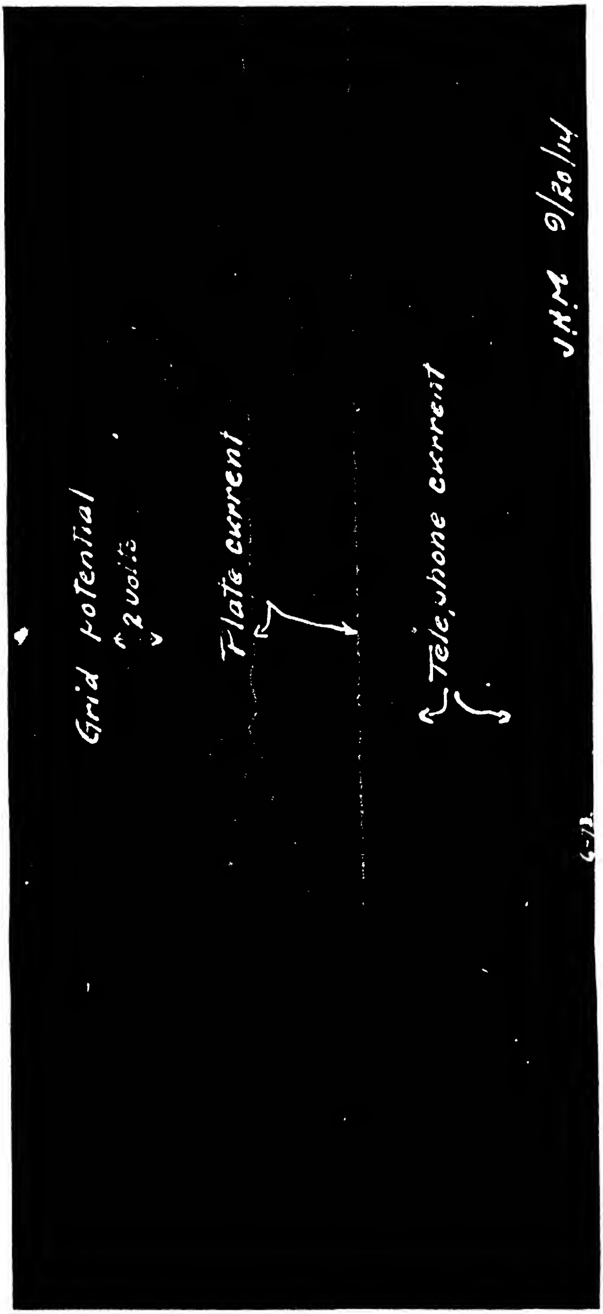


Fig. 94.—Rectifying action of tube adjusted for normal grid potential as at *B*, Fig. 92.

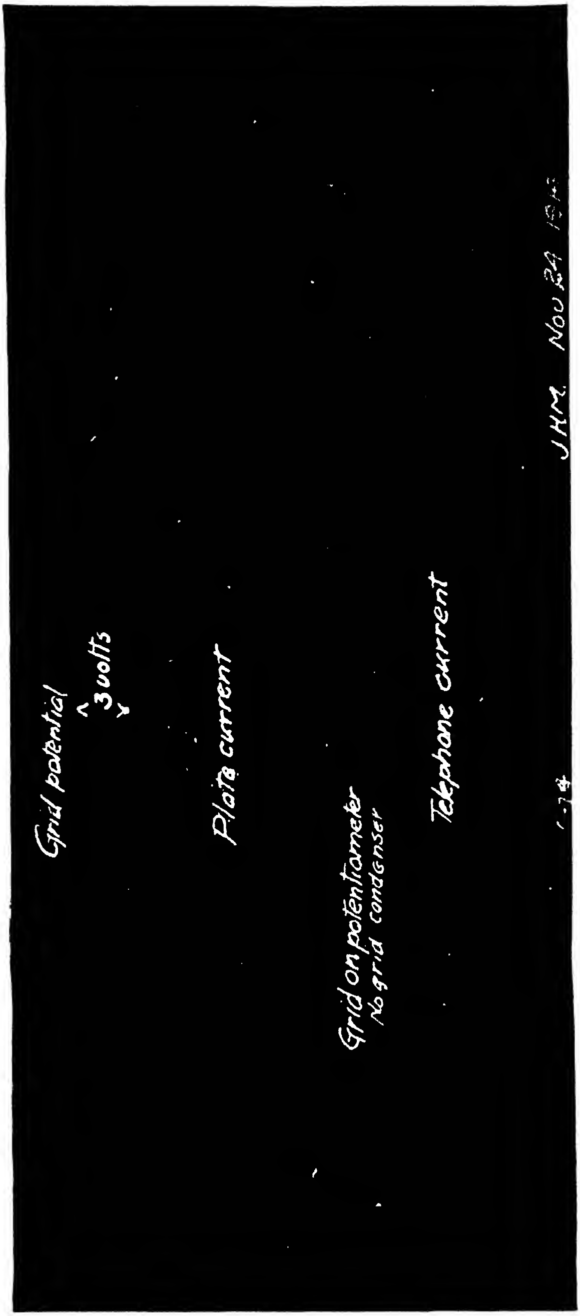


FIG. 95.—Rectifying action of tube adjusted for normal grid potential as at C, Fig. 92.

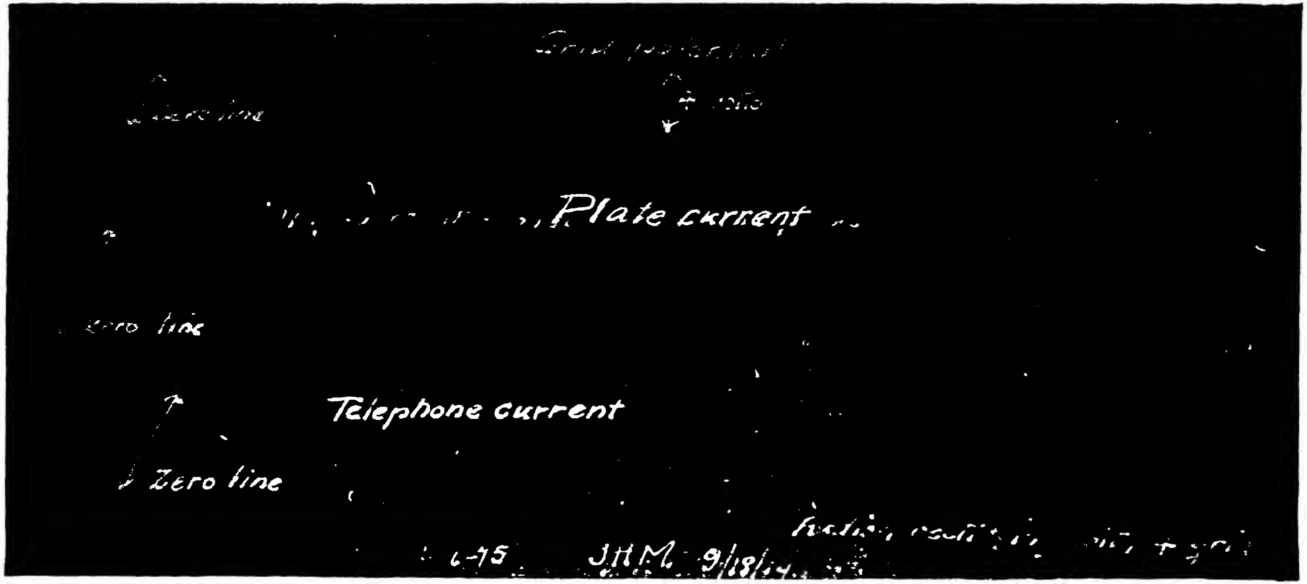


Fig. 96.—Rectifying action of tube adjusted for normal grid potential as at *D*, Fig. 92. Note that for the highest value of grid potential occurring during the application of the signal the plate current actually decreases.

If ΔE_g is periodic the average value of the first term, $\Delta E_g \dfrac{dI_p}{dE_g}$, is zero, so that the average value of ΔI_p becomes equal to the average value of $\dfrac{\Delta E_g^2}{2} \dfrac{d^2 I_p}{dE_g^2}$. Now, if ΔE_g is a sine function of time of the form, $E \sin pt$, we have for the average value of the change in plate current

$$\Delta I_p = \frac{d^2 I_p}{dE_g^2} \frac{1}{2T} \int_0^T E^2 \sin^2 pt d(pt) = \frac{E^2}{4} \frac{d^2 I_p}{dE_g^2}. \quad . \quad . \quad . \quad (17)$$

T to be taken as an even number of cycles.

The increment in plate current therefore varies with the *square* of the signal strength, a defect practically all rectifying devices have. That this property of the tube is a defect becomes apparent when we consider its application. The increment (or decrement) in plate current in what makes the radio-frequency signal audible to the operator, so we may say that the increment in plate current per volt of signal is a measure of the efficiency of the tube as a detector. The efficiency then increases directly proportional to the signal strength; Eq. (17) shows that the rectified current increases with the square of the signal strength.

It is to be remembered, of course, that in deriving Eq. (17) it was assumed that the plate current varies with the square of the grid voltage; that is, we assumed it was a parabolic curve. In so far as this is not so the conclusions we draw from Eq. (17) are incorrect. In general, it will be found that all types of tubes do follow Eq. (17) for weak signals, say small fractions of 1 volt, but for signals measured in volts the rectified current in the plate circuit increases less rapidly than Eq. (17) predicts. Thus, it is said, that for strong signals a suitably arranged tube gives *linear detection*; it is true only when the detector is handling powerful signals. Such a detector is often referred to as a *power detector*; the term is generally used only when the detector input is arranged to function properly with input signals measured in volts, say up to 10 or even 20 volts.[1]

Automatic Grid Bias Control.—Frequently a detector is arranged to have its grid bias automatically increase as the signal impressed on the grid is increased. This may be done by inserting a resistance in the plate circuit, next to the cathode, and using the drop across this resistance as the grid bias voltage. One scheme is indicated in Fig. 97; the resistance R of perhaps several thousand ohms, has an IR drop which increases with the average value of the plate current, that is, with the signal strength.

[1] A very thorough analysis of the detecting action of a triode, using the curvature of the plate-current curve to obtain rectification (so-called plate detection), has been given by Woods, as Engineering Research Series 28 published by the University of Texas as Bulletin 3114, April 8, 1931. Woods discusses the action of the tube for both weak and strong signals.

In Fig. 98 the effect of such a resistance is shown by curves B, C, and D. The two condensers of Fig. 97 are to by-pass the radio and audio frequencies to prevent the loss in rectification which would occur if these alternating currents had to traverse the resistance R. Condenser C_1 has perhaps 0.001-μf capacity, and condenser C_2 has 1 μf. The best values depend upon the amount of resistance in R and upon the type of tube used.

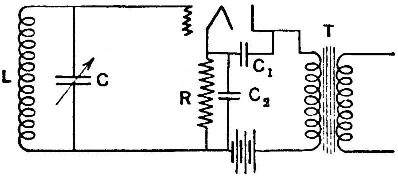


FIG. 97.—Detector circuit arranged for automatically controlled grid bias; the plate current, flowing through R, provides the bias.

It is to be pointed out that the condensers C_1 and C_2 in series form a by-pass around the primary of transformer T, and it may well be that for the higher audio frequencies these condensers which aid the rectifying action of the triode may actually diminish the amplifying action by forming a low impedance shunt around the transformer primary, or other high impedance load used in the plate circuit.

The curve of Fig. 91 is obtained by maintaining plate voltage constant; if there is a high resistance or reactance in series with the B battery, the effect is to straighten out the characteristic curve, and so decrease the value of d^2I_p/dE_g^2 throughout the whole extent of the curve as shown by the dotted line curve in Fig. 93. The reactance of a pair of phones, for radio-frequency current may be very high, hence the effect just mentioned might exist; to eliminate it a condenser should be used in shunt with the phones, thus furnishing a low impedance path for the high-frequency current and so maintaining the plate voltage essentially constant as the grid potential fluctuates. In Figs. 92, 94, 95, and 96, a condenser "by-pass" around the phones was used, its impedance for the frequency used was very much lower than that of the phones, so that practically all of the high-frequency pulsations took place through the condenser, the telephone current changing only as the average value of the plate current decreased.

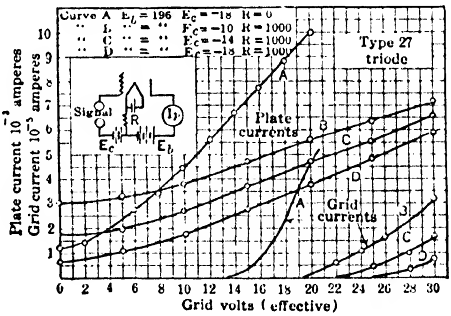


FIG. 98.—Action of a type 27 detector with automatic grid bias.

Plate Rectification of Typical Triode.—The rectifying action of a triode may be shown by impressing an alternating voltage signal, of adjustable strength, in the grid-cathode (input) circuit and reading the plate current with a c.c. ammeter. The amount of rectification is shown directly by this ammeter in the plate circuit; this will not respond to the alternating component of the plate current.

In Fig. 98 is shown the rectifying action of a type 27 triode, arranged with a large grid bias so that comparatively large signals may be impressed on the grid before this draws current. According to Eq. (17), the increment in plate current, ΔI_p, should vary as the square of the signal voltage, and the plate current curve of Fig. 98 follows this law approximately. The grid should not draw current unless it becomes positive; in Fig. 98 it starts to take current when the signal voltage equals 13.5 volts (effective). Such an amplitude of signal is just sufficient to make the grid go positive for a small faction of the cycle.

New types of tubes are continually appearing. Thus the type 27 triode, using 1.75 amperes at 2.5 volts in its filament is now replaceable by the type 56, which uses only 1.0 ampere at 2.5 volts. This triode used as grid bias detector, employs 250 volts on the plate, and sufficient bias to reduce the plate current to 0.0002 ampere when no signal is impressed on the grid. This requires about -20 volts on the grid. Used as a grid rectification detector the plate voltage is only 45; a grid condenser of 0.00025 μf and leak of 1 to 5 megohms is used.

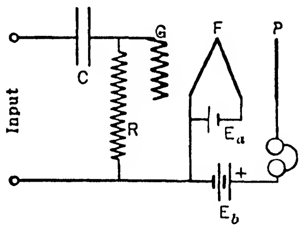


FIG. 99.– Arrangement of three-electrode tube for detection by use of a condenser in series with the grid.

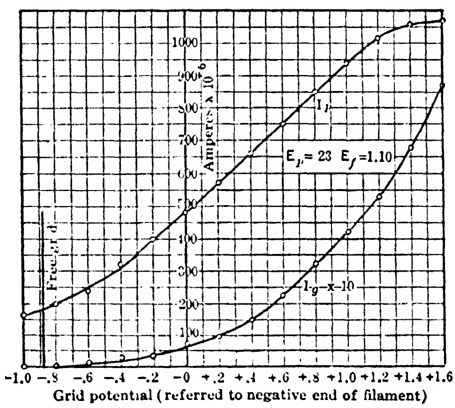


FIG. 100.—Plate and grid current for a typical triode.

The manufacturer suggests the use of this triode as a diode (Fleming valve). The plate and cathode are connected, and the signal is impressed between cathode and grid. A signal of 40 volts (eff.) may be rectified by this scheme.

Effect of Grid Condenser.—The average three-electrode tube will give better rectifying action if the curvature of the I_g–E_g curve is used instead of that of the I_p–E_g curve. The use of a suitable condenser in series with the grid enables us to utilize the curvature of the grid-current curve; the ordinary connection is shown in Fig. 99, the resistance R being about 1 megohm for the average tube. It is called the "leak" resistance and its function will be explained shortly. The potential of the grid (when signal is coming in) depends upon the value of the leak resistance, the form of the I_g–E_g curve, and upon the potential of the point to which the ground end of R is connected.

The form of the I_g–E_g curve for a typical detecting tube is shown in Fig. 100; the curves are shown for comparatively large change in the grid potential, much larger than generally occurs when the tube is being used. As would naturally be supposed, the free grid potential is that for which the grid current becomes zero in the graph; when free the grid potential

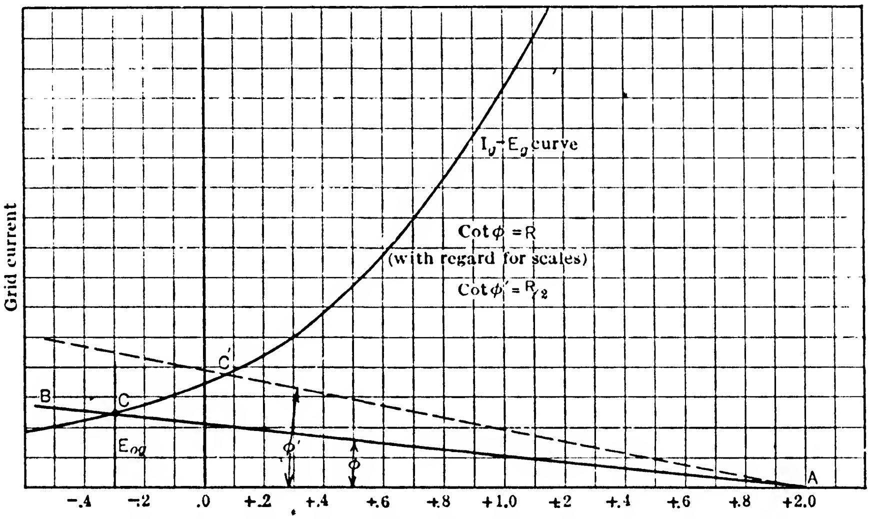


FIG. 101.—A diagram for determining the normal grid potential of a tube connected as in Fig. 99; the leak resistance is supposed to be connected to the positive end of the filament and the IR drop in the filament is assumed as 2 volts.

will decrease to such a potential that no more electrons tend to accumulate on it.

When using such a tube in the connection scheme shown in Fig. 99 the first point to be examined is the potential at which the grid will set itself when no signal is being impressed on the grid. It is common practice to connect the end of resistance R to the positive end of the filament, and we will so assume it in finding the normal grid potential. In Fig. 101 is shown the grid current (with enlarged scale for I_g); it is supposed that the IR drop in the filament is 2 volts. The straight line AB is drawn through the point $E_g = +2$ and at an angle such that cot $\phi = R$. The point C, where this line intersects the I_g–E_g curve, fixes the normal grid potential, E_{og}. This follows from the fact that whatever current flows to the grid must return to the filament (positive end) through the resistance R and so cause in this a drop of I_gR; furthermore this drop, added to the normal grid potential E_{og}, must give a voltage equal to $+2$ volts, the potential of the positive end of the filament.

If the leak resistance is 10^6 ohms, cot ϕ must be 10^6 when the scales of potentials and currents are in corresponding units as, e.g., volts and amperes. As the scale of current in Fig. 101 is 10^5 smaller than that of potential, the angle ϕ in this diagram is so drawn that cot $\phi = 10$. If a leak resistance of only 5×10^5 ohms were used, the normal value of grid potential E_{og} would be as shown at C', obtained by making cot $\phi = 5$.

When an alternating e.m.f. is now impressed on this input circuit, the grid will start to fluctuate about its normal value of potential, E_{og}; its potential will be increased and decreased from the value E_{og} equally for the first cycle. Due to the form of the I_g-E_g curve, however, the increase in current, when the impressed e.m.f. is positive, is greater than the decrease in current when the impressed e.m.f. goes negative, and this rectifying action tends to increase the number of electrons accumulated on that side of the condenser C (Fig. 99), which is connected to the grid. But this accumulation of electrons must depress the potential of the grid below its normal value, and so cause a decrease in the plate current. The amount of this decrease in plate current for a given alternating e.m.f. impressed on the input circuit, is a measure of the efficiency of the tube as a detector, so we shall investigate this point more fully.

Before starting this analysis it is well to point out that whereas a tube may detect by either an increase or decrease in plate current when no grid condenser is used, with the grid condenser a signal always produces a decrease in plate current, never an increase.

At the end of the wave train the grid condenser C (Fig. 99) will be charged (negatively on the side connected to the grid) and this charge must leak off before the next wave train arrives, otherwise the tube will not respond to a signal as well as it should. The time taken for the charge to leak off from C depends upon the magnitude of C and the leak resistance R, in fact, can be calculated directly from these two quantities. In a time equal to RC, 63 per cent of the charge will have leaked off; if the tube is to operate efficiently as a detector therefore the product RC must be small compared to the time between the successive wave trains of the signal.

On the other hand, C must be as large as feasible and R also must be large, otherwise a large fraction of the signal voltage will be used up in C, and thus be of no service in producing sound in the telephones. The input circuit of Fig. 99 may be represented as in Fig. 102; C is the external condenser used in series with the grid, C' is the capacity of grid-to-ground inside the tube and r is the leakage inside the tube itself. The values of C' and $1/r$ (tube conductance) for various tubes were given in Figs. 71–83; the impedance between D and B. Fig. 102, is therefore calculable when R is given. Designating this impedance by Z_t, we then find that the voltage impressed on the grid of the tube is equal to the input voltage (across points A–B, Fig. 102) multiplied by the fraction

$$\frac{Z_t}{\left(Z_t+\frac{1}{\omega C}\right)},$$

the addition and division being carried out vectorially.

In addition to the features just analyzed we must remember that the impedance between points A–B is to be kept high as this input circuit is connected directly across the tuning condenser of the receiving set. With the detecting tubes commonly used it seems that $C=5\times10^{-10}$ and

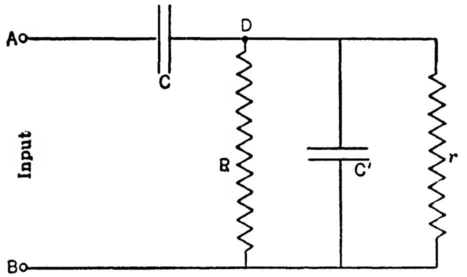


FIG. 102.—A circuit equivalent to the input circuit of a detector tube; C' represents the effective capacity of the input circuit and $1/r$ represents the conductance of the input circuit. These quantities for different tubes were shown in Figs. 71–83.

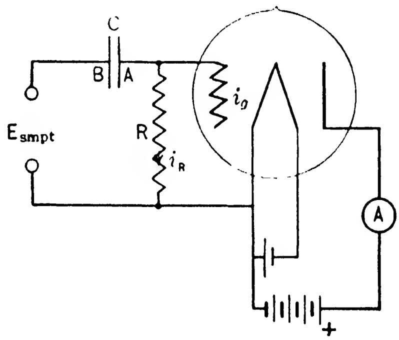


FIG. 103.—When detecting by grid condenser, the decrease in average grid potential is brought about by an increase in the current i_R.

$R=10^6$ give the best results. For tubes having smaller internal capacity lower values of C and higher values of R are better suited; thus the detecting tube shown at J, Fig. 31, is generally used with $C=4\times10^{-10}$ and $R=4\times10^6$.

The magnitudes of R and C must be determined by a certain compromise involving sensitivity and selectivity of the receiver, and so-called frequency distortion, to be analyzed in a following section.

Analysis of Detector Action with Grid Condenser.—The arrangement we are to discuss is shown in Fig. 103. When the signal $E \sin pt$ is absent the plate current has a certain value, I_{op}, and when the signal is impressed this current *decreases* by a certain amount, ΔI_p, which we wish to calculate.

First analyzing the action qualitatively we have the following facts to consider. The potential of the grid, with no signal, E_{og}, must be controlled by the resistance R, in fact it is equal to $I_R R$. With no signal $I_g = I_R$ and both of these currents are constant in value. If $R = 0$, $E_{og} = 0$ and if $R = \infty$ then E_{og} is equal to the "free grid potential" of the tube, ordinarily a volt or two lower than the negative end of the filament.

When the signal is impressed a part of its magnitude is used up as IX drop in the condenser C but we will neglect this for the moment and suppose the whole signal voltage $E \sin pt$ appears between the grid and filament. Now the shape of the grid current-grid voltage curve results in a greater increase in grid current when the signal is positive than decrease in grid current when the signal is negative. That is, the grid current experiences the same action as does the plate current in Fig. 91. This net increase in grid current first accumulates a charge of electrons on the grid and on plate A of condenser C, thus making the grid go more negative than when no signal is impressed. After the condenser C is charged as much as the signal can accomplish the steady state is reached and the average value of i_g must now be the same as the average value of i_R. The average value of the grid voltage with signal, E'_{og} will be more negative than the value of E_{og} because of the charge on condenser C. The signal voltage will now make the grid fluctuate up and down around the potential E'_{og} whereas when the signal is first impressed it fluctuates about E_{og}.

As the average potential of the grid has been depressed the average value of the plate current will have been *decreased*, we conclude. However, it is not decreased as much as we might suppose because of the curvature of the plate current-grid potential curve.

When the grid fluctuates about its potential E'_{og}, the average value of the plate current is *greater* than the value indicated by the grid potential E'_{og} because of the action shown in Fig. 91.

So the detector action with grid condenser involves two opposing actions, grid rectification tending to *decrease* the average plate current and plate rectification tending to *increase* the average value of the plate current. The former action always predominates so that detection with grid condenser is always accompanied by a decrease in the average value of plate current when a signal is impressed. In Fig. 104 this detector action is shown; in the diagram the curvature of the grid current has been somewhat exaggerated (over that of the actual curve in Fig. 100) to make it possible to clearly represent very small differences. Thus the change in average grid voltage, ΔE_g, is shown one-third the value of the signal voltage E. Actually the grid current curve may have much sharper curvature than this, and does have in some triodes. In Fig. 105 are shown the forms of grid current-grid voltage curves for a type 201 A triode and for a type 27 triode. It is seen that the latter, a heater-type tube, has very sharp curvature in its grid current characteristic.

Before the signal comes in the grid current is steady (Fig. 104), at the value, I_{og}, and this current is flowing through R, holding the grid at potential E_{og}. As soon as the signal, $E \sin pt$, is impressed the grid current fluctuates from a to b, having an average value shown by the dashed line A. The excess of this current A, over I_{og}, flows partly into condenser C and partly through R. In the steady state, with signal impressed, no current can

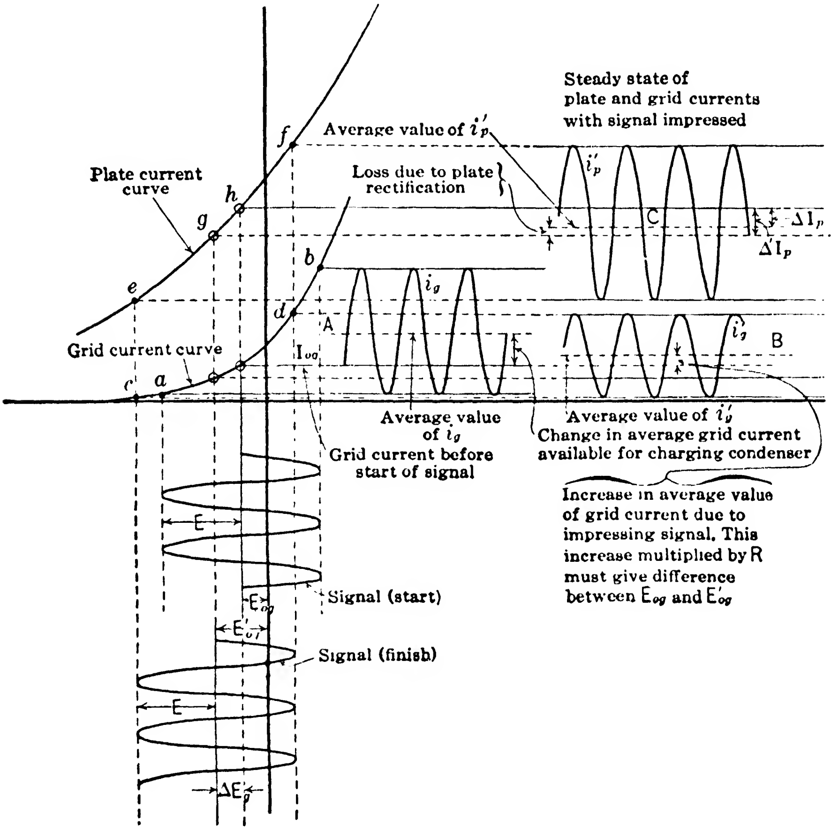


FIG. 104.—Showing the action of the triode when detecting by the action of the grid condenser; the net rectification is a result of two opposing actions.

flow through the condenser, but the charge it has accumulated from the effect of the signal, has depressed the average grid potential by ΔE_g, thus making the new average grid potential E'_{og}.

The signal still being impressed, the grid current now fluctuates from c to d, having an average value shown by the dashed line B. The grid current (average) has therefore been *increased* by the signal by the difference between I_{og} and the value indicated by line B. This increase in average grid current, flowing through the resistance R, must give voltage drop equal to the ΔE_g shown on the diagram (Fig. 104).

In the steady state, with signal impressed, the plate current evidently fluctuates from e to f on the plate current curve, giving a plate current as shown by i'_p. The average value of this current is shown by the dashed line C. With no signal the plate current has the value shown at h so that the plate current has decreased by the amount ΔI_p as a result of the signal.

It is to be noted that for the average grid voltage E'_{og} the plate current should be as large as shown at g. This would result in a somewhat greater decrease in current than actually occurs, as shown at $\Delta' I_p$. The difference between $\Delta' I_p$ and ΔI_p is caused by the upward curvature of the plate current curve.

We see then that grid rectification by itself gives a change in plate current equal to $\Delta' I_p$ but that rectification due to plate current decreases this to ΔI_p. This effect will be shown in the analysis to follow.

As mentioned above the various quantities are not shown in their proper proportions in Fig. 104; it has been necessary to exaggerate some of them to keep the diagram clear.

When the signal is impressed, we have

$$e_g = E'_{og} + E \sin pt. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

As grid current is a smooth curve, we may put

$$i_g = I_{og} + \frac{dI_g}{dE_g}(e_g - E_{og}) + \frac{1}{2}\frac{d^2 I_g}{dE_g^2}(e_g - E_{og})^2 + \quad . \quad . \quad . \quad . \quad (19)$$

Now $i_R = e_g/R$ and in average values $\bar{i}_R = \bar{i}_g$. For the average value, $\bar{i}_g$, we use the average value as given by Eq. (19) and for the average value $\bar{i}_R$ we use average e_g/R and write the average value of e_g from Eq. (18). Thus

$$\frac{1}{R}(\overline{E'_{og} + E \sin pt}) = I_{og} + \frac{dI_g}{dE_g}(\overline{E'_{og} + E \sin pt - E_{og}})$$

$$+ \frac{1}{2}\frac{d^2 I_g}{dE_g^2}[(E'_{og} - E_{og})^2 + 2(E'_{og} - E_{og})E\,\overline{\sin pt} + \overline{E^2 \sin^2 pt}] + \ldots \quad (20)$$

But $\overline{\sin pt} = 0$ and $\overline{\sin^2 pt} = \frac{1}{2}$ and $(E'_{og} - E_{og}) = \Delta E_g$. So

$$\frac{1}{R}E'_{og} = I_{og} + \frac{dI_g}{dE_g}\Delta E_g + \frac{1}{2}\frac{d^2 I_g}{dE_g^2}\left(\Delta E_g^2 + \frac{E^2}{2}\right).$$

In ordinary triodes $\frac{E^2}{2} >> \Delta E_g^2$; also $\frac{1}{R}E_{og} = I_{og}$

so we may write

$$\frac{1}{R}\Delta E_g = \Delta E_g \frac{dI_g}{dE_g} + \frac{E^2}{4}\frac{d^2 I_g}{dE_g^2}$$

or

$$\Delta E_g = \frac{E^2}{4} \frac{\dfrac{d^2 I_g}{dE_g^2}}{\dfrac{1}{R} - \dfrac{dI_g}{dE_g}}. \quad \ldots \ldots \ldots \quad (21)$$

In the steady state $e_g = (E_{og} + \Delta E_g) + E \sin pt$.
And

$$i_p = I_{op} + \frac{dI_p}{dE_g}(e_g - E_{og}) + \frac{1}{2}\frac{d^2 I_p}{dE_g^2}(e_g - E_{og})^2 + \ldots$$

So

$$i_p - I_{op} = \frac{dI_p}{dE_g}(\Delta E_g + E \sin pt) + \frac{1}{2}\frac{d^2 I_p}{dE_g^2}(\Delta E_g^2 + 2\Delta E_g \times E \sin pt + E^2 \sin^2 pt)$$

So the *average* change in plate current is

$$\Delta I_p = \frac{dI_p}{dE_g}\Delta E_g + \frac{1}{2}\frac{d^2 I_p}{dE_g^2}\left(\Delta E_g^2 + \frac{E^2}{2}\right)$$

and as

$$\frac{E^2}{2} >> \Delta E_g^2$$

we have approximately

$$\Delta I_p = \frac{dI_p}{dE_g}\Delta E_g + \frac{1}{2}\frac{d^2 I_p}{dE_g^2}\frac{E^2}{2}. \quad \ldots \ldots \ldots \quad (22)$$

Combining (22) and (21), we finally get

$$\Delta I_p = \frac{E^2}{4}\left(\frac{\dfrac{d^2 I_g}{dE_g^2}}{\dfrac{1}{R} - \dfrac{dI_g}{dE_g}} \frac{dI_p}{dE_g} + \frac{d^2 I_p}{dE_g^2} \right). \quad \ldots \ldots \quad (23)$$

And with ordinary tubes and circuit constants this may be written without much error

$$\Delta I_p = -\frac{E^2}{4} \frac{\dfrac{d^2 I_g}{dE_g^2}}{\dfrac{dI_g}{dE_g}} \frac{dI_p}{dE_g}. \quad \ldots \ldots \ldots \quad (24)$$

It must be remembered that E is not the voltage impressed upon the input circuit (i.e., the signal voltage), but something less due to the drop in the grid condenser. The solution in Eq. (24) supposes the signal has persisted long enough for the steady state to be reached; if a damped sine wave is impressed, the detection efficiency will depend upon the decrement, size of grid condenser, etc., as analyzed on p. 554.

In Fig. 105 are shown the grid currents of two modern detecting tubes and in Fig. 106 curves for two older types. If no grid condenser were used with these tubes we find (using Eq. (17) and measurements of slope, etc., from the curves) that to produce an increase of 1 microampere in the average value of the plate current requires an alternating voltage of 0.15 volt on the grid for tube A and 0.19 volt for tube B. These values were calculated on the assumption that the normal grid potential is zero, which means that the input circuit is connected to the *negative* end of the filament.

If the grid condenser were used with these tubes, having leak resistances of 1 megohm, these leaks being connected to the positive end of the filaments, the normal grid potentials would be as indicated by the large circles on the I_g–E_g curves of Fig. 105. Using Eq. (24) we find that, to produce a decrease in the plate current of 1 microampere, for tube A requires an alternating voltage on the grid of 0.059 volt, and for tube B it requires 0.052 volt. Both these tubes would therefore be much better detectors with grid condensers than without them, and such was found true experimentally.

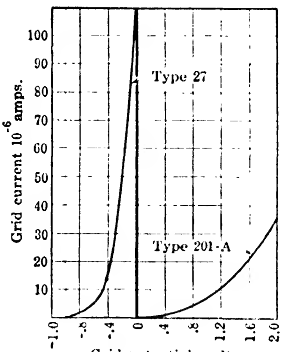


FIG. 105.—Grid current-grid potential curves of two triodes much used as detectors.

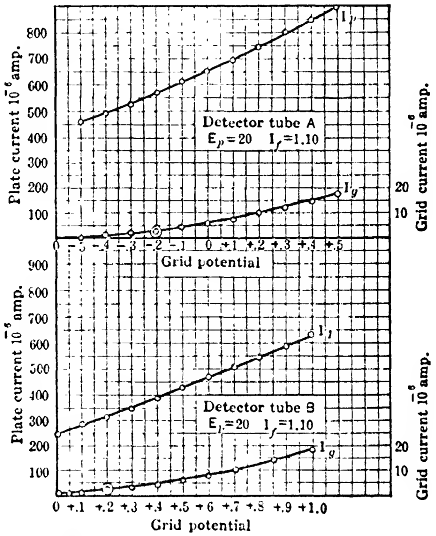


FIG. 106.—Characteristic curves of two detector tubes. Using Eq. (24) it is found that to change the average plate current by 1 microampere requires a signal voltage of 0.059 volt for tube A and 0.052 volt for tube B. Without grid condensers the tubes require about three times as much grid voltage for the same change in plate current.

In Fig. 107 is shown the great difference in the form and magnitude of I_g and I_p when the junction of grid-filament circuit is changed from the negative end of the filament to the positive end; the difference is very much exaggerated here because of the low value of the voltage of the battery in the plate circuit.

In Fig. 108 are shown the characteristics of an old Deforest detecting bulb, the filament current being at the rated value for this type of bulb. It will be readily appreciated that such tube would act peculiarly as different adjustments were made. Thus with a plate voltage between 30 and 50 the tube would not detect, with or without grid condenser. With 20 volts in the plate the tube gave very good detection with or without grid condenser; with 10 volts on the plate the tube gave fair detection with grid condenser and none at all without grid condenser.

Effect of Frequency and Decrement of Signal.—The previous analyses have not taken into account the amount of electricity available for charging condenser C; only relative reactances, etc., have been considered. But it is evident that if the condenser is to be charged the grid current

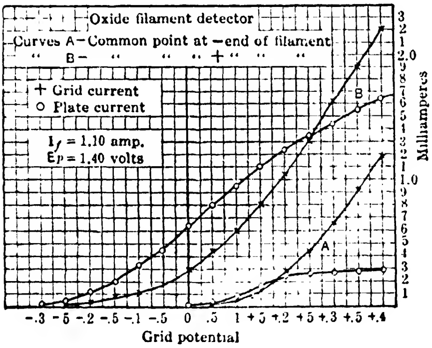


FIG. 107.—With low plate voltages it makes a great deal of difference whether the grid is connected to the positive or negative end of the filament; plate current indicated by circles and grid current by crosses.

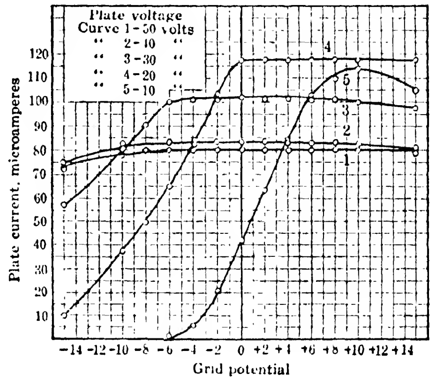


FIG. 108.—Peculiar characteristics of an old Deforest audion detector; such a tube detects in very erratic fashion, probably owing to the considerable amount of gas left in the tube.

must supply the electrons required, and it may be that the current is not sufficiently large to do this, in the short time the signal is impressed.

Suppose the signal voltage has the form shown in Fig. 109; it reaches its maximum in three cycles and then rapidly decreases. If possible the grid condenser C should be charged up to a potential fixed by the *maximum* value of this signal. To make the problem simple we will suppose the amplitude of the voltage to have its maximum value during the first three cycles and examine the possibility of the condenser C having reached the value of potential fixed by Eq. (21).

If the condenser is to have its potential changed by ΔE_g the required quantity of electricity is $(\Delta E_g \times C)$. The current available for charging the condenser is very nearly $(E^2/4)\,(d^2I_g/dE_g^2)$ and for three cycles this makes available a quantity of electricity $q=(3TE^2/4)\ d^2I_g/dE_g^2$. Now if $1/R$ is negligible compared to dI_g/dE_g we get from Eq. (21)

$$\Delta E_g=\frac{E^2}{4}\frac{d^2I_g}{dE_g^2}\frac{dE_g}{dI_g}. \quad \ldots \quad (25)$$

So, as $q=C\Delta E_g$, we may put

$$C\frac{E^2}{4}\frac{d^2I_g}{dE_g^2}\frac{dE_g}{dI_g}=\frac{3TE^2}{4}\frac{d^2I_g}{dE_g^2}, \quad \ldots \quad (26)$$

from which we conclude that the largest condenser which can be used, and still be fully charged, is fixed by the relation $C=3T\ (dI_g/dE_g.)$

If $dI_g/dE_g=5\times10^{-6}$ (which is about the value obtained from Fig. 100, when E_{og} is -0.6 volt) and T is 2×10^{-6} (which is the period of a 600-meter wave), the maximum value of capacity should be 30×10^{-12} farad. But such a low value for C would result in a very small fraction of the signal voltage being impressed on the grid, so that a much larger condenser must be used and the value of dI_g/dE_g must be made larger.

By using a lower value for the leak resistance, R, the normal grid potential E_{og} may be made higher, which will result in a higher value of dI_g/dE_g; in Fig. 100 when $E_{og}=0$, $dI_g/dE_g=15\times10^{-6}$. If the decrement of the signal is low, we may allow more than three cycles for the condenser to charge without greatly decreasing ΔE_g, because the amplitude of signal voltage will still be nearly its maximum value.

From the foregoing discussion it is evident that the predetermination of the best value for C is somewhat involved; moreover it will be found experimentally that C may be varied over a wide range without appreciably changing the efficiency of the tube as a detector, probably due to compensation among the different effects just mentioned. A large C will not charge completely, e.g., but it will permit a greater fraction of the input voltage to act on the grid than would a smaller one which would charge more completely.[1]

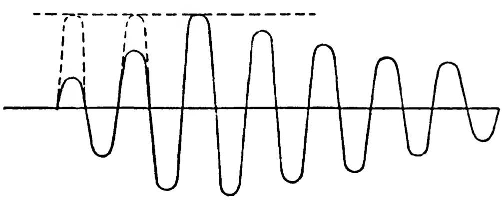


FIG. 109.—In analyzing the effect of the decrement of the signal on the detecting action we assume the first three cycles of a wave train have the same amplitude, the maximum value of the signal voltage.

The action of the tube as detector with grid condenser is well shown in Fig. 110, a film taken by the author in 1914. In the upper part of the

[1] Analysis also shows that as the decrement of the signal increases the most suitable value of grid condenser capacity decreases.

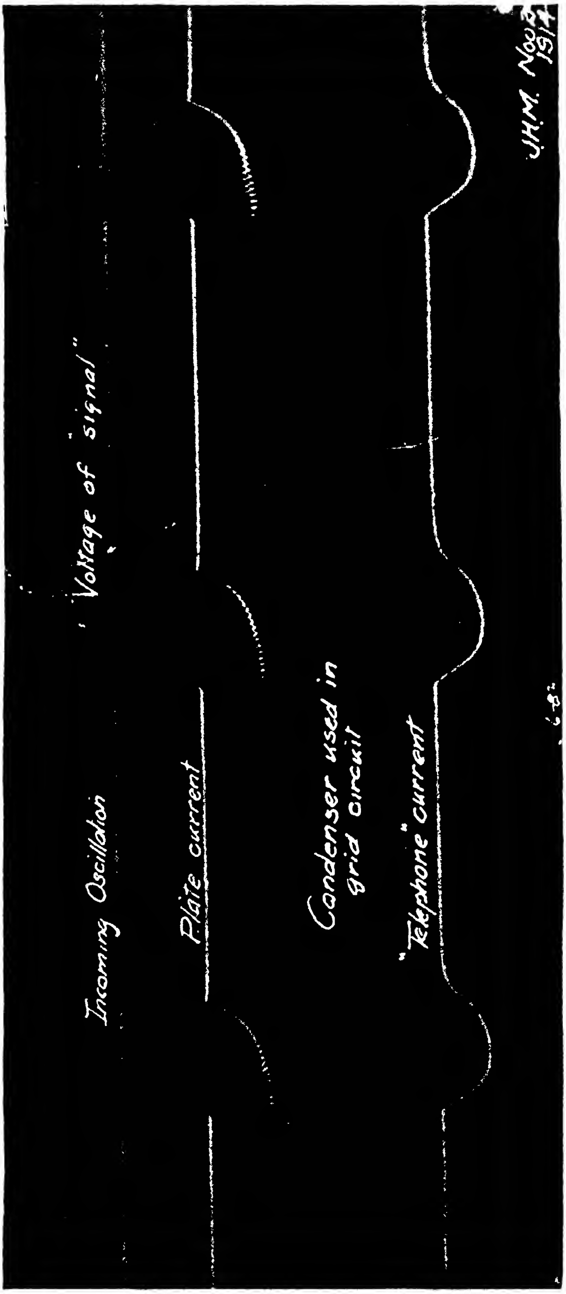


FIG. 110.—Oscillogram showing the action of the tube as detector when a condenser (with suitable grid leak) is used in series with the grid.

figure is shown the input voltage; the second curve shows the plate current, having pulsations of the same frequency as the signal voltage, but having also a large average decrease due to the grid condenser becoming charged; the " telephones " (in this case a coil of high inductance) were shunted by a large capacity so that the " high-frequency " fluctuations in plate current did not pass through them, but the low-frequency change in plate current did pass through them, giving a current of the form shown.

By increasing the value of the leak resistance about three times the time required for the grid condenser to discharge was increased and so the plate current was held at its lowered value for a longer interval of time; the currents then had the forms shown in Fig. 111.

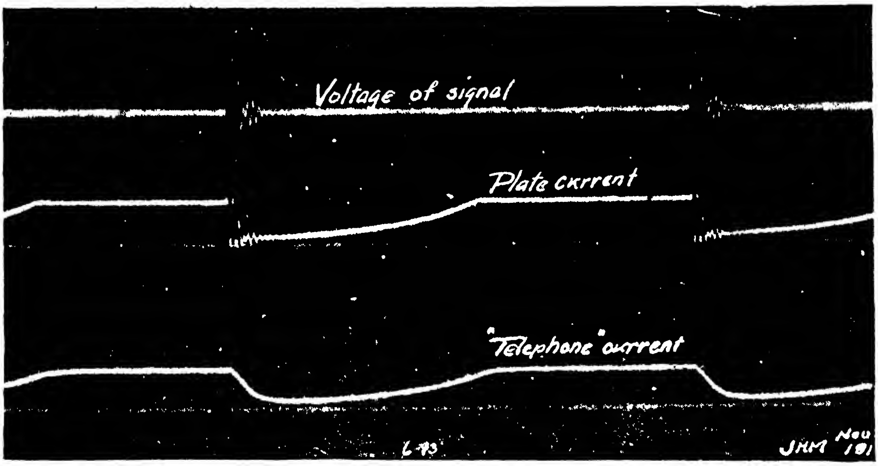


FIG. 111.—By increasing the value of the grid leak the form of the plate-current curve may be changed; in this film all conditions were the same as those of Fig. 110 except the value of the grid-leak resistance had been approximately trebled.

It is to be noted that the mean potential of the grid can be no longer depressed when the fluctuations in grid potential due to the signal do not carry it to a potential more positive than the value it has when no signal is coming in. Unless its potential exceeds this potential it will not attract any excess electrons (i.e., more than it attracts when no signal is coming in) and hence cannot depress the average potential of the grid. In this respect many writers have shown the action of the three-electrode tube incorrectly; in Fig. 112 curve (*a*) shows the voltage impressed on the grid due to the signal and in (*b*) is shown correctly the resulting grid potential, the average potential being shown by the dotted line. After the third cycle the signal voltage is not of sufficient magnitude to carry the grid potential higher than its no-signal value; after this time, therefore, the average grid potential must rise due to the accumulated charge escaping through the leak resistance.

In curve (c) is shown the grid potential as frequently given in texts; the average potential is shown as decreasing further even when, during the previous cycle, the grid potential did not rise as high as its no-signal value; this illustration is incorrect.

The term "accumulative amplification" has been used in describing the action of a tube with grid condenser, but it is to be noticed that there is no true amplification; the grid potential is in no case depressed by an amount *in excess of the amplitude of the signal e.m.f.*, as it is when regenerative amplification is used.

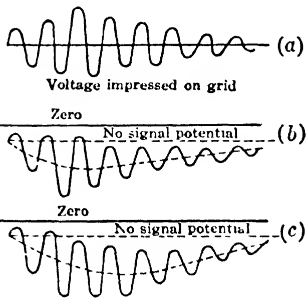


FIG. 112.—This diagram shows in (b) a correct representation of the grid potential when signal (a) is impressed and in (c) an incorrect representation. The average potential of the grid will not be further depressed unless during the previous cycle the grid forced to a potential higher than its "no signal" potential. In case the grid leak is connected to the positive end of the filament the potential of the grid with no signal coming in is higher than the free grid potential, so that curve (b) might possibly start from a positive value instead of the negative value as shown.

"Equivalence Theorem" of Detector Action. — From the analyses given in the foregoing pages it is evident that when a radio-frequency signal, of varying amplitude, is impressed on the input circuit of a triode arranged for detection by the action of a grid condenser and associated leak resistance there is developed in the grid circuit, a voltage of the same frequency as that of the amplitude variations. In the foregoing analyses we did not stop at the grid voltage but carried them through to see what change in plate current was brought about by these modulation frequency voltages produced in the grid circuit.

In the last few years there has been much discussion of the detector action of a triode in terms of an "equivalence" theorem (due to Carson), and although nothing new is shown by this type of analysis (over that already given) a résumé of the method is given here for the sake of completeness of the discussion of detector action.

In Fig. 113 *a*, is shown the usual detector circuit; the radio signal, of varying amplitude, is set up in resonant circuit L C_1, and so impresses on the triode input circuit, through the C–R combination, this varying amplitude voltage. This voltage, acting between grid and filament, as a result of the peculiar slope of the I_g–E_g curve, develops across the C–R combination a certain amount of voltage of frequency equal to that of the amplitude variation, i.e., the modulation frequency. Now the impedance of L for this modulation frequency (practically always an audio frequency) is very low, so that, so far as the action of the audio frequency is concerned, the arrangement of Fig. 113 *b* is the same as that of *a*. The equivalence theorem states:

" For small signal voltage the rectified grid current produced by the grid circuit rectification is exactly the same as would be produced by a suitable generator G acting between the grid and filament in series with the grid-filament resistance of the tube. The amplitude of voltage of this hypothetical generator depends upon the grid bias, nature of signal, etc. The current which this generator would produce, if it were present, is the rectified current that the applied signal actually does produce in the actual circuit."

To care for the distortion which rectification generally produces, the theorem in its more general form provides for a series of generators, of harmonic frequencies, and adjustable phases.

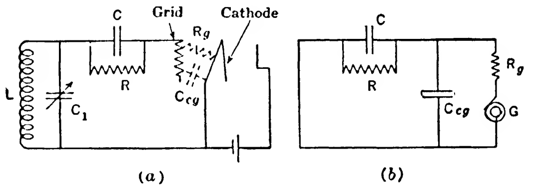


FIG. 113.—Circuits illustrating the "equivalence theorem" of detector action.

To find what voltage these equivalent generators must have an analysis similar to that already carried out must be made. The voltage is generally given in terms of a tube constant, called the *detector voltage constant*, and represented by v. Letting R_g represent the a.c. resistance of the input circuit, that is dE_g/dI_g, v is given by the relation

$$v=\frac{2R_g}{\dfrac{dR_g}{dE_g}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

This, it will be noticed, is merely a symbol for part of the expressions already derived for detector action; expanding it we see that

$$v=\frac{2\dfrac{dE_g}{dI_g}}{\dfrac{d^2E_g}{dI_g^2}},$$

and this is found in both Eqs. (23) and (24).

From the analysis already given, and one to follow (see pp. 585–588), we can write with the magnitude of voltages of these equivalent generators for three typical signal voltages:

(*A*) An unmodulated radio signal $e=E\sin\omega t$

Component of rectified voltage is

$$\frac{E^2}{2v}. \qquad (28)$$

This component has zero frequency.

(*B*) A modulated radio-frequency signal (signal of varying amplitude), $e=E(1+m\sin pt)\sin\omega t$.

Component of rectified voltage (zero frequency) is

$$\frac{\left(E^2+\frac{\overline{mE}^2}{2}\right)}{2v}. \qquad (29)$$

Component of modulation frequency

$$=\left(\frac{mE^2}{v}\right)\sin pt. \qquad (30)$$

Component of double modulation frequency (distortion)

$$=\left(\frac{\overline{mE}^2}{4v}\right)\sin 2pt. \qquad (31)$$

(*C*) Two superimposed radio voltages of different frequencies, $e=E_1\sin 2\pi f_1t+E_2\sin 2\pi f_2t$.

Component of rectified voltage (zero frequency)

$$=\frac{E_1^2+E_2^2}{2v}. \qquad (32)$$

Component of rectified voltage of frequency (f_1-f_2)

$$=\left(\frac{E_1E_2}{v}\right)\sin 2\pi(f_1-f_2)t. \qquad (33)$$

It will be noticed that the voltage of the hypothetical generators for all cases varies inversely as v, so we conclude that a good detector should have a low value of v. Formulas (32) and (33) are introduced without proof; but this is developed on pp. 585–588. The significance of Formula (31) is discussed in p. 707.

Terman has done much work in the analysis of detector action from this " equivalence " point of view and has measured the value of the detector voltage constant v for all the ordinary detector tubes.[1]

It is evident that v depends directly upon the slope of the I_g–E_g curve, and as the slope and curvature of this vary with grid bias, we should expect

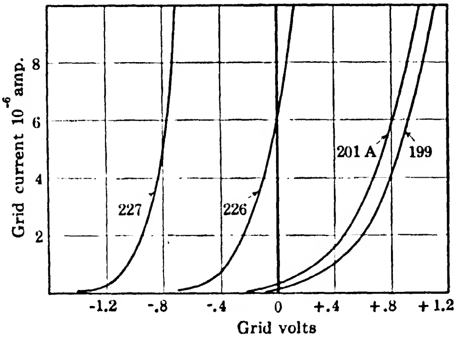


Fig. 114.—Grid current curves of modern tubes.

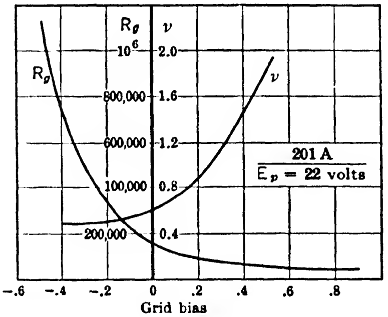


Fig. 115.—Detector voltage constant and grid circuit resistance of a detector tube.

v to vary with the grid bias. In Fig. 114 are shown grid current curves for four of the ordinary detector tubes. In Fig. 115 are shown the grid resistance curve and values of v for a 201 A tube; and in Fig. 116, the values of v for this tube plotted against R_g, instead of grid bias. In Fig. 117 the rela-

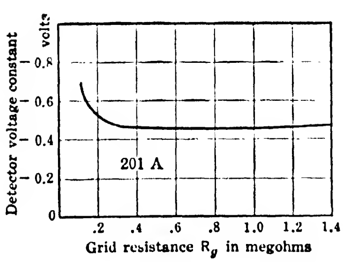


Fig. 116.—Dependence of detector voltage constant upon grid circuit resistance.

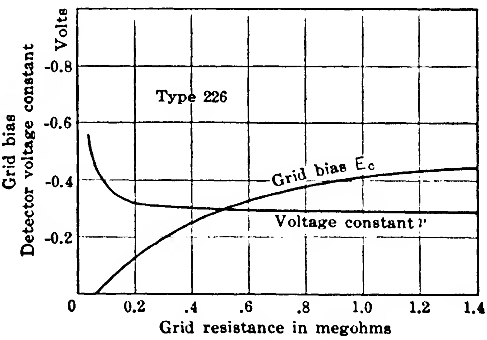


Fig. 117.—Detector voltage constant and its dependence upon grid circuit resistance.

tions between grid resistance, grid bias, and the detector voltage constant, for a type 26 tube, are indicated.

[1] See Terman, I.R.E., Oct., 1928, p. 1384; and Terman and Googin, I.R.E., Jan., 1929, p. 149.

In the ordinary range of values for grid condenser leaks (which largely determines the normal grid bias) there is not a great deal of variation in the value of v even for many different types of tube; furthermore, for all tubes v shows a reasonably constant value for considerable changes in grid bias. Terman and Googin measured this value of v for many different tubes, with a special bridge arrangement; a few of their values are tabulated here. The values of v correspond to the flat part of the v–R_g curve. (See Figs. 116–117.) The value of R_g given in the table is the value of this resistance which corresponds to that grid bias where the v curve starts to become flat. The significance of the column marked F is explained below under the topic of frequency distortion.

Type	v	R_g	μ	R_p With $E_p = 45$ Volts	F
201A.........	0 47	150,000	8	14,000	3,500
226..........	0 29	150,000	8	9,000	3,500
227	0 23	50,000	8	10,000	11,500
240..........	0 47	150,000	30	150,000	3,500
171A.........	0 28	200,000	3	2,500	2,600

Harris[1] has given a method for determining the necessary factors for evaluating the detector voltage constant of a tube, and in Fig. 118 are shown typical curves obtained by his method; the values are these for a type 27 tube.

Frequency Distortion.—Referring to Fig. 113 *b*, and Eqs. (28)–(33), it will be seen that a tube with a low v and low R_g will produce the largest rectified voltage across the C–R combination. Now evidently R must be large compared to R_g, otherwise most of the rectified voltage (supposedly generated by alternator G) will be used up internally in the R_g of the tube, and a relatively small part will appear across R; however, it is only the change in voltage across R that effects the plate current, and so makes available the rectified signal.

Now C must be large enough to let the radio-frequency signal in the L–C_1 circuit act on the grid without appreciable diminution; referring to Fig. 113 *a*, it is apparent that if the impedance of the C–R combination is high the radio-frequency voltage impressed on the grid will be smaller than that in L–C_1 circuit. But the voltage of the hypothetical generator G (Fig. 113 *b*) depends upon the voltage impressed upon the grid, hence to get good rectification the radio frequency drop through the C–R impedance must be small. The means that C must be large compared to C_{cg}, the grid-

[1] I.R.E., Aug., 1929, p. 1322.

cathode capacity of the tube. A reasonable proportion gives to C a capacity about ten times the capacity of C_{cg}.

Now if the generator G (Fig. 113 b) acts on the circuit C–R with different frequencies it will be evident that for a given voltage of G there will be increasingly smaller voltages across R with increasing frequency, because of the diminishing reactance of condenser C and the internal resistance R_g. At low frequencies (say 100 cycles per second), R being much larger than R_g and C only 100–300 $\mu\mu f$, practically all the voltage of G will appear across R and so be available for controlling plate current. But at 10,000 cycles, for example, with $R = 5 \times 10^6$ ohms, $R_g = 2 \times 10^5$ ohms, $C = 2 \times 10^{-10}$ farad, and $C_{cg} = 10 \times 10^{-12}$ farad, there will be across R only 37 per cent of the generator voltage. At 5000 cycles there will be about 62 per cent, and at 1000 cycles about 97 per cent. It is thus evident that grid-circuit detection acts less efficiently, in producing audio-frequency changes in plate current, at high audio frequencies than at lower ones; this discrimination in favor of the lower ones is called frequency distortion.

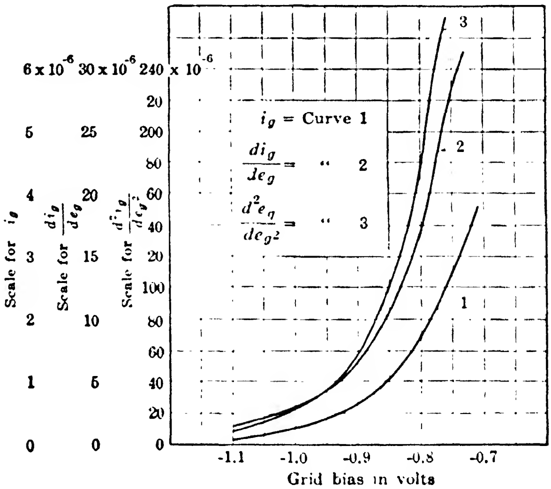


FIG. 118.—Grid current and its grid-voltage derivatives as functions of grid bias.

The column headed F in the table on p. 562 give the highest audio frequency which the tube detects with 70 per cent the efficiency it has for the low frequencies. The values have been calculated for the different tubes with the same capacity of grid condenser, namely 300 $\mu\mu f$. It is quite likely that different tubes would give better performance if each had its most suitable condenser.

Nelson[1] gives curves showing the effect of this detector frequency distortion on an actual radio receiver, and some of his results are reproduced in Figs. 119 and 119-A. In curve A of Fig. 119-A is shown the audio-frequency response from the grid of the detector to the grid of the second audio stage; no detector frequency distortion was present, because the tube was not functioning as a detector, the audio frequency itself being impressed on the detector grid. In curves B and C the response includes that of the

[1] I.R.E., March, 1929.

detector; instead of using an audio-frequency signal as was done for curve *A*, a modulated radio-frequency signal was used, thus adding to the audio amplifier defect that of the detector. Curve *B* was for a 0.5-megohm leak and curve *C* for a 2-megohm leak. Curves *B* and *C* have been arbitrarily pro-rated to give the same response as *A* for the low frequencies. Of course, with a 0.5-megohm leak the detector actually rectified less efficiently than with the 2-megohm leak, as could be surmised from the curves of Fig. 119.

Evidently the selection of a proper-sized grid condenser and grid leak is a compromise; too low a leak resistance means low amplification and poor selectivity of the tuned radio-frequency circuit connected to the detector input, and too high a resistance means too much frequency distortion.

Distortion Due to Non-linear Response of Detector.—The rectified current in the plate circuit does not vary directly with the ampli-

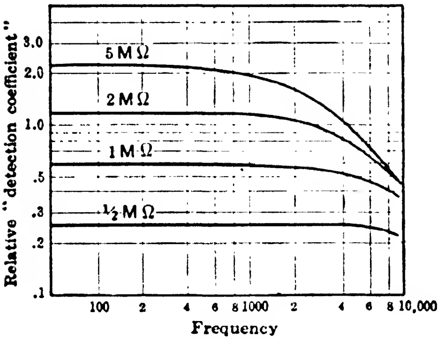


Fig. 119.—Frequency distortion produced by detector, as a function of frequency and the resistance used as grid leak.

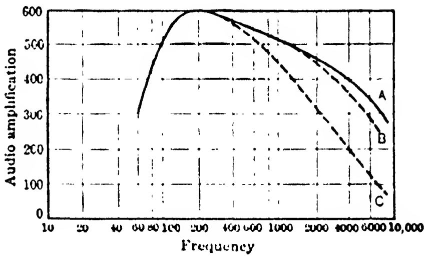


Fig. 119A.—Frequency distortion in an actual receiver.

tude of the modulated r.f. voltage impressed in its grid, especially for weak signals; here it varies nearly as the square of the signal strength. (See Figs. 98 and 122.) According to some experimental results recently published,[1] a good deal of distortion is caused by detector action, even at very low input voltages. A resonance bridge was used to balance out from the output the detected voltage of the same frequency as the modulation frequency, and the apparatus measured the residual voltage, principally the second harmonic of the modulation frequency. That is, a radio signal, modulated at 800 cycles, gave in the detector output a rectified current of 1600 cycles as well as some higher multiples of 800 cycles. As this 1600-cycle frequency was not present in the original modulation frequency it must have come in by reason of the detector action. (See Eq. (31).)

[1] Brown, Pickels, and Knipp in Radio Engineering, Jan., 1932, p. 21.

In Fig. 120 are shown the output voltage, of fundamental frequency (800 cycles), of a 1000-kc., 50 per cent modulated, 0.15-volt signal acting

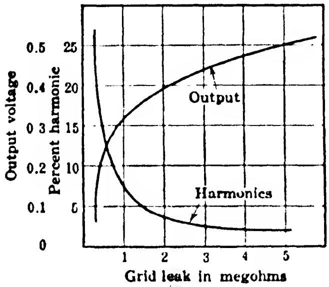


FIG. 120.—Harmonic distortion, due to non-linear detector action, as grid leak resistance is varied.

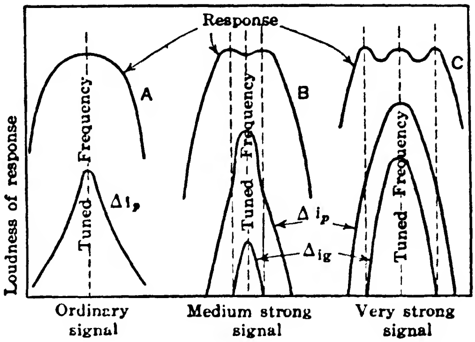


FIG. 121.—Peculiar response of a radio set having an over loaded detector.

in a 201A triode arranged for grid circuit detection. The grid condenser was 200 $\mu\mu f$ and the leak resistance was varied, with the results shown. As resistance increased, the output voltage (modulation frequency voltage in the plate circuit) increased, and the proportion of harmonics decreased. This discrimination against the harmonic frequencies, as leak resistance is increased, is to be expected from the analysis of frequency distortion, discussed above.

Detector Overloading.—Another source of distortion produced by detector action is that due to overloading. This is especially likely to occur in sets using grid rectification without automatic bias control. Some of the factors involved in overloading are discussed by Harris,[1] who gives a curve sheet

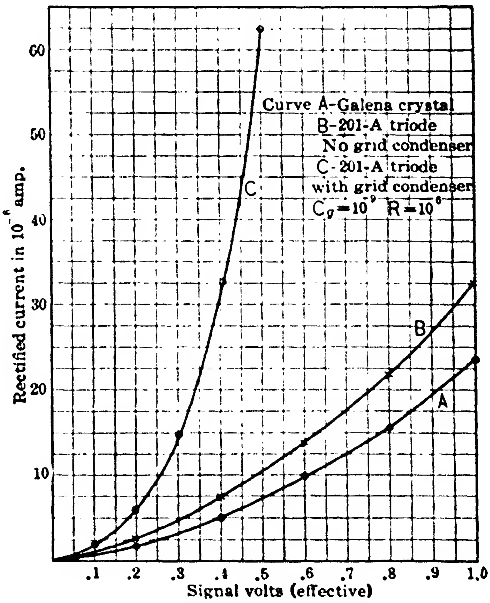


FIG. 122.—Comparison of rectifying efficiency of crystal and triode, with and without grid condenser.

[1] I.R.E., Oct., 1929, p. 1834.

(reproduced in Fig. 121) showing the loud-speaker response, as tuning is varied through the resonant frequency, for ordinary signals and two others too strong for the detector. Not only is the quality poor, when the detector is overloaded, but the response has the peculiar form showing in curves B and C of Fig. 121. For ordinary signals the grid draws inappreciable current, but for the stronger signals it becomes large and it is the action of this current that causes the peculiar response.

Comparison of Crystal and Triode as Detectors.—In Fig. 122 are shown the comparative rectifying properties of a good crystal and a triode, used with and without a grid condenser. It may easily be seen that for all three curves the rectified current varies nearly as the square of the impressed voltage, as the analyses predict should be the case for low-voltage signals.

The triode is somewhat more efficient than the crystal, used in either of the two possible ways, but it is evidently much better when used with a grid condenser.

Effect of Plate Circuit Impedance on Rectification.—In general, we may say that any rectifying device, having an impedance in series with it, will show comparatively poor rectification unless this impedance is by-passed by a low reactance condenser. The idea is illustrated in Fig. 123, where a high-frequency generator impresses a sine wave of voltage on the circuit made up of impedance Z and rectifier R in series. In curve a of Fig. 123 there is shown the shape of rectified current when the alternating voltage is impressed directly across the rectifier, and in curve b is shown the form of rectified current and voltage across the rectifier when the circuit is as shown. During the positive alternation, when current flows through the rectifier, there is a drop through the impedance, resulting in a peculiar voltage form across the rectifier; the tops of the positive alternations of voltage have been cut off. (In curve b the dashed curve voltage is the normal sine wave.) But these flat-topped voltage alternations draw through the rectifier a comparatively small current, as indicated, with the result that the value of the rectified current is very much less than it should be. It will be noticed in Fig. 97, p. 544, that the radio-frequency fluctuations in plate current do not have to traverse either the resistance R or the transformer primary winding. In Fig. 124 are shown two curves, giving the rectifying action of a type 27 triode. In curve A the external impedance in the plate circuit was by-passed with a suitable condenser, and in curve B it was not so shunted. The great improvement in rectifying action obtained by properly shunting the impedance is evident.

Measurement of Detecting Efficiency of a Three-electrode Tube.—It is possible to determine experimentally the detection coefficient of a tube by such a scheme as that originated by Van der Bijl;[1] in his treatment it is shown that the strength of signal given by the telephone varies as the fourth power of the voltage impressed on the grid. This follows at once also from Eq. (17) p. 543 in which it is shown that the increment of plate current varies with the square of the voltage impressed on the grid; as the amount of noise given off from the telephone varies with the square of the current through it, it is evident that the noise varies with the fourth power of the grid voltage.

Using a receiver which required 3×10^{-12} watt input to produce the "least audible signal," Van der Bijl found that the ordinary detector tube (without a condenser in series with the grid, depending only on shape of plate current curve for rectification) required a signal voltage of 0.025. Unless some very radical change is made in either telephone receiver or detecting tube, it may be assumed that, for a readable signal, it is necessary to impress on the grid of a detector a voltage (high frequency) of between 0.01 and 0.05 volt.

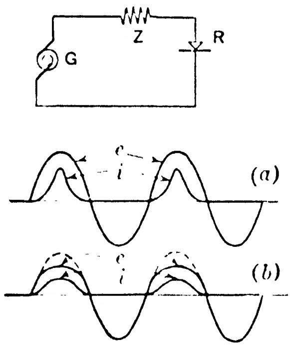


FIG. 123.—The action of a rectifier depends greatly upon the impedance in series with it.

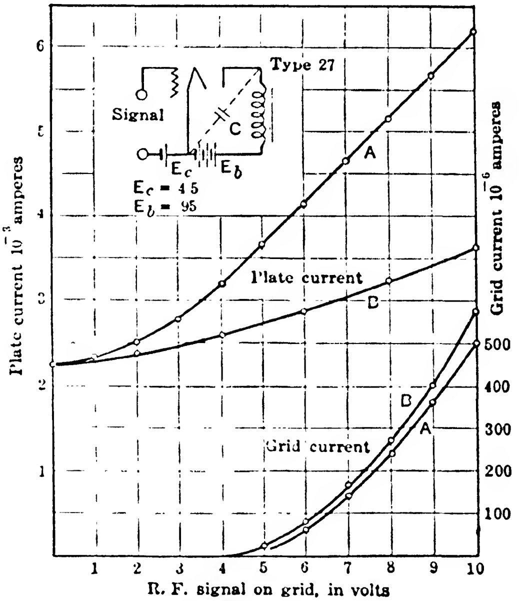


FIG. 124.—A condenser across the primary of the plate circuit transformer greatly increases the rectifying efficiency of the triode.

The Three-electrode Tube as a Source of Alternating Current. General Field of Application.—A three-electrode tube, if supplied with the proper c.c. power and if connected to a circuit having a natural period of

[1] H. J. Van der Bijl, "On the Detecting Efficiency of the Thermionic Detector," Proc. I.R.E., Dec., 1919. See also "Detecting Characteristics of Electron Tubes," by Freeman, Proc. I.R.E., Vol. 13, No. 5, Oct., 1925.

oscillation, will, under certain conditions, generate a.c. power of the frequency fixed by the L and C of the circuit to which it is connected. The action is nearly analogous to that of a violin bow; although the force and velocity of the bow are essentially constant the peculiar friction between the bow and string enables the string to absorb more power from the bow when string and bow are moving in the same direction than is given back to the bow by the string when the motions of bow and string are in opposite direction. If the frictional force between string and bow is plotted as a function of the relative velocity of the two, the graph will have the form given in Fig. 125; curve (a) is for the bow without resin and curve (b) shows the change in this friction after resin has been put on the bow.

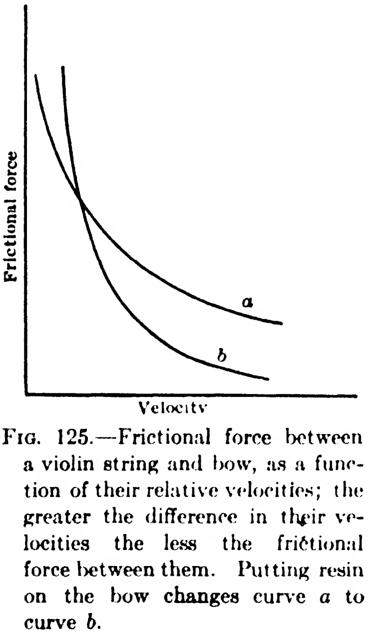


FIG. 125.—Frictional force between a violin string and bow, as a function of their relative velocities; the greater the difference in their velocities the less the frictional force between them. Putting resin on the bow changes curve a to curve b.

The muscles of the arm actuating the bow constitute a source of continuous power; it is obviously impossible for an arm muscle to supply (directly) power to a string vibrating 1000 times a second. The arm supplies energy to the bow at an essentially constant rate, the reactions between the bow and string serve to utilize this power to maintain the string in a state of rapid vibration.

The system which drives the balance wheel of a watch is also somewhat analogous; the mainspring furnishes power by a continuous force, but the escapement system serves to feed energy into the moving balance wheel in such a way as to maintain it in a state of oscillation, the period being fixed by the mass of the wheel and stiffness of the hairspring.

It is to be noted that neither the balance wheel nor violin string will vibrate if the damping of the oscillating member is too high; if too much friction occurs in the bearings of the balance wheel the watch will stop. The same effect exists in the oscillating-tube circuits to be described later.

The efficiency of the three-electrode tube as a generator of a.c. power is normally rather low; in small tubes such as used for aeroplane telephony the circuits are generally arranged to give an efficiency [1] of

[1] In speaking of the efficiency of an electron tube the "input" does not ordinarily include the power necessary for heating the filament; it is the power supplied in the plate circuit only.

about 25 per cent, and in the larger tubes to get an output of 150 watts requires an input of about 300 watts. In a later section of this chapter the efficiency of a tube is discussed and analyzed in detail. In spite of its rather low efficiency it will probably be always used when a small amount of power is desired at a frequency of 10 kilocycles or more, because there is no other simple method of generating power at these frequencies. Above about 100 kilocycles the vacuum tube has no competitor as a generator, and for small amounts of power, the frequency of which is preferably variable, the vacuum tube is probably better than any other device, no matter what this frequency may be.

There are two general fields in which the oscillating vacuum tube is used in radio, as a source of high-frequency power for a continuous-wave transmitting station, and as a necessary part of any station receiving continuous-wave signals, by means of the heterodyne or " beat " method. When used as part of a continuous-wave receiving set the oscillating tube is required to generate but a very small fraction of a watt and the smallest type of detecting tube will suffice.

For general laboratory use the small oscillating vacuum tube is of great service, as a source of a few watts of a.c. power for bridge measurements of frequency adjustable to any degree desired, or as a source of complex a.c. forms, from which exact octaves are obtainable; in combination with a suitable sound generator, such as piezo electric crystal or untuned telephone receiver, it is invaluable in a laboratory for measurements on sound.

Elementary Analysis of the Operation of a Three-electrode Tube as Generator of Alternating-current Power.—The three-electrode tube may be used as a self-exciting device or the power required to excite its grid circuit may come from some other source. This latter scheme is often used when it is desired to get maximum possible power from several tubes, operating in parallel. Their input circuits (grid-filament) are all connected in parallel and excited from some other, self-exciting, vacuum-tube circuit, the power capacity of which may be small compared to that of the tubes excited.

The operation of the separately excited tube is extremely simple; if an alternating voltage is applied on the input circuit, the plate current, which is fixed by Eq. (7), must rise and fall as this grid potential alternately increases and decreases. The simplest circuit to be considered for using the a.c. power generated in the plate circuit is that shown in Fig. 126; a choke coil L, having a negligible resistance, is put in series with the machine M, furnishing the plate circuit voltage, the value of L being large enough so that its reactance is large compared to the a.c. resistance R_p of the plate-filament circuit. Shunting this choke coil is the output circuit or load circuit of the tube; it consists of a condenser C, having a reactance, the magnitude of which is small compared to R_p, in series with a resistance R of about the same value as R_p. A condenser C_1 shunts the machine M to make the reactance of this part of the circuit negligible compared to R. The condenser C is generally necessary to keep the continuous component of the plate-current from flowing through the load circuit ammeter; the ammeter used in the load circuit should measure only the alternating current supplied to the load.

With the conditions named (large L, C, and C_1), the external impedance of the plate circuit will consist essentially of R only. As the voltage of the grid goes alternately positive and negative the plate current will fluctuate about its normal value, I_{op}, and the plate voltage will also fluctuate about its normal value, E_{op}. The actual plate current may be considered as made up of the constant value I_{op} which flows through L, and does not appreciably vary as E_g is varied, and an alternating component I_p, which flows in the plate circuit by the path C, R, C_1. The plate voltage similarly will be considered as made up of a constant term E_{op} on which is superimposed the alternating voltage E_p; at any instant the actual plate voltage will be equal to E_b-i_pR, where i_p is the instantaneous value of I_p. As the magnitude of E_g is increased the maximum value of I_p increases until it is practically equal to I_{op}. Under this condition, the actual current through the plate will fluctuate between $2I_{op}$ and zero and the value of plate voltage will fluctuate between $2E_b$ and zero if R is chosen of proper value. (Actually this amount of current and voltage fluctuation is not reached; the values named are limiting values.) The characteristic curves are shown in Fig. 127.

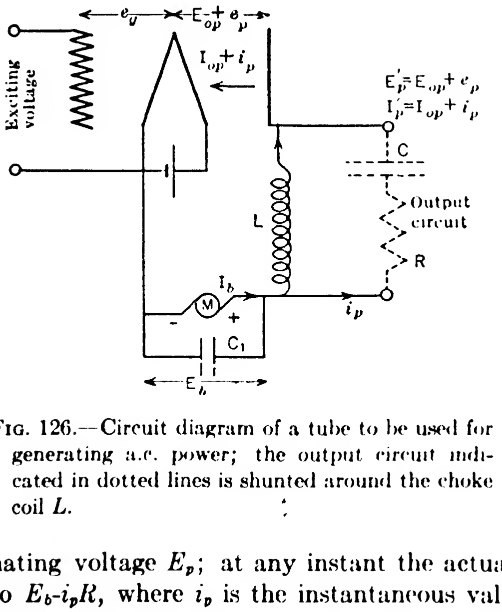


Fig. 126.—Circuit diagram of a tube to be used for generating a.c. power; the output circuit indicated in dotted lines is shunted around the choke coil L.

If the excitation is still further increased and the circuit L, C, R, is properly adjusted, the forms of E_p and I_p may be made to differ very materially from the sinusoidal forms here shown. It is, however, difficult to write the theory of the various circuits for any but sinusoidal functions, and we shall assume that I_p and E_p are such, unless specific mention is made to the contrary. We shall call the oscillations normal when I_p is sinusoidal or approximately so, that is for $(I_p)_{max} = \text{or} < I_{op}$.

Output, Efficiency and Internal Losses, for Normal Oscillation.—The effect of the load resistance R on the output of a tube could be predicted by noticing that the alternating current I_p really flows through R and the tube resistance, R_p, in series; as R is decreased I_p increases, just as the load current from any alternator increases when the resistance of its load circuit is decreased. The voltage E_g impressed on the grid

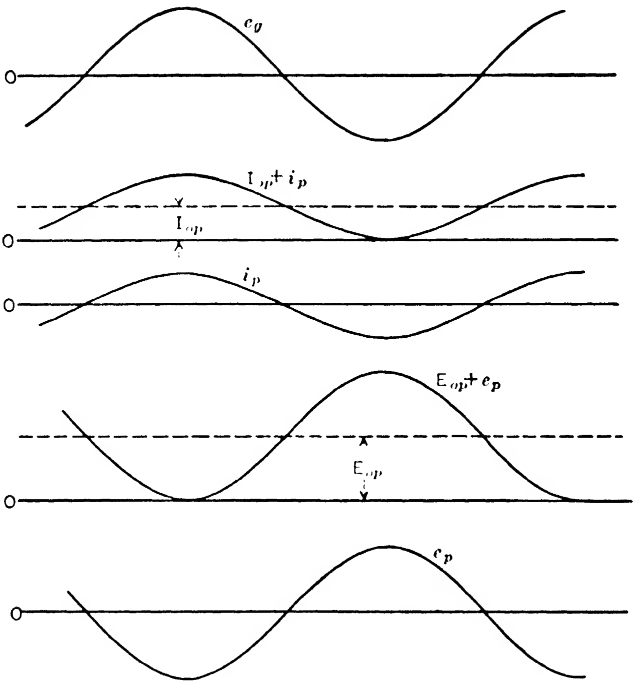


FIG. 127.—Theoretical curves of voltages and currents in a tube; actually the plate voltage does not go through such wide variations.

circuit generates in the plate circuit an alternating current through R_p and R in series.

If sufficient excitation is supplied to the grid circuit to force the actual plate current to vary between zero and $2I_{op}$, the maximum value of the alternating current, I_m, though R and R_p (in series) is I_{op}. The a.c. power delivered to the external circuit is

$$P = \frac{I_m^2}{2}R = \frac{1}{2}\left(\frac{\mu E_{mg}}{R_p+R}\right)^2 R, \quad \ldots \ldots \quad (34)$$

E_{mg} being the maximum value of the voltage impressed on the grid.

If now $R=R_p$, we have

$$P=\frac{(\mu E_{mg})^2}{8R}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (35)$$

The generated voltage, μE_g, is used up in overcoming the two drops I_pR_p and I_pR, so we have

$$I_{mp}R_p+I_{mp}R=\mu E_{mg}$$

and if

$$R_p=R,$$

$$2I_{mp}R=\mu E_{mg}.$$

But the maximum possible value of the drop across R is E_{op}; we therefore have $\mu E_{mg}=2E_{op}$. Hence from Eq. (21), we get

$$P_m=\frac{4E_{op}^2}{8R}=\frac{E_{op}^2}{2R}=\frac{E_{op}}{2R}(I_{mp}R)=\frac{E_{op}}{2R}I_{op}R=\frac{E_{op}I_{op}}{2}. \quad . \quad . \quad . \quad (36)$$

But on the assumption that the resistance of the choke coils is negligible, the input to the plate circuit is $E_{op}I_{op}$, the value of which we assume, is independent of the magnitude of the external resistance R. It therefore follows that a *separately excited tube having sinusoidal variations in the plate current has a maximum efficiency of 50 per cent, and that this occurs for the same condition as gives maximum output, i.e.,* $R=R_p$.

This theoretical limit of efficiency is never reached, because the plate current cannot be made to execute harmonic changes and still be forced to zero value. The reason for this is the variation in μ when the plate voltage becomes very small and the grid voltage large (in positive value); neither does μ hold constant when the plate voltage is very high and with a high negative potential on the grid.

Of course the efficiency factor of 50 per cent neglects the losses in the grid, or exciting circuit, which really should be charged up to the tube, and also the power required to heat the filament. These two factors very materially reduce the possible efficiency of the tube as a generator.

As mentioned above, this limiting figure of 50 per cent for efficiency holds only for sinusoidal plate current; it is possible to so operate the tube that the plate current is much distorted and at the same time the efficiency is increased to perhaps 85 per cent or more. This case will be taken up later in this chapter.

A large power tube was connected as indicated in Fig. 126 and the effect of variation in R was noted. The grid excitation E_g was kept sufficiently low so that the tube was not being worked near its limiting output for any value of R used.

The results are given in Fig. 128, and serve well to show how the power output varies with the resistance of the load circuit; the magnitude of the alternating current generated by the tube is also shown on the curve sheet. It is apparent that this tube should be used with a load circuit resistance close to 1000 ohms if maximum power is to be obtained.

The effect of continued operation on the characteristics of the tube is shown by the dotted curve; it shows the output (for exactly the same conditions as were used for the solid curve) after the tube had been operating for twenty minutes. The temperature of the filament depends not only on the filament current, but also on the temperature of the plates;

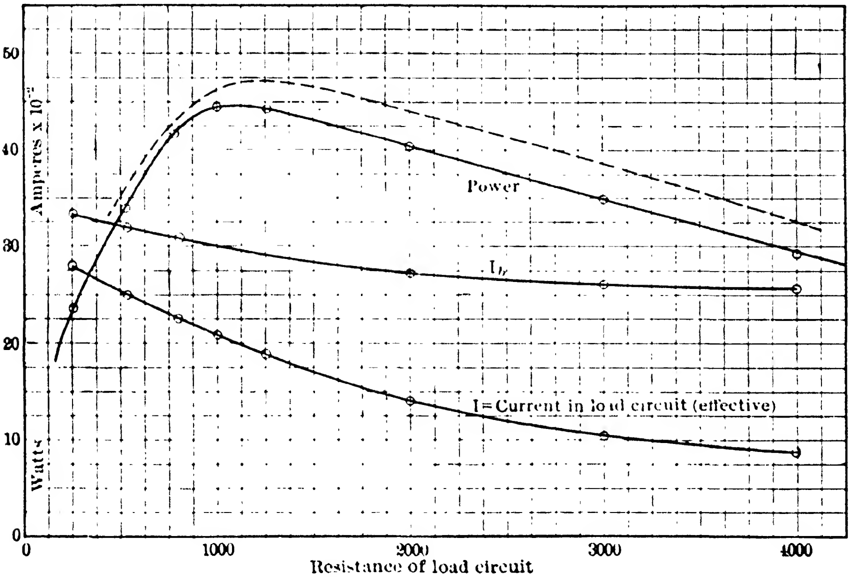


FIG. 128. Variation in output of a power tube as the resistance in the load circuit is varied, grid excitation remaining constant; circuit connected as shown in Fig. 126, with 600 volts used in the plate circuit.

the hotter the plates the higher will be the filament temperature for a given filament current, and of course the more will be the emission of electrons.

For the lower values of R (less than 500 ohms) it will be noticed that the alternating current exceeds 0.707 of the current supplied by the machine in the plate circuit; with sinusoidal current in the load circuit this condition could not occur, it must therefore be that the current in the load circuit was distorted in form when the lower values of load circuit resistance were used.

The curves do not show faithfully the characteristics of the tube as a generator for the higher values of the load circuit resistance because the choke coil used in the plate current circuit had an impedance of only 8000 ohms, so that the supply current was far from constant for the higher values of R. This supply circuit acted as a partial short-circuit for the load circuit, more so as R increased in value.

Heating of the Plates of a Tube.—The safe limit of operation of a power tube is fixed by the allowable heating of the plates; with no oscillations taking place (no excitation of input circuit) the total power delivered by the plate circuit battery or generator must be used in heating the plate, the resistance of the choke coil L (Fig. 126) being negligible. When the tube is oscillating to the extent indicated by the curves of Fig. 127, one-half the input E_bI_{op} is delivered to the output circuit R, *hence only one-half of* E_bI_{op} *is used in heating the plates, whereas if the excitation is removed the heating of the plates is given by* E_bI_{op}. If, therefore, a tube is rated as 250 watts on the plate, the product E_bI_{op} must not exceed 250 when the tube is not oscillating, but if the tube is generating a.c. power, and conditions are adjusted for maximum output (sinusoidal variations of I_p assumed) the input, E_bI_{op}, may be safely increased to practically double the rating, or 500 watts.

Another way of obtaining the amount of power used on the plate is to write the expression for $E'_pI'_p$ from Fig. 126 (where E'_p and I'_p are the voltage between plate and filament, and current through tube, respectively) and find its average value. It is

Power expended on plates $=\frac{1}{T}\int_0^T E'_pI'_pdt$. Now as E_p and I_p are 180° out of phase (high plate voltage occurring at the same instant as low plate current occurs for resistive load), we have for the power used on the plates (where E'_p and I'_p are fluctuating as much as shown in Fig. 127,

$$\text{Power } =\frac{1}{T}\int_0^T (E_b+E_b\sin pt)(I_{op}+I_{op}\sin(pt+\pi))d(pt)$$

$$=E_bI_{op}-\frac{1}{T}\int_0^T (E_bI_{op}\sin^2 pt)d(pt)$$

$$=E_bI_{op}-\frac{E_bI_{op}}{T}\int_0^T\left(\frac{1-\cos 2pt}{2}\right)d(pt)=E_bI_{op}-\tfrac{1}{2}E_bI_{op}=\tfrac{1}{2}E_bI_{op}. \quad (37)$$

If the circuit is not adjusted to give maximum output the proportion of the input power which is used in heating the plates is increased, so the input power must be reduced if safe rating of the plates is not to be exceeded. The input can in general be cut down by decreasing either the filament current or plate voltage, or by introducing a suitable battery or other device into the grid circuit so as to lower its normal potential.

As a black surface radiates much more effectively than does a polished one it is customary to give the plates of the large air-cooled tubes a dull black finish. This very much increases their power capacity.

Phase Relations of Voltages and Current in a Vacuum-tube Generator.—We have previously mentioned the fact that there is no appreciable lag or lead of the electron current in a vacuum tube with regard to the electric field causing the current to flow; this is true for the highest frequencies ever generated by vacuum-tube circuits. As the electric field is produced by both the plate and grid acting together we have the fundamental fact expressed by Eq. (7), which holds for the instantaneous value of the current as well as for the steady state. If then e_p designates the instantaneous value of the alternating component of the plate voltage, e_g, the same for the grid voltage and i_p is the instantaneous value of the alternating component of the plate current (grid current to be neglected in this discussion), we have,

$$i_p = A(e_p + \mu e_g). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)$$

This equation holds true only if the value of $(e_p + \mu e_g)$ is sufficiently low to produce changes in plate current directly proportional to the voltage $(e_p + \mu e_g)$; this means that the fluctuation in plate current must be so small that the portion of the I_p-E_g curve coming into play, is essentially a straight line. Under this condition it is evident that the constant, A, is the reciprocal of the a.c. plate circuit resistance, which we have previously called R_p. The value of this R_p will depend upon the constant values of plate and grid potentials, E_{op} and E_{og}, increasing with a decrease of either of them.

The viewpoint from which the action of the triode may be most easily grasped treats it as an ordinary alternator. The steady quantities, I_{op}, E_{og}, E_{op}, etc., do not enter into consideration at all except as they fix the plate circuit resistance, etc. The plate current then is an alternating current (actually of course it is a fluctuating one, never reversing) and the voltages we consider are alternating voltages.

The voltage which produces the alternating plate current is impressed on the grid; this grid voltage produces changes in the plate current just as though the voltage had been introduced directly in the plate circuit. But the grid voltage is much more effective in controlling the plate current than would be the same voltage introduced directly into the plate circuit, and is greater by the factor μ.

We then consider a triode, having a voltage e_g impressed on the grid, as a diode having a voltage, μe_g, impressed in the plate circuit. We then forget the grid exists, except when it is necessary to consider the amount of current taken by the grid, this sometimes having an important effect on the operation of the tube.

This viewpoint is illustrated in the two diagrams of Fig. 129. In (*a*) is shown the actual circuit, with the *B* battery forcing the plate current to flow. This current is made to increase and decrease by the action of the grid voltage e_g. The equivalent circuit is shown in (*b*) and this is the one we use in deriving the relations of current and voltages of the triode. We suppose there is an a.c. generator between the filament and plate generating a voltage μe_g and that the internal resistance of this generator is R_p. This R_p depends for its value on the steady voltages used in the grid and plate circuits.

Actually in a triode the value of R_p varies with the magnitude of the plate current (alternating) not only increasing when too much current is drawn from the tube, but actually varying during the cycle. This is just the same action as occurs in an alternator having its magnetic circuit of iron with a very small air gap. The internal impedance of such an alternator not only varies with increase in load current but it too varies during

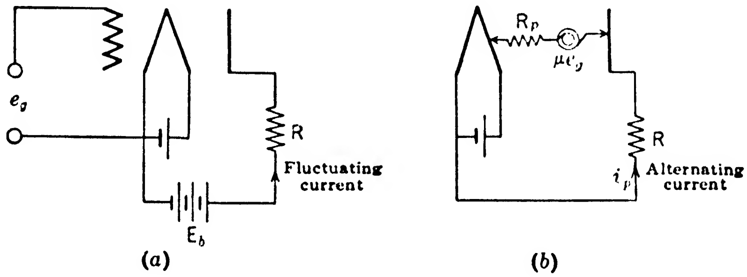


FIG. 129.—The three-electrode tube is most easily analyzed as a generator of a.c. power by supposing the grid absent, but a voltage acting in the plate circuit equal to μe_g.

the cycle. The effect of variation in R_p will be briefly mentioned later on.

Now referring to Fig. 129 (*b*), we see

$$\mu e_g = i_p(R_p + R). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (39)$$

This is exactly the same relation we have for an ordinary generator when we say the generated voltage is equal to the sum of the internal resistance drop and the external drop. The external drop is evidently the voltage impressed on the load, i.e., the terminal voltage of the generator.

If we construct the ordinary reaction diagram for the triode (using Fig. 129 (*b*)), we get results as shown in Figs. 130–132. In each case the current is taken for the reference vector. The resistance reaction due to the internal resistance of the tube is of course shown just opposite in phase to the current vector *OI*. When the load is resistive only its reaction acts in phase with that of the tube. The generated voltage must be just equal and opposite to the sum of the reactions (for any kind of an electrical circuit) and so for Fig. 130 it falls in the same phase as the current.

By reference to Fig. 129 (*a*) we see that the voltage between the filament and plate of the tube must be equal to the voltage of the *B* battery minus the drop in the load circuit. But the "drop" in the load circuit is just equal and opposite to the reaction in the load circuit so that the voltage E_p for any case is equal to the *B* battery voltage plus the load circuit reaction. Thus for the inductive load of Fig. 131 the plate voltage E_p should have a phase lagging behind the current by $180° - \phi$. And for the condensive load the plate voltage, E_p, should lead the phase of the current by $180° - \phi$.

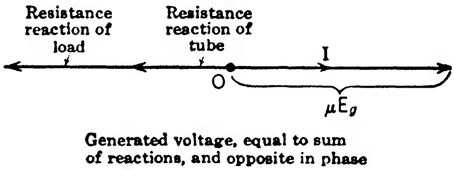


FIG. 130.—Vector relations of the triode working into a resistive load.

In the oscillograms of Figs. 133–135 these relations are well brought out. In these three films the plate current is shown *below* its zero line so that maximum positive value of *I* occurs when the plate current line has the value *lowest* on the film.

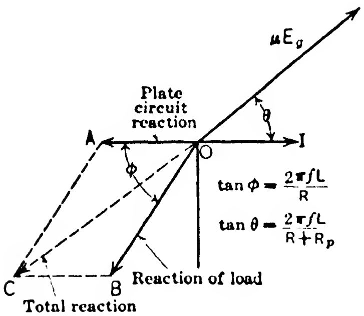


FIG. 131.—Vector relations of the triode working into an inductive load.

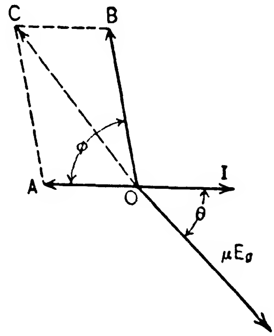


FIG. 132.—Vector relations of the triode working into a condensive load.

For Fig. 133 then the generated voltage, e_g, and current (i_p minus the steady value I_{op}) are evidently in phase. The voltage e_p should be equal to the *B* battery voltage plus the load resistance reaction (which is opposite in phase to *I*) and so should have its minimum value when the current has its maximum value and this is well shown on the film. The phases cannot be measured with any great accuracy because the curves of i_p and e_p are not pure sine waves even though a sine wave was used for the exciting voltage, e_g.

For inductive load (Fig. 131) we concluded that the plate voltage e_p should have its phase lagging $180° - \phi$ behind the current. On the film of Fig. 134 this is seen to be so, as nearly as the distorted waves permit the phase to be determined. A similar conclusion is reached by inspection of Fig. 135.

It will be noted by comparing Figs. 134 and 135 that the conclusions reached in Chapter I regarding the action of coils and condensers is well shown in the action of these tube circuits. In the inductive circuit of

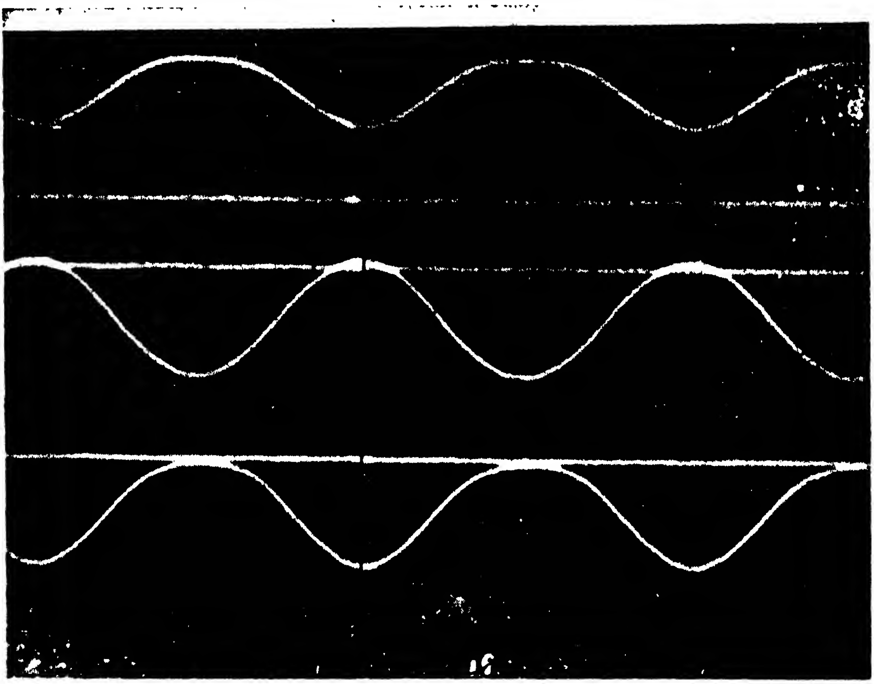


FIG. 133. Oscillogram of grid voltage, plate voltage, and plate current, corresponding to conditions of Fig. 130.

Fig. 134 the distortions, which the internal reactions of the tube bring about, occur in the voltage, e_p, the current i_p being nearly a sine wave. For the condensive load (Fig. 135) the opposite is true; the voltage, e_p, appears as a pretty fair sine wave and the distortions occur in the current, i_p.

Quantitatively we have the following simple relations. If the frequency of the grid voltage is f and magnitude E_g, the tube resistance R_p, and load impedance R, $R+j2\pi fL$, or $R-1/j2\pi fC$ as the case may be

Resistive load $$I=\frac{\mu E_g}{R+R_p}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (40)$$

Inductive load $$I=\frac{\mu E_g}{\sqrt{(R_p+R)^2+(2\pi fL)^2}}. \quad . \quad . \quad . \quad . \quad . \quad (41)$$

Capacitive load $$I=\frac{\mu E_g}{\sqrt{(R_p+R)^2+\left(\frac{1}{2\pi fC}\right)^2}}. \quad . \quad . \quad . \quad . \quad . \quad (42)$$

Effect of Phase Relations on the Possible Power Output of a Tube Generator.—From the foregoing analysis it is evident that a tube generator can act on its output circuit with a voltage E_p, the maximum value

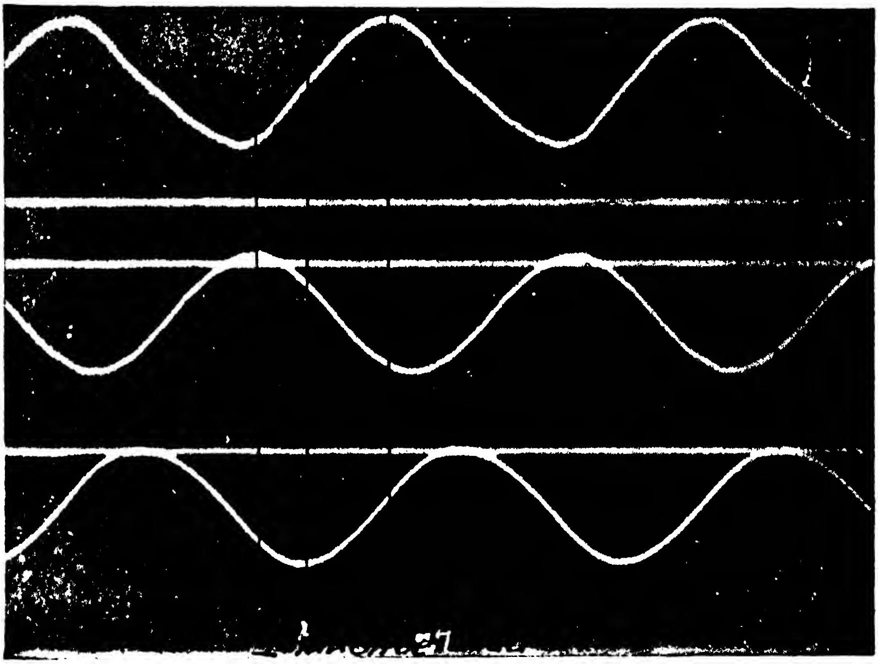


FIG. 134. Oscillogram of plate voltage, grid voltage, and plate current, corresponding to conditions of Fig. 131.

of which is somewhat less than the normal plate voltage [1] E_{op}; also that it can supply to the output circuit an alternating current I_p, the maximum value of which is somewhat less than the normal plate current I_{op}. If I and E represent the effective values of voltage and current which the tube furnishes to its output circuit, it is evident that the maximum power output will occur when the load circuit is such as to bring I and E in phase and this power is equal (in case of maximum output) to $\frac{1}{2}E_{op}I_{op}$.

[1] For resistive load this normal plate voltage is approximately one-half the voltage of the machine used in the plate circuit.

Now if the load circuit is such that E and I are in phase, it is evident that its impedance must be resistance only, furthermore the value of this resistance must be equal to E/I, which is also the a.c. resistance of the plate circuit of the tube. The truth of this statement was shown in Fig. 128.

For such a circuit as that given in Fig. 126, the magnitude of current I_p must be directly proportional to E_g, as indicated by Eq. (40). Due to the non-sinusoidal currents and voltages, however, this relation does not hold good, except for low values of grid excitation; the alternating

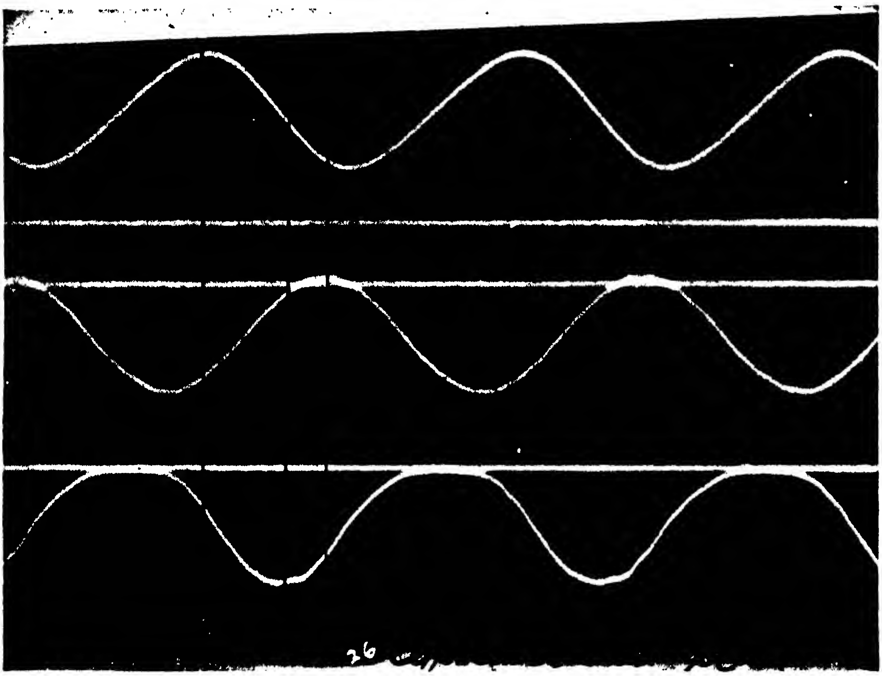


FIG. 135.—Oscillogram of plate voltage, grid voltage, and plate current corresponding to conditions of Fig. 132.

current does not change as rapidly as indicated by Eq. (40). In Fig. 136 are shown curves of load circuit power and current as functions of E_g, the resistance of the load circuit having been adjusted equal to the tube resistance at low excitation. The power output increases with the square of the grid voltage for very low grid voltages only and for the higher values of excitation the output is increasing at a rate lower even than the first power of the grid voltage.

There is shown also in Fig. 136 the value of the current taken by the grid; as long as the grid was not forced positive with respect to the filament the reading of the c.c. ammeter in the grid circuit was zero, but when the value of alternating voltage impressed on the grid exceeded $1/\sqrt{2}$ of the normal negative grid potential E_{og}, the grid was positive for a small portion of the cycle and so took current. The variation of the reading of the grid ammeter is shown by the curve marked I_g; it was zero until E_g reached a value of 80 volts (effective) which is approximately equal to 70 per cent of the voltage, E_{og}, and for higher voltages rose to a value of several milliamperes. This is the average value of the grid current, because a c.c. ammeter reads average values; the grid current flows for only a small fraction of the cycle, so that the actual maximum value of I_g is probably ten times as large as the value given on the curve sheet.

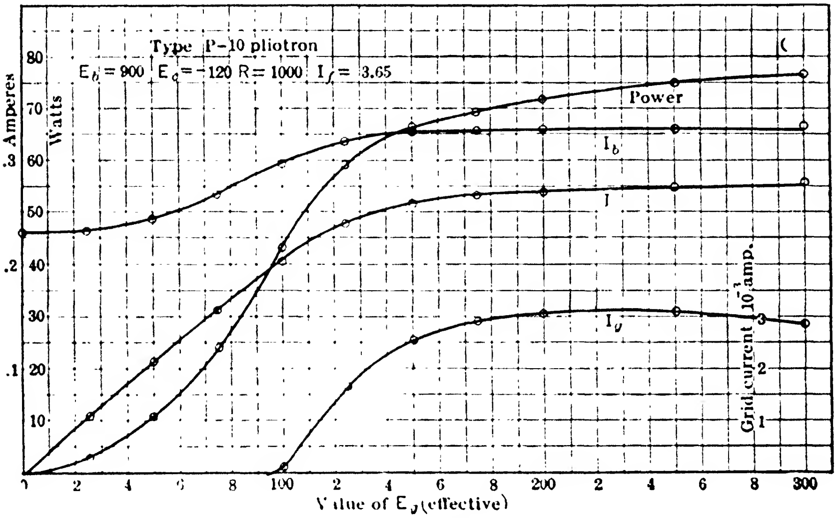


FIG. 136. —Variation of output current and power as the exciting voltage on the grid is increased; the circuit was arranged as shown in Fig. 126. Variations of I_b and the average grid current are also shown.

An accurate analysis of these voltages and currents will be given in a later section of this chapter.

General Analysis of the Conditions Necessary for Self-excitation.— From the analysis given so far it is evident that if a vacuum tube is going to operate efficiently as a generator of a.c. power, it is necessary to have in the plate circuit a load having a resistance equal to that of the tube; it is also necessary to have such reactions occurring in the circuit to which the tube is connected that when the plate current undergoes sinusoidal variations the plate potential and grid potential both undergo sinusoidal variations of potential in opposite phases, and that the relative magnitudes of these two potential variations be properly adjusted for the tube being used. The fundamental requirements of the problem can be readily specified. In Fig. 137 are shown the filament, plate, and grid terminals 1, 2, and 3; the filament battery and plate circuit battery (or machine) are omitted, as they do not enter directly into the determination of the conditions for self-excitation of the tube.

If the normal plate voltage and plate current are E_{op} and I_{op} respectively, we know that the tube can, when operating properly, generate an amount of power somewhat less than $\frac{1}{2}E_{op}I_{op}$; let us call this available power P. If this power is supplied to a circuit of L, C and R, in series, it will produce a current fixed in magnitude by the relation $P=I^2R$. This current, I, will produce an alternating voltage between the terminals of L, the effective value of which is equal to $I\omega L$, where ω is nearly equal $1/\sqrt{LC}$.

When generating the amount of power P the potential of point 2 must be fluctuating in voltage (with respect to the filament) by an amount approximately equal to E_{op}. The potential of point 3 must be fluctuating, with respect to the filament, by an amount E_{mg}, such that μE_{mg} is about equal to $2E_{op}$, as shown in Fig. 130.

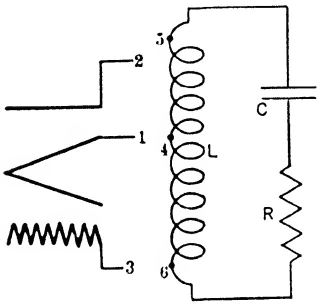


FIG. 137.—Diagram to show the conditions required for self-excitation; an oscillatory circuit, of suitably low resistance, must be connected to plate, filament, and grid about as indicated. (Filament and plate circuit batteries not shown.)

It is then evident (referring to Fig. 137), that if we connect the filament to point 4 of the coil, we must connect points 2 and 3 to points in the coil (*on opposite sides of point* 4, such as 5 and 6) such that the maximum value of voltage between 4–5 is equal to E_{op} and the maximum voltage between points 4–6 is equal to $(2/\mu)E_{op}$.

If the resistance drop in the coil is negligible compared to the reactance drop, we must have (neglecting the effect of mutual induction between L_{4-5} and L_{4-6}).

$$\sqrt{2}I\omega L_{4-5}=E_{op} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (43)$$

$$\sqrt{2}I\omega L_{4-6}=\frac{2}{\mu}E_{op}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (44)$$

The current flowing through the coil L between points 4 and 5 is really the combination of the actual pulsating plate current and the current I which is flowing in the oscillatory circuit. If I is large compared to the alternating component of the plate current I_p, the error made in assuming the drop between points 4 and 5 as due to I only is small. In a typical radio circuit I was 0.50 ampere and the effective value of I_p (alternating component of plate current) was only 0.03 ampere. It will be noticed, however, that if the resistance of the oscillating circuit is large I decreases in value so that the assumption is no longer justified.

If we now use the approximate relation

$$P=\frac{E_{op}I_{op}}{2} \text{ and } P=I^2R,$$

we get from Eq. (43)

$$\omega L_{4-5}\sqrt{\frac{E_{op}I_{op}}{R}}=E_{op},$$

or

$$\omega L_{4-5}=\sqrt{\frac{RE_{op}}{I_{op}}}=\sqrt{2RR_p},$$

on the assumption that $R_{op}=2R_p$.[1]

If we now make the assumption that $\omega=1/\sqrt{LC}$ (which would be nearly true if the input and output circuits of the tube had negligible capacities) we find

$$L_{4-5}=\sqrt{2RR_pLC}. \qquad \text{(45)}$$

And for the proper grid excitation, we should have

$$L_{4-6}=\frac{2}{\mu}\sqrt{2RR_pLC}. \qquad \text{(46)}$$

As an illustration of how these approximate relations are applied to an actual circuit, we suppose a set designed to generate a frequency of 600,000 cycles per second (500-meter wave), the capacity C being 0.0004 μf. The total resistance of the oscillating circuit is 10 ohms, the μ of the tube to be used is 4 and R_p is 3000 ohms. Using the relation

$$\lambda=1885\sqrt{LC},$$

we find

$$LC=0.0705,$$

and therefore

$$L=176\mu h.$$

[1] In case R_p is greater than specified by this relation, the equation for L_{4-5} must be correspondingly changed; thus if $R_p=R_{op}$ then we must have

$$\omega L_{4-5}=\sqrt{RR_p}.$$

As previously mentioned R_p varies with the amount of excitation on the grid; a curve showing the variation in R_p is given in Fig. 151, p. 604.

Then we find, $L_{4-5}=65\mu h$ and for L_{4-6}, we find $32\mu h$. If the tube can supply 4 watts of power, the current in this oscillating circuit would be

$$\sqrt{\frac{4}{10}}=0.632 \text{ ampere.}$$

If $R_p=3000$, we have $E_{op}/I_{op}=6000$, and also we have $E_{op}I_{op}=8$. From these two equations, we find that $E_{op}=220$ and $I_{op}=0.037$. From the above values, we have

$$\omega L_{4-5}I=(2\pi 6\times 10^5)\times(65\times 10^{-6})\times 0.632=153,$$

this being the effective value of the voltage impressed on the plate. But this is equal to $E_{op}\div\sqrt{2}$, as we have already assumed necessary for generating a power equal to $E_{op}I_{op}/2$.

This elementary analysis serves for an approximate solution of the circuit; the filament would be connected to point 4, somewhat lower than the middle of the coil and points 5 and 6 should be adjustable by multipoint switches. Normally there should be $65\mu h$ between points 4 and 5 for the plate connection, and $32\mu h$ between points 4 and 6 for the grid connection.

The foregoing calculations have been made on the assumption that the a.c. output of the tube was 50 per cent of the input. Actually on a small tube like this 25 per cent efficiency would be more likely than 50 per cent; this would decrease the value of I and so require an increase in the required values of L_{4-6} and L_{4-5}.

As the alternating component of the plate current I_p is practically 90° out of phase with the power circuit current I, the required phase difference of 180° between E_p and E_g will not be obtained if I_p is appreciable compared to I. This shift in phase of E_p as the ratio I_p/I increases, very materially reduces the possible output of the tube. This has led to the conclusion that the circulating power in the oscillatory circuit should be at least 10 times the actual power drawn from the circuit by the load.

If it should happen that R and R_p are so high that the L_{4-5} required is more than about two-thirds of the whole coil L, the conditions required by Eqs. (43) and (44) could not be satisfied by this circuit, so it would not oscillate.

The case has many features in common with a shunt-wound self-excited generator. In such a machine maximum output is reached when the external resistance is equal to the internal resistance of the machine (if the generated e.m.f. is kept constant); also if there is too much resistance in the shunt-field circuit of the machine, it will not excite itself, or " build up," as it is called. This corresponds somewhat to a tube circuit having too high a resistance in the oscillating circuit.

The Oscillating Tube as a Detector of Undamped Waves.—From the explanation of the action of the three-electrode tube as a detector of high-frequency currents, given on pp. 534 et seq., it is evident that the amplitude of the high-frequency current must vary with audible frequency if an audible response is to be given by the telephone. In continuous-wave telegraphy the signal current received by the antenna does not have variations in amplitude; a dot, e.g., might consist of 5000 cycles of a 50,000-cycle current, the amplitude of the current being constant for the duration of the 5000 cycles.

If the input circuit of the detecting tube is continually excited by a locally generated frequency of 49,000 cycles, when the signal comes in the input circuit is excited by both 49,000 cycles and 50,000 cycles, the result being a high-frequency excitation the amplitude of which varies

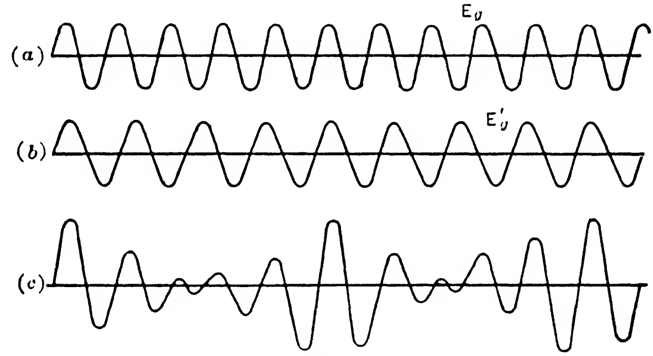


FIG. 138.—Conventional diagram of a continuous-wave signal voltage E_g, a locally generated voltage E'_g, of slightly different frequency, and the sum of the two, which is the voltage acting on the grid.

1000 times a second. This high-frequency, variable amplitude, input voltage will give a 1000-cycle note in the telephones, connected in series with the plate circuit of the tube. In case the locally generated, high-frequency current is produced by the detecting tube itself, it has been called *autodyne reception*, in case some device other than the detecting tube is used for impressing the local high-frequency current on the grid the scheme is called *heterodyne reception*.

The excitation of the input circuit when no signal is arriving is due to the voltage E'_{mg} sin ωt, and when the signal, E_{mg} sin pt, is being received the actual excitation of the grid circuit is (E'_{mg} sin $\omega t+E_{mg}$ sin pt) as indicated in Fig. 138.

Detection with no Grid Condenser.—We have previously shown that if the grid is actuated by a voltage E'_{mg} sin ωt and if the plate current varies as the square of the grid potential, the increase in plate current is given by $\frac{1}{2}$ (average value of e_g^2)$\times d^2I_p/dE_g^2$. Hence when the excitation is such as given by curve, Fig. 138, the increase in plate current is

$$\Delta I_p = \text{average value of } \frac{(E'_{mg}\sin\omega t + E_{mg}\sin pt)^2}{2}\,\frac{d^2I_p}{dE_g^2}$$

$$= \left\{\frac{E_{mg}^2}{4} + \frac{E'^2_{mg}}{4} + \text{average value of } (E_{mg}E'_{mg}\sin\omega t\sin pt)\right\}\frac{d^2I_p}{dE_g^2}. \quad (47)$$

The first two terms give the increase in the plate current which is constant, as long as the excitation is applied; their effect would produce an increase in the value of the plate current as read by a c.c. ammeter in the plate circuit, but they would not produce a readable signal in the phones, giving only a slight click in the phones when the excitation is put on the grid and another when it is taken off.

Whatever audible signal is obtained must come from the third term; this may be written in the expanded form

$$\tfrac{1}{2}E_{mg}E'_{mg}[\cos(\omega-p)t - \cos(\omega+p)t]\,\frac{d^2I_p}{dE_g^2}.$$

The average value of both these cosine terms is zero, but $\cos(\omega-p)t$ may fluctuate so slowly as to produce an audible signal in the phones, and it is this term which is useful in continuous-wave detection. The strength of signal is then measured by this term.

$$\Delta I_p(\text{of audible frequency}) = \frac{E_{mg}E'_{mg}}{2}\cos(\omega-p)t\,\frac{d^2I_p}{dE_g^2}. \quad . \quad (48)$$

The frequency of this fluctuation in the plate current, which is the note heard in the phones, is adjustable by the operator, as he can make the value of ω anything he may desire. The ear and phone are both most sensitive at a frequency of about 800 cycles per second, so ω is generally adjusted to give $(\omega-p)/2\pi = 800$, or $(p-\omega)/2\pi = 800$.

It is to be noticed that whereas the response of the tube detector is proportional to the square of the signal strength for damped wave signals Eq. (17), it is proportional to the first power of this signal strength when used for continuous-wave receiver. This fact makes the tube a better detector of signals for undamped, than for damped, waves, its sensitiveness not decreasing with the strength of signal so rapidly for one as it does for the other. Eq. (48) shows also that the response to a given signal varies with E'_g the amplitude of the local oscillations, so long as the variation of E'_g does not change the value of d^2I_p/dE_g^2.

This increase in response with the strength of the local oscillations is similar in character to the increase in response of a telephone receiver due to the use of the permanent magnet. It is not a characteristic peculiar to a vacuum tube, but holds for any detecting device in which the response varies with square of the impressed force (when a single-frequency is impressed). A crystal rectifier has a nearly parabolic relation between the current through it and the impressed voltage (see Fig. 34, p. 443) and the curve of response as a function of the signal strength is as shown in Fig. 139, when it is used to detect spark signals. If, however, the crystal is used to detect continuous-wave signals by use of an auxiliary source of continuous wave excitation (Fig. 140), its response follows the same law as obtained for the vacuum-tube receiver, given in Eq. 48; its response will be proportional to the first power of the voltage of the received signal, not as the square. This is indicated in Fig. 141.

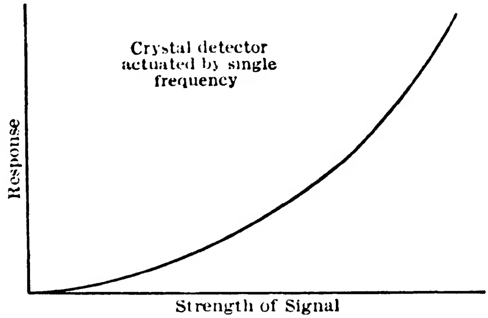


FIG. 139.—Rectifying action of a crystal actuated by a single frequency.

The crystal rectifier will also act like the vacuum tube in that the response of the rectifier, for a given signal strength, will vary as the first power of the locally generated e.m.f. until the value of this e.m.f. is such that the region of parabolic rectification is exceeded. The response of the crystal, for given signal voltage, as the value of E'_0 is varied, is about as shown in Fig. 142; the response is proportional to E'_0 for a certain range, then ceases to increase with E'_0 and if E'_0 is still further increased the response falls and may reach practically zero for excessively large values of E'_0.

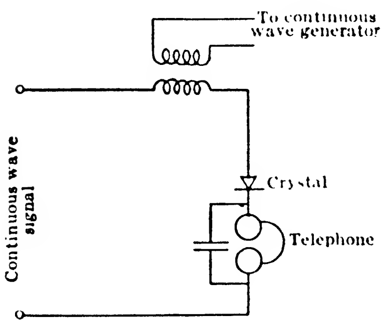


FIG. 140.—A possible scheme for hearing continuous-wave signals with a crystal detector.

This same characteristic holds for the vacuum tube used as a beat receiver, the static characteristic of a tube being as indicated in Fig. 143; if the amplitude of the locally generated e.m.f. is OC, the response (for given signal strength) will be about twice as strong as if E'_0 had the amplitude OB only, whereas a value of E'_0 equal to OD would result in a signal perhaps less than for $E'_0=OB$. If E'_0 is increased to the value OE the response to the signal will be practically zero.

Detection with Grid Condenser.—In case a condenser is used in series with the grid of the tube being used as a beat receiver, Eq. (23)

must be used in predicting the detection efficiency. The question is somewhat more involved than for the tube with no grid condenser, because the normal grid potential (average value with no signal coming in) varies with the value of E'_g, the potential decreasing as E'_g is increased in value. As all four of the derivatives used in Eq. (23) vary somewhat as the normal grid potential is varied, an exact expression for the detection factor must be rather complex. As the tube is used in practice the most sensitive condition is easily found as will be described in a succeeding paragraph, dealing with the self-excited, oscillating tube as detector.

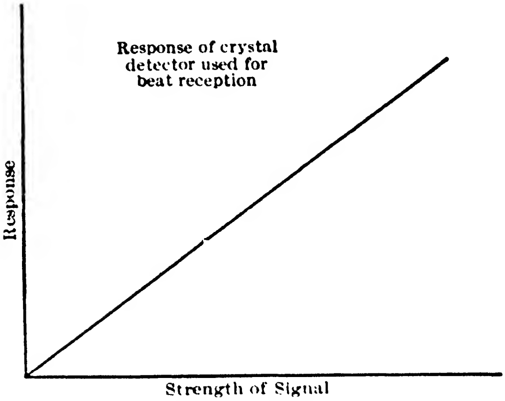


FIG. 141.—Rectifying action of a crystal used as indicated in Fig. 140.

Analysis of Some of the More Commonly Employed Circuits for the Self-excited Vacuum-tube Oscillator.—The first circuit to be analyzed is one having a tuned plate circuit, the grid being excited by magnetic coupling with this tuned circuit; the connections are as indicated in Fig. 144. The actual plate current is $I_{op}+i_p$, actual plate voltage $E_{op}+e_p$, and actual grid voltage $E_{og}+e_g$. We suppose that $(E_{op}+\mu E_{og})$ is of such a value that I_{op} is about one-half of the saturation current of the tube, as shown in Fig. 145; when the tube is in the oscillating state the value of $(E_p+\mu E_g)$ fluctuates between OF and OE, and the plate current fluctuates between the values CF and DE. Through this range

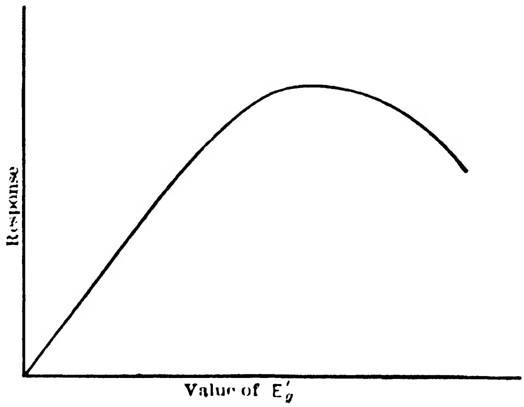


FIG. 142.—Rectifying action of a crystal detector as a function of the amplitude of the locally impressed voltage, the signal voltage being of constant amplitude.

of fluctuation in the plate current, the relation between I_p and $(E_p+\mu E_g)$ is nearly linear so that

$$i_p=A(e_p+\mu e_g), \quad \ldots \ldots \ldots \quad (49)$$

A being a constant which, upon inspection, must be the inverse of a resistance. This relation is not true of course for the extreme values of i_p, the proportionality factor decreasing as the value of i_p approaches the values of CF and DE.

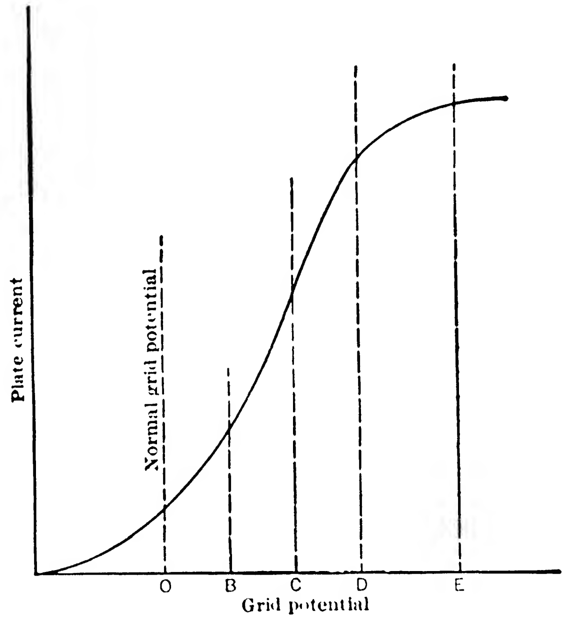


FIG. 143. The response of a tube used for receiving continuous-wave signals will vary with the strength of the local oscillations, the same as the crystal; a local oscillation of amplitude OC would give (for a fixed incoming signal) response about twice as great as if the amplitude were OB, but an amplitude OE would give but very little response.

This point is illustrated in Fig. 146, in which a sine wave of $(e_p+\mu e_g)$ is shown and below it in full line the alternating plate current i_p; the dotted additions to this curve serve to show how i_p differs from a sine wave. The curve i_p shown in Fig. 146 is symmetrical about the zero axis, a condition rarely obtained in an actual tube circuit. It occurs only if the plate-current curve shown in Fig. 145 is symmetrical about the point A. This supposes that the upper part of the curve (caused by saturation) is of the same form as the lower part of the curve (caused by effect of space charge) and also that $(E_{op}+\mu E_{og})$ has been properly adjusted to bring point A to the middle part of the curve.

For such a special state of affairs two effects exist which are practically never found in practice; the value of plate current, as read by a c.c. ammeter, does not change when oscillations begin and the alternating component of the plate current contains no even harmonics. Generally the plate current does change when oscillations are started and the plate current has very pronounced even harmonics.

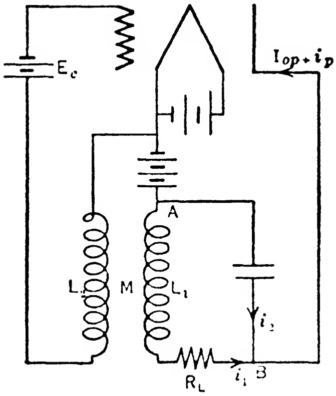


FIG. 144.—A commonly employed circuit for producing oscillations; the frequency is fixed by the constants of the plate circuit and the value of M is the critical factor for production or non-production of oscillations. Grid biasing battery should be preferably between L_2 and the filament connection.

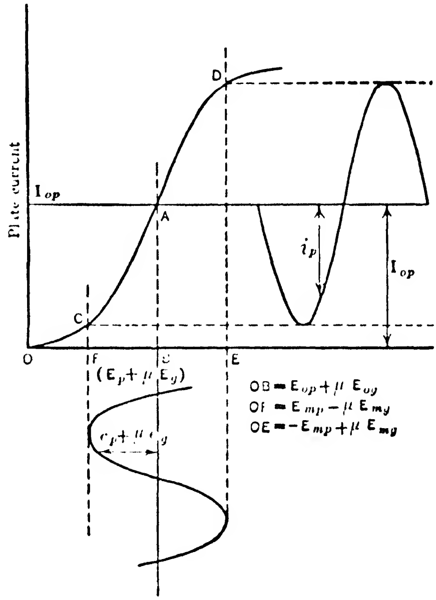


FIG. 145. A sinusoidal variation in grid potential will produce a sinusoidal variation in plate current only if the fluctuation in plate current occurs over a limited range. With a curve symmetrical about the point A the plate current will contain no even harmonics, only odd ones being present. In case the upper and lower bends are not symmetrical about A (which is generally the case), both even and odd harmonics will occur in plate current.

On the basis of Eq. (49), we can write at once, putting $A = 1/R_p$,

$$R_p i_p = e_p + \mu e_g. \quad \ldots\ldots\ldots \quad (49)$$

From the notation given in Fig. 144

$$i_p = i_1 + i_2,$$

i_1 being the alternating component of current through L_1.
Also

$$e_p = -i_1 R_L - L_1 \frac{di_1}{dt}, \quad \ldots\ldots \quad (50)$$

and

$$e_g = -M \frac{di_1}{dt}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (51)$$

To make Eq. (51) true the grid circuit must be so adjusted that no current flows in it as the voltage, e_g, goes through its cycle of values; this requires that at no time throughout the cycle must the grid be positive with respect to the filament, and that the capacity of the grid-filament circuit is negligible. The first condition can be brought about by using a proper value of E_c (Fig. 144), but the second condition cannot be brought about by any adjustment of the tube circuit. In some cases this capacity is of extreme importance; for very high-frequency circuits it may be one of the limiting factors of operation. It must be borne in mind that the capacity to be considered is not the geometrical capacity of the tube, but the *effective capacity* as explained on pp. 527 et seq. (An oscillating tube furnishes maximum power when the external resistance in the plate circuit is equal to R_p, so in calculating the probable effect of the grid-filament capacity of an oscillating circuit the proper value of α to use in Eq. (15), p. 528, is $\mu/2$.)

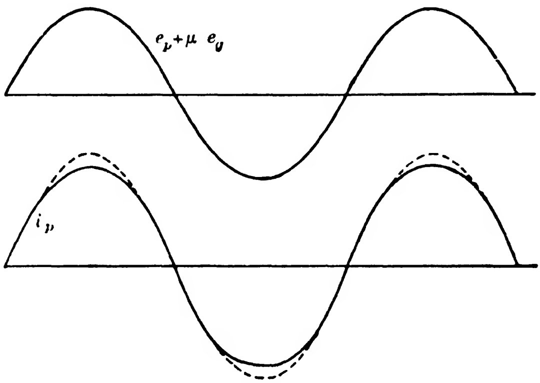


FIG. 146.—Due to the upper and lower curves of the plate current curve of Fig. 145 the actual alternating component of the plate current is flat topped (sine-wave shape shown by dotted lines).

We have also

$$i_2 = -C \frac{de_p}{dt},$$

which by using (50) becomes

$$i_2 = CR_L \frac{di_1}{dt} + CL_1 \frac{d^2 i_1}{dt^2}.$$

By combining the foregoing equations to eliminate e_p, e_g, i_p, and i_2, we get

$$L_1 \frac{d^2 i_1}{dt^2} + \left[R_L + \frac{1}{CR_p}(L_1 + \mu M)\right]\frac{di_1}{dt} + \frac{i_1}{C}\left(1 + \frac{R_L}{R_p}\right) = 0. \quad . \quad . \quad (52)$$

Such a differential equation is satisfied by writing i_1 as an exponential function, the form of the function (trigonometric, hyperbolic, etc.), depending upon the relative values of the various constants in Eq. (52) The roots of Eq. (52) are real if we have [1]

$$\left[R_L+\frac{1}{CR_p}(L_1+\mu M)\right]^2-\frac{4L_1}{C}\left(1+\frac{R_L}{R_p}\right)>0. \quad . \quad . \quad . \quad . \quad (53)$$

For this condition, the current i_1 must be non-oscillatory, the circuit is aperiodic. If $R_L+\frac{1}{CR_p}(L_1+\omega M)>0$, the exponential function is a decreasing one and in case some disturbance occurs in the circuit the disturbance soon disappears and the circuit resumes its normal condition. This can evidently occur if M is positive and also if M is negative provided its value is less than $(1/\mu)(L_1+CR_LR_p)$.

If $R_L+\frac{1}{CR_p}(L_1+\mu M)<0$, any disturbance set up in the system tends to increase itself; this occurs if M is negative and its value such that $M>\frac{1}{\mu}(L_1+CR_LR_p)$. The plate current then tends to increase or decrease (according to the sense of this disturbance) and does so as long as the characteristic curve (Fig. 145) is straight.

In the case the constants of Eq. (52) are such as to make its roots imaginary, we have

$$\left[R_L+\frac{1}{CR_p}(L_1+\mu M)\right]^2-\frac{4L_1}{C}\left(1+\frac{R_L}{R_p}\right)<0. \quad . \quad . \quad . \quad (54)$$

The current i_1 must then be of the form,

$$i_1=A\epsilon^{\alpha t}\sin(\omega t+\theta), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (55)$$

in which we must have

$$\alpha=-\frac{1}{2L_1}\left[R_L+\frac{1}{CR_p}(L_1+\mu M)\right], \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (56)$$

and

$$\omega=\frac{1}{2L_1}\sqrt{\frac{4L_1}{C}\left(1+\frac{R_L}{R_p}\right)-\left[R_L+\frac{1}{CR_p}(L_1+\mu M)\right]^2}. \quad . \quad . \quad (57)$$

The exponential in Eq. (55) is decreasing if $R_L+\frac{1}{CR_p}(L_1+\mu M)>0$, which

[1] For an analysis of an equation of this kind see the first few pages of Chapter III.

is true for all positive values of M, or if M is negative, but its value is such that $M < \frac{1}{\mu}(L_1 + CR_LR_p)$.

For such conditions any shock on the circuit will produce oscillations, of frequency as determined from Eq. (57), but the oscillations will die away because of the negative value of α.

The last, and most important, case to consider is given by

$$R_L + \frac{1}{CR_p}(L_1 + \mu M) < 0,$$

that is, when M is negative and its absolute value is such that,

$$M > \frac{1}{\mu}(L_1 + CR_LR_p). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (58)$$

For this condition any disturbance to the circuit will start oscillations, and these oscillations will increase in magnitude until the straight part of the curve in Fig. 145 is exceeded. In the average radio circuit L_1 is 5 to 10 times as large as CR_LR_p, so that Eq. (58) may be approximately written $M > L_1/\mu$.

The effect of making the plate current fluctuate through such large values is to make R_p variable throughout the cycle, resulting also in an increase in the average value of R_p; the oscillations will therefore increase in amplitude, after once being started, until the value of R_p is increased to such an extent that the inequality given in Eq. (58) is changed to an equality. When this condition is brought about the value of α becomes zero, and the exponential in Eq. (55) reduces to unity, giving neither increase nor decrease in the amplitude of the current.

From the foregoing it is evident that if the circuit of Fig. 144 is to produce oscillations, M must be somewhat greater than its *critical value* given by the relation

$$-M = \frac{1}{\mu}(L_1 + CR_LR_p). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (59)$$

If M exceeds (in absolute value) this value oscillation will start; if oscillations are already present and M is made slightly less than this value (in absolute magnitude), the oscillations will stop, hence the use of the term *critical value*.

The frequency of the oscillations is obtained from Eq. (57) and is

$$f = \frac{1}{4\pi L_1}\sqrt{\frac{4L_1}{C}\left(1 + \frac{R_L}{R_p}\right) - \left[R_L + \frac{1}{CR_p}(L_1 + \mu M)\right]^2}, \quad . \quad . \quad (60)$$

and if M is adjusted to its critical value,

$$f=\frac{1}{2\pi}\sqrt{\frac{1+\dfrac{R_L}{R_p}}{L_1C}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (61)$$

which in the average radio circuit is practically the same as

$$f=\frac{1}{2\pi\sqrt{L_1C}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (62)$$

If the coupling between the grid and plate circuits is made tighter than required for the limiting value, the frequency is somewhat decreased.

If we suppose Eq. (62) to give the frequency, the critical value of M may be written in the form

$$-M=\frac{L_1}{\mu}\left(1+\frac{R_LR_p}{(\omega L_1)^2}\right)=\frac{C}{\mu}\left(\frac{1}{(\omega C)^2}+R_LR_p\right). \quad . \quad . \quad . \quad . \quad (63)$$

From this relation it is evident that if a circuit is oscillating with a value of M equal to the critical value any decrease in the frequency, accomplished by varying either L_1 or C, must be accompanied by an increase in the coupling, otherwise the oscillations will cease.

Self-excited Triode Circuits Always Generate Harmonics.—It is shown in a previous section that the exponential factor, ϵ^{at}, of Eq. (55) is an increasing function (instead of decreasing) if M has greater than a certain critical value. But such an exponential would lead to a current having an infinite value, evidently an absurdity. When the oscillations have built up to a certain value the factor ϵ^{at} must reduce to unity, which means that the circuit actually does something which our simple analysis does not predict.

Eq. (49), from which the solution was obtained, assumes a linear relation between current and voltage, but Fig. 145 shows this to be true for only a small range of current changes. Now if the relation of voltage to current is non-linear, harmonics of the fundamental frequency will be generated, and these harmonics become proportionately greater as the oscillations become more violent. Thus the oscillations predicted by the condition of Eq. (58) will build up in intensity, after they start, and will not stop increasing until the plate current is swinging to such an extent as to cover the upper and lower curved parts of the characteristic given in Fig. 145.

The conclusion to be reached from this simple analysis is that *any self-excited triode circuit will produce many harmonics of the fundamental frequency, and that the tighter the coupling between plate and grid circuits the* *greater will the harmonics be in comparison with the amplitude of the current of fundamental frequency.*

Prediction of Oscillatory Condition by Putting Total Resistance Equal to Zero.—The condition for oscillations in any circuit can be expressed by putting the total resistance of the circuit equal to zero. In Fig. 144 the external plate circuit has a resistance R_{A-B} and this is in series with the effective tube resistance R'_p. This R'_p is the relation between e_p and i_p, *when the values of* i_p *are determined not only by* e_p *but by the simultaneously acting* e_g; as e_g may be 180° out of phase with e_p and of such value that $\mu e_g > e_p$, it is possible to have i_p 180° out of phase with e_p, so that R'_p may be negative, whereas R_p is always positive.

This difference may be formulated by writing

$$R_p = \frac{d(e_p + \mu e_g)}{di_p} \text{ with } e_g \text{ held constant,}$$

and

$$R'_p = \frac{d(e_p + \mu e_g)}{di_p} \text{ with } e_g \text{ and } e_p \text{ both allowed to vary.}$$

In the latter expression we have to remember that when e_p changes i_p changes and thereby may induce voltages which affect e_g. That is, e_g is a function of e_p, and may exert such a controlling effect that R'_p turns out negative. In fact, by using the relation $i_p = A(e_p + \mu e_g)$, upon which all this theory of oscillating circuits is based, we find that, with e_g held constant $di_p = A de_p$ so $de_p/di_p = 1/A = R_p$. Now if e_g is allowed to vary,

$$di_p = A(de_p + \mu de_g) \text{ or } \frac{de_p}{di_p} = \frac{1}{A} - \mu\frac{de_g}{di_p} \text{ or } R'_p = R_p - \mu\frac{de_g}{di_p}.$$

We have then, as the condition for self-sustained oscillations

$$R_{A-B} + R'_p = 0. \qquad (64)$$

But from Eq. (59) of Chapter I we know that the resistance of the parallel circuit, at resonance, is given by

$$R_{A-B} = \frac{L_1}{C}\frac{1}{R_L}. \qquad (65)$$

Now $de_g = -\omega M di_1$, de_p is nearly equal to $-\omega L_1 di_1$ and

$$di_p = \frac{1}{R_p}(de_p + \mu de_g).$$

Then

$$R'_p = \frac{de_p}{di_p} = \frac{\omega L_1 di_1}{\frac{1}{R_p}(\omega L_1 di_1 + \mu\omega M di_1)} = \frac{R_p L_1}{L_1 + \mu M}. \qquad (66)$$

Hence using Eq. (64), (65) and (66) the critical value of M may be obtained from the relation

$$\frac{R_pL_1}{L_1+\mu M}+\frac{L_1}{C}\frac{1}{R_L}=0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (67)$$

If we use the approximate relation $C=\frac{1}{\omega^2L_1}$, Eq. (67) yields the solution

$$-M=\frac{L_1}{\mu}\left(1+\frac{R_LR_p}{(\omega L_1)^2}\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (68)$$

which is the same as obtained in Eq. (63).

Experimental Check of Theory.—It is possible to measure the resistance of the oscillatory circuit of Fig. 144, in a Wheatstone bridge to see if its resistance actually does vary in the fashion predicted. The experiment shows that the resistance of the circuit actually does diminish in accordance with the theory; with a critical value of M the resistance becomes zero, and if the impressed frequency (at which the zero resistance has been measured) is the same as the natural frequency the circuit will start to generate oscillations. By inserting more resistance in series with it, however, the oscillations may be prevented and measurements made for values of M greater than the critical value.

In Fig. 147 is shown the experimental arrangement for testing the theory and the theoretical and experimental values of resistance for various values of condenser C, and, various degrees of coupling between the plate and grid circuits.[1] These curves show for example that if the condenser in the plate circuit has a value of 0.5095 microfarad the circuit will start to oscillate by itself when the mutual inductance between the grid and plate coil has been increased to 5.5 millihenries.

It will be noticed that there is remarkable agreement between experimental and theoretical values. The test, however, was carried out at a frequency of about 1200 cycles, far from the range of radio frequencies. This was done advisedly as certain errors are likely to occur when making the measurements at high frequencies and the theory is apparently disproved. To bring out this point there is given on the following page a set of data on the action of the circuit of Fig. 144 when using smaller inductance and capacity than was used for Fig. 147. The experiment also determined the values of M required to start oscillations, to check Eq. (58).

[1] For detailed analysis of this test see "An Analysis of Two Triode Circuits," by John H. Morecroft and Axel Jensen, Proc. I.R.E., Vol. 12, No. 5, Oct., 1924.

The inductances L_1 and L_2 were each 173 microhenries; R_p was 3600 ohms, and μ was 6.7. The values of C and R in the oscillating circuit were varied as shown in the accompanying table.

Capacity C	Wave Length	Resistance R	M in μh Experimental	M in μh Calculated from Eq. (58)
$13{,}400 \times 10^{-12}$		4 8	60.5	60.6
	2870	5.2	63.3	63.3
		5.6	66.7	66.2
		4.8	55.3	54.0
10,800	2580	5 8	59.2	59.6
		6.8	64.8	65.3
		6 8	53.3	54 3
7,770	2185	8 3	59 2	60.5
		9 8	64 8	66.7
		8 4	48 6	48 8
5,100	1770	11 4	56 6	57.0
		14 4	64 0	65.3
		12 0	31.6	42.4
2,570	1258	24 0	46 6	59 0
		36 0	61 8	75.6
		16 5	25 5	34.3
943	761	56 5	41 5	54 5
		96 5	59 4	74 8
		20 3	20 3	30 8
445	522	120 0	34 7	54.4
		220 0	49 7	78 5

It will be seen that for the higher value of capacity the agreement between theoretical and experimental values is very good, as it is in Fig. 147, but that large discrepancies occur for the smaller values of capacity. This is caused by the capacitive coupling of the elements of the triode; when the capacity in the oscillating circuit is low this triode capacity very materially assists the magnetic coupling in producing oscillations so that the experimentally determined value of M is about 60 per cent of the theoretical value. The missing 40 per cent of the theoretical is supplied (in effect at least) by the grid-plate capacity coupling of the triode itself.

This capacity coupling, due to the triode, is more effective at the higher frequencies, hence the increasing discrepancy between theory and experiment with increasing frequency (decreasing C). The theory is correct, but it has not considered all of the factors entering into the experiment.

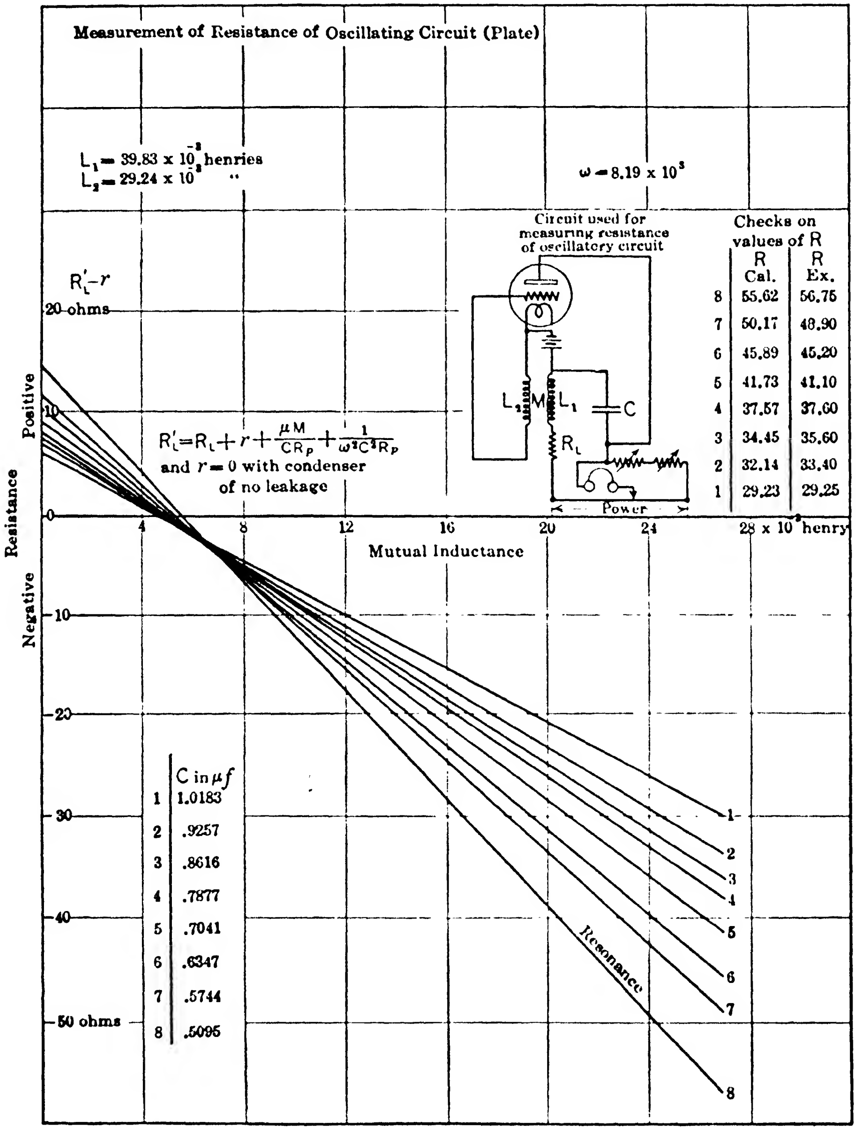


FIG. 147.—Showing measured values of resistance and its dependence upon the coupling between grid and plate circuits.

Phases of Voltages and Currents in the Steady State.—When the value of M, being increased from zero, slightly exceeds the critical value as determined by Eq. (59), oscillations start and build up to a certain steady value; how quickly they reach the steady state depends upon how much M exceeds its critical value, when the oscillations start and on the value of R_L. The steady state is reached when R_p has sufficiently increased (in average value) to reduce (59) to an equality.

When
$$R_L+\frac{1}{CR_p}(L_1+\mu M)=0,$$

Eq. (52) reduces to the form

$$L_1\frac{d^2i_1}{dt^2}+\frac{i_1}{C}\left(1+\frac{R_L}{R_p}\right)=0. \qquad (69)$$

The solution of this is $i_1=I_1 \sin \omega t$, in which

$$\omega=\sqrt{\frac{1+\frac{R_L}{R_p}}{L_1C}}.$$

The grid voltage $e_g=-M\frac{di_1}{dt}=-\omega MI_1 \cos \omega t.$

But as M is negative, this may be written

$$e_g=\omega\|M\|I_1 \sin\left(\omega t+\frac{\pi}{2}\right), \qquad (70)$$

in which $\|M\|$ represents the *magnitude* only of M, irrespective of its algebraic sign. This equation makes the grid voltage lead the current I_1 by 90°.

The plate voltage

$$e_p=-R_Li_1-L_1\frac{di_1}{dt}=-R_LI_1 \sin \omega t-\omega L_1I_1 \cos \omega t$$
$$=-I_1\sqrt{R_L^2+(\omega L_1)^2}\sin(\omega t+\phi),$$

in which
$$\tan \phi=\frac{\omega L_1}{R_L}.$$

In practically all radio coils $\omega L_1/R_L$ is so large that ϕ may be put equal to 90° without much error, so that,

$$e_p=-I_1\sqrt{R_L^2+(\omega L_1)^2}\sin(\omega t+\pi/2). \qquad (71)$$

From (70) and (71) it is evident that e_p and e_g are practically 180° out of phase, a condition we have previously shown necessary for oscillation.

The plate current is fixed by the condition

$$e_p = -R_L i_1 - L_1 \frac{di_1}{dt}.$$

And $R_p i_p = e_p + \mu e_g$, and $e_g = -M \frac{di_1}{dt}$, and $i_1 = I_1 \sin \omega t$.

We have

$$i_p = \frac{1}{R_p}(e_p + \mu e_g) = -\frac{1}{R_p}\left(R_L i_1 + L_1 \frac{di_1}{dt} + \mu e_g\right)$$

$$= -\frac{I_1}{R_p}(R_L \sin \omega t + \omega L_1 \cos \omega t + \mu\omega M \cos \omega t)$$

or

$$i_p = -\frac{I_1}{R_p}\sqrt{R_L^2 + \omega^2(L_1 + \mu M)^2} \sin(\omega t + \psi)$$

in which $\tan \psi = \dfrac{\omega(L_1 + \mu M)}{R_L}.$

As M is negative and μM is greater in absolute value than L_1, the angle ψ is nearly $-\pi/2$.

Then

$$i_p = -\frac{I_1}{R_p}\sqrt{R_L^2 + \omega^2(L_1 + \mu M)^2} \sin(\omega t - \pi/2)$$

$$= \frac{I_1}{R_p}\sqrt{R_L^2 + \omega^2(L_1 + \mu M)^2} \sin(\omega t + \pi/2). \quad . \quad . \quad (72)$$

It is therefore evident that the plate current leads the current in L_1 by practically 90°.

Amplitude of Oscillation in the Steady State.—The greatest current is generally obtained in a low resistance oscillating circuit with the least coupling that can be used to maintain oscillations. The lower the value of M the greater must I_1 be to maintain the required grid excitation and I_1 will vary with M about as shown by the full line curve in Fig. 148; if the value of M is decreased beyond the critical value oscillations will generally cease entirely. In certain circuits it is possible, however, to get maximum current with somewhat greater coupling than that at which oscillations start; in that case the curve between M and I_1 has the form indicated by the dotted line. The form depends somewhat upon the static characteristic of the triode; in Fig. 149 are shown two possible curves. The full-line curve corresponds with the full-line curve of Fig. 148, and the dotted curve of Fig. 149 corresponds to the dotted curve in Fig. 148. If the value I_{op} (no oscillations) is so adjusted that it is equal to one-half the saturation current, then the maximum possible value of i_p is I_{op}. But from Eq. (72) we have the maximum value of i_p given by the relation.

$$I_{mp}=\frac{I_{m1}}{R_p}\sqrt{R_L^2+\omega^2(L_1+\mu M)^2},$$

I_{mp} and I_{m1} being the maximum possible values of the two currents, i_p and i_1, and this value of I_{mp} must be equal to I_{op}. So we put

$$I_{op}=\frac{I_{m1}}{R_p}\sqrt{R_L^2+\omega^2(L_1+\mu M)^2},$$

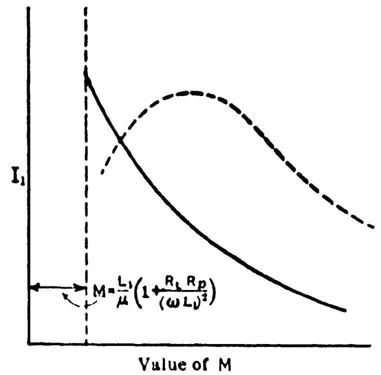


FIG. 148.—Showing possible relations between amplitude of oscillatory current and value of M in Fig. 144.

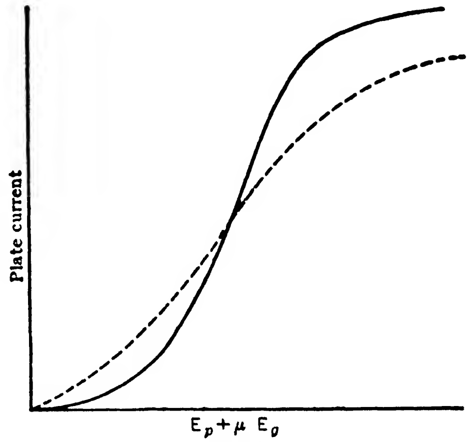


FIG. 149.—A tube having a plate current curve as shown by the full line will behave as indicated for the full line of Fig. 148; similarly for the dotted lines.

or

$$I_{m1}=\frac{I_{op}R_p}{\sqrt{R_L^2+\omega^2(L_1+\mu M)^2}},$$

and by substituting the condition

$$(L_1+\mu M)=-CR_pR_L$$

and assuming

$$\omega=\sqrt{\frac{1}{L_1C}}.$$

$$I_{m1}=\frac{I_{op}}{R_L}\frac{1}{\sqrt{\frac{C}{L_1}+\frac{1}{R_p^2}}},$$

which for the average tube is practically the same as,

$$I_{m1} = \frac{I_{op}}{R_L}\sqrt{\frac{L_1}{C}}. \quad \quad (73)$$

For this condition we conclude that increasing C must result in a decrease in I_1, but in trying out this relation experimentally we often find that I_1 may be increased by increasing C. There must be in the circuit some other limitation which must also be considered in using the relation of Eq. (73). Indeed this is at once evident, because Eq. (73) would lead to a value of I_1, approaching infinity as C is made to approach zero.

By examining the possible values of e_p we find the other limiting factor; it is evident that the maximum value of e_p is E_{op}, so that we have as another limiting condition on the amplitude of the oscillating current,

$$E_{op} = I_{m1}\sqrt{R_L^2 + (\omega L_1)^2}.$$

This follows from Eq. (71), putting the maximum value of e_p equal to E_{op}. We then have

$$I_{m1} = \frac{E_{op}}{\sqrt{R_L^2 + (\omega L_1)^2}},$$

which, for critical value of M, and assuming

$$\omega = \sqrt{\frac{1}{L_1 C}}$$

gives,

$$I_{m1} = \frac{E_{op}}{\sqrt{\frac{L_1}{C} + R_L^2}},$$

and as the value of R_L^2 is ordinarily small compared to L_1/C, we have as another limitation on the value of the oscillating current,

$$I_{m1} = E_{op}\sqrt{\frac{C}{L_1}}. \quad \quad (74)$$

Eqs. (73) and (74) then constitute two limits on the possible amplitude of I_1; whichever gives the lower value will determine the maximum value of I_1. The best condition makes the two limits the same which occurs when

$$\frac{E_{op}}{I_{op}} = \frac{1}{R_L}\frac{L_1}{C}. \quad \quad (75)$$

The symbols I_{op} and E_{op} have been used to indicate the limiting values of i_p and e_p, so that Eq. (75) is properly written, using maximum values of alternating voltage and current.

$$\frac{\text{maximum value } E_p}{\text{maximum value } I_p}=\frac{1}{R_L}\frac{L_1}{C}. \qquad (76)$$

But from Eq. (60), Chapter I,

$$\frac{1}{R_L}\frac{L_1}{C}=R_{A-B},$$

the external resistance of the plate circuit, and

$$\frac{\text{maximum value of } E_p}{\text{maximum value of } I_p},$$

is really R_p, the internal resistance of the tube.

The foregoing analysis therefore yields the same result as obtained on p. 572, namely, for maximum output the external resistance of the tube circuit should be equal to the tube resistance itself.

By comparing Eqs. (76) and (75), it is seen that the resistance of the tube for maximum output is equal to E_{op}/I_{op}, which we previously called R_{op}, the c.c. resistance of the tube. We also showed that R_p, the a.c. resistance, was generally about one-half the c.c. resistance R_{op} Eq. (12) and Fig. (61). This apparent discrepancy arises from the fact that R_p is really a variable quantity, depending for its value upon the amount of change in the plate current. The discussion of R_p on pp. 518 et seq. and the measurements recorded in Fig. 128 had to do with R_p for very small variations in plate current, and in such a case R_{op} is about twice as great as R_p.

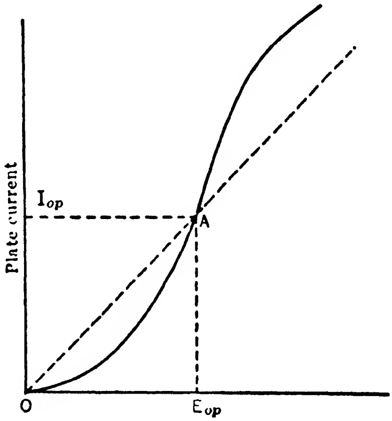


FIG. 150.—If the R_p of a tube is to be constant the relation between I_p and $(E_p+\mu E_g)$ must be a straight line as indicated in the dotted graph; actually the solid line curve gives the plate current, hence it is evident that R_p must vary with the magnitude of fluctuation of $(E_p+\mu E_g)$.

For the conditions obtaining when Eqs. (75) and (76) are applicable, the plate current is supposed to vary from zero to 2 I_{op}, and furthermore the relation between I_p and $(E_p+\mu E_g)$ is supposed to be linear; for such conditions $R_p=R_{op}=E_{op}/I_{op}$. The difference in R_p with weak excitation and strong excitation is indicated in Fig. 150; the full-line curve represents the actual relation between I_p and E_p, when there is no resistance in series with the plate circuit and the dotted curve shows the assumed relation on the basis of Eq. (49). The dotted-line curve gives for R_p (at point A)

$$\frac{\Delta E_p}{\Delta I_p}=\frac{E_{op}}{I_{op}},$$

whereas the full-line curve gives for R_p at the point A a value of $\Delta E_p/\Delta I_p$, about half as great as E_{op}/I_{op}.

Of course, it is not possible to excite a tube to the limits set by Eqs. (73) and (75), so R_p actually never increases to the value

$$R_p=\frac{E_{op}}{I_{op}};$$

as the intensity of the oscillations varies; the value of R_p for the ordinary tube will undergo changes about as shown in Fig. 151.

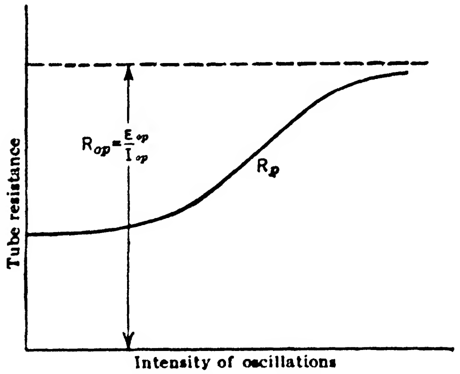


FIG. 151.—The resistance of the plate circuit of a tube varies with the fluctuation of $(E_p+\mu E_g)$ about as shown here; intense oscillations require large fluctuations in $(E_p+\mu E_g)$.

Stability of Oscillations.—In the average low-resistance circuit the value of the oscillating current is greatest when the coupling is as weak as can be permitted and still maintain oscillations. For this condition, however, the stability of the circuit is very poor; the slightest decrease in either I_f or E_b is likely to stop the oscillations. Also for this condition it is necessary to readjust the coupling for every change in the oscillating circuit; if either R_L, L_1, or C is increased the oscillation will cease. To make this circuit stable it is necessary to have the coupling at a setting considerably in excess of its critical value, perhaps twice as much. This of course will diminish somewhat the magnitude of the oscillating current, but the increased reliability of the generating action of the tube generally compensates for this.

It many times happens that the critical value of coupling for starting oscillations is greatly different from the critical value to stop oscillations; in a certain circuit this critical value of coupling for starting the oscillations was 17 per cent, whereas it could then be decreased to 12 per cent before the oscillations ceased. This is due to the variable value of R_p, as brought out in a previous paragraph; when the oscillations start their amplitude is necessarily small and R_p is determined by the slope at the value of I_p, as shown in Fig. 152 at A. After the oscillations are started, the plate current fluctuates between zero and BC, and the average resistance between these limits is less than the value of R_p at A. The plate current for such oscillations would be very complex and so the behavior of the circuit could not be predicted from the analyses previously given, which have assumed sine waves of current.

Starting and Stopping Oscillations.—It is sometimes necessary to give a circuit some sort of a shock to start oscillations; if normal filament current and plate voltage are impressed and then the coupling gradually increased, it will be found that M may greatly exceed its critical value without causing the tube to oscillate. If, however, the plate circuit is opened and then closed, thus giving a pulse to the circuit, oscillations will start.

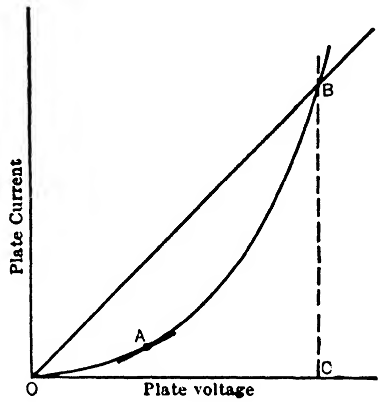


Fig. 152.—When a tube starts to oscillate its resistance is fixed by the slope of the I_p-E_p curve, and this resistance may be very different from the value when intense oscillations are occurring.

In case a tube is used for generating power in a transmitting station, the oscillations must be continually started and stopped, as the signals are sent out by the key. The vacuum-tube generator permits this operation to be carried out readily; a small hand key properly introduced in the grid circuit may control kilowatts of power with imperceptible sparking. Probably the most convenient scheme for "keying" a tube generator is that shown in Fig. 153; with the key open the grid is forced to such a negative potential by the battery E_c (which can be small dry cells) that the circuit stops oscillating and when the key is closed the coil L_2 is connected to ground which is its normal connection for oscillation. Of course when the key is closed, the battery E_c is short-circuited through the resistance R, but this will do them no harm if R is chosen sufficiently high. Condenser C_0 shunts the contact points of the key; this condenser must be of sufficiently small capacity, otherwise the set will continue to oscillate for an appreciable time after the key is opened; about 0.1 microfarad seems satisfactory when R is 20,000 ohms.

Effect of Oscillation on the Grid and Plate Currents.—In such a circuit as that shown in Fig. 144, the grid current is very nearly zero until oscillations start; when the tube is oscillating the grid becomes positive during part of the cycle and so takes current. The value of the grid current is larger than would be at first supposed, because, although the grid potential does not reach high positive values, the plate potential is low at the time the grid is positive.

In a small power tube designed for $E_b = 300$ and $E_{op} = 0.04$ ampere the average value of I_g when the tube is adjusted for maximum value of power output is about 0.003 ampere. The maximum value of the grid current, when its average value is 0.003, is probably from 0.02 to 0.05 ampere. In a large power tube, excited for maximum power output, I_g may be considerably greater than the values given above.

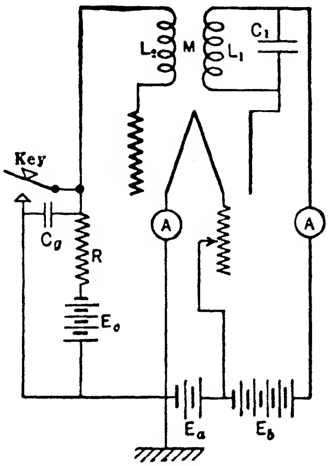


FIG. 153.—This diagram shows a convenient method of "keying" a tube circuit; if the proper values of R and E_c are chosen a small hand key may control kilowatts of power with imperceptible sparking.

If the plate current I_{op} has been adjusted equal to half the saturation current, for the values of I_f and E_{op} used, a c.c. ammeter will indicate no change in the value of the plate current when oscillations start. In general, however, there will be a change; when oscillations start the average plate current will generally increase if the circuit is such that no condenser is used in series with the grid and will decrease if such a condenser is used. Conditions may occur in which this general statement is not true.

Adjustments to Give Maximum Output of Tube.—With a circuit arranged as in Fig. 144, there are two adjustments to carry out before the tube will give its maximum output; the grid must have the proper excitation and the plate circuit resistance must equal the tube resistance. The circuit of Fig. 144 is reproduced, with slight modification, in Fig. 154. The oscillating circuit L_1, R_L, C, is many times an antenna, with loading coil, so it is evident that R_L itself is not adjustable, yet the resistance between points A and O must be made equal to the tube resistance.

The plate circuit inductance is made with taps as indicated in Fig. 154; point B is adjusted to give the right frequency to the oscillating circuit, and then point A is adjusted to give the plate circuit the right resistance. Neglecting the effect of the plate current compared to the oscillating current (an ordinary radio set makes I_p equal to about 1/20 of I), we have

$$R_{O-A} = R_L\left(\frac{\omega L_p}{R_L}\right)^2 = \frac{\omega^2 L_p{}^2}{R_L} = \frac{L_p{}^2}{R_L L_1 C}. \quad . \quad . \quad . \quad . \quad (77)$$

If R_{O-A} is to be equal to $R_p = E_{op}/I_{op}$ (for conditions of maximum power) we so adjust tap A that

$$\frac{E_{op}}{I_{op}} = \frac{L_p^2}{R_L L_1 C},$$

or

$$L_p = \sqrt{\frac{E_{op} R_L L_1 C}{I_{op}}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (78)$$

This required value of L_p may be either greater or less than L_1.

It may well be that, in circuits like that shown in Fig. 154, capacities in the tube itself become important. Or it may be that instead of oscillating in the CR_L circuit the oscillations occur in the circuit made up of the capacity between the grid and plate, together with the inductance L_2 and inductance between points O-A. Illustrating the effect of tube capacity we give the case of a circuit having two turns from O to B, and 14 turns from O to A. The actual value of capacity used in C was $450\mu\mu f$ but from the measured wave length and known inductance, the capacity in the oscillating circuit was $1000\mu\mu f$. The plate-filament capacity was taking part in fixing the frequency of the oscillations and was much magnified by the step-up ratio of the auto-transformer made up of coils OB and OA.

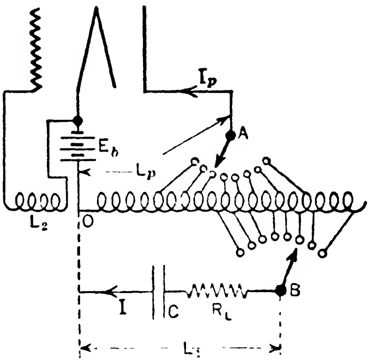


FIG. 154.—To make the circuit shown in Fig. 144 useful, the inductance in the oscillatory circuit must be fitted with two sets of taps as indicated here; the mutual induction between the two coils L_2 and L_p must also be adjustable. For short waves (say 150 meters) maximum current will probably be obtained with no mutual induction at all.

A general precaution to take in circuits of this kind is to make the grid coil, L_2, of as few turns as possible. Otherwise the grid circuit may be the one which sets the frequency of the oscillations, using the capacity of the tube itself.

In a certain radio circuit $R_L = 3$ ohms, $C = 4 \times 10^{-10}$ farad, $L_1 = 150\mu h$, $E_{op} = 300$ volts, $I_{op} = 0.03$ ampere. Using Eq. (78), the required value of L_p proves to be $42\mu h$. The tap A would therefore be made between O and tap B.

The current in the oscillating circuit can be calculated from the relation, $I = \omega CE$, where E is the effective voltage across OB, which is equal to

$(E_{op}/\sqrt{2}) \times L_1/L_p$.[1] Using these relations and also remembering that $\omega = \sqrt{1/L_1C}$, we get,

$$I = \frac{E_{op}}{L_p}\sqrt{\frac{L_1C}{2}} \text{ and so, using Eq. (78) } = \sqrt{\frac{E_{op}I_{op}}{2R_L}}. \quad . \quad . \quad (79)$$

Substituting the values above gives a value for I of 1.23 amperes. Actually 1.05 amperes was the maximum obtainable from the circuit.

The resistance R_L was then increased to 50 ohms, and it was found experimentally that tap A was outside of tap B for maximum output. By calculation, Eq. (78), we find the proper value for $L_p = 171\mu h$. The oscillatory current, from Eq. (79), should be 0.31 ampere, whereas only 0.26 was actually obtained.

After the right position for tap A has been found the coupling between L_2 and L_1 is reduced until the critical value of M is nearly reached, and then a slight readjustment of tap A may be necessary. It will be found that varying M and the position of tap A will have only minor effects on the frequency being generated by the tube.

The value of L_2 should be kept as low as possible; if it should happen that the natural period of L_2, combined with the capacity of the input circuit of the tube is about the same as the period of the L_1C circuit, trouble may be experienced in making the tube oscillate because of the unexpected phase of the voltage impressed on the grid; the voltage changes its phase nearly 180° as the natural frequency of the L_1C circuit is made to pass through the natural frequency of the grid circuit.

Oscillations at Other than the Desired Frequency.—It may happen that if the grid circuit has its natural frequency in the neighborhood of the frequency of the L_1C circuit, the tube will generate power of the frequency of the grid circuit, instead of that of the L_1C circuit.

To remedy this trouble the grid circuit is sometimes tuned to nearly the same frequency as the L_1C circuit. Another method of ensuring the desired frequency of oscillations is to couple the grid, not to the plate coil, but to a coil in the L_1C circuit, which is so placed that no current flows in it unless the main circuit is oscillating. This idea is depicted in Fig. 155. Coil L_1 will carry no current unless the main circuit, including L_1 and C, is oscillating.

The difficulty occurs principally when the resistance of the main

[1] This relation is approximate only, because of the mutual induction between the inductance between points O–A (Fig. 154) and that between points A–B. If the coil is short, so that the turns are all close together, the effect of this mutual induction will be considerable and the relation given above is more accurately written $(E_{op}/\sqrt{2}) \times N_1/N_2$ where

$$N_1 = \text{number of turns between points } O\text{–}B$$

and

$$N_2 = \text{number of turns between points } O\text{–}A.$$

oscillating circuit is high so that the current I is relatively small; to sufficiently excite the grid in this case requires a comparatively large value of L_2, which of course lowers the natural period of the grid circuit.

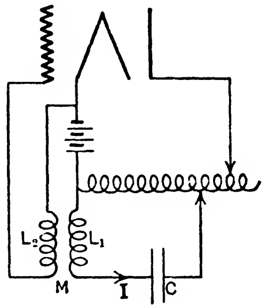


FIG. 155.—To prevent spurious oscillations it is advisable to couple the grid circuit to some part of the main oscillatory circuit through which the plate current does not flow; the grid is then not so likely to be excited unless the main circuit is oscillating.

Oscillating Current Comparable in Value with Plate Current.—When the resistance of the oscillating circuit gets very high the oscillating current I may decrease to such an extent that it is of about the same value as I_p or even less. In this case it is not easy to produce oscillations, because the e.m.f. for the excitation on the grid tends to get the wrong phase. The scheme of Fig. 155 may not work because L_1, which must be small compared to L_p (because of the high value of R_L), may not induce a sufficient voltage in L_2, so resort must be had to coupling L_2 with L_p. Now the oscillatory current in L_p is ordinarily 90° out of phase (nearly) with I_p, and such condition results in a correct phase for the voltage E_g. But if now I_p is comparable with I, the actual current in L_p (which produces the magnetic field affecting L_2) tends to come into phase with I_p, that is, shift its phase 90° from its normal value. But such a shift in phase will result in such a phase for E_g that oscillations cannot be

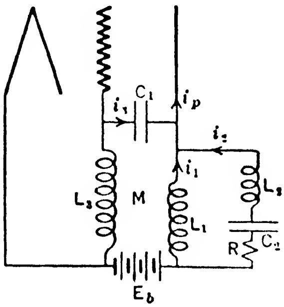


FIG. 156.—Another oscillatory circuit in which the grid is excited by inductive coupling between L_1 and L_3 as well as by the capacitive coupling produced by C_1.

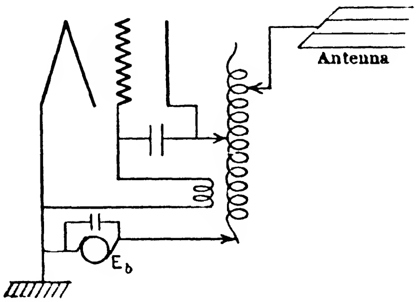


FIG. 157.—Showing how the circuit of Fig. 156 is applied to an actual radio circuit.

maintained; in fact a comparatively small shift will materially cut down this possible power output of the tube.

For a condition of this sort it is better to use separate excitation for the tube, instead of trying to make it self-exciting. Another tube circuit, having a low resistance, is self-excited at the desired frequency, and from this circuit a suitable voltage may be obtained (either directly or magnetically) for excitation of the tube furnishing power to the high-resistance circuit.

Coupling between Grid and Plate Circuit by Capacity.—In the foregoing discussions of a self-excited tube the voltage for excitation of the grid has been obtained by a magnetic coupling with the oscillating circuit, but it is of course possible to use electrostatic coupling, or even a combination of both. Such a circuit is shown in Fig. 156; in order to make the discussion more general only a part of the inductance in the oscillating circuit is included in the plate circuit. The extra inductance L_2, in combination with C_2 and R, represents an antenna, thus making the circuit the direct equivalent of the actual circuit shown in Fig. 157. From Fig. 156 using directions of current shown in the diagram, we have,

$$i_p = i_1 + i_2 + i_3. \qquad (80)$$

$$-L_1 \frac{di_1}{dt} - M \frac{di_3}{dt} = -L_2 \frac{di_2}{dt} - Ri_2 + e_{c2} = -L_3 \frac{di_3}{dt} - M \frac{di_1}{dt} + e_{c1}. \qquad (81)$$

Also we know that e_{c1} and e_{c2} are fixed by the relation $-C_2 \frac{de_{c2}}{dt} = i_2$, and $-C_1 \frac{de_{c1}}{dt} = i_3$. By the use of these relations, and deriving Eq. (81), we get the equations,

$$L_1 \frac{d^2 i_1}{dt^2} + M \frac{d^2 i_3}{dt^2} = L_2 \frac{d^2 i_2}{dt^2} + R \frac{di_2}{dt} + \frac{i_2}{C_2} \qquad (82)$$

$$(L_1 - M) \frac{d^2 i_1}{dt^2} = (L_3 - M) \frac{d^2 i_3}{dt^2} + \frac{i_3}{C_1}. \qquad (83)$$

We can write

$$R_p i_p = e_p + \mu e_g; \quad e_p = -L_1 \frac{di_1}{dt} - M \frac{di_3}{dt}; \quad e_g = -L_3 \frac{di_3}{dt} - M \frac{di_1}{dt}.$$

Using these conditions, and using Eq. (80), we get

$$R_p(i_1 + i_2 + i_3) + (L_1 + \mu M) \frac{di_1}{dt} + (\mu L_3 + M) \frac{di_3}{dt} = 0. \qquad (84)$$

Eqs. (82), (83) and (84) permit the precise determination of i_1, i_2, and i_3, but it is evident that the solution would be tedious and the solution can be easily guessed. If oscillations occur at all they will be sinusoidal and as they are all supplied with power from the same source (the plate circuit) we can write

$$i_1 = I_1 \sin \omega t, \quad i_2 = I_2 \sin (\omega t + \phi_2), \quad i_3 = I_3 \sin (\omega t + \phi_3).$$

By deriving these expressions and substituting in Eqs. (82), (83) and (84), and for each equation thus obtained, equating the coefficients of cos ωt and sin ωt, we find

$$R_p(I_1+I_2\cos\phi_2+I_3\cos\phi_3)-\omega(\mu L_3+M)I_3\sin\phi_3=0. \quad . \qquad (85)$$

$$R_p(I_2\sin\phi_2+I_3\sin\phi_3)+\omega(L_1+\mu M)I_1+\omega(\mu L_3+M)I_3\cos\phi_3=0. \qquad (86)$$

$$\omega LI_1+\omega MI_3\cos\phi_3=\left(\omega L_2-\frac{1}{\omega C_2}\right)I_2\cos\phi_2+I_2R\sin\phi_2. \quad . \qquad (87)$$

$$\omega MI_3\sin\phi_3=\left(\omega L_2-\frac{1}{\omega C_2}\right)I_2\sin\phi_2-I_2R\cos\phi_2. \quad . \quad . \quad . \qquad (88)$$

$$\omega(L_1-M)I_1=\left\{\omega(L_3-M)-\frac{1}{\omega C_1}\right\}I_3\cos\phi_3. \quad . \quad . \quad . \qquad (89)$$

$$\left\{\omega(L_3-M)-\frac{1}{\omega C_1}\right\}I_3\sin\phi_3=0. \quad . \quad . \quad . \quad . \qquad (90)$$

Eq. (90) shows that unless $\omega(L_3-M)-\frac{1}{\omega C_1}=0$, sin ϕ_3 must be equal to zero, which means that e_3 and i_3 are either in phase or 180° out of phase. In case $\omega(L_3-M)-\frac{1}{\omega C_1}=0$, we have resonance in the L_3-C_1 circuit.

Using Eq. (89) and (90) to get values of $I_3\sin\phi_3$ and $I_3\cos\phi_3$, then using Eqs. (85) and (86) to get values of $I_2\sin\phi_2$ and $I_2\cos\phi_2$ and putting these values in Eqs. (87) and (88), we get the two equations

$$\left\{\omega(L_1+L_2)-\frac{1}{\omega C_2}\right\}+\left\{\omega(L_2+M)-\frac{1}{\omega C_2}\right\}\frac{\omega(L_1-M)}{\omega(L_3-M)-\frac{1}{\omega C_1}}$$
$$+\frac{R}{R_p}\left\{\omega(L_1+\mu M)+\omega^2\frac{(L_1-M)(\mu L_3+M)}{\omega(L_3-M)-\frac{1}{\omega C_1}}\right\}=0, \quad . \quad . \qquad (91)$$

and

$$R-\frac{\omega L_2-\frac{1}{\omega C_2}}{R_p\left[1+\frac{(L_1-M)}{\omega(L_3-M)-\frac{1}{\omega C_1}}\right]}\left\{\omega(L_1+\mu M)\right.$$
$$\left.+\omega^2\frac{(L_1-M)(\mu L_3+M)}{\omega(L_3-M)-\frac{1}{\omega C_1}}\right\}=0. \quad . \quad . \qquad (92)$$

The frequency might be calculated from Eq. (91), and this frequency carried into Eq. (92) would permit the calculation of the critical coupling for oscillations. From inspection of Fig. 156 it is evident there will be two possible frequencies and of course each of these must be used in solving Eq. (92). This general solution is lengthy, so we will investigate only two of the more important cases.

In case $C_1=0$ and $L_2=0$, the circuit degenerates into that of Fig. 144 and so our general Eqs. (91) and (92) should reduce to the simpler forms obtained for this case. Eq. (91) becomes (if we put $C_1=L_2=0$)

$$\omega L_1-\frac{1}{\omega C_2}+\frac{R}{R_p}\omega(L_1+\mu M)=0,$$

which we previously obtained, and if R/R_p is small enough to be negligible,

$$\omega=\sqrt{\frac{1}{L_1C_2}},$$

and if this value of ω is substituted in Eq. (92), in addition to the condition that $C_1=L_2=0$, we find as the condition for oscillation,

$$R+\frac{1}{R_pC_2}(L_1+\mu M)=0,$$

which we have already obtained from the circuit of Fig. 144.

In case $M=0$ and $L_2=0$, Eq. (91) becomes

$$\left(\omega L_1-\frac{1}{\omega C_2}\right)-\frac{L_1}{C_2}\,\frac{1}{\omega L_3-\dfrac{1}{\omega C_1}}+\frac{R}{R_p}\left(\omega L_1+\frac{\omega^2\mu L_3L_1}{\omega L_3-\dfrac{1}{\omega C_1}}\right)=0,$$

and if again R/R_p is negligibly small, we find,

$$\frac{1}{\omega L_1}-\omega C_2+\frac{1}{\omega L_3-\dfrac{1}{\omega C_1}}=0. \quad \text{. . . . . (93)}$$

This is evidently the condition for zero reactive current in the three-branched plate circuit, one branch having L_1, another having C_2, and the third having L_3 and C_1 in series. Eq. (93) may be put into the form

$$\frac{1}{\omega^4}-[L_1C_2+(L_3+L_1)C_1]\frac{1}{\omega^2}+L_1C_2L_3C_1=0. \quad \text{. . . . (94)}$$

If we put $[L_1C_2+(L_3+L_1)C_1]=a$, and $L_1C_2L_3C_1=b$ we can write the two positive roots of this equation,

$$\frac{1}{\omega}=\sqrt{\frac{a}{2}\pm\sqrt{\frac{a^2}{4}-b}}$$

Of these two roots for ω one is greater than $\sqrt{\dfrac{1}{L_3C_1}}$ and the other is less than $\sqrt{\dfrac{1}{L_3C_1}}$. We shall show that the only possible oscillation is the lower one of the two.

If we substitute $M=0$ and $L_2=0$ in Eq. (92), we find that the critical conditions for maintaining oscillations is given by,

$$R+\frac{L_1}{R_pC_2\left(1+\dfrac{\omega L_1}{\omega L_3-\dfrac{1}{\omega C_1}}\right)}\left(1+\frac{\omega\mu L_3}{\omega L_3-\dfrac{1}{\omega C_1}}\right)=0, \quad . \quad . \quad . \quad (95)$$

the condition for oscillation making the left-hand member less than zero.

The condition for oscillations is then determined by the inequality

$$\frac{\omega L_1}{\omega L_3-\dfrac{1}{\omega C_1}}(RR_pC_2+\mu L_3)+L_1+RR_pC_2<0. \quad . \quad . \quad . \quad . \quad (96)$$

This can evidently be satisfied only by having

$$\omega L_3-\frac{1}{\omega C_1}<0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (97)$$

which shows that the circuit cannot sustain oscillations at a frequency which makes ωL_3 greater than $1/\omega C_1$. This bears out the prediction made above that of the two roots of Eq. (94), only that one having a value less than $\sqrt{1/L_3C_1}$ is a possible frequency for the oscillations because the conditional inequality (97) may be written

$$\omega<\sqrt{\frac{1}{L_3C_1}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (98)$$

Equation (96) also serves to further limit C_1, because from it we get the relation,

$$C_1>\frac{1}{\omega^2}\frac{L_1+R_pRC_2}{L_1L_3(\mu+1)+R_pRC_2(L_1+L_3)}. \quad . \quad . \quad . \quad . \quad (99)$$

So we have C_1 fixed by the double condition

$$\frac{1}{\omega^2L_3}>C_1>\frac{1}{\omega^2}\frac{L_1+R_pRC_2}{L_1L_3(\mu+1)+R_pRC_2(L_1+L_3)},$$

ω being fixed as the smaller of the roots of the equation given on p. 612. The relation given in Eq. (99) shows that if $\mu L_3>L_1$ (which will generally be the case), $\delta C_1/\delta R$ is positive, so that as the resistance of the oscillating circuit is increased, the value of C_1 must also be increased to maintain the oscillations. By similar reasoning, we see that C_1 must be increased as the frequency of the oscillations is diminished.

If we consider both magnetic and static coupling as given in Fig. 156, we can much simplify the general equations obtained—(91) and (92)—by supposing L_2 absent and R/R_p negligibly small. Eq. (91) then becomes,

$$\omega L_1 - \frac{1}{\omega C_2} + \left(\omega M - \frac{1}{\omega C_2}\right)\frac{\omega(L_1 - M)}{\omega(L_3 - M) - \dfrac{1}{\omega C_1}} = 0, \quad . \quad . \quad (100)$$

and Eq. (92) becomes

$$R + \frac{(\mu+1)\omega(L_1L_3 - M^2) - \dfrac{L_1 + \mu M}{\omega C_1}}{R_pC_2\left\{\omega(L_1 + L_3 - 2M) - \dfrac{1}{\omega C_1}\right\}} = 0. \quad . \quad . \quad (101)$$

The capacity coupling serves to increase the magnetic coupling if M is negative and if $\omega(L_3 - M) - 1/\omega C_1 < 0$. Even if M is *positive* the condition for oscillations may be still maintained by using sufficient capacity coupling.

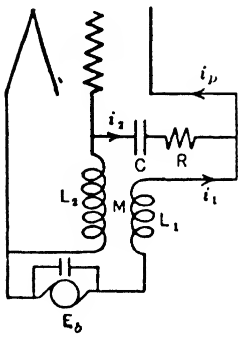


FIG. 158.—This circuit is similar to that of Fig. 156, but simplified by eliminating the dummy antenna circuit.

It is to be noted that even if no actual condenser C_1 is used in the circuit, there is always such a capacity present in the tube itself, due to capacity between the actual grid and plate, as well as that of the lead-in wires connecting to them. At very high frequencies this internal tube capacity may very seriously affect the behavior of the tube; in triodes designed for oscillating at very high frequencies the grid lead is always brought out of the tube as far from the plate lead as possible, with the idea of keeping the grid-plate capacity low.

Another circuit which may be used is shown in Fig. 158. For this case, we have

$$e_p = -L_1\frac{di_1}{dt} - M\frac{di_2}{dt}.$$

$$e_g = -L_2\frac{di_2}{dt} - M\frac{di_1}{dt},$$

and as

$$R_pi_p = e_p + \mu e_g,$$

we have the relation

$$R_p(i_1+i_2)+(L_1+\mu M)\frac{di_1}{dt}+(\mu L_2+M)\frac{di_2}{dt}=0. \quad . \quad . \quad (102)$$

When the reactance across the machine or battery furnishing the plate voltage is negligible (it should always be made so by shunting with a large capacity, if necessary), we have

$$-L\frac{di_1}{dt}-M\frac{di_2}{dt}=-L_2\frac{di_2}{dt}-M\frac{di_1}{dt}-Ri_2+e_c, \quad . \quad . \quad . \quad (103)$$

and as

$$i_2=-C\frac{de_c}{dt},$$

we can write Eq. (103) in the form

$$(L_1-M)\frac{d^2i_1}{dt^2}=(L_2-M)\frac{d^2i_2}{dt^2}+R\frac{di_2}{dt}+\frac{i_2}{C}. \quad . \quad . \quad . \quad (104)$$

From this i_1 might be eliminated and so enable a solution of i_2 to be obtained. Instead of this formal procedure, we guess at the solution and put,

$$i_1=I_1\sin\omega t \quad \text{and} \quad i_2=I_2\sin(\omega t+\phi).$$

Using these two values and substituting in Eq. (104), and using Eq. (102), we get

$$\omega^2\frac{R}{R_p}(\mu L_2+M)(L_1-M)-\omega(L_1-M)\left[\omega(L_2-M)-\frac{1}{\omega C}\right]$$
$$-R^2-\left[\omega(L_2-M)-\frac{1}{\omega C}\right]^2=0, \quad (105)$$

and

$$R_pR(L_1-M)+\omega(L_1-M)(\mu L_2+M)\left[\omega(L_2-M)-\frac{1}{\omega C}\right]$$
$$+(L_1+\mu M)\left[R^2+\left(\omega(L_2-M)-\frac{1}{\omega C}\right)^2\right]=0. \quad (106)$$

If in Eq. (105) we neglect the terms R/R_p and R^2, we get for the natural period of the circuit.

$$\omega=\frac{1}{\sqrt{(L_1+L_2-2M)C}}. \quad . \quad . \quad . \quad . \quad . \quad (107)$$

And using this value of ω in Eq. (106), which determines the critical condition for oscillation,

$$R[R_p(L_1-M)+R(L_1+\mu M)]<\frac{[\mu L_2-L_1-(\mu-1)M](L_1-M)^2}{(L_1+L_2-2M)C}. \quad (108)$$

This conditional inequality requires

$$\mu L_2 > L_1 + (\mu - 1)M.$$

If we suppose there is no magnetic coupling $M=0$ and the frequency of oscillation becomes

$$\omega = \frac{1}{\sqrt{(L_1+L_2)C}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (109)$$

and the condition for oscillation

$$R(R_p+R) < \frac{L_1(\mu L_2 - L_1)}{(L_1+L_2)C}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (110)$$

Oscillating Circuit in the Grid Circuit.—When a three-electrode tube is used as detector for continuous waves, it is necessary to have an additional tube for producing the local oscillations or else to use the detecting tube itself to generate the local oscillations. While any arrangement which makes the tube oscillate will serve for the purpose, the one which is probably used more frequently than any other is shown in Fig. 159; the tuned circuit is now associated with the grid, being coupled to the plate circuit by a coil in the plate circuit, L_1. This is generally called the "tickler" coil.

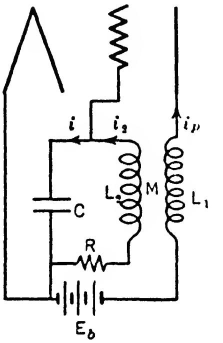


FIG. 159.—This is the circuit generally used when an oscillating tube is used to receive a continuous-wave signal; the oscillatory circuit is here associated with the grid.

If we make the same assumption as has been made for all the other circuits so far considered, namely, the grid takes no current, then $i_2 = i$ and the equations of the circuits are,

$$e_p = -M\frac{di_2}{dt} - L_1\frac{di_p}{dt},$$

and

$$e_g = -L_2\frac{di_2}{dt} - Ri_2 - M\frac{di_p}{dt}.$$

Also

$$R_p i_p = e_p + \mu e_g = -(M+\mu L_2)\frac{di_2}{dt} - \mu R i_2 - (L_1+\mu M)\frac{di_p}{dt}$$

or

$$R_p i_p + \mu R i_2 + (\mu L_2 + M)\frac{di_2}{dt} + (L_1 + \mu M)\frac{di_p}{dt} = 0. \quad . \quad . \quad (111)$$

Now

$$-L_2\frac{di_2}{dt} - M\frac{di_p}{dt} - Ri_2 = e_c \text{ and } i = -C\frac{de_c}{dt}.$$

By taking the derivative of the above equation and substituting value of i, then eliminating i_p between the resulting equation and Eq. (111) we get (substituting the symbol i for both i and i_2, which are the same)

$$\frac{1}{R_p}(L_1L_2-M^2)\frac{d^3i}{dt^3}+\left(L_2+L_1\frac{R}{R_p}\right)\frac{d^2i}{dt^2}+\left(R+\frac{L_1+\mu M}{CR_p}\right)\frac{di}{dt}+\frac{i}{C}=0. \quad (112)$$

Guessing the solution to be $i=I\sin\omega t$, substituting the proper derivatives in Eq. (112), we get for the period of oscillation

$$\omega=\frac{1}{\sqrt{C\left(L_2+L_1\frac{R}{R_p}\right)}}, \quad \ldots\ldots \quad (113)$$

which is practically the same as $1/\sqrt{CL_2}$.

For the limiting condition of oscillations, we find,

$$R+\frac{1}{R_pC}\left(L_1+\mu M-\frac{L_1L_2-M^2}{L_2+L_1\frac{R}{R_p}}\right)=0. \quad \ldots \quad (114)$$

Eq. (114) can be written in the form

$$R+\frac{1}{C\left(L_2+L_1\frac{R}{R_p}\right)}\left\{\frac{1}{R_p}\left[\left(L_2+L_1\frac{R}{R_p}\right)(L_1+\mu M)-L_1L_2+M^2\right]\right\}=0,$$

from which, using (113) and neglecting terms involving R/R_p, we get

$$R+\frac{\omega^2M(\mu L_2+M)}{R_p}=0, \quad \ldots\ldots \quad (115)$$

and this can be satisfied only if M is negative and its absolute value is greater than μL_2. The condition imposed by Eq. (115) will be satisfied if M is negative and its absolute value lies between the two roots of Eq. (115). So the absolute value of M is limited by the relation

$$\frac{\mu L_2}{2}-\sqrt{\left(\frac{\mu L_2}{2}\right)^2-\frac{RR_p}{\omega^2}}<M<\frac{\mu L_2}{2}+\sqrt{\left(\frac{\mu L_2}{2}\right)^2-\frac{RR_p}{\omega^2}}. \quad . \quad (116)$$

The condition is evidently different from that existing when the oscillating circuit was in series with the plate. In that case if M exceeded its critical value the value of the oscillating current was reduced, but there was no upper limit for the permissible value of M. With the oscillating circuit in series with the grid, however, the oscillations will cease if the absolute value of M exceeds a certain critical value, which is nearly μL_2.

In the paper previously referred to[1] the circuit of Fig. 159 was analyzed more carefully than is done here, and it is shown there that the resistance of the grid circuit as measured in a bridge should be

$$R' = R_L + r + \frac{R_p}{R_p^2 + \overline{\omega L_1}^2}\left(\overline{\omega M}^2 + \frac{\mu M}{C}\right), \quad . \quad . \quad . \quad (117)$$

in which r is the equivalent series resistance of the condenser produced by its leakage. The third term of the above equation for R' is the resistance introduced by the triode as a result of the coupling between grid and plate coils, M. Now as M may be negative this last term may evidently be negative. It is seen, however, that as M occurs in the term $\overline{\omega M}^2$ as well as $\mu M/C$, and as M^2 must always be positive, and increase faster than M itself, the third term must again become positive (after being negative) for the larger values of M.

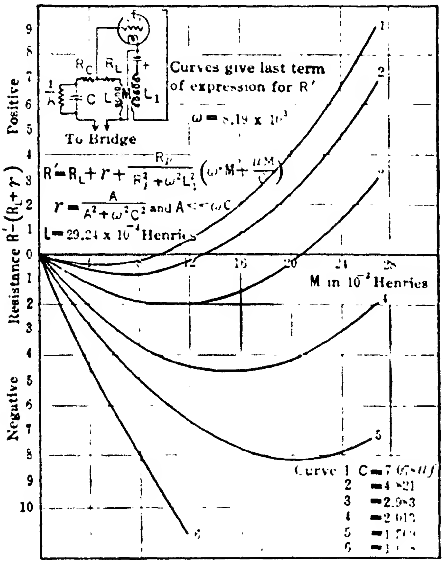


FIG. 160.—Measured values of the resistance of the oscillatory circuit showing its dependence upon the capacity in the circuit and the value of M. In the test R_c was zero and the value of L_1 was 39.83 millihenries.

In Figs. (160) and (161) are shown the measured values of

$$\frac{R_p}{R_p^2 + \overline{\omega L_1}^2}\left(\overline{\omega M}^2 + \frac{\mu M}{C}\right)$$

for various values of M and C. All of these curves have a certain negative region and then become positive again. However, for the smaller values of C (Fig. 161) the value of M required is so large that the curves seem to be going increasingly negative, without returning above the axis. If M could be obtained large enough, however, they all would eventually curve back to the positive region.

Circuits of Very High Frequency.[2]—Vacuum-tube circuits will generate

[1] "An Analysis of Two Triode Circuits," by Morecroft and Jensen, Proc. I.R.E., Vol. 12, No. 5, Oct., 1924.

[2] Many other circuits than the few here analyzed have been designed and used. The reader is referred to an article by L. A. Hazeltine in Proc. I.R.E., April, 1918, one by W. C. White in G. E. Review for September, 1916, and one by G. C. Southworth in the Radio Review for September, 1920. Southworth was able to obtain frequencies as high as 3×10^8 cycles per second.

any frequency between one per second or less to many millions per second; the low frequencies require very high values of L and C, and tight coupling, but are comparatively easy to produce. To get the very high frequencies, it is necessary to consider carefully all the capacity in the circuit, especially that in the tube.

The circuit shown in Fig. 162 will generate perhaps as high as 10^8 cycles per second, if the internal capacity of the tube is low. The oscillating circuit is indicated by the arrow, and must be made with very short leads; the condensers C_1 and C_2 should each be several milli-microfarads, and the values of L_1 and L have to be properly selected for maximum oscillating current.

These very high-frequency currents often occur when neither expected nor wanted. Thus in the connection scheme shown in Fig.

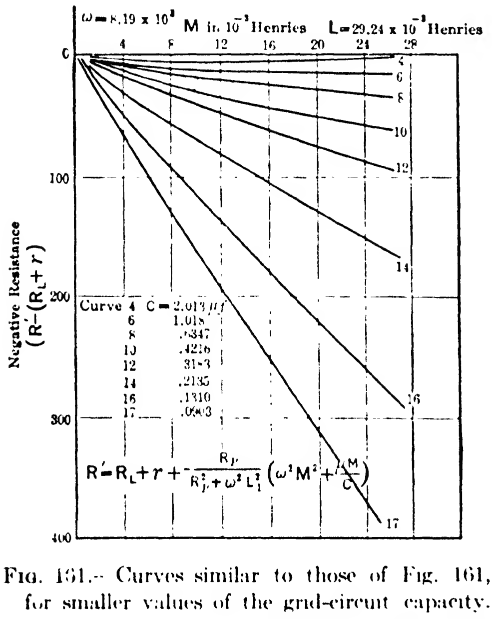


FIG. 161.—Curves similar to those of Fig. 161, for smaller values of the grid-circuit capacity. In the test R_c was zero, and the value of L_1 was 39.83 millihenries.

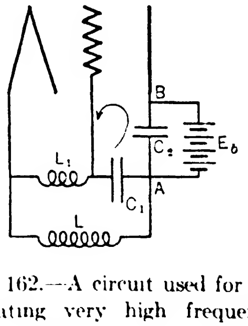


FIG. 162.—A circuit used for generating very high frequency; the oscillatory circuit is indicated by the arrow.

163, the circuit in which oscillations are desired is made up of L_g, L_p, R, and C, the current being indicated by ammeter A_1. If either R or C is too large, the conditions for oscillations in the main oscillatory circuit may not be satisfied, but the adjustment may serve to maintain oscillations in the circuit indicated by the arrow. That the tube is oscillating is known by indication of the c.c. ammeter which is nearly always used in the plate circuit, but ammeter A_1 shows nothing. If, however, a hot-wire meter of low resistance be inserted in the grid lead, as shown at A_2, it will be found that a comparatively large current is being generated in the local path.

A similar condition may occur in the circuit of Fig. 164; the main oscillating circuit L–C may show no current at all, but oscillations of very high frequency may be flowing through L_g and L_p, as indicated by the arrow, the dotted condenser really being the internal capacity of the tube.[1]

Englund[2] has investigated the possibility of generating short waves with the commercial tubes available and has concluded that to generate

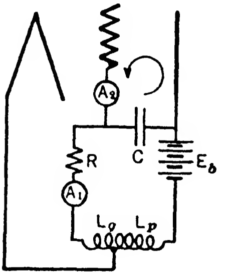


FIG. 163.—In a circuit such as this the oscillatory circuit is made up of R, L_G, L_p, and C in series; the circuit is very likely, however, to set up spurious high-frequency oscillations in the circuit including grid, plate, and C as indicated by the arrow.

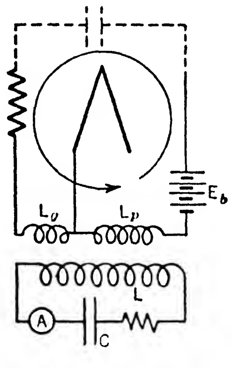


FIG. 164.—In a circuit of this kind (often called a Meissner circuit) spurious oscillations may be set up in the circuit indicated by the arrow, the main oscillatory circuit remaining unexcited.

wave lengths of less than 5 meters (frequency higher than 6×10^7) require special care and assembly of apparatus.

With a 5-watt tube of low internal capacity, with a circuit about as shown in Fig. 162, he was able to generate as low as 3.5 meters, but there was practically no control of the wave length in the circuit. By building the oscillating circuit right in the tube itself (thus making it non-adjustable) shorter connections may be used and the wave length lowered. By using a push-pull scheme as shown in Fig. 165, with Western Electric N tubes ("peanut" tubes), he was able to get a small fraction of 1 watt of power at 1.67 meter wave length.

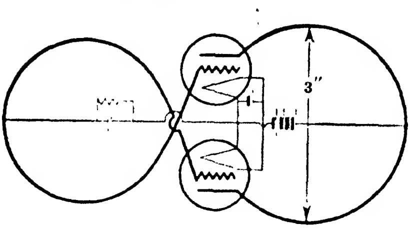


FIG. 165.—Using this push-pull scheme, and "N" tubes (very small triodes) it is possible to generate measurable power at 1.7 meter wave length.

[1] The circuit shown in Fig. 165, without the main oscillating circuit, (L–C–A) is frequently used to produce oscillations of high frequency in a receiving set. The values of L_g and L_p must be adjustable for different frequencies. A very complete discussion of this circuit is given by A. S. Blatterman in Vol. 1, No. 13, of the Radio Review.

[2] I.R.E., Nov., 1927, p. 914.

It is reported [1] that a triode has been built with small interelectrode capacity which is capable of generating 15 kw. of power at 5×10^7 cycles (6 meters wave length) and 0.8 kw. at 7×10^7 cycles. It will be noticed how rapidly the available power decreases when the wave length is less than 5 meters. The general type of construction used in these short-wave-length triodes is shown in Fig. 166; this is a 50-watt tube. It will be noticed that the grid and plate leads are kept far apart, so there is not much grid-plate capacity. McArthur and Spitzer [2] give a good summary of the high frequency oscillator, using both triodes and split anode magnetrons (Fig. 170). Using a water-cooled tube of the construction shown in Fig. 167 with plate voltages varying from 2 to 8 kv. they got an output of nearly 3 kw. at 5-meter wave length and 1 kw. at 2-meter wave length. The proper plate voltage to use varied with the wave length, and of course the efficiency of conversion also varied; in Fig. 168 are shown the performance curves of this special triode. Its capacities are low for a tube of this rating; C_{pg} is 8 $\mu\mu f$, C_{gf} is 5 $\mu\mu f$, and C_{fp} is 1.5 $\mu\mu f$.

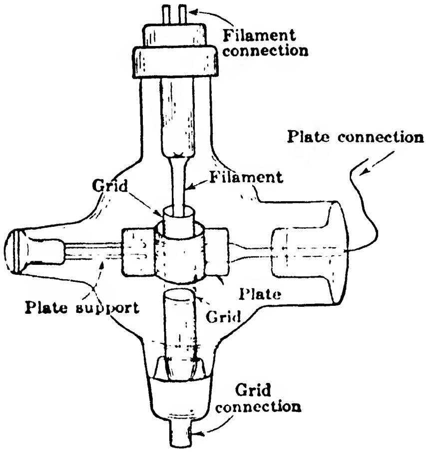


FIG. 166.—Triodes designed for generating very high frequency currents generally have the grid and plate leads well separated. This is a 50-watt triode used principally by amateurs.

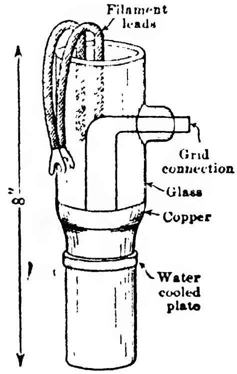


FIG. 167.—A compactly designed water-cooled triode, for generation of high frequency power.

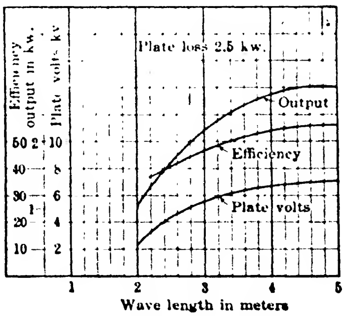


FIG. 168.—Performance of the triode shown in Fig. 167.

Acheson and Dart [3] have reported on a tube of this form, designed for

[1] G. E. Review, Jan., 1929, p. 37.
[2] I.R.E., Nov., 1931, p. 1971.
[3] I.R.E., March, 1932, p. 449.

high power at high frequency. The filament uses 52 amperes at 22 volts and gives 9.5 amperes of emission. Its ratings are different for different frequencies as given herewith.

Frequency	Plate Volts	Output	Frequency	Plate Volts	Output
1,500 kc.	20,000 volts	27 kw.	30,000 kc.	12,500 volts	14 kw.
10,000	18,000	24	40,000	10,000	9
20,000	15,000	20	50,000	8,000	5

When the tube was being used at the higher frequencies parts of the body were warmed to fever heat when even as far as 10 ft. from the oscillator.

Yager[1] has reported that with low-capacity tubes and a circuit arrangement as shown in Fig. 169, consisting of two Lecher systems of short parallel wires with movable bridges B, B', connected to grids and plates, he was able to get measurable power at a wave length of 60 cm. The wave length was controlled by the distance between tubes and bridges. By using a specially constructed magnetron (the operation of a magnetron is described on p. 695) having a split anode around a filament, as shown in Fig. 170, he reports measurable power at 30–40 cm. wave length.

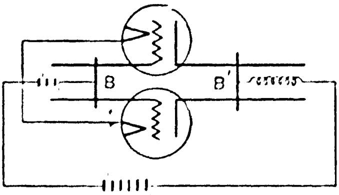


FIG. 169.—A push-pull scheme using Lecher wires to fix the frequency; small amounts of power at 60 cm. wave length may be generated if the internal capacity of the triodes is sufficiently low.

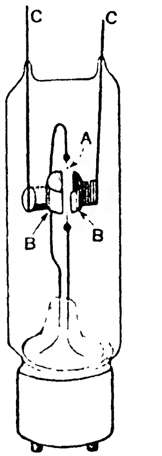


FIG. 170.—A special type of tube (magnetron) by the use of which very short waves (30–40 cm.) have been generated.

Apparently very short waves, nearly independent of the circuit to which the triode is attached, were first noticed by Barkhausen; they are called Barkhausen oscillations. Referring to Fig. 171, we suppose the grid to be held at a positive potential, say 100 volts, and the anode A at a low negative potential, say 10 volts. This, it will be noticed, is an entirely different potential distribution from that of the ordinary triode circuit, in which the grid is held negative and the plate positive.

[1] I.R.E., June, 1928, p. 715.

An electron starting from the filament at a may go directly into a grid wire, one starting at b may shoot between the grid wires and then, being slowed down by the grid and repelled by the plate, be sent into the grid on its return trip. And possibly one starting at C may go through the grid wires and back and forth through the grid several times before it enters the grid. We will calculate the probable frequency of such an oscillation.

It is shown on p. 461 that an electron, falling through a potential difference of 1 volt, gains a velocity of about 5×10^7 cm. per second. Hence falling from filament to gird of Fig. 171 (100 volts) it will have a velocity of 5×10^8 cm. per second as it passes the grid. Now if the oscillatory motion is sinusoidal the average velocity is $2/\pi$ of the maximum velocity, that is, $2/\pi$ of the velocity which the electron has as it passes the grid. This gives an average velocity for the harmonic oscillation of $1/\pi\times10^9$ cm. per second. If the grid is 0.5 cm. away from the filament the total distance covered by an electron in one cycle of its motion is about 2 cm., so that the time for one complete oscillation is $2\div10^9/\pi=2\pi\times10^{-9}$ second, and the frequency is thus 1.6×10^8 cycles per second. This corresponds to a wave length of 140 cm. It is thus evident that in the region of lower frequency which Englund found possible, the time taken for the electron to go from filament to plate was an appreciable part of the cycle.

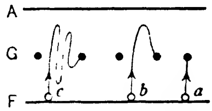


Fig. 171.—With high positive grid, and slightly negative plate, the electrons may execute gyrations as shown here; this produces what are called Barkhausen oscillations.

Of course these oscillations of the electrons are entirely haphazard in their relative phases, so that if the plate and grid of the tube actually had constant potentials, the effect of these heterogeneous oscillations in any outside circuit must be nil. However, it appears that, owing to the surging back and forth of the electrons in the plate and grid leads, their potentials must go up and down and thus react to some extent on the motions of the oscillating electrons. The net result is that the electrons "pull into phase" to some extent, giving a somewhat coordinated motion, thus giving to the plate and grid an alternating voltage, the frequency of which is determined primarily by the natural period of oscillation of the electrons.

We would expect the frequency of these oscillations to depend upon the grid potential, the same as the period of a pendulum swing depends upon the force of gravity. Hollman[1] has reported on these Barkhausen oscillations and gives a rough formula for determining their wave length

$$\lambda_{cm}=\frac{1000d_a}{\sqrt{E_g}}. \qquad \qquad (118)$$

in which d_a is the diameter of the anode, and E_g is the voltage of the grid.

[1] I.R.E., Feb., 1929, p. 229.

Using a triode arranged as in Fig. 172, the grid and plate were connected to a Lecher system with a movable bridge; this bridge short-circuited the wires for the high frequency, but kept them insulated for continuous currents.

With the bridge very close to the triode no oscillation occurred, but at a distance of about 20 cm. high-frequency oscillations were set up, which showed but little change in frequency as the bridge was moved along until the distance was about 50 cm. Here the wave length suddenly dropped and then increased linearly as the bridge was moved along. In Fig. 173 are shown roughly the results obtained; in the *A* region the wave length is nearly independent of the bridge position—these are the Barkhausen oscillations. For greater bridge distances the wave length suddenly dropped and then rose gradually with increasing bridge distance—these are sometimes called Gill-Morrel oscillations. The intensity of the Barkhausen oscillations decreases gradually as the bridge is moved along, and the shorter waves start suddenly when the critical bridge distance is reached.

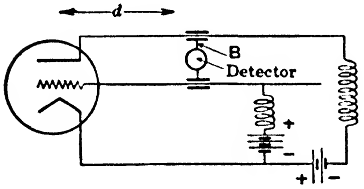


FIG. 172.—One scheme of utilizing Barkhausen oscillations.

For a given grid potential the wave length decreases to some extent as the anode potential is made more negative; this is undoubtedly because the amplitude of swing of the electron in the interelectrode space is less.

Circuits with Two Frequencies.—As mentioned above, many circuits have their coils and capacities so disposed that two natural frequencies exist for the combination. In the steady state a triode cannot generate two frequencies, it must oscillate at one or the other. Now as one of the condensers is decreased the frequency may gradually change until a certain point is reached and the oscillation suddenly jumps to that of the other circuit frequency. Now if the condenser is increased the oscillation will not jump back to the first circuit at what was the critical value of condenser before; it will hold on to the second circuit until the condenser is increased a good deal above the former critical value and then it will jump back suddenly to the first circuit oscillation.

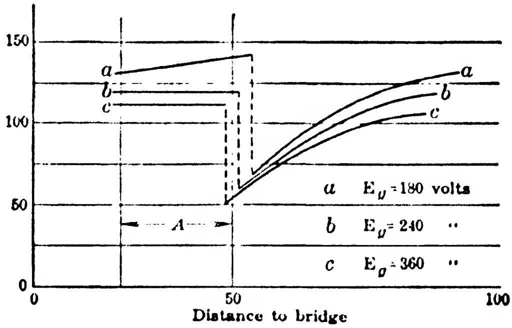


FIG. 173.—Showing the wave lengths generated by the circuit of Fig. 172; ordinates are centimeters.

Elimination of Undesired Frequencies.—These undesired high-frequency currents are sometimes troublesome, but may, in general, be eliminated by suitable precautions. Thus in a circuit used with a Type *P* pilotron the arrangement of apparatus was nearly as shown in Fig. 158; in series with R and C was another inductance, the actual connection being as shown on the curve sheet given in Fig. 232, p. 687.

With L_p, of this figure, below a critical value, the main circuit, L_g-L_p-C-L-R, will not oscillate; it is quite likely, however, that when the main circuit is not oscillating, high-frequency currents will be generated in the circuit made up of L_g and L_p in series with the internal capacity of the tube. Thus, with $L_g = 200\mu h$, $L_p = 400\mu h$, $L = 8000\mu h$, $C = 0.002\mu f$, the ammeter I (Fig. 232) gave no indication, but the meter I_p showed that the tube was oscillating violently. Test with wave meter showed the circuit, L_g-L_p-tube-capacity, to be generating a complex current of fundamental wave length equal to 800 meters; this is about the natural frequency of the circuit.

The desired wave length, of about 6000 meters, was not started until L_p was adjusted in excess of $1200\mu h$; the frequency changed suddenly from one value to the other, as L_p was varied through its critical value.[1] There is a tendency in such a circuit, however, to maintain the existing oscillation; thus if L_p was increased, the high-frequency oscillation persisted until L_p exceeded $1200\mu h$. As L_p was decreased, however, the high-frequency oscillation did not start until L_p was made less than $1000\mu h$, so that with $L_p = 1100\mu h$, either 900-meter or 6000-meter oscillations might exist, depending upon whether L_p had been decreased from a high value to $1100\mu h$, or had been brought up to the value from something lower.

An interesting condition was found in this test: if the condenser across machine E_b was taken out the high-frequency oscillation was very persistent, whereas the 6000-meter oscillation would not start, no matter what value L_p might have. Evidently for the lower frequency the machine offered a high inductive reactance and resistance, whereas for the high-frequency current it acted like a condenser of low impedance.

The undesired high-frequency current for the circuit above described was completely eliminated by introducing a suitable resistance directly in series with the grid, as indicated at A in Fig. 232; 100 ohms sufficed to diminish their amplitude considerably and 2000 ohms at this point resulted in such high losses for the 800-meter wave that it could not sustain itself. This high resistance had a negligible effect on the 6000-meter oscillation, because of the comparatively small charging current flowing to the grid at this frequency. A little thought will reveal many other schemes which will offer a high dissipation for the spurious oscillations, but have practically no effect on the main oscillation.

[1] See also article by Möller, "Jahrbuch der Drathlosen Telegraphie," December, 1920.

Constancy of Frequency of an Oscillating Tube.—The foregoing formulas for frequency of oscillation of a tube circuit have all been derived on the assumption that the grid current was zero, and do not involve any characteristics of the tube, except μ and R_p. It is a fact, however, that there is an appreciable capacity between the grid and filament of a tube, and that the value of this capacity varies with any factor which affects the amplification (not μ)[1] of the tube and circuit, as shown on page 523 et seq. This grid-filament capacity is always shunted across a part of the oscillating circuit and so must have an effect on the frequency of oscillation of the circuit.

It is therefore evident that any factor which influences either tube resistance or grid-filament capacity must also affect, to some extent, the frequency of oscillation, and such is found to be the case. A change in either of the filament current or plate voltage will produce variations in frequency, the variation sometimes amounting to 1 per cent or 2 per cent, without excessive change in either I_f or E_b.[2]

However, if batteries are used for filament and plate circuits and the set has been operating an hour or two to get the batteries in a "steady" condition and to warm up the coils and condensers as much as they are going to, the frequency stays remarkably constant. Thus in one such circuit the frequency of a 50-kc. power set in the laboratory varied less than 5 cycles in two hours. In a smaller set, which gave practically no drain on the batteries, and did not appreciably heat the apparatus the change in frequency was about 1 cycle in 100 kc., in a half-hour run.

Of course if the output of the circuit is changed it may well be expected that the frequency will, in general, also change. Changing the plate circuit impedance changes the amplification factor of the circuit and this in turn changes the effective input capacity. This effect can be eliminated however, by the use of a neutralizing circuit,[3] which prevents the plate circuit from reacting in the grid circuit.

Eller [4] has given a complete analysis of the factors affecting the frequency of an oscillating triode, and has shown that, with reasonable precaution in holding voltages and circuit conditions constant, the frequency remains fixed within about 1 part in 20,000. This assumes that temperature conditions are reasonably constant. He found the tuned plate

[1] Changing the plate-circuit impedance changes the effective value of the tube capacity (and hence its effect on the frequency of oscillation), because the amplification of the tube and circuit has been changed; the μ of the tube, however, has not been altered by changing the plate-circuit impedance.

[2] Transmitter tubes should never have their frequency fixed by the capacity of the antenna, which varies as it swings in the wind; a small master oscillator, working into a closed oscillating circuit, should be used for setting the frequency of the big tubes.

[3] Discussed in Chapter X.

[4] I.R.E., Dec., 1928, p. 1706.

circuit, Fig. 144, with grid condenser and leak somewhat more stable than the tuned grid circuit, Fig. 159.

With a 201 A triode having plate circuit tuned to 1000 cycles, grid condenser and leak of 0.025 μf and 0.5 $M\Omega$ respectively he reports that a 60 per cent change in E_p gave a frequency change of only 0.09 per cent, and a 30 per cent change in I_f gave a frequency change of only 0.03 per cent. He found that adding resistance to the oscillating circuit sometimes increased the frequency and sometimes diminished it, also that, in general, any change which increased the current drawn by the grid, decreased the frequency.

Llewellyn[1] has reported results somewhat similar with circuits tuned to 10^6 cycles. Diminishing the plate battery by 60 per cent, or the filament current by 30 per cent, changed the frequency less than 100 cycles in 1,000,000.

Fixing Frequency by Piezo-electric Crystal.—There are certain crystals, notably quartz and Rochelle salts, which show the phenomenon of piezo-electricity, or development of electric charge as a result of pressure. A suitably crystallized piece of Rochelle salts will show a difference of potential on two of its faces as high as several hundred volts, when vigorously twisted.[2] A piece of quartz crystal, properly cut, will develop a few volts difference of potential between its opposite faces, when squeezed. These crystals develop a charge when their shape is changed and as the phenomenon is a reversible one, they change their shape when charged.

This peculiar action makes it possible to control the frequency of oscillation of a triode, by the mechanical vibration of a piece of quartz. Now quartz is, mechanically, a very perfect material; it is extremely elastic, having almost no loss due to viscosity. A piece of quartz, started into mechanical oscillation by a blow, will vibrate for a very long time, as compared to steel and other elastic materials. Quartz then requires but very little energy supply to maintain it in mechanical oscillation. Weather conditions and temperature have practically no effect on the elastic properties of quartz so that the natural period of vibration of a piece of quartz is practically a constant; certainly it is constant to a very small fraction of 1 per cent.

The application of the quartz resonators for fixing the frequency of oscillation of a station has been developed mostly by W. G. Cady.[3] He

[1] I.R.E., Dec., 1931, p. 2063.

[2] See Nicolson, Proc. A.I.E.E., Vol. 38, p. 1315, 1919; also Electrical World, June 12, 1920, page 1358. See also an article on Rochelle Salts by Sawyer, Phys. Rev., Feb. 1, 1930, giving its mechanical expansion, saturation electric field, dielectric constant, etc.; and another by Sawyer on using this crystal for generating sound, in I.R.E., Nov., 1931, p. 2020.

[3] See "The Piezo-electric Resonator," by W. G. Cady, Proc. I.R.E., Vol. 10, No. 2, April, 1922. Also "Uses and Possibilities of Piezoelectric Oscillators," by August Hund, Proc. I.R.E., August, 1926.

shows how the piece of piezo-electric crystal can be treated as a special electric circuit, having peculiar changes in resistance and reactance in the region of frequencies for which it is mechanically in resonance. This gives one of the most striking illustrations of the ideas brought out in Chapter II, regarding the general viewpoint from which complex electric circuits must be viewed. Thus a thin slab of quartz, having tinfoil pasted on its two opposite faces, is certainly a condenser, from the physical viewpoint. It is in fact a very perfect condenser, in general, the specific inductive capacity of quartz being about 4.5 and its insulation resistance extremely high. Yet this slab of quartz, when tested in an a.c. bridge, instead of showing a constant capacity, of negligible series resistance, shows a changing capacity; at one frequency it becomes a pure resistance and within a certain narrow band of frequencies it shows inductive reactance!

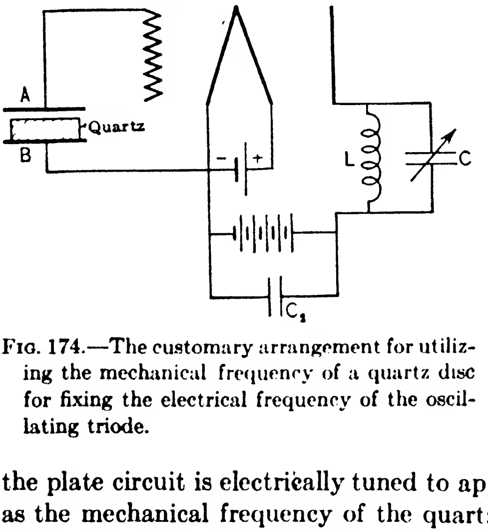


FIG. 174.—The customary arrangement for utilizing the mechanical frequency of a quartz disc for fixing the electrical frequency of the oscillating triode.

In Fig. 174 is shown a common arrangement for a standard frequency oscillator.[1] A small disc of quartz, perhaps as big as a dime, is loosely held between two metal plates A and B, forming a minute grid condenser. This piece of quartz will have a certain mechanical period of vibration (in fact three of them; see Hund's paper) and when the plate circuit is electrically tuned to approximately the same frequency as the mechanical frequency of the quartz the circuit will oscillate quite violently. The tuning of the plate circuit is merely to facilitate the transfer of energy from plate to grid circuit, not to set the frequency. There may be a considerable range of the variable condenser C which sustains oscillations, but it may be observed that as C is varied the frequency of oscillation does not vary at all. A small dry cell tube, with only 20 volts on the plate, will operate this circuit, if the quartz slab is good, piezo electrically.

This scheme is used nowadays in setting the frequency of a broadcasting station; by exciting a 5-watt tube from this small master oscillator, and then a 50-watt tube from the 5-watt tube, etc., the frequency of the main circuit, of many kilowatts of power, is controlled accurately by the minute quartz disc.

[1] A Mathematical Analysis of this Circuit is given by Wheeler, I.R.E., April, 1931 p. 627.

Characteristics of Piezo-electric Quartz.—A normal quartz crystal is somewhat hexagonal in form, more or less pointed at one end. The longitudinal axis of the crystal is called its optical axis; polarized light sent through the crystal along this axis will have its plane of polarization rotated, either clockwise or counter-clockwise, as the crystal is a right- or left-handed one. Sometimes a crystal, during its period of formation, starts to grow right-handed and then stops and grows the other way. Such crystals are said to be "twinned." This twinned structure is not evident to the eye but must be investigated by the action of polarized light or similar test. It is possible to find a crystal which has grown in several layers, alternate layers having opposite rotations. Such crystals make very poor oscillators, unless the plate is small enough to be cut from one layer only.

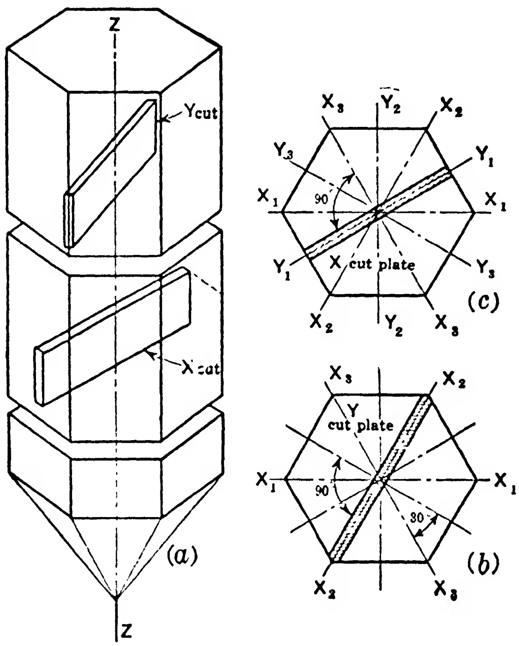


FIG. 175.—Conventionalized drawing of a quartz crystal; they are practically never found with such perfect form. In the cross sections (*b*) and (*c*), are shown the two ways of cutting plates from the crystal.

In Fig. 175 (*a*) is shown a crystal much more uniform in structure than is usual, and in *b* and *c* are shown two cross-sections. The so-called electric axes are marked by X_1, X_2, and X_3; it is along these axes that the crystals show most piezo activity. The axes perpendicular to the crystal faces, marked Y_1, Y_2, and Y_3, are called the mechanical axes of the crystal. For plates having a frequency of vibration less than 200 kc. it is customary to use rectangular slabs taken from the crystal as shown by the "Y cut" in Fig. 175 (*a*); for frequencies above 400 kc. it is customary to use Y cut slabs, which are generally elliptical in form, somewhat like a small spectacle lens. The low-frequency rectangular slabs vibrate principally in the direction of their long dimension and so must have their length accurately determined. The elliptical slabs for higher frequency oscillate along the direction of their thickness; they must be accurately ground so that their thickness is everywhere the same.

Various names have been applied to plates taken from the crystal in the two ways shown in Fig. 175 (*b*) and (*c*). X cut plates are called "zero angle plates," "face perpendicular cut," or "Curie cut"; Y cut plates are called "30° cut" or "face parallel cut." The 30° cut plates generally oscillate more readily than the zero angle ones. The Curies were the first to use the piezo-electric properties of quartz, and as they used X cut plates their name has been given to this type; they used quartz slabs in calibrating electrometers, etc.

Whereas the slabs are shown taken from the center of the crystal in Fig. 175 this is not necessary or customary; the whole crystal can be cut up in layers parallel to the pieces shown in Fig. 175, (*b*) and (*c*), and all of the slabs will act alike.

The velocity of a compression wave in quartz is about 5×10^5 cm. per second; for a slab to vibrate in the half wave length mode at, say, 50 kc., it must be about 5 cm. long; the wave length of a 50-kc. frequency will be $5\times10^5\div5\times10^4$ or 10 cm., so that a half wave length plate (half wave length is the normal type of oscillation used, as the quartz is free at both ends) must be 5 cm. long. The length might vary appreciably from this value because the velocity of wave travel is different in the Y axis direction from that along the X axis.

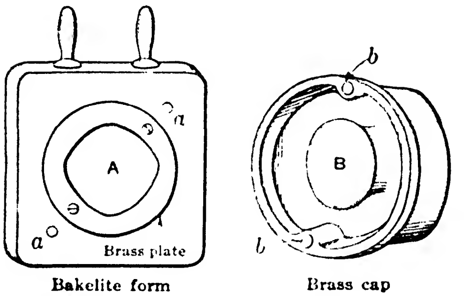


FIG. 176.—The small oscillating discs are held in mountings similar to this one. The two electrical terminals connect to the two metal plates, *A* and *B*, which come almost in contact with the faces of the quartz slab.

For plates vibrating along their thickness dimension the half wave length frequency is about 2×10^6 cycles per second for one millimeter thickness of Y cut plates and more nearly 3×10^6 cycles per second for one millimeter thickness of X cut plates. This corresponds to a radio wave length of 140–150 meters per millimeter thickness for Y cut plates and 100–110 meters per millimeter thickness for X cut plates. Thus a plate to vibrate 1800 kc. is about the size and thickness of a dime; it must be the same thickness throughout to perhaps better than 0.001 mm.

Mounting of Oscillating Plates.—The small plates (high frequency) are generally loosely mounted in a container as shown in Fig. 176. A molded form is fitted with a brass plate having a shallow recessed space *A*, somewhat larger than the quartz plate. A cover of brass has a finished surface *B* which almost touches the crystal plate, when the cover is screwed down tight on the molded form by screws through holes *b*, *b*, into the threaded holes *a*, *a*. The two brass plates, top and bottom, are of course insulated from each other by this construction; each is electrically connected to one of the plugs.

Hund[1] has studied the effect of the mounting on the action of the quartz plate and finds that the size of the air cavity between the quartz and upper plate has a powerful effect upon the crystal oscillation. Fig. 177 shows how a crystal acted as the air gap between it and the upper plate was varied. Evidently when the air gap was half wave length or full wave length long the crystal was loaded so much it would not oscillate. Apparently when the air column was an integral number of half wave lengths the air was set into violent agitation and hence tended to use up so much power that the crystal action could not supply it. The longer the air gap the greater was the amplitude of oscillation of the quartz (except at the critical length). As would be expected, the length of this air gap has much influence on the frequency of oscillation of quartz; the air gap gives an elastic force and load to the quartz plate, both varying with the length of gap. Because of this effect, an accurate quartz oscillator is always calibrated in its mounting, and this is sealed so as to be non-adjustable after the calibration is made. The length of the air gap may be used to vary slightly the oscillation frequency of the plate, thus doing away with the necessity of continued grinding operations. As shown later, temperature may be used in the same way.

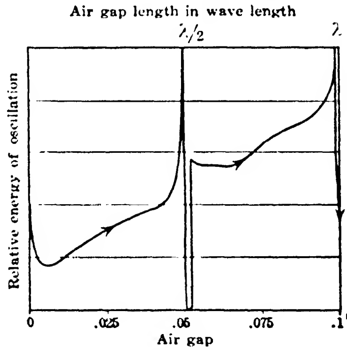


FIG. 177.—The length of air gap between the face of the quartz disc and the metal plate *B* (Fig. 176) has a pronounced effect on the oscillation.

Temperature Effects in Quartz.—The velocity of sound (compression and rarefaction waves) in quartz varies with temperature, as it does in practically every other sound-conducting medium. But if the velocity changes the natural frequency of oscillation must change; strangely enough the effect is opposite in the two types of plates. An X cut plate has a negative temperature coefficient, i.e., the frequency of oscillation decreases as the temperature rises; it amounts to between 10 and 25 parts per million per degree centigrade. The Y cut plates have a positive coefficient, the natural frequency rising with increasing temperature. The effect raises the natural frequency from 25 to 100 parts per million per degree centigrade.

Lack[2] has investigated the temperature effect for various styles of

[1] I.R.E., Aug., 1928, p. 1072. See also "Quartz Plate Mountings and Temperature Control for Piezo Oscillators," Heaton and Lapham, I.R.E., Feb., 1932, p. 261.

[2] I.R.E., July, 1929, p. 1123.

cutting, etc. One experimenter reports that a 2700-kc. plate had a temperature effect of 61 cycles per degree at 65° C., decreasing linearly to 4 cycles per degree at −189° C.

Marrison [1] has ingeniously cut plates from the crystal in such a way that the two temperature effects nearly neutralize one another. Fig. 178 shows his method of cutting these oscillators; from a Y cut slab is cut out, by a hollow drill, a life-saver-shaped piece, which oscillates along its thickness dimension. Whereas a solid disc, of the same dimensions as the outside dimensions of the ring, showed a frequency change of 30 parts per million per degree centigrade, the ring showed a change of only 1 part in a million under the same conditions. Marrison says that these ring-shaped pieces are good to 1 part in 100,000 even with no temperature control. They are mounted as shown in Fig. 179. Two brass plates

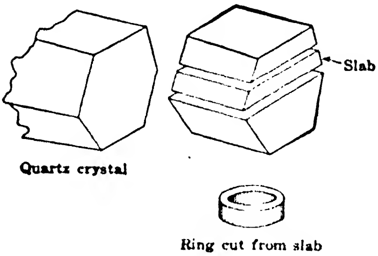


FIG. 178.—Properly cut toroidal-shaped pieces of quartz show oscillation frequencies almost independent of temperature.

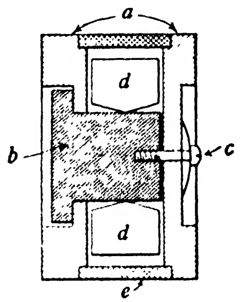


FIG. 179.—The rings of quartz (Fig. 178) are mounted in holders of the form shown here; the metal plates, *a*, serve as the electrical terminals. The quartz ring is shown in section, at *d*-*d*.

a, *a*, are held together by a core of insulating material, *b*, and screw *c*. The ring-shaped quartz *dd*, the inner surface of which has been tapered from both ends to the middle, fits loosely over the core and loosely between the end plates. A ring, *e*, of insulating material, closes the ring from dust, etc.

This crystal ring in its holder, mounted on a shock-absorbing support, is kept in a temperature-controlled oven, which holds its temperature at about 50° C. within 0.1°. Marrison estimates that such a crystal standard will not vary from its mean frequency by more than 1 part in 10 million, over long periods of time.

Clapp [2] has given a good summary of the various factors acting to make a quartz oscillator change its frequency, and summarizes them as

[1] I.R.E., July, 1929, p. 1103.
[2] General Radio Experimenter, Oct. and Nov., 1930.

follows, allowing what seems to be a reasonable variation in the factor considered. He assumes the circuit given in Fig. 174, using a type 112A triode, temperature of oven at 50° C., plate circuit condenser set at its lowest value to produce oscillation, filament volts 5 and plate volts 45. The plates were Y cut discs, not rings, like Marrison's.

Variable	Range of Variation	Frequency Change Parts in One Million
Temperature	±0.1° C.	±5.0
Plate capacity	±1.0%	±1.0
Plate voltage	±1 0 volt	±0.5
Filament volts	±0.1 volt	±0.5
Various tubes	(Average)	±2.0
Vibration	Heavy shocks	±3.0
Total	....	±12 0

There is not much chance of all the variables acting in the same direction at the same time; it is more likely that some will act to increase and others to decrease the frequency, thus making the probable change not more than 8 or 10 parts per million. This degree of constancy of the quartz oscillator is the one factor which makes it possible for the Federal Radio Commission to require a station to hold to its assigned frequency within 50 cycles, whereas before 1932 a departure of 500 cycles was permitted.

It is to be remembered that for a quartz crystal to act as shown in the above table no appreciable power is to be drawn from the electric circuit and the vibration of the quartz plate is comparatively weak. If the same crystal is used in a 5-watt tube, and this is used to control the frequency of an amplifier, the vibration of the quartz will be more violent, the quartz will heat more, and the variation of frequency will probably be considerably in excess to those given in the table. The values there given are for a station " monitor," the circuit serves merely as a comparison circuit, for testing station frequency by a heterodyne method.

The Bureau of Standards uses many quartz oscillators to maintain their frequency standards; the manner in which they are used to check against each other and to check the broadcasting stations is described by Hall.[1]

Hollman [2] has given some interesting data in the use of quartz oscillators in maintaining station frequency within the limits specified by the Radio Commission.

[1] I.R.E., March, 1930, p. 490.
[2] Radio Engineering, Feb., 1932, p. 15.

Electric Circuit Equivalent of a Quartz Disc.—A quartz disc, having electrode above and below it, and mounted so that it is free to oscillate, is apparently nothing but a small condenser; but when it is in a state of oscillation the piezo-electric effect produces electric charges on the faces of the disc, which act something like the back voltage of an a.c. synchronous motor. As the impressed electric frequency is increased through that value which gives mechanical resonance to the quartz disc, this back voltage varies greatly in phase and magnitude, causing the crystal to act like a condenser, resistance, or coil at the various frequencies. It is equivalent to a series resonant circuit and is so shown in Fig. 180.

Lack has calculated the constants of a circuit, equivalent to a quartz disc 1 mm. thick and the size of a dime. The piezo-electric effect (steady) of the quartz was such that 3000 volts (c.c.) impressed on the disc caused a thickness change of 6×10^{-7} cm. In Fig. 180 C_2 is the capacity of the air between the electrodes and the disc, C_1 is the capacity of the quartz, at frequencies remote from that of mechanical resonance, the L-C_0-R circuit is the equivalent of the piezo-electric effect of the vibrating quartz. For the disc mentioned above, the equivalent circuit has $C_0=0.06\ \mu\mu f$, $C_1=1.0\ \mu\mu f$, $C_2=0.06\ \mu\mu f$, $R=100$ ohms, $L=0.5$ henry. This gives the L-C_0-R circuit a decrement, $R/2fL$ of 10^{-4}, which means that the reactance of either the coil or condenser, at resonant frequency, is 31,000 times the resistance! This indicates how sharply resonant the mechanical vibration of quartz is; a very good radio circuit might have the reactance 300 times the resistance, so that the mechanical resonance curve of the piezo active quartz is 100 times as sharp as that of a very good electric circuit.

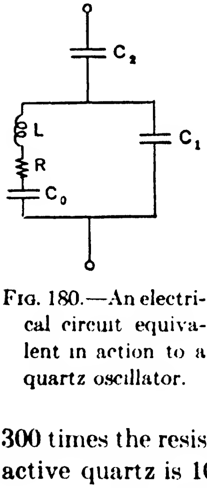


FIG. 180.—An electrical circuit equivalent in action to a quartz oscillator.

Van Dyke[1] gives a somewhat higher decrement for a 1100-kc. disc; he found the equivalent circuit to have $L=0.33$ henry, $C_0=0.065\ \mu\mu f$, and $R=5500$ ohms. Terry[2] gives for a slab 3.33 cm. long, 2.75 cm. wide, and 0.636 cm. thick: $L=3.65$ henries, $R=9045$ ohms, and $C_0=0.031\ \mu\mu f$. From the last two sets of data it would seem that the crystal tested by Lack was exceptionally free from friction effects.

In Fig. 181 is shown the general form of the reactance-frequency curve of a piezo-electrically active piece of quartz. It will be noticed that within a narrow frequency range it actually acts like a coil, having positive reactance.

The Oscillating Tube as Detector of Continuous Waves—Autodyne.—The circuit given in Fig. 159 is generally used for exciting a tube used

[1] I.R.E., June, 1928, p. 742.
[2] I.R.E., Nov., 1928, p. 1486.

as autodyne receiver; with no grid condenser, as shown in Fig. 182, the detecting efficiency of the tube is indicated by Eq. (48). The antenna circuit L_3C_a is tuned to incoming signals and the circuit L_2C is tuned to a frequency differing from this signal frequency by about 800 cycles per second, so as to give a beat note for which both the ear and ordinary telephone receiver are sensitive.

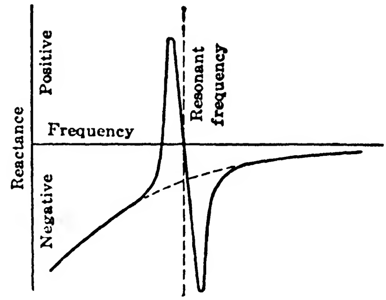


FIG. 181.—Electrical characteristics of a piece of piezo active quartz, as the impressed electrical frequency passes through the mechanical resonant period of the quartz.

From Eq. (48) it seems that the more violently the tube is oscillating, thereby making E'_g as large as possible, the more sensitive will the tube act as detector, and so it does as long as d^2i_p/de_g^2 remains constant. This term d^2i_p/de_g^2 is really a measure of the asymmetry of the change in plate current when E_g is positive and when it is negative. In other words, it measures the excess of the increase of plate current for positive E_g over the decrease for negative E_g. So long as the relation between I_p and E_g is parabolic the value of di^2_p/de_g^2 is constant, but for this condition the tube resistance R_p is also constant. We have previously shown, however, that to make a tube oscillate, the coupling (of whatever kind is used) must be increased beyond a certain critical value, and that after this value is past the oscillations start and automatically increase in amplitude, until the plate resistance R_p is sufficiently increased to restore a certain balance which was destroyed by increasing M. This change of resistance was analyzed in discussing Fig. 150. The plate current in an autodyne receiver will fluctuate over the straight part of the full-line curve of this figure if the value of M (between L_1 and L_2 of Fig. 182) is kept sufficiently low: if it is increased much beyond its critical value the fluctuation in plate current will extend over the upper and lower bends of the curves.

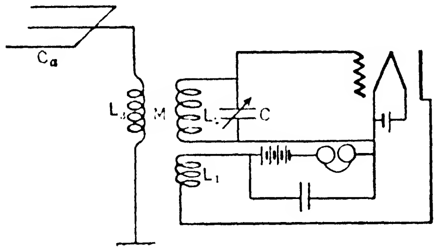


FIG. 182.—This is the arrangement generally used when an oscillating tube is to act as detector for continuous-wave signals. The frequency of the local oscillations is fixed by the values of L_2 and C, the tickler coil, L_1 serving to make the tube oscillate.

The tube will act best as a detector of continuous-wave signals for that coupling of L_1 and L_2 (Fig. 182) which results in the greatest product of $E'_g(d^2i_p/de_g^2)$. This product will generally be a maximum for the weakest coupling which will maintain the tube in the oscillating state; such is nearly always found to be the case in practice. If the coupling between L_2 and L_3 is held constant and the coupling between L_2 and L_1 is diminished, the signal strength will be a maximum for the weakest possible coupling. In carrying out this test it is necessary continually to change C to keep the beat note of constant pitch, because of the effect of L_1 on the value of the effective self-induction of L_2.

Three possible conditions of the adjustment of a beat receiver are shown in Fig. 183. In (*a*) the coupling is so adjusted (tight) that the grid potential, with no incoming signal, fluctuates between A and B; the plate current fluctuates with a frequency nearly the same as that of the signal, between the values AG and BH, its average value being OI. This current OI flows through the phones and the high-frequency alternating component of the plate current is carried by the condenser shunting the phones. In case no actual condenser is used to shunt the phones this current will utilize the capacity of the phone cords or the distributed capacity of the windings to by-pass the high inductance circuit of the windings themselves.

When the signal voltage E_g is superimposed on the grid it alternately increases and decreases the amplitude of the grid fluctuations of potential; the value of grid potential now fluctuates with variable amplitude, the amplitude being fixed by the limiting values EF and DC, the frequency of these cycles of variation of amplitude being equal to the difference in frequency of E_g and E'_g.

The plate current will now be of the form shown in the right-hand part of the diagram, and the average value of this high-frequency plate current will be as shown by the dashed line shown at K, L, M, etc., and it is this pulsating current which, flowing through the telephone receivers, gives the signal.

In diagram (*b*) of Fig. 183 is shown the effect on the signal strength of reducing somewhat the amplitude of the locally generated oscillations E'_g, which occurs as a result of decreasing the coupling between L_1 and L_2 in Fig. 182 (dotted line of Fig. 148). Although E'_g is less than in diagram (*a*), the value of the signal current (shown again by the dashed line) is greater for (*b*) than it is for (*a*).

In diagram (*c*) of Fig. 183 is shown the result of still further decreasing the value of the local oscillation E'_g; it is likely that M could not be sufficiently reduced to make the tube oscillate in this fashion without stopping the oscillations altogether. The signal current is, however, greater for this condition than for either of the two other values of E'_g shown at (*a*) and (*b*).

An accurate analysis of the " detected current " in this case, however, will show it to be non-sinusoidal, a distortion effect which has some importance in radio telephony. It will be discussed in that chapter of the text.

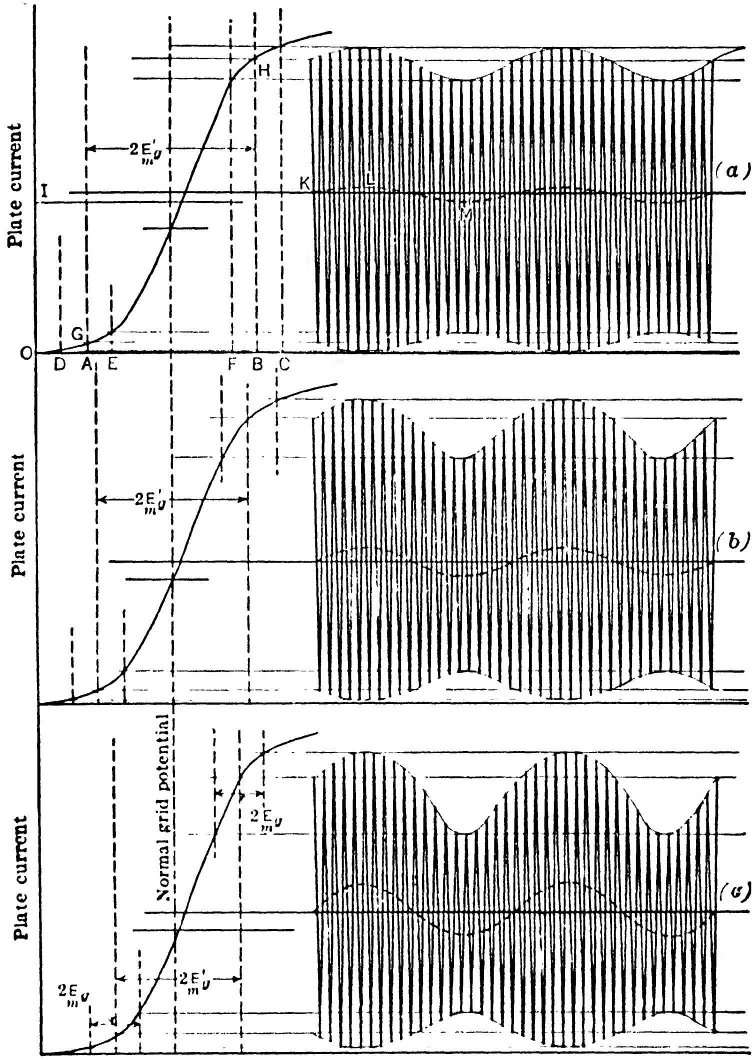


FIG. 183.—This diagram shows the effect of the strength of the local oscillations on the signal strength; the audio-frequency current through the phones, which gives the audible signal, is indicated by the wavy dashed line in each diagram. In (*a*) the local oscillations are too violent to give a good signal, in (*b*) the signal is somewhat improved and in (*c*) it is best. It is doubtful if the local oscillation could be cut down as much as indicated in (*c*) without stopping the oscillations altogether. For all three diagrams the amplitude of the high-frequency signal voltage is the same.

Use of a Separate Tube for Generating the Local Oscillations.—In order to use the vacuum tube as detector most efficiently it is necessary to have the amplitude of the voltage E'_g under control, and this can best be done by using a separate tube for generating the voltage E'_g, in addition to the detecting tube. The scheme of connection is then as shown in Fig. 184. The local oscillations are generated in tube B, their frequency being fixed approximately by L_4, L_5, and C_1, and intensity by the coupling between L_5 and L_6. This coupling should be considerably greater than the critical value, so that as conditions in the circuit are changed the oscillations of tube B are not stopped.

The value of E'_g impressed on the grid of the detecting tube A can be controlled by varying the mutual inductance between L_3 and L_4, either by moving the coils with respect to one another or by changing

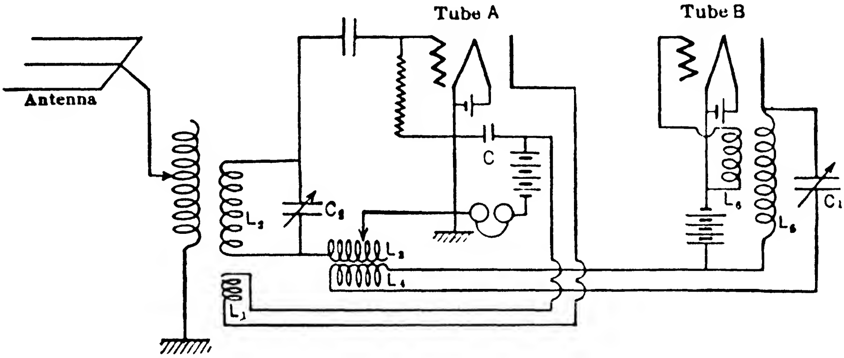


FIG. 184.—In order to control easily the strength of the local oscillations impressed on the detecting tube it is best to have a separate oscillator and couple this properly to the detector, Tube A. In this diagram Tube B is the oscillator; it is coupled to the detector by the two coils L_3 and L_4.

the value of either of them. The value of M should be so adjusted that the condition obtained is that shown in Fig. 183, diagram (c).

The antenna circuit and L_2C_2 circuit are each tuned for the frequency of the incoming signal, and the coupling between L_1 and L_2 is adjusted as near the critical value as possible. We have shown that the effect of the coupling between L_2 and L_1 is to decrease the resistance of the L_2C_2 circuit, and this resistance may be made to approach zero, if the coupling is suitably adjusted. Further, the L_2C_2 circuit can be exactly tuned for the incoming signal, so that the reactance is zero also, hence *the impedance of the L_2C_2 circuit may be made to approach very close to zero, so that the current caused to flow by a weak signal may be perhaps a hundred or more times greater than it would be if the coupling L_1–L_2 were not used.*[1]

[1] An experimental investigation of the magnification obtainable in such circuits was carried out by E. H. Armstrong and reported in Proc. I.R.E., Vol. 5, No. 2, April, 1917.

The impedance of the L_2C_2 circuit, as a function of the impressed frequency, has the form shown in Fig. 185; it is evident from this curve that not only is the circuit of Fig. 184 one to amplify signal strength, but also that this amplification is very selective. With a low-resistance coil for L_2 and a well-insulated condenser, and the grid circuit of the tube adjusted to absorb but little power, the selectivity is extremely sharp.

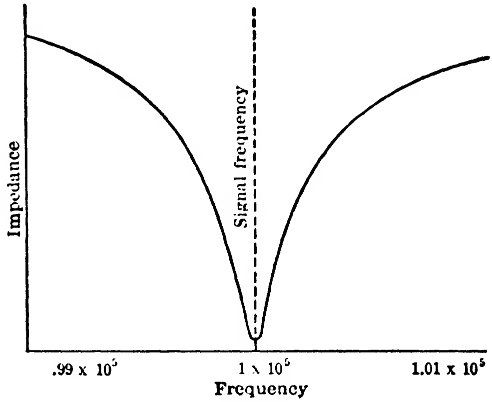


FIG. 185. By properly adjusting the coupling of coils L_1 and L_2 of Fig. 184 (keeping the coupling too low to produce oscillations in L_2-C_2) the resistance of the circuit L_2 C_2 may be made to approach zero. This curve shows how the impedance of the L_2 C_2 circuit will then vary with frequency of impressed signal.

Effect of Condenser in Series with the Grid on the Critical Coupling. In the foregoing analyses of the conditions required for self-excitation of tubes no mention was made of the effect of a condenser in series with the grid, as affecting the possibility of oscillation. In some common oscillating circuits it is necessary to use a grid condenser to insulate the grid from a high positive potential; such a one is shown in Fig. 186.

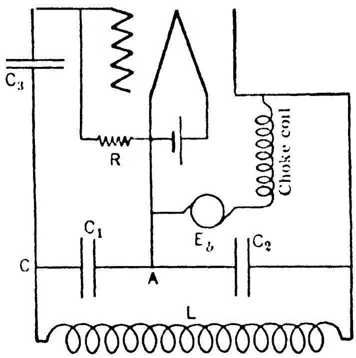


FIG. 186. In a circuit of this kind it is necessary to use a condenser C_3 to insulate the grid from the high continuous voltage impressed on coil L by machine E_b.

The oscillating circuit is made up of L with C_1 and C_2 in series, and the tube is connected to it as shown. The excitation for the grid is supplied by the drop of potential between the points A C and the plate voltage is fixed by the drop across condenser C_2. The scheme of connection results in the coil L being at plate potential, i.e., it is positive with respect to the filament by an amount equal to E_b; if the grid were connected directly to point C, the tube would at once burn out, due to excessive plate and grid currents.

The grid is therefore insulated (in so far as continuous voltage is concerned) by the condenser C_3; a suitable leak resistance R serves to hold the grid at a proper average potential.

The excitation impressed on the grid is now not equal to the potential drop between points $A-C$, but somewhat less than this due to the drop across the condenser C_3; moreover, due to the absence of the leak path across C_3 and the presence of such a leak across the grid-filament circuit, the phase of the voltage impressed on the grid is not the same as that of the voltage across condenser C_1.

The effect of the drop across C_3 is to require a higher drop across $A-C$ than would otherwise be required; if it should happen that the capacity of C_3 is equal to the capacity of the input circuit of the tube, then C_1 must be made only one-half as large as it would otherwise have to be.

Effect of Oscillations on the Magnitude of the Plate Current.—If an oscillatory circuit has a condenser in series with the grid, the average value of the plate current will always decrease after oscillations are set up; this is due to accumulation of electrons on the grid forcing its average potential more negative when oscillations start, causing an accompanying decrease in the plate current. This effect is shown by the decrease in the reading of a c.c. meter in series with the plate.

However, careful observation will generally show that the plate current, as indicated by a c.c. meter, will momentarily increase just as oscillations start, and then as the oscillations persist the plate current will decrease and go to a value lower than it has when the tube is in the non-oscillatory state. This momentary increase (which generally lasts only a fraction of a second) is particularly noticeable if the grid condenser is comparatively large.

As was mentioned when analyzing the detector action of a triode with grid condenser there are two rectifying actions taking place in the triode, which actions have opposite effects on the average value of the plate current. Due to the concave shape of the E_g I_p curve the fluctuations of grid potential tend to *increase* the average value of plate current, whereas the accumulation of electrons on the grid (due to the concave form of the E_g-I_g curve) tends to *decrease* the average value of plate current.

But it takes an appreciable time to charge the grid condenser and during this time the *increase* tendency may predominate. The larger the grid condenser the longer it takes to accumulate its charge and the more pronounced the above-mentioned effect becomes.

When the oscillating circuit is such that no grid condenser is required and none is used, the reading of the c.c. meter in the plate circuit will generally increase when oscillations start, the increase being more the greater the excitation of the grid. This statement is not universally true; it is possible to so adjust the conditions that when oscillations start the average value of the plate current stays the same, or even decreases. This effect can be noted if, with all other conditions constant, the filament current is varied throughout a sufficient range of values; with high filament current, the plate current will increase when oscillations start, and with low filament current it will decrease. The case is similar to the action of the tube as a detector, without grid condenser as described in p. 534 et seq.; it is there shown that the effect of an incoming signal may be to either increase or decrease the average value of the plate current.

Criterions of the Oscillating Condition of a Detecting Tube.—In the case of a power tube the oscillatory condition is indicated by the meters used either in the grid circuit, plate circuit, or oscillating circuit. In the case of a small tube used for the detection of continuous-wave signals there are generally no meters in the circuit to indicate oscillations; it is, however, extremely important that the operator should know at all times whether or not his tube circuit is oscillating, because if it is not oscillating he cannot possibly hear the signal for which he is listening. The only method of testing for oscillations in the ordinary continuous-wave detecting set is to properly interpret the noises in the telephone receivers; to an experienced operator they serve as well as do the meters on a power set.

When no condenser is used in series with the grid it is very easy to tell when the tube is oscillating and when not; when grid condenser is used the determination is not so easy. There are two methods of testing for oscillations; *first*, by making the coupling of the tickler coil (or other type of coupling) so weak that the circuit is not generating oscillations and then gradually increasing the coupling past the critical value, listening for the characteristic noise which occurs when the critical coupling is exceeded and, *second*, by properly interpreting the noises heard in the receivers when the grid terminal is grounded by putting the thumb or one finger, on the negative end of the filament circuit and touching the grid terminal with another finger. These two schemes may be called the *coupling test* and *finger test*.

As has been noted above when a tube circuit starts to oscillate the plate current practically always changes its average value, generally increasing when no grid condenser is used. The change in the plate current is not extremely rapid because, with the critical value of coupling, it takes many cycles before the steady state is reached; the result of the slow change in plate current is to produce a peculiarly soft quality of click in the receiver.[1] This noise resembles, perhaps more than anything else, the "plucking" of a loose violin string and, when once noted, is very easy to recognize.

In the case of no grid condenser this coupling test is very reliable and easy to make. If grid condenser is used the distinctness of this plucking sound is by no means as pronounced as is the case for no grid con-

[1] When listening for this noise the coupling must not be increased too slowly; with a very slow increase in the coupling the noise is so soft that it may not be heard at all. When first listening for this noise the tickler coupling should be changed quite rapidly.

denser; for some values of capacity and leak resistance it is almost impossible to hear it at all, even though the critical coupling is known and especial care is used in listening.[1]

In the case of no grid condenser the finger test gives very distinct indication of the oscillating condition; with the moistened thumb placed on a filament connection (binding post) a finger is touched to the grid connection of the tube, thus grounding the grid to an extent sufficient to stop oscillations.[2] The cessation of oscillations is accompanied by a sharp click in the receivers and when the finger is removed from the grid connection the starting of oscillations, with accompanying change in plate current, is indicated by another click, generally less distinct than the first. For coupling of the tickler coil considerably in excess of the critical value, the two clicks (starting and stopping oscillations) are of about the same intensity.

With grid condenser and leak the finger test does not give reliable results, except to the experienced operator; even with no oscillations two clicks are heard when the finger is touched to the grid connection and when it is removed therefrom. With the tube not oscillating the grid is practically always positive, with respect to the potential of the negative end of the filament; when the grid is grounded by the finger, thus suddenly bringing it to the same potential as the filament,[3] a sudden change occurs in the plate current with resultant click in the receiver; when the finger is removed the grid at once resumes its normal positive potential and so again gives a change in plate current and click in the phones. As has been previously noted, when grid condenser is used the grid leak resistance is best connected by the positive end of the filament; such has been assumed in statements just made.

The same two clicks are observed if the tube is oscillating, and there is not much difference between the clicks in the two cases. This is especially true if the grid condenser is small and electron supply in the vicinity of the grid plentiful; if, for example, with an ordinary detecting tube the grid condenser is 100 $\mu\mu f$ (a commonly used value) and filament temperature normal, even a good operator may not distinguish any difference in the clicks for the oscillatory and non-oscillatory condition.

[1] The distinctness of the noise depends upon the rapidity of change in plate current; if a large condenser is used it charges slowly and hence the change in plate current is slow, with corresponding indistinctness in the sound in the telephones.

[2] On most receiving sets it will be found that, even though the grid connection directly at the tube is not accessible, some screw or binding post connected to the grid, is available.

[3] It may possibly happen that when the tube is oscillating the average potential of the grid is the same as the negative end of the filament; in this case no change in the average value of the plate current occurs when the grid is touched and so no noise is heard in the phones.

If, however, the grid condenser is much larger, say 5000 $\mu\mu f$ or larger, there is a marked difference to be noticed; with oscillations the two clicks have nearly the same intensity, but with no oscillations the click heard upon removing the finger from the grid connection is much softer than the one heard when making contact with the grid. When the tube is not oscillating it takes an appreciable time to charge the grid condenser to its normal potential and the accompanying change in plate current is slow, thus giving a weak sound; the larger the grid condenser and the lower the filament temperature, the longer will this charging time be and correspondingly weaker is the click in the receivers.

The uncertainty of the finger test, in the circuit employing grid condenser, disappears, however, if, *instead of touching the grid side of the grid condenser the other side of this condenser is touched with the finger.* This side of the grid condenser (the one not connected to the grid) has the same average potential as the filament, because it is connected to the filament through the inductance of the tuned input circuit. Hence with no oscillations the potential conditions of the circuit are not disturbed when the condenser plate and filament are connected by the hand—they are already at the same average potential. With oscillations, however, the oscillations are stopped by the finger so that the plate current is changed and a click is heard.

From this it is seen that, with a grid condenser, by applying the finger test to the non-grid side of the condenser the test shows up the condition of the circuit in the same way as though no condenser was used.

The tests for the oscillating condition can then be summarized as follows:

Coupling Test.—*No Grid Condenser.*—Distinct sound (plucking string) when critical coupling is exceeded.

With Grid Condenser.—The click occurring when critical coupling is exceeded is not distinct especially when the grid condenser is large (several millimicrofarads) and the filament temperature subnormal.

Finger Test.—*No Grid Condenser.*—Two distinct clicks when tube is oscillating and none at all when tube is not oscillating.

With Grid Condenser.—Two distinct clicks of nearly equal intensity if tube is oscillating; if tube is not oscillating the click upon touching the grid connection is more pronounced than that when releasing the grid, the distinction being more pronounced with larger grid condensers. In case the non-grid side of the grid condenser is touched, instead of the grid side, the test is very reliable; no click at all if the tube is not oscillating, and a click if the tube is oscillating.

Peculiarities of Adjustment of Oscillating Detectors.—When first working with oscillating detectors certain apparent discrepancies will be encountered. Thus if the tuned grid circuit uses one of the coils of a loose coupler and the other coil of the coupler, or a section of it, is used for the tickler coil, it may be found that when the coils are separated, oscillations occur, *no matter which way the tickler coil is connected in the plate circuit.* It may also be found that oscillations occur when the coils are quite widely separated and that as the coils are brought nearer together the oscillations cease, an apparent contradiction to the analysis previously given.

With the coils arranged as shown in Fig. 187, it is apparent that the magnetic coupling of L_1 and L_2 is weak, but it may well be that the two coils of the coupler permit enough electrostatic coupling of the plate and grid circuits to produce oscillations, and this even if the connection of L_1 is reversed. Now if the sense of the magnetic coupling of L_1 and L_2 is incorrect for producing oscillations, the electrostatic coupling of the two circuits will be neutralized as the two coils are brought closer together, and when they get close enough, the coupling due to both effects will be less than the critical value and so oscillations will stop. In case the tickler coil consists of only a few concentrated turns this effect will not be noticed.

When the coupling of plate and grid circuits is accomplished by rotating one coil inside the other, it will often be found that setting the coils at right angles to one another, which of course makes $M = 0$, will not stop oscillations and that the coils must be rotated considerably past the 90° point before the oscillations stop. This is because of the electrostatic coupling introduced by the proximity of the two coils; enough reversed magnetic coupling must be introduced so that the total coupling, inductive plus capacitive, is less than the critical value for the circuit. This effect is mentioned, and analyzed on p. 610 et seq.

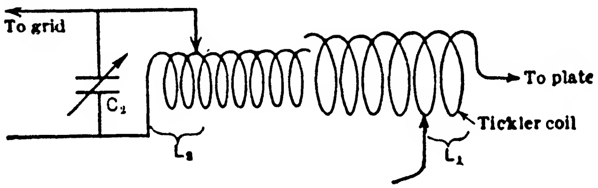


FIG. 187.—If an ordinary coupler is used in making tests for oscillations some peculiar results may be obtained.

Peculiar Noises Occurring in an Oscillating Detector Circuit.—If the oscillating detector circuit has no condenser in series with the grid its behavior is very regular, but if a grid condenser is used all sorts of queer noises may be heard in the phones, unless the adjustment is carefully carried out. The noise may vary from a series of regular "clicks," separated from each other by several seconds, to a high shrill signal; on carrying out further adjustments, the note may become so high as to be inaudible, so that the operator has no convenient way of telling that the action of the tube is irregular and that readjustment is required.

The condition practically always occurs as a result of too tight coupling of the tickler coil, too high a resistance for the grid leak, or a combination of both. The noise is due to *the starting and stopping of oscillations*, the musical pitch having nothing to do with the frequency of oscillation, but being fixed by the rapidity with which one group of oscillations follows the next.

The oscillations start, thus charging the grid condenser and reducing the mean potential of the grid and so changing the R_p of the tube; but the condition for oscillation for the circuit given in Eq. (116) depends upon R_p, and it is evident from inspection of this equation that if R_p increases, the value of M required for oscillation is increased. In Fig. 188 is shown the relation between E_p and I_p, for two values of E_{og}; the curve OA is for $E_{og}=0$, and the curve DB is for E_{og} at some negative value. The slope of this curve serves as a measure of R_p, the value of R_p being actually given by the cotangent of the slope, when the scales for E_p and I_p are the same; if not the same, the value of R_p obtained by measurement in the curve must be multiplied by the ratio of the "volts per inch" of this graph to the "amperes per inch."

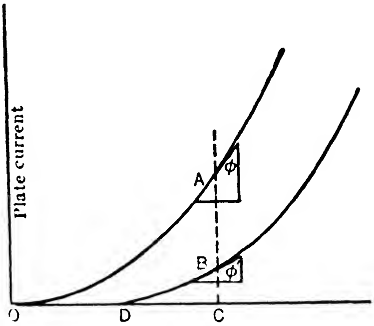


FIG. 188.—When oscillations start, in a circuit using a condenser in series with the grid, the plate-current curve may change from OA to DB, due to the decrease in average potential of the grid, when oscillations start.

Let us suppose that when oscillations start the normal potential of the grid is zero, the plate voltage being given by OC, Fig. 188; the value of R_p is then tan ϕ, and M is adjusted to such a value that oscillations start. The grid is then forced negative so that the E_p I_p curve changes from OA to DB, thus increasing the value of R_p to tan ϕ', and so increasing the coupling requirement, as given by Eq. (116), that the value of M is not sufficient to maintain oscillations, the circuit then stops oscillating.

During the oscillations, however, the grid condenser has become charged, and before oscillations can again start the charge must leak off sufficiently to bring the plate current from CB to CA, Fig. 188. The time required is fixed by the magnitude of the charge on the condenser and the time constant of the grid-condenser, grid-leak circuit. The adjustment might be such, for example, that 90 per cent of the charge in the condenser must leak off before oscillations again start. If $(1-\epsilon^{-\frac{t}{RC}})$ is to be 0.90, we must have $t/RC=2.3$; if then $C=500\ \mu\mu f$ and $R=2$ megohms, we have $t=2.3\times 510^{-10}\times 2\times 10^{6}=0.0023$ second. The starting and stopping of oscillations in the circuit would then occur about 500 times a second and a musical note of 500 vibrations a second would be heard in the telephone receivers.

If the leak resistance is greater in value, or the condenser of greater capacity, the note will be of lower pitch, and it will be higher if either the leak resistance or capacity is decreased. When the terminals of the vacuum tube are well insulated from one another and no external grid leak resistance is used, it is possible to so adjust a tube circuit that the interval between successive groups of oscillations is a minute or more, thus producing a series of clicks in the telephone receivers separated from each other by that interval of time.

The pitch of these disturbing noises may be practically always sent beyond the audible limit by lightly touching the grid connection and filament connection of the oscillating tube; if the finger and thumb making the connection are pressed down too tightly the leak resistance will be lowered to such an extent that the tube will stop oscillating altogether. The pitch of the note may be varied by changing either the plate voltage or filament current, both of these having influence on R_p and thus on the critical value of the coupling M; they also affect to some extent the grid leak resistance.

The squealing noise will nearly always be produced if, after the proper value of M has been obtained for a certain setting of the tuning condenser C (Fig. 159) the capacity of this condenser is much decreased. Decreasing C increases ω and so, according to Eq. (116), makes a lower value of M permissible; with the ordinary detecting-tube circuit, having grid condenser, it is practically always necessary to use the lowest value of M compatible with the requirements of Eq. (116) if steady oscillations are to be produced. A value of M much greater than this will not only cut down the sensitiveness of the tube as a detector, but is always likely to produce noises.

In general if the tuning condenser is decreased when the tube is giving off a squeal, the pitch of the squeal is diminished. This is due to the fact that with decreased C the excess of M over its critical value is increased. This means that when the oscillations start, they start with more violence, produce larger oscillatory currents in the L C circuit which of course take correspondingly longer to die out, after the triode has ceased to maintain the oscillation (due to increased R_p). These greater oscillatory currents result in a greater negative charge building up on the grid and this, of course takes correspondingly longer to leak off. Hence there are fewer groups of oscillations per second and the note is lowered.

Use of Regenerative Circuit for Spark Reception.—A tube circuit arranged with "tickler" or other form of coupling for the detection of continuous-wave signals is also adapted for the reception of spark, or damped-wave, signals; with the antenna circuit and the local circuit (L_2–C

of Fig. 182) tuned accurately to the incoming signal the tickler coupling can be increased to a value slightly less than that required for producing oscillations. The intensity of the signal, by using a suitable value of coupling, can be increased hundreds of times over the value it would have if no tickler coupling were used.

It has been shown that the effect of the coupling is to reduce the resistance of the L_2-C circuit to a very low value (Fig. 185) so that a certain e.m.f. impressed on this circuit, from the antenna circuit, will produce a current perhaps 100 times as great as would normally be the case. The change in the plate current (which gives the signal in the phones) is proportional to the square of the voltage impressed on the grid, as given in Eq. (24), and so will increase greatly as the resistance of the L_2-C circuit is made to approach zero by suitable tickler coupling. If e.g., the actual resistance of L_2-C is 10 ohms and by means of tickler coupling the effective resistance is reduced to 0.1 ohm, the current in L_2-C is increased 100 times, the voltage impressed on the grid is increased 100 times, and the signal current, ΔI_p is increased 10^4 times.

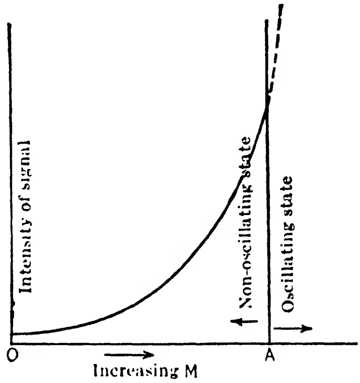


FIG. 189.—If the tickler coil is used when receiving spark signals (of sufficiently low decrement) the signal will increase very rapidly as tickler is increased until this passes its critical value; the tube starts to oscillate and then the signal, although very loud, loses its characteristic musical note and becomes "mushy" in quality.

The effect on the signal strength as the mutual inductance between L_1 and L_2 (Fig. 182) varies is shown in Fig. 189; as M is increased the signal intensity rapidly increases, retaining its normal musical quality, until such a coupling is reached, OA, that oscillations start. The resulting noise in the telephone when the tube is oscillating, is of "scratchy" quality being caused by a kind of beat phenomenon between continuous waves locally generated and the incoming damped waves; as the phase relations between the successive wave trains and the continuous oscillations of the tube are of haphazard values, and as the amplitude of the spark signals is variable throughout each wave train, the resulting variation in amplitude of the plate current is of very irregular character, thus producing the scratchy note for couplings indicated by the dotted line in Fig. 189.

Amount of Amplification by Regeneration.—The amount of amplification available before the triode starts to oscillate has been investigated by van der Pol,[1] who shows that weak signals may be amplified much more than strong ones. He uses the term "grid saturation volts," V_{g_0}, to indicate the change in grid voltage necessary to bring the plate current from zero to saturation value, V_{g_1} to mean the grid volts with no regeneration, and V_{g_2} to mean the grid volts with critical regeneration, and shows that

$$\frac{V_{g_2}}{V_{g_1}}=\left(\frac{V_{g_0}}{V_{g_1}}\right)^{2/3}. \qquad (119)$$

Using a certain detector tube, and a receiving circuit having a ratio $\omega L/R$ of 40, he assumes various signal strengths induced in the grid circuit, calculates the grid volts with and without regeneration, and finds:

Signal Volts Induced in Grid Circuit	Volts across Grid With No Regeneration	Volts across Grid With Critical Regeneration	Voltage Amplification Due to Regeneration
10^{-6}	4×10^{-5}	0.31	7700
10^{-5}	4×10^{-4}	0.66	1600
10^{-4}	4×10^{-3}	1.4	360
10^{-3}	4×10^{-2}	3.1	77
10^{-2}	4×10^{-1}	6.6	16

Regenerative Circuit for Short-wave Spark Reception.—For short-wave reception, say less than 400 meters, probably the most satisfactory type of circuit is one which uses no other coupling between the grid and plate circuits than that due to the capacity coupling in the tube itself. In this scheme the "tickler" coil of Fig. 182 is replaced by a small variometer, not coupled magnetically to the L_2 C circuit at all; the required amount of inductance in this variometer varies with the wave length, type of tube, etc., but is generally less than 1 millihenry. It is best to add in the L_2-C circuit another variometer about the same as that used in the plate circuit, thus making it possible to tune the closed circuit with very small value of C.

The L_2-C circuit is carefully tuned to the incoming signal and the regenerative action is brought to its maximum permissible value by suitably adjusting the variometer in the plate circuit. As the plate inductance is increased, a slight further adjustment of the closed tuned circuit is generally required in order to get maximum sensitiveness.

Behavior of a Regenerative Receiver Regarding Sound of Signal, etc.—There are many interesting phenomena connected with the adjustments of this regenerative circuit other than those already mentioned. When M is made *just great enough* to produce oscillations (slightly greater

[1] I.R.E., Feb., 1929, p. 339.

than OA, Fig. 189) the detecting efficiency of the circuit is greatly increased, so much so that spark signals so weak as to be entirely inaudible with tickler coupling just less than the critical value become quite loud when oscillations start. In this case the listening operator gets no clue to the identity of the sending station from the spark note because the signal is inaudible until the tube is oscillating, and then the distinctive spark note is not present.

If such a weak signal is coming in and the closed circuit is properly tuned to it (with tickler coupling about equal to its critical value), a peculiar effect is produced by the adjustment of the antenna circuit. With this circuit much detuned of course the signal is inaudible; as the antenna loading coil, or similar adjustment, is increased, the signal becomes audible with a scratchy quality (tube supposedly oscillating); as this adjustment is continued the signal increases in intensity until at a certain value of loading it disappears completely; on continuing the adjustment, however, the signal reappears at a certain point and gradually decreases as the adjustment is further carried out. Upon investigation it will be found that this narrow region where the signal is inaudible is caused by the *cessation of oscillations in the tube circuit.* The antenna circuit introduces a resistance effect into the oscillating tube circuit which varies with the relative tuning of the two, being a maximum when the antenna and closed circuit are tuned alike; hence a tickler coupling which is just sufficient to cause oscillations with an antenna somewhat mistuned is insufficient for the tuned condition. A quantitative idea of this change of resistance of the oscillating circuit, due to variation in antenna tuning is given in Fig. 123, Chapter I.

For the best reception of the signal the adjustment of the antenna should be set at the midpoint of the silent region and the tickler coupling increased just sufficient to produce oscillations for this condition.

In case a continuous-wave signal is being received, the following effect of the antenna tuning on the reactance of the oscillating circuit may be noted. With antenna and closed circuit normally adjusted, a certain note is heard in the telephone; this note may be observed to vary over a considerable range as the tuning of the antenna is changed, *the variation in note being caused by the change in the effective inductance in the closed oscillating circuit by the reaction of the antenna.* The amount of change in note obtainable depends upon the coupling between antenna and closed circuit; some idea of its magnitude may be had by inspection of Fig. 123, Chapter I.

Equations (108) and (109), p. 125, permit quantitative prediction of the amount of change in resistance and reactance of the oscillating circuit, as affected by the antenna circuit.

The foregoing ideas in regeneration have been presented from a very elementary viewpoint. For those especially interested in this phase of the triode behavior we refer to an excellent article dealing with the question in a more exact fashion, by Landon and Jarvis.[1]

Operation of Power Tubes in Parallel for Greater Power Output.—Vacuum-tube generators or converters operate very well in parallel, remaining synchronized automatically;[2] the proper division of the load may be most easily accomplished by variation of the filament currents. All filaments may be lighted in parallel from the same source, but each filament should have its own rheostat; it is also best to have an ammeter in series with each plate and grid. A suitable connection scheme is shown in Fig. 190; the same scheme of connection can be used for any number of tubes.

The adjustments of the circuit must of course be changed as more tubes are put into operation, because the effective resistance of the load must equal the plate-circuit resistance of the battery of tubes for maximum output, and the combined tube-circuit resistance varies inversely as the number in operation.

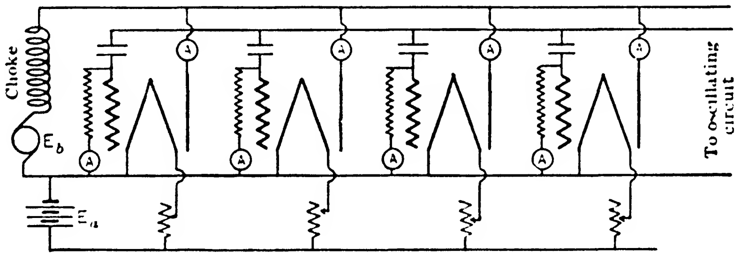


Fig. 190.—Connection of several power tubes for parallel operation.

The voltage required for the filament of the ordinary power tube is about twenty; it is limited by the fact that too much power must not be used in the filament circuit, yet the filament current must be fairly large because the plate current must not be more than about 15 per cent of the filament current, and this plate current must be a considerable fraction of an ampere or more unless excessively high voltages are used.

[1] "Analysis of Regenerative Amplification," Proc. I.R.E., Vol. 13, No. 6, Dec., 1925

[2] An interesting demonstration of the inherent tendency of tube circuits to synchronize with each other is easily obtained. If a small power tube is set into oscillation in the laboratory and an autodyne detector circuit in the same room is used for listening, it will be found that as the beat note is decreased from high value there will be a certain lowest audible note obtainable. Thus perhaps the detector adjustment is such as to give a beat note of 200; upon attempting to bring this detector more nearly into synchronism with the power tube, lowering the beat note, this note will completely disappear, and it will seem as though the autodyne had stopped oscillating, but it will be found that the beat note has disappeared because the detector tube has *pulled into synchronism* with the power tube. The closer the two circuits are together the higher will be the lowest beat note obtainable.

As the ordinary electrical power supply is 110 volts, it might seem that if tubes are to operate in a group several filaments might be connected in series, thus saving in power consumption; thus five 20-volt filaments might be operated on a 110-volt line and still leave enough voltage for a control rheostat. *But such a connection of filaments is impossible.* The filament current of each tube would differ from that of the next one in series by an amount equal to the plate current of that tube as shown in Fig. 191. If each plate current is 0.3 ampere, the currents in the filament circuit would be as shown, but such a condition is impossible because a filament which will safely carry 4.0 amperes (tube A) would have practically no electron emission with current less than 3.40 amperes, so that tubes C and D of the series would be dead in so far as electron current is concerned; the plate current of each of these tubes would be nearly zero instead of 0.3 ampere as shown. When oscillations start,

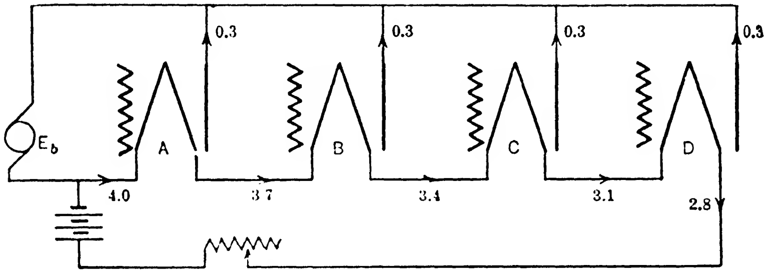


FIG. 191.—It is impossible to work tubes in parallel, with their filaments in series, because of the greatly different filament currents resulting in the different tubes.

the maximum value of plate current of tube A would be about 0.8 ampere, thus rendering even tube B more or less ineffective.

Use of a Separate Exciter for a Group of Tubes.—It is generally not possible to get as much power from a self-excited tube as from a separately excited one, and the adjustment for such a condition is critical; if it is used, the tube may stop oscillating when a slight drop in plate voltage or filament current occurs. This is a dangerous condition, because unless the operator notices at once that the tube is not oscillating the plates will rapidly become overheated and the tube perhaps spoiled. To avoid this contingency a separate exciting tube may be used, this tube furnishing only enough power to operate its own circuit and supply the losses in the grid circuits of the group of power tubes. Such a scheme is shown in Fig. 192; the exciter tube A is adjusted with tight coupling between L_2 and L_1, so that it oscillates under any condition which may occur, and the power tubes M, N, etc., are each excited by a common connection to tube A. By adjustment of the condensers C_1–C_2, etc., the output of each power tube may be controlled, this control being in addition to that afforded by the filament current. The frequency of the exciter circuit must of course be that required for resonance in the load circuit of these power tubes.

In case the individual control of the excitation is not desired a common adjustable condenser may be inserted in the exciter lead where indicated by the dotted lines at X; this condenser should have a reactance about equal to the impedance of the combined input circuits of the power tubes.

In the early radio-telephone experiments the transmitter consisted of hundreds of tubes all operating in parallel. With the advent of the water-cooled tube, however, the necessity and advisability of such an arrangement disappeared. At the present time not more than two triodes in parallel are advisable and preferably only one is used. The water-cooled

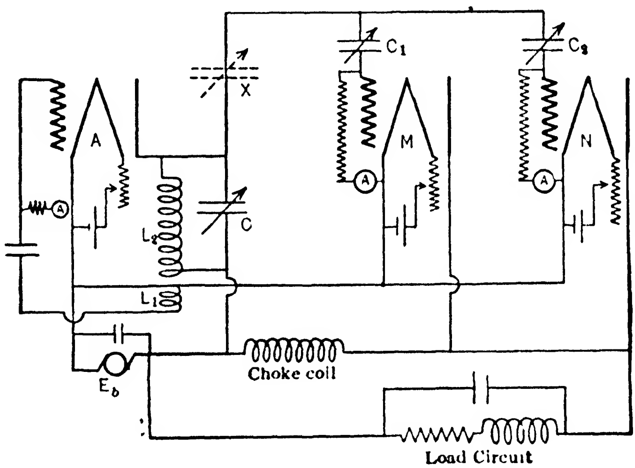


FIG. 192.—When many tubes are to operate in parallel it is generally best to excite them from a separate tube A, self-oscillating, controlling the amount of excitation by condensers, C_1–C_2, etc.

triode, of any capacity up to 100 kw. or more, is perfectly feasible and practicable, so the necessity of parallel operation does not exist.

Another scheme for using an exciter tube for maintaining the power tube in oscillation is shown in Fig. 193. In this case the exciter tube is not a self-exciting unit, but operates in conjunction with the power tube; the frequency of output is determined entirely by the L–C circuit of the power tube.

The amount of excitation furnished to the grid of the exciter tube depends upon the relative magnitudes of C_2 and C_3; ordinarily C_3 should be many times as great as C_2. The value of the resistance R may vary widely, a suitable value being equal to the resistance of the plate-filament circuit of the exciter tube.

This arrangement is a very useful one if it is desired to vary the frequency of the output circuit over a wide range; in a typical case the frequency of the output circuit was varied (by changing L and C) from 500 to 300,000 cycles per second without changing the adjustment of the exciter tube and a wider range could have been covered without any other adjustments than those of L and C, had it been so desired.

Power Required for Excitation.—From the characteristics of triodes already analyzed it is evident that, if the grid is always maintained at a potential more negative than the cathode, practically no grid current will flow, and of course, if no grid current flows no power is used in the grid circuit. Of course the circuit used to impress the voltage on the grid may be carrying current and so use power, but the grid itself will use none.

Now to get much power from a triode used as an oscillator, or amplifier, it is necessary to use such excitation on the grid that for a short part

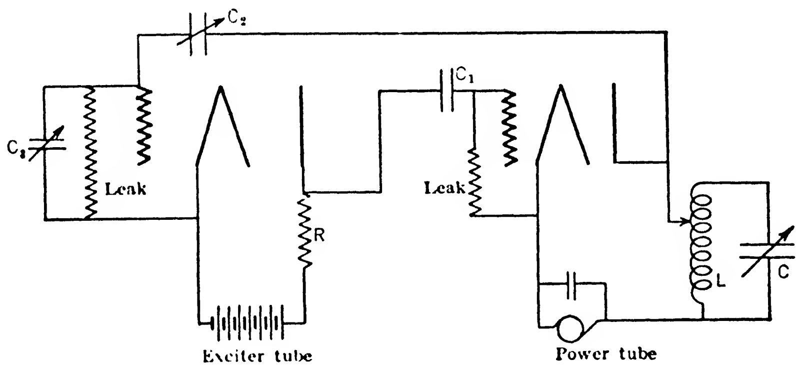


FIG. 193.—A scheme for using an untuned exciter tube; this scheme is a good one if the set is to oscillate with very wide variations in the values of L and C.

of the cycle the grid swings positive with respect to the cathode, and of course it will then draw current. This current, made up of electrons leaving the cathode and impinging on the grid, represents power; all the kinetic energy which the electrons gain in falling from the cathode to the grid is given up as heat on the grid. It frequently happens that this current has to flow through a resistance of some sort, to return to the cathode, and more power is used up here.

Spitzer[1] has analyzed and measured the amount of power used on the grid itself, for various types of power tubes; the measurements were made

[1] I.R.E., June, 1929, p. 985.

with a suitable wattmeter. Some of his results are tabulated below, the amount of grid excitation, grid bias, grid power, etc., being given for full rated output of the tube.

Type	E_b	I_b	E_c	I_c (10^{-3} Amp.)	E_g	L/RC of Load	Watts Output	Plate Circuit Efficiency, Per Cent	Watts for Excitation	$E_b/2I_b$	Excitation Loss Constant, A
211	1000	0.175	−100	19	160	3100	122	70	3.9	2860	0.039
203A	1000	0.175	−75	35	145	3100	123	70	6.7	2860	0.036
204A	2000	0.275	−175	131	328	4030	405	74	56.0	3640	0.053
849	2000	0.350	−200	57	240	2860	562	80	29.0	3500	0.029
851	2000	0.900	−200	300	274	979	1420	79	111.0	1110	0.024
861	3000	0.35	−200	40	455	4300	665	64	22.5	4290	0.107

He finds for each tube a relation between power lost on the grid, and the grid current as read on a c.c. ammeter, as follows:

$$\text{Watts loss} = A I_c^{1.34} \qquad (120)$$

in which I_c is the grid current in milliamperes and A is a constant as shown in the table above. A varies from 0.02 to 0.10 for various tubes of a kilowatt or less capacity.

It can be seen from the above table that a triode requires an excitation loss (for full power output) of from 5 to 15 per cent its rating. When the

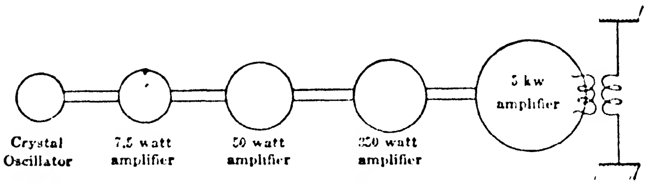


FIG. 194.—Schematic diagram of train of tubes required for a crystal-controlled 5Kw. transmitter.

power used in the tuned plate circuit of the exciter (circuit L_2-C of Fig. 192) is added to that lost on the grid of the excited tube the total is seldom less than 10 per cent of the output rating of the excited tube. This means that to excite a 50-watt tube a 5- to 10-watt tube is required; to excite a 250-watt tube a 50-watt tube is required, etc.

Thus a crystal-controlled 5-kw. transmitter might use a train of tubes about as shown in Fig. 194. The 7.5-watt tube would be neutralized[1] to prevent a variation of its output from reacting on the crystal oscillator. One of the tubes will have a potentiometer feed from its plate circuit

[1] See p. 1016 for explanation of neutralized triode.

to the grid circuit of the following tube, for controlling the output of the 5-kw. tube.

Special Arrangement of Triode.—In a special form of three-electrode tube, apparently first advocated by A. W. Hull and called by him the dynatron, the phenomenon of *secondary emission* is utilized. If an electron traveling at high speed collides with a metallic surface, the giving up of its energy at the surface is likely to " jar " other electrons out of the metal at the point where the collision occurs; the emission of the electrons from this surface, caused by the colliding electron, is called *secondary emission.* The number of electrons emitted depends upon the speed of the colliding electron; it may be none at all and may be as much as a dozen or more.

The speed with which the impinging electron must strike the metallic surface to get secondary emission depends greatly upon the condition of the surface. Thus tungsten with a clean surface requires the speed given to an electron by the fall through 180 volts difference of potential before another electron is splashed out but if there is a layer of gas covering the tungsten (adsorbed) secondary electrons may appear when there is only 30 volts difference of potential between the filament and plate.

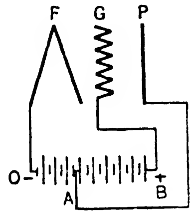


FIG. 195.—Connection of a three-electrode tube to get the characteristics of the dynatron.

Ordinarily, these electrons, due to secondary emission, will at once reenter the surface from which they have been emitted, but, if there happens to be in the vicinity of the surface an electrode of higher potential, these secondarily emitted electrons will not reenter the surface from which they came but will go to the higher potential electrode, thus causing electron current *away* from the surface to which the first electron is traveling.

The number of electrons taking part in this reversed current depends upon the number caused by the secondary emission and upon the potential of the surface attracting them. Suppose the arrangement of electrodes as given in Fig. 195; the grid is at higher potential than the plate and so attracts most of the electrons caused by the normal thermal emission from filament F. However, some of these electrons will go through the interstices of G and impinge on P, causing secondary emission where they strike. As G is at higher potential than P, the electrons due to secondary emission are likely to go to G instead of reentering P.

If the potential of G is held constant (contact B remaining fixed) and the potential of P is gradually increased from zero by moving contact A to the right, the various happenings will be about as shown in Fig. 196, Curve O–A shows the electron current to P due to emission from F; curve O-B shows the amount of secondary emission from P, due to electrons of current OA; curve C shows the fractional part of the secondary emission which is attracted to G; curve O-D shows the electron current away from P due to secondary emission, and curve $OEFGH$ shows the actual electron current to P, all of these curves being plotted for increasing plate potential.

The peculiarity of that part of the curve from E to G is the basis of action of the dynatron; an increasing plate potential results in a decrease in plate current, in other words, an a.c. test of the resistance of the plate-filament circuit in this region of operation would show a negative resistance.

The dynatron has thus practically the same characteristics as an ordinary three-electrode tube with the regenerative connection of plate and grid circuits, and it may be used for similar purposes.

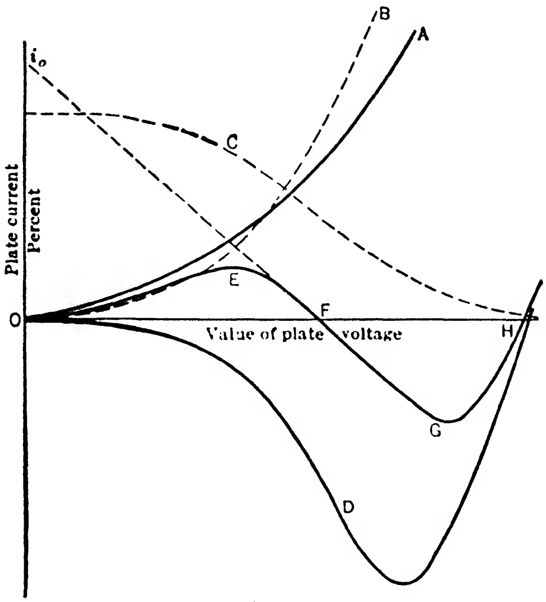


FIG. 196.—Curves of various currents occurring in the operation of the dynatron. Of course only the actual current, $OEFG$, is normally measurable.

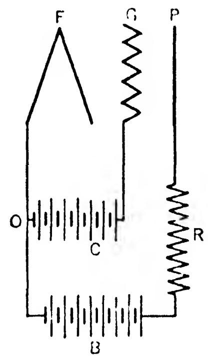


FIG. 197. Connection of three-electrode tube as a dynatron.

The current curve of Fig. 196, between points E and G, can be expressed by the equation

$$i = i_0 - \frac{v}{r}. \qquad (121)$$

where i = plate current;

i_0 = value of plate current obtained by projecting the curve GFE back to $v=0$ as shown in Fig. 196;

r = internal resistance of the tube, determined from the slope of the GFE curve.

Transposing the terms of Eq. (121) we have

$$v = r(i_0 - i).$$

If the voltage of the battery B (Fig. 197) is E and drop across the resistance R is V, then

$$E = V + v = Ri + r(i_0 - i) = (R - r)i + ri_0 = \frac{(R-r)V}{r} + ri_0.$$

Then

$$\frac{dV}{dE} = \frac{R}{R-r}. \qquad (122)$$

As $R-r$ may be made small, it is evident that a small increase in E, the voltage used in the plate circuit, may result in a much larger change in the voltage drop across R. It has been possible to regulate the tube so that an increase of 1 volt in E has resulted in a change of the potential difference across R of 100 volts, thus giving a voltage amplification of 100 times.

The dynatron may be used as regenerative detector, oscil-

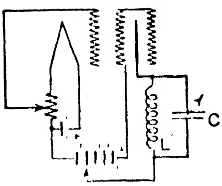


Fig. 198.—Connections of the plio-dynatron.

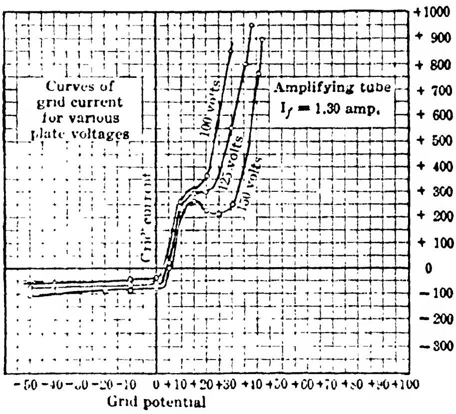


Fig. 199.—Dynatron characteristics in an ordinary telephone repeater tube.

lating detector of continuous waves, or as a generator of a.c. power just as can the ordinary three-electrode tube; it is not evident, however, that it has any advantage over the three-electrode tube as ordinarily used.

It is possible to use the secondary emission effect, that is, the dynatron principle, and add a normal control grid as well. The ordinary screen grid tetrode is such a tube, and by operating it in a region where the plate current decreases as the plate voltage increases (i.e., in the region of $E_p = 15$–40 volts of Fig. 234, p. 689) the dynatron characteristics may be obtained in addition to the action of the normal control grid. Such a connection of a screen grid tube is shown in Fig. 198; this scheme is frequently used as a source of known adjustable frequency. The frequency of oscillation of the L–C circuit is more independent of changes in tube parameters (filament current, grid bias, etc.) than any of the other oscillating circuits shown in the previous sections. By using a carefully calibrated L–C circuit, the device may be used as a laboratory frequency standard, good to much better than 1 per cent.

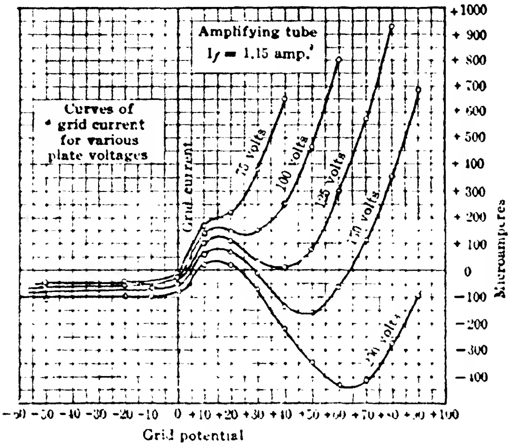


FIG. 200.—Dynatron characteristics in an ordinary telephone repeater tube.

The special forms of plate current curve for the dynatron given in Fig. 196, may be duplicated to some extent by any three-electrode tube; in Figs. 199, 200, and 201 are shown the curves of grid current of an ordinary telephone amplifying tube operated outside its normal range. This tube normally operates with a negative grid, but by carrying the grid through sufficiently high positive potentials the form of its current is made to resemble that of the dynatron very closely. The tube was not pumped to as high a vacuum as are the dynatron and pliotron, so that there was more gas present in this tube, but the regularity of the curves and the fact that they could be duplicated as many times as desired shows that however much gas there was present, it was probably playing a minor role in the action of the tube.

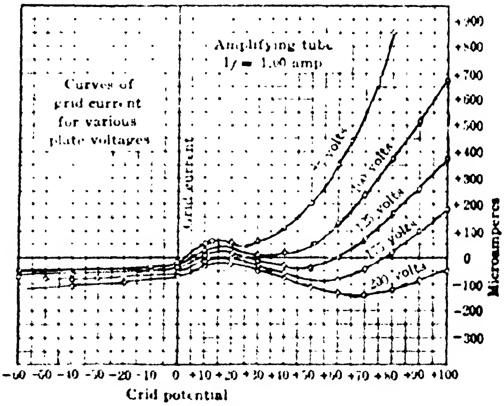


FIG. 201.—Dynatron characteristics in an ordinary telephone repeater tube.

Detailed Study of the Three-electrode Tube as a Power Converter.—The foregoing analyses of the conditions for oscillation of a three-electrode tube have all been based on the assumption that the plate current in the oscillatory condition could be sufficiently well represented by a constant current with a sine-wave current superimposed, and on this basis we have shown that the theoretical maximum output of the tube was one-half of the input; the fact was also mentioned that the conditions demanded for this efficiency of 50 per cent could not be realized, so that we were forced to conclude that the maximum efficiency of a tube generator was about 40 per cent.

The author with the assistance of Mr. H. Trap Friis carried out a detailed study of the tube generator for both separate and self-excitation,[1] and it was found that the efficiency might become very much higher when the proper adjustments were made; part of the results of this study will be given here, as they show exactly how a tube functions. The notation used in this analysis is somewhat different from that used so far because the previous symbols are not applicable. The plate current cannot be represented by $I_{op}+I_{mp}\sin\omega t$, as has been previously assumed; it consists of a series of pulses so that an infinite series of sine terms would

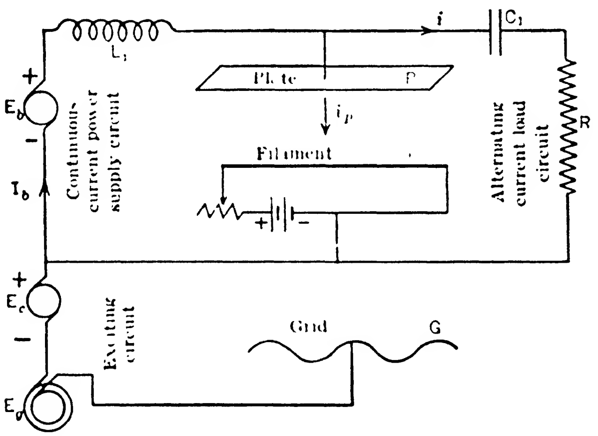


FIG. 202. Connection of power tube for study of its characteristics.

be required to represent the alternating component. The plate voltage also does not have exactly a sinusoidal variation. We therefore represent the instantaneous value of the *actual* plate voltage by e_p, grid voltage by e_g, plate current by i_p, grid current by i_g, etc., instead of representing each by a constant plus a sine term.

Oscillograms were taken to show the various quantities entering into the operation of the tube and circuit, the frequency of the alternating current being between 100 and 200 cycles; later the circuit constants were diminished sufficiently to raise the frequency to 100,000 cycles, to show that the results obtained at the lower frequency (which allowed accurate oscillographic records to be obtained) were valid at radio frequencies.

[1] Proceedings of A.I.E.E., Vol. 38, No. 10, Oct., 1919.

The first effect studied was the change in form of e_p and i_p as the excitation of the grid was increased, using a separately excited circuit as indicated in Fig. 202. The reactance of C_1 was 62 ohms and of L_1 was 8700 ohms; the value of R was 1000 ohms and the resistance of L_1 was 190 ohms. The μ of the tube used was 3.9.

With comparatively low values of E_c and E_g a record was taken of e_g, e_p, and i_p, and is given in Fig. 203; it is seen that the fluctuations in e_p and i_p were nearly sinusoidal so that the results of the previous analysis would hold good for this condition. Upon increasing e_g to six times its value the forms of e_p and i_p are made to differ widely from sine forms, however, as shown in Fig. 204.

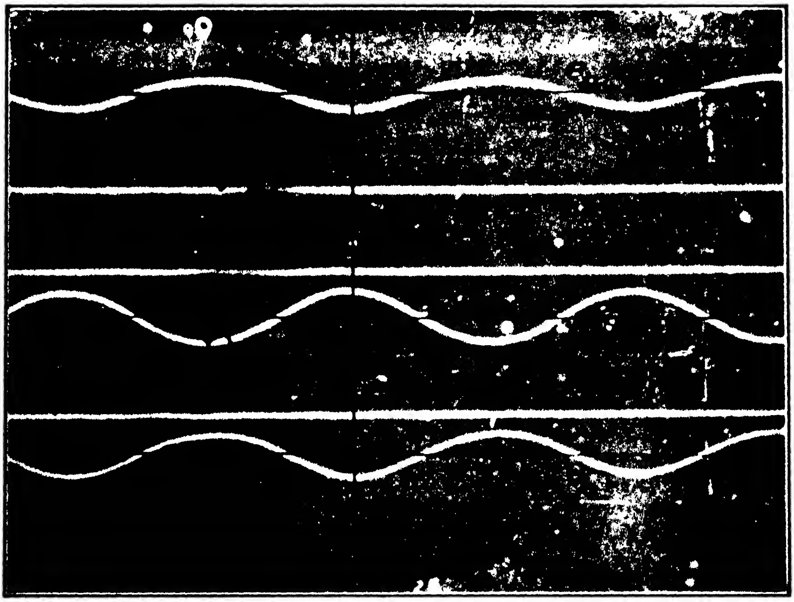


FIG. 203.—Nearly sinusoidal variations in e_p and i_p for low grid excitation. $E_b = 500$, $I_b = 0.25$, $E_c = 120$, $E_g = 50$, Frequency = 140, $R = 1000$, $C = 18.4$ microfarads.

An interesting point is shown by the film; the value of R used was 1000 ohms and this is the value which gives, for this tube, maximum output for low values of E_g, as shown in Fig. 128. This value 1000 ohms must therefore be the tube resistance R_p for the low value of E_g. But with large excitation used in Fig. 204 the plate current evidently fluctuated as much as possible (from zero to saturation current) and the fluctuation in e_p is less than half of E_b, indicating that R should be more than doubled if maximum output is to be obtained from the tube. This it will be remembered has been predicted as necessary when i_p fluctuates between zero and saturation current, and e_p fluctuates between the limiting values of zero and

2 E_{op}. With a resistance load of the kind shown in Fig. 202 it is evident that such a wide variation in the value of e_p is impossible; the load circuit must contain inductance and capacity to cause e_p to fluctuate so widely. This point is taken up later on in this section.

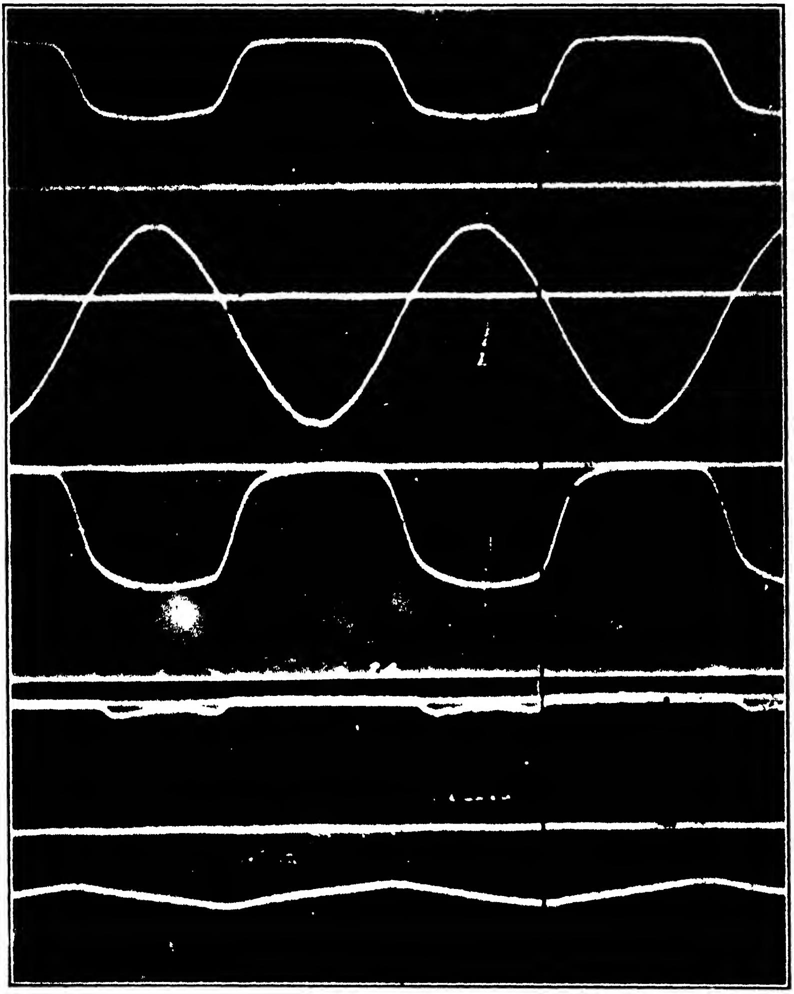


FIG. 204.—Distortions occurring with higher grid excitations. $E_p = 900$, $I_b = 0.34$, $E_c = -120$, $E_g = 300$, $f = 140$, $R = 1000$, $C = 18.4\ \mu f$.

If R is still further reduced the distortion in e_p and i_p will appear with much lower values of E_g; in Fig. 205 is shown a record for a value of E_g of 100 volts with R only 100 ohms. The fluctuation in e_p is now hardly noticeable although i_p fluctuates, with distorted form, from zero to saturation current as before. The current taken by the grid in Figs. 203 and 205 was zero; in Fig. 204 the grid swings positive 300 volts so we might expect a large grid current, but it is shown to be small. This is due to the fact that the plate is at rather high potential (650 volts) during the time the grid is positive, so that but few electrons go to the grid.

Fig. 204 shows also the fluctuation of I_b; in spite of the large inductance L_1 (which was 10 henries) there is considerable variation in I_b. This must of course always be the case; the value of $L_1\, di_b/dt$ must at any instant be equal to the difference between E_b and e_p.

In Fig. 206 are shown the curves of e_p, i_p, and e_g for two values of R, all other conditions being the same; it may be seen that the amount

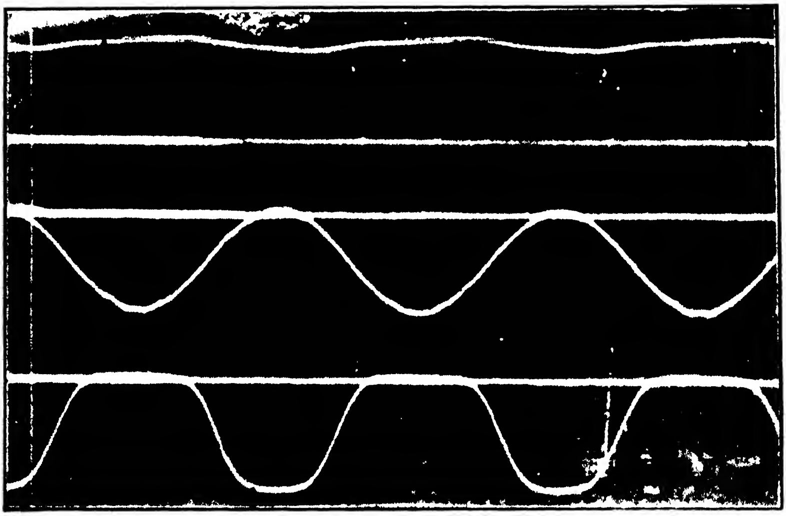


FIG. 205.—With low-load circuit resistance distortions occur for even low-grid excitation, $E_b = 900$, $I_b = 0.34$, $E_c = 120$, $E_g = 100$, $f = 140$, $R = 100$, $C = 18.4$ μf.

of distortion in i_p is reduced as the value of R is increased. In getting these two films the value of L_1 was kept constant, with the result that a larger percentage of the generated alternating current of the tube went through this path, with the higher value of R, instead of through the load circuit, C_1-R.

Attempts were then made to see what adjustments of the tube and associated circuits gave best efficiency; the importance of high efficiency will be at once appreciated when it is mentioned that a given tube (the one used in these tests) has an output of about 200 watts in normal operation whereas if the efficiency could be raised to 90 per cent, the safe output would increase to 2250 watts.

The tests carried out involved an adjustment with separate excitation to find the conditions for maximum output and then transferring the grid connection to a proper point of the circuit to get self-excitation, recording for each condition the forms and phases of currents and e.m.f.s. The tests were run at low frequency so that oscillograph records might be obtained; the results obtained were duplicated later in a high-frequency run.

Fig. 207 shows the circuit used; simpler ones may be used, but the laboratory apparatus at hand was best suited to this one. The diagram also shows where the oscillograph vibrators were introduced and the direction of currents assumed as positive; if, on a film, a current is shown

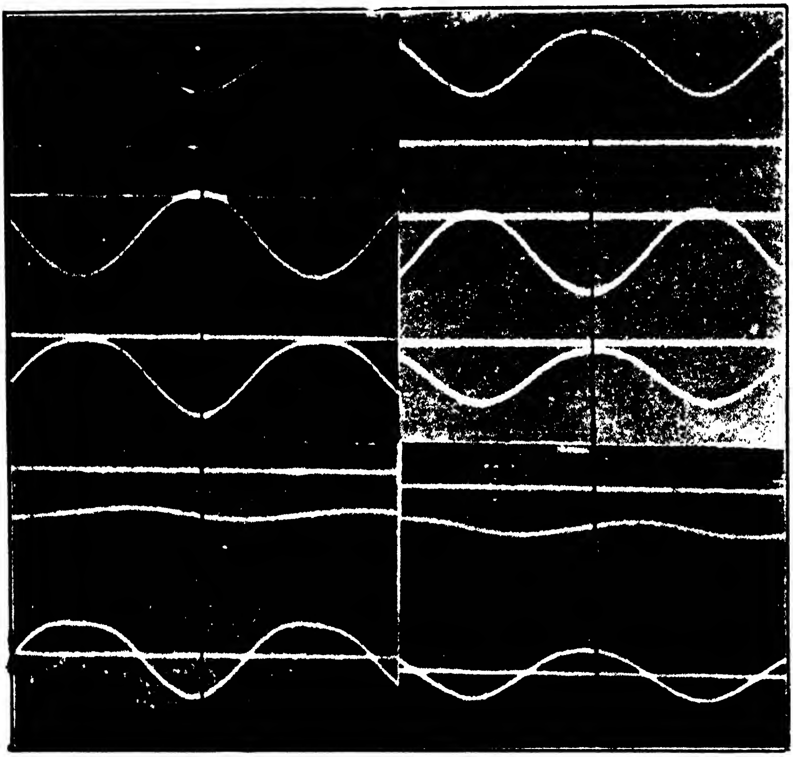


FIG. 206 Showing effect of load resistance on forms of voltage and current, other conditions constant. For both films $E_b = 900$, $E_c = 120$, $E_g = 100$, and $f = 140$. For left-hand film $I_b = 0.295$, $R = 1000$. For other, $I_b = 0.272$, $R = 2010$.

below its zero line, it was flowing in the opposite direction to that shown in the diagram. If the frequency of the exciting voltage, E_g, is chosen the same as the resonant frequency of the load circuit

$$f = \frac{1}{2\pi\sqrt{L\left(\frac{C_1C_2}{C_1+C_2}\right)}},$$

the impedance of this circuit between the two points M and N, where the tube is attached, will be resistive only, its magnitude being equal to $1/\omega^2 C_1^2 R$ ohms. This is calculated by the method given in Chapter I, p. 95.

The quantities to be considered are shown conventionally in their proper phases in Fig. 208; the current i_1, which flows in the resonant load circuit, may be several times as large as the current i, furnished by the tube. The two important things in this diagram are shown in the lower part of the figure, namely, the curves of $e_p i_p$ and $e_p i$. These curves

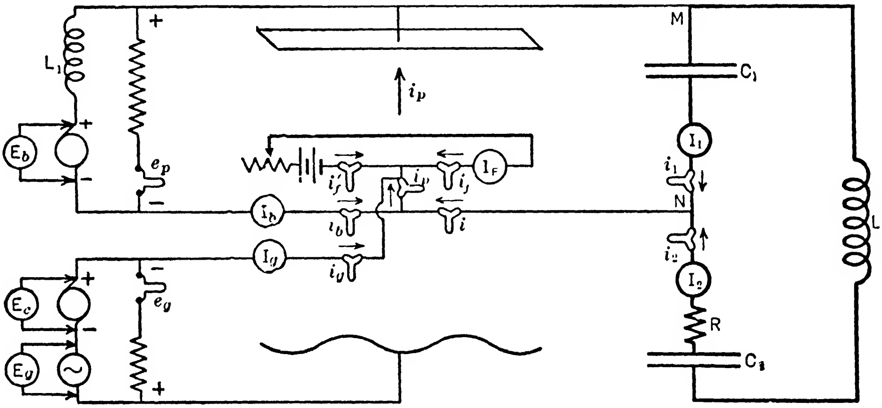


FIG. 207.—Connection of the power tube to a tuned output circuit, showing where oscillograph vibrators were introduced and directions of current assumed as positive (above zero line in oscillograms).

give the power loss on the plate and the power supplied by the tube to the load circuit, respectively. It is at once evident that

$$\text{Energy loss on plate per cycle} = \int_0^{2\pi} e_p i_p dt = \text{Area } A.$$

$$\text{Energy supplied to load circuit per cycle} = \int_0^{2\pi} e_p i dt = \text{Area } C - \text{Area } B.$$

It is evidently desirable to make the latter as large as possible and the former as small as possible, if the tube circuit is to operate efficiently. Any ordinary scheme of analysis, using the relation given in Eq. (7), p. 513, must fail because the relation does not hold good for those values of e_p and e_g, which are the most important ones in the cycle of operation, namely, low e_p with positive e_g, and very high values of e_p with large negative e_g.

The ordinary so-called static characteristics of the tube used are given in Fig. 209; they are not of much service in predicting the behavior of the tube when the output is forced as high as possible. They did bring out the fact, however, that the filament ammeter, if a c.c. instrument, does not read correctly the filament current when the tube is generating a.c. power. The ammeter indicated 3.65 amperes when getting the curves of Fig. 209 and the total emission for such a current is evidently about 0.5 ampere. Now when the tube was oscillating, the filament ammeter reading 3.65 amperes, the total emission was about 0.8 ampere, showing that the filament temperature was much hotter than when not oscillating. Holding the voltage across the filament constant (approximately the condition when the tube is oscillating) the set of curves given in Fig. 210 was obtained. The grid was held at a positive potential of 100 volts and the plate voltage suitably varied. The electron current to the plate increases the filament current at one end and decreases it at the other; the relative values of increase and decrease will be determined largely by the resistance used in series with the filament battery. It can be seen that even with the larger filament current as great as 3.75 amperes the emission was only 0.5 ampere.

From some preliminary oscillograph records we knew that in operation the total emission was about 0.8 ampere

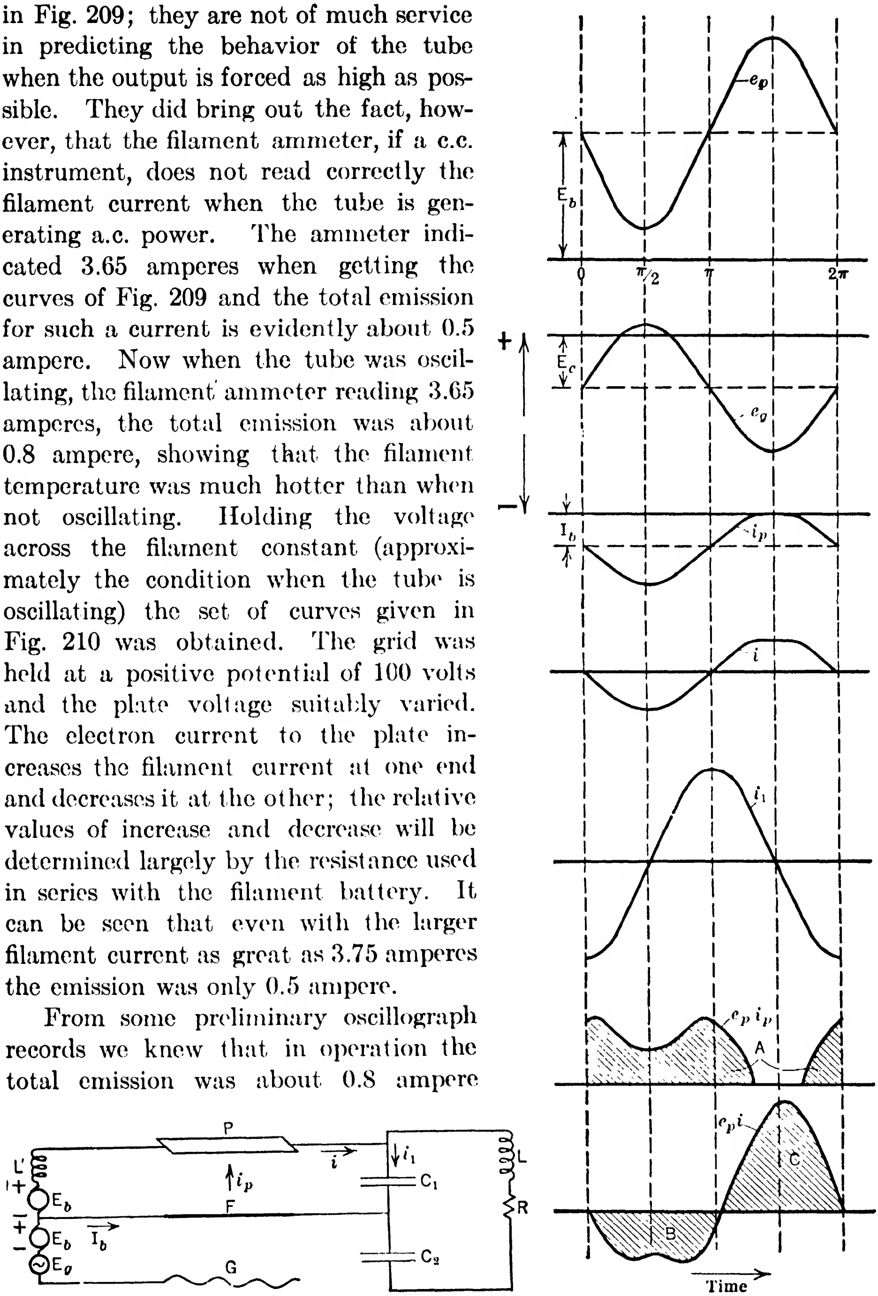


FIG. 208.—Showing the important variables to be studied in determining tube efficiency.

when the filament ammeter read 3.65 amperes. A brief test showed that the filament current required to give this much emission was 4.00 amperes, but this seemed like an excessive current at which to carry out a test, so the characteristics were obtained by extrapolation. In Fig. 211 is shown a set of curves showing the variation of plate and grid currents for various filament currents, and grid and plate potentials, they being extrapolated for the higher filament currents. From this set of curves the results given in Fig. 212 were obtained; as these are important curves they were verified for correctness of form by actually getting them for a lower filament current. These are given in Fig. 213 and are of just the same form as those of Fig. 212.

It is well to point out here that even if it had been possible to get the curves of Fig. 212 with a filament current of 4.00 amperes they would not have given the proper values of i_p and i_g for the tube in operation. While getting these static characteristics the plate and grid get very hot, much hotter than when the tube is in operation as a generator. The emission from the filament is fixed by the filament temperature, and this in turn is fixed by the filament current and the temperature of the plate; if this is hotter when getting the static characteristics than when the tube is generating, the values of i_p and i_g obtained would probably be too large. By the use of an oscillograph it would

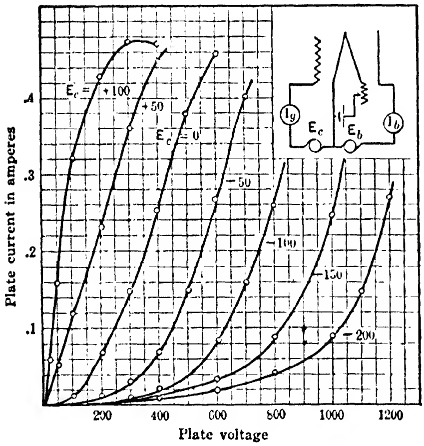


FIG. 209.—Static characteristics of the triode used in making tests.

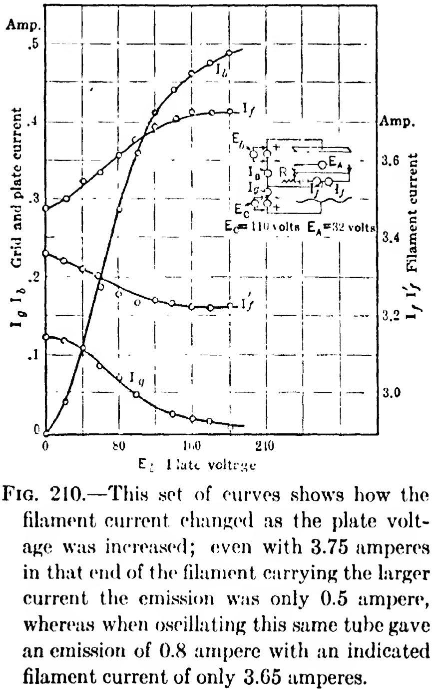


FIG. 210.—This set of curves shows how the filament current changed as the plate voltage was increased; even with 3.75 amperes in that end of the filament carrying the larger current the emission was only 0.5 ampere, whereas when oscillating this same tube gave an emission of 0.8 ampere with an indicated filament current of only 3.65 amperes.

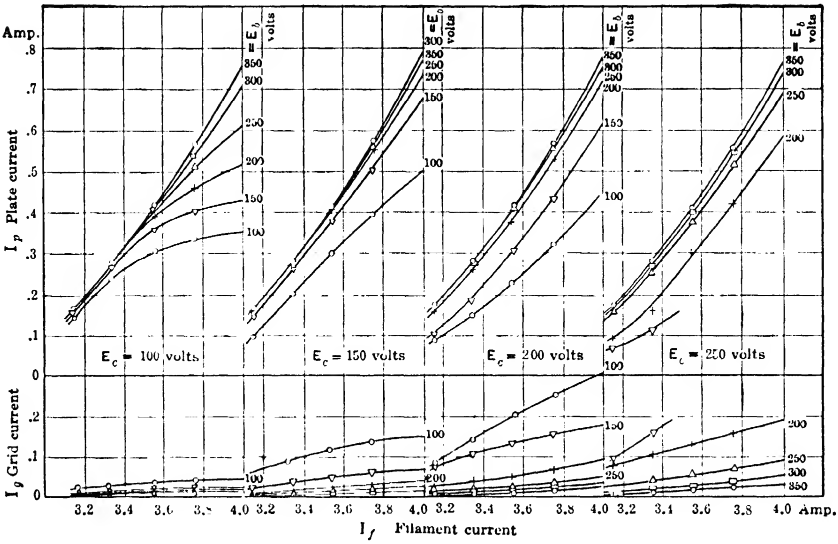


FIG. 211.—Plate currents and grid currents of the tube for various grid and plate voltages, filament current being varied.

have been possible to actually measure the quantities which are obtained in Fig. 212 by extrapolation.[1]

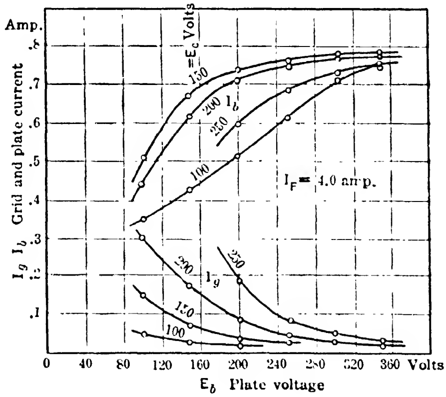


FIG. 212.—Grid and plate currents for various fixed grid potentials and variable plate voltage, filament current at 4.0 amperes; these curves were obtained by extra-polation in Fig. 211. The numbers noted on the individual curves signify the grid potential (positive), the lower set of curves being grid current and upper set plate current.

The curves of Fig. 212, in connection with Fig. 208 enable us to at once give the minimum potential to which the plate should drop and the maximum positive potential for the grid. In order to make the area A (Fig. 208) small, the plate potential, at time $\pi/2$, should be as low as possible. This minimum will be controlled, however, by the other requirement that the area C should be large. If, during the time when e_p is low, i_p does not have its maximum possible value (saturation current), then the positive alter-

[1] Such use of the oscillograph is described by Kalin in the Univ. of Washington, Eng. Bulletin No. 30, Dec. 1, 1924.

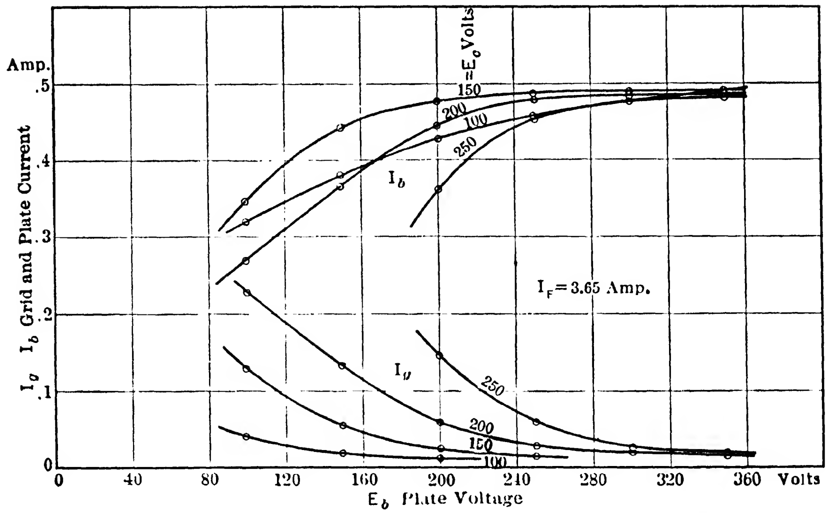


FIG. 213.—As the curves of Fig. 212 are important, and they were obtained by extrapolation, they were verified for correctness of form by picking off from Fig. 211 a similar set of curves for a filament current of 3.65 amperes; evidently these curves are of the same form as those of Fig. 212.

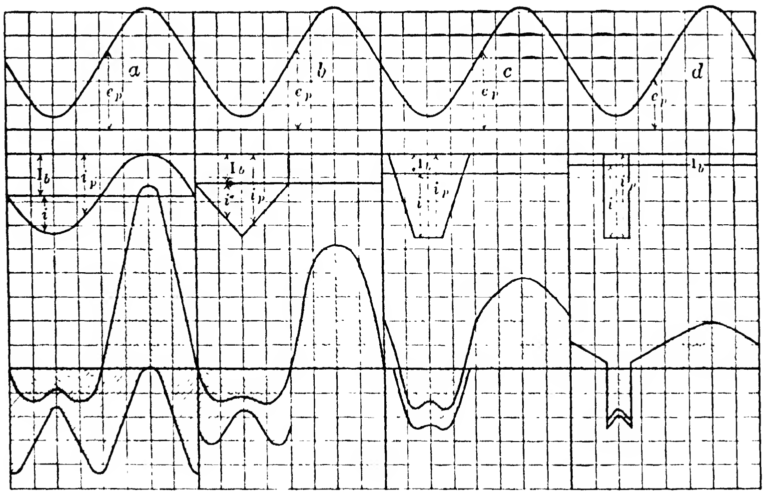


FIG. 214.—In this figure various forms of plate current have been assumed and (plate voltage remaining fixed in shape) the resulting efficiencies calculated.

$$a\left\{\begin{array}{ll}\int e_p i_p dt = 72 & \int e_p i dt = 39 \\ \multicolumn{2}{c}{\text{Efficiency} = 35\%}\end{array}\right. \qquad b\left\{\begin{array}{ll}\int e_p i_p dt = 35 & \int e_p i dt = 31 \\ \multicolumn{2}{c}{\text{Efficiency} = 47\%}\end{array}\right.$$

$$c\left\{\begin{array}{ll}\int e_p i_p dt = 20 & \int e_p i dt = 28 \\ \multicolumn{2}{c}{\text{Efficiency} = 59\%}\end{array}\right. \qquad d\left\{\begin{array}{ll}\int e_p i_p dt = 7 & \int e_p i dt = 23 \\ \multicolumn{2}{c}{\text{Efficiency} = 77\%}\end{array}\right.$$

nation of i will not be as large as it should be and if this is not large the power input to the load circuit, determined principally by the area of C, will be lower than its proper value.

As the average value of i must be zero, if its positive loop is to be as large as possible, and the area of A to be kept as small as possible, the conditions should evidently be so adjusted that, at minimum plate potential, saturation current should flow, and this flow should last for a short time only. During the rest of the cycle the plate current should be zero.

Fig. 214 shows the calculated losses on the plate and input to the load circuit for four different forms of plate current, the plate voltage having the same form for each. It will be seen that both the losses and the output of the tube are greatest for the sinusoidal plate current, but the efficiency for this condition is only 35 per cent; as the form of plate current approaches a short pulse the efficiency increases, being 77 per cent for the form shown in curve (d). The trapezoidal form shown at (c) resembles very closely the form generally employed; our test actually gave about 60 per cent efficiency for a plate current of this form.

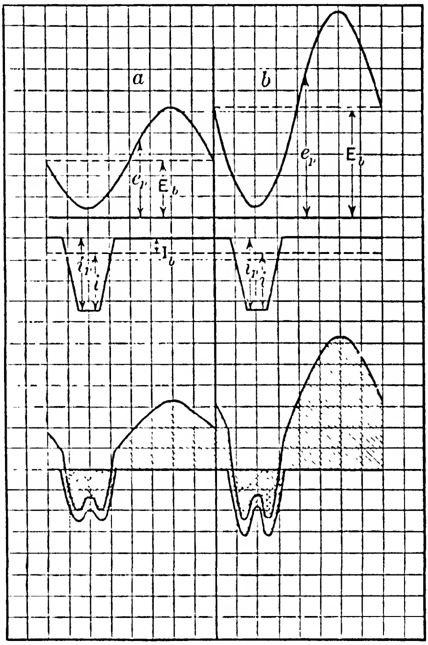


FIG. 215.—$a \begin{cases} \int e_p i_p dt = 15 \qquad \int e_p i dt = 34.5 \\ \text{Efficiency} = 70\% \end{cases}$ $b \begin{cases} \int e_p i_p dt = 19 \qquad \int e_p i dt = 71.5 \\ \text{Efficiency} = 79\% \end{cases}$

Assuming a fixed form of plate current, and fixed minimum of plate potential, it is seen that the efficiency rises as the voltage (E_b) used in the plate circuit is increased.

All four curves are drawn with the maximum plate current the same, supposedly the saturation current for the filament current used; by carrying out other constructions it will be evident that any other condition would result in poorer operation.

By now referring to Figs. 208 and 212 it may be seen that for the tube we were using the plate potential should not fall lower than 200 volts, and that this time the grid should have a positive potential of 150 volts. With greater or less grid potential, the plate potential being 200 volts, the plate current would be less than saturation value; with less plate potential, the current (at time $\pi/2$, Fig. 208), would be less than saturation value, and with greater voltage than 200 volts the loss on the plate would be greater than necessary.

It is to be noted that the efficiency will increase for all the cases given in Fig. 214, if the voltage of the power supply, E_b, is increased, providing that conditions are suitably changed to have the same minimum plate voltage as given in Fig. 214. This is shown by Fig. 215; the two cases given suppose the same form of plate current and same minimum value of plate voltage, but in the second the voltage E_b is about twice as large as in the first case. It is seen that the loss on the plate is increased only

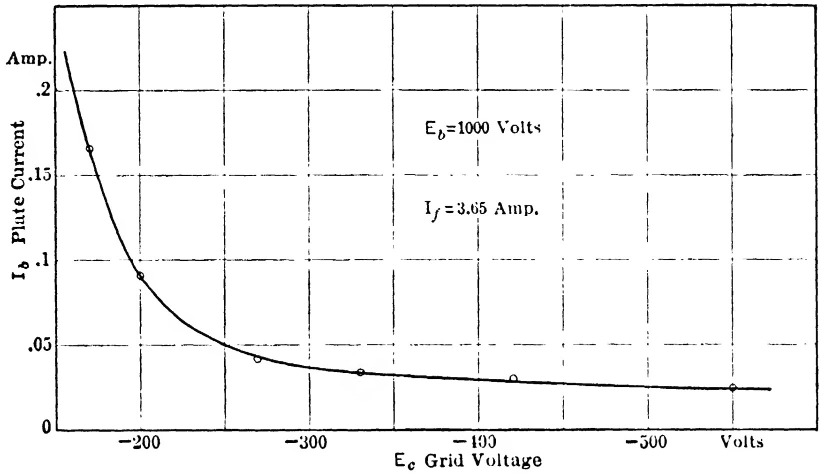


FIG. 216.—The tube used in the tests did not have a constant value for μ; theoretically a negative potential of 260 volts should have reduced the plate current to zero. This tube would have required about 1000 volts (negative) on the grid to completely cut off the plate current. The μ of this triode was variable because of the irregular spacing of the grid wires, a construction feature intentionally incorporated in the present type '35 tubes.

25 per cent whereas the input to the load circuit has been more than doubled. The higher the value of E_b the higher is the efficiency, the limit being fixed only by the safe voltage for the tube.

In the tube we used the efficiency did not rise as high as might be expected, due to the fact that it took excessively high negative potential on the grid to bring the plate current to zero. The oscillograms showed this effect, so a static characteristic curve was taken to investigate this point; it is shown in Fig. 216. If Eq. (7) were valid for this tube, a negative potential of 260 volts would have brought the plate current to zero, whereas it took about 1000 volts; although the plate current is small with a grid negative more than 300 volts this small current has a marked effect on loss of power on the plate, because of the very high plate voltage during that part of the cycle when this small current is flowing to the plate.

Effect of Secondary Emission in High-voltage Tubes.—It was mentioned when discussing the dynatron that if an electron impinges on a metallic surface with sufficient velocity some electrons will be splashed out of the metal, and that if another electrode, charged highly positive, is near by, these secondary emission electrons will leave the plate from which they have been splashed and flow to the neighboring electrode. The electrode from which they come will of course be left positively charged unless other electrons can flow in to replace those lost.

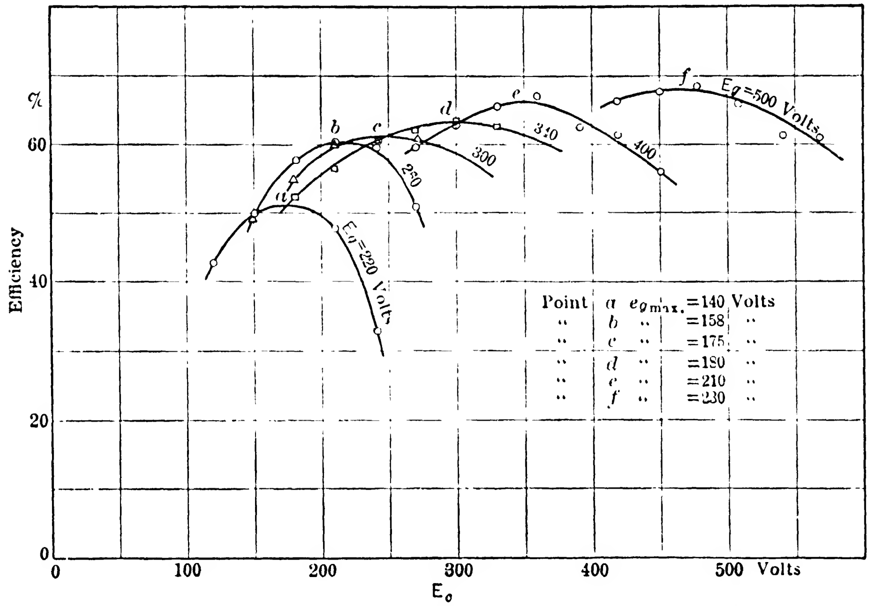


FIG. 217.—Efficiency curves plotted from Table I; the values of positive grid potential for maximum efficiency in each run is calculated and recorded. This agrees well with the predicted "best grid potential."

In high-powered triodes, using thousands of volts on the plate, the grid is very likely to be bombarded with sufficient energy to experience much secondary emission, the electrons splashed off from the grid at once flowing to the highly charged plate. Now, if electrons cannot readily flow on to the grid to replace these lost the grid assumes a positive potential. Such may well be the case if a circuit arrangement is employed which utilizes grid condenser and leak. The leak may impede the flow of electrons on to the grid to such an extent that the grid becomes positively charged; this will correspondingly neutralize the space charge, and very likely increase the plate current to such an extent that the triode is ruined by excessive heat of the plate.

The effect, it will be noticed, is cumulative. More secondary emission means a more positively charged grid, with correspondingly higher plate current and thus more vigorous bombardment of the grid. But this increased bombardment means more secondary emission, etc. Thus if the grid of a high-powered tube starts to swing positive, as a result of the actions analyzed above, the triode is sure to be ruined unless the plate voltage is immediately lowered.

In high-power triodes it is advisable, therefore, to use a sufficiently low-resistance grid leak or still better, use some other scheme than a grid condenser to get the requisite negative bias.

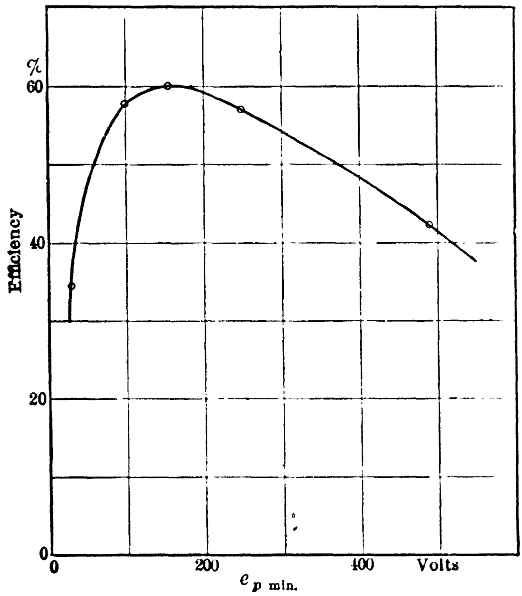


FIG. 218.—From the results given in Table II the efficiency curve shows that maximum efficiency occurs for minimum plate potential of 160 volts, the proper value predicted from Fig. 212.

Experimental Proof of Foregoing Theory.—To test the validity of the ideas presented in the foregoing analysis of the triode converter action a series of runs were made with the tube, using the circuit given in Fig. 207 and the results therefrom are shown in Table I. The frequency was kept at the resonant value for the output circuit and each time a set of readings was taken the value of R was changed properly to maintain the current in the oscillating circuit constant. This was necessary in order to keep the form of the voltage e_p constant as the values of E_c and E_g were varied. While it was not thus pointed out in discussing the current forms of Figs. 214 and 215, the values of E_c and E_g are the factors which bring about the change of current form as the form of e_p is maintained constant. The form of current shown in (a) Fig. 214 was obtained with relatively low E_c and E_g, the value of each of these being increased for the succeeding diagrams of the figure.

In Fig. 217 are shown the efficiency curves for the various runs of

TABLE I

$E_b = 1000$ volts. $C_1 = 2\mu F$. $C_2 = 3.91\mu F$. $\backsim = 140$. $L_1 = 9.8H$. $I_f = 3.65$ Amp.

Run	E_c Volts	E_g Effective Volts	Input Watts	I_2 Effective Amperes	R Ohms	Output RI_2^2 Watts	$\eta = \frac{\text{Output}}{\text{Input}}$ Per Cent
E	120	220	334	0.98	149	143	42.8
	150	220	298	1.00	149	149	50.0
	180	220	214	1.02	119	124	51.5
	210	220	186	1.00	89	89	47.8
	250	220	119	0.96	42	39	32.8
	150	260	302	1.00	149	149	49.3
	180	260	273	1.03	149	158	58.0
	210	260	241	0.99	149	149	60.5
	240	260	197	0.99	119	117	59.5
	270	260	161	0.96	89	82	51.0
	150	300	302	1.00	149	149	49.3
	180	300	283	1.02	149	155	54.8
	210	300	261	1.02	149	155	60.0
	240	300	246	1 00	149	149	60.8
	270	300	212	0 98	134	129	60.8
A	180	340	291	1.01	149	152	52.3
	210	340	278	1.03	149	158	56.8
	240	340	265	1.04	149	161	60.8
	270	340	244	1 01	149	152	62.4
B	300	340	229	0 99	149	146	63.8
	330	340	186	0.99	119	117	63.0
	270	400	260	1.02	149	155	59.7
	300	400	250	1.03	149	158	66.3
	330	400	235	1.02	149	155	66.0
	360	400	222	1.00	149	149	67 3
	390	400	197	0 96	134	124	63.0
	420	400	150	1.02	89	93	61.8
C	450	400	126	0.98	74	71	56.3
D	410	460	228	1.02	149	155	68.0
	420	500	245	1.05	149	164	67.0
	450	500	237	1.04	149	161	68.0
	480	500	222	1.01	149	152	68.6
	510	500	195	1.04	119	129	66.2
	540	500	176	1.02	104	108	61.5
	570	500	157	1.04	89	96	61.2

Table I, and on the curve sheet are given the calculated values of the maximum positive grid potential for that condition in each run which gave maximum efficiency, as indicated at *a*, *b*, *c*, *d*, etc. For the comparatively low value of current in the oscillating circuit which obtained during these tests the form of plate voltage is somewhat different from a sine wave, and the variation of best grid potential may have been due to this cause. The increase in efficiency with increase of E_g and E_c is as would be expected from the analysis given for Fig. 214.

A series of runs was then carried out (results given in Table II) to study the effect of varying the value of the minimum plate voltage, other conditions remaining the same; this was accomplished by varying R, thus cutting down the value of the oscillating current and hence the variation of voltage across the condenser C_1, Fig. 207. The variation of potential

TABLE II

$E_b = 1000$ volts. $C_1 = 2\mu F$. $C_2 = 3.91\mu F$ $\sim = 138$. $L_1 = 9.8H$. $I_f = 3.65$ amp.

Run	e_p Min. Volts	E_c Volts	E_g Effective Volts	Input Watts	I_2 Effective Amps.	R Ohms	Output $= RI_2^2$ Watts	$\eta = \frac{\text{Output}}{\text{Input}}$, Per Cent
A	30	270	300	134	1.12	37	46.5	34.7
	100	270	300	179	1.10	85	103	57.5
	160	270	300	204	1.02	117	122	59.8
	250	270	300	217	0.91	149	123	56.8
B	490	270	300	255	0.60	297	107	42.0

across this condenser, it will be noticed, is what controls the fluctuation of plate voltage.

The value of minimum plate voltage can be calculated by subtracting from E_b the resistance drop through L_1 (which was very small for most of our tests) and from this subtracting the maximum value of the alternating potential drop across C_1. These calculations were made and the results are shown in the curve of Fig. 218; the results verify, better than might be expected, the conclusions reached from theory. With the exception of the first value of e_p (min.) the calculated values agreed with the values measured from the films; the value of 30 was obtained by measurement of the film, the calculated value not agreeing very well.

For various of the runs given in Table I oscillograms were taken; for conditions of run *A* the curves of e_p, e_g, and i_p are given in Fig. 219. From this film, as for the succeeding ones, the first thing to be noticed is that the grid voltage and plate voltage are just 180° out of phase, showing that the load circuit was resistive only. The maximum positive potential of the grid measures on the film 296 volts and the corresponding value of plate potential measures 220 volts. By reference to the curves of Fig. 212 it may be seen that for these respective voltages a large part of the electron current is drawn to the grid, resulting in the peculiar double humped curve of plate current. The maximum negative grid potential was 650 volts, but even this was not sufficient to make the plate current zero. Its values follow, as exactly as can be measured, the values given by the curve of Fig. 216.

For run *B* a set of oscillograms was taken to show all of the quantities involved in the operation of the tube; it required five oscillograph records to get all the quantities wanted. These five films were combined to make

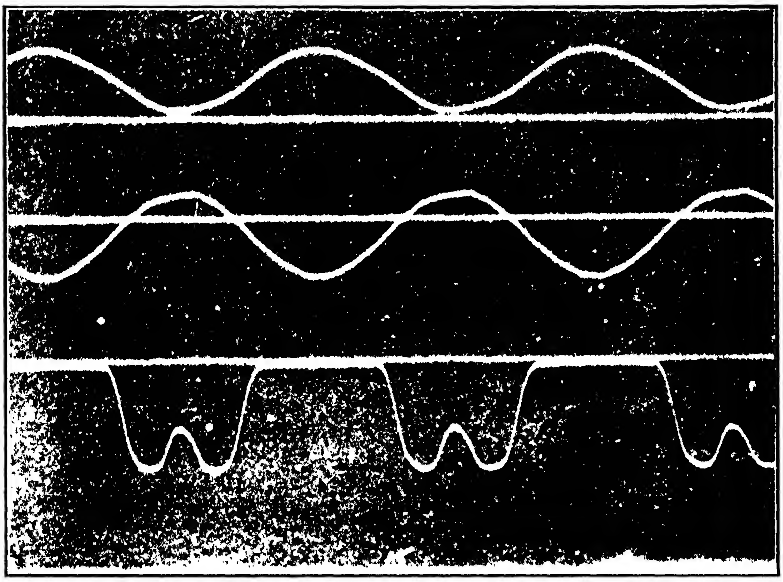


FIG. 219.—Oscillogram for conditions of Run *A*—Table I; evidently the minimum plate potential is too low.

the record shown in Fig. 220; in fitting the various films together care was taken to see that they had their proper respective phases. The white line drawn vertically through all the records gives a line of equiphase.

This set of curves gives the complete story of the circuit and tube. The plate current is very nearly the form shown in Fig. 215, and the plate potential is nearly of the form shown in condition (*a*) of the same figure. The slight depression in the peak value i_p is due to the grid taking some current, this depression coinciding in time with the peak of grid current. The form of the positive alternation of the i curve is not like those previously given, owing to the fact that it has been assumed that I_b was constant whereas it actually had considerable fluctuation, as shown in the record. If the coil used for L_1 had more inductance, this variation in I_b would be diminished; we had only 10 henries with a resistance of 189

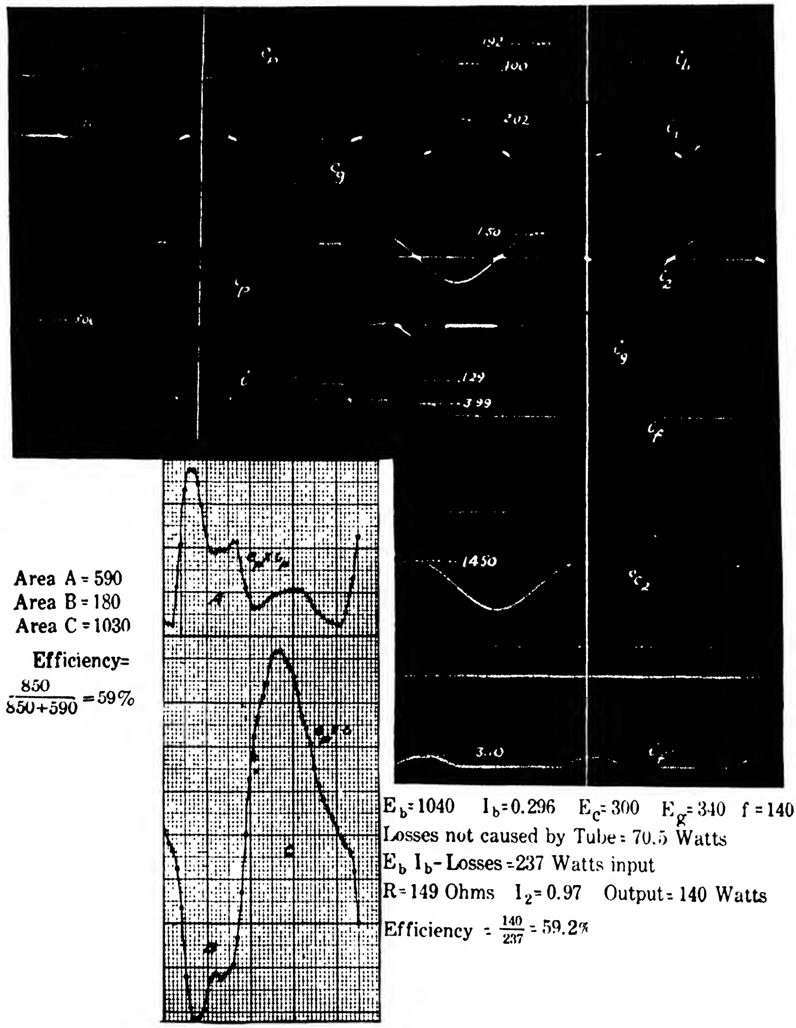


FIG. 220.—Oscillogram of all the currents and voltages for Run B—Table I; the white line represents the same phase on all films. The calculated efficiency from the e_pi_p and e_pi areas agrees well with the calculated efficiency.

ohms, the coil being air core. In practice an iron-core coil of greater inductance would be used, but we did not want to introduce any other sources of distortion than the tube itself.

The form of current in condenser C_1 differs from that in condenser

C_2 because of the uneven distribution of the upper harmonic currents which go to make up the form of i. A glance at Fig. 207 will show that current i can divide, one path being directly through C_1 and the other through R_1, C_2, and L in series. The fundamental component of i will divide essentially equally between these two paths because, for the fundamental frequency, their reactances are equal in magnitude and the effect of resistance compared to reactance is negligible.

For the upper harmonics the two paths do not offer equal reactances, however. Thus for the second harmonic the reactance of the C_2–L branch is seven times that of the C_1 branch and for the third harmonic it is seventeen times as much. So the current through C_2–L will be nearly a pure sine wave and that through C_1 will have all the distortion which the irregular form of i causes.

The grid current has just the form and magnitude predictable from Fig. 212; the amount of current taken by the grid in this test and the values of E_g and E_c used, caused a loss of power on the grid (due to bombardment) of about 10 watts.

The two filament currents, i_f and i'_f, have forms which might be predicted from curves similar to those given in Fig. 210; in that end of the filament carrying the larger current the c.c. ammeter measuring the current indicated only 3.65 amperes, whereas the current actually went as high as 3.99 amperes when the plate was taking its maximum current. The exact amount of emission from the filament when the tube is acting as a generator cannot be predicted from the static characteristic; the temperature distribution in the filament which exists in the oscillating condition of the tube cannot be duplicated in a static test, and it is this temperature distribution which determines the total emission.

The drop across the condenser C_2 was taken to see whether or not it had the right magnitude and phase to serve for excitation of the grid when the tube was run self-exciting; the value of C_2 had been adjusted with this point in mind.

The scheme of getting the efficiency indicated in Figs. 214 and 215 was tried on this record of e_p, i_p, and i, the power curves of e_pi_p and e_pi being shown in Fig. 220; the value obtained, 59 per cent, agrees within the precision of the test with that measured by the meters in the test. The value of 63.8 per cent given in Table I was the value obtained when the oscillograph circuits were not connected, the closing of the circuits changed the conditions enough to drop the efficiency to 59.5 per cent.

Fig. 221 shows the form of i_p which is predicted from Fig. 212 after the forms and magnitudes of e_p and e_g have been assumed; this form of i_p is very close to the actual form given in the oscillogram of Fig. 220.

The result of our tests and analysis have then shown that the efficiency of a tube as a converter can be accurately predicted from the three sets of curves given in Figs. 209, 212, and 216 after we have determined, from the curves of Fig. 212, what the best minimum plate potential is and also what the maximum positive potential of the grid should be.

To get a fair efficiency (60 per cent or better) the value of I_b should not be greater than 25 per cent of the saturation current of the tube; with the efficiency known and the safe radiation of power from the plate being known the proper value of E_b is fixed.

Self-excited Tube.—Using the circuit and constants used in getting the records of Fig. 220 an attempt was made to run the tube self-exciting by changing the connections slightly as shown in Fig. 222. The choke coil L_2 serves to prevent the grid from being short-circuited to the filament (for the a.c. excitation) through the machine E_c. The voltage for excitation was obtained from the drop across the condenser C_2, the insulating condenser C_3 being necessary to prevent short-circuiting the machine E_b. With this connection the grid does not get quite as much excitation as shown by the curve e_{c_2} in Fig. 220, because an appreciable part of this voltage is used in overcoming the reactance drop in C_3. (In this calculation the capacity of the grid circuit of the tube itself must be considered; in some of the power tubes this capacity is as high as 500 $\mu\mu f$, when the load circuit has its proper impedance for maximum output—see p. 529.)

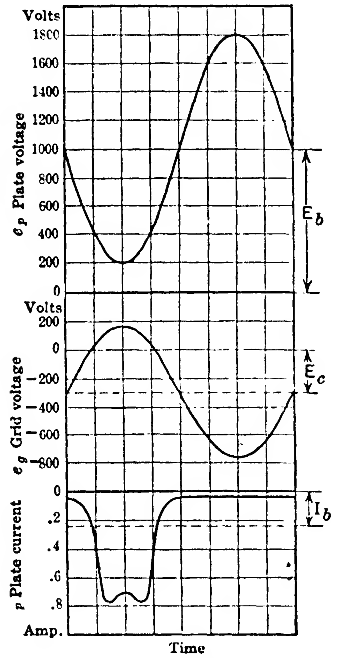


FIG. 221.—Form of plate current curve predicted from Fig. 212; it agrees well with the form actually obtained in Fig. 220.

The circuit of Fig. 222 refused to act as it did for the separate excitation, giving a small output at a low efficiency; a more careful examination of the record in Fig. 220 gave the reason. The alternating components of e_g and e_p must be exactly 180° out of phase if the maximum output and efficiency are to obtain, as becomes at once evident if the construction of Fig. 214 be carried out for any other than the 180° relation. Measurement of the film of Fig. 220 shows e_{c_2} to be 33° out of the 180° phase with e_p and that much phase displacement is sufficient to completely upset the conclusions so far reached. It was therefore necessary to change the relative phase of e_p and e_{c_2}.

Exciting Voltage of Adjustable Phase.—A possible scheme is conventionally indicated in Fig. 223; a rotating field is produced by proper connection to the load circuit and a rotatable coil placed in this rotating field serves for the grid excitation. We had a simpler scheme at hand so did not try this one.

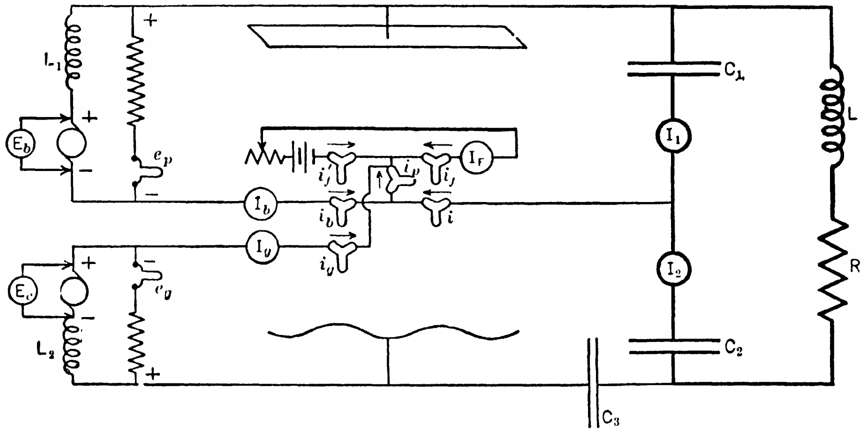


FIG. 222.—Arrangement of the tube circuit for self-excitation; the machine E_c maintains the grid at the proper average potential.

Two other schemes for getting a rotating field are shown in Fig. 224. In the arrangement of Fig. 224 (*a*) the two circuits L_1-C_1 and L_2-C_2 are tuned to resonance with the power supply, by the readings of their respective ammeters A_1 and A_2. Now C_1 is *increased* sufficiently to drop its current to $1/\sqrt{2}$ of the resonance value and the current in the circuit will now lag 45° behind the resonance phase. Next C_2 is *decreased* sufficiently to reduce its current to $1/\sqrt{2}$ of its resonance value and the current in this circuit now leads its resonance phase by 45°. The two coils, A and B, then deliver two voltages 90° apart. They are used to excite the grids of two triodes, in the plate circuits of which are placed the two right-angle coils of Fig. 223.

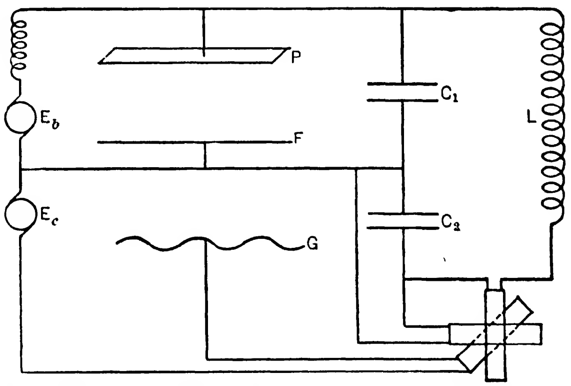


FIG. 223.—A possible arrangement of self-excitation, in which the phase of the voltage impressed on the grid is adjustable.

In Fig. 224 (*b*) is shown a very simple scheme for getting two voltages 90° apart in phase. In this scheme, as in that of diagram (*a*), the triodes, to the grids of which the coils A and B supply excitation, must employ sufficient negative bias that no current is drawn from the coils, otherwise phase shifts of the excitation voltages will occur.

Results Obtained with Self-excited Circuit.—The difference in phase in the voltages across C_1 and C_2 (Fig. 222) comes from the effect of the current i, present in C_1 to a greater extent than in C_2. By making the effect of this current small its disturbing effect may be reduced, and this can be done by increasing the value of C_1 and C_2, and decreasing the value of R, the value of L being properly reduced to maintain the same resonant frequency. The increase in capacity will increase the value of the oscillatory current i_1, and as i remains constant its effect on the relative phases of e_{C_1} and e_{C_2} becomes proportionately less as the capacity is increased.

The arrangement of apparatus remaining as in Fig. 207, the constants

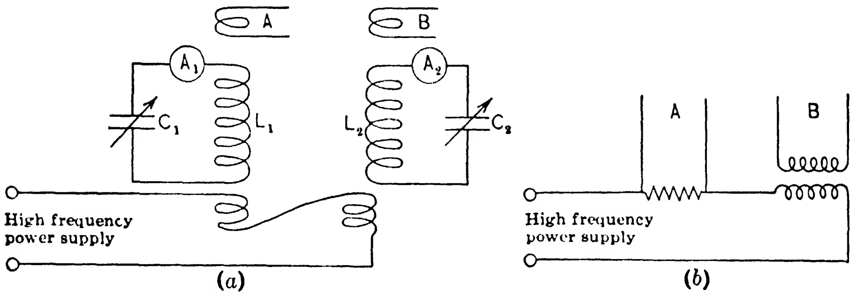


FIG. 224.—Showing two schemes for getting excitation voltage of an adjustable phase.

were readjusted for efficient operation and a set of readings were obtained as follows: $E_b = 900$ volts, $E_c = 230$ volts, $E_g = 310$ volts, frequency = 143, $L_1 = 9.8$ henries, $I_b = 0.321$ ampere, $C_1 = 9.2$ microfarads, $C_2 = 19.4$ microfarads. The resistance of the load circuit was 7.80 ohms and the oscillatory current produced was 4.30 amperes, giving an a.c. output of 143 watts. The input to the tube circuit is obtained from the product I_bE_b after certain losses, not chargeable to the tube circuit, have been deducted.

The condensers C_1 and C_2 each consisted of two condensers connected in series because of the high potentials occurring in the circuit. In order to make the two individual condensers divide the voltage E_b, equally it is necessary that their insulation resistances be alike, a condition seldom encountered. That condenser having the higher resistance (the better one) will take practically all of the E_b voltage as well as its share of the alternating voltage of the circuit, resulting in its probable breakdown. To prevent this occurrence leak resistances were used across each of the condensers making up C_1 and C_2, the leaks each being 21,000 ohms, making

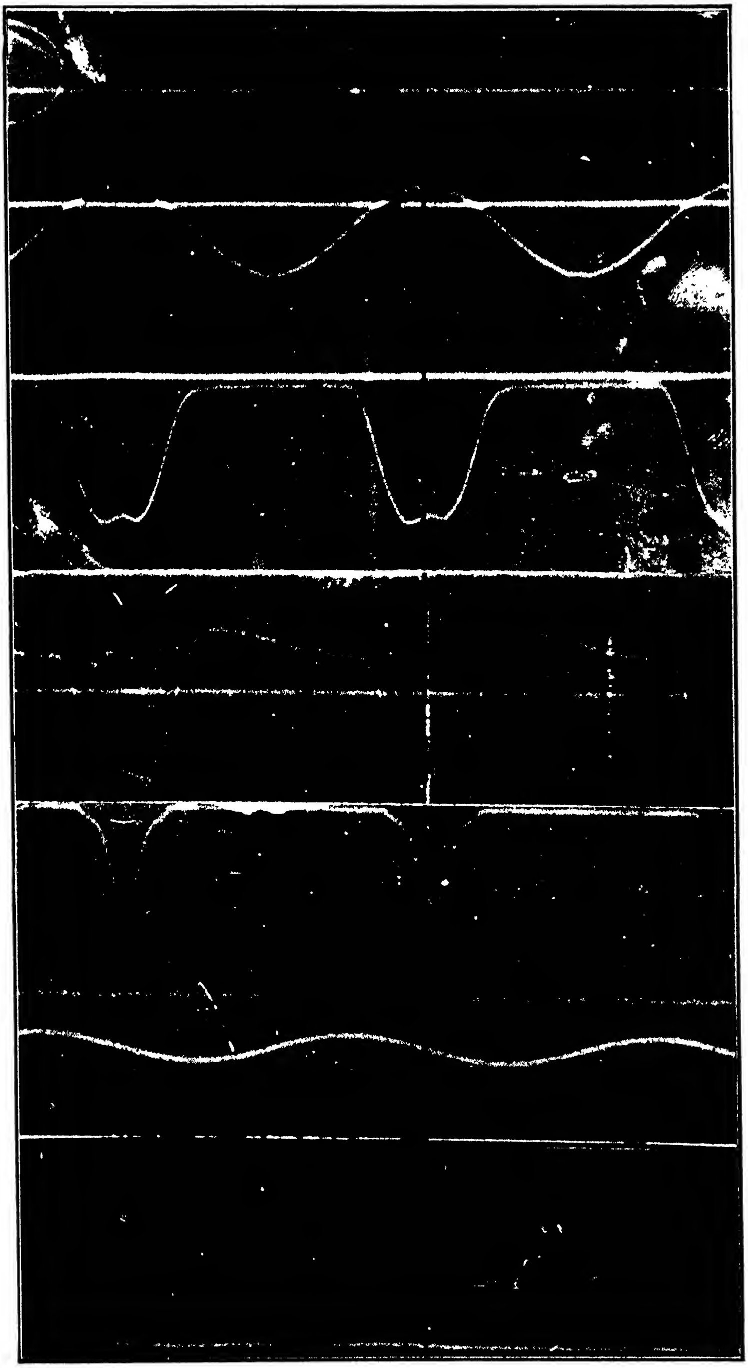


FIG. 225.—Conditions occurring in the self-excited tube; the plate current did not drop to zero as it should do, because there was not enough negative potential on the grid.

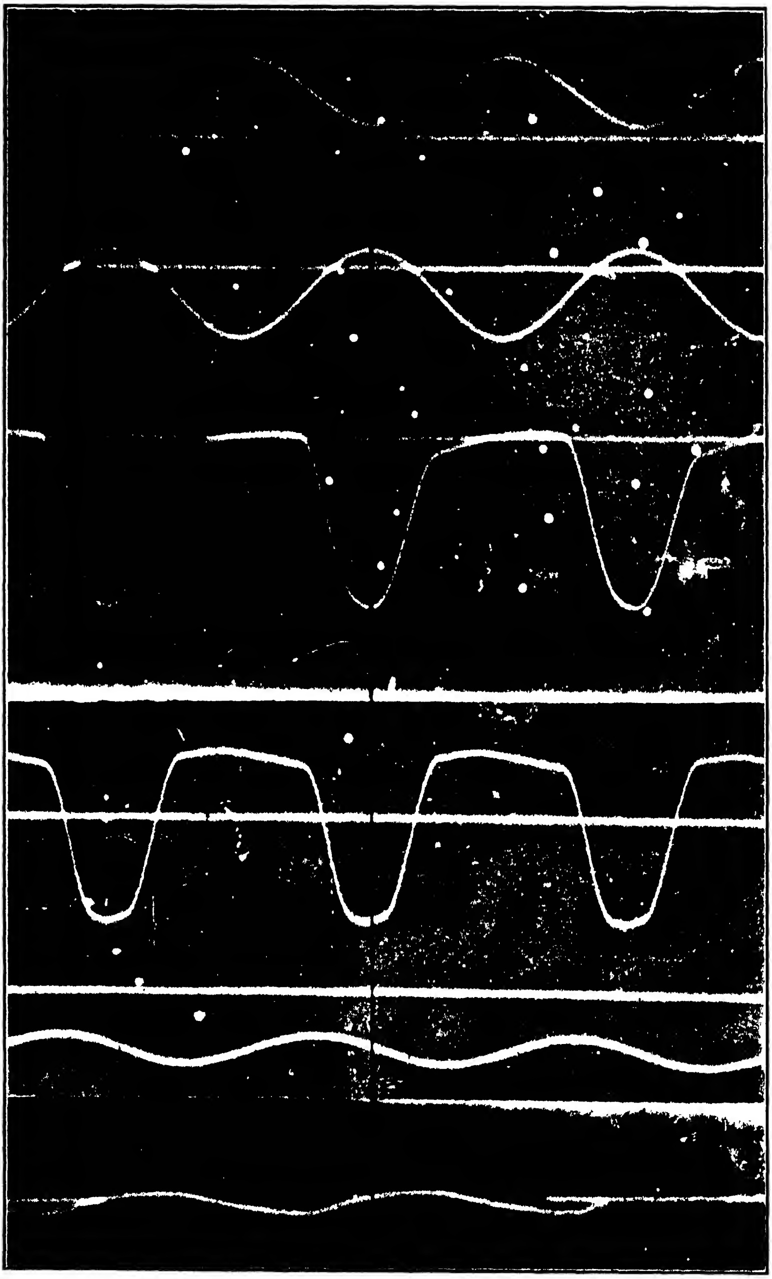


FIG. 226.—In this case of the self-exciting tube, the plate voltage did not fall sufficiently low to give best efficiency; it measures on the film 300 volts whereas Fig. 218 shows the proper minimum plate potential to be 160 volts.

the leak resistance of C_1 and C_2 each 42,000 ohms. Subtracting the I^2R losses in these leaks as well as the I^2R losses in the choke coil L_1, gives

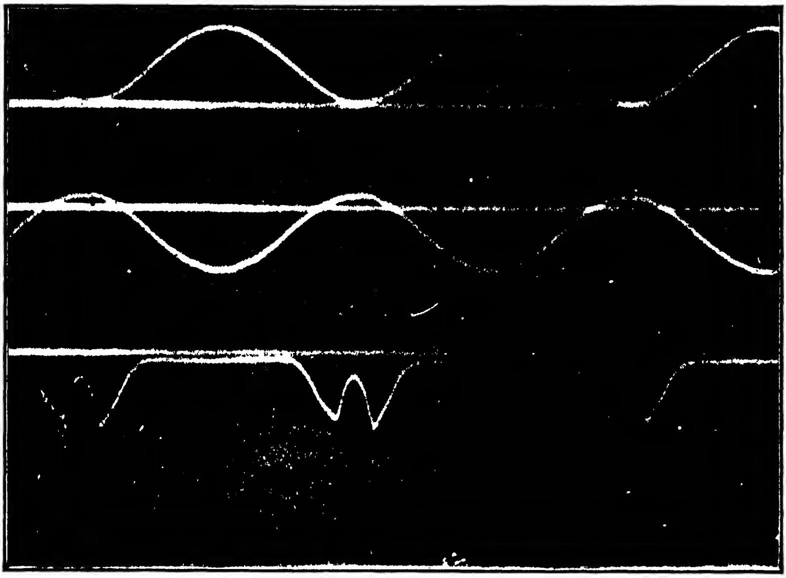


FIG. 227.—Separately excited tube—$E_b = 1040$, $I_b = 0.170$, $E_c = 270$, $E_g = 300$, $R = 37$. Input = 106 watts. Output = 50, Efficiency = 47%.

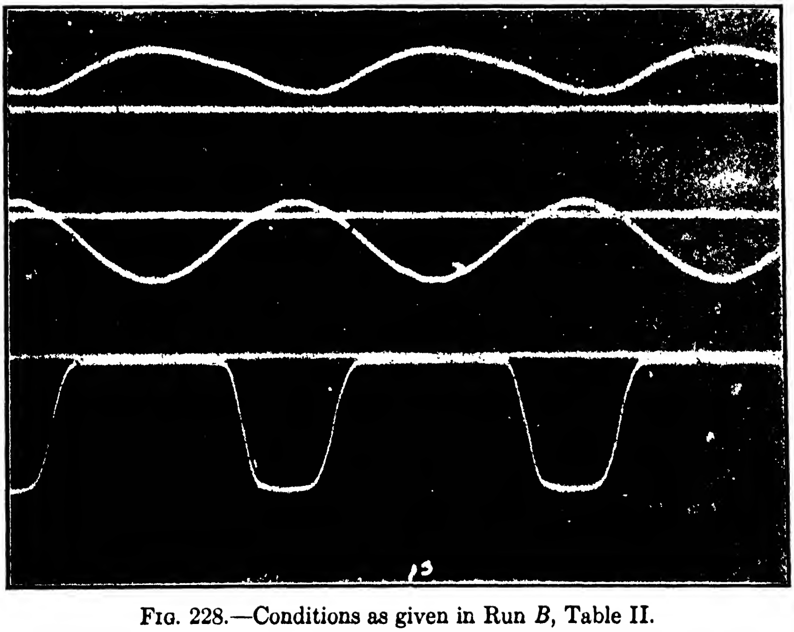


FIG. 228.—Conditions as given in Run B, Table II.

the input to the tube circuit 229 watts; the efficiency was thus 62.7 per cent.

Oscillograms taken of the currents in this circuit are given in Fig. 225. It is evident that the values of E_g and E_c might well have been greater, resulting in a higher efficiency because of the resultant smaller minimum plate current. Although the plate current is small the plate voltage is large and so results in a high unnecessary loss on the plate.

The phase of e_{C_2} is now practically coincident with that of E_g, and it should therefore serve as a source of excitation. The circuit did not give as much power, however, when made self-exciting, as it should, so the constants were changed slightly to get more power. As finally tested

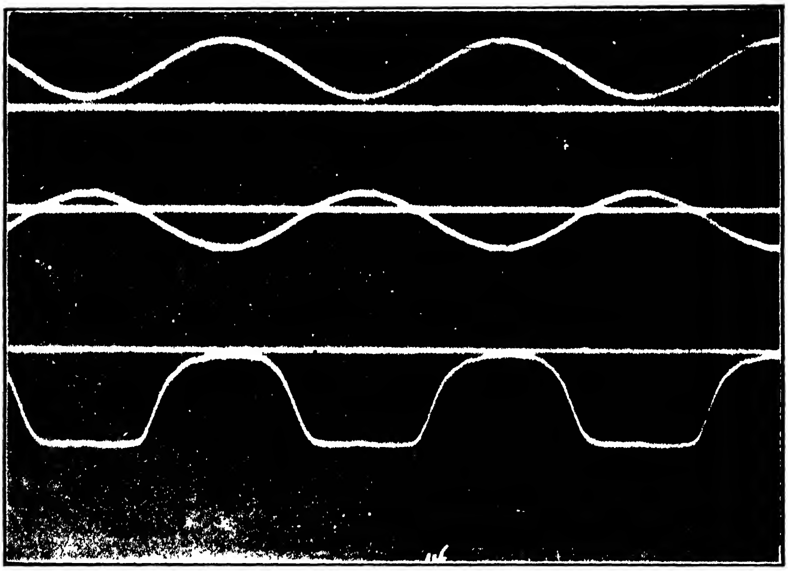


Fig. 229.—Conditions as given in Run E, Table I.

the self-exciting circuit had the constants and performance given herewith: $E_b=1000$ volts, $I_b=0.335$ ampere, $C_1=7.36$ microfarads, $C_2=13.8$ microfarads, $L=0.201$ henry, $L_1=9.8$ henries, $L_2=9.0$ henries, $E_c=230$ volts, $R=8.0$ ohms. The current produced in the oscillating circuit was 4.40 amperes, resulting in an efficiency of 57 per cent.

Fig. 226 shows the currents and voltages in this self-exciting circuit, and it is at once evident why such a comparatively low efficiency was obtained; the minimum plate voltage, instead of being 160 volts, as it should for this tube, was 300 volts. For this figure the curve of plate current included also the alternating component of the grid current, hence the absence of the depression at the peak value.

The current through the plate-current vibrator reversed during part of the cycle, due to the fact that this vibrator carried in addition to the

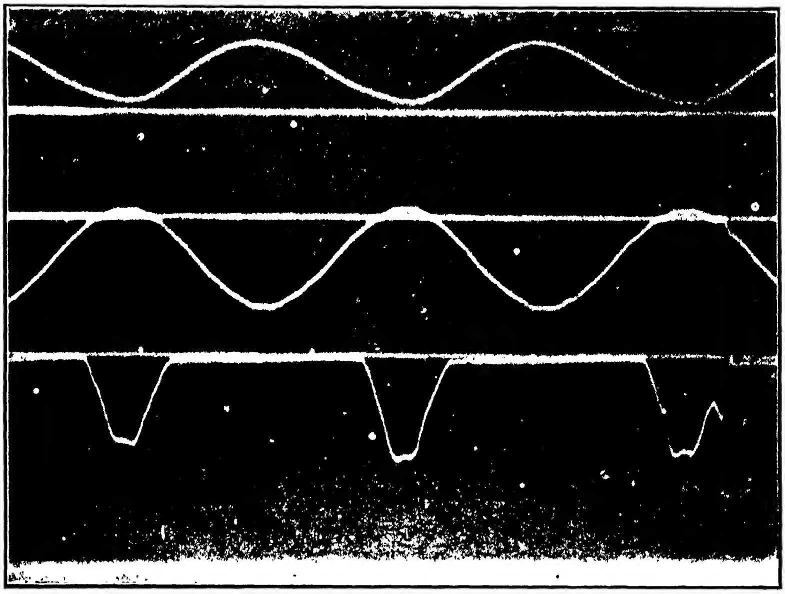


FIG. 230.—Conditions as given in Run C, Table I.

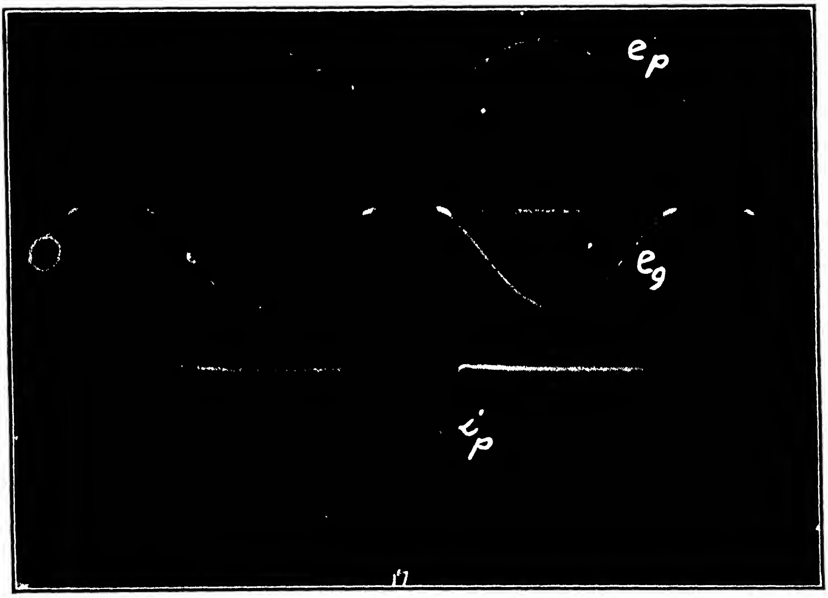


FIG. 231.—Best conditions for high efficiency. $E_b = 1040$, $I_b = 0.286$, $E_c = 410$, $E_g = 460$, $R = 149$, input $= 227$ watts, output $= 155$, efficiency $= 68\%$.

plate and grid currents, an alternating current which resulted from the voltage across the condenser C_2 acting through the reactance of coil L_2 and condenser C_2, Fig. 222. This current is shown as i_x in Fig. 226; when the plate current is corrected by this small amount it is seen that the plate current does not reverse, as we know it cannot with the conditions as they existed in this test.

Action of the Tube at High Frequency.—It was desired to show that the action of the tube was just the same at high frequency as at the low frequencies used, so a circuit was arranged similar to that of Fig. 222, with smaller values of capacity and inductance. The choke coils L_1 and L_2 used in the previous tests would act as condensers of comparatively low reactance at the high frequency to be used so they also had to be changed. The constants of the circuit used were: $E_b = 1000$ volts, $I_b = 0.285$ ampere, $C_1 = 0.0144$ farad, $C_2 = 0.0284$ microfarad, frequency $= 98{,}500$, $L_1 = 0.023$ henry, $L_2 = 0.016$ henry, $E_c = 240$ volts, $R = 6.16$ ohms (high-frequency determination). There were no leaks used with the condensers in this circuit, so that the product E_bI_b, after subtracting the I^2R loss on the choke coil L_1, gives the input. It is found to be 284 watts, and as the output to the load circuit was 160 watts the efficiency was 56.2 per cent, which is in fair agreement with the results obtained at 166 cycles.

In Figs. 227–231 are shown some special oscillograms of the plate current, plate voltage, and grid voltage, all for the separately excited tube with the circuit shown in Fig. 207; the conditions of the circuit were as noted in Tables I and II.

The conditions obtaining when Fig. 231 was taken show the best adjustments for efficiency which we were able to get with this type of tube; the high efficiency was obtained without unduly decreasing the output. If this form of plate current could be maintained and the value of E_b be increased to 3000 volts the calculated efficiency becomes 85 per cent; this is probably as good as could be done with sine wave shapes of e_p and e_g, but it seems as though, by suitably deforming both of them, giving them both flat tops, the efficiency could be considerably increased over this value.

Tests similar to those described in this paper were carried out using a much smaller tube, that styled by the U. S. A. Signal Corps, type VT-2. The results obtained with the large tube were duplicated almost exactly in so far as efficiency was concerned. It was found possible to so adjust the values of E_c and E_g that the tube gave an output of 6.3 watts with an efficiency of 70 per cent, the voltage used in the plate circuit being the rated value, namely 300 volts. It was found possible to get over 7 watts output with the plate loss considerably lower than its safe rated value; if the plate voltage had been increased to perhaps 400 volts the tube output might have been raised to 10 watts while still having the plate loss within its safe value.

These tests were all carried out with a separately excited tube; with the tube self-excited, the efficiency was not obtained higher than 61 per cent, with a plate voltage of 300. This run gave an output of 5.6 watts output with a current I_b of 0.0305 ampere; the frequency was 400,000 cycles, the value of R was 53 ohms, the oscillating current 0.325 ampere, C_1 and C_2 being 1360 and 770 micro-microfarads, respectively. The value of E_c was 40 volts.

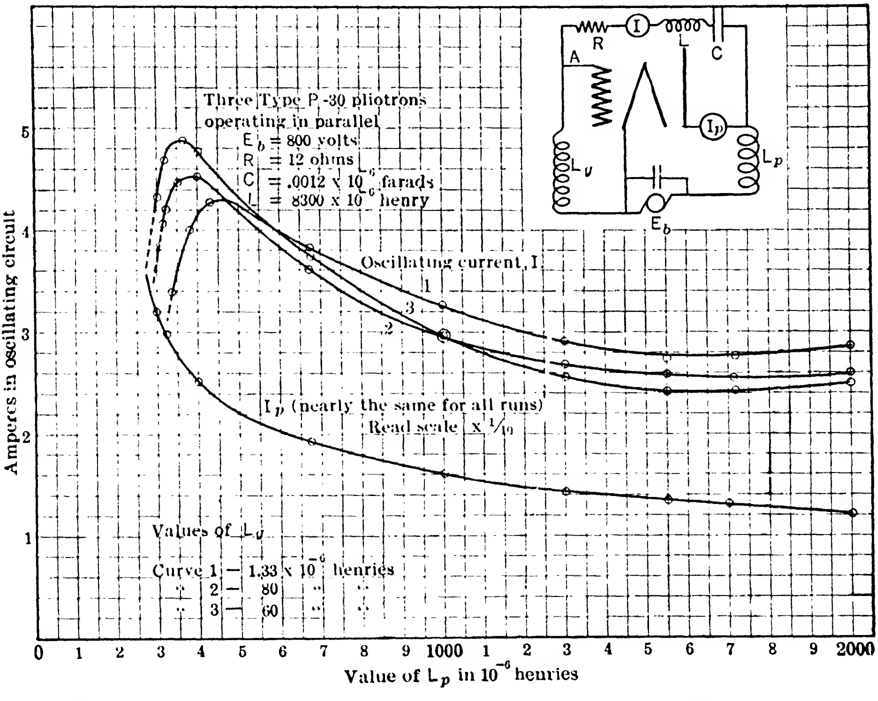


FIG. 232.—Amount of high-frequency power obtainable from three 250-watt triodes in parallel and effect of changes in L_g.

With the conditions of a self-excited circuit adjusted for the best conditions as previously outlined difficulty may be encountered in starting the circuit to oscillate, a shock of some kind being generally required to start oscillations. Because of this possible difficulty it may be the best practice to run these tubes separately excited, using one tube (so adjusted that it oscillates readily) for exciting others as indicated in Figs. 192–193, pp. 652–656. It may well be that with a higher resistance in the oscillating circuit more output can be obtained from two tubes if one only is used as a generator, the other being used as exciter only. Certainly if more than two tubes are to be used it will be well to use one as exciter for supplying the grid voltage for the others. The small exciter tube should be set with coupling much greater than the critical value and it will " key " readily and quickly. The power tube which is being excited should be arranged with sufficient grid bias to get the form of plate current shown on p. 685; the amount of excitation imparted to the grid of the power tube must, of course, also be properly adjusted.

Characteristics of the Circuit of Fig. 158.—Using the circuit shown in Fig. 158 (p. 614), a series of runs was made to investigate the effect of

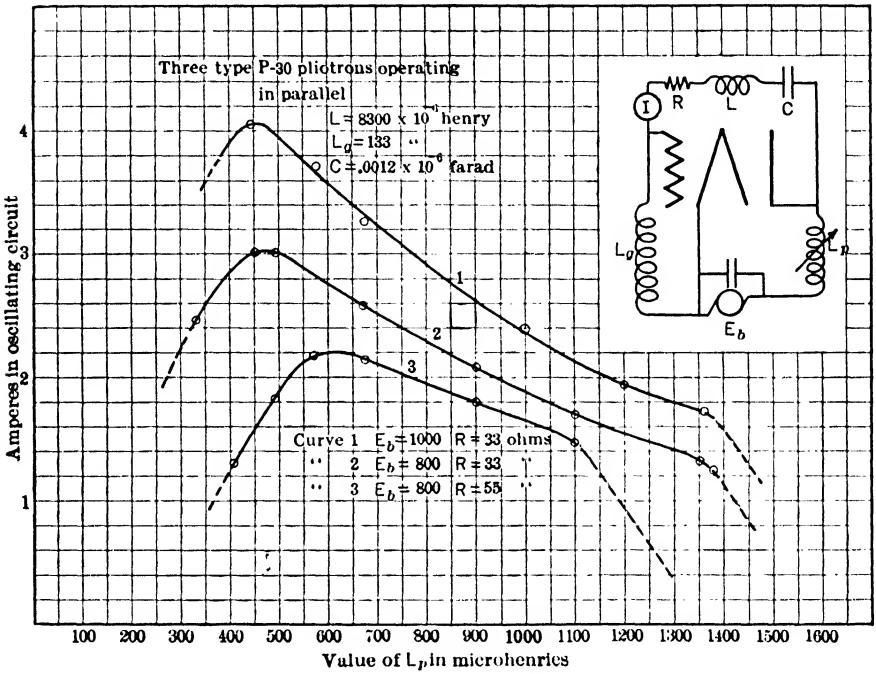


FIG. 233.—Showing the effect of varying the plate circuit inductance with fixed value of L_g. Region of oscillation indicated by solid lines; outside these limits tubes refused to oscillate.

changing the constants of the circuit and some of the results are shown in Figs. 232 and 233; the legends and diagrams on the curve sheets make them self-explanatory. For these tests three 250 watt power tubes were operated in parallel, all grids being connected together as also were the plates; the plate current recorded on the curve sheets is that of one tube.

It may be seen, from Fig. 232, curve 3, for example, that for efficient coils and condensers of the type used here it is possible to get as much as 16,000 volt amperes from one tube, with only 800 volts on the plate and normal filament current. The behavior of the tubes, as regards stability, conditions for maximum output, etc., agree fairly well with the theoretical predictions.

Characteristics of Screen-grid Tubes.—The capacity between the plate and grid of a triode is frequently very disturbing; it not only results in a high effective input capacity of the triode but it also tends to make it unstable if used in a high-gain amplifier. As shown in Figs. 84 and 85, a shield grid may be put around the plate and suitably connected to the *B* battery, this prevents the electric field of the grid from " reaching through " to the plate, and vice versa. In addition to this desired effect, however, the screen grid gives to the tube some very unusual and sometimes undesired characteristics.

In Fig. 234 are shown the ordinary characteristics of a type 224 tetrode;

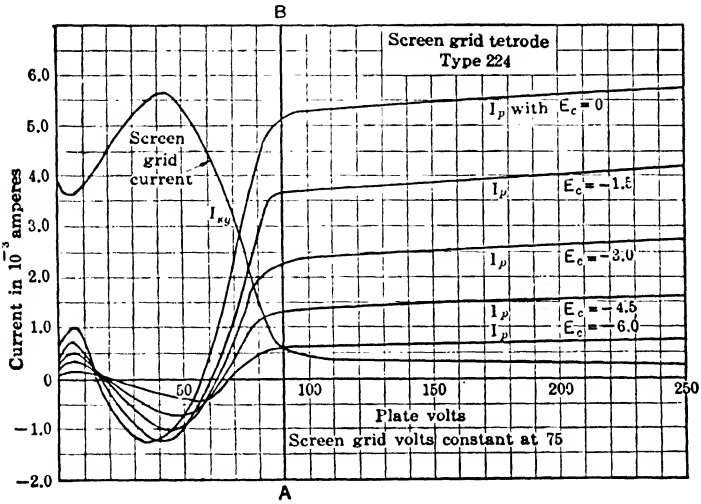


FIG. 234.—Normal characteristics of a screen-grid tetrode.

this is a screen-grid, a.c. heater type of tetrode used primarily as a radio-frequency amplifier. It is appreciated at once that these curves are entirely different from those of the ordinary triode. For this special tube the recommended screen-grid voltage is 75, so in obtaining the plate-current curves with different bias on the control grid the screen grid was maintained at 75 volts.

With the control grid held at cathode potential ($E_c = 0$) the screen-grid current and plate current follow the peculiar variations shown. The sum of I_p and I_{sg} is nearly constant, but the division of this total current changes greatly as the plate voltage is raised. The screen-grid current performs no useful function, so we will not bother to discuss its form.

The plate current I_p increases from zero, with increasing E_p, for a few volts, and then diminishes and actually reverses in the region 15 to 60 volts. This is the dynatron region; the secondary emission from the plate is sufficient to more than offset the electron stream from the cathode, giving a net flow of electrons *from* the plate. It will be noticed that in this region the current I_{sg} increased even though screen voltage was held constant.

When the plate voltage has been raised to about the same voltage as the screen grid, whatever electrons are splashed out (secondary emission) are pulled back by the plate; in addition, the plate is able to pull away from the screen grid a greater share of the electrons which have left the hot cathode. With greater than 90 volts in the plate the change in I_p with E_p is very small, being, however, practically constant for plate voltages from 100 to 250; this is the operating region of this tetrode. It may be noticed that in this region the screen-grid current drops almost as rapidly as the plate current rises.

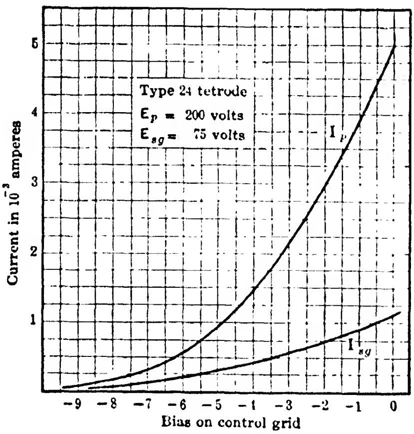


FIG. 235.—Variation of plate and screen-grid currents, as control-grid potential is varied.

With various values of control-grid bias the plate current goes through corresponding changes as the plate voltage is raised from zero, the value of I_p, of course, being greater for the lower values of bias.

In Fig. 235 are shown the changes in I_p and I_{sg} as the control-grid bias is altered; these curves are for a different tube from that used in getting the curves of Fig. 234, hence the values of one do not quite check with those of the other.

Plate-circuit Resistance of the Screen-grid Tube.—Plate-circuit resistance is defined as the ratio of change in plate voltage to the corresponding change in plate current, the control grid and other factors being maintained constant. For $E_c=0$ the value of ΔI_p for a plate-voltage change from 150 to 250 is 0.27 milliampere. This gives $R_p=100/0.00027=370{,}000$ ohms. With $E_c=-1.5$, the same change in E_p gives ΔI_p of 0.00024 ampere, so that $R_p=417{,}000$ ohms. Similar high values of R_p are found for all other values of grid bias.

The plate-circuit resistance of a screen-grid tube is extremely high, so high, indeed, that it is difficult to make an external circuit that properly matches the tube resistance. This high resistance is caused by the inability of the plate to "reach through" the screen grid and attract many of the electrons which are evaporating from the cathode. The difficulty of matching the tube resistance may be understood by the help of Fig. 236. The capacity of this tube from plate to screen grid is about 10 $\mu\mu f$. The wiring may increase this by another 10 $\mu\mu f$, unless it is carefully disposed, so that there may be a capacity from plate to ground of 20 $\mu\mu f$. Now if a radio frequency of 1000 kc. is being amplified (or to be amplified) the reactance from plate to ground is 8000 ohms. Now if we use a resistor, R, of 400,000 ohms, to match the tube resistance, there will be a reactance of only 8000 ohms connected in parallel with it, *thus practically short-circuiting the plate circuit.* The amplification would be very small, so in a radio-frequency amplifier a resistance coupler of this kind could not be used even if it were desired. An amplifier for television purposes should take frequencies as high as 100 kc., but even for this comparatively low frequency the capacity reactance to ground would be only 80,000 ohms, and this again would have a serious shorting effect on a 400,000-ohm resistor. It is evident that if a resistance is to be used in the plate circuit, this should be taken as low as feasible, and the wiring should be done in such a way as to minimize the capacity from plate to ground.

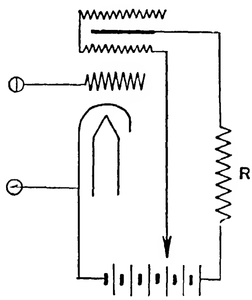


FIG. 236.—Simplified circuit of screen-grid tube.

Amplification Factor of a Screen-grid Tube.—The amplification factor μ is defined as it was for the triode, the ratio of $\Delta E_p/\Delta E_g$ to maintain I_p constant. It is also defined as the ratio of $\Delta I_p/\Delta E_g$ to $\Delta I_p/\Delta E_p$, and we can get these values from the curves of Fig. 234. With $E_c = 0$, a change in E_p of 100 volts (150 to 250) gives a change in I_p of 0.00027 ampere. At $E_p = 200$, a change of 1.5 volts in E_c (0 to − 1.5 volts) reduces the value of I_p by 0.00159 ampere. This gives $\Delta I_p/\Delta E_g$ of 0.00106 ampere per volt and $\Delta I_p/\Delta E_p$ of 0.0000027 ampere per volt, so that $\mu = 0.00106/0.0000027 = 392$. Thus the amplification factor of this type of tube is very high; even if the external circuit is much too low in impedance to match the tube, so that only a small fraction of the theoretically possible voltage amplification is obtained, still the voltage amplification is very high compared with what can be obtained with other types of tubes.

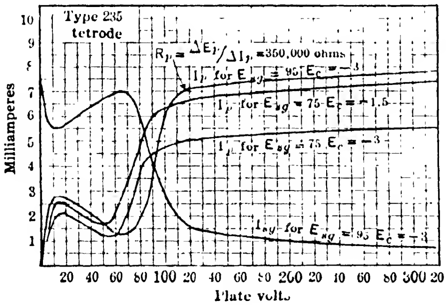


FIG. 237.—Characteristics of screen-grid tube with variable μ.

An exact mathematical treatment of the screen-grid tube seems to be very complex; such an analysis has been attempted,[1] but it leads to extremely cumbersome expressions which are of little real service.

Screen-grid Tube with Variable μ.—By constructing the control grid of a spiral with variable spacing between adjacent turns the amplification factor of the tube is made variable; it depends greatly upon the amount of bias used on the control grid. Thus μ and s_m (the transconductance), which are reasonably constant in the working range of the ordinary triode, have wide variations in this type of tube.

In Fig. 237 are shown the ordinary characteristics of this variable μ tube, and in Fig. 238 are shown the variations in I_p and I_{sg} as the control-grid bias is changed. With 250 volts on the plate the value of s_m is 1050 micromhos with 3-volt bias; it is 15 micromhos with 40-volt bias and about 2 with 50-volt bias. There is thus a gradual change in s_m with increasing bias, making the tube useful for controlling the amplification of a set. It is used in just this manner; the control grid is connected to a potentiometer capable of changing its bias from zero to 50 volts.

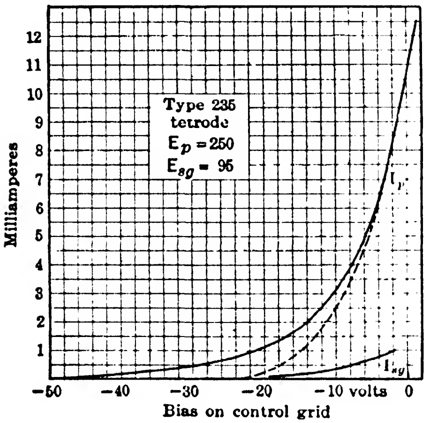


FIG. 238.—Showing how shape of I_p-E_c curve, with variable μ grid, differs from that of a constant μ tube.

Screen-grid Tube a Low-power Tetrode.—If it is desired to get much power output from a tube *the plate current must be high and the circuit must be arranged to permit wide fluctuations in plate voltage.* It is the product of the change in plate volts and in plate current that indicates how much power the tube is delivering. Now inspection of Fig. 234 shows that large plate current, cannot be obtained unless the control grid is made positive, and this is prohibited because of the resultant distortion and loss of selectivity of the amplifier.

But it might be possible to get plenty of plate current if the screen grid were operated at higher voltages; the effect of doing this is shown in Fig. 239.

Larger plate currents are obtained when the screen grid is made more positive, but the range over which the plate voltage can swing is decreased because the lower voltage limit of the operating range (where the plate current falls off rapidly) is pushed up as the screen-grid potential

[1] Brainerd, "Mathematical Theory of Tetrode," I.R.E., June, 1929, p. 1006.

is raised. Thus with 100 volts on the screen the lowest permissible swing of the plate voltage is 110; with 140 volts on the screen it is 160, and with 180 volts on the screen it is 200 volts. Thus it is impossible to get

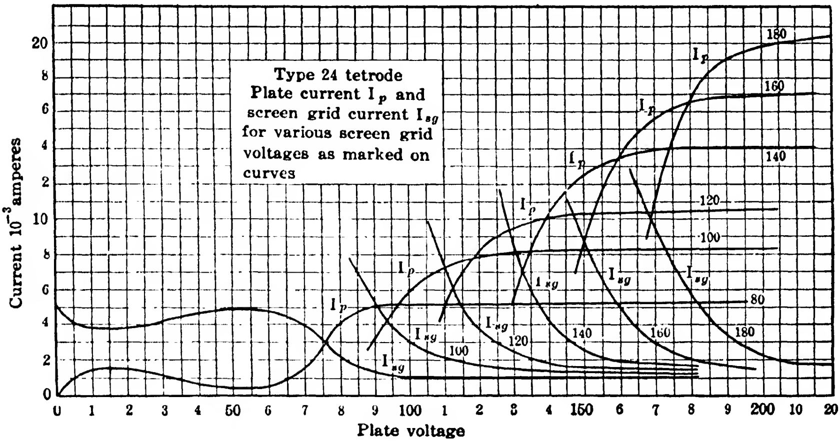


FIG. 239.—This set of curves shows why it is impossible to get much power from the ordinary screen-grid tube.

much power out of this screen-grid tube; as the plate current is raised the useful range of plate voltage is reduced. However, by using a fifth electrode properly connected, this limit on the power output is removed.

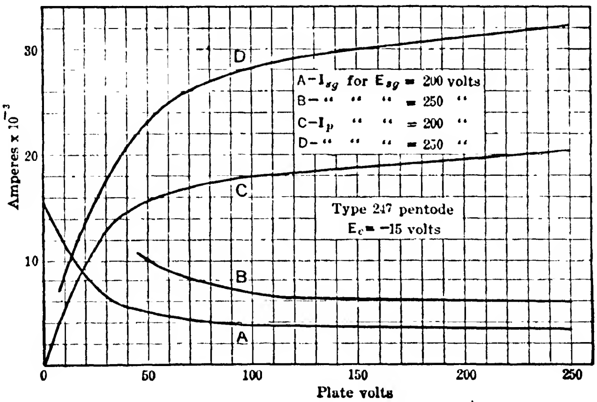


FIG. 240.—The "suppressor" grid of the pentode, extends the plate current "plateaus" of Fig. 239, down to low-plate voltages, making possible greater variations in plate voltage thence greater output.

The Pentode.—The pentode (type 247) is a tube with oxide-coated filament, a control grid, a screen grid, plate, and *suppressor grid.* This extra grid is put between the screen grid and the plate and is connected electrically inside the tube, to the middle point of the filament. Its effect is to completely suppress the phenomenon of secondary emission from the plate, hence its name. Thus the lower limit of plate-voltage swing, which was high for the tube shown in Fig. 239 (due to secondary emission from the plate lowering the plate current prohibitively) is reduced to very low voltages, hence the plate voltage may swing over wide variations.

With the structure which this tube has at present it is advisable to hold the screen grid at high voltages (as was attempted in Fig. 239). In Fig. 240 the ordinary characteristics of the pentode are shown. The plate circuit resistance is reasonably low (compared to that of the screen-

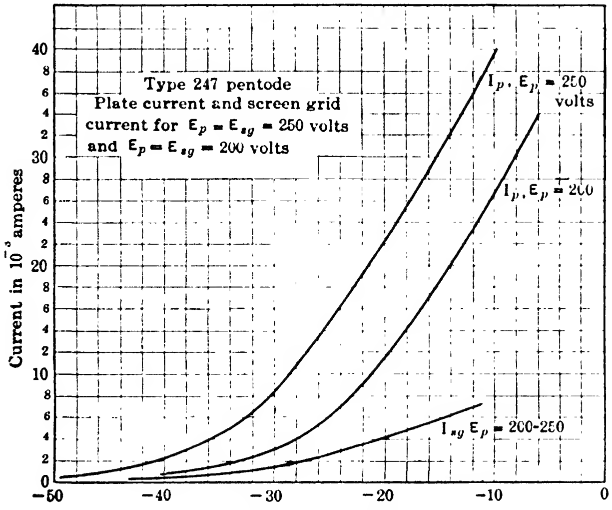


FIG. 241.—Action of the control grid of the pentode.

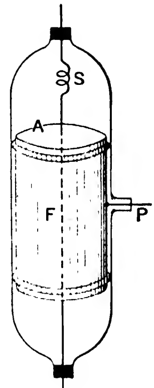


FIG. 242.—The magnetron consists of a straight filament axially located in a cylindrical plate.

grid tube); in Fig. 240 we find that a change in plate volts of 150 (100 to 250) gives a change in plate current of 0.0044 ampere, which gives a plate-circuit resistance of 32,000 ohms. In Fig. 241 are shown the changes in plate current and screen-grid current as the control-grid bias is altered. With 250 volts on the plate a change of 1 volt in grid bias produces a change in I_p of 0.0017 ampere. But a change of 1 volt on the plate, with $E_p = 250$, gives a change in I_p of only 0.000020 ampere, so that the amplification factor μ is equal to 0.0017/0.000020, or 85.

Hence the pentode has a very high voltage amplification. With a load resistance of 7000 ohms (the value recommended by the manufacturer), 250 volts on plate and screen grid, and control-grid bias of 16, a signal of 10 volts (effective) will deliver about 2 watts of signal, much more than any similar tube. This pentode is being used as the output tube for all modern sets.

In a new type of tube just appearing (type 58) the suppressor grid is not connected to the filament inside the tube, but is brought out to a separate terminal; this requires a six-prong base and socket. With this type of tube the suppressor grid can be adjusted to any desired potential, and its characteristics correspondingly altered. Thus the type 58 is a variable μ control grid, screen grid, suppressor grid tube, with special suppressor grid prong on the socket; similarly requiring a six-prong base is the type 57, which is somewhat similar to the type 24 screen-grid amplifier and detector. These newer tubes are smaller than their predecessors and have no screen grid outside the plate. The metal can, which is used over all tubes of the screen-grid type, is itself used to act as the outer screen, in conjunction with a metal ring inside the tube, connected with the cathode.

The Magnetron.—This type of tube, a diode, depends upon the deflecting action of a magnetic field on an electron moving across this field. Cylindrical in structure, it has a centrally located filament and a cylindrical plate. The construction is shown in Fig. 242; a tungsten filament, F, is stretched tight by the spring S and held central in the cylindrical tube. A metal cylinder (generally molybdenum) A is held securely by the spiral springs fitting tightly between the outside of the metal cylinder and the inner wall of the glass tube. If the filament of such a tube is sufficiently heated to evaporate plenty of electrons, the current to the anode is given by

$$i = 14.65 \times 10^{-6} \frac{l}{r\alpha} V^{3/2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (123)$$

in which i = current in amperes;
l = length of anode in centimeters;
r = radius of anode in centimeters;
V = voltage between filament and anode;
α = a constant depending upon relative radii of filament and anode. It is practically unity.

Now if a uniform magnetic field is passed axially through the tube (parallel to filament) it will be found that when the field strength reaches a critical value the anode current will suddenly fall to zero. The value of this critical field strength is given by

$$H = \frac{6.73}{r} V^{1/2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (124)$$

in which H is the field strength in gausses.

In Fig. 243 are shown the characteristics of a magnetron having an anode 4 in. in diameter and 12 in. long. With 400 volts in the anode the current shuts off with a field strength of 27 gausses, and with 1600 volts it shuts off with a field strength of 53 gausses. This satisfies the relation given in Eq. (124).

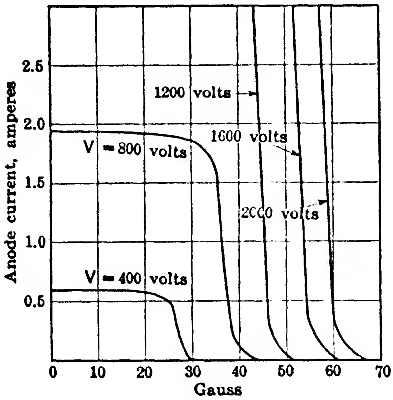


FIG. 243.—A magnetic field, parallel to the axis of the magnetron, controls the flow of electrons from filament to plate as shown here.

It is possible to make a diode of this kind produce oscillations, if connected up in some such fashion as shown in Fig. 244.

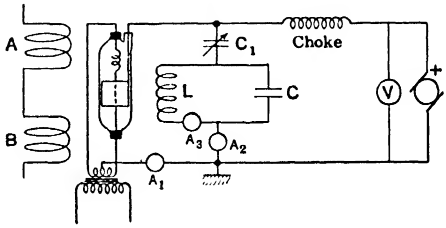


FIG. 244.—The circuit arrangement for producing alternating current by magnetron action.

Coils A–B are Helmholz coils, actually surrounding the magnetron and properly designed and placed to produce a parallel uniform field over a cylindrical space larger than the magnetron. The coil L is also around the magnetron inside the others. The outer coils are adjusted to give a magnetic field strength nearly great enough to shut off the anode current, then oscillations will be set up in the L–C circuit.

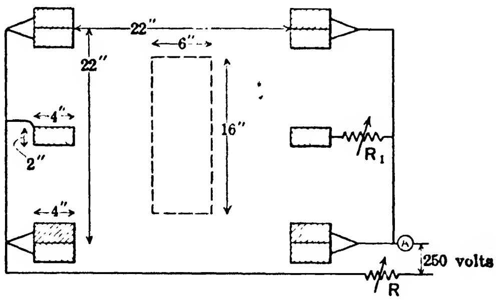


FIG. 245.—An arrangement of Helmholz coils for neutralizing the earth's magnetic field. Inside the cylindrical space 6″ in diameter and 16″ long, the earth's field can be reduced to zero within 1 per cent.

Elder[1] has analyzed this circuit and shows the conditions for oscillations, amounts of power obtainable, etc. He designed a set of Helmholz coils which gave him a uniform field over a volume 6 in. in diameter and 16 in. long, using 5 coils of 4000 turns each of 0.04-in. diameter copper wire. The coils were arranged as shown in Fig. 245. Each coil had a winding cross-section of 2 in. by 4 in.; each

[1] I.R.E., April, 1925.

weighed 125 lb. and had 172 ohms resistance. The center coil had an extra resistance R, used for equalizing the field more completely. In the 6 in. by 16 in. volume the field was uniform to better than 1 per cent. He states that a group of coils like this gives a flux density of 45.3 gausses per ampere drawn from the line (about 0.2 ampere in each coil). Using a magnetron giving 4 amperes of emission from its filament, having an anode 4 in. diameter by 12 in. long, with the tuned circuit adjusted for 12,000 meters, he was able to get powers and efficiencies as in the accompanying table:

Volts on anode	Amperes to anode	Input in kw.	Output in kw.	Losses on anode	Efficiency
3000	1.0	3.0	1.55	1.45	52
5000	1.0	5.0	3.25	1.75	65
7000	1.1	7.7	5.10	2.60	66
9000	1.3	11.7	8 05	3.65	69

Power-supply for Plate Circuits of Power Tubes.—Of course it is possible to build high-voltage generators to deliver the power required by the plate circuits of high power tubes; generators of 20-kw. capacity are available to give voltages as high as 10,000. It is now found more economical, however, both in first cost and maintenance cost, to use rectified a.c. power, instead of high-voltage c.c. generators.

Raguet [1] estimates that for a 100-kw. power supply, used 20 hours a day, the rectifier outfit saves over $4000 a year when compared with motor generator sets. The rectifier outfit is also cheaper to buy than the motor generator sets.

There are two characteristics every rectifier should have: plenty of current capacity, and a low voltage drop in the conductive direction and a high " breakdown " voltage in the reverse direction (this reverse voltage is generally called the *inverse* voltage). The ordinary high-vacuum diode is a perfect rectifier so far as rectification is concerned, but it requires a high voltage across the tube even when carrying current in the normal direction. This " space charge drop " may be only a few volts for a small low-powered valve, but for valves used in high-power circuits it may be thousands of volts.

The hot cathode mercury vapor rectifier is much more satisfactory; it requires but small drop through the tube, and its breakdown voltage in the reverse direction, though not as high as that of a vacuum tube, may be made as high as necessary by regulating the mercury vapor pressure. The relation between mercury vapor pressure and temperature is shown

[1] I.R.E., Jan., 1930, p. 49.

in Fig. 246. In Fig. 247 is shown in one curve the drop through one of these hot filament mercury vapor rectifiers, when carrying rated current, and another curve giving the "breakdown" or "flash-back" voltage, both curves being given as functions of the temperature of the coolest part of the tube. The temperature of this coolest part determines the mercury vapor pressure. The drop through the tube, in the conductive direction, must not exceed about 20 volts, and the flash-back voltage must be considerably greater than the tube will be called upon to stand. These two limits determine the temperature limits of operation of the tube, as indicated in Fig. 247. If the tube gets too cool the tube drop exceeds 20 volts and the filament is subject to destructive bombardment by positive ions; if the tube gets too hot its flash-back voltage is lowered, possibly to a value which permits reversed current to flow with probably a resultant short circuit in the power supply. This type of rectifier is now available in the following ratings:

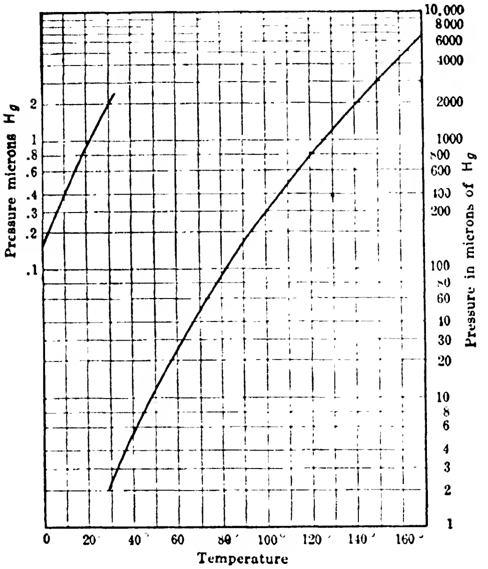


FIG. 246.—Pressure-temperature curve of mercury vapor.

Type	Filament		Peak anode current	Flash-back voltage
	Volts	Amperes		
866	2.5	5	0.6	5,000
872	5	10	2.5	5,000
869	5	20	5.0	20,000
857	5	60	20.0	20,000

Six of the type 869 properly connected will deliver 100 kw. and six of the 857 properly connected will deliver 400 kw. of rectified power. The life of these tubes is expected by the manufacturer to be many thousands of hours.

In Fig. 248 is shown a simple single-phase full-wave rectifier and the voltage across one tube during one cycle of the power supply voltage. A small transformer N with good insulation between its primary and secondary, supplies filament current for both tubes; the center point O of its secondary serves as one terminal of the rectified power supply. Transformer M furnishes a secondary voltage (effective) about $2\frac{1}{2}$ times as much as the desired voltage of the rectified power supply. Valve A carries current when end a of the transformer M is positive and the drop through valve A is only 15–20 volts. Hence by inspection of this connection diagram the drop across valve B must be the total voltage of the secondary winding, minus the few volts' drop through A. Thus the peak inverse voltage for this circuit is equal to $(E\sqrt{2}-20)$, the average voltage across the load, $E_{av.}$, is the average of a series of sine wave alternations, all in the same direction as shown in Fig. 249. If we allow 20 volts' drop in the valve the average value is $(2/\pi)\{(E_{max}/2)-20\}\cong 0.31\ E_{max.}\cong 0.44\ E.$ Ordinarily a low-pass filter of one or two sections would be used between the valves and the load, and of course this would decrease $E_{av.}$ by the RI drop in the choke coils.[1]

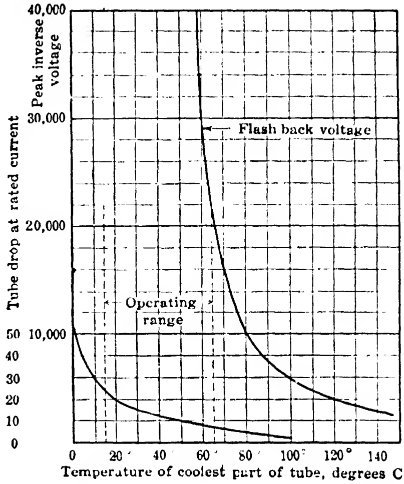


Fig. 247.—Action of mercury vapor in a hot-cathode, mercury vapor rectifier.

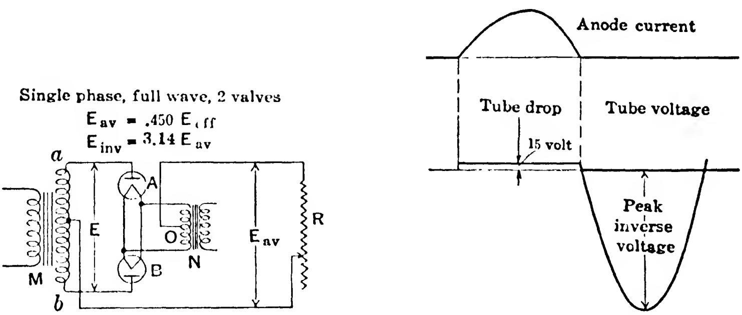


Fig. 248.—Action of a single-phase full-wave rectifier.

[1] For a good analysis of this or the following circuits see article by Steiner and Mason, I.R.E., Jan., 1930, p. 67. Also one by R. W. Armstrong, I.R.E., Jan., 1931, p. 78.

In Fig. 250 is shown a single-phase, full-wave, four-valve rectifying circuit. By this scheme the ratio of peak inverse voltage to the load voltage is much reduced from the value it has for the two-valve arrangement of Fig. 248. Thus for the given type of valves the allowable load voltage is twice as much as it is for the two-valve circuit; furthermore, the plate-circuit transformer is much more efficiently used in Fig. 250 than in Fig. 248. In other words, for a given load current, the heating of the secondary of the plate-circuit transformer is much less for the four-valve than for the two-valve circuit. To appreciate this point it is necessary to consider the heating effect of currents consisting of various shaped pulses, all giving the same average value. It will be found that the shorter the pulse (expressed as a fraction of a cycle) the greater will be the heating value, for a given average value.

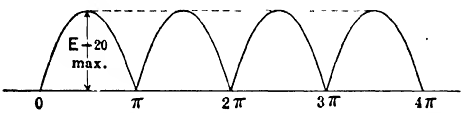


FIG. 249.—Load voltage of the circuit of Fig. 248.

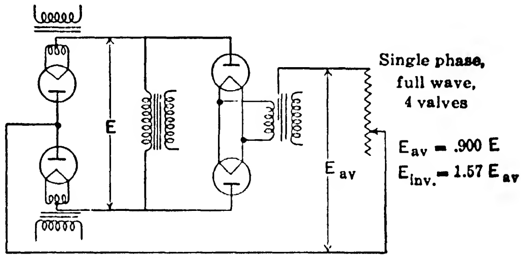


FIG. 250.—Single-phase full-wave rectifier using four valves. Its inverse voltage is only half that of Fig. 248, in terms of load voltage.

Fig. 251 shows a three-phase, half-wave, three-valve circuit; the evident advantage of the three-phase, over the single-phase, is a smaller fluctation in the pulsations of the rectified current, hence smaller and cheaper filter required to give a non-fluctuating load current.

Fig. 252 shows a three-phase, half-wave, double Y arrangement using six valves. Each secondary coil of the three-phase transformer is in two parts, and the load current divides equally between these two halves and flows through them in such a way that the m.m.f. due to the load current is reduced to zero. This prevents the rectified current from magnetizing the transformer core, an effect which is very troublesome in some of the other circuits.

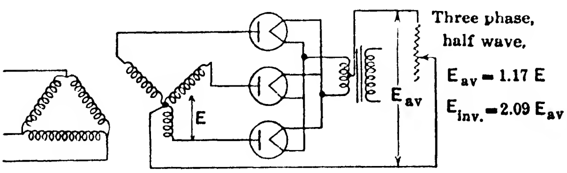


FIG. 251.—Three-phase, half-wave, rectifier circuit.

Fig. 253 shows a three-phase, full-wave, six-valve circuit which is generally used with mercury vapor rectifiers; its advantage lies in the very low inverse voltage impressed on the valves, thus permitting a load voltage almost as high as the flash-back voltage of the valve.

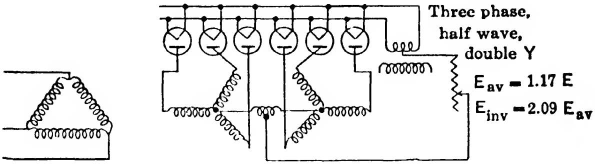


FIG. 252.—Three-phase half-wave having such connections that the rectified current does not magnetize the transformer cores.

In Fig. 254 are given curves showing the active parts of the various voltage cycles (in heavy lines) and the valves which participate in the production of load current for the part of the cycle. It is assumed that

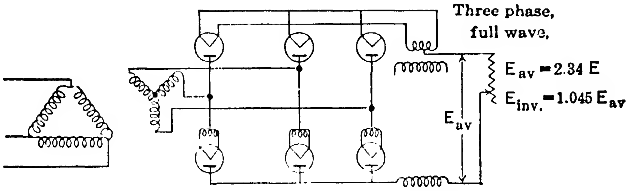


FIG. 253.—Three-phase full-wave rectifier. Here the inverse voltage is not much greater than the load voltage.

voltage of the power supply follows the phase sequence shown in the curves, namely C–A leads B–C, and this leads A–B.

In Fig. 255 is shown the form of current for a single-phase, full-wave

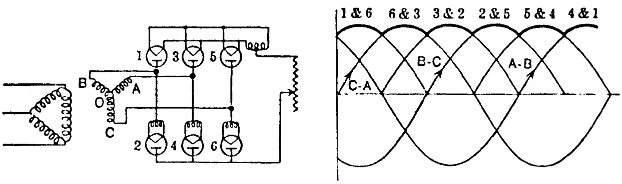


FIG. 254.—This diagram shows the successive action of the various valves used in Fig. 253.

rectifier having a one-section filter, the condenser being next to the rectifier; the peak value of anode current is much greater than the average load current. In Fig. 256 are shown the current and voltage curves for one valve of the circuit of Fig. 253, when a choke coil is used in series with the load; the effect of the coil is to prevent the anode current from changing so rapidly so that it is more rectangular in form than for Fig. 255. The value of the maximum anode current in this case does not greatly exceed the load current.

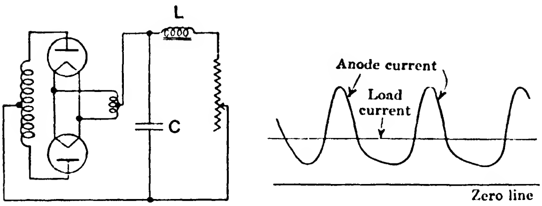


FIG. 255.—A simple filter used with a single-phase full-wave rectifier. The maximum current through the valve is much greater than the load current.

The comparative merits of the hot cathode mercury vapor valve, compared with the high-vacuum valve, are given by the following table, which shows the performance of the two types. The last row of figures shows the performance with the vapor valves at rated capacity.

No valves	Type circuit	Tube drop		Losses		Output			Efficiency
		Volts at amps		Filament	Plate	E	I	Kw	
6	3 ϕ Double Y	1560	6	6.9 kw.	18 7 kw.	15,000	12	180	87 5
6	3 ϕ Full wave	15	12	1 8	0 36	15,000	12	180	98 8
6	3 ϕ Full wave	15	20	1.8	0 6	19,100	20	382	99 4

In addition to the better efficiency of the mercury vapor valves the voltage regulation is very much better. From no load to full load the load voltage drops 20 per cent with the high-vacuum valves and only 8–10 per cent with the vapor valves. This regulation includes that of the transformers.

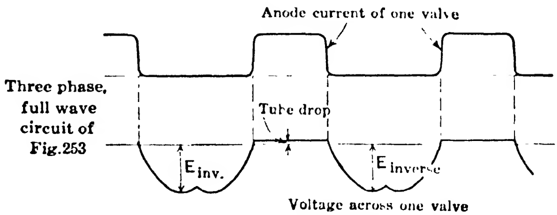


FIG. 256.—Voltage and current relations of one valve of Fig. 253.

Because of the effect of rectified current wave forms, some of the transformers used in the rectifying circuit arrangements shown above have only a fraction of their capacity for normal a.c. operation. The "utilization factor" is only 68 per cent for some of the connections and as much as 96 per cent for others. The transformers used in the three-phase, full-wave arrangement of Fig. 253 will carry 96 per cent as much power as they would in normal a.c. operation.

A Special Type of Oscillating Circuit.—It is possible to produce a type of oscillating circuit with no inductance at all; one of these schemes, called the "multi-vibrator," originally due to Abraham and Bloch, is shown in Fig. 257. In such a circuit the two condensers alternately charge and discharge periodically; the voltage across one condenser, and plate current of one tube, having about the forms shown in Fig. 258. These square pulses of current follow one another at very definite time intervals. Watanabe[1] attempts an analysis of the action of this circuit and derives a relation between circuit constants and periodicity. Armagnat has given the period, if we take $C_2R_2 = C'_2R'_2$, as $T \cong C_2R_2$, whereas van der Pol calculates it to be more closely represented by $T = (\pi/2)\ C_2R_2$ for the same circuit constants.

Because of the very sharp rise and fall of plate current this type of circuit furnishes a power source with a long series of harmonics; thus it is possible to detect energy even as high as the fiftieth harmonic when the pulse has sharp corners as indicated in Fig. 258. Thus if the fundamental

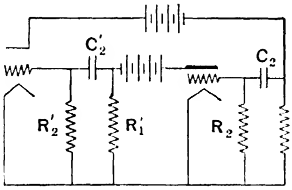


FIG. 257.—The so-called "multi-vibrator" circuit arrangement using two triodes.

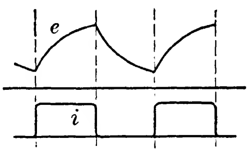


FIG. 258.—Voltage and current relations of the circuit shown in Fig. 257.

is set at some measurable, audible value, this circuit furnishes a scheme for obtaining known radio frequencies.

Characteristics of the Three-electrode Tube as an Amplifier.—As the voltage impressed on the input circuit of a tube causes a change in the plate current which may be flowing through an inductance or resistance in the external plate circuit, and as it is evident that the drop across this external circuit may be many times greater than the e.m.f. impressed on the grid, the device may be used as a voltage amplifier. The amplified voltage in the output circuit will, under proper conditions, have very nearly the same form as the input voltage, sufficiently so that the currents due to speech may be amplified many times (1,000–10,000) and the reproduction of the voice be almost perfect. The circuits used in amplifiers and arrangement of apparatus are taken up in a later chapter; in this section we shall consider only the amplifying characteristics of the tube itself.

As noted before, a voltage of $E_{m_g} \sin \omega t$ introduced in the input circuit of a tube is equivalent to a voltage of $\mu E_{m_g} \sin \omega t$ introduced into the plate

[1] I.R.E., Feb., 1930, p. 327.

circuit, this voltage causes an alternating current to flow in the plate circuit, the magnitude and phase of which depend upon the external impedance in the plate circuit and the resistance of the tube itself. If we call the alternating component of the plate current I_p we have the relation,

$$I_p=\frac{\mu E_g}{R_p+R}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (125)$$

in which R is the external resistance of the plate circuit. The drop across R (which is the only available part of the amplified voltage, the rest being used up inside the tube itself) is I_pR and this is evidently given by

$$I_pR=E_p=E_g\mu\frac{R}{R_p+R}.$$

From this we get the actual voltage amplification due to the tube, which is designated as α,

$$\alpha=\mu\frac{R}{R_p+R}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (126)$$

This factor α will be constant (independent of the magnitude of the input voltage E_g) only for such value of E_g as give constant R_p. This can be seen at once from the static characteristic of a tube, showing the relation between plate current and grid potential, this curve to be taken with the proper value of R in the plate circuit; throughout that part of this curve which gives uniform slope the factor μ is constant and the amplification is *distortionless*, a very necessary feature of an amplifier used for speech amplification, but of little importance for telegraph signal amplification.

This point is indicated in Fig. 259; two different tubes (or different arrangements of the same tube) might have characteristics as shown at M and N and the form of the plate current produced for a sine wave of voltage impressed on the grid as shown by curves m and n in the same figure. With curve N the value of dI_p/dE_g is greater the more positive the grid becomes, resulting in a lower R_p; from Eq. (126) it may be seen that for a given value of R the factor α becomes greater the smaller R_p. A sine wave of voltage impressed on the grid, therefore, does not produce a sine wave of current in the plate circuit and so will not produce a sine wave of voltage across a resistance in the plate circuit.

From Eq. (126) it is evident that if the amplifying power of a tube is to be efficiently used the value of R must be at least as large as R_p and should really be much larger. In Fig. 260 is shown the measured amplification constant of a small tube taken under various conditions. It is seen that the factor α increases as R increases, for all conditions.

Curve 1 was taken with a constant "B" battery voltage; under this condition the plate voltage decreased as R was increased due to the resistance drop in R. But a decreased plate voltage resulted in an increase

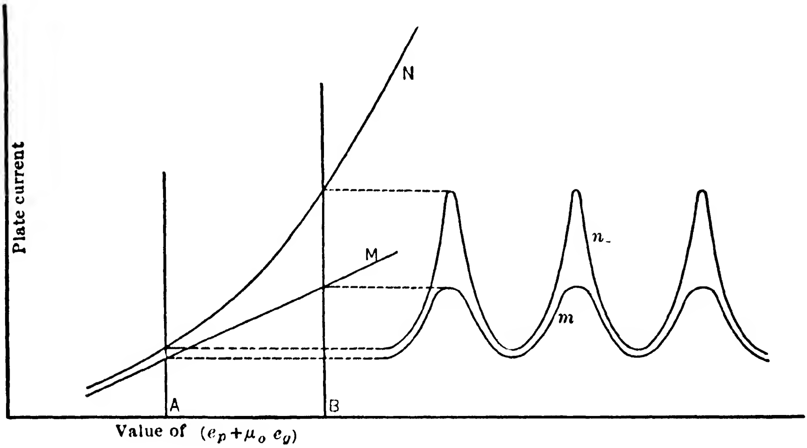


FIG. 259.—Two tubes having different plate-current characteristics as indicated in M and N will give amplified currents having more or less distortion, N giving more distortion than M.

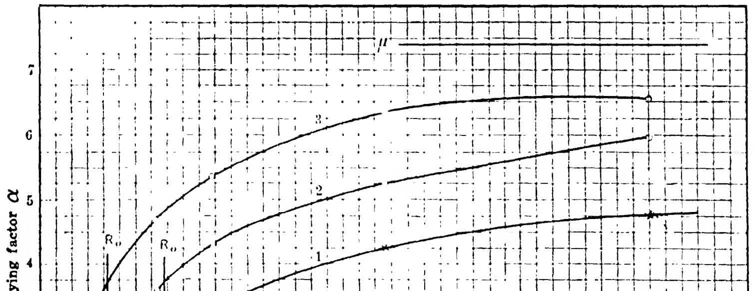


FIG. 260.—Amplifying power of a small receiver tube; for curve 1 the plate voltage was allowed to fall, E_b being constant at 20 volts; for curves 2 and 3 the value of E_b was variable and so adjusted to give a plate voltage of 20 or 30 for all values of external resistance with the plate circuit.

in R_p, so that for this condition as R was increased, it approached R_p very slowly due to the increase in R_p with increase in R.

Curve 2, compared to curve 1, shows the effect on α of increasing the "B" battery voltage sufficiently to compensate for the IR drop in the plate circuit, maintaining the average plate voltage constant; it is seen that the increase of α with R is much more rapid. Curve 3 shows the effect of maintaining the plate potential at 30 volts instead of 20 volts.

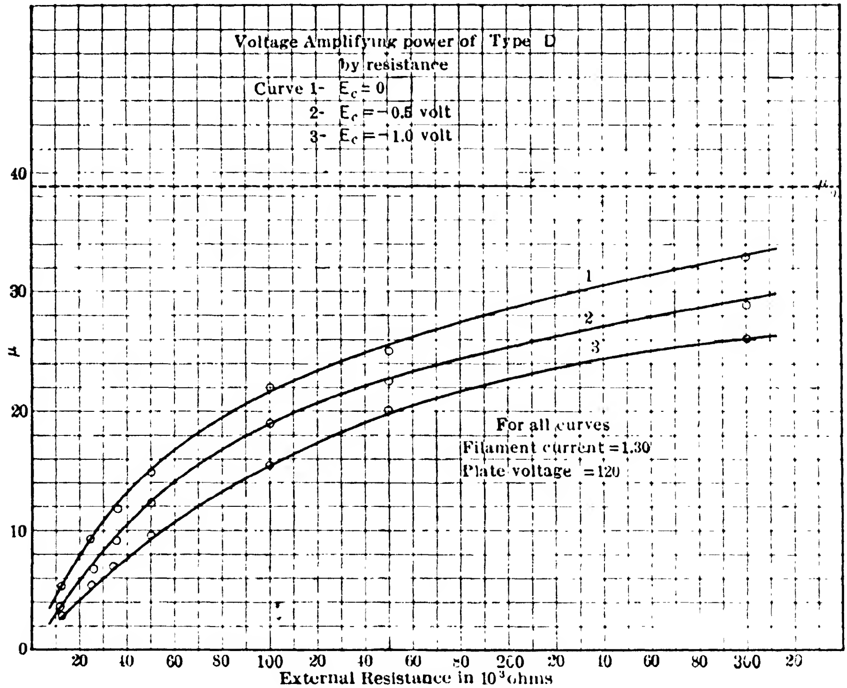


FIG. 261.—Showing effect on the amplifying power of a tube of holding the grid at different average potentials; making the grid more negative increases the tube resistance, hence requiring a higher external resistance to get the same amount of amplification.

The a.c. resistance of the plate circuit of the tube R_p was measured for curves 2 and 3 and is indicated in the curves; it is seen for each of them that when $R = R_p$, $\alpha = \frac{1}{2}\mu$, as it should from Eq. (126).

In using a tube as an amplifier it is customary to maintain the grid at such a negative potential that, for any probable input voltage, the grid will not become positive; maintaining a negative grid increases the value of R_p, so that for a given R, α is decreased. This effect is shown in Fig. 261, which gives the behavior of a tube having a higher value of μ than is customary. Fig. 262 shows the variation of the amplification factor as both plate circuit voltage, E_b, and grid potential, E_c, were varied. It is evident that this tube could be used effectively for only small values of input voltage.

If the plate current of a tube is expressed by the relation,

$$i_p = A\left\{E_p + \mu(E_c + E_{mg}\sin\omega t)\right\}^2, \quad . \quad . \quad . \quad . \quad . \quad (127)$$

we get after expansion

$$i_p = A(E_p + \mu E_c)^2 + 2A\mu(E_p + \mu E_c)E_{mg}\sin\omega t$$
$$+\mu^2\frac{AE_{mg}^2}{2}\cos(2\omega t + \pi) + \frac{\mu^2 AE_{mg}^2}{2}. \quad . \quad . \quad (128)$$

The first term gives the steady value of plate current with no input voltage, the second the true amplification current, the third a double-frequency distortion current, and the fourth a steady increase in the value of i_p, while $E_{mg}\sin\omega t$ is acting. The third term has the same coefficient as the fourth and the fourth term will register on a d.c. ammeter in the plate circuit. Hence the quality of amplification of a tube (distortionless or not) may be judged by the indication of the plate ammeter as the input voltage is impressed. Fig. 263 shows this effect and also the effect of added resistance in the plate circuit in decreasing the distortion. With 150,000 ohms added in the plate circuit, this tube would give essentially distortionless amplification for input voltage as high as 5 volts.

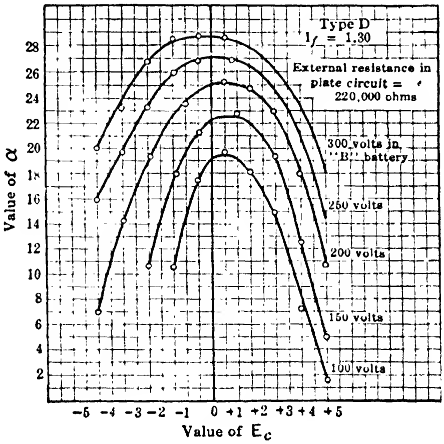


FIG. 262.—Variation in amplifying power with different grid potentials and different plate circuit voltages.

In case a reactance is used in the plate circuits, for repeating, instead of resistance, it will be found that the value of α is greater than for a corresponding value of resistance. Thus if an inductive reactance (of negligible resistance) is used in the plate circuit, the value of the reactance being equal to the tube resistance, a value of α is obtained equal to 0.7μ instead of 0.5μ as was obtained for resistance. This follows at once by considering the voltage relations in a tube circuit, as given in Fig. 131, p. 577.

Characteristics of Tubes Used at Present.—As indicative of practice in tube construction to-day there is appended a table of typical tubes used in receiving sets.

DETECTORS AND AMPLIFIERS (RCA)

Tube	Use	C	E_f	I_f	S	E_b	E_c	E_s	I_p	R_p	g_m	μ	r_p	P
UX112-A	G. L. Det.	OF	5.0	0.25	D. C.	90	+F							
	Amp.					90	4.5		5.2	5,600	1500	8.5	5,600	30
	Amp.					135	9.0		6.2	5,300	1600	8.5	8,700	115
UX199	G. L. Det.	TF	3.3	0.06	D. C.	45	+F							
	Amp.					90	4.5		2.5	15,500	425	6.6	15,500	7
UX200-A	Det.	TF	5 0	0 25	D. C.	45	−F		1.5	30,000	666	20		
UX201-A	G. L. Det.	TF	5.0	0.25	D. C.	45	+F							
	Amp.					90	4.5		2.5	11,000	725	8.0	11,000	15
	Amp.					135	9.0		3.0	10,000	800	8.0	20,000	55
UX222	A. F. A.	TF	3.3	0.132	D. C.	180	1.5	22.5	0.3	2,000,000	175	350	250,000	
	R. F. A.					135	1.5	45	1.5	850,000	350	300		
	R. F. A.					135	1.5	67.5	3 3	650,000	480	290		
UY224	B. Det.	OH	2.5	1.75	A. D. C.	275	5	20–45	0.1				250,000	
	A. F. A.					250	1 0	25	0.5	2,000,000	500	1000	200,000	
	R. F. A.					180	1.5	75	4 0	400,000	1050	420		
	R. F. A.					180	3.0	90	4.0	400,000	1000	400		
	R. F. A.					250	3 0	90	4 0	600,000	1025	615		
UX226	Amp.	OF	1.5	1.05	A. D. C.	90	5 0		3 8	8,600	955	8.2	9,800	30
	Amp.					135	8 0		6 3	7,200	1135	8.2	10,500	80
	Amp.					180	12 5		7 4	7,000	1170	8.2	10,500	180
UY227	G. L. Det.	OH	2.5	1.75	A. D. C.	45	Cath.							
	B. Det.					275	30		0 2				50,000	
	Amp.					90	6 0		2 7	11,000	820	9.0	14,000	30
	Amp.					135	9 0		4 5	9,000	1000	9.0	13,000	80
	Amp.					180	13 5		5.0	9,000	1000	9.0	18,700	165
	Amp.					250	21.0		5.2	9,250	975	9.0	34,000	300
UX230	G. L. Det.	OF	2.0	0.06	D. C.	45	+F							
	Amp.					90	4 5		1.8	13,000	700	9.3	15,000	16
UX232	B. Det.	OF	2.0	0 06	D. C.	175	6	67.5	0.2				100,000	
	A. F. A.					180	1 0	22.5	0.25				250,000	
	R. F. A.					135	3.0	67.5	1.4	1,150,000	505	580		
UY235[1]	R. F. A.	OH	2.5	1.75	A. D. C.	180	1 5	75	5.8	350,000	1100	385		
	R. F. A.					250	3.0	90	6.5	350,000	1050	370		
UY236	R. F. A.	OH	6.3	0.3	D. C.	90	1 5	55	1.8	200,000	850	170		
	R. F. A.					135	1 5	67.5	3.0	300,000	1050	315		
UY237	G. L. Det.	OH	6.3	0.3	D. C.	45	Cath.							
	Amp.					90	6.0		2.6	11,500	780	9.0	17,500	30
	Amp.					135	9.0		4.3	10,000	900	9.0	14,000	80

UY239[2]	R. F. A.	OH	6.3	0.3	D. C.	90	3.0	90	4.4	375,000	960	360		
	R. F. A.					135	3.0	90	4.4	540,000	980	530		
	R. F. A.					180	3.0	9.0	4.5	750,000	1000	750		
UX240	Volt A.	TF	5.0	0.25	D. C.	135	1.5		0.2	150,000	200	30	250,000	
	Volt A.					180	3.0		0.2	150,000	200	30	250,000	
						POWER AMPLIFIERS (RCA)								
UX112-A	A. F. P. A.	OF	5.0	0.25	A. D. C.	135	9 0		6.2	5,300	1600	8.5	8,700	115
	A. F. P. A.					180	13 5		7.6	5,000	1700	8.5	10,800	260
UX120	A. F. P. A.	TF	3.3	0.132	D. C.	90	16 5		3.0	8,000	415	3.3	9,600	45
	A. F. P. A.					135	22 5		6.5	6,300	525	3.3	6,500	110
UX171-A	A. F. P. A.	OF	5.0	0.25	A. D. C.	90	16.5		12.0	2,250	1330	3.0	3,200	125
	A. F. P. A.					135	27.0		17 5	1,960	1520	3.0	3,500	370
	A. F. P. A.					180	40 5		20.0	1,850	1620	3.0	5,350	700
UX210	A. F. P. A.	TF	7.5	1.25	A. D. C.	250	18 0		10 0	6,000	1330	8.0	13,000	400
	A. F. P. A.					350	27 0		16 0	5,150	1550	8.0	11,000	900
	A. F. P. A.					425	35 0		18.0	5,000	1600	8.0	10,200	1600
UX231	A. F. P. A.	OF	2.0	0.13	D. C.	135	22 5		6.8	4,950	760	3.8	9,000	150
UY233[3]	A. F. P. A.	OF	2.0	0.26	D. C.	135	13 5	135	14 0	50,000	1500	75	7,000	700
UY238[3]	A. F. P. A.	OH	6.3	0.3	D. C.	135	13 5	135	9.0	102,000	975.	100	13,500	525
UX245	A. F. P. A.	OF	2.5	1 5	A. D. C.	180	33 0		27.0	1,900	1850	3.5	3,500	780
	A. F. P. A.					250	48 5		34.0	1,750	2000	3.5	3,900	1600
	A. F. P. A.					275	54 5		36.0	1,670	2100	3.5	4,600	2000
UY247[3]	A. F. P. A.	OF	2.5	1.75	A. D. C.	250	15 0	250	32.0	35,000	2500	90	7,000	2500
UX250	A. F. P. A.	OF	7.5	1.25	A. D. C.	250	41 0		28 0	2,100	1800	3.8	4,300	1000
	A. F. P. A.					350	59 0		45 0	1,900	2000	3.8	4,100	2400
	A. F. P. A.					400	66 0		55 0	1,800	2100	3.8	3,670	3400
	A. F. P. A.					450	80 0		55 0	1,800	2100	3 8	4,350	4600

NOTES

C—Type of cathode
TF—thoriated filament
OF—Oxide-coated filament
OH—Oxide-coated heater
E_f—Filament voltage
I_f—Filament current (Amp.)
S—Filament supply
A.D.C. = A.C. or D.C.
E_b—Plate supply voltage
E_c—Negative grid bias voltage
This is with respect to $-F$, or cathode, and should be increased by one-half E_f for A.C. Filaments.

E_s—Screen-grid voltage
I_p—Plate current (MA)
R_p—A.C. plate resistance (ω)
g_m—Mutual or trans-conductance (μ_M)
μ—Voltage amplification factor
r_p—Plate load or coupling res. (ω)
P—Maximum "undistorted" power output (MW)
Amp.—General purpose amplifier
A. F. A.—Audio-frequency amplifier

R. F. A.—Radio frequency amplifier
Det.—General purpose detector
G. L. Det.—Grid leak detector
B. Det.—Biased or power detector
Volt. A.—Voltage amplifier:
[1] Super-control or "variable μ" screen-grid tube
[2] Super-control or "variable μ" pentode
[3] Power pentode

Multi-Function Tubes.—Due primarily to the desire for compactness in receiving sets it is evident that tubes having sufficient elements to perform two or more functions are sure to appear. Thus the type 55, just announced by the manufacturer, is described as a duplex-diode triode; it has an indirectly heated cathode which furnishes electrons for three separate circuits, performing entirely different functions.

Near one section of the cathode are located a grid and plate in their normal relationship, this part of the tube serving as a normal audio-frequency amplifier. One plate, near another part of the cathode, serves with the cathode as a Fleming valve, that is, a two-element detector tube. As mentioned in the section on detectors, the Fleming valve is not as efficient a detector as the three-element one but it has the advantage that it can handle high-voltage signals without danger of overloading, and, if sufficient resistance is put in the plate circuit, the rectification is reasonably linear over a wide range of signal strength.

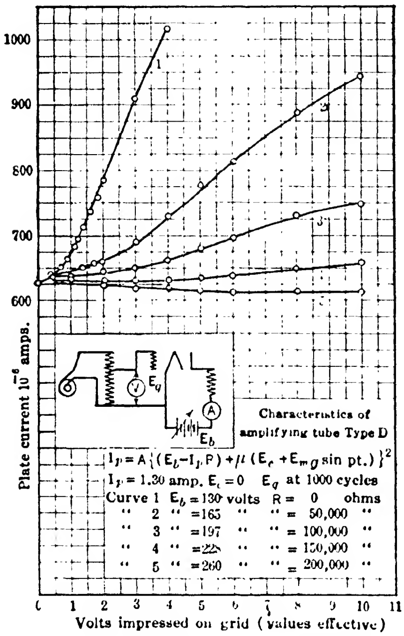


FIG. 263.—The quality of amplification (distortionless or not) is shown by a test of this kind. The tube used requires an external resistance in series with the plate of at least 150,000 ohms to give distortionless amplification.

Another plate acts as a volume control circuit serving to give a variable grid bias for the preceding radio-frequency amplifying tubes, the bias increasing with signal strength. Thus the type 55 tube serves as a two-element "power detector," an automatic volume control device, and a normal triode audio-frequency amplifier.

CHAPTER VII

CONTINUOUS-WAVE TELEGRAPHY

Advantage of Continuous-wave Telegraphy.—Continuous-wave telegraphy possesses several distinct advantages over damped-wave systems which may be summarized as follows:

1. *Greater Selectivity.*—This advantage is due primarily to the fact that energy radiated by a spark transmitter is sent out in damped-wave trains. These wave-trains, striking the receiving antenna, induce therein an electromotive force, and if the circuit is tuned to the incoming wave, maximum current and signal strength are obtained. However, even if the circuit is somewhat de-tuned, the damped-wave train will excite the circuit to a considerable extent, causing it to oscillate at its own frequency, as well as at the frequency of the signal wave.[1] In other words the selectivity of reception of a spark signal is fixed, not only by the decrement of the receiving circuit, but also by the decrement of the wave-train itself, which, of course, is that of the transmitting station; thus more or less interference always exists between spark stations, if the wave lengths are close to one another.

If we consider the effect of continuous waves at the receiving station, the conditions will be somewhat different. The incoming energy forces the receiving circuit to oscillate at its own signal frequency, except at the beginning, when the forced and natural oscillations are coexistent for a few cycles. Therefore if this circuit is not tuned to resonance with the incoming signal and does not possess abnormal resistance values (which would flatten out its resonance curve), the current flowing will be very small and the signal strength extremely weak, under all conditions of adjustment except that of resonance. Thus the selectivity is good, and the station will receive no messages except those for which it is tuned.

2. *Increased Range of Transmission.*—This follows from the fact that with continuous-wave transmission, the energy is radiated at and concentrated into, essentially one wave length, instead of being spread over a number of wave lengths, as indicated by the energy distribution curves discussed in Chapter V, p. 426. The greater the amount of energy we can thus concentrate into one wave length, the further will be the distance penetration or propagation of this energy, and stations may be

[1] See Chapter III, p. 351.

reached at much greater distances from the sending station than with the spark transmitter.[1] Also, for the same range, less power is required than with the spark transmitter, and the transmission efficiency thus improved.

3. *Antenna Voltages Decreased.*—Since the energy is radiated in a continuous stream, when a signal is being sent, and not in groups, it follows that for a given power in the antenna the amplitude of the oscillations need not be so great. For example, if we assume 1000 sparks per second, a decrement of 0.1, and a 300-meter wave, the time per second during which energy is radiated [2] is:

$$1000\times\frac{4.7}{0.1}\times 10^{-6}=0.047 \text{ second}$$

$$=4.7 \text{ per cent of the total time;}$$

whereas with continuous-wave transmission the time would be 100 per cent. It is therefore obvious that if much power is to be radiated by the damped wave-transmitter, comparatively high oscillation amplitudes must be used, that is, the energy associated with a group of waves, for a given amount of energy radiated per second, must be high, since energy is radiated only during a small fraction of the time. Thus a given antenna will have a greater possible energy radiation on continuous waves, since the energy may be radiated continuously. An advantage of thus decreasing the required amplitude of oscillation for a given radiation is the reduction in required voltage, thus decreasing the construction difficulties encountered in extremely high-voltage apparatus and antennas (due to corona losses, insulation requirements, etc.).

4. *Adjustment of Signal Note.*—With damped-wave transmission this characteristic is a fixed quantity which cannot be adjusted by the receiving operator, and is determined entirely by the transmitter group frequency. With the undamped-wave receivers (described on p. 634), this can be varied, over wide limits, to a value most suitable to the operator for distinguishing from strays. The adjustability of the note of the received signal also serves to a remarkable degree to eliminate interference from other stations; because of this feature another signal, differing in frequency from the true signal by perhaps 1 per cent, is actually inaudible.

Summary.—The above advantages combine to give to a continuous-wave transmitter a wonderful degree of selectivity and efficiency of trans-

[1] The statement is true primarily because of the greater sensitivity of the receiving circuit adjusted for continuous-wave reception. The attenuation which occurs as a wave travels over the surface of the earth is probably the same for continuous, as for damped, waves.

[2] On the assumption that radiation ceases when the current in the antenna has dropped to 1 per cent of its original value.

mission, very much higher than could be obtained with the damped-wave type. In addition may be mentioned the very important part which continuous waves have played in the development of radio telephony (see Chapter VIII), for which it is essential.

High-frequency Undamped-wave Generators.—Continuous high-frequency oscillations may be produced by any one of the several schemes described below; all of these have been commercially developed and applied. The development and importance of vacuum tubes as generators of high-frequency oscillations has been very rapid within recent years, and it is certain that this source of high-frequency power will soon replace all others.

The several means of high-frequency power generation are as follows:

(1) Poulsen Arc;
(2) Alexanderson Alternator;
(3) Goldschmidt Alternator;
(4) Medium frequency alternator and frequency transformer;
(5) Oscillating Tubes.

Poulsen Arc.—A great deal of work has been done in an effort to determine with exactness the action and theory of this type of generator, the best presentation being that of P. O. Pederson,[1] to which the reader is referred. In the discussion which follows, we have referred largely to his paper and to certain earlier theory as developed by Barkhausen, to which Pedersen also makes reference. Much of the laboratory work done in the past is not applicable to the modern arc generator, due to the wide divergence of the test arc and the arc generator as designed and constructed for commercial service.

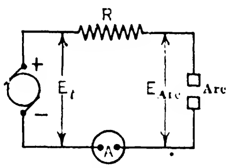


FIG. 1.—The ordinary arc has to have a stabilizing resistance in series with it, or else it is inoperative.

Elementary Theory—Instability of the Arc.—Consider the ordinary arc circuit indicated in Fig. 1 with the resistance R omitted for the present. The conduction of current through the arc is simply a case of conduction through an ionized gas, in this case vaporized carbon or copper at a very high temperature. Initially, this arc stream of ionized gas is not present, so that to start the current flow in the above circuit, it is necessary to bring the two electrodes in contact. The intense heat developed by the current passing through the point of contact vaporizes some of the electrode material, and as the electrodes are separated a vapor stream or arc of ionized gas is produced which forms a conducting path for the current.

[1] "On the Poulsen Arc and Its Theory," Proc. I.R.E., Vol. 5, p. 255, 1917.

The ionization is assisted by the high temperature of the arc and the bombardment of the negative electrode (cathode) by the positive ions, this dissociating the electrode into positive ions and electrons, the latter then being attracted to the positive electrode (anode).

As is the case for practically all gaseous conductors, the resistance of the arc decreases as the current increases. This will be apparent when it is recalled that the resistance of the arc depends on its state of ionization which, in turn, depends on the heating or vaporizing forces which act on the electrodes, caused by the current flowing through the circuit. Stated briefly, the more current we pass, the more ionized vapor we have, and the more ionized vapor, the " fatter " the arc and the lower the resistance. If the resistance decreases with current fast enough, the IR drop will decrease, even if the current increases, and this is always the case in the actual arc. The current and voltage relations of such a conductor would appear as shown in the curve of Fig. 2, known as the " static characteristic " of the arc; this name arises from the fact that it is obtained from a series of fixed (static) current values together with the corresponding voltages. (See Figs. 42 and 43 of Chap. II.)

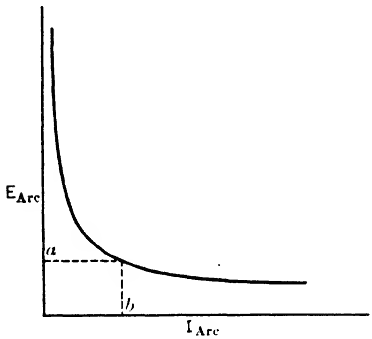


FIG. 2.—Relation of current and voltage across an ordinary arc.

Thus, if an arc were connected directly across a constant-potential supply, a short-circuit condition might be immediately attained, the voltage impressed always being above the value required for equilibrium and the current thus continually increasing to make $IR=E$. Since R is very small for a large current, the condition would be equivalent to a short-circuit. On the other hand, if the voltage impressed corresponded to a, Fig. 2, and the current started to decrease below the value b, then the arc current would continually decrease until the arc was extinguished.

Stabilizing Effect of Resistance.—The above phenomenon represents an unstable condition, in which the IR drop decreases automatically with increase in current. A stable circuit is one in which the IR drop increases with current, and to stabilize the arc it is necessary to add additional resistance R in the circuit, as indicated in Fig. 1. This resistance is the familiar " ballast " resistance used on all arc lamps, and the conditions now existing in the circuit are shown in Fig. 3, an inspection of which will indicate how the circuit has been stabilized.

Considering curve A, the operation is stable for all currents greater than I, and is unstable for currents below this value. Thus, on the stable portion of the curve, a decrease of voltage results in a decrease of current, and thus a decreased IR drop across the resistance. The voltage across the arc therefore rises, and initial conditions are restored. The current value may be controlled by varying the terminal voltage, as shown by the two current values a and b on the characteristic curve A, or by changing the amount of ballast resistance used, as shown by the current values a and a' for two different characteristics A and B for the same terminal voltage. If impressed voltage is reduced to the minimum value (E' for curve B) then the arc may operate or may go out. This point is therefore called the point of "indifferent" stability.

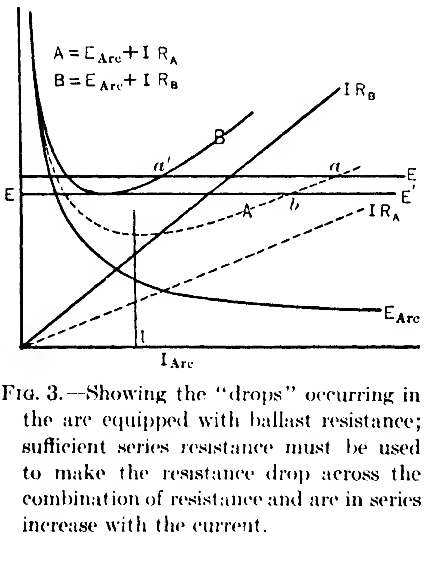


FIG. 3.—Showing the "drops" occurring in the arc equipped with ballast resistance; sufficient series resistance must be used to make the resistance drop across the combination of resistance and arc in series increase with the current.

The function of the ballast resistance is thus to stabilize the operation of the arc for slow changes of voltage. Commercial

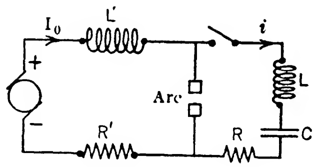


FIG. 4.—If a circuit of L and C in series is shunted around an arc supplied from a continuous-current generator through choke coil L', an alternating current will flow in the circuit made up of L, C, R, and arc in series.

arc generators have this resistance short-circuited when operating steadily, to increase the efficiency, the inductance and inherent resistance of the circuit being sufficient to stabilize the circuit.

Effect of Inductance—Choke Coils.—If a very high inductance is inserted in the circuit as shown at L' (Fig. 4) very quick changes of generator current are minimized and prevented to a large extent. Thus if a sudden increase of generator voltage occurred, the current would tend to increase, *and would increase slightly*, setting up a counter e.m.f. of self-induction in L', which would minimize the variation of current. It is important to note that the inductance, in order to be effective, requires a *variation* in the current flowing through it, and the arc supply is therefore not strictly a constant-current source. This variation may be made extremely small, however, by using large inductance values.

A Simple Explanation of the Operation of the Oscillating Arc.—If a condenser, in series with an inductance, is connected to a source of electric energy, of voltage E, the current which flows after closing the switch is an oscillatory one,[1] its frequency being fixed by the natural period of the oscillatory circuit and its magnitude depending upon the voltage E, and the ratio C/L. This oscillatory current dies away due to the damping, and the condenser is finally charged to a potential difference E, and there is no current in the circuit.

Suppose an arc, connected as in Fig. 4, is burning steadily (switch in the oscillatory circuit being open), with a difference of voltage across the two electrodes equal to E. When the switch is closed the current flowing in the L, C, R circuit is given by [2]

$$i=E\sqrt{\frac{C}{L}}\epsilon^{-at}\sin\omega t. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Now by inspection of Fig. 4 it is evident that this current must be robbed from the arc, because the value of L' is always chosen large enough to prevent rapid variations of current from the generator. Hence when the current i starts to flow the arc current starts to decrease, being always equal to I_0-i. But we have seen that it is a characteristic of the arc that when its current decreases the voltage across it increases; hence Eq. (1) does not correctly represent the current into the condenser, unless E is made to depend on i for its value. Actually the current is greater than indicated by Eq. (1), because of the increase in E during the time i is flowing in the direction indicated in Fig. 4.

The increase in E during the first alternation is not great, because the amount of current taken by the condenser is only a small fraction of the arc current. Thus if the arc is burning with 50 volts across the gap and a current of 10 amperes is flowing, the decrease in current upon first closing the switch (Fig. 4) will probably be less than 10 per cent of the arc current. The value of $\sqrt{L/C}$ used in the oscillating circuit should not be less than about 50; the normal value is perhaps 200. This value substituted in Eq. (1) shows that during the first alternation of the oscillatory state the maximum value of condenser current will be less than 1 ampere. The condenser will charge up to a voltage about twice that of the arc [3] and then start to discharge; the current during discharge adds to the current through the arc and thus gives it greater than normal value. This results in a *decrease in voltage across the arc*, thus tending to facilitate the discharge of the condenser, thus producing a greater discharge current than would have occurred if the arc voltage had held constant.

[1] For analysis of this action see Chapter III, p. 330.
[2] Eq. (11), p. 290; see also Fig. 39, p. 331.
[3] See Chapter III, p. 330.

It will thus be seen that the voltage-current characteristic of the arc tends to give a greater current in the condenser during both the charge and discharge periods, than would occur if the arc voltage were independent of the current through the arc.

Now the current flowing into the oscillatory circuit is supplied when the arc voltage is *higher than normal* and the current flows out of the oscillatory circuit (against the influence of the arc voltage) when the arc voltage is *less than normal.* As energy is being supplied to the oscillatory circuit during the charge and extracted during the discharge, from the conditions just cited it is evident that *more energy is supplied to the oscillatory circuit from the arc supply circuit during the charge than is given up*

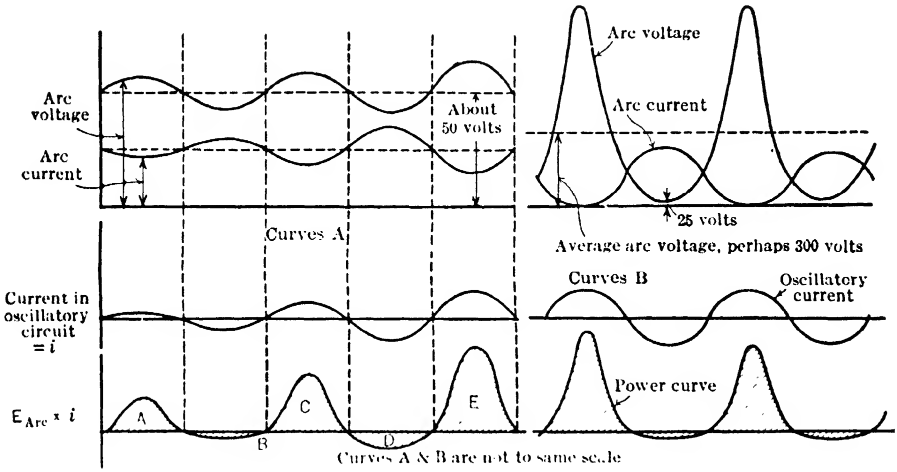


FIG. 5.—A simple explanation of the oscillating arc can be obtained from these curves; curves *A* show the start of oscillations and curves *B* give the conditions when the oscillations have reached the steady state.

to the arc during the discharge. Unless too great a resistance is present in the oscillatory circuit this action results in a building up of the current in the oscillatory circuit, and this building up will increase until the maximum value of the oscillatory current is practically equal to the generator current, I_0.

This action is shown by curves *A* of Fig. 5, which are nearly self-explanatory; the current in the oscillatory circuit is taken as positive when it is flowing in the direction indicated in Fig. 4. The lower curve of Fig. 5 is the product of the current i and the voltage acting on the oscillatory circuit. Area *A* gives the energy supplied to the oscillatory circuit by the arc during the first alternation, and area *B* gives the energy supplied by the oscillatory circuit to the arc during the second alternation. *The* *difference, A–B, gives the energy supplied to the oscillatory circuit during the complete cycle, and if this is greater than the I^2R loss in the oscillatory circuit during the cycle, the oscillatory current will continue to increase until some other factor controls the action.*

The excess of area A over area B depends upon the arc characteristic, being greater as the characteristic curve (Fig. 2) becomes steeper; as to whether or not the excess is sufficient to build up oscillations depends upon the resistance of the oscillatory circuit. These two factors control completely the operation of the arc; it must be remembered, however, that the relation between arc voltage and arc current used in plotting Fig. 5 must be determined from the oscillatory state because the static characteristic gives too great a difference in areas A and B. The variation between the static characteristic and dynamic characteristic increases with frequency, in such a way that at high frequency (say 500,000 cycles per second) the difference between areas A and B is not sufficient to produce much oscillatory power.

A simple arrangement of apparatus which has nearly the same action as the arc is shown in Fig. 6. A source of e.m.f. is connected to a resistance R which is fitted with a sliding contact, B. Between the lower point of the resistance R and the contact B is connected an oscillatory circuit consisting of L and C in series.

Suppose that, with B in the middle of R, switch S is closed; current will immediately start to flow as indicated by i, charging condenser C. Now as C starts to charge contact B is moved up on R, thereby increasing the voltage impressed on the L-C circuit. The motion of B is so regulated that it reaches B' in an interval just equal to one-quarter of the natural period of L-C; it then starts to move down on R and reaches point B'' in an interval equal to one-half of the natural period of L-C. Thereafter, the contact oscillates between B' and B'', making a complete cycle in a time equal to the natural period of L-C.

Such an arrangement will result in the building up of a large oscillating current in the L-C circuit, the magnitude being limited only by the voltage E and the resistance of the oscillatory circuit.

Types of Oscillation.—The types of oscillation which may be generated have been arbitrarily designated by Zenneck on the basis of the minimum arc current value as follows:

Type I. Minimum current is greater than zero.
Type II. Minimum current is equal to zero.
Type III. Type II with immediate re-ignition, resulting in production of trains of damped oscillations.

This classification is based upon the fundamental action of the arc; for details see the reference noted.

Normal Poulsen Arc.—The three classes of oscillations mentioned above do not in any case exactly apply to the operation of the modern Poulsen arc. Present generators of all capacities up to 1000 kw. (input) utilize what has been designated as the "normal Poulsen arc." In this arc the ratio of direct current in the supply circuit to the radio frequency current (effective value) is always nearly equal to the $\sqrt{2}$, or:

$$I_{dc} = \sqrt{2} I_{ac}$$

where I_{ac} is the effective value of the current in the oscillatory circuit.

The normal arc therefore represents the division limit between oscillations of type I and type II, its characteristics being somewhat similar to those of type I, as shown in Fig. 7.[1]

Professor Pedersen in his paper, previously referred to, emphasizes the importance of the extinction voltage on the characteristics of the normal arc. As the arc current approaches the zero value, the arc voltage suddenly rises, as shown by the lower curve of Fig. 7. The arc must be able to develop this voltage if operation is to be efficient. The early theory neglected this portion of the characteristic, principally because investigators worked with comparatively long arcs.

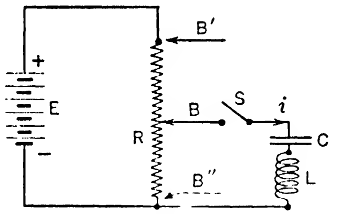


FIG. 6.—A simple circuit which may be made to operate the same as an oscillating arc.

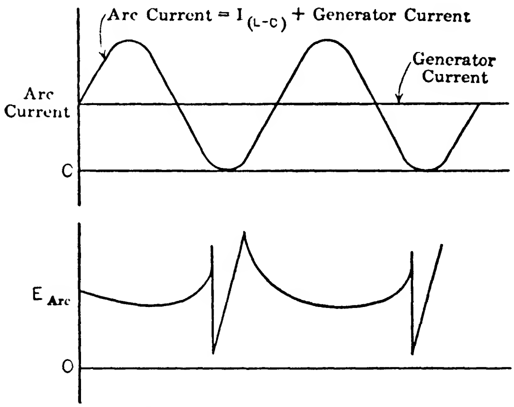


FIG. 7.—Voltage and currents for the "normal" arc.

It seems to the author that the curves of Fig. 5 represent the condition better than the curves shown in Fig. 7. If the action shown in Fig. 5 is continued for many cycles the difference between the arc voltage during charging of the oscillatory circuit condenser and that during discharge continually increases; this is indicated by curves B of Fig. 5, which represents the condition perhaps 100 cycles after the oscillations start.

It is to be noted that here the current has reached a steady value (fixed amplitude) and that the arc voltage pulsates between perhaps 25

[1] Zenneck, "Wireless Telegraphy," p. 236.

volts and a very high value, that corresponding to practically zero current in Fig. 2. Fig. 5 brings out a relation seldom mentioned by writers, that the reading of a c.c. voltmeter across the arc is about 50 volts before oscillations begin, but immediately jumps to 300 or more when oscillations start; in fact the reading may be as much as perhaps 500 volts if a sufficiently high-voltage power supply is used. A c.c. voltmeter reads average values, hence the change in reading from 50 volts to 300 volts indicates that the maximum voltage across the arc may be 1000 or more; as the duration of this high voltage is only a small fraction of a cycle its value may be three or four times the average value, i.e., three or four times as much as indicated by the reading of the c.c. voltmeter across the arc.

Practical Construction of the Arc Generator or Converter.—To increase the power, efficiency, and constancy of frequency of the arc generator, several special devices are employed. These devices are the result of extended investigations carried out by V. Poulsen, and their application has been primarily responsible for the rapid development and commercial success of this type of generator. Previous to this time many investigators had utilized the fundamental arc circuit, but had not succeeded in obtaining sufficient high-frequency energy to permit the operation being considered a practical success. Poulsen's investigations offered the first solution and demonstrated that the simple arrangement of Fig. 4 would give undamped oscillations at constant radio frequencies and sufficient energy for the purposes of radio telegraphy and telephony if modified as follows:

1. The arc is caused to take place in hydrogen, or a gas rich in hydrogen.

2. The positive electrode is kept as cool as possible, and therefore is constructed of some material having a high heat conductivity, usually copper, cooled by circulating water. The negative electrode is of carbon and is rotated slowly on its axis while the arc is in operation, to improve the regularity of the oscillations.

A water-cooled shoe, placed in close sliding contact with the cylindrical surface of the carbon electrode at its tip, has also been found [1] to improve the regularity and steadiness of the arc by limiting its travel along this electrode. This arrangement also permits "one peak" operation as shown in Fig. 5 instead of "two peak" operation as shown in Fig. 7. A purer wave is thereby radiated, minimizing interference caused by harmonics.

3. The arc is acted upon by a transverse magnetic field, which assists in the rapid de-ionization (scavenging) of the gases in the arc. The electromagnets supplying this magnetic field are sometimes connected directly in the supply circuit as indicated schematically in Fig. 8.

The strength of this field affects the characteristics of the arc to a

[1] See "Some Improvements in the Poulsen Arc." P. O. Pedersen, I.R.E., Oct., 1921.

considerable extent, and if not of the correct value, inefficiency and inconstancy of oscillation result. As expressed by Professor Pedersen: "The arc should burn in the weakest field in which it works normally, only igniting once a period, and always on the electrode edges. Both stronger and weaker fields require excessive supply voltage. This is therefore the most suitable field intensity—the one giving the highest efficiency and the most constant behavior of the arc."

If the field is made too strong, the increase in resistance, due to the stretching out of the arc as it is being extinguished, causes the extinction voltage (the potential across the arc at the instant the arc current has fallen to zero) to reach excessive values. This excessive voltage may cause a re-ignition of the arc across a shorter path and interfere with the constancy of the oscillations.

If the field is too weak, the conditions for successive arcs (in successive cycles) are not constant, due to the fact that ignition does not take place from the same point of each electrode, but from the points where the preceding arc existed at the instant of being extinguished. The arc length thus grows successively longer and longer until the arc can no longer ignite across the longer distance.[1] Ignition then takes place between the electrode tips again and the process is repeated. This also causes a variation in the frequency as the conditions for successive arcs are not constant (arc resistance, charging period, etc.). Since the proper action of the field consists in blowing out the arc and allowing a new arc to form at the beginning of the next period, it is evident that its intensity will depend on the frequency to be generated. Pedersen has found the proper field intensity to be approximately proportional to the frequency. Thus with an arc drawing about 20 amperes from the supply line, and an oscillatory circuit with a ratio of $\sqrt{L/C}$ about 300, Pedersen found the most suitable field strength to be given by the relation $(H+400)\lambda=5000$, in which H is in gausses and λ in kilometers. With a hydrogen atmosphere it seemed that a field about one-fifth as large as this was proper.

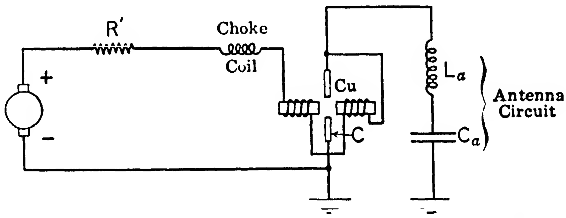


FIG. 8.—Connection scheme for producing alternating current in an antenna by an arc.

Action of the Gaseous Atmosphere.—The hydrogen or coal gas in which the arc usually operates assists in cooling the electrodes, and thus when the arc current falls to zero, the cooling action of the gas promotes

[1] The reader is referred to the photographs in Pedersen's paper for evidence of the correctness of these statements.

a rapid increase in the arc resistance (de-ionization). It also affects the static characteristic, making it steeper than in air, as shown in Fig. 43, p. 206.

The reason for the hydrogen atmosphere thus steepening the curve is not known, but the effect of this increase in slope upon the arc operation is evidently to cause the arc voltage variation (which in turn acts to charge the condenser) to be more sensitive and of greater amplitude for a given arc current variation. The radio-frequency energy input is thus increased.

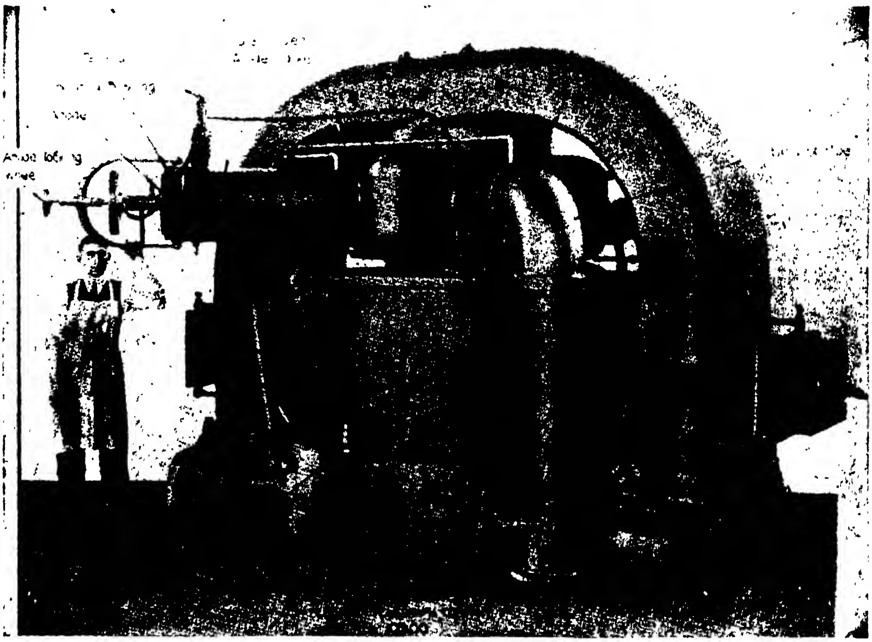


FIG. 9.—A 500-kw. Poulsen arc converter; over the operator's head is the anode terminal and on the right is shown the pipe for carrying off the exhaust gases from the arc chamber. (Proc. I.R.E., Vol. 7, No. 5.)

The foregoing features of construction are embodied in all modern arc generators. Fig. 9 illustrates a 500-kw. arc (input rating), which is much less than the maximum capacity to which generators of this type have been built up to the present time.

Generators of 1000-kw. capacity are of the same general construction, but somewhat larger in size. The arc chamber is equipped with a water-cooled jacket to assist in cooling the chamber, while the copper anode has circulating water supplied to it by means of flexible pipe connections. The negative electrode is usually of carbon, although graphite is being largely used for the higher capacity arcs. The anode, as shown in the figure, is equipped with handwheels to permit the accurate adjustment of the gap length. The smaller wheels shown are used for clamping the electrode into its proper position. The enormous size of magnetic circuit apparently required for these large capacity generators is indicated in Fig. 9 as well as in Fig. 10, which shows the generator with the electrodes and arc chamber removed. The circuit shown in the latter figure is for a 500-kw. arc; the upper pole piece has been removed.

Arc generators are most efficiently used at the longer wave lengths and are therefore usually operated at 3000 meters or above, 6000 meters being the wave length generally used. In some cases the wave length is as high as 18,000 meters. The capacities range from 100 kw. or less up to 1000 kw.; 350-kw. arcs are generally used for high-power land stations.

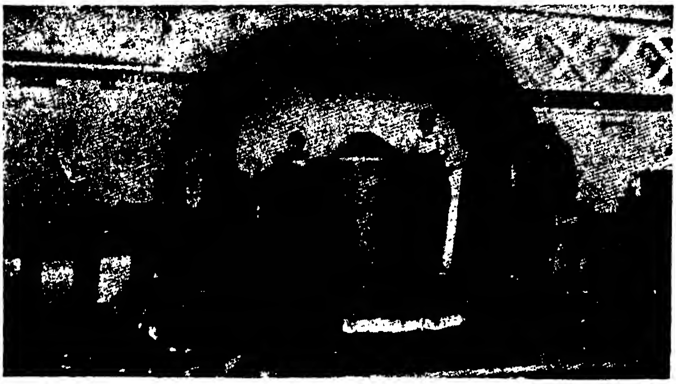


FIG. 10.—Magnetic field structure for a 500-kw. converter; the proper form of pole piece has been the subject of considerable study. (Proc. I.R.E., Vol. 7, No. 5.)

Small-capacity arcs, with a capacity of 20–30 kw., have been in successful use on board ship. The usual wave length is 4000 meters, and the sets have a transmission range of perhaps 2000 miles in the daytime. There have been a few small arcs (2-kw. input) constructed, which when fed from a 600-volt line seem to operate reasonably well for wave lengths as short as 800 meters.

Typical Installations.—The Annapolis station of the U. S. Navy Department is a typical example of an arc converter installation. Two 350-kw. converters, each capable of supplying 400 amperes of radio-frequency current to the antenna circuit, are installed; the maximum voltage impressed on these arcs is 1500 volts. A unique feature of this station is a provision for melting the ice from the antenna wires by current supplied from the two 400-kw. d.c. generators (connected in multiple) which are normally used to supply the converters.

Another typical plant is the Lafayette station located in France. Two arc converters, each capable of supplying 500 kw. of high-frequency power to the antenna, are installed, each supplied by a d.c. generator rated at 1000 kw. The normal wave length is 20,000 to 25,000 meters (15 to 12 kc.), a signal speed of 50 words per minute being attained.

Arc converter stations are also located at Lyons, San Diego, Pearl Harbor, and Cavite; these representing the largest and more recent installations. The general features are similar to those of stations already described.

For detailed installation and construction details of a large arc station the reader is referred to an article by DeGroot in Proc. I.R.E., Vol. 12, No. 6, Dec., 1924, in which the high-power Dutch station at Malabar, Java, is described.

High-frequency Alternator.—The generation of high-frequency currents by means of machines similar in their principles of construction to the huge alternators which supply the modern central-station loads, has doubtless occurred to the student. The extremely high frequencies required, however, necessitate machines of special design which require an exceptionally high grade of engineering skill in their construction. Alternators for supplying loads of commercial frequency may be any one of the three following types:

I. The armature is the rotating element, the d.c. field being stationary. This arrangement is similar to that employed on all d.c. generators but is rarely used on alternators, particularly the large sizes.

II. The field rotates with respect to the armature, which is fixed in position. This construction possesses several advantages over type I, particularly due to the lesser insulation requirements of the field winding and its greater simplicity as compared to the armature winding. This construction is universal on all modern alternators.

III. Both the field and armature windings are stationary in space, the flux linking the armature winding being periodically varied by means of an inductor, revolving in the air gap. This inductor is essentially a disc whose periphery has been divided into sections, alternate sections possessing a high magnetic reluctance, while the intermediate sections, which are made of steel, possess a relatively low reluctance. This type is practically unused in the low-frequency machines of commercial engineering, but possesses several inherent advantages which make it the most satisfactory of the alternators designed for high-frequency generation. Since both windings are fixed in position, their proper insulation is much simplified. Very serious difficulty is encountered when it is attempted to place an insulated winding on the revolving member (rotor), due to the high peripheral speeds and consequently high centrifugal stresses involved.

Design of the High-frequency Alternator.—That a special construction and design is required may be seen from the following: if we consider a machine of the inductor type having a maximum permissible speed of 20,000 r.p.m. and a required frequency of 100,000 cycles per second, the rotor diameter being assumed 30 centimeters, the distance through which a point on the rotor moves in generating one cycle is

$$\frac{\pi \times 30 \times \frac{20,000}{60}}{100,000} = 0.31 \text{ cm.}$$

Therefore, in this small space we must have a section of high reluctance (for instance, bronze) and a section of low reluctance (steel) so that a complete cycle of flux variation from minimum to maximum and back to minimum occurs while the inductor moves through this space. Special precautions in design and materials used must be observed if the hysteresis and eddy-current losses are to be minimized, as these become very large at the higher frequencies.

Construction.—The construction of the Alexanderson high-frequency alternators (first suggested, and first ones built, by R. A. Fessenden), is indicated in Fig. 11.

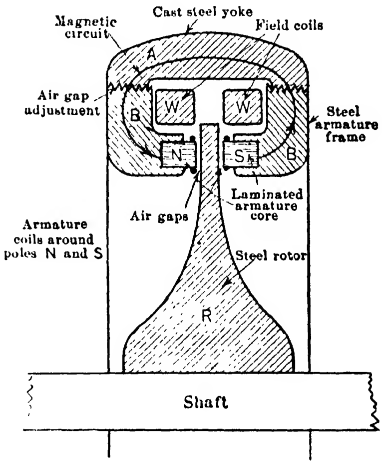


FIG. 11.—Simplified cross-section of an Alexanderson alternator.

The pole pieces BB are threaded into the yoke A as indicated, the air gaps being thus accurately adjustable. The pole tips NS are of finely laminated high resistance steel to reduce eddy-current losses and are slotted for the armature winding. The field windings WW are installed as shown, and with a steady direct current flowing through them, set up a flux as indicated in the diagram. It is evident that the reluctance of this circuit will vary as the inductor R rotates between the two faces of the air gap. This inductor is properly designed for the high stresses which exist when it is operated at its rated speed of 20,000 r.p.m. The rim velocity under this condition is about 300 meters per second and the centrifugal force at the periphery is 68,000 times the weight of metal there.

A developed view of the winding and rotor is shown in Fig. 12. If we consider the loop formed by conductors 6–7, the flux linking it will be a maximum with the inductor in the position shown. As the inductor moves to the right, the tooth *e* is replaced with a non-magnetic insert and the flux decreases to a minimum value. Then as the inductor continues to move to the right the tooth *d* enters the loop 6–7 and the flux increases again to a maximum. The variation of flux and corresponding induced voltage ($E=d\phi/dt$) is indicated on the figure. It will be noted that a complete cycle is generated while the rotor travels a distance represented by the tooth pitch (distance of a point on a tooth to similar point on adjacent tooth). The frequency is thus equal to the number of inductor teeth multiplied by the revolutions per second.

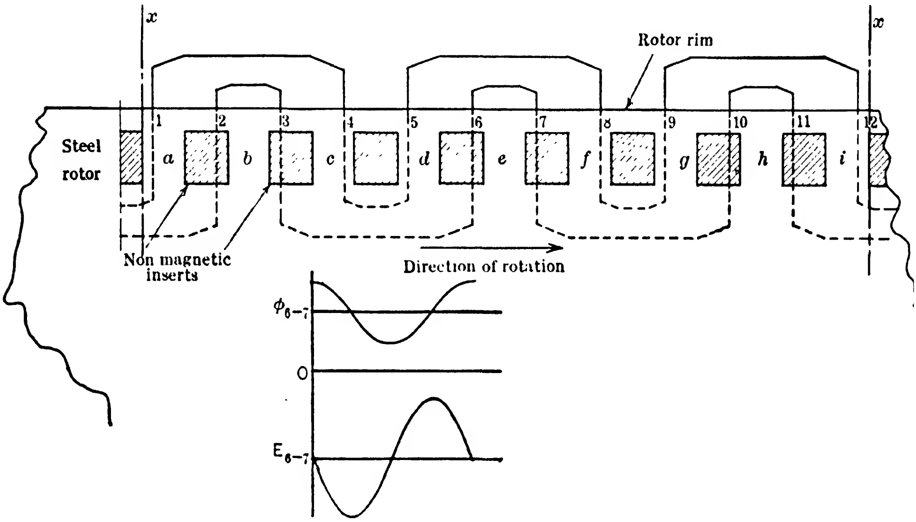


FIG. 12.—Developed view of the winding and rotor of the machine shown in Fig. 11.

The effective number of poles for this type of machine is evidently twice the number of inductor teeth or spokes. Thus to generate 100,000 cycles we would require

$$N=\frac{f}{rps}=\frac{100{,}000}{\frac{20{,}000}{60}}=300 \text{ spokes on the inductor.}$$

An examination of the winding indicates that loops 2–3, 10–11, etc., pass through the same flux variations as loop 6–7, and as these loops are all connected in series the voltage will add up around the periphery. The same analysis holds for loops 4–5, 8–9, 12–1, etc. The windings are brought out to separate terminals and may thus be connected in series or parallel, whichever may be most suitable for the conditions involved.

It should be remembered that two similar windings are placed in the pole tip on the other side of the air gap, which are not shown in the figure. Thus the operator has four or more separate windings which he may connect in any arrangement most desirable for his conditions. On the typical alternator, each coil has its terminals brought out, there being 64 such coils.

On the normal inductor generator, the number of armature slots is always equal to the effective number of poles. In the Alexanderson machine the number of armature slots may be two-thirds the number of poles. This is a distinct advantage, as more space and more thorough insulation is thus permitted for the winding. Thus in the figure, we have between the lines xx, twelve armature slots and nine inductor spokes, which represent eighteen effective poles, or the armature slots are two-thirds the effective poles.

The greater the flux variation ($d\phi/dt$) the greater will be the generated voltage. By decreasing the air gap the effect of the inductor on the flux becomes more pronounced and thus the generated voltages increase. On a certain machine, with a minimum permissible air gap of 0.004 inch (for each of the two gaps), the voltage generated was nearly 300 volts. With the air gaps increased to 0.015 inch each, the voltage decreased to 150. (Eccles.) Similarly, the output capacity increases as the air gap is decreased and vice versa.

The highest frequency for which these machines have been constructed at the present time (1927) is 200,000 cycles per second with a capacity of about 1 kw. Machines of 50 kw. and greater have been constructed, the frequency, however, being lower for these higher capacity machines, from 50 kilocycles to 20 kilocycles. A 2-kw., 100,000-cycle generator is indicated in Fig. 13, showing the driving motor, normally operating at 2000 r.p.m., connected through special 1 : 10 gears to drive the alternator shaft at 20,000 r.p.m. This general arrangement is followed on all alternators of this type, the gear ratio decreasing as the capacity increases. On some machines the driving motor is connected to the low-speed gear shaft by means of a chain connection. A view of a piece of one armature of a 2-kw., 100,000-cycle machine is shown in Fig. 14.

As might well be supposed, the high-speed machines are not as reliable in operation or as easy to maintain as a low-frequency machine of the same power. The bearings of the machine shown in Fig. 13 are flexibly fastened to the bed plate of the machine so that as the armature shaft expands each bearing will move away from the rotor disc equally, thus maintaining the two air gaps equal. Forced oil feed must be used for the bearings and for the larger machines, pipes carrying cooling water are liberally distributed throughout the structure of the machine.

The high peripheral speed of the disc results in a very rapid circulation of air through the two air gaps, causing considerable noise and power consumption. The small machine shown in Fig. 13 requires about 7 h.p. to turn the disc at rated speed, the machine not being loaded.

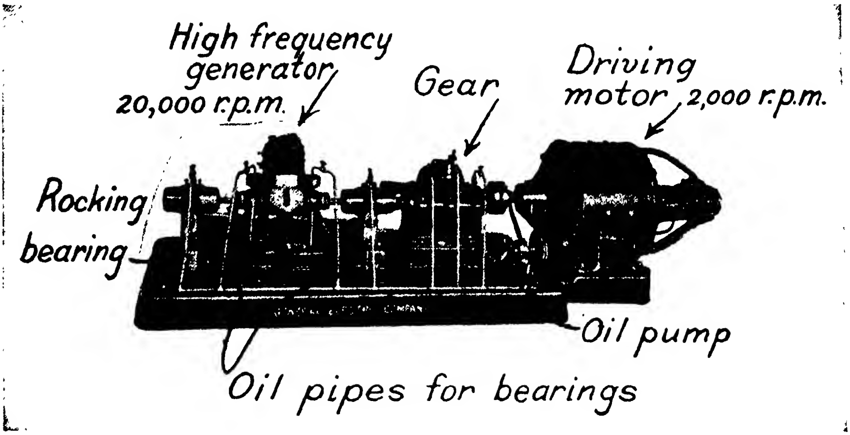


FIG. 13.—View of a small 100,000-cycle Alexanderson alternator.

These high-frequency inductor alternators require suitable tuning condensers to neutralize their internal reactance before they can deliver

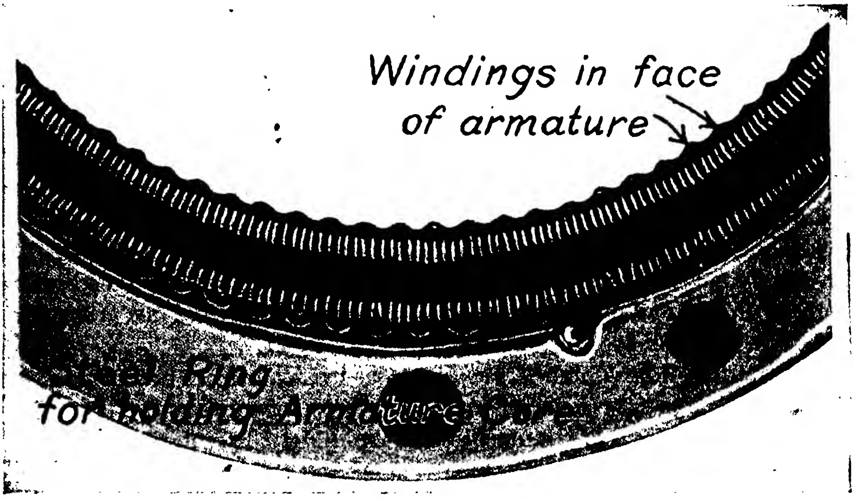


FIG. 14.—A view of a section of one armature of the machine in Fig. 13.

appreciable power; a small 200,000-cycle machine will scarcely deflect a voltmeter across its terminals unless a proper condenser is connected across the armature terminals to neutralize the effect of its armature inductance.

Connection to the Antenna.—The armature winding may be directly connected to the antenna as shown in Fig. 15(*a*), or it may be inductively coupled as shown in Fig. 15(*b*), by using an oscillation transformer. In either case the antenna circuit must be tuned to the frequency of the alternator if maximum output is to be obtained. If the 2-kw. 100,000-cycle machine is considered, its reactance at this frequency with normal air-gap, may be taken as 5.4 ohms or

$$L=\frac{5.4}{2\pi\times100{,}000}=8.58\ \mu h.$$

Thus the antenna capacity (Fig. 15(*a*)) must be such that

$$3000=1885\sqrt{LC}=1885\sqrt{8.58C}$$

or $C=0.3$ microfarad. It is evident that the direct connected scheme (Fig. 15(*a*)) could not be used unless a suitable loading inductance (L') were added, as antennas are not built with such a high capacity. The arrangement utilizing an oscillation transformer (Fig. 15(*b*)) would most probably be used in any case. The maximum continuous load of this machine is 30 amperes at 70 volts, or the equivalent antenna circuit resistance as measured at the terminals of the generator must thus be 70/30=2.3 ohms.

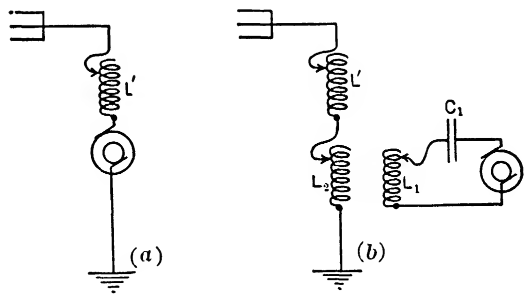


FIG. 15.—Two schemes for connecting a high-frequency generator to an antenna; that shown in (*b*) is generally used.

Application.—A few years ago the Alexanderson alternator had attained a position of great commercial importance, particularly the lower-frequency machines (20,000–50,000 cycles) of large capacity. The inherent disadvantages of this type of generator due to the high speeds and complications attendant thereto, such as lubrication, etc., apparently prevent it from representing a practical means of generation on board ship, or for small-power installations or portable field service. For high-power land stations, however, engaged in transoceanic and transcontinental service, it has proven very successful in application and performance. The earliest successful station utilizing this type of generator was located at New Brunswick, N. J., where a 200-kw. set was installed. This equipment is shown in Fig. 16. For a further description of this machine and station the reader is referred to a paper by E. F. Alexanderson.[1]

One of the chief difficulties in the operation of a high-frequency alternator is the accurate control of its speed. An almost imperceptible change in alternator speed will result in the pitch of the signal note at the receiving station changing several octaves. That the Alexanderson generators are controlled in speed to better than 0.1 per cent can be told at once by noting the constancy of pitch of the signal received from one of these machines. An ingenious arrangement of relays operate on the driving motor so as to make its speed essentially constant.

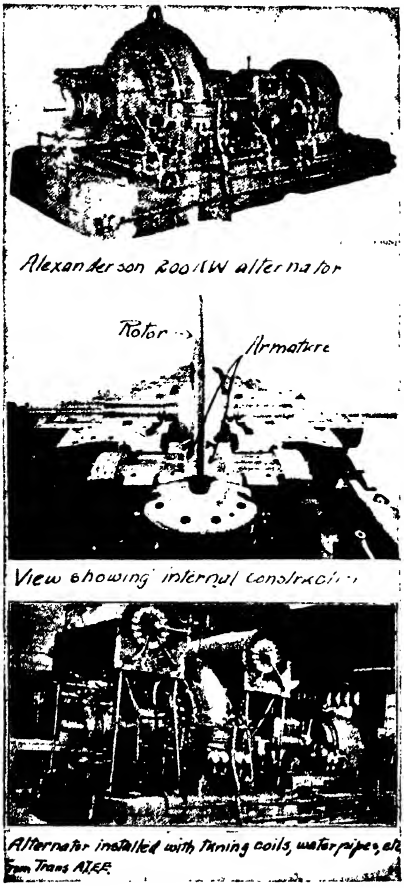


FIG. 16.—Views of a high-powered Alexanderson generator; over the alternator (in the lower view) may be seen the oscillation transformers which connect the generator to the antenna.

The speed must be held constant regardless of voltage or frequency fluctuations of the supply to the driving motor, as well as load fluctuations caused by operation of the sending key. Fluctuations due to the latter are minimized by a system of relays operated synchronously with the telegraph key, these relays in turn varying the voltage applied to the driving motor and motor rotor resistance so that the motor torque is made equal at all times to the load. This assumes a polyphase induction motor as the prime mover. With a d.c. driving motor the scheme indicated in Fig. 39, p. 754 is successfully used.

Speed variations due to fluctuations in supply are minimized as follows: the alternator supplies a small tuned circuit (in addition to the antenna circuit) which is tuned to a frequency slightly different from that at which the set normally operates. With normal frequency a certain current will flow in this circuit; if the frequency increases or decreases,

[1] "Transatlantic Radio Communication," Proc. A.I.E.E., Oct., 1919.

however, a correspondingly large change in this current will occur. Part of this local circuit current is rectified and, through relays, is made to control the supply voltage. The response is practically instantaneous to changes in frequency and in practice a variation of about 3 meters, when sending at 15,000 meters, is normal.

Typical Installations.—At the Nauen (Germany) station a 400-kw. high-frequency alternator, generating 1200 amperes, 450 volts, 6000 cycles at 1500 r.p.m., supplies the antenna system through a step-up transformer and two frequency-doubling transformers. The normal frequency supplied to antenna is thus 24,000 cycles per second and radiated wave length is 12,500 meters. The overall efficiency (antenna power/driving motor input) is reported to be 66 per cent for $\lambda = 12.6$ km., 56 per cent for $\lambda = 8.4$ km., and 53 per cent for $\lambda = 6.3$ km. Signaling speed is 50 words per minute. A speed change of $\frac{1}{2}$ per cent (7.5 r.p.m.) will reduce the radiated power to 64 per cent of its maximum value; the Telefunken governor, with which the set is equipped, responds to a speed change of 0.01 per cent.

At Port Jefferson, L. I., two 200-kw. units of the Alexanderson type are installed. Wave lengths of 15,800 to 20,000 meters are used, a signal of 100 words per minute being attained. At the Paris Radio Central, three 500-kw. units of French design are installed for intercontinental transmission. Two different transmissions may be carried on simultaneously; at a signaling speed of 100 words per minute the station output is 12,000 words per hour. The total capacity, transmittible and receivable, is about 2,000,000 words per 24-hour day. Photographic recording apparatus is used for receiving signals at this station.

Comparison of the Arc Converter and High-frequency Alternator.—Although the oscillating tube is rapidly increasing in capacity, for high-power applications, the arc and high-frequency alternator form the two principal means for providing large amounts of high-frequency power at the present time. The following comparison of these is, therefore, of interest:

The disadvantages of the arc may be stated as follows:

(*a*) Efficiency is lower than for the alternator if compensating method of sending is used, viz., less than 40 per cent versus 50–60 per cent. Space signal radiation is wasted and interference is increased.

(*b*) Oscillations are more or less irregular, producing upper harmonics and increasing interference. These upper harmonics are particularly troublesome when they correspond to natural frequencies of the antenna system being supplied. Thus the resonance curve of an antenna showed a pronounced peak at the 7th harmonic frequency. The arc converter supplying the antenna contained this harmonic in its voltage wave and marked interference effects were experienced by short-wave stations operating in the vicinity, due to the large value of 7th harmonic current flowing in the transmitting antenna. In this case the interference may be reduced by coupling to the antenna circuit an absorbing circuit, tuned to the 7th harmonic, thereby inserting a resistance in the antenna, effective for the 7th harmonic current only.

(*c*) The arc requires cleaning more frequently and more extensively than the alternator. Duplicate equipment must be provided.

(*d*) Signaling speed is only about 60 words per minute, whereas the alternator permits a speed of 100 or more words per minute. The primary advantage of the arc lies in greater flexibility as to wave-length adjustment. However, it operates best at a certain value of λ, and efficiency must be sacrificed if the wave length differs from this best value. This wave-length flexibility would be of importance in military applications but is of relatively little value in commercial work.

The foregoing would appear to indicate that the alternator is superior to the arc. One must not overlook the fact, however, that the high-frequency alternator is a precise and delicate mechanism, requiring a high grade of supervision and care. Many other considerations, including financial and economic aspects, would have to be taken into account if a just and fair comparison is to be made.

The Goldschmidt, or reflection type, Alternator—Theory.—This generator, first brought out in commercial form by Dr. R. Goldschmidt, is based on principles which are radically different from those involved in the Alexanderson machine. These are:

1. The magnetic field produced by an alternating current of frequency n, may be considered as consisting of two component fields, the magnitude of each of these fields being one-half the magnitude of the resultant total field and considered as rotating in opposite directions at frequency n.

This is simply the theory of conjugate vectors and is illustrated in Fig. 17. Fig. 17*A* represents the normal derivation of a sine curve from a rotating vector while Fig. 17*B* utilizes the principles of conjugate vectors; the horizontal components of ϕ' and ϕ'' neutralize one another, while the resultant vertical component is at all times as indicated by the sine curve to the right. Similarly, the current I could be represented in the same manner. The construction illustrates graphically the principle stated above.

2. If a coil (rotor) is revolved in this alternating magnetic field at synchronous speed, it is apparent from the foregoing that the component

fields will induce e.m.f.s of different frequencies, since they are rotating in opposite directions. If we assume the coil to revolve in a counter-clockwise direction, flux ϕ' will rotate with it, thus cutting no conductors and generating no e.m.f. in the coil. The other component ϕ'' is moving against the rotation of the coil. Thus the frequency will be twice what it would be if the coil were standing still. Reviewing the above, two frequencies may be considered as being generated in the coil, viz.,

$$f_1 = N - N_r \text{ (produced by } \phi')$$

$$f_1 = N + N_r \text{ (produced by } \phi'')$$

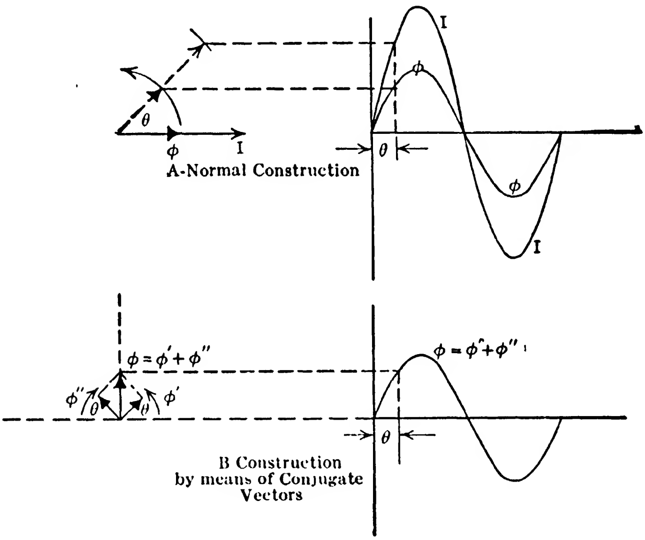


FIG. 17.—A stationary, pulsating, magnetic field may be represented by two rotating fields each constant in strength, rotating in opposite directions. Each rotating field has one-half the strength of the stationary pulsating field.

since

$$N_r = N,$$

this becomes

$$f_1 = 0$$
$$f_2 = 2N.$$

In these expressions N is the frequency of the alternating field, while N_r is the rotational speed of the coil expressed in cycles per second. Therefore, if the terminals of the rotating coil are connected together, a current will flow in the circuit whose frequency is $2N$, and a doubling of the initial frequency N has been obtained. It is evident that this double-frequency current could be led to a second fixed coil, with a revolving coil placed in the influence of its magnetic field and a further doubling of frequency would result, if the coil is rotated at a frequency $2N$. If the speed of rotation is N_r as for the first case, the frequencies would be

$$f_1 = N - N_r = 2N - N = N$$

$$f_2 = N + N_r = 2N + N = 3N$$

A practical example of these effects is found in the double-frequency component which appears in the d.c. field circuit of a single-phase alternator when the machine is carrying load. This is illustrated by the ondograph record shown in Fig. 18, in which the 60-cycle armature e.m.f. and the 120-cycle ($2N$) e.m.f. induced in a search coil on the pole are both shown. The field winding revolves at synchronous speed in an alternating

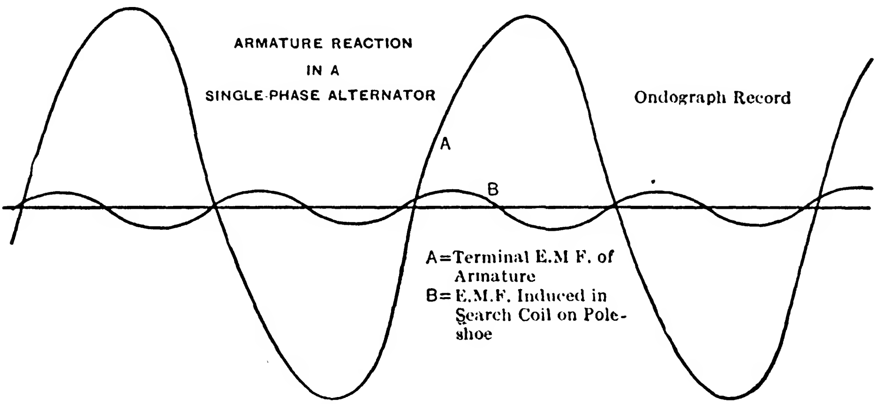


FIG. 18.—Ondograph record from a single-phase alternator; curve B, obtained from a search coil on the pole face, shows a frequency twice as great as that generated in the armature.

field produced by the current in the armature winding. We thus have, as for the first case cited above

$$f_1 = N - N_r = 0$$

$$f_2 = N + N_r = 2N.$$

It has already been indicated how additional increases in frequency might be obtained by using a number of machines (consisting of a fixed and movable coil) connected in cascade. However, instead of having the rotor currents excite a distant stator, it is more practical and economical to have it react back on its own stator. The fundamental connections would then be as shown in Fig. 19.

Connections for Increasing Frequencies in the One Machine.—The source of power, A, supplies current of frequency N to the stator winding S and the rotor winding R rotates in this field at synchronous speed (in r.p.m. $= N \times 120$/no. of poles). There will thus be induced in R an e.m.f. of frequency $N - N = 0$ and an e.m.f. of frequency $N + N = 2N$. If the terminals are joined by a low-impedance path, a current of this frequency will flow in the rotor circuit. Associated with this current is a magnetic field whose frequency is $2N$. Recalling that we may represent this field by two oppositely rotating fields, whose frequency is $2N$, it is evident that one component (the one revolving in the same direction as the rotor) will cut the stator at a frequency $2N + N = 3N$, while the other component will cut it at a frequency $2N - N = N$, corresponding e.m.f.s being induced in the stator circuit. If the terminals of the stator coil be joined by a suitable circuit, currents of these frequencies will flow. Furthermore, the field set up by the triple-frequency current will induce in the rotor e.m.f.s of frequency $3N - N = 2N$ and $3N + N = 4N$, and currents of these frequencies will flow in the rotor circuit. In turn, the current of frequency $= 4N$, will induce in the stator e.m.f.s of frequency $4N - N = 3N$ and $4N + N = 5N$. Thus, if we assume an initial supply frequency to the stator of 10,000, we have transformed it by means of so-called "electrical reflections" outlined above, into a frequency of $5N = 50{,}000$, which may be employed for radio-telegraph and radio-phone transmission.

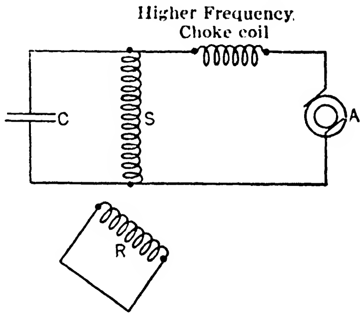


FIG. 19.—Conventional diagram of rotor coil R, and tuned stator circuit S-C, supplied with magnetizing current through choke coil.

These several e.m.f.s may be tabulated as follows:

Stator	Rotor
N	$2N$ and 0
$3N$ and N	$4N$ and $2N$
$5N$ and $3N$	

If, instead of supplying alternating current to the stator winding, we employ direct current, the results are, similarly:

Stator	Rotor
0	N
$2N$ and 0	$3N$ and N
$4N$ and $2N$	

This is the usual arrangement for commercial machines, the source of supply being a storage battery or d.c. generator.

Application of Tuned Circuits.—In discussing the reflection of frequencies we have indicated the coil circuits as being completed in the rotor by a short-circuiting resistance while the stator circuit is completed by a condenser. Since the coils themselves possess considerable impedance at the high frequencies involved, this must be compensated for by suitable capacities, so that the circuit may be in resonance for the frequency of the induced e.m.f., i.e.,

$$2\pi fL = \frac{1}{2\pi fC}.$$

By thus employing tuned circuits, the magnitude of the current flow produced will be a maximum and is limited only by the effective resistance of the circuit. This effective resistance includes the losses due to hysteresis and eddy currents as well as dielectric losses.

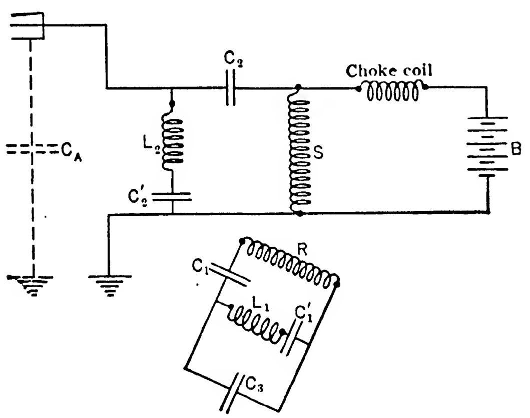


FIG. 20.—In order to get currents of appreciable amplitudes of the various frequencies generated in a "reflection" type machine the rotor and stator must be supplied with suitably tuned circuits, one for each frequency generated.

Since e.m.f.s of several frequencies are concerned, circuits must be available which are tuned to each of these frequencies. Fig. 20 indicates the arrangement employed by Goldschmidt, for a quadrupling of the generated frequency, direct current being supplied to the stator.

The rotor R revolves at the required speed in the d.c. field produced by the stator winding S, supplied by means of the storage battery B. There is thus induced in R an e.m.f. of frequency N, the value of which is given by

$$N = \frac{N_p \times RPM}{120}, \text{ where } N_p = \text{the number of poles.}$$

This e.m.f. will cause a current of corresponding frequency to flow in the circuit R, C_1, L_1, C'_1, the values of the capacities and L_1 being adjusted so that the circuit is tuned to this frequency. C_1 compensates for the inductance of the rotor, while L_1 and C'_1 compensate each other, and the drop across them is thus very small. This current induces an e.m.f. of frequencies $2N$ and 0 in the stator circuit S, C_2, L_2, C'_2, in which the values of C_2, L_2 and C'_2 are adjusted to resonance for the frequency $2N$. C_2 compensates for the inductance of the stator winding, while L_2 and C'_2 compensate each other, and therefore practically no drop exists across this portion of the circuit. The double-frequency current induces two e.m.f.s of frequencies $3N$ and N in the rotor, and triple-frequency current flows in the circuit R, C_1, C_3, which is tuned to resonance for this frequency. Practically no current of frequency N will pass through C_3, since L_1, C'_1, represents almost a short-circuit path across C_3 for this frequency.

The triple-frequency current flowing in the rotor circuit R, C_1, C_3, induces in the stator e.m.f.s of frequencies $4N$ and $2N$, and currents of corresponding frequencies flow in the circuits S, C_2, C_A and S, C_2, L_2, C'_2, respectively, each of which is tuned to resonance. The condenser C_A represents the antenna through which we thus have a current flowing whose frequency is four times the frequency (N) of the current initially generated. If it were desired to utilize the triple-frequency current, the antenna would be connected to the rotor in place of C_3, while L_2 and C'_2 could be omitted from the stator circuit. By suitably arranging other circuits higher frequencies could be obtained but such an arrangement is not employed to any extent commercially, as the quadruple-frequency machine is more efficient and fulfills all requirements.

For the complete Goldschmidt machine as described in the preceding discussion, we may tabulate the generated frequencies, as before:

Stator	Rotor
0	N
$2N$ and 0	$3N$ and N
$4N$ and $2N$	

An exact analysis of all the actions in this machine is complicated and would be out of place here. The amplitudes of the various frequencies, it must be pointed out, however, are not the same; for all e.m.f.s of zero frequency the amplitude is zero, while the other amplitudes depend for their values upon the tightness of the magnetic coupling between the rotor and stator circuits.

A fairly good idea of this reflecting action may be obtained by examination of the cut in Fig. 21 which shows the stator and rotor currents of a single-phase induction motor excited by a 60-cycle supply and running near synchronous speed. The rotor current evidently shows the two frequencies $(f+f')$ and $(f-f')$, f being the impressed frequency and f' the frequency of rotation. The rotor circuit and stator circuit were not tuned, otherwise more frequencies might have been accentuated, and the stator current would, for example, show $f+2f'$ and $f-2f'$ frequencies.

In Fig. 22 is shown the effect of running of the rotor at practically

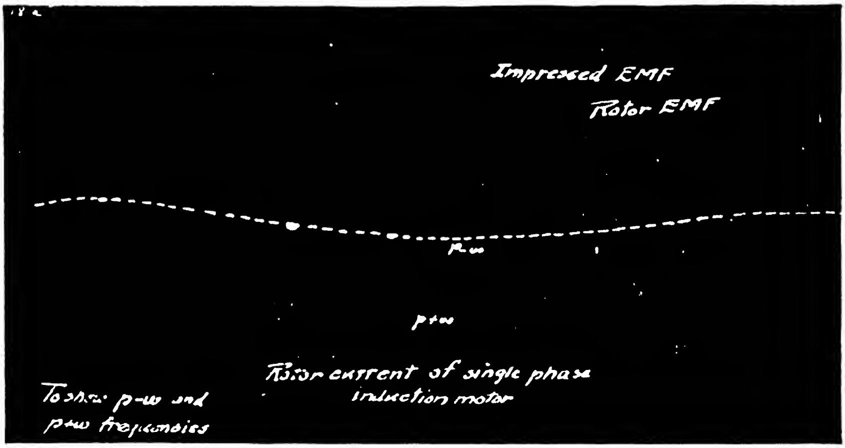


FIG. 21.—Rotor and stator e.m.f s of a single-phase induction motor. The rotor e.m.f. may be separated into its two components as shown by the dashed line. One frequency is equal to that impressed on the stator plus the frequency of rotation and the other frequency is the difference of the two.

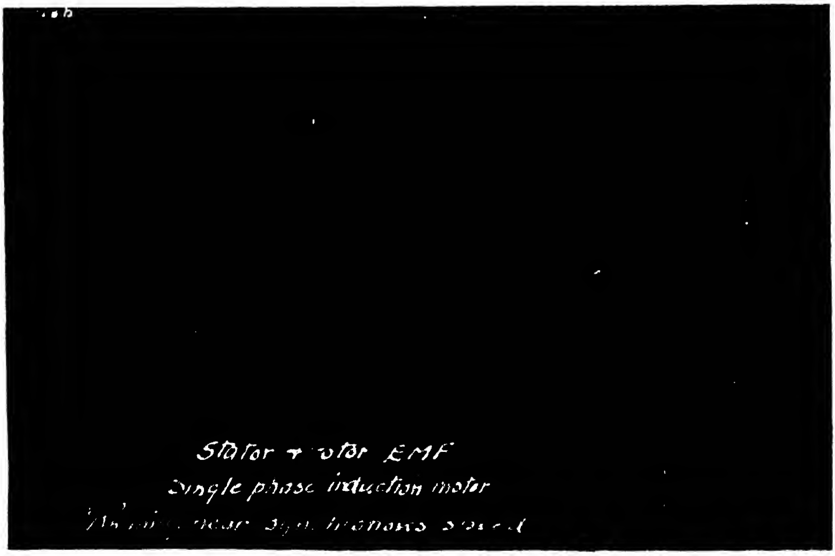


FIG. 22.—When the rotor was run at practically synchronous speed the amplitude of the differential frequency was practically zero, leaving in the rotor only the additive frequency.

synchronous speed. Here the amplitude of the differential frequency $(f-f')$ is so small that the film does not show it, although it can be seen from the film that the rotor frequency is slightly more than twice the stator frequency.

Construction.—If we assume a required frequency of 40,000 cycles per second, and a frequency transformation ratio of 4, the initial frequency generated is 10,000. From the general equation for frequency

$$f=\frac{\text{No. of poles}\times\text{r.p.m.}}{120}$$

it is readily seen that a large number of poles will be necessary. Assuming a maximum desirable speed of 3000 r.p.m., we have

$$\text{No. of poles}=\frac{120\times 10{,}000}{3000}=400.$$

Considering a maximum safe peripheral speed of 200 meters per second, we obtain 1.25 meters as the maximum diameter permissible for the rotor.

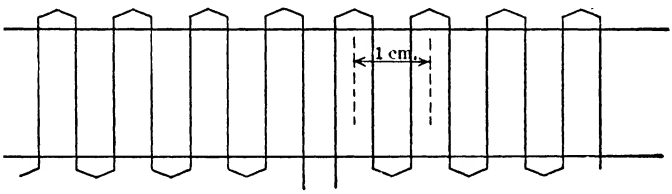


FIG. 23.—Developed winding of a Goldschmidt alternator.

This gives a circumference of 400 cm., and thus the space available per pair of poles is 1 cm. The windings are therefore laid out similar to the armature winding of the Alexanderson machine, and consist on both rotor and stator (the windings are identical) of the simple zig-zag winding shown in Fig. 23. The windings are split up into sections which can be connected in series or parallel arrangement to secure the most suitable voltage for the resistance of the antenna used. The conductor is made up of a number of very fine strands, about No. 40 A. W. G., each insulated individually to reduce skin effect.

To reduce the iron losses, the rotor and stator are constructed of very thin laminations of high-resistance iron. These laminations are 0.05 mm. in thickness and are separated by paper about 0.03 mm. thick. When the assembled material is compressed the volume of paper is more than one-third the total volume. To further decrease the iron losses, the iron is worked at low densities.

It is evident from the foregoing discussion on the action of this machine, that the air gap must be made as small as possible so that the magnetic leakage between the two windings shall be a minimum, since the induced voltages decrease for each succeeding reflection, this decrease being fixed by the amount of magnetic leakage and the losses. Excessive magnetic leakage also tends to prevent neutralization of the intermediate currents and thus additional losses are caused which may be minimized by reducing the gap.

On the largest machines we thus find extremely small air gaps, being about 0.08 cm. on a 100-kw. machine (normal rating). The rotor of this machine weighs about 5 tons and is 1.25 meters in diameter, which indicates the extreme precision and care required for the proper construction of this type of generator. Trouble may be experienced if the rotor expands under the effects of temperature rise produced by continuous operation. This will cause an increasing output (almost inversely proportional to the gap length) as the gap decreases until the rotor suddenly makes contact with the stator and jams tight, resulting in the destruction of the machine.

It is also important that the rotor and stator slots be strictly parallel to the shaft and to each other. That is, there should be no skewing, as otherwise the e.m.f.s induced throughout the length of one conductor of the winding will not be in the same phase, and a decreased voltage (and thus a decreased output) results. A divergence of 1 mm. in 1 meter length would cause a decrease of 20 per cent in the total output.

Typical Installation.—The largest alternators of this type have a maximum output of 200 kw. with a normal output of 100 kw., one of which is located at Hanover, Germany. This machine is of the four-reflection type, with d.c. supply to the stator and having 400 poles. For an output frequency of 50,000, which means an initial frequency of 12,500, the motor drives the generator at 3750 r.p.m. This motor is rated at 4000 r.p.m., 250 h.p., 220 volts, and is supplied from two d.c. generators in Ward Leonard connection, to secure the necessary flexibility of speed control and ease of starting. The generator is directly connected to this motor by means of a flexible coupling.

The antenna at the Hanover Station consists of a double-cone system, the aerial wires being supported by a single steel tower 250 meters high. The aerial system is made up of 36 bronze cables of 8-mm. diameter, the outer ends of these cables being attached through insulators to poles 12 meters high which are arranged in a circle around the tower, the radius of this circle being about 450 meters. The tower is insulated at the base and half way up by means of heavy glass insulators and is supported by steel guy wires, sectioned by insulators.

Frequency Transformation.—The design and construction of such alternators as described above, which provide at their terminals, e.m.f.s of frequency sufficiently high to be used directly for radio-transmission, require the highest type of enginéering skill, if the many complex problems involved in their construction are to be solved successfully. Alternators of somewhat lower frequency, however, say 10,000 cycles per second, can be built with comparative ease with consequent reduction in initial cost, and increased reliability of operation. Therefore instead of using the high-frequency alternator directly supplying the antenna circuit, we may substitute a lower-frequency machine, and step-up the frequency to the required value by means of a frequency changer or transformer. Efficient frequency transformation thus presents a means for supplying undamped radio-frequency current, and the action of some of the various frequency changers which have been proposed for this purpose is therefore of interest.

Types of Frequency Changers.—Frequency changers may be static, constructed similarly to the ordinary modern power transformer, or may contain a revolving element. In the latter type, utilizing one rotor winding and one stator winding, the frequency is raised by electrical reflections, four, five, or even higher transformations being accomplished. This type is illustrated by the Goldschmidt machine as described above, which may thus be considered as a generator and frequency changer in one, as it generates a current of frequency N, and this initial frequency is then transformed to some higher frequency at the output terminals.

The simplest type of frequency changer utilizing a rotating element is illustrated by the large-capacity frequency changers used for the interchange of power between systems of different frequencies as for instance, a 25 and a $62\frac{1}{2}$ cycles system. The machine consists of two rotors and two stators, the rotors being mounted on a common shaft, the one element operating as a synchronous motor and the other as an alternator, and vice versa, depending on the direction of energy flow. By means of apparatus of this type it is thus possible to transform a commercial frequency supply of 25 cycles to a frequency of 10,000, or less, which may be further transformed by one of the static transformers described below.

Many forms of the static type of frequency changer have been suggested, differing mainly in the manner of their connections, and the number of frequency transformations. Thus there are frequency doublers and triplers, which may be further connected in cascade, resulting in additional increase of frequency. Fundamentally, nearly all of them operate on the same phenomenon, namely, the asymmetrical variation of flux with magnetizing force in saturated iron cores. The explanations to be given below will consider only this feature of the circuit although a rigid analysis would undoubtedly require an investigation of the variation in resistance throughout the cycle, as well as these peculiar flux changes.

Frequency Doubler.—An arrangement for doubling the frequency first suggested by Epstein in 1902 and subsequently developed by Joly and

Valauri, is indicated in Fig. 24. Both transformers are identical, and each is equipped with a tertiary circuit, supplied from the storage battery B.

The steady current flowing through these windings is adjusted to bring the transformer fluxes just to the point where saturation occurs, i.e., at the knee of the curve as indicated in curve A, Fig. 24. If the two primaries, in series, are connected to an a.c. supply, it may readily be seen from the figures, that on a positive half cycle the flux in T_1 (wherein the m.m.f. of the primary winding assists the d.c. winding), will change very little while the flux in T_2 will decrease considerably, since the primary m.m.f. opposes the m.m.f. of the d.c. winding. On the negative half cycle, the reverse is true. These conditions are indicated in Fig. 25, A and B. It should be noted that an asymmetrical variation of flux thus occurs in each

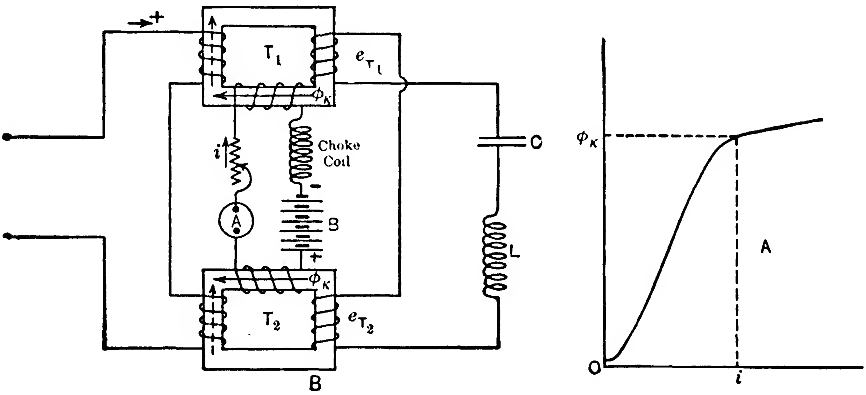


FIG. 24.—Use of two saturated cores for frequency doubling.

transformer, the flux of one transformer having a large variation during one alternation while the flux of the other transformer changes only slightly and on the next alternation, these conditions are reversed. These fluxes exist in separate cores and do not combine to form a double frequency flux. The e.m.f.s which they induce in their respective secondary windings are indicated in Fig. 25C, and since the secondaries have been connected reversed with respect to each other, the difference of the voltages must be taken to obtain their resultant. This is indicated in Fig. 25D.

This method for obtaining a doubling of the initial frequency has found some commercial application to high-frequency work, having been developed by Count von Arco for the Telefunken Company, and known as the Joly-Arc System. It was employed at the U. S. Radio Station at Sayville, Long Island, for doubling an initial frequency of 15,000.[1]

[1] Bucher, "Practical Wireless Telegraphy," p. 273.

It may be shown [1] that by proper choice of the direct and alternating primary current, the double frequency voltage may be made a pure sine wave. This occurs when the maximum flux decrease on the negative half cycle is 1.6 to 1.8 times the maximum flux increase on the positive half cycle. In practice, however, to secure a greater output, the direct current is adjusted to the knee of the magnetization curve, and the effective value of the alternating current is made equal to this direct current. A peaked wave is then produced, but the iron losses are not excessive as rather low frequencies are usually involved. At very high frequencies, the sine-wave system is preferable.

Fig. 26 illustrates circuit connections utilizing frequency doubling transformers (Alexanderson). A high-frequency alternator is shown at the left, which provides the fundamental frequency f_1. The iron cores are provided with d.c. excitation of proper value, double frequency thus being produced. By means of the trap circuits, installed and tuned to the frequencies indicated, each frequency is confined to a definite portion of the circuit. The traps are merely examples of parallel resonance, possessing a high joint impedance to currents of the frequency to which they are tuned.

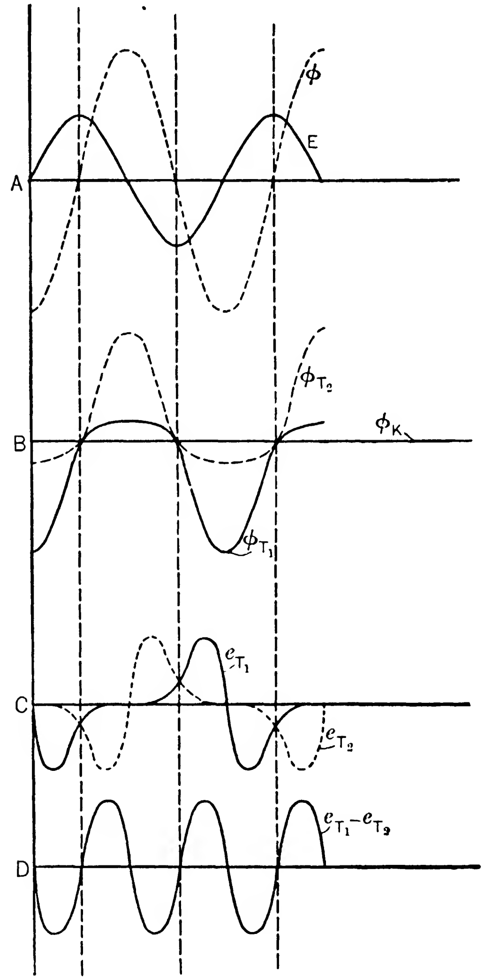


Fig. 25.—Analysis of the action of the arrangement of Fig. 24. The flux ϕ_{T2} in curves *B* is shown in reversed phase.

Frequency Tripler.—An arrangement for tripling the initial frequency of a three-phase supply, as developed by A. M. Taylor, is indicated in Fig. 27.

[1] "Some Characteristics of the Frequency Doubler as Applied to Radio Transmission," T. Minohara, I.R.E., Dec., 1920.

The three chokers R_1, R_2, and R_3 are saturated early in the cycle, at a relatively low value of current, while the core of the transformer T remains unsaturated at all times. Considering one of the elements, for instance that between a and b, and assuming a sine wave of voltage, the voltage and flux conditions which must exist are as indicated in Fig. 28A.

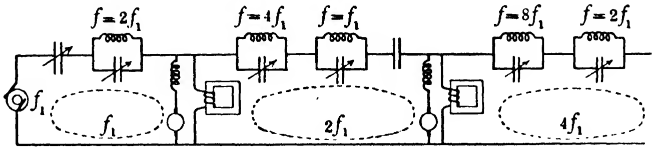


FIG. 26.—Suitably arranged tuned circuits are required for efficient frequency transformation.

(Fig. 27A indicates the circuit detail under analysis.) As the primary current increases, a point is reached where the choker becomes saturated. When this occurs, the impedance of the circuit decreases materially, and the primary current increases rapidly, as indicated by the current curve (Fig. 28B). This causes very little change in the choker flux, which has already reached saturation, but does cause a variation in the transformer

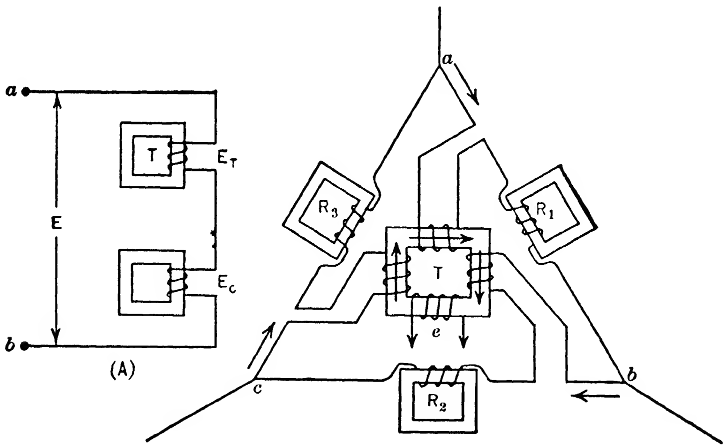


FIG. 27.—A frequency tripler, working from a three-phase supply.

flux. The transformer flux, although varying proportionately to the primary current I, does not reach large amplitudes, since most of the flux required to produce the proper sine wave of counter e.m.f. is already existent in the core of the choker. The choker and transformer flux must add to give a resultant equal to the sine wave of flux shown in Fig. 28A, based on the assumed sine voltage. The transformer flux will induce in the secondary a voltage the form of which is shown in Fig. 28*C*. A similar e.m.f. wave will be induced by each of the other two phases in exactly the same way as outlined above.

The voltages e_{bc} and e_{ca} will differ in phase from e_{ab} by 120° and 240° (electrical degrees), respectively, as the primary supply voltages, E_{ab}, E_{bc}, E_{ca} differ in phase from one another by this amount. These three induced voltages e_{ab}, e_{bc}, e_{ca}, exist simultaneously in the secondary winding of T and thus add up to give the resultant triple-frequency voltage indicated in Fig. 28*D*. It would be possible to employ nine chokers, and a nine-phase supply, to produce a nine-fold transformation, if this were desirable. In this case a sine-wave alternator could not be used, due to interference effects in the high-frequency circuits, and a machine of special design would be required.

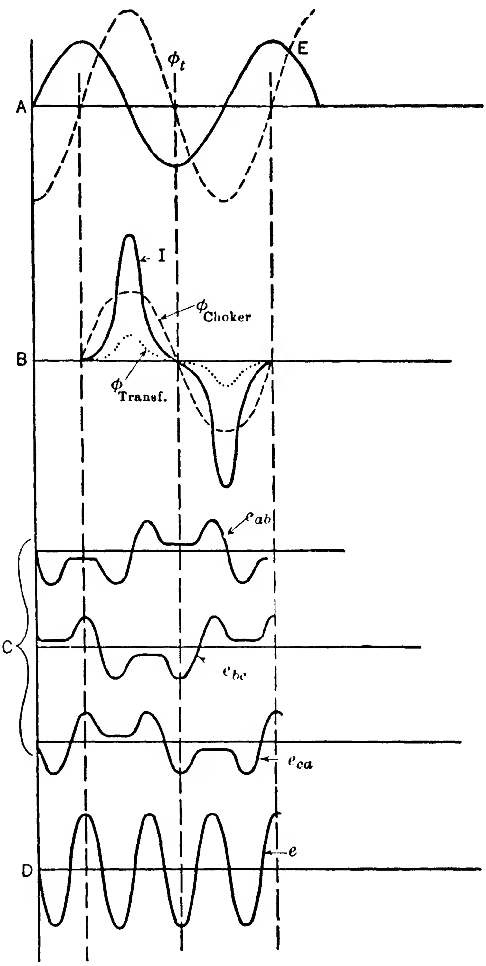


Fig. 28—Curves of flux and e.m.f. explaining the action of the apparatus shown in Fig. 27.

The "Wabbling Neutral" as a Means of Tripling the Alternator Frequency.—It is a well-known fact that the line currents of a 3-phase system are 120° out of phase and their algebraic sum is equal to zero. Their third harmonics differ therefore by $3 \times 120°$, or 360°, i.e., a complete period, and are in phase with each other. Since, in all cases, the instantaneous (algebraic) sum of the alternator currents must be zero,[1] it is evidently impossible for the line currents to contain

[1] Delta-connected load assumed, or if Y-connected, the neutral point of the load is supposed ungrounded.

third harmonics. If we impress a sine wave of voltage on three Y-connected transformers (their secondaries being open-circuited, and hence not shown in Fig. 29 as they can have no effect when open), the third harmonic component, which normally predominates in the exciting current of an iron-core transformer, is suppressed, and the magnetizing current is a sine wave.

The line voltages are sine waves, but the voltage to neutral must contain a strong third harmonic, due to the suppression of the third harmonic component in the exciting current, which must be present if the c.e.m.f. is to be a sine wave. Therefore, the wave of magnetization cannot be of sine form, but will be flat topped (somewhat as indicated in Fig. 30, curves 1, 2, and 3) due to the saturation of the iron. The induced e.m.f.s. will thus have the wave form indicated, and may be resolved into their fundamental and third harmonic components as shown in curves 4, 5, and 6.

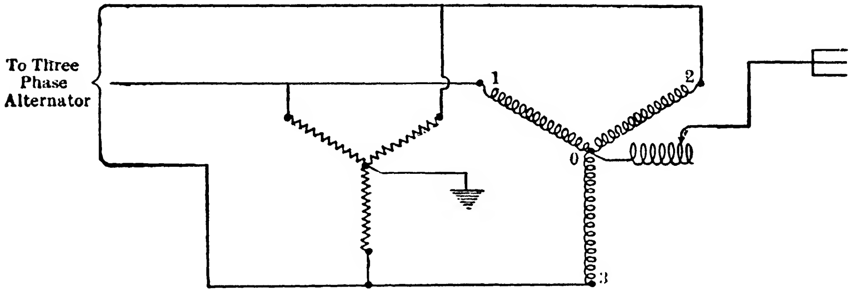


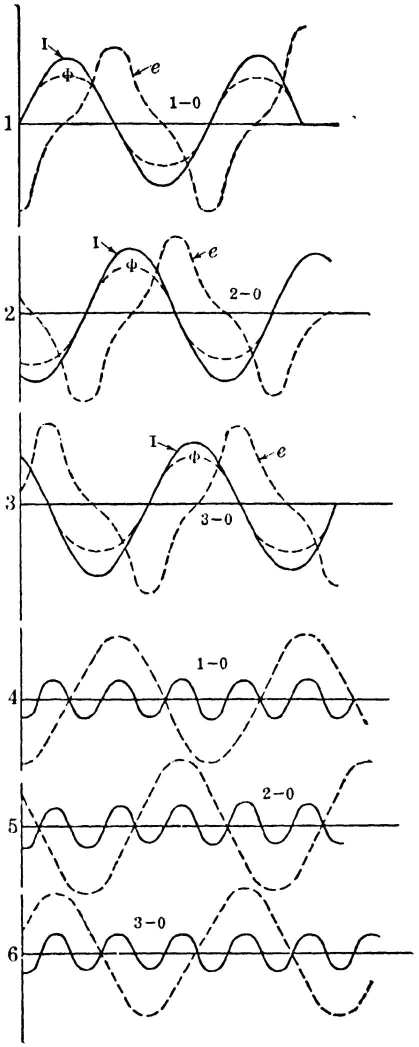
FIG. 29.—The "wabbling neutral" scheme of tripling frequency; the center point of 3 Y-connected iron core coils is connected to the antenna and the center point of 3 Y-connected air core coils is connected to ground. The three-phase power supply is otherwise ungrounded.

It will be noted that the third harmonic components in all three phases are in phase. Thus the potential of point 0, Fig. 29, will fluctuate at triple frequency as shown in curve 7 of Fig. 30. This triple-frequency e.m.f. may be impressed on an antenna circuit as shown in Fig. 29.

The voltages indicated in the above curves exist across the transformer windings, and add up to give a sine wave of c.e.m.f. at the line terminals. This is shown in curve 8, wherein the third harmonic potentials neutralize one another, only the fundamental components combining. This curve is evidently a sine wave, which is as it should be, if it is to neutralize the impressed voltage, which has been assumed as a sine wave.

In Fig. 31 is shown an oscillograph record of this scheme of frequency conversion; three transformer primaries (secondaries open) were connected in Y to a three-phase power line and another Y-connection was made with three air-core coils. The two neutrals were then connected together, and the resulting current in the connection was nearly a pure sine wave of triple frequency.

Losses of Static Frequency Changers.—The above methods of frequency transformation which utilize static transformers possess the disadvantage of excessive iron losses, even though special precautions are taken in the construction of the iron cores, as at the higher frequencies these losses are very large; dielectric losses in the insulation may also be excessive. Probably the most practical would be the Joly arrangement for doubling the frequency, using two of these doublers in cascade to quadruple the frequency, and making the delivered energy thus suitable to the requirements of radio telegraphy and telephony. With every arrangement it is highly



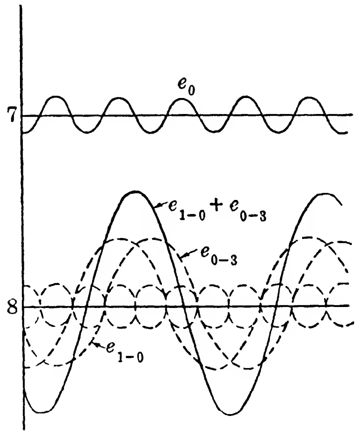


FIG. 30.—Curves of flux and e.m.f. to analyze the action of the wabbling neutral; in curves 8 the voltage forms e_{0-1} and e_{0-3} are shown without their third harmonics, these being shown separately on the X-axis.

important that the secondary circuit be tuned to the desired upper harmonic, as otherwise the higher-frequency current will be relatively small and current of fundamental frequency will probably predominate.

Oscillating Tubes.—Within the last few years many improvements in the design and construction of vacuum tubes have been made and their applications are continually growing more varied and important. At the present time they have not been used to replace the largest arc and a.c. generator equipments but quite possibly if these large stations were to be rebuilt today triodes would be employed. A single water-cooled tube can be built to supply as much as 100 kw. to the antenna and such an amount of power is sufficient for trans-oceanic communication except under adverse atmospheric con-

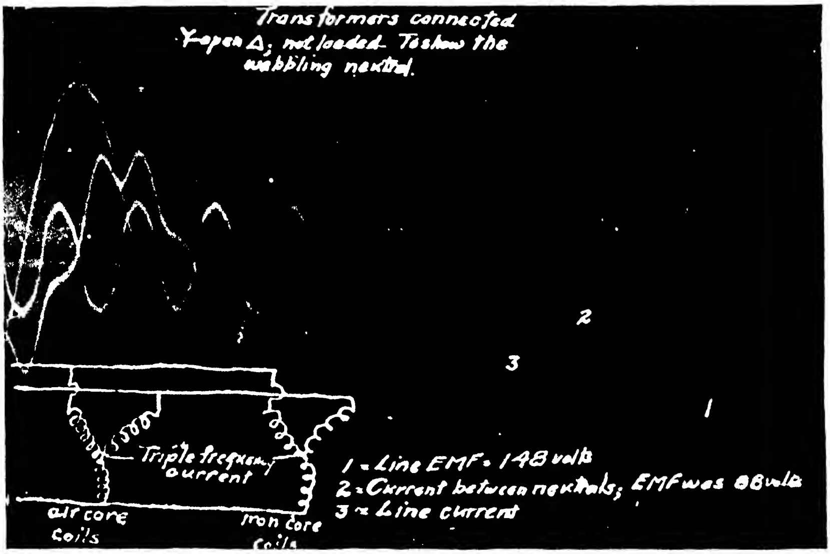


FIG. 31.—Oscillogram showing the third harmonic obtainable from the circuit of Fig. 29.

ditions. Furthermore if there was sufficient demand 500-kw. triodes could be developed.

To supply the high voltage c.c. power to the triodes, banks of two-electrode tubes (diodes), arranged to rectify three-phase (or six-phase) 60-cycle power supply, are used. Thus a high-powered transmitting station, which today may use machines weighing many tons, will require seven or eight tubes instead, weighing possibly a hundred pounds. Even with their water-cooling accessories the total weight is probably not more than a few hundred pounds.

The hundreds of radio broadcasting stations in operation today use many water-cooled tubes, the average rating being a few kilowatts.

Probable Efficiencies of above Apparatus.—*Poulsen Arc.*—Assuming sine waves, a theoretical efficiency of 50 per cent is possible, but probably an actual arc does not give greater than 40 per cent. For instance, the cooling water of a certain 25-kw. arc carried away 14 kw. of heat. For arc oscillations of the third type (p. 718) efficiencies much greater than 50 per cent are conceivable, but as this type of oscillation is seldom used, we assume the efficiency of the normal arc less than 50 per cent.

Measurements made on the arc transmitter at the Eiffel Tower station in Paris indicated an efficiency of 29 per cent; the musical spark equipment gave 47 per cent. With no compensating wave employed the arc efficiency rose to 45 per cent.

Alexanderson Alternator.—No data are obtainable regarding the efficiency of the large Alexanderson alternators, but it seems likely that it is not better than 50 per cent. Examination of the construction of a modern machine shows the likelihood of high iron losses and the care taken to provide adequate cooling [1] facilities substantiates this idea. In the smaller sets, the efficiency may be extremely low: a 200,000-cycle machine, for example, having a maximum output of 500 watts, requires a 10-h.p. driving motor. A large part of the motor output is apparently used in windage losses caused by the high rotative speeds.

Goldschmidt Alternator.—Although one would judge that the efficiency of this type of machine could not be very high, the great care taken in the construction of both the magnetic and electric circuits evidently keeps the losses as small as possible. It is stated by Eccles [2] that a 12½-kw. machine of this type (one of the first to be built) had an efficiency of 80 per cent.

Static Frequency Changers.—It is estimated by the inventor of one of these schemes using iron cores that a 28-kw. transformer will have an efficiency of about 86 per cent.[3] It seems that these devices use about 1 lb. of iron per kilowatt of output and an attempt to calculate the probable eddy-current and hysteresis losses gives a value of perhaps 1 kw. per pound of core used, which would indicate an efficiency in the neighborhood of 50 per cent. It must be pointed out, however, that attempts to calculate the core loss from the ordinary formulas are probably inaccurate, because of the peculiar magnetic cycles to which the iron is subjected.

Oscillating Tube.—The efficiency of an oscillating tube varies a great deal with the adjustments of the circuit, and may have any value between 25 per cent and 95 per cent, neglecting the amount of power used for

[1] See Alexanderson, "Transatlantic Radio Communication," Proc. A.I.E.E., Oct., 1919.

[2] See Eccles, "Wireless Telegraphy and Telephony," p. 230.

[3] Ibid., p. 235.

heating the filament. This point is discussed in detail in Chapter VI, p. 658 et seq. It will be noted in comparing tubes with an arc that they generally consume power only when actually used for transmission whereas the arc is using full power whether the key is up or down.

Methods of Signaling with Continuous-wave Transmitters.—The generators described above will supply a continuous-power input to the antenna circuit and with no changes in the antenna or supply circuit, a continuous undamped high-frequency current will flow through the antenna. The power radiation from the transmitter is therefore constant in magnitude and frequency. Three methods may be utilized for varying this radiated energy in accordance with a prearranged code, and thus transmit intelligence to the distant receiving station. The three methods of sending may be stated as follows:

1. The total interruption of energy radiation during a " space." This is known as the " cut-in " method.

2. Continuous radiation of energy throughout the sending of a message, the space and signal differing only in the wave lengths at which the energy is transmitted. This is called the " compensated " method.

3. The total interruption of energy radiation during a space period, with the radiation rapidly varied by means of a chopper or other scheme, during the signal period. This is known as the " modulated " method of sending, and possesses advantages under certain emergency conditions as described later.

The high-frequency current flowing in the transmitter antenna for each of three methods of sending is indicated in Fig. 32.

Signaling Devices.—For transmitting by means of the above methods one of the following devices may be used, depending on the type of generator used:

I. Chopper or buzzer (choice will depend on the amount of current to be interrupted).
II. Wave length changing switch.
III. Switching to dummy antenna.
IV. Control of excitation of machine.

The application of these devices to the several types of high-frequency generators, previously described will now be considered.

Methods of Sending Applicable to the Poulsen Arc Generator.—This generator depends for its operation on an uninterrupted supply to the arc and antenna circuit (which is the sole natural-frequency circuit of the transmitter). Therefore the means indicated under II and III only can be applied, namely, changing the wave length or switching to a dummy antenna. A change in wave length may be secured by simply connecting the transmitting key so as to short-circuit one or more turns of the antenna inductance when a signal is being transmitted, as indicated in Fig. 33 and to cut in these turns during the space interval.

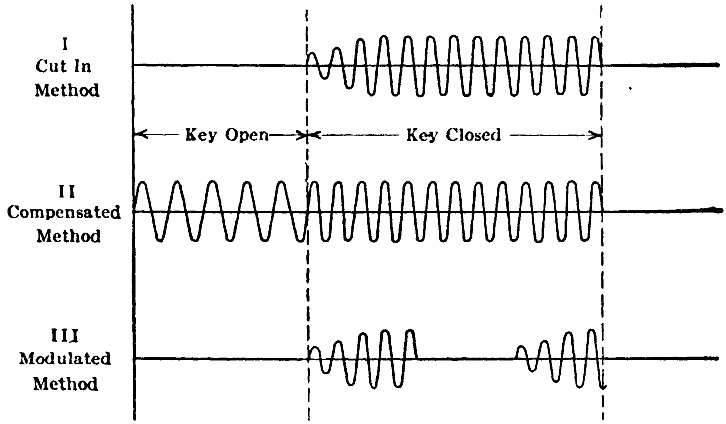


FIG. 32.—Methods of transmitting signals from continuous-wave stations.

This arrangement is practically universal on present arc installations. On the higher power sets, the key does not directly short-circuit the inductance, but operates an auxiliary relay, which in turn actuates the solenoid-operated contactor at the coil. This is required due to the heavy current which must be broken, and rapid signaling would be impossible with the heavy and massive key required if it were attempted to operate it manually. Sometimes, instead of short-circuiting a turn of the antenna load coil, an independent circuit of one or two turns, connected to the antenna load coil by mutual induction, is short-circuited by the relay key.

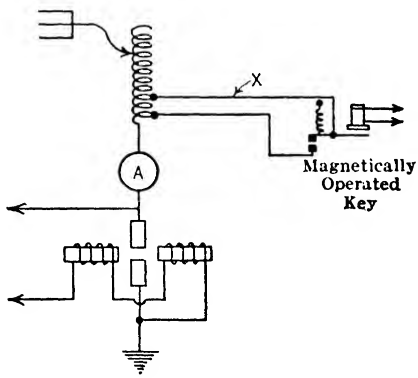


FIG. 33.—The ordinary method of signaling with a Poulsen arc; by short-circuiting a small part of the loading coil the wave length radiated is changed slightly and with a suitable receiving circuit the signal becomes audible.

At high powers the independent circuit may be interrupted at several points simultaneously to reduce the volt amperes interrupted by each contactor, the contactors being actuated by the key as described above.

The connections for utilizing a dummy antenna are shown in Fig. 34.

In this case the key simply acts to transfer the arc circuit to the radiating antenna when it is desired to send a signal. At other times the arc supplies the dummy antenna and practically no energy is radiated. The energy radiation would thus be as shown in Method I, Fig. 32.

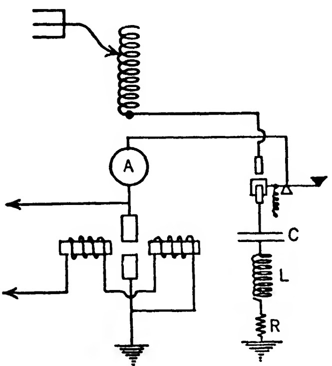


FIG. 34.—Another scheme which has been tried with the Poulsen arc is to switch the arc to a dummy antenna. The switch is of course not a manually operated one, but is magnetically energized from a relay key. High resistances are generally shunted around the switch contacts, to eliminate sparking.

This method is relatively little used, but illustrates the application of switching to a dummy antenna to secure "cut-in" radiation. The constants of the dummy circuit should be identical with the constants of the radiating antenna circuit, so that the conditions at the arc are constant.

Referring to Fig. 33, we may place an interrupter or chopper at X, and thus secure a combination of Methods II and III. The antenna current would then have the form indicated in Fig. 35.

Methods of Sending Applicable to the High-frequency Alternator.—With this generator the frequency is fixed by the speed of the machine. Therefore, transmission by Method II cannot be used (a variation in antenna inductance simply causing a decrease in the amplitude of the antenna current), but Methods I and III are applicable, the former usually being used.

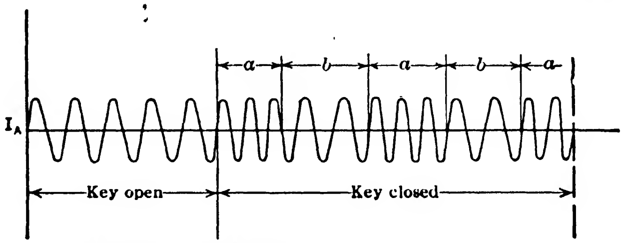


FIG. 35.—A possible type of radiation from a Poulsen arc using the circuit of Fig. 33, with an interrupter of some kind at X.

Signaling is most easily accomplished, however, by control of the excitation, which may simply involve a key in the generator field circuit as indicated in Fig. 36. A resistance may, with advantage, be inserted in the field circuit, to decrease the time constant (L/R) of the circuit and minimize any tendency toward sluggishness which may prevent the signals from being clean cut and distinct, and thus limiting the sending speed.

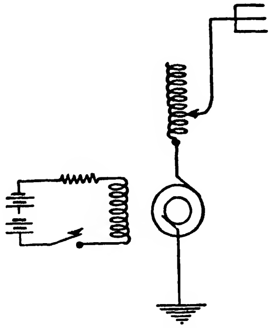


Fig. 36.—The simplest possible transmitting scheme using a high-frequency alternator.

A method has also been developed to control the radiated energy by means of a shunting circuit across the alternator terminals, the impedance of this circuit being controlled by the sending key. The connections are indicated in Fig. 37. When the key is raised (contacts closed) the current flowing through L_2 saturates the iron cores a a, and the reactance of L_1 decreases accordingly. This effectively spoils the tuning of the alternator load circuit and hence brings the alternator ouput to practically zero.

When the core is saturated, the impedance of the shunt circuit is so low as to amount almost to a short circuit on the alternator, under which condition the alternator voltage is very small and is able to send but very little current through the antenna circuit. Therefore the radiated energy will decrease to a very small value, essentially zero. When the key is depressed (open position), the iron is no longer saturated and the impedance of L_1 increases to a high value. The alternator current will then flow through the antenna circuit in preference to the shunt circuit, and energy will be radiated.

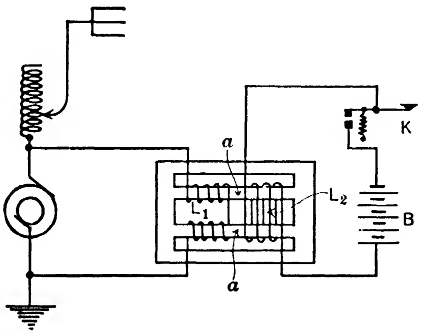


Fig. 37.—A method of sending by generator which employs a magnetically controlled short-circuit on the machine.

In practical installations this variable, iron-cored impedance may be connected in a tuned circuit (tuned when the key is open) which is coupled to the antenna and alternator as shown in Fig. 38. When the key is closed, the local circuit is detuned and the energy input into this circuit becomes very small, the major portion of the energy thus being diverted to the antenna circuit. The transformer indicated in the figure is an integral part of the alternator and is shown supported in two sections above and on either side of the alternator in Fig. 16, p. 730.

For either scheme of control, the energy radiation is essentially as shown in Fig. 32–I. The use of a chopper or buzzer in the exciter circuit may not be entirely satisfactory, due to the inability of the machine voltage to follow accurately the rapid variations of field current produced. There is no doubt, however, that satisfactory results could be obtained by inserting the interrupter in the key circuit of Fig. 38. Radiation in this case would be nearly as indicated in Fig. 32–III.

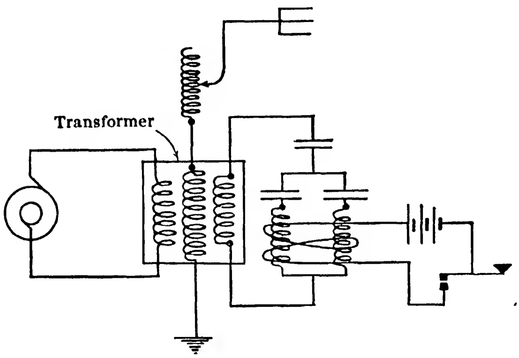


FIG. 38.—In the application of the scheme indicated in Fig. 37 it is found advisable to use the circuit arrangement shown above.

Methods of Sending Applicable to the Goldschmidt Alternator.—As with the Alexanderson alternator, the generated frequency for this machine is fixed by its speed, and therefore wave-changing methods are not applicable. Signaling is accomplished by means of the "cut-in" method using field excitation control, the connections are indicated in Fig. 39. In addition to opening and closing the exciter circuit, the key also simultaneously cuts out or in a portion of the driving motor field resistance. Thus, any tendency of the alternator to suffer a drop in speed, when the exciter key is closed, and the load applied, is compensated for by the cutting in of a certain amount of motor field resistance, which will tend to raise the speed. In addition, the heavy weight and inertia of the rotating element effectually aid in maintaining constant speed, and under operating conditions the variation in wave length is claimed to be less than one-tenth of 1 per cent.

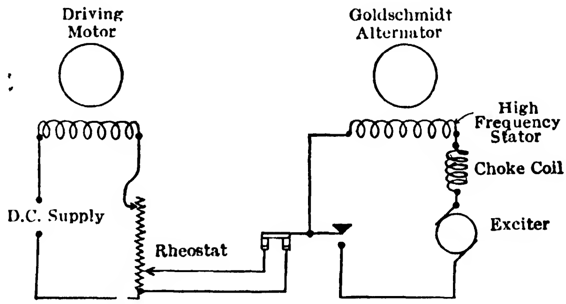


FIG. 39.—Scheme for sending signals with the Goldschmidt alternator using a motor speed control in addition.

The above discussion describes the only method which has yet been used for controlling the output of this alternator. Switching to a dummy antenna, or some form of shunt circuit, as described for the Alexanderson machine, would also be applicable, but the present method seems to be completely satisfactory.

Methods of Sending which May be Used when Frequency Transformers are Used.—Since these transformers must be associated with some form of high-frequency alternator, whose frequency is rigidly fixed by its speed, the same methods as described above for the Alexanderson and Goldschmidt machines will apply. On low-power sets the key may be connected to open the supply circuit directly, while on large-power sets the circuit may be opened indirectly by auxiliary relays actuated by the sending key. The antenna current would then be as shown in Fig. 32-I. The key may also have associated with it some form of interrupter or chopper, resulting in current variation as shown in Fig. 32-III.

For the larger installations, the energy would be controlled by means of the exciter supply due to the smaller power involved. Cut-in sending would be the most feasible, although switching to a dummy antenna or connecting a variable impedance across the alternator terminals as in the case of the Alexanderson machine could also be used.

Control of Radiated Energy when the Oscillating-tube Generator is Used.—The radiation of energy from an antenna supplied from an oscillating-tube generator may be in accordance with any one of the three methods indicated in Fig. 32. The method employed for small sets is usually a direct opening of the antenna circuit by means of the key, which may or may not be associated with an interrupter (usually a buzzer for small field sets) to obtain the modulated method of sending. The wave length change may be obtained by short circuiting a portion of the antenna circuit inductance. Since the power generated by these circuits is as yet comparatively small, there is no necessity for auxiliary relay equipment to be associated with the key. The most feasible control scheme, however, is one which controls the "grid potential" of the oscillating tube; by making this sufficiently negative the tube stops generating power as described in Chapter VI, and illustrated in Fig. 153, p. 606.

Fig. 40 shows the diagram of connections for a small oscillating tube set, which utilizes three of the above-named methods of sending. The oscillating circuit involved was described in Chapter VI, p. 620, and the student is referred there for a discussion of their action.

Referring to the diagram and assuming the switch S thrown upward, i.e., open, it will be noted that the antenna circuit will be completed by the closing of the key K. Therefore, if the key is open, the antenna circuit is open, the tube does not oscillate, and no energy is radiated. When the key is closed, completing the antenna circuit, oscillations will start and be maintained, if the proper conditions have been fulfilled. Thus energy will be radiated as long as the key is held closed and will cease when the key is opened. Therefore with the switch S in the "up" position, transmission is on the "cut-in" method.

If the switch S is thrown to the right so as to make contact with terminal a, then transmission will be by the "compensated" method. This may be seen from the following: with the key open, the antenna circuit is completed to ground through L_x; the tube will therefore oscillate and the antenna radiate energy at a wave length determined by the constants of the circuit, including L_x; the wave length of the energy radiated while a signal is being sent is therefore less than the wave length of the energy radiated during a space interval. Transmission is thus in accordance with Fig. 32–II.

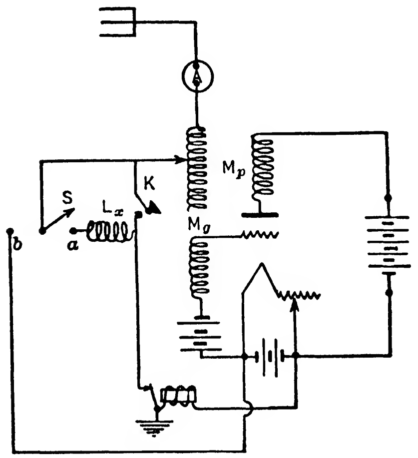


FIG. 40.—An arrangement whereby the output of this small tube transmitter can be controlled by either one of three methods.

Throwing the switch S to the left so as to make contact with terminal b, will permit sending on the "modulated" method. With the key open, the antenna circuit includes the buzzer winding and so the set will not oscillate. When the key is closed, two results are produced: First, the buzzer circuit is completed through the filament battery and the buzzer will vibrate as long as the key is down; second, the vibrating buzzer armature alternately makes and breaks the antenna circuit; therefore, when it completes the circuit, oscillations occur and energy is radiated, while during the break no oscillations are possible. The energy radiated is thus as shown in Fig. 32–III. It should be noted that when the buzzer armature is in the open position, the antenna circuit is not actually opened, but is completed through the filament battery and buzzer winding to ground. Due to the high impedance and resistance of the latter to the flow of high-frequency currents, oscillations are prevented as effectively as though an actual break has occurred in the antenna circuit.

Use of Radiophone Transmitting Set for Undamped-wave Telegraphy.—An effective scheme for transmitting low power undamped wave signals is shown in Fig. 41.

The operation and action of the radiophone set shown is discussed in detail in a later Chapter (see Chapter VIII), and it is there shown that when no sound waves strike the transmitter diaphragm, a high-frequency current of constant amplitude flows in the antenna, and constant power is radiated. When the transmitter is spoken into, the amplitude of the antenna current (and radiated power) varies in accordance with the intensity and frequency of the sound waves set up by the speaker.

Similarly, a buzzer, placed in front of the transmitter, would set up sound waves of constant frequency and intensity, and cause the radiated power to vary in accordance with the pitch of the buzzer note. By using a high-pitch buzzer of proper construction, a very clear transmission, possessing a high degree of selectivity, may be easily obtained.

Another scheme which has had considerable application for vacuum tube transmitters utilizes an alternating voltage for the plate supply. The generator normally used for plate circuit power is replaced by the secondary winding of a step-up transformer, the primary of which is supplied with 110-volt, 500-cycle power. The tube generates oscillations only when the plate is positive; the high-frequency current in the antenna, therefore, consists of a series of wave trains similar to Fig. 32–III. The 500-cycle frequency is used because the old spark sets, which are being replaced by triode transmitters, had 500-cycle alternators and transformers, which could thus be easily adapted for use in the modern transmitter.

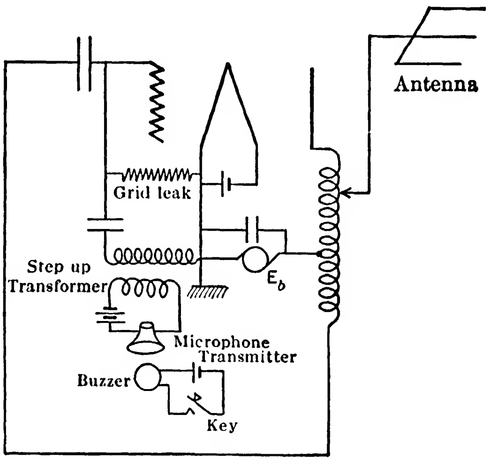


FIG. 41.—A scheme for using a radiophone set to send continuous-wave telegraph signals. Of course, the buzzer and switch may be inserted directly in the circuit, in place of the microphone, if more complete modulation of signals is desired.

Advantages and Disadvantages of the Different Methods.—The advantages and disadvantages of the several methods of sending described above may be summarized as follows:

"CUT-IN" METHOD. *Advantage.*—Only one wave length is radiated after the antenna current has risen to its normal effective value, and energy is radiated only when signal is being sent. Signal is easily received, and permits of a high degree of selectivity.

Disadvantage.—Not suitable for such generators as the Poulsen arc, which do not operate well until the "steady state" has been reached.

"COMPENSATED" METHOD. *Advantage.*—The oscillations are continuous. This method must be employed for the Poulsen arc (neglecting the dummy antenna as an alternative). Transmission is reliable, because the change in wave length, as the key is operated, is positive and certain.

Disadvantage.—The power efficiency is comparatively low because the set requires full power all the time, whether radiating a "signal" or not. A more serious disadvantage arises from the fact that each sending set "uses up" two different wave lengths. This latter feature is especially undesirable when long wave lengths are employed; the difference in frequency of the signal wave and compensation wave should be about 800 cycles per second, and thus the number of arcs which can be used in one district, in the long wave-length range may be seriously limited.

"MODULATED" METHOD. *Advantage.*—The primary advantage of the modulated method is that the signals can be received by means of an ordinary crystal or non-oscillating vacuum-tube receiving set. Thus, if the special continuous-wave receiver is out of service for any reason, the ordinary receiver may be used for reading the message. Radiation occurs only while key is closed, thus increasing efficiency.

Disadvantage.—Less energy is radiated since the energy is broken or chopped into groups. A continuous stream of energy, with given maximum potential on the antenna, sends off more power than a series of "trains" and when utilized in a proper receiving set, permits communication over a greater distance than the modulated signal. With the modulated signal the selectivity is poorer than that obtainable by means of the cut-in method under similar conditions; thus the number of neighboring stations, operating in a given wave-length range, without serious interference, is less.

A Modern Radio Telegraph Short-wave Transmitter.—As was shown in Chapter IV, it is possible to send radio messages completely around the world with comparatively low power, if the proper high frequency is used. However, the commercial communication companies do not depend upon a few watts of power but have many kilowatts available in their short-wave transmitters. Modern short-wave transmitting stations require equipment that permits a rapid change from one wave length to another; the best wave for day transmission is never the best night wave. The outfit must permit accurate maintenance of assigned frequency and assure protection of equipment and personal. Furthermore, there should be a minimum of adjustments.[1]

The equipment is divided in general into two parts, a crystal-controlled oscillator and amplifier terminating with 1 kw. output, and a main power amplifier using four water-cooled tubes capable of delivering 5 to 10 kw. each.

Referring to Fig. 42, we first analyze the power supply. The main rectifier consists of a three-phase, full-wave rectifier using six 872 mercury vapor tubes. With 1520 volts across each secondary of the Y-connected transformer bank the voltage at the output is 3550 volts. The peak inverse voltage in any valve is 3700, and as the valve is rated at 5000 volts inverse voltage there is a reasonable factor of safety. About 1.5 amperes are delivered to the load circuit. The full-load regulation of the rectifier is 10 per cent over all.

The crystal oscillator and the buffer amplifier have a special rectifying circuit; it is necessary that the circuit of the oscillator be free of the voltage disturbances produced by keying and this separate power supply accomplishes that purpose; the rectifier utilizes two 866 tubes. The crystal is maintained at $45° \pm 0.25°$, and the frequency is constant to better than 0.025 per cent.

The crystal oscillator uses a 210 type triode, and the buffer tube is an 860-type tetrode. This screen-grid tube gives some amplification of power at crystal frequency and also prevents the keying effects from repeating back into the crystal circuit, hence its name "buffer"; the 860 type tube has a 75-watt rating. The crystal oscillates at 2500 kc. and gives in its tank circuit about 50 volts of this frequency, and the buffer amplifier raises this to about 400 volts at the same frequency.

The first frequency doubler uses a 160 type tube, which draws 75 milliamperes of plate current when key is down; its control grid bias is 400, enough to bring the plate current to nearly zero when the key is up. When the key is down the average grid current of this tube is 0.012 ampere. The tank circuit has such condensers that it can be tuned from 2500 to 6000 kc. A few turns of the coil are short-circuited for the highest frequencies. Of course the crystal oscillator cannot have its frequency changed; there must be as many crystals as there are desired separate and independent transmitting frequencies. (With one crystal frequencies in multiple ratios only are available.)

The tank circuit of this doubler is tuned to just twice the frequency of the crystal; the crystal generates harmonics, and by tuning the output circuits of the succeeding tubes to higher harmonics the frequency is

[1] See description of typical station (abstracted here) by Byrnes and Coleman, I.R.E., March, 1930, p. 422.

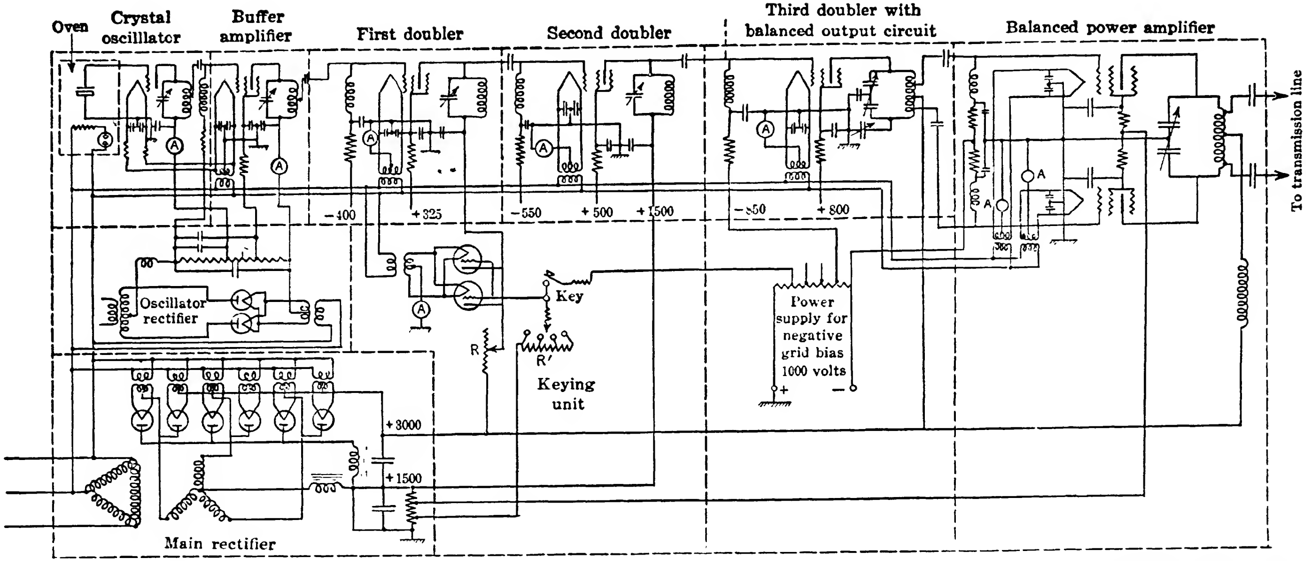


FIG. 42.—Complete circuit diagram of a modern, crystal controlled, 1-kw. telegraph transmitter. By frequency doubling, the output is raised to 20 megacycles, although the crystal oscillator is generating only 2.5 megacycles.

raised. Thus the output circuit of the first doubler is tuned to 5000 kc., and this frequency current flows in its tank circuit. Now the principal excitation in its grid is 2500 kc., so the main component of the plate current will have this frequency. As the tuned tank circuit has a very low impedance for 2500 kc., practically all the power of this frequency must be dissipated on the plate. It is this consideration that requires a 75-watt tube to generate a few watts of double frequency power.

The second frequency doubler is a type '60 tetrode and takes the 5000 kc. from the first, and using the second harmonic of this, generates in its output circuit 10,000 kc. Here again most of the tube's power is used on its plate.

The third doubler uses a '61 type tetrode, of 500 watts rating, and raises the frequency to 20,000 kc., and uses in its plate circuit a specially arranged tank, tuned for 20,000 kc. and grounded at its mid-point. This gives a balanced output circuit for supplying excitation to the balanced "power amplifier." The small variable condenser on the lower end of this tank circuit is for balancing the output circuit for plate-ground capacity of the tetrode; the condenser has to be set to a value equal to the tube capacity.

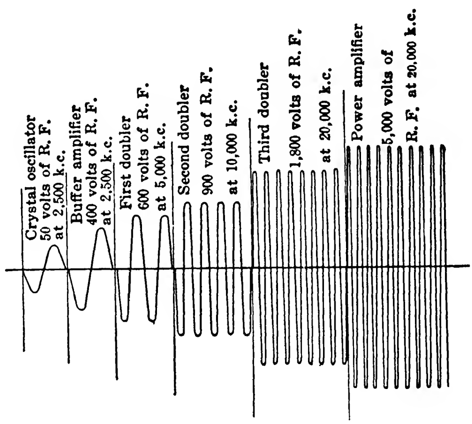


FIG. 43.—In this diagram the change in frequency and power, of the arrangement given in Fig. 42, is indicated.

The balanced power amplifier uses two '61 type tubes; the output circuit is electrically balanced to ground, and the two points where the power is taken off are so selected that the transmission line to which they connect draws a maximum power from the output circuit. To deliver 1 kw. of power at 20,000 kc. the plate voltage required is 3500 and the plate current is 0.45 ampere.

In case transmission is desired on 10,000 kc. the last frequency doubler and the power amplifier each have their tank circuits tuned to this frequency instead of 20,000 kc.

In Fig. 43 is shown conventionally the operation of the different steps of the exciter unit; it shows the gain in frequency and gain in amplitude furnished by the various stages.

Directly at the transmitting antenna there is the last stage of the amplifier, consisting of four type 207 tubes working into a balanced output circuit, two tubes in parallel on each side. A special rectifying outfit, using twelve type 869 rectifiers in a double, three-phase, full-wave circuit, furnishes 10 amperes at 12,000 volts. The grid bias on the amplifier tubes is 1200 volts.

At the high frequency (20,000 kc.), this stage will give 25 kw., using 7500 volts on the plates and 1.3 amperes per tube; when operating at 10,000 kc. it is possible to get 50 kw. of power by raising the plate voltage to 12,000 and the current to 1.7 amperes.

The signaling with this transmitter is accomplished by a special tube arrangement working into the plate circuit of the first doubler. Two 50-watt triodes (type '11) in parallel draw their plate current from the 3000-volt tap of the power supply through a resistance R. The first doubler tetrode also draws its power from the same source through the same resistance. Normally the grid bias on the keying triodes is slightly negative, by connection to the proper point on resistance R'. The amount of plate current thus taken by the two '11 triodes gives such a drop in the resistance R that practically no power is generated in the tank circuit of the first doubler. Now when the sending key is depressed an extra negative bias is put on the grids of the keying triodes so that their plate current drops, and owing to the decreased voltage drop in R the first frequency doubler gets enough plate voltage to operate at normal power. This keying scheme is said to permit the sending of 1000 words a minute.

The filament voltage of the larger tubes is applied in two steps, with a 10-second delay relay to permit the filaments to heat up before full voltage is applied. A copper disc rotatable within the coil of the tank circuit permits vernier tuning; the action of such a disc is discussed on p. 234. The cooling water is tested for purity, and not allowed to show less than 10,000 ohms per cm.[3]

Two separate transmission lines are provided from the exciter unit to the main power amplifier, each being terminated with the right transformer to permit maximum energy transfer.

The rectifier of the large amplifier set is provided with a one-step filter; a 3-microfarad condenser is provided across the line next to the load, and between the rectifiers and this condenser there is a 0.5-henry choke coil.

Special relays permit the selection of any one of eight taps provided on the primary side of the transformer bank supplying the rectifier, thus making possible the use on the plates of the power tubes any voltage between 40 and 105 per cent of normal. Other relays take off the plate voltage if the cooling water gets too hot, or if the filament voltage falls more than 5 per cent. A 30-second delay relay prevents application of plate voltage until the filaments are all up to proper temperature. Other relays " kill " all high-voltage circuits by the opening of a door which has to be passed to permit inspection or adjustment of parts of the apparatus.

Transmission Line between Station and Antenna.—It may sometimes be inconvenient to make direct connection between the power set and the antenna, thus requiring a short transmission line between station and antenna. At radio frequencies, however, the ordinary ideas about transmission lines must be somewhat modified; the reactance due to inductance is so high that the transmission must be thought of in terms of waves rather than in terms of impedance drop.

Radio-frequency power may be sent over lines thousands of feet long, and with reasonable efficiency, provided only that the impedance of the connected load has the proper value. The load must have an impedance equal to the *surge impedance* of the line, which is about 700 ohms for lines of ordinary construction. Neglecting resistance and leakage it can be shown that the surge impedance is equal to $\sqrt{L/C}$ where L and C are the inductance and capacitance per unit length.

As the load will practically never have this resistance (an antenna has

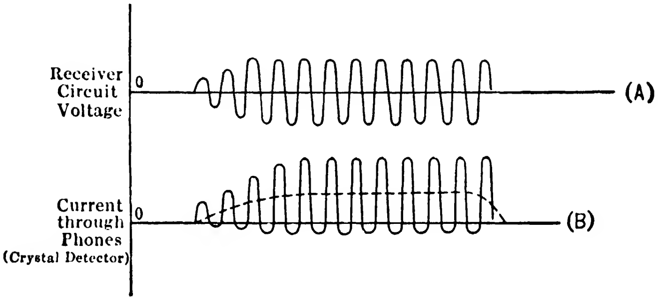


FIG. 44.—Action of crystal detector receiver on continuous-wave signal being sent by the "cut-in" scheme.

only a few ohms) a transformer must be used between the antenna and line having a turn ratio equal to the square root of the ratio of the surge impedance to the antenna resistance.

Reception of Continuous-wave Signal. Necessity for Special Receiving Sets.—That some special means must be provided for the reception of continuous-wave signals, in addition to the simple rectifying device, i.e., a crystal or vacuum tube, will be evident from the following analysis. If we consider an undamped wave-generator transmitting on the "cut-in" method, and this energy being received by a simple crystal or vacuum-tube receiver, the potential across the receiver circuit will have the form indicated in Fig. 44, curve A.

The rectifying action of the crystal or tube produces an asymmetrical change in current through the phones, the mean current being indicated by the dotted line, Fig. 44B. Since the diaphragm is only actuated to give a click when a sudden variation of the mean current through the phone is produced, the result is a slight click at the beginning and end of each signal. Evidently, the message received would be unintelligible.

If we consider signal transmission by the " compensated " method, the results are similar and may be even worse, depending on the sharpness of tuning at the receiving station. Conditions would be as shown in Fig. 45.

It is evident that if the signal and compensation waves are practically

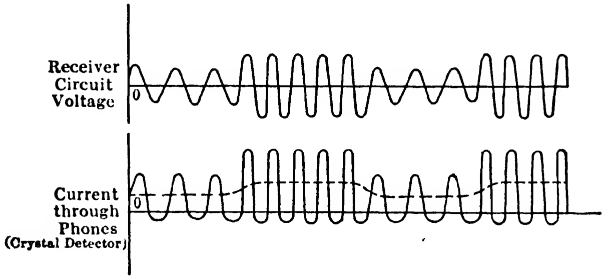


FIG. 45.—Action of crystal detector for compensated continuous-wave signal.

equal in amplitude (as they may be under broad tuning conditions), no clicks at all will be heard in the phones.

If the set is sending on the " modulated " method, the signal is received exactly as in the case of a spark signal. The action is indicated in Fig. 46. This shows the mean current through the phones to vary at audio frequency when the key is held closed, and the signal is thus made audible to the observer.

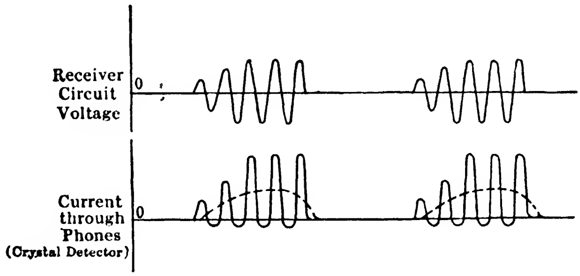


FIG. 46.—Action of crystal detector on a modulated continuous-wave signal.

Action of Continuous-wave Receivers.—The above curves and discussion indicate that, in order to receive continuous-wave signals, some device must be used, which, when interacting with the incoming signal, will give an audio-frequency component. This component, in turn, after rectification, causes pulses of audio frequency to occur in the phones, the corresponding note being heard by the observer as long as the incoming signal energy continues. With the compensated method of sending the space and signal note may also be differentiated by their difference in pitch, as described later.

Continuous-wave Receivers.—In the past, several different devices have been used successfully to break up the received continuous stream of energy, to give an audio-frequency effect. Among these may be mentioned the "tikker" and "chopper" arrangements, which were essentially revolving switches, to open and close the circuit periodically. The Goldschmidt Tone Wheel was a high-speed rotating switch, accomplishing about the same effects as the heterodyne receiver described below. It was limited, however, to relatively low frequencies, say 50 kc., and represented a delicate and rather expensive receiver. Periodic tuning, obtained by rotating the plates of a variable condenser, was also one of the earlier schemes used. The pitch of the received note was fixed by the rotational speed of the condenser.

Heterodyne Receiver or "Beat" Receiver.—The receivers described above have all been superseded by receivers involving the generation of local high-frequency currents by means of oscillating vacuum tubes. The advantages of this type of receiver over the earlier schemes are:

1. Ease of operation.
2. Simplicity.
3. Greater selectivity and sensitiveness.
4. Lower cost.
5. Small space requirements and portability.

Its operation is based on the idea of combining two currents of different frequencies to produce a resultant current, the amplitude of which varies periodically (first used by R. A. Fessenden), the frequency of this amplitude variation being the difference between the two component frequencies.[1] This method is known as the heterodyne or "beat" method, of which two schemes may be used, known as the separate heterodyne and self-heterodyne (autodyne), depending on whether the detecting device is distinct from the local high-frequency generator, or whether the two functions are performed by the same piece of equipment, i.e., a vacuum tube. The former is sometimes simply called the "heterodyne" method, while the latter may be called the "self-heterodyne" or "autodyne" method of reception.

Self-heterodyne Receiver or Autodyne.—The self-heterodyne receiver, utilizing an oscillating vacuum tube as a generator and detector, is undoubtedly one of the most important developments in the field of radio, and will be described somewhat in detail. A possible connection for the receiving set is indicated in Fig. 47.

If the various oscillation requirements of the tube have been satisfied,

[1] See Chapter VI, p. 634 et seq. for mathematical analysis.

the tube will oscillate at a frequency determined by the constants of the local circuit, L_1, L_2, C, and a current of this frequency will flow in the local

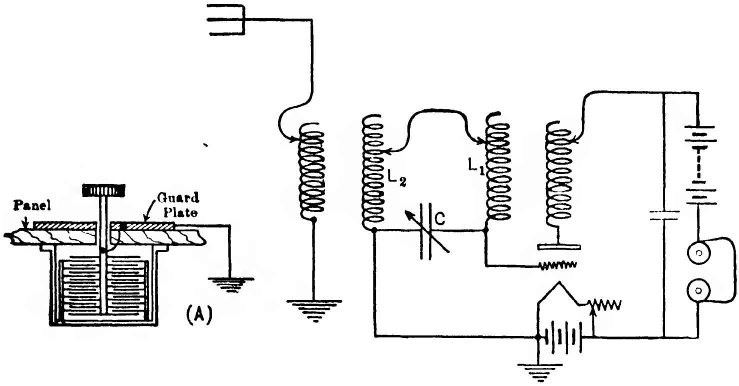


FIG. 47.—The oscillating tube as receiver; it uses the beat note idea and is used to-day universally. Instead of shielding merely the rotor plates of the variable condenser, the modern receiver is built inside a copper-lined box, the copper lining being grounded.

circuit.[1] This is known as the local high-frequency current, and is indicated by curve a, Fig. 48. Assume its frequency to be 1,000,000 cycles/sec. Now consider that the transmitter is operated on the "cut-in" method and is radiating at a frequency of 999,000 cycles per second. A portion of this energy strikes the receiving antenna, which is tuned to it, and a maximum current is caused to flow in the antenna. This, in turn, induces an e.m.f. in the coil L_2 and causes a current whose frequency is 999,000 to flow in the local oscillating circuit. This current is called the incoming high-frequency current and is shown in curve b. (It should be noted that the antenna and local oscillating circuits are slightly detuned, the antenna being tuned to the signal frequency and the L_1L_2C circuit tuned to a frequency about one kilocycle higher or lower.)

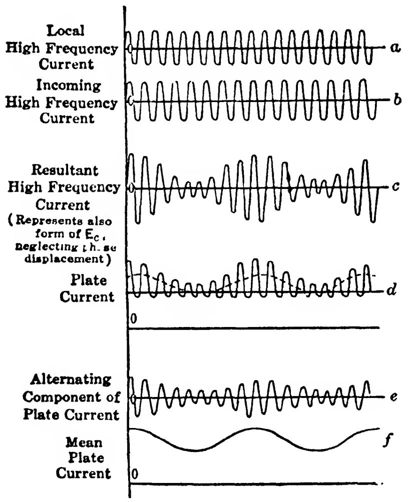


FIG. 48.—Action of the tube as a beat receiver.

[1] For analysis of conditions required for oscillation see Chapter VI, p. 616.

The two high-frequency currents, flowing in the same circuit, combine to give the resultant current indicated in curve *c*, which shows the periodic variation in amplitude produced. These periodic variations in amplitude are called " beats," and the beat frequency is always the difference between the component frequencies. (A " beat cycle " consists of one complete rise and fall in amplitude.) For the values assumed above the beat frequency would thus be $1{,}000{,}000-999{,}000=1000$ cycles per second. It is to be particularly noted that the frequency of alternation of this resultant current [1] is the mean of the two component frequencies, namely, 999,500 for the values assumed. The resultant current is therefore a radio-frequency current.

The drop across condenser C will have the same form as the current curve ($E_c=I/2\pi fC$) and is identical with the variation in grid voltage E_g.

The effect of this variation in grid voltage upon the plate current depends on the point of the characteristic curve at which the tube is being operated. If it is assumed that operation is on the lower bend, the plate current will vary as shown on curve *d*. This variation may be resolved into two components as shown in curves *e* and *f*, *e* flowing through the bridging condenser, while *f* flows through the phones. The latter component varies at beat frequency, and if this frequency is high enough, a musical note is produced in the phones, which is maintained as long as the key is held closed at the transmitter. Opening the key of the transmitter leaves only the local high-frequency current flowing and no variation of plate current at beat frequency is produced, hence no note is heard in the phones. If the tube stops oscillating and the incoming signal is maintained, the same result is obtained.

If it is assumed that the tube is oscillating symmetrically with respect to the upper and lower bends of its characteristic curve, the mean plate current remains unchanged (giving no current of audible frequency) although a beat-frequency variation in amplitude is produced. This means that the tube must be operated on a rectifying part of the curve if a signal is to be heard. Of course if a condenser is used in series with the grid, a signal will be heard, no matter what part of the curve the tube is operating on, as pointed out in Chapter VI, p. 548.

The above discussion indicates that the receiving tube must perform simultaneously the functions of oscillation and rectification. Failure of either would result in no signals being received. These functions, which are performed by the one piece of apparatus in the self-heterodyne receiver described above, may evidently be performed by two different tubes or a tube and high-frequency alternator. Connections for a

[1] On the basis of measuring frequency by the time between successive zero values. At the points of minimum amplitude the phase reverses as explained in Chapter III, p. 321 et seq.

"separate heterodyne" receiver utilizing two vacuum tubes is indicated in Fig. 49.[1]

Control of the Beat Frequency or Pitch of the Signal Note.—It is evident that the local high frequency may readily be controlled by the variable condenser C of Fig. 47. If the incoming high frequency is 1,000,000, and condenser C is of too large a value, the local frequency may be low, for instance 900,000 cycles per second. The beat frequency is thus 100,000 cycles per second, which is above the audible limit. As the value of C is decreased, the local frequency increases, the beat frequency decreases, and as the audible values are reached, the pitch of the note heard in the phones (i.e., the beat note) will change from a very high pitch to lower and lower values, until, when the two frequencies coincide, the beat frequency is zero and no sound is heard in the phones. (In this case we have the addition of two currents of the same frequency, producing a resultant current of constant amplitude. The mean plate current thus has no periodic variation in amplitude; i.e., the beat effect is absent.) As the capacity continues to be decreased, the local frequency increases, and the difference between the local and incoming frequencies again increases; i.e., the pitch of the beat note in the phones again rises until it disappears at the limit of audibility. The above phenomenon is illustrated by Fig. 50. (Curve A.)

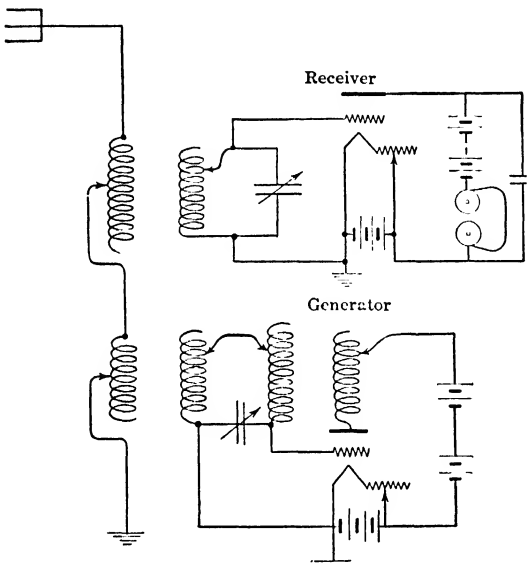


FIG. 49.—Instead of using the detector tube to produce the local oscillations for beat reception, a separate oscillating tube may be coupled to the antenna, or more suitably, to the input circuit of the receiver tube.

In connection with the foregoing discussion it may be noted that in the practical installation or assembly of a heterodyne receiving set, the handle of the variable condenser C should be on the ground side, thus grounding the moving plates. If the apparatus is assembled in a containing

[1] For more detailed study of the action of this type of receiving circuit see Chapter VI, pp. 585 et seq.

case, a metal plate should be placed in front of the condenser and electrically connected to the moving plates. See Fig. 47*A*. This precaution prevents any change in frequency due to the proximity of the observer's hand or body near the condenser and is extremely important on short-wave-length receivers.[1]

Effect of Upper Harmonics.—Since the vacuum tube does not generate a pure sine current of fundamental frequency, but also produces upper harmonics, a unique phenomenon is observed when the heterodyne receiver is close to an oscillating tube transmitter, as may be the case in the laboratory.

Referring to Fig. 50 the combination of the fundamentals will produce the pitch curve designated as *A*, this note being assumed as becoming audible [2] when the condenser is set to the 100° graduation on the condenser

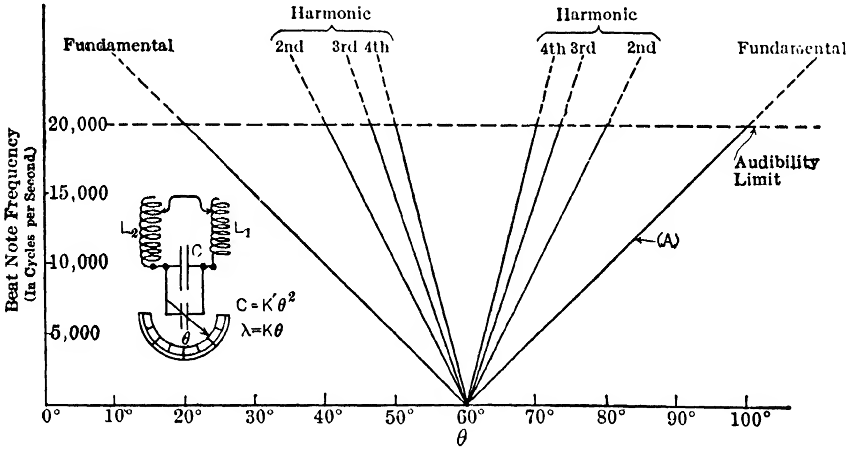


Fig. 50.—A diagram for analyzing the peculiar noises heard when an oscillating tube receiver is close to a continuous-wave transmitter.

scale. As the condenser value is decreased from 100°, a value is reached (at 80°) when the combination of the second harmonics produce a just audible beat note, this note as well as the fundamental beat note being heard simultaneously as the condenser value is further decreased. At certain smaller values (73.3° and 70° on the scale) of condenser capacity, the interaction of still higher harmonics (third and fourth) produces additional beat notes. Thus, in the figure, four beat notes will be heard simul-

[1] Of course a much better scheme is to mount all the parts of the receiving circuit inside of a copper box, grounded; heavy copper mesh is sometimes used.

[2] The upper limit of audibility here assumed, is much higher than that of the ordinary person; generally an adult cannot hear a note higher than 14,000–15,000 complete vibrations per second.

taneously in the phones at condenser adjustments between 70° and 60°. At 60° the fundamental beat note and the upper harmonic beat notes all pass through zero frequency simultaneously, and as the condenser value is further decreased, the beat notes increase in pitch and successively become inaudible again as shown. These effects are summarized in the following tabulation:

	Harmonic	Transmitter	Receiver	Beat Note
C = 65°	Fundamental	1,000,000	997,500	2500
	2d Harmonic	2,000,000	1,995,000	5000
	3d Harmonic	3,000,000	2,992,500	7500
	4th Harmonic	4,000,000	3,990,000	10000
C = 60°	Fundamental	1,000,000	1,000,000	
	2d Harmonic	2,000,000	2,000,000	
	3d Harmonic	3,000,000	3,000,000	
	4th Harmonic	4,000,000	4,000,000	
C = 55°	Fundamental	1,000,000	1,002,500	2500
	2d Harmonic	2,000,000	2,005,000	5000
	3d Harmonic	3,000,000	3,007,500	7500
	4th Harmonic	4,000,000	4,010,000	10000

In actual reception the upper harmonics generated by the receiver are always very much weaker than the fundamental, and when adjustments are made so that the beat frequency heard is one resulting from the combination of an upper harmonic of the local oscillation and the incoming signal, the signal strength and clearness are very greatly reduced. Thus, in adjusting to receive a 1,000,000-cycle wave, the operator may adjust his receiving circuit to a fundamental frequency of 500,500, tuning to the second harmonic (frequency = 1,001,000) for a 1000-cycle beat note, or he may adjust his set to a fundamental frequency of 300,333, tuning to the third harmonic. Similarly he may tune to the fourth or higher harmonics, if present, reception becoming increasingly inefficient and difficult, due to the smaller and smaller amplitudes of these higher harmonic components. This may be seen from inspection of the curves of Fig. 48, p. 766; if the local high-frequency amplitude is small, little change in amplitude occurs in the resultant current, which in turn determines the strength of signal.

Upper harmonics may also be produced by the transmitting set as already noted.[1] In this case the receiving set may have its fundamental frequency adjusted to these upper harmonics, and again weakness of

[1] See "Suppression of Radio Frequency Harmonics in Transmitters," by Labus and Roder. I.R.E., June. 1931, p. 949.

signal and inefficiency result. This possibility, however, is relatively small, since:

1st. The upper harmonics radiated by the transmitter are weak and ineffective unless the transmitter is close to the receiver as assumed in the detailed description above.

2d. The receiving antenna is not tuned to these upper harmonics, still further decreasing their effect on the receiving circuit.

For illustration, assume the fundamental transmitter frequency as 1,000,000, with a second and third harmonic also being radiated. The receiver in this case may be adjusted to a fundamental frequency of 2,000,000 or 3,000,000, but the beat note in either case will be very much weaker than if the fundamental incoming frequency had been utilized. The operator, when in doubt, should vary his local frequency over wide limits, and select that adjustment giving maximum signal strength.

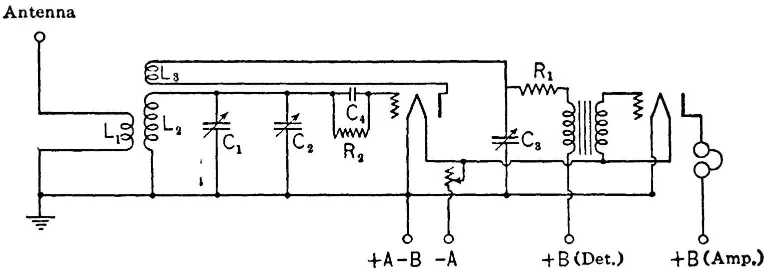


Fig. 51.—Circuit arrangement of a standard short-wave receiver.

Arc transmitters send out many harmonics, as mentioned before, and it often happens that an arc station signal being transmitted on a wave length of perhaps 10,000 meters, may be read on an oscillating receiver which is adjusted for perhaps 1000 meters. The receiver is "picking up" the tenth harmonic of the arc.

Short-wave Receivers.—For those who have not done experimental work with the very high frequency signals being widely investigated today we give in Fig. 51 the circuit arrangement of a beat note receiver of one of the well-known radio manufacturers, adapted for wave lengths from 10 to 200 meters. This means a frequency range from 1500 kc. to 30,000 kc., a band in which an unbelievably large number of continuous-wave channels can be utilized. For continuous-wave telegraph signals the channels may be much closer without producing troublesome interference than is the case for radio telephone communication. A separation of a few hundred cycles, with signals of equal strength, is sufficient to eliminate bad interference when the beat method of reception is being used.

The regeneration control in the set of Fig. 51 is by means of condenser

C_3; the 25,000 ohms resistance in the plate circuit of the oscillator is said to make the regeneration effect more uniform over the range of condenser C_1, which is the main tuning condenser of the set. The beat note varies so rapidly as condenser C_1 is changed that the vernier condenser C_2 is placed in parallel with it. The various condensers used in this set have very small capacities, the range of C_1 is from 6 to 130 $\mu\mu f$; the range of the vernier C_2 is 3 $\mu\mu f$, and the range of C_3 is 200 $\mu\mu f$.

When a 201-A tube (or its equivalent) is used in this circuit the frequency ranges available, with various coils for L_2 and L_3, are as given in the accompanying table. All coils are wound on celluloid tubes 3 inches in diameter, the wire used being No. 16, spaced ten turns per inch.

COILS FOR SHORT-WAVE RECEIVER

Number turns in L_2	Number turns in L_3	Frequency range in mega cycles
1	2	17 -35
3	2	10 -19
8	4	4.8-10.3
18	6	2.7- 5.3
49	18	1.4- 2.8

The antenna coil, of a few turns, is arranged for variable coupling with coil L_2 but the antenna itself is not a tuned circuit. Coils L_2 and L_3 are wound on the same spool there being only 0.1 inch space from the end of one to the beginning of the other. The control of regeneration is by means of variable condenser C_3; the coupling of tickler coil L_3 with respect to coil L_2 remains fixed.[1]

Using receivers of this kind very remarkable transmission has been reported. Small transmitting sets, using one dry cell tube as the source of power, have been sufficient to establish communication over 1000 miles or more distance. An idea of what one of these minute transmitters looks like is obtained from Fig. 52. The tube used is type 199, using only 0.06 ampere in the filament and a few milliamperes at about 100 volts in the plate circuit. With a total input to the tube of 1 watt, and a sensitive receiver, communication has been reported with some consistency over a distance of 1000 miles.

Effect of Frequency Variation in Short-wave Reception.—When receiving on short wave lengths, say 3000 kc. (100 meters or less), it is evident that both the incoming and local oscillations must be very constant indeed if a beat note of constant frequency is to be obtained. Thus, with

[1] For further discussion of the action of detectors in short-wave reception, see article by Karplus, "Communication with Quasi-optical Waves," I.R.E., Oct., 1931, p. 1715.

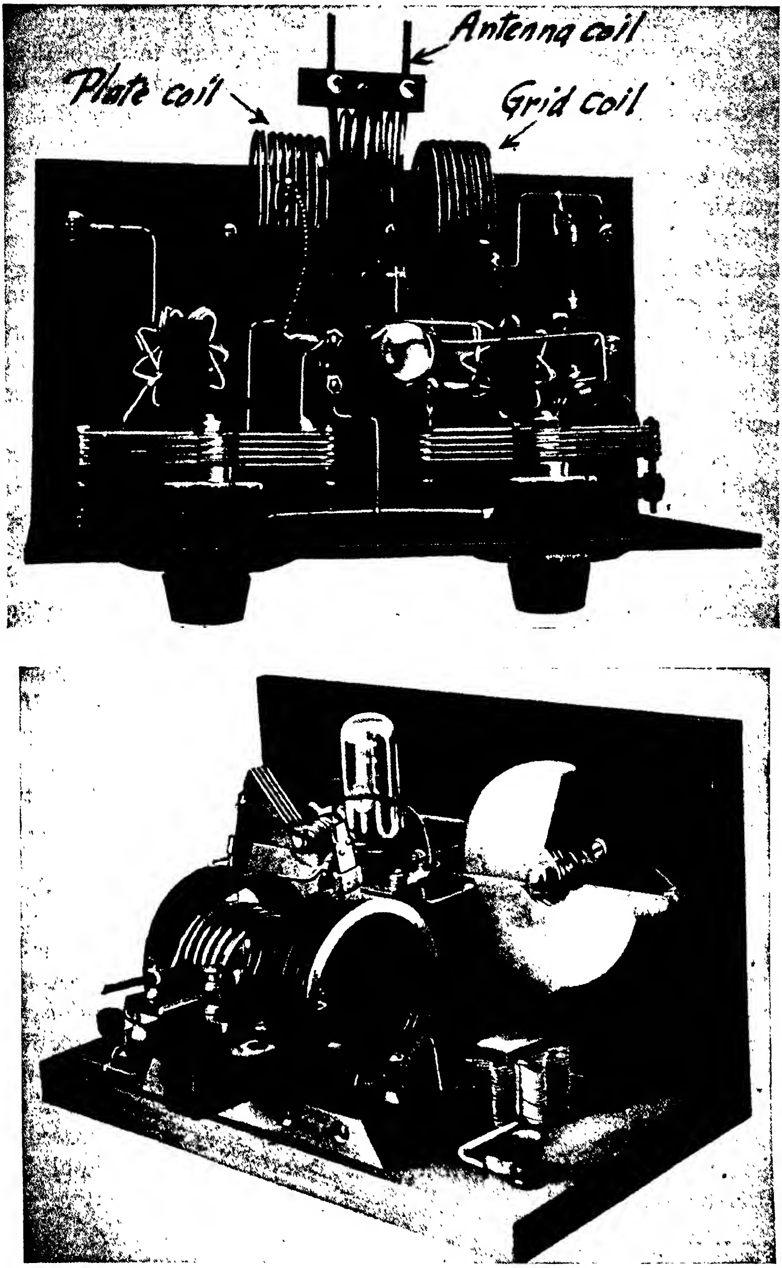


FIG. 52.—Two views of a low-power, short-wave transmitter. Type 199 tube is used as a generator.

1000 kc. incoming, and a local frequency of 999 kc., a 0.1 per cent change in the local frequency will cause a 100 per cent change in the beat note pitch. With a small receiver, it may be impossible to hold the local frequency sufficiently steady. In this case it has been suggested that a periodic variation (say 1000 times per second) be made in the local frequency by using a revolving plate condenser in shunt with the main condenser in the oscillating circuit.[1] This auxiliary condenser varies the local frequency about $\frac{1}{2}$ per cent above and below the incoming frequency (incoming frequency assumed as 3000 kc.), thus the audible beat note range is passed through 1000 times per second, and a signal note of 1000 cycles per second is heard in the phones. Any small variation in frequency say 0.01 per cent, which might make the reception almost impossible with the usual receiver, is thus made negligible in effect by the larger periodic variations imposed by the revolving plate condenser.

At the longer wave lengths, where the above difficulties are not present to such an extent, the following scheme [2] may be used to improve selectivity. The receiver is first tuned to oscillate at the incoming frequency, and then adjusted to be *just on the verge of oscillating.* The incoming signal will then supply enough energy to the grid circuit to cause oscillations; these oscillations exist only during the duration of the signal. This high-frequency oscillation could be coupled to a second tube oscillating continuously and received exactly as described in the normal heterodyne method. Another scheme employs an a.c. bridge arrangement, supplied by a 1000-cycle source, viz., tuning-fork oscillator; normally the bridge is balanced and no signal is heard in the phones. When the receiving tube oscillates, its mean plate potential changes; this changes the grid potential of a tube placed in one arm of the bridge, thereby unbalancing the bridge and causing a 1000-cycle note in the phones.

Radiation from Receiving Sets.—Referring again to the circuit shown in Fig. 47, with the circuits adjusted to the point where oscillations almost occur, maximum regenerative amplification is obtained; this scheme has been very widely used in radiophone reception of broadcast messages. In securing this adjustment, however, it is inevitable that oscillations will occur. The set then acts as a low power transmitter; nearby receivers closely tuned to this same oscillation will hear an interfering beat note produced by the combination of the broadcasting station oscillation and the local receiver oscillation. This interference may be eliminated by making the receiver *non-radiating.* One arrangement for accomplishing this [3] is shown in Fig. 53.

[1] C. S. Franklyn, "Wireless Age," July, 1919.

[2] "Wireless Age," July, 1921.

[3] "Non-radiating Receiving Circuit for Damped and Undamped Waves," J. Scott Taggart, Electrical Review, Nov. 14, 1919.

An inspection of the circuits indicates the scheme of operation. Tube A acts as an amplifier, while tube B acts as a detector when switch S is thrown to point a (damped-wave reception). With S thrown to point b and proper coupling adjustments made, tube B acts as oscillator and detector (self-heterodyne) for reception of continuous-wave signals. Practically no local frequency energy, however, reaches the antenna circuit, due precautions being taken to keep C_3, L_3, and L_4 remote from the closed receiving circuit and antenna.

As a matter of fact there is still some slight amount of power sent from tube B into the antenna, due to the capacity between the plate and grid of tube A. This effect is discussed further in Chapter X under the topic of neutralization, p. 1016.

When using separate tubes for oscillation and detection (regenerative) a certain adjustment of the oscillator gives best reception; this is not improved by changing the coupling on the second tube.[1] If the oscillator adjustment is not at best (optimum) value, then increase of detector coupling will improve the signal strength up to best value. It seems that this best signal is obtained with the same strength of local oscillations for a given circuit and wave length, independent of incoming signal strength (audibility range extending from 1 to 5000).

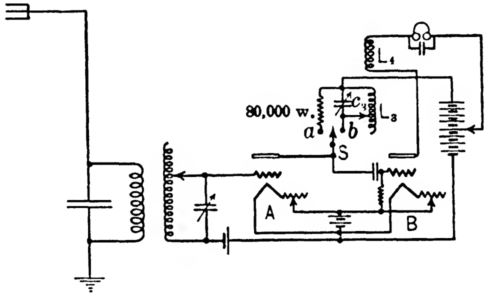


FIG. 53.—A receiving circuit using an oscillating tube which gives practically no radiation from the antenna.

The signal current in the telephone is directly proportional to the antenna current and not to the square as in the case of crystals and non-oscillating tubes. Variations in the ratio of $C : L$ in the local oscillating circuit have no effect on sensibility provided the local oscillations are at optimum value.

Possibility of Receiving Undamped-wave Signals with an Ordinary Crystal.—An ordinary damped-wave receiver, using a crystal or simple vacuum-tube circuit, may, under certain conditions, receive an undamped wave signal. The possibility arises when two undamped-wave transmitters are operating simultaneously at practically the same wave length. Thus, if station A sends at 6000 meters (50,000 cycles), while B sends at 6060 meters (49,500 cycles), currents of these frequencies will simultaneously flow in the receiving antenna, giving a resultant current having

[1] "Notes on Beat Reception," L. W. Austin and W. F. Grimes, Washington Academy of Sciences, March 19, 1920.

a frequency of 500 cycles, which will cause a note of this frequency to be heard in the phones. It is evident that for the signals of either station to be correctly received the second transmitting station must be radiating continuously, acting simply as a high-frequency generator. Thus, if B is sending while A is tuning his set, with key down, operators with crystal

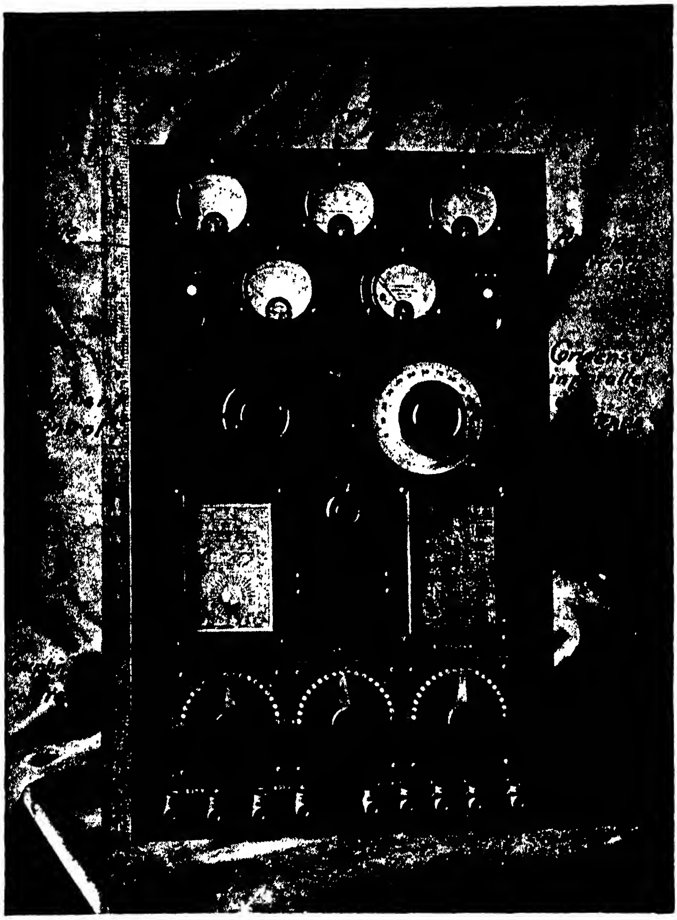


FIG. 54.—Front view of a small continuous-wave transmitter; the high-frequency power is generated by four five-watt tubes.

detector sets adjusted for receiving a frequency close to that used by A and B will be able to read A's signal.

This phenomenon is responsible for the familiar whistling note heard in many broadcast receivers. Two transmitting stations are radiating carrier frequencies which have a beat frequency in the audible range, and this beat note is heard, in spite of the fact that the receiver itself is not oscillating.

Use of Grid Condenser.—It will be recalled that the tube (in the self-heterodyne circuit) must perform the functions of oscillation and detection.

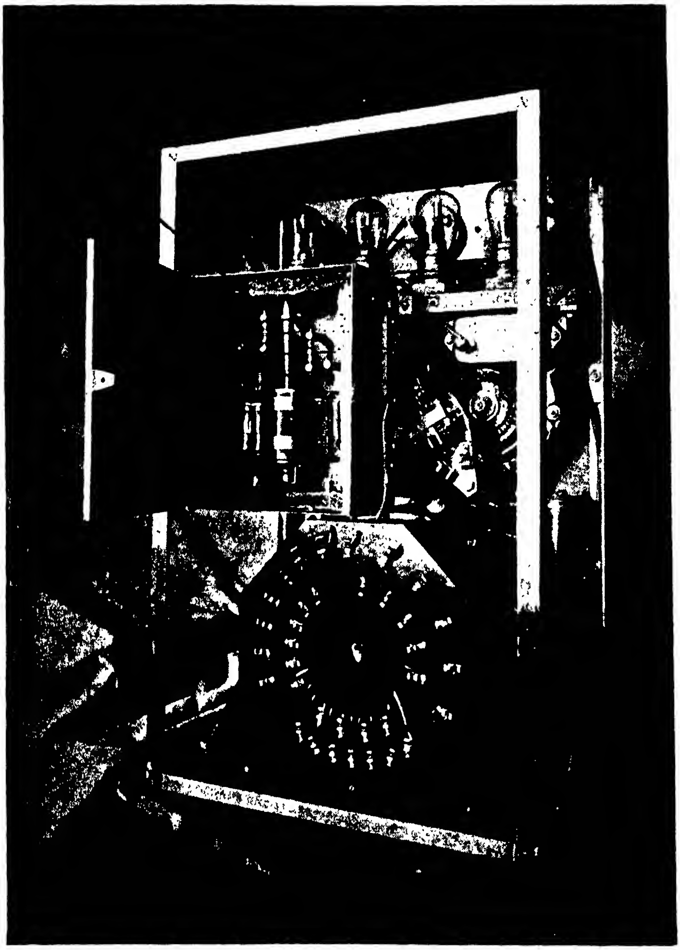


FIG. 55.—Back view of the set shown in Fig. 54; toroidal transmitting coils were used to eliminate local interference. The magnetically operated key is seen in the opened box.

In Chapter VI it was shown that the best point for oscillating is on the straight part of the characteristic curve, while the best point for detection is on the bend. It has also been noted that the use of a grid condenser improves the detecting action and does not require that the tube be operated on the bend of the curve. In fact, the detection is generally best when the tube is operated on the straight portion. For these reasons the grid condenser is also used in connection with the heterodyne receiver.[1]

The one disadvantage of using a grid condenser is the possibility of the tube "squealing" or "clicking" and thus obscuring or preventing entirely the reception of signals. This action has been described in a previous chapter[2] and means employed for its prevention were considered. These means are not uniformly successful in getting rid of the trouble, however, and it is doubtful if the use of a grid condenser would always be desirable.

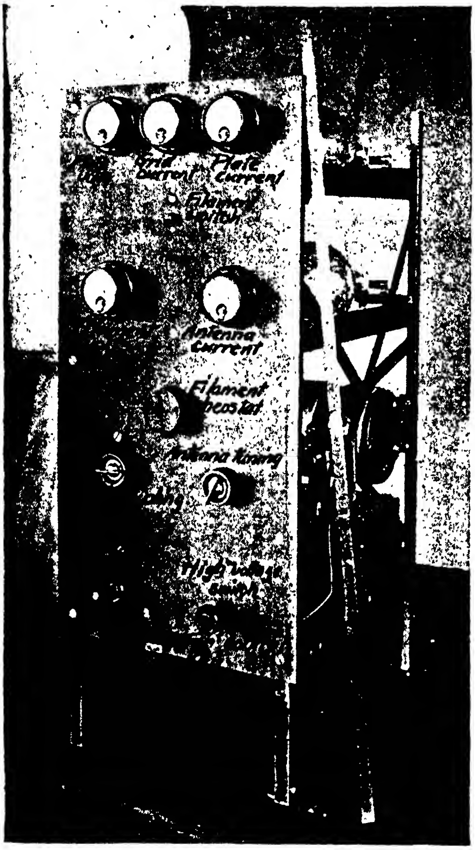


FIG. 56.—Continuous-wave transmitter using four 250-watt triodes; ammeters are supplied for plates and grids, and voltmeter for filament control.

Arrangement of Apparatus in Tube Transmitting Sets.—The exact arrangement of apparatus on a vacuum-tube transmitting set depends of course in general upon the use to which the set is to be put. In so far as possible all the apparatus should be assembled on one board, with suitable instruments, rheostats, etc. Figs. 54 and 55 show front and rear views of a set having an output of about 15 watts; it is intended for laboratory use, so that extreme compactness was not necessary. To eliminate as far as possible disturbances to and from other circuits the coils of the set are made toroidal. An electrically operated key is shown in the rear view, this serving to connect the receiving amplifier and telephones whenever the sending key (which operates the relay) is not depressed. For convenience the filaments of the tubes are arranged for power from the 110-volt c.c. laboratory supply. As the antenna load coil is not adjustable (being toroidal) the frequency of the output is

[1] For analysis see Chapter VI, p. 587.
[2] See Chapter VI, p. 644.

regulated by an adjustable condenser, in parallel with the antenna. A Meissner circuit was used in this set, the plate and grid coupling being adjustable so that maximum output might be obtained, no matter what the resistance of the load might be.

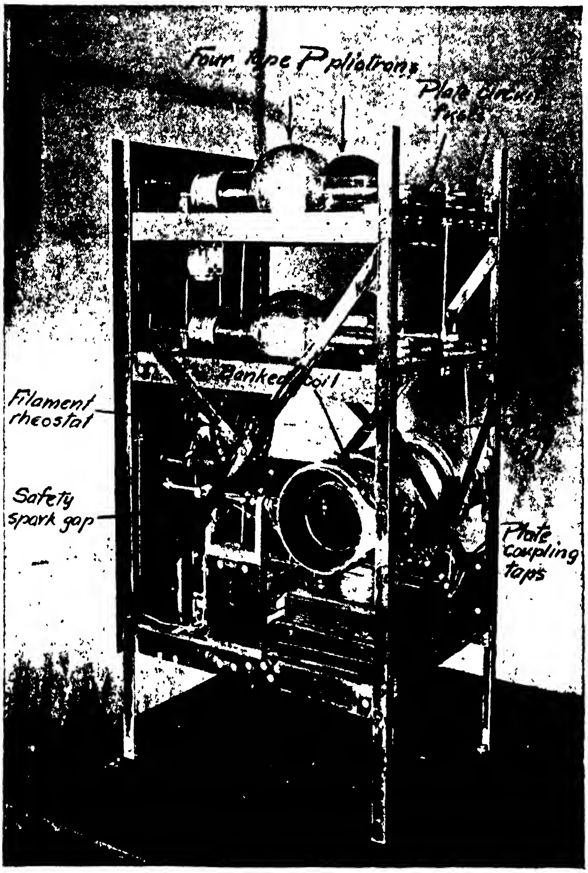


Fig. 57.—Back view of the set shown in Fig. 56; with 1500 volts supplied to the plate circuit this set generates 1 kw. of high-frequency power.

In Figs. 56 and 57 are shown two views of a higher-power set, this using four 250-watt triodes in parallel, and having an output of one kilowatt at 6000 meters.

CHAPTER VIII

RADIO-TELEPHONY

Field of Use.—The radio-telephone supplements the radio-telegraph in the same manner that the wire telephone supplements the wire telegraph. The advantages of the radio-telephone over the radio-telegraph are that, while the latter requires an experienced operator who is familiar with the code, the former does not, and, therefore, the conversation may be carried on directly between the interested parties. In other words, the same factors operate in favor of the radio-telephone over the radio-telegraph as operate in favor of the wire-telephone over the wire-telegraph.

A comparison between the radio-telephone and the wire-telphone is exactly similar to that between the radio-telegraph and the wire-telegraph. The radio-telephone's accepted field of use is from ship to ship, ship to shore, also from airship to airship and from airship to ground, from one moving train to another and from train to station, and, again, in places over land and over water where it would be either impossible or extremely uneconomical to use wires. An example of this last application would be the speech transmission by radio-phone over the ocean, in which case the length of the cable and the impossibility of repeating amplifiers make wire-telephony at this time entirely out of the question; the same is true over a desert or other undeveloped region where it would be far more economical to use the radio-telephone than the wire-telephone. The above does not, however, mean that these two systems of telephony are antagonistic; on the contrary, they supplement each other. A subscriber to a wire-telephone system is now able to communicate with passengers on board ships equipped with radio-phone, the transmission of speech being accomplished by wire overland to a central radio station and therefrom by radio to the ship; it is expected that the same will soon apply generally to airships. Thus, the two divisions of the telephone art will work hand in hand rather than in any way conflict with each other.

For broadcasting purposes radio-telephony is in a field by itself; at the present time an audience of twenty million people is available to the speaker who has something worth while listening to. Because of the vast field, and consequent importance, of radio-telephony a tremendous amount of research and development has taken place recently in the analysis and reproduction of the human voice. Ordinary wire-telephony has consequently profited greatly by the progress of radio-telephony but it has not generally been able to keep pace with the latter. The voice of a speaker today may reach his vast audience over the radio channel with an accuracy of reproduction far greater than that existing over a mile of ordinary telephone line with its associated apparatus. The economic reason becomes evident when we remember that the great expense involved in faithful reproduction of the voice is required at only one station in radio-telephony, namely the transmitting station. It could not possibly pay to bring the ordinary wire-telephone channel to the same degree of perfection as exists today in a good transmitting station.

Outline of Principle of Operation.—The two elements necessary for radio-telephony are, of course, the transmitter and the receiver. We will consider the transmitter and the receiver separately and at first in their simplest forms.

The Transmitter.—Consider Fig. 1, in which the high-frequency alternator, such as an Alexanderson, or Fessenden, alternator, is connected in series with the loading inductance L, the antenna, and the microphone transmitter T. The microphone transmitter may be one of the ordinary carbon granule type, the construction of which is fully explained on p. 790; without going into details, it will suffice to state here that such a microphone consists simply of an elastic diaphragm bearing against a mass of carbon granules enclosed in a suitable chamber: the carbon granules form part of an electrical circuit (in the case of Fig. 1 the circuit of the alternator). When the microphone is not being spoken into the diaphragm remains stationary and exerts a constant pressure upon the carbon granules, the resistance of which remains, therefore, constant. On the other hand, when the diaphragm is set vibrating, as is done by speaking into the microphone or through a noise or sound reaching it, the pressure exerted by the diaphragm against the carbon granules changes, and this change of pressure causes the resistance of the carbon granules to increase or decrease in accordance with the displacement of the diaphragm from its position of rest.

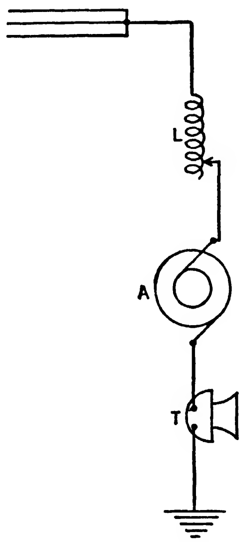


FIG. 1.—The simplest scheme for radio-telephony utilizes a source of high frequency A, and a microphone T, in series with the antenna and a tuning inductance, L.

In the case of Fig. 1, when the microphone is not being spoken into, the alternator produces a high-frequency current of *constant amplitude*, i.e., an undamped current; the amplitude of this current is adjusted to the maximum by adjusting the inductance L so as to make the natural frequency of the circuit equal to the frequency of the alternator. The current flowing through the antenna under these conditions may be represented by Fig. 2, which simply shows an alternating current of constant amplitude, I_o.

Now, assume, for the sake of simplicity, that a vibrating tuning fork is placed in front of the microphone. The harmonic vibrations of the tuning fork will cause harmonic vibrations of the microphone diaphragm, and these will produce variations in the resistance of the microphone. Since no other part of the circuit of Fig. 1 is undergoing any change, it is plain that a variation of the microphone resistance will produce a corresponding variation in the *amplitude* of the high-frequency antenna current. Thus, when the diaphragm is displaced inwardly the resistance of the microphone and, therefore, of the entire alternator circuit, decreases, and the amplitude of the current supplied by the alternator must necessarily increase; the reverse takes place when the diaphragm is displaced outwardly.

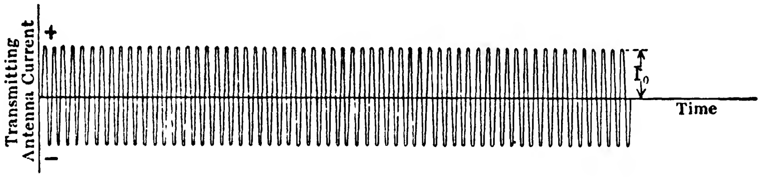


FIG. 2.—When no sound impinges on the microphone the amplitude of the high-frequency current supplied to the antenna is constant.

The antenna current under these conditions would be as shown in Fig. 3, where the curve of the displacement of the microphone diaphragm is also given.

It will be noted that the frequency of the antenna current (*as determined by time between successive zero values*) must remain the same whether the microphone diaphragm is operating or not, since it is solely determined by the frequency of the alternator; but the amplitude of this high-frequency current is made to vary in accordance with the tuning-fork vibrations, in so far as this amplitude changes from the maximum of A_1F_1, corresponding to the maximum inward diaphragm displacement of B_1H_1, to the minimum of C_1G_1, corresponding to the maximum outward diaphragm displacement of D_1L_1. The time between the maximum current amplitudes at A_1 and A_2 or between the minimum amplitudes at C_1 and C_2 is the same as that between the maximum positive diaphragm displacements at B_1 and B_2 or between the maximum negative diaphragm displacements at D_1 and D_2. *Or, in other words, the frequency, with which the antenna current amplitude changes from maximum to minimum* *and back to maximum, is the same as the frequency of the microphone diaphragm and of the tuning-fork vibrations.* When the displacement of the microphone diaphragm is zero, as after the point K, the antenna current becomes the same as in Fig. 2, i.e., of unvarying amplitude.

The antenna current represented by Fig. 3 is said to be "*modulated.*" The high frequency is known in this case as the "*carrier frequency,*" and the frequency of the microphone diaphragm, which is impressed upon the antenna current, is known as the "*modulating frequency.*"

It now remains to show how the modulated antenna current represented by Fig. 3, when received by the receiving antenna, may be made to

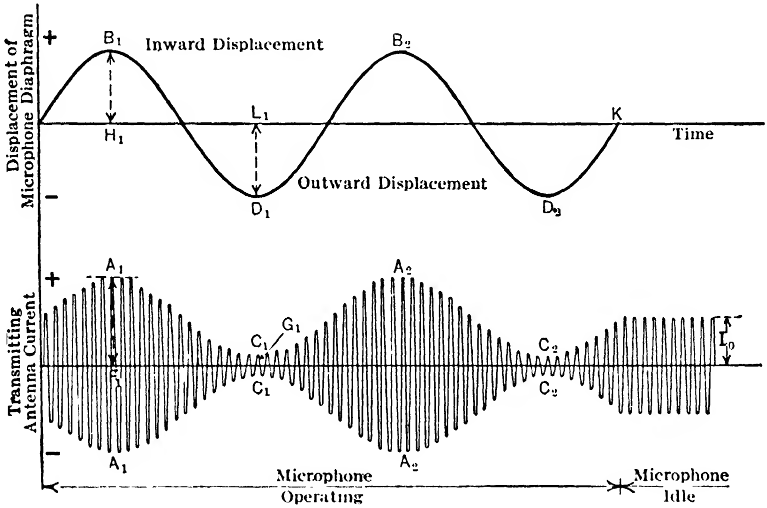


FIG. 3.—If a sound wave actuates the microphone, its inward and outward displacement, varying the resistance in the antenna circuit, results in a high-frequency current in the antenna of *variable amplitude*, called a *modulated* high-frequency current.

so affect the diaphragm of a telephone receiver as to reproduce the note emitted by the tuning fork.

Types of Modulation.—The modulated current depicted in Fig. 3 has its **amplitude** varying at modulation frequency; this type of modulation is universally used today. However, it is possible to carry on radio telephony by other modulation schemes. The *phase* of the carrier current can be shifted back and forth at the modulation frequency, and the *frequency* of the carrier current can be varied in accordance with the modulation. These are called *phase modulation* and *frequency modulation* respectively.

Both phase and frequency modulation give an infinite number of **side bands** (explained later in this chapter) causing interference in neighboring radio channels and so are never used in practice. Neither of them will give an intelligible signal if supplied to a detector circuit; they must

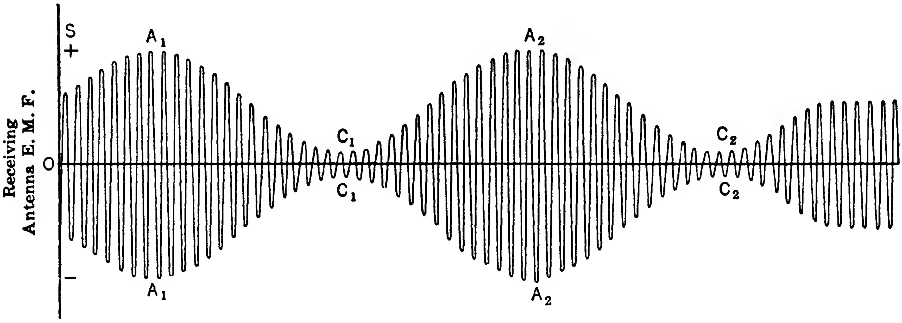


FIG. 4.—The current flowing in the receiving antenna is a modulated high-frequency current, similar in form to the current in the transmitting antenna (providing certain conditions, outlined later, are fulfilled).

first be supplied to a circuit tuned to the unmodulated carrier current. The action of this tuned circuit is to produce a **variation in amplitude** of the current; thus before being supplied to the detector both of these types of modulation must be effectively changed to amplitude modulation.

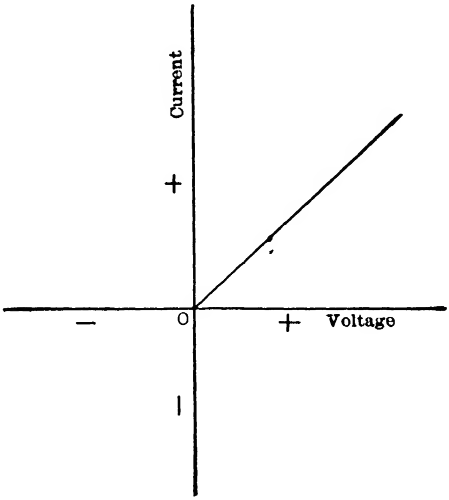


FIG. 5.—To make the discussion of the received signal simple a rectifier with this simple rectification characteristic is assumed.

Whenever the term modulation is used in this text, **amplitude** modulation is assumed.

The Receiver.—The simplest possible receiver is exactly the same as used for spark telegraphy, as shown in Fig. 26 of Chapter V, p. 436. Whereas in this figure a crystal rectifier has been shown, a vacuum detector would probably be used.

The manipulations necessary for the operation of this receiver are the same as for any spark receiver; the antenna circuit and the closed circuit must be tuned to the incoming high frequency, and the coupling between the antenna circuit and the closed circuit should ordinarily be made loose.

It is plain that the e.m.f. impressed upon the receiving antenna, due to the electromagnetic waves emanating from the transmitter, will produce a current in the receiving antenna which will be a reproduction of the current in the transmitting antenna; let it be represented by the curve of Fig. 4, wherein the part between S and K corresponds to a period of action of the distant microphone diaphragm and the rest of the curve corresponds to a position of rest of the microphone diaphragm. Assume, for the sake of simplicity, that the rectifier used in the receiving circuit has the characteristic represented by Fig. 5; i.e., a characteristic such that a negative e.m.f. impressed upon the circuit of the rectifier produces no current whatsoever and a positive e.m.f. produces a current which varies *directly* with the e.m.f.[1]

It is then plain that the e.m.f. impressed upon the receiving antenna and transferred to the rectifier circuit by suitable coupling coils will produce a current in the rectifier circuit of the form shown in Fig. 6. The

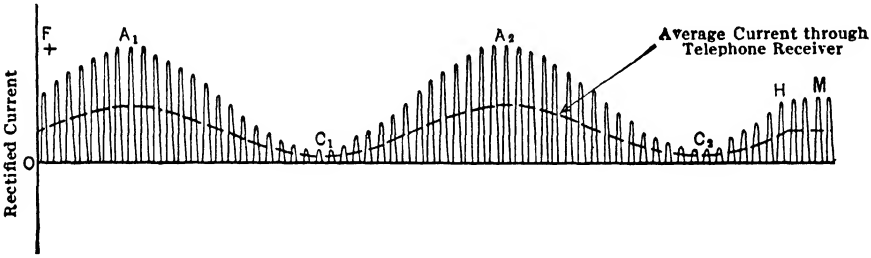


FIG. 6.—The form of current flowing through the crystal and "by-pass" condenser around the phones; the current through the phones is the average value of this unidirectional high-frequency current, shown by the dotted line.

current of Fig. 6, though unidirectional, is yet one which changes at high frequency, and as such it cannot flow through the high-impedance winding of the telephone receiver; therefore, the current in the receiver will be the average current shown by the dotted curve entered in Fig. 6.

It will be noted that the current in the telephone receiver between F and H, Fig. 6, which corresponds to a period of activity of the microphone at the distant transmitting station, is one which changes periodically, between a maximum and a minimum, at the "modulating frequency"; on the other hand, the current between H and M corresponding to a

[1] Rectifiers used in radio work (such as crystal detectors and tubes with or without grid condenser) have a characteristic such that the current varies approximately with the *square* of the e.m.f.; the action of these rectifiers in connection with spark telegraph reception is fully discussed on pp. 441 et seq. The assumption of a rectifier with linear characteristic, as in Fig. 5, does not involve any change in the fundamental principle of radio-phone reception, and is here made purely for the sake of presenting this matter in the simplest possible manner.

period during which the microphone transmitter is idle, is constant. The result is that during this latter period the receiver diaphragm will suffer a constant displacement represented by D_0 in Fig. 7; while during the period of activity of the transmitting microphone the displacement of the receiver diaphragm will change somewhat as shown by B_1-D_1-B_2-D_2 on Fig. 7, or, in other words, the receiver diaphragm will be caused to vibrate at the modulating frequency, i.e., the frequency of the tuning fork, at the transmitting station. Thus, the vibrations of the tuning fork and the sound produced thereby will be duplicated by the vibrations of the receiver diaphragm at the receiving station. It will, of course, be understood that the amplitude of the vibrations of the receiver diaphragm, and hence the volume of sound emitted thereby, will depend upon the strength of the electromagnetic field on reaching the receiving antenna, the height and construction of the receiving antenna, and upon the receiver and detector sensitiveness, etc.

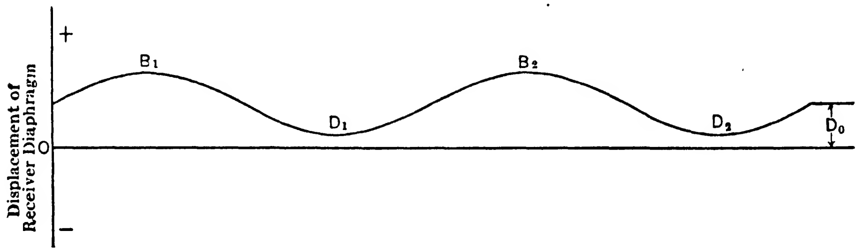


FIG. 7.—Displacement of receiver diaphragm produced by dotted current of Fig. 6; it is assumed that the direction of current through the telephone receiver is such as to weaken the field of the permanent magnet, thus letting the diaphragm move farther out as the current increases.

It now remains to show that such a radio-phone system as was discussed above will transmit speech. That is, it is necessary to show that, if we speak into the transmitting microphone and thereby cause its diaphragm to vibrate in accordance with the complex air vibrations produced by speaking, the diaphragm of the telephone receiver at the distant receiving station will vibrate in such a manner as to reproduce speech.

Frequencies Required for Speech and Music.—To begin with, the very complex vibrations of the microphone diaphragm, due to speech, may be resolved into an infinite number of *harmonic* components of different frequencies, different amplitudes, and bearing certain phase relations to one another. Experimental investigation has shown, however, that, while the number of these components is theoretically infinite, yet, practically, only the components having frequencies between about 100 and 10,000 cycles per second need be considered, since the amplitude of the others is so small as to be negligible.

For music, however, a somewhat wider range of frequency is required. In Fig. 8 there is shown a chart giving the frequency range of all ordinary musical instruments, as well as the human voice. In Fig. 9 the ranges are presented in a different manner, and at either end of the line showing the

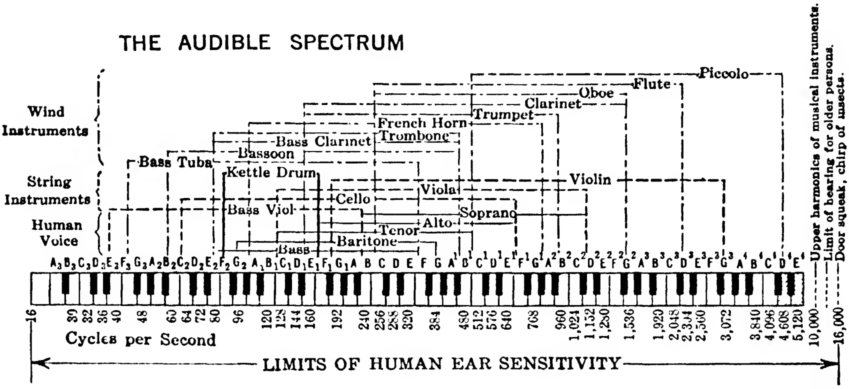


FIG. 8.—Chart showing frequency ranges of the human voice, and various musical instruments, compared to the piano key board (Electronics).

frequency range is shown a cross, marking what might be called the " necessary frequency range " for that instrument. This limit was found by cutting off the upper and lower frequency limits to such an extent that the average listener could tell that the quality, or timbre, of the musical instrument had been affected. If can be seen from this figure that for good rendition of speech and music the frequencies between 40 and 15,000 should be reproduced.

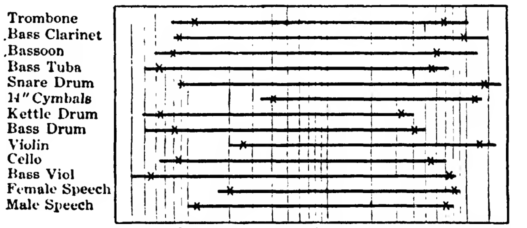


FIG. 9.—The complete frequency range of an instrument is not necessary for its appreciation by the average listener; the upper and lower frequencies may be eliminated as far as the points marked on the various lines, before the average listener can detect any change in the quality.

Phase of Harmonics not Important.—It has been proved [1] that, as long as the amplitudes of the harmonic components of the microphone diaphragm vibrations are reproduced in the vibrations of the receiver diaphragm in the *same ratio* as they have for the transmitter diaphragm, *without any reference whatever to phase relations,* then the speech which caused the vibrations of the

[1] See Bureau of Standards Scientific Paper No. 127, by Lloyd and Agnew.

microphone diaphragm will be faithfully reproduced, *without any distortion,* by the receiver diaphragm. In other words, without paying any attention to phase relations, it is sufficient for transmitting speech that if, as is generally the case, the simple components of the vibrations of the microphone diaphragm are reproduced by the receiver diaphragm with *changed amplitudes, the percentage change be alike for the amplitudes of all the component frequencies.* This principle is of very great practical importance not only in radio-telephony, but in wire-telephony as well.

This idea is illustrated in Fig. 10; here in dotted lines are shown two simple sine waves, one of three times the frequency of the other. This higher frequency is called the *third harmonic* of the lower frequency. For the two conditions, A and B of Fig. 10, the two component waves have the same amplitude, but in B the phase of the third harmonic is reversed from the case of A. The resultant wave for each case is shown by the solid-line curves. Now a pressure wave of shape A, acting on the ear, would produce just the same impression as would the pressure wave shown in B. The two musical tones would seem to have just the same quality.

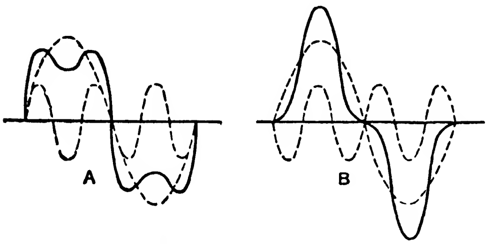


FIG. 10.—Although there still is some difference of opinion it is generally conceded that a sound wave gives the same impression to the brain no matter what the relative phases of its harmonics are; wave B, made up of the same two components as wave A, will sound just the same as wave A.

A musical note made up of a 1000-cycle fundamental and a 5000 to 10,000 harmonic does not sound distorted if the higher frequency is delayed, with respect to the fundamental, as much as 5 to 10 milliseconds. If there is a 50-cycle note as well, however, this can be delayed, with respect to the 1000-cycle note, as much as 75 milliseconds, without sounding badly distorted.

All Modulation Frequencies Attenuated Alike.—We have already shown how harmonic vibrations of the microphone diaphragm having a single frequency, such as those caused by a tuning fork, may be reproduced in the receiver diaphragm. It is plain that the amplitude of the displacement of the receiver diaphragm depends upon the intensity of the electromagnetic field on reaching the receiving antenna and upon the constants of the receiving system, including the construction of the antenna, the sensitiveness of the rectifier and of the telephone receiver, as well as the amount of coupling between the open and closed circuits, the damping thereof, and also whether the rectified current is amplified by a suitable amplifier or not. Of course the intensity of the electromagnetic field at the receiving antenna is a function of the distance between the transmitting and receiving antennas, the wave length corresponding to the carrier frequency, the height of the two antennas and the absorption of energy due to the intervening medium, which is in turn a function of the wave length; hence, no matter what the value of the modulating frequency or the frequency of the transmitter diaphragm, the per cent change in amplitude of the carrier current as related to the displacement of the receiver diaphragm must be the same for all values of modulating frequency,[1] because the percentage of radiated energy which reaches the receiving antenna is dependent upon the *carrier frequency* and *not upon the modulating frequency.* Again, as regards the effect of the constants of the receiving circuit upon the amplitude of the receiver diaphragm displacement the receiving circuit may be so chosen and adjusted that it will affect all modulating frequencies within the speech range to approximately the same extent.

It follows from the above that, if the transmitting diaphragm be spoken into, the displacement of the diaphragm corresponding to each of the possible harmonic components of its vibrations will be reproduced in the receiver diaphragm with practically the *same percentage change in amplitude,* and hence speech will be correctly reproduced.

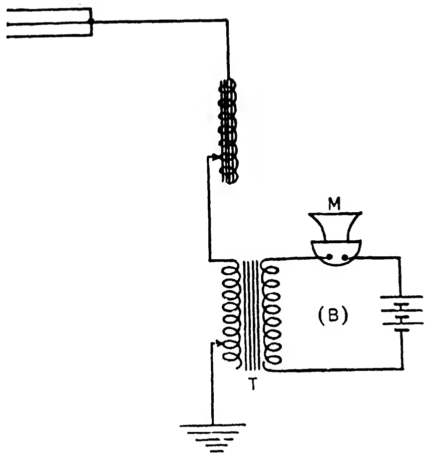


FIG. 11.—Such a scheme as this, dispensing with the carrier frequency, cannot be used because practically no power can be radiated from an antenna with currents of voice frequency.

The carrier frequency should be much higher than the highest important speech frequency, which is in the neighborhood of 5000 cycles per second; therefore, the carrier frequency should be at least above, say, 15 kc. per second, and, as a matter of fact, in actual practice it is seldom lower than 100 kc. per second, and a frequency as high as 40 megacycles per second has been used.

It might be thought that this carrier frequency may be dispensed with and the vibrations of the telephone diaphragm may be caused to produce antenna currents of audio frequency, by means of a circuit arrangement somewhat as shown in Fig. 11, where the microphone M would, on

[1] In case interference between ground wave and sky wave is serious this statement is not true; see p. 385.

being spoken into, produce audio-frequency currents in the antenna, through the means of the transformer T. This system would fail, because it would require a prohibitively large antenna in order that the audio-frequency currents might cause sufficient energy to be radiated for successful transmission over a reasonable distance; hence the use of the "high-frequency carrier." [1]

Sources of Power.—It is hardly necessary to emphasize the fact that the generator of the high-frequency carrier must be such as to cause by itself no change in the amplitude of the high-frequency carrier; otherwise this would be heard in the receiver, together with the speech, and would interfere with the latter. In other words, the high-frequency generator must not interfere with the modulation of the high-frequency current as brought about by the microphone transmitter.

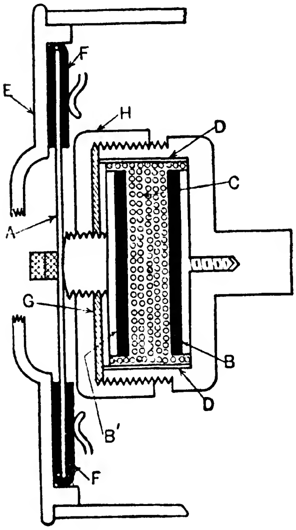


FIG. 12.—Internal construction of the ordinary microphone; the carbon granules between plates B and B' are the seat of the variable resistance.

The sources of power which may be used are those which will produce undamped high-frequency currents. (See p. 713, Chapter VII.) Of these various sources the following have been used for radio-telephony:

The Poulsen Arc.
The Alexanderson or Fessenden Alternator.
The Oscillating Vacuum Tube.

All of the above have been fully described in Chapters VI and VII, and we shall, in this chapter, study the manner only in which each of them may be connected for successful radio transmission of speech.

Before going any further we will first briefly describe various types of telephone transmitters and will later discuss the manner of using them in radio-telephone circuits.

Microphones or Transmitters.—The microphones used in radio at present are generally of the carbon granule type, called the solid back carbon transmitter. These microphones may use single or double cells of carbon granules, the latter being used almost exclusively in broadcasting stations. The single-cell microphone, used in ordinary wire telephony, has essentially the construction indicated in Fig. 12.

[1] It is shown in Chapter IX, that the power radiated from a simple antenna *increases with the square of the frequency.*

It consists of an elastic diaphragm A mounted upon the rubber ring FF, which is in turn held against E, the diaphragm being mechanically connected to the carbon block B'. B' is placed opposite another carbon block B in a chamber filled with small carbon granules C; this chamber is closed by means of the mica washer G and the insulating nut H. The two carbon blocks B and B' form the two electrical terminals of the transmitter; the wall of the chamber containing the granules is covered with a strip of paper designated by D; if a source of e.m.f. be connected to B and B' it will send a current from B through the carbon granules and to B', or vice versa. On speaking into the transmitter the diaphragm is caused to vibrate, and these vibrations are mechanically transferred to the block B' so that the latter's pressure upon the carbon granules is made to vary; this varies the resistance between B and B', and hence it varies also the current in the circuit wherein the transmitter is connected.

Such an arrangement is very sensitive to changes in pressure on the diaphragm and is known as a microphone transmitter. The current carried by such a transmitter is very small because of the fact that a limit is soon reached beyond which " arcs " are developed between granules, the contact points of which become red hot, and the transmitter becomes useless. The current-carrying capacity of an ordinary transmitter is about 0.1 ampere, and its average resistance when not spoken into is 50 to 100 ohms, so that the power capacity is a maximum of $0.1^2 \times 100$ or 1 watt. Some special microphone transmitters " low resistance," may be obtained which have a resistance of 10 to 20 ohms and a current-carrying capacity of 0.5 ampere, or a maximum power capacity equal to $0.5^2 \times 20$ or 5 watts.

The maximum motion of the diaphragm of the ordinary microphone should not exceed about 0.0001 in., otherwise the a.c. output of the circuit will not resemble the sound-pressure wave of input. That is, if the sound pressure is a pure sine wave, such as might be given off from a tuning fork, the fluctuation of current through the microphone will also be of sine-wave form if the to-and-fro motion of the diaphragm does not exceed the value given; if this is exceeded the current wave will depart from sine form being distorted by the presence of both even and odd harmonics. The distortion becomes increasingly greater as the sound-pressure wave increases in intensity.

Double-Button Transmitter.—By using two cells of carbon granules and making connections as shown in Fig. 13 some of the defects of the simple microphone are done away with. As shown in Fig. 13 this microphone consists of a diaphragm stretched between the two chambers containing the carbon granules. Evidently a sound wave which increases the pressure in one chamber will decrease it in the other so that the current in one side decreases as that in the other side increases. This idea of balancing an increasing effect by a decreasing effect has been much used in radio apparatus; it will in general get rid of the distortion produced by even harmonics. To utilize the effect of such a double-cell microphone a special transformer must be used, having a split primary. It may be seen by following the direction of windings and currents that the opposition effects of the microphone do not balance out, but add together in the transformer. Thus the voltage induced in the secondary winding is twice as much as would be given by only one side of the microphone, and is much more free from distortion than the single-cell microphone would be. Furthermore, the steady value of current flowing through the microphone does not magnetize the core of the transformer at all; it is only the change in currents I_1 and I_2 that produce flux in the core and hence voltage in the secondary coil. It will furthermore be noticed that changes in the voltage of the battery have no effect at all on the secondary voltage; such a change affects both I_1 and I_2 equally and so the circuit remains balanced whatever E may become.

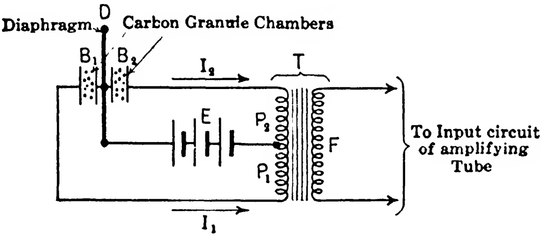


FIG. 13.—In the double-button microphone the motion of the diaphragm compresses the granules in one cell and decreases the pressure in the other cell; this construction, in combination with the split primary transformer does much to prevent distortion of the speech-current forms.

The diaphragm is very tightly stretched, and placed a very short distance from a flat metal plate. The stretching makes the natural period very high and placing the diaphragm close to the metal plate gives a high damping effect. Both of these features of construction reduce the sensitiveness of the microphone so that more amplification must be used and, of course, the amplification itself may bring in some distortion. The engineer has to balance these two effects. His knowledge of the characteristics of amplifiers is such that it pays to make the microphone very insensitive, with the concomitant feature of faithful reproduction.

In Fig. 14 is shown a cross-section of one of these double-button microphones; the carbon granules are shown at F, in one cell above the diaphragm and one below it. This diaphragm is 0.0017 in. thick, of duralumin, and is stretched (by the rings shown in the cut) until its natural frequency is about 6000 per second. The back plate I is only 0.001 in. from the diaphragm when this is in its final position.

Each granule chamber has a volume of only 0.06 cc. and contains about 3000 granules of carbon made from anthracite coal. This is crushed and sifted; the granules to be used must pass through a 60-mesh and be caught by an 80-mesh sieve. These selected particles are treated with hydrofluoric and hydrochloric acids and given a certain heat treatment.

The duralumin diaphragm is gold plated on both sides where it passes through the granule chambers; this is to prevent oxidation of the metal which would introduce an uncertain resistance in the circuit. The gold is put on by the "sputtering" process, from gold cathodes, in a vacuum of about 0.1 mm. About 5 mg. of gold are used, making the gold plate about 0.001 mm. thick.[1]

The granules are held in their chambers by an ingenious gasket construction shown at *G*, Fig. 14. A lot of paper rings (about 30) each 0.0004

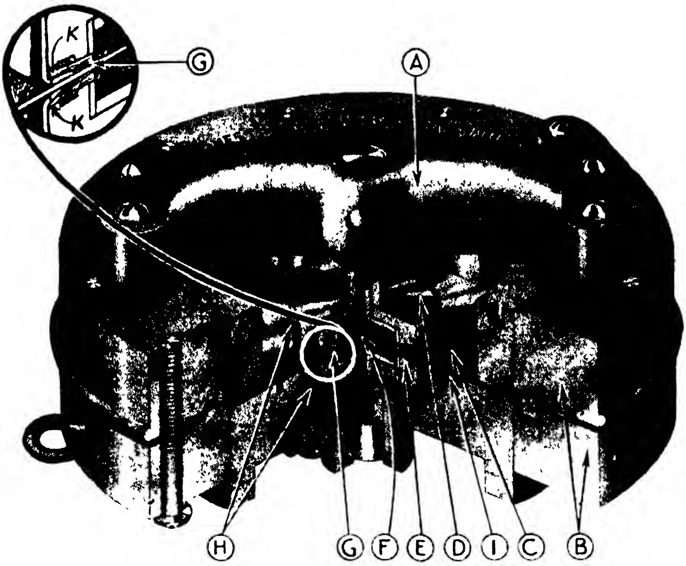


Fig. 14.—Cross-section of a modern double-button microphone.

in. thick are clamped tight at their outer edges, at *KK* of the insert, and spread out at their inner edges, *G*, sufficiently to effectively close the two granule cells. However, these light paper gaskets put practically no mechanical load on the diaphragm, so that it can move back and forth freely. As made today the back plate *I* is properly grooved, to prevent the stretched diaphragm from showing any marked resonance characteristics.

The normal microphone of this type uses a continuous current in each button of about 0.025 ampere. Each button has a resistance of 100 to 400 ohms. It has been found that, when connected up in the simple circuit shown in Fig. 11, the granules are likely to stick together, making the

[1] B.S.T.J., April, 1932, p. 283.

button nearly inoperative. By the use of a simple filter (Fig. 15) to prevent electrical "kicks" from getting into the microphone, the packing of the granules is practically prevented.

Rating of a Microphone.—A microphone is rated in terms of the alternating voltage output generated by a known sound pressure. Naturally this voltage output must depend upon the resistance of the load circuit into which the microphone is sending its power. If the microphone is intended to feed into a vacuum tube amplifier its output transformer will have a high step-up ratio, and if it is to work into a telephone line it will have a somewhat lower ratio.

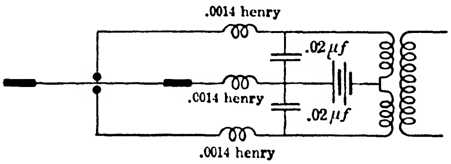


FIG. 15.—A simple one section filter is used to prevent arcing and sticking of the carbon granules.

To make the ratings of microphones comparable it is customary to rate them as of zero decibel level if they give 1 volt output on open circuit, with a sound pressure of 1 bar (one dyne per square centimeter). A microphone that is 20 decibels down, i.e., −20 DB, would be capable of delivering 1/100 as much power as the hypothetical standard one; and as power varies as the square of the voltage, this microphone must generate 1/10 as much voltage as the standard, or 0.1 volt per bar. Most microphones give a response between −50 DB and −60 DB, and so generate a few millivolts per bar. The voltage output is of course measured with a 1–1 transformer. The standard zero level microphone is supposedly measured the same way; the voltage is that developed by the microphone itself, and not affected by the associated apparatus.

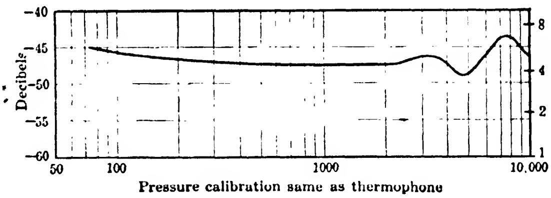


FIG. 16.—Response of a double-button microphone for different frequencies; a thermophone was used to generate the sound wave.

The double-button microphone shown in Fig. 14 has a calibration as shown in Fig. 16; the response scale is given in both decibels and millivolts. It will be seen that the response is remarkably uniform, at 4 millivolts, from 50 to 10,000 cycles per second. This is about 1/100 as much response as the microphone of the ordinary telephone set; the response of the latter, however, does not extend over so wide a frequency range, and is not so free from distortion effects, as the double-button stretched-diaphragm we described above.

In order to prevent jars from giving loud microphone response this is always spring-suspended in a light metal frame to make the microphone independent of floor vibrations, etc.; the frame is generally covered with silk, to keep out dirt and improve its appearance.

The Microphone as a Circuit Element.—The microphone is merely a variable resistance, the variation presumably following the sound pressure waves which impinge upon it; the circuit diagram is given in Fig. 17 (a). Here the resistance of the microphone is given by R_m and that of the rest of the circuit by R. This R is the equivalent primary circuit resistance of the transformer from which the microphone modulated power is taken. It will, of course, not be a true resistance, as there will also be an inductive reactance involved. However, this reactance should be kept as small as possible in the actual circuit to prevent frequency discrimination by the microphone; it may be made quite small compared to the resistance component of the transformer impedance, so we neglect it and treat the transformer as a pure resistance R.

We will assume that the resistance of the microphone may be written

$$R_m = R_0 + R_1 \cos \omega t. \quad (1)$$

This assumes that the resistance of the microphone varies sinusoidally as a sine wave of sound pressure is put on the diaphragm. A possible curve of resistance variation of a microphone is given in Fig. 17 (b); the relation is not a linear one, so that if the pressure is sinusoidal the resistance variation cannot be sinusoidal. However, if a very small pressure change is considered the variation between increments of pressure and increments of resistance is nearly linear, although in an inverse sense. In other words, if the pressure increases 1 per cent the resistance decreases 1 per cent, and vice versa. Thus for a small sinusoidal pressure variation on the microphone the resistance may be assumed to have a sinusoidal variation as is given in Eq. (1). But it must be remembered that for large pressure variation Eq. (1) is not correct and to that extent any deduction reached by the help of Eq. (1) will be correspondingly in error.

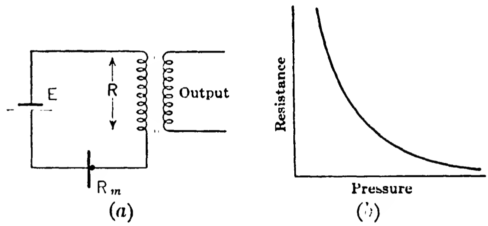


FIG. 17.—The microphone as a circuit element.

Another very important point is involved in this simple analysis. In general, we use in electric circuit theory the principle of superposition; having solved a circuit separately for various harmonic forces, of different frequencies, we add the separate solutions and conclude this to be the proper answer to the problem when all the harmonic forces act on the circuit simultaneously. *This principle of superposition cannot be used when* *any of the so-called constants of the circuit L, C, and R, are not constant.* If therefore we derive from Eq. (1) solutions for various frequencies of sound pressure, say 10 bars of 1000 cycles, 3 bars of 2000 cycles, and 1 bar of 3000 cycles, we cannot add these solutions and get the proper answer as to what happens when a complex pressure wave, made up of these three forces, acts on the diaphragm.

Evidently a general solution to this microphone problem will be extremely involved, so we attempt only the simplest case, that of a single frequency acting in accordance with Eq. (1).

The total fixed resistance in the circuit we call R_2. Evidently

$$R_2 = R + R_0. \qquad (2)$$

Then we have

$$i = \frac{E}{R_2 + R_1 \cos \omega t}. \qquad (3)$$

The solution of an equation of this type is given in any good table of integrals and is

$$i = \frac{E}{R_2\sqrt{1-(R_1/R_2)^2}} + \sum_1^{\infty} \frac{(-1)^n 2E}{R_2\sqrt{1-(R_1/R_2)^2}} \left(\frac{1-\sqrt{1-(R_1/R_2)^2}}{R_1/R_2}\right)^n \cos n\omega t. \qquad (4)$$

If we assume that R_1/R_2 is less than 10 per cent (as in practice it should not exceed this if distortion is to be avoided), then we have as an approximate solution

$$i = \frac{E}{R_2} - \frac{E}{R_2}\frac{R_1}{R_2}\cos \omega t + \frac{E}{2R_2}\left(\frac{R_1}{R_2}\right)^2 \cos 2\omega t - \frac{E}{2^2 R_2}\left(\frac{R_1}{R_2}\right)^3 \cos 3\omega t + . \ . \qquad (5)$$

Whereas for the first term R_2 is the total circuit resistance offered to the flow of continuous current, the resistances R_1 and R_2 for the other terms should be the resistance offered to the flow of alternating current, of the frequency involved in that term.

From the solution of Eq. (5), which was obtained by neglecting some factors that produce distortion, we see that the output is still distorted; putting on a sound wave of sine form gives a microphone output having an infinite number of frequencies. Inspection of Eq. (5) shows that *to be reasonably free from distortion a microphone must be very inefficient.* In other words, the pressure changes on the diaphragm must be small, hence the change in current delivered by E must be small, hence but a very small fraction of the power delivered by battery E is changed into alternating current power resembling the voice wave impinging on the diaphragm. This is apparently a fundamental principle in devices of this kind; to give a distortionless response the device must be operated to give an output which is a very small fraction of its input.

Moving Coil Microphone.—Jones and Giles[1] have applied the principle and general construction of a moving coil loud speaker to make a moving coil microphone. The light diaphragm carries a light coil which moves back and forth in an annular air gap and so generates a voltage proportional to the velocity of the moving coil. Its response is reasonably uniform from 40 to 10,000 cycles but is low, being about −80 DB. Its voltage is only a small fraction of that given by the carbon microphone.

Desirable Characteristics of a Microphone.—It will be realized from the ideas so far presented in this chapter that it is the function of the microphone to produce electric currents which have the same form as the sound waves impinging on its diaphragm. It may well be called the heart of a broadcast transmitting station.

If the electrical engineer once gets currents of exactly the same shape as the sound-pressure waves of the radio performer, he can send out perfectly modulated radio-telephone waves. If, however, the microphone does not faithfully perform its function the radiated signal will, to that extent, be unrecognizable as the performer's voice. The following characteristics should evidently be possessed by a good microphone:

(1) Proportionality between secondary voltage (of transformer of Fig. 13 for example) and the amplitude of the sound waves, this proportionality to hold good over large variations in sound intensity.
(2) For a given sound pressure a given secondary voltage, independent of frequency.
(3) Sensitivity high and constant, irrespective of weather conditions, etc.
(4) Critical damping of its moving parts so that there is no mechanical resonance tending to accentuate some frequencies more than others.
(5) Freedom from noises due to irregularities within itself, such as " carbon noises " and " breathing," a kind of whispering sound which the carbon microphone gives off to some extent.

Condenser Microphone.—This is probably the most perfect transmitter used in broadcasting today but it is somewhat troublesome to maintain.[2]

[1] Jour. Soc. Motion Picture Engineers, Dec., 1931, p. 977.

[2] See "Electrostatic Transmitter," by E. C. Wente, Phy. Rev., July, 1917, and May, 1922. Also article on air damping of condenser microphone, by Crandall in Phys. Rev. for June, 1918.

It is essentially a condenser consisting of two steel plates with air for dielectric. One of the plates is heavy and rigid and forms the solid mechanical frame work of the transmitter. The other plate of the condenser is a thin, tightly-stretched duralumin diaphragm about 0.002 in. thick, insulated from the other plate and supported 0.001 in. away. The diaphragm is stretched sufficiently to give it a natural frequency of several thousand a second, and the back plate is properly grooved, and provided with holes for air flow, to prevent any marked natural frequency, or resonance effects.

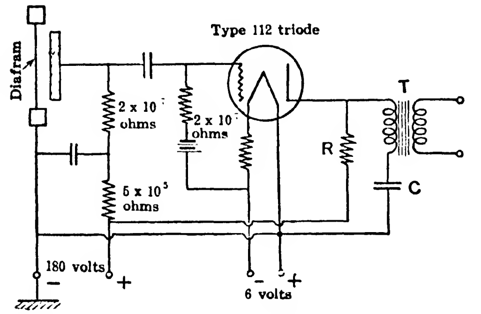


FIG. 18.—Circuit diagram of a condenser microphone.

As the sound waves move the diaphragm back and forth the capacity of the device changes and it thus varies the amount of charge it takes for a given voltage. The condenser is connected to a high-voltage battery, through a high resistance, and the variable charging current drawn through this resistance, produces a varying *RI* drop which is supplied to the grid of an amplifier. As the capacity is only a few micro-microfarads the varying charging current will evidently be extremely small. The insulation of this type of transmitter must be extremely high so that any leakage current will be negligible compared to the charging current. After the microphone is assembled, and joints made tight, it is filled with dry nitrogen, to keep it from rusting, etc.

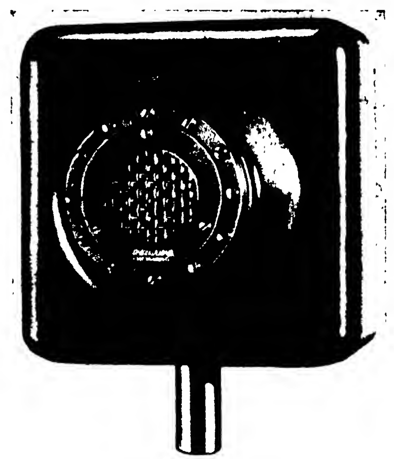


FIG. 18A.—Appearance of a condenser microphone; inside the case, with the vibrating diaphragm, is one stage of transformer repeating amplifier, generally using a low output impedance triode. (Microphone manufactured by Jenkins-Adair Co.)

In Fig. 18 is shown a typical circuit arrangement for a condenser microphone and the external appearance of the microphone is shown in Fig. 18-*A*. The 180-volt battery charges the condenser (i.e., the microphone) through a 20-megohm resistance; when the capacity changes due to sound waves changing the distance between the diaphragm and back plate, a variable current flows through this resistance and so gives a variable voltage which is impressed on the grid of a type 112 triode. The plate circuit of this is excited by the same 180-volt battery, through a resistance R. In this special circuit a transformer T, designed to feed into a 200-ohm circuit, in series with condenser C, is shunted around the resistance R. With the secondary of the transformer open, this amplifying circuit of itself has a frequency response as shown in Fig. 19. The voltage amplification is about two, being somewhat higher in the low-frequency range owing to resonance of the condenser and transformer. This effect would not be there when the proper load circuit is connected to the transformer. This audio-frequency amplifier is practically flat in its response from 40 to 10,000 cycles per second. If it were not like this, it would be of no use to employ the condenser microphone, having a wide frequency response.

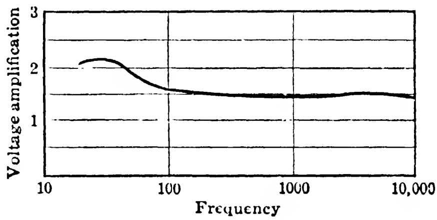


Fig. 19.—Response of the amplifier stage of the circuit of Fig. 18.

In Fig. 20 are shown two calibration curves for a condenser microphone, one taken in the laboratory with a thermophone for sound generator, and the other taken outdoors, to prevent reflections. The voltage given on these curve sheets is that developed across the 20-megohm resistance in series with the 180-volt battery and the microphone. The peak at about 3000 cycles, in the field measurement curve, is probably due to resonance

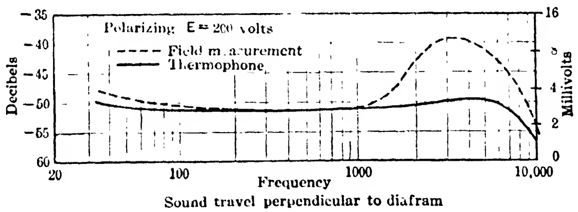


Fig. 20.—Calibration of a condenser microphone sound waves traveling perpendicular to diaphragm.

in the shallow cavity in front of the vibrating diaphragm. When getting this curve the sound was traveling at right angles to the microphone (normal condition); in Fig. 21 is shown the corresponding curve when the microphone had been rotated 90°, so that the sound was traveling parallel to the diaphragm.[1]

[1] A theoretical analysis of methods of calibrating microphones by Sivian, is given in B.S.T.J., Jan., 1931, p. 96.

These calibration curves show the condenser type to be somewhat less sensitive than the carbon button microphone; however, because of the fact that it is more free from back ground noise (" carbon noise ") it is generally used for the highest grade of sound reproduction. It requires, of course, somewhat more amplification.

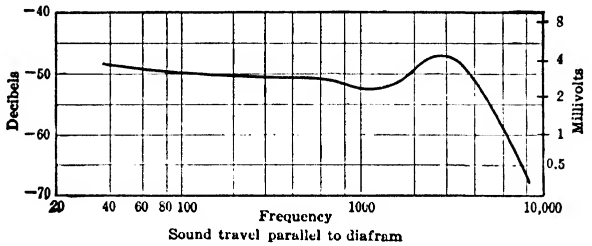


FIG. 21.—Calibration curve of condenser microphone sound waves traveling parallel to diaphragm.

Form of Current Generated by a Condenser Microphone.—The circuit of the condenser microphone is shown in Fig. 22. The condenser C is made to change its capacity by the voice waves impinging upon the stretched diaphragm used for one of its plates and thus changing the distance between the plates. Its capacity is thus a variable, and we will assume that when a sine wave of sound pressure is .cting, the capacity follows the law

$$C=C_0(1+a \sin \omega t) \qquad (6)$$

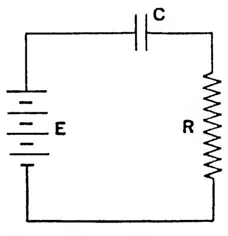


FIG. 22.—Circuit essentials of a condenser microphone.

in which C_0 is the capacity when the microphone is inactive and a is the fractional part by which C_0 changes under the influence of the sound waves.

We may obtain an approximate solution very simply when the ordinarily used values of R and C are considered. Thus $C=500$ micro-microfarads and $R=20$ megohms represent typical practice, so the time constant of the circuit RC is 0.01 second. This means that if the capacity changes, thus requiring the charge to change if the voltage across the condenser is to remain equal to the battery voltage E, this change in charge will be only $1/e$ (63 per cent) complete in 0.01 second. Hence if a 500-cycle, or 1000 cycle, sound pressure wave actuates the diaphragm *the charge in the condenser will not have time to change appreciably.* But if the charge on the condenser remains constant and the capacity decreases by the fraction a the voltage across the condenser must increase by the same factor. In other words, the voltage on the condenser being E when it is at rest, when the capacity suddenly decreases by an amount aC the voltage in the condenser must corresponding increase by an amount aE. However, inspection of the circuit (Fig. 22) shows that the voltage across the condenser, plus that across resistance R, must always be equal to E, the constant battery voltage. Hence when the voltage across the condenser increases by an amount aE *there must be developed across* R *an equal and opposite voltage,* $-aE$. Thus we conclude that as long as the frequency is reasonably high there will be developed across R a voltage equal to aE, where a is the fractional part by which the capacity of condenser C is changed as a result of the sound waves.

More exactly, we attempt the solution of the problem by calculating the magnitude and form of the current from the fundamental circuit equation.

We have
$$Ri+\frac{q}{C}=E \quad \ldots \ldots \quad (7)$$

which becomes
$$Ri+\frac{\int i\,dt}{C_0(1+a\sin\omega t)}=E. \quad \ldots \ldots \quad (8)$$

The solution of this equation is not simple and requires the investigation of the properties of certain infinite series. The solution yields an infinite number of harmonic terms, of magnitudes as follows:[1]

$$i_1=aE\frac{1}{\sqrt{\frac{1}{\omega C_0{}^2}+R^2}} \quad \ldots \ldots \quad (9)$$

$$i_2=aE\frac{2\frac{a}{\omega C_0}}{\sqrt{\left(\frac{1}{\omega C_0{}^2}+R^2\right)\left(\frac{1}{\omega C_0{}^2}+\overline{2R}^2\right)}} \quad \ldots \ldots \quad (10)$$

$$i_3=aE\frac{3\left(\frac{a}{\omega C_0}\right)^2}{\sqrt{\left(\frac{1}{\omega C_0{}^2}+R^2\right)\left(\frac{1}{\omega C_0{}^2}+\overline{2R}^2\right)\left(\frac{1}{\omega C_0{}^2}+\overline{3R}^2\right)}} \quad \ldots \quad (11)$$

$$i_n=aE\frac{n\left(\frac{a}{\omega C_0}\right)^{n-1}}{\sqrt{\left(\frac{1}{\omega C_0{}^2}+R^2\right)\left(\frac{1}{\omega C_0{}^2}+\overline{2R}^2\right)\ldots\left(\frac{1}{\omega C_0{}^2}+\overline{nR}^2\right)}} \quad . \quad (12)$$

[1] Solution by John B. Russell, of our department.—Author.

To get the voltage across R, which is the voltage impressed on the grid of the amplifying tube (after effect of grid resistance and grid leak have been considered), we multiply each of these terms by R. Thus Eq. (9) becomes

$$i_1R=aE\frac{R}{\sqrt{\frac{1}{\omega C_0}^2+R^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

and if R is large compared to $\frac{1}{\sqrt{\omega C_0}}$ this becomes

$$i_1R=aE, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

which agrees with the conclusion reached from the elementary analysis given in the first part of this section. The amplitude of the second harmonic becomes correspondingly $a^2E\ (1/\omega C_0R)$, the third harmonic amplitude becomes $a^3E\ (1/2\overline{\omega C_0R}^2)$, etc.

The amount by which the capacity of condenser C changes when voice waves of ordinary intensity strike it is extremely small. Thus we know (see curve of calibration of condenser microphone, p. 799) that one bar of pressure gives about 0.003 volt across R at ordinary voice frequencies. But the voltage E of the battery is about 200 volts, so we find from Eq. (14)

$$0.003=a\times 200 \quad \text{or} \quad a=0.000015.$$

For a pressure of 1 dyne per square centimeter the value of a is thus about 10^{-5}; that is, the distance between the condenser plates changes by about 1 part in 100,000. But, as the separation normally is only 0.001 in., this means that the movable diaphragm of the condenser microphone moves back and forth about 10^{-8} in., that is, one hundredth millionth of an inch. This is less than one two-thousandth of the wave length of yellow light. Hence even with very large sound pressures the motion of the diaphragm is only a small fraction of the wave length of light.

The Nature of Speech and Music.—The faithful reproduction of the voice naturally requires a knowledge of what the voice waves are and has thus stimulated workers to find the answer to the question. In wire-telephony the knowledge is not so important because the wire-telephone channel does not attempt to carry all the frequencies of the voice. Just enough of the frequency band is transmitted to enable the listener to make an intelligent guess as to what is being said. But in radio broadcasting it is economically feasible to send out practically perfect speech and for this an intensive study of speech has been made.[1]

[1] See "The Nature of Language" by R. L. Jones, Journal A.I.E.E., April, 1924; also "Transmission and Reproduction of Speech and Music," by Martin and Fletcher, Journal A.I.E.E., March, 1924.

The vowel sounds are in general of low frequency; these are easy to amplify and transmit and ordinary telephone communication consists very largely in the transmission of vowel sounds. The consonants, especially those like p, f, th, etc., never reach the listener; his mind supplies these consonants in the received speech and ordinarily the listener thinks he heard them over the telephone wire. Over a good radio channel the consonants are carried with remarkable clarity. Summing up the results of much research work we can conclude:

(1) The frequencies encountered in human speech are within the range of 100 to 9000 complete vibrations per second.

(2) The energy contained in speech is carried almost completely by frequencies below 500 but the quality and intelligibility of speech is determined very largely by the frequencies higher than 500.

(3) The average power output of the average normal voice is about 75 ergs per second or 7.5 microwatts.[1]

(4) The average male voice exerts a pressure of about 10 dynes per square centimeter at a distance 3 cm. from the mouth of the speaker.

(5) The human ear can detect sounds, at a frequency of about 1000 cycles, if the sound pressure is as low as 0.001 dyne per square centimeter.[2] If the pressure exceeds about 1000 dynes per square centimeter at this frequency, the ear is practically paralyzed in so far as sound is concerned and the sensation is one of feeling rather than hearing.

(6) The ratio of peak power in the voice (accented syllable) to average may be 200 to 1. Thus an average voice of 10 microwatts shows peaks of 2000 microwatts.

(7) The range of frequencies encountered in orchestral selections is from 40 to 15,000 cycles per second.

(8) The range of intensity in orchestral selections is at least 100,000 to 1, but if the range is reduced by a monitoring system to 1000 to 1 the reproduction of the program is reasonably satisfactory.[3]

In Fig. 23 is given the compilation of the results of many experimenters

[1] See "Speech, Power and Energy," by C. F. Sacia. Bell System Technical Journal, Oct., 1925.

[2] This pressure per square centimeter corresponds to the weight of a piece of human hair about one-third as long as its diameter.

[3] An excellent summary of the constants and fundamental ideas involved in speech analysis is given by Harvey Fletcher in the Bell System Technical Journal for July, 1925.

on the sensitiveness of the normal human ear. The upper frequency limit of hearing decreases with increasing age; thus a child hears 20,000 vibrations per second whereas an adult may cease to hear above 10,000 vibrations. This effect is entirely different from ordinary deafness because the adult hearing for the lower frequencies may be quite normal.

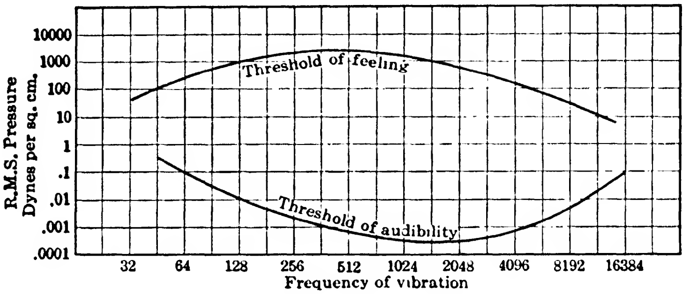


FIG. 23.—Showing the relation between pressure and frequency, for notes just audible and notes so intense as to be painful to the listener.

In Fig. 24 is shown the energy distribution throughout the frequency range for ordinary speech and in Fig. 25 is shown how intelligibility depends upon the frequencies present. It will be seen that speech is nearly perfect (in so far as intelligibility is concerned) if all the frequencies below 1000 are eliminated, yet the curve of Fig. 24 shows that practically all the energy of the voice is below the 1000-cycle frequency.

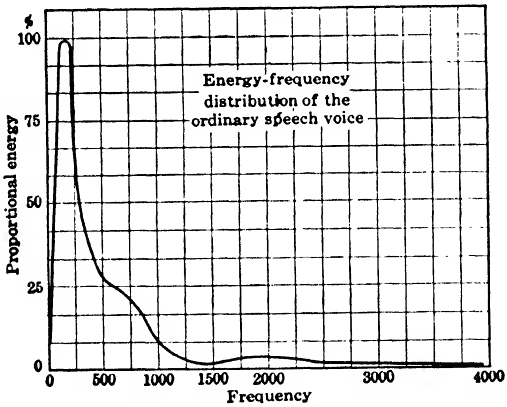


FIG. 24.—Showing the distribution of energy of speech.

Amount of Sound Power Required.—By noticing the manner in which different people adjust their radio receivers it is evident that some desire a much louder reproduction than others, so that a figure on the amount of power required for satisfactory audition must evidently be a general average. For the average home room, a sound power (acoustic power) of 1 milliwatt is plenty unless the program is to be heard through a lot of interfering noise. This figure is an average one; a program will vary in intensity from perhaps 10 microwatts to several milliwatts. For larger rooms, like auditoriums, more power is required, the average being about as shown in Fig. 26.[1]

Conditions for Best Modulation.—We will again note that the speech transmission is brought about simply by changing the amplitude of the transmitting antenna current (modulation of antenna current); in other words, if the amplitude of the antenna current should be changed but little by the operation of the telephone transmitter, speech would be transmitted but poorly, and vice versa. In other words, the range and quality of transmission does not depend upon the amount of current in the transmitting antenna, but upon the *change in this current* or the extent of the modulation. Hence, a radiophone system should be so designed as to enable the telephone transmitter, when spoken into, to produce the *maximum possible change* in the antenna current. This corresponds to a condition where the antenna current amplitude is caused to reach a minimum of zero, and a maximum which should be dependent upon the characteristic of the rectifier in the receiving circuit. It is therefore reasonable to investigate at this point the effect of the rectifier characteristic upon the best conditions for modulation.

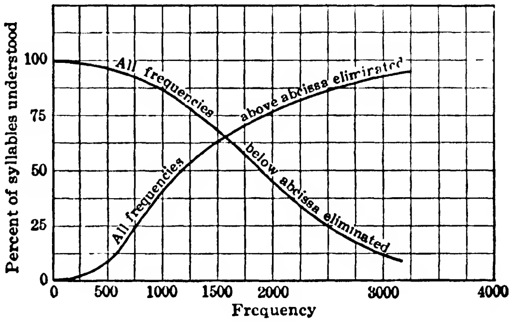


Fig. 25.—Showing how the intelligibility of speech varies as various frequencies are eliminated. If all the frequencies below 1000 vibrations a second are eliminated, the speech is still 85 per cent intelligible, yet most of the speech energy has been thrown away, as evident from Fig. 24.

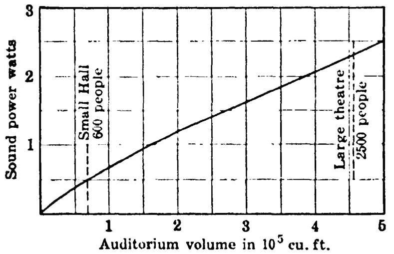


Fig. 26.—Sound power required for rooms of various sizes.

Analysis of Modulation.—Assume, as before, the simple transmitting circuit represented by Fig. 1 and the simple receiving circuit represented by Fig. 4; and let us again suppose that a harmonically varying sound pressure is impressed upon the microphone diaphragm by means of, say, a tuning fork placed in front of it. We then desire that the telephone receiver in the receiving circuit shall give off a pure sine-wave tone of the tuning fork frequency.

[1] This curve is from a paper by Blattner and Bostwick, Journal of the Society of Motion Picture Engineers, Feb., 1930, p. 161.

We will first investigate the case where the amplitude of the transmitting antenna current is made to change by the action of the microphone from *a maximum of twice that corresponding to the microphone idle to a minimum of zero,* this being what is known as a "completely modulated current." Fig. 27 shows the curve of the e.m.f. produced in the

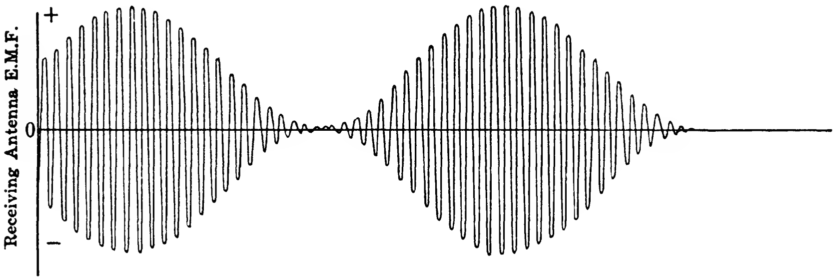


FIG. 27.—A completely modulated antenna current, having a sine-wave envelope.

receiving antenna circuit by the flow of the completely modulated current in the transmitting antenna; the e.m.f. across the rectifier in the receiving circuit of Fig. 4 will be of the same form as Fig. 27, though, of course, reduced in amplitude.

If we assume, as we did before, a rectifier giving a rectified current

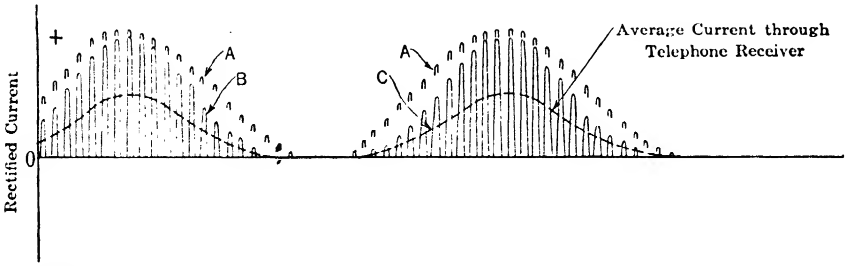


FIG. 28.—The current of Fig. 27, in combination with such a rectifier as that assumed in Fig. 6, would give a rectified current as shown by the points A; its average value would be a sine-wave current. If an ordinary rectifier is used the rectified current is as shown by the solid line curves, the average value of which is shown by the dotted curve which is not a simple harmonic current but is more complex in form.

proportional to the first power of the impressed voltage, then the harmonically modulated e.m.f. of Fig. 27 would produce a rectified current the amplitude of which would vary harmonically, as shown by the points marked *A* in Fig. 28, and the result would be that the average current in the telephone receiver would also vary harmonically and cause this to give off a pure harmonic note of the same pitch as that of the tuning fork.

But, as already pointed out in Chapter VI, practically all rectifiers give a rectified current proportional to the square of the impressed voltage; hence the harmonically modulated e.m.f. of Fig. 27 would produce the rectified current represented by curve B in Fig. 28, and the curve of the average current in the telephone receiver would be as represented by the dotted curve C of Fig. 28, and would, evidently, not vary harmonically. So that, in this case, the receiving circuit telephone would give off a note, which, though of the same pitch as that impinging upon the transmitting microphone, would be of a more *complex quality.*

To remedy the objectionable condition brought about by the combination of a harmonically modulated transmitting antenna current and a rectifier giving a current proportional to the square of the impressed voltage it is necessary that the transmitting antenna current be differently

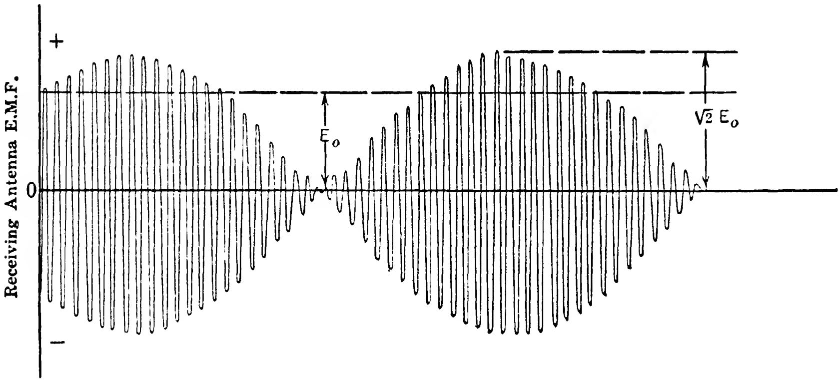


FIG. 29.—A type of modulated current in which the *square of the amplitude* varies as a sine wave.

modulated. Thus, assume that the transmitting antenna current is modulated in such a manner that the difference between the square of its amplitude when modulated and that when not modulated varies with the sine of a uniformly varying angle, so that, if I_0 = amplitude of antenna current with microphone idle, the maximum amplitude will be $\sqrt{2}\ I_0$ and the minimum zero. Then, the curve of the e.m.f. acting upon the receiving antenna will be as represented by Fig. 29, the rectified current will be as represented by curve A, Fig. 30, and will have amplitudes which will vary harmonically, and, therefore, the average current through the telephone receiver, represented by curve B, Fig. 30, will vary harmonically. The result will be the reproduction in the telephone receiver of the tuning-fork note without any change in the quality of the sound.

From this analysis it follows that, if the sound at the receiver is to be similar in quality to that acting at the transmitter, either of the two following conditions must be satisfied:

(*a*) If the receiver circuit *rectifies proportionally to the first power of the voltage impressed upon it*, then the difference between the amplitude of the antenna current with the microphone in operation and that with the microphone idle should vary *in direct proportion* to the pressure of the sound waves on the microphone, or, in symbols

$$I-I_0=kp, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

where I = amplitude of the antenna current with the transmitter in operation;
I_0 = amplitude of the antenna current with the transmitter idle;
k = a constant of proportionality;
p = the pressure of the sound waves upon the microphone.

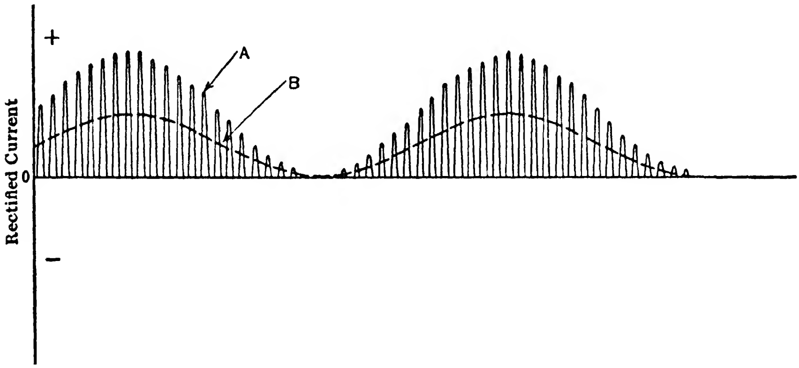


FIG. 30.—Such a current (as that shown in Fig. 29) in the transmitting antenna, with an ordinary type of detector will give in the receiving circuit a rectified current as shown by curves A, the average value of which is curve B, a sine-wave current.

(*b*) If the receiver circuit *rectifies proportionally to the square of the voltage impressed upon it* then the difference between the *square* of the amplitude of the antenna current with the microphone in operation and that with the microphone idle should vary in direct proportion to the pressure of the sound waves on the microphone, or, in symbols:

$$I^2-I_0^2=kp, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

where I, I_0, p have the same significance as in Eq. (15) and k = the constant of proportionality.

Of course, in practice, neither of the two conditions set forth above is fully and entirely satisfied throughout the entire range of pressures impressed upon the microphone diaphragm, and the speech transmission is, therefore, never ideal.

Percentage of Modulation.—The percentage of modulation is expressed by means of the following equation:

$$M=\frac{D_1}{I_0}\times 100, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (17)$$

where I_0 = amplitude of antenna current with microphone idle;
D_1 = difference between I_0 and minimum antenna current amplitude;
M = percentage of modulation.

In the ideal case of a "completely modulated" antenna current $D_1=I_0$ and $M=100$ per cent.

Of course, in designing a radio-phone transmitter, the aim is to make the percentage of modulation as large as possible without, at the same time, interfering with the quality of the transmission.

In view of this the idle resistance of the telephone transmitter and the change in the resistance must be carefully chosen with respect to the resistance of the rest of the system. Thus, considering the simple circuit represented by Fig. 1, p. 781, it is plain that if the idle resistance of the telephone transmitter were much lower than that of the balance of the circuit, any change in the former could not appreciably affect the total resistance and, hence, could not effectively modulate; on the other hand, if the reverse were the case the antenna power radiation would be very small, since the largest percentage of the alternator power output would be absorbed by the microphone. Thus, there must be a best telephone transmitter resistance and this was shown by Seibt to be equal to that of the rest of the antenna circuit.

In many of the systems of radio-telephony the telephone transmitter is not placed directly in the antenna circuit, but, in practically every case, it is so connected that, by speaking into it, an effect is produced which is equivalent to changing the resistance of the antenna circuit; the telephone transmitter resistance may, in these cases, be transferred in the form of an equivalent resistance, to the antenna circuit. Hence in practically every case, whether the telephone transmitter is directly in the antenna circuit or not, it should be observed that the optimum idle resistance of the telephone transmitter is such that, when transferred to the antenna circuit, it should be equal to the rest of the antenna-circuit resistance.

As regards the variation in the telephone transmitter resistance it will be noted that, considering the simple system of Fig. 1, and assuming the idle resistance equal to the balance of the antenna circuit, then in order to obtain 100 per cent modulation with a harmonic sound wave and with a receiver rectifying proportionally to the square of the impressed voltage the resistance of the telephone transmitter should change from

41 per cent of its idle resistance to infinity. For, when it is 41 per cent of the idle resistance the antenna current amplitude would be $\sqrt{2}\ I_0$, and when it is infinity the amplitude of the antenna current would be zero. Of course, it will be easily realized that such extreme variation of resistance is almost impossible of practical accomplishment by means of the microphone; hence 100 per cent modulation by this scheme is impossible.

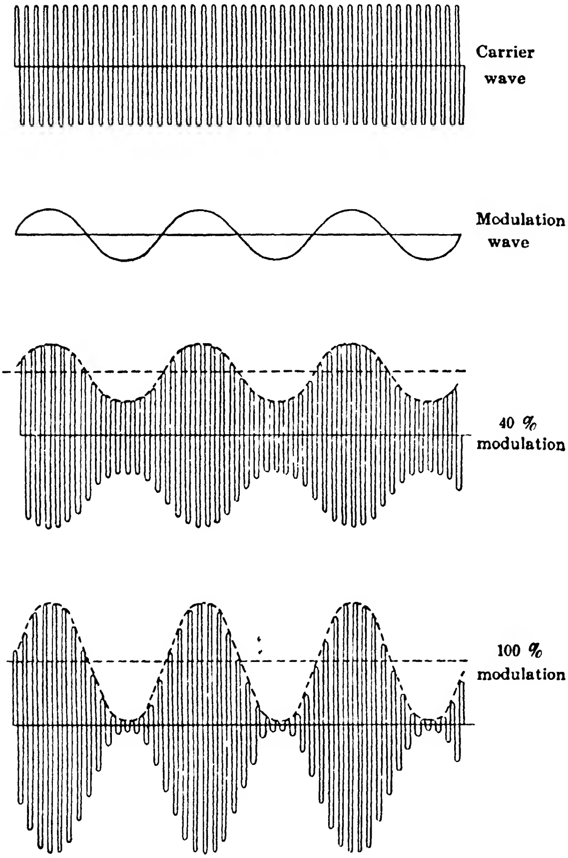


FIG. 31.—This diagram gives the form of antenna current for a 40% modulation and that for 90% modulation.

In Fig. 31 there is shown a 40 per cent modulated wave and a 90 per cent modulated wave. It will be appreciated that in a station designed for even 100 per cent modulation, owing to the great variation in intensity of either voice or music, and owing to the fact that, even at peak value of speech or music intensity, the modulation must not exceed 100 per cent, the average amount of modulation *produces an almost imperceptible variation in amplitude of the carrier wave.*

Measurement of Modulation.—It is evidently important for a station operator to know to what extent his carrier wave is being modulated; some performers talk loudly, close to the microphone, and others talk softly farther away. The microphone response might well be one hundred times as great for the first as for the second, so that one would cause much over-modulation or the other would be much under-modulated.

Fig. 32 shows a scheme for measuring the amount of modulation, by suitable adjustments. Coil L_1 and condenser C make available a voltage of any suitable magnitude for impressing on the grid of the triode. With no current in L_1, and R_1 set at zero, resistance R_2 is adjusted to make the plate current zero, as read on milliammeter A. Now unmodulated carrier is put through L, and of course ammeter A will read. With R_2 left fixed, R_1 is adjusted to make A again read zero and the reading of voltmeter V is taken; call it V_1. Now the desired modulation is put on the carrier and A will again read, because of the increased amplitude of the positive alternations impressed on the grid. Now R_1 is increased until A again reads zero and reading of volt meter V is taken; call it V_2. Then the percentage modulation is given by $M = 100\ (V_2 - V_1)/V_1$.

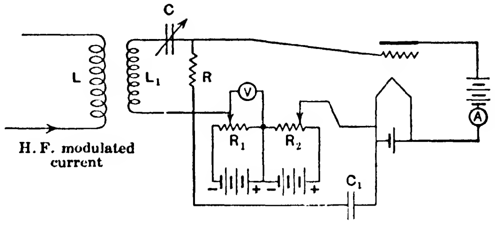


FIG. 32.—The circuit diagram of a simple scheme for measuring modulation.

Evidently it is possible to calibrate ammeter A in percentage modulation, and for station use this is done. With no modulation R_1 is adjusted to make A read zero and R_1 is left at this setting. Now A will give increasing readings as the modulation is increased; its scale may be calibrated directly in per cent modulation.

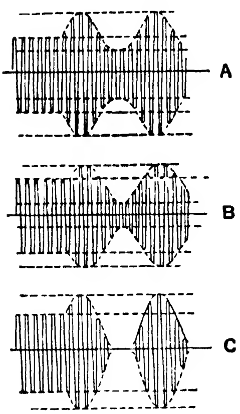


FIG. 33.—Illustrating three types of distortion due to defective modulation.

Distortion in Modulation.—When a sine wave is used to modulate a carrier its amplitude should follow a sine wave, but it frequently does not. The amount of departure from a sinusoidal amplitude variation indicates the amount of distortion present.

Fig. 33 shows three kinds of distortion occurring in modulation. Case A has greater increases in r.f. amplitude than decreases, and case B has the opposite condition; both of these conditions will cause the receiver to pick up audio tones which were not in the original modulation. If, the original modulation frequency was 100 cycles the listener will hear not only the frequency but also a large component of 200-cycle frequency, as well as a series of others.

Case C represents over-modulation; the signal received from such a station will be unusually loud, but of disagreeable quality. Many notes and tones will be heard by the listener which were not in the sound in the studio of the transmitting station.

The Vacuum Tube in Radio-telephony.—Fig. 34 shows an elementary type of circuit which has been very seldom used but which illustrates the principle very well. It consists of the oscillating circuit illustrated by Fig. 164, p. 620, with the telephone transmitter connected directly in series with the antenna. Its principle of operation is exactly the same as that of the simple system illustrated by Fig. 1 except that the tube oscillator has replaced the alternator. This particular tube oscillator circuit is known as the Meissner circuit.

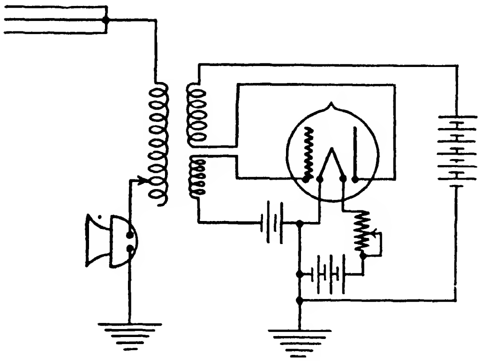


FIG. 34.—Simple circuit for telephone transmitters using a vacuum tube for power; this has been sometimes used in low-power sets.

Fig. 35 illustrates a method of connecting the telephone transmitter in the grid circuit of the oscillator. In this case if the telephone transmitter is idle the amplitude of the antenna current will be constant, but if the transmitter is excited by sound waves a changing current will be pro-

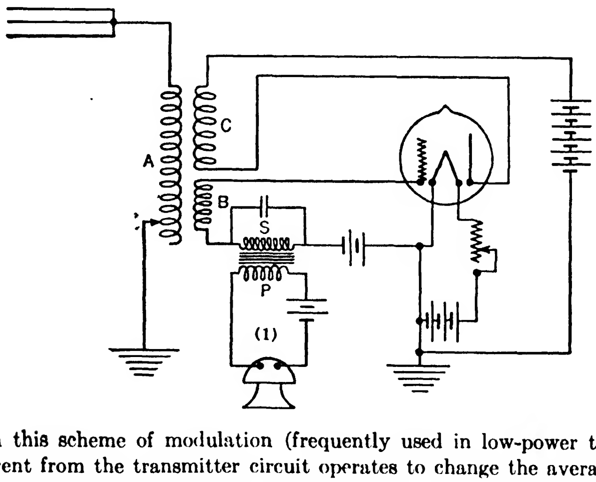


FIG. 35.—In this scheme of modulation (frequently used in low-power transmitters) the current from the transmitter circuit operates to change the average potential of the grid of the oscillating tube, thus effectively modulating the antenna current. The condenser shunting S must have a low reactance for the radio frequency and should have at least 10^5 ohms for the highest voice frequency.

duced in the circuit of (1), which will produce an e.m.f. across the terminals of the coil S, this e.m.f. being a function of the vibrations of the telephone diaphragm. Thus the grid of the oscillator will have impressed upon it not only the high-frequency e.m.f. due to the interactions of the coils A–B, but also the low-frequency e.m.f. due to the speech; the effect of this low-frequency e.m.f. is to increase or decrease the grid potential above or below what it would otherwise be if the transmitter were idle; and since the grid potential reacts upon the antenna current by means of the tube and the coils C–A, it is plain that the amplitude of the antenna current will, instead of being constant, be changed in accordance with the e.m.f. of S, or of the vibrations of the telephone diaphragm. In this type of connection it is plain that the telephone transmitter may be of very low power capacity; this is such an evident advantage that practically all of the modern radiophone tube systems have their telephone transmitters connected in some such way to the grid of a tube.

The above two combinations of telephone transmitter and tube oscillator are typical, in so far as, while they have been shown for a certain type of oscillator circuit, they may be applied in an exactly similar manner to any type of tube oscillator.

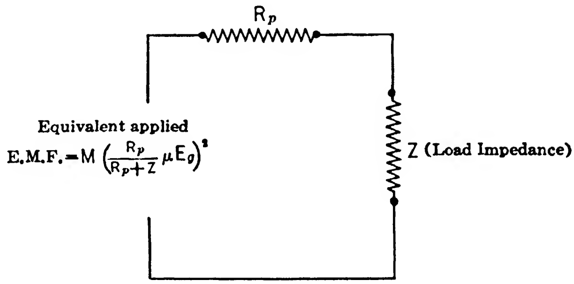


FIG. 36.—Simplified circuit for a triode amplifier.

In analyzing the action of the circuit of Fig. 35 we notice that the voltage impressed on the grid of the triode is expressible as

$$e = E_m \sin \omega t + E'_m \sin \omega_1 t, \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

in which E_m = max. value of high-frequency voltage;
E'_m = max. value of low-frequency voltage;
ω = angular velocity of high-frequency voltage, fixed by the period of the oscillatory circuit (the antenna);
ω_1 = angular velocity of low-frequency voltage, from microphone.

The tube can then be treated as the simple circuit[1] of Fig. 36.

The circuit can then be treated just as an amplifier tube, in which the grid is eliminated from consideration, and a voltage μE_g is supposed to be impressed in the plate circuit; for the tube of Fig. 35, however, we must

[1] See "The Equivalent Circuit of the Vacuum Tube Modulator," by J. R. Carson Proc. I.R.E., June, 1921.

suppose the voltage impressed in the plate circuit is equal to (see Carson's paper)

$$E = M\left(\frac{R_p}{R_p+Z}\mu E_0\right)^2, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (19)$$

in which M = the modulation factor $= (1/2R_p)\partial R_p/\partial E_p$;
R_p = alternating current resistance of plate circuit $= \partial E_p/\partial I_p$;
Z = impedance of the load circuit.

By applying this idea to the case of a grid e.m.f. of the form given by Eq. (18) Carson shows that if the load circuit impedance Z is very low for

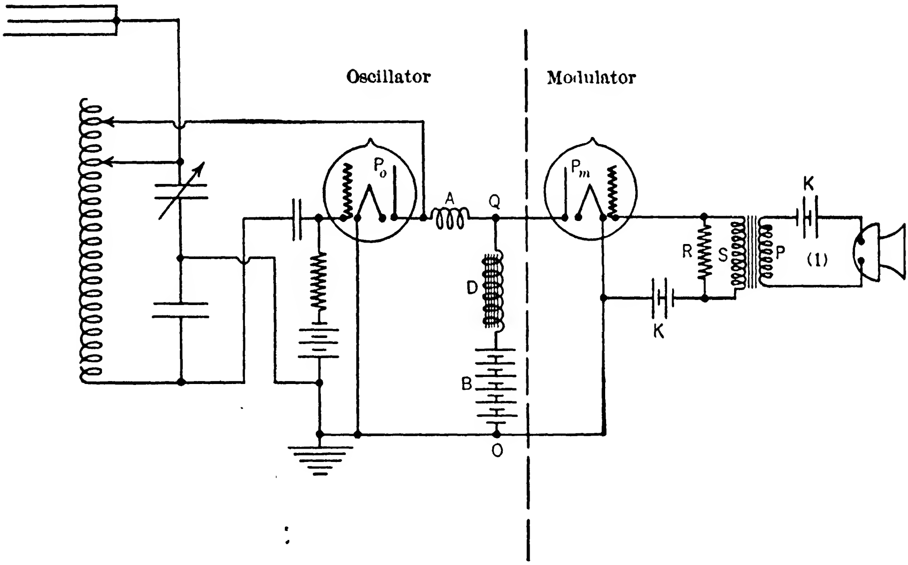


FIG. 37.—A scheme of modulation due to Heising in which a separate tube is used to accomplish modulation; the scheme has been extensively used in small transmitters.

the voice frequencies then the current in the plate circuit of the tube will have the frequencies $\omega/2\pi$, $\omega-\omega_1/2\pi$ and $\omega+\omega_1/2\pi$. Such a current is called a modulated current and is analyzed more fully later in this chapter.

If the load circuit furthermore is tuned, giving a resistance only ($Z=R$) then the modulated power is a maximum when $R=R_p/3$.

Plate Circuit Reactance Modulation.—We will now discuss another type of radiophone tube connection apparently due to Heising, wherein two tubes, or two sets of tubes, must be used; it is called the "plate control" or "constant current" system and is arranged as illustrated in Fig. 37. In this system the part to the left of the dotted line represents the oscillator circuit, which has been discussed on p. 659, Chapter

VI, with the exception of the coil D, which is merely a high inductance choke. When the telephone transmitter is not operative the potential difference across the points Q and O is constant, and hence the amplitude of the high-frequency antenna current, as well as the plate current for the modulator tube, is constant. However, if the telephone transmitter is spoken into, e.m.f.s are induced in the coil S, which change the potential of the modulator grid in accordance with the vibrations of the microphone; this changes the plate current of the modulator or the current between the points Q and P_m, this change taking place at speech frequency, or audio frequency. In virtue of this the battery B will be called upon to supply a current varying at audio frequency, which current must flow through the iron-core inductance D; since the impedance of this is very high at audio frequency, it follows that it will cause a large audio-frequency drop of potential over itself, and thus the potential difference between the points Q and O will be varied at audio frequency and in accordance with the vibrations of the telephone transmitter. Again, since the potential difference impressed upon the plate of the oscillator (i.e., that across Q and O) is being varied, it finally follows that the amplitude of the antenna current will thereby be varied, since the amplitude of the antenna current increases with increase of the plate voltage. Thus, the vibrations of the telephone transmitter are finally reproduced in the antenna as variations in the amplitude of the antenna current or, in other words, the antenna current is thereby modulated.

The function of the coil D may be more clearly seen if the coil were assumed to be short-circuited. Under these conditions, no matter how much the modulator plate current were caused to vary by the action of the transmitter, the potential difference across the points Q and O would remain constant, and no change would be effected in the amplitude of antenna current.

The function of the choke coil A, which should be an air-core coil, is to prevent the plate circuit of the modulator tube from taking from the antenna circuit any of the high-frequency power which the oscillator tube is supplying to it; the proper amount of inductance for coil A depends upon the types of tubes used, but, in general, its reactance should be considerably greater than the plate-circuit resistance of the modulator tube.[1]

Analysis of Heising Scheme of Modulation.[2]—This scheme of modulation is probably better than any other so far suggested, and we are therefore giving a more complete analysis of its operation.

Let us first suppose that the coil D, Fig. 37, has so much reactance

[1] See "Modulation in Radio Telephony," by R. A. Heising, Proc. I.R.E., Aug., 1921; and "Radio Telephone Circuits and Modulation," by W. A. MacDonald, Radio Review, Aug., 1921.

[2] See Radio Review, Feb., 1922, p. 110.

that *no appreciable change* of current through it occurs due to the action of the microphone. We will assume, as has been done before, that the microphone is actuated by a sine wave of sound, and furthermore, that the sine wave of sound gives a sine wave of e.m.f. across the secondary terminals of the transformer S. (In order that the possible variation in the impedance of the grid-filament circuit of the modulating tube may not produce distortion of the terminal voltage of the transformer secondary, a high resistance R of constant value is permanently connected across the secondary to give the load circuit of the transformer an essentially constant impedance.) The potential variations of the modulator grid will cause its plate current to pass through sinusoidal variations, and will thus make the plate circuit of the modulator behave like a variable resistance connected across the points Q and O in multiple with the plate circuit of the oscillator tube. This is schematically indicated in Fig. 38 where R_{mod} represents the variable resistance of the modulator plate circuit and R_{osc} represents the resistance of the oscillator plate circuit.

Let I_{mod} = current in plate circuit of modulator; I_{osc} = current in plate circuit of oscillator; I_b = current supplied by the plate battery.

If we now suppose that I_{mod} is due to the vibrations of the microphone diaphragm, caused to change from zero to twice its average value, then, since we have assumed that the coil L has such reactance as to keep I_b essentially constant, it follows that the current I_{osc} must increase and decrease about its average value to the same extent as does I_{mod}. Of course, as the value of I_{osc} is changed in response to the vibrations of the microphone diaphragm, the power given to the antenna in the form of high-frequency current must be changed and so must the amplitude of the antenna current; in other words, modulation of the antenna current is made to take place.

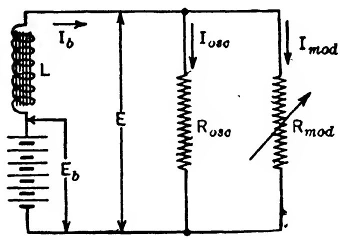


FIG. 38.—Simple representation of the Heising scheme of modulation.

The variations of some of the quantities involved in this scheme of modulation are represented in Fig. 39, where the various curves are self-explanatory; the current I_{osc} for any instant is obtained by subtracting from the essentially constant I_b the value of I_{mod} at that instant.

Now, if we investigate the variation in the power supplied to the oscillator, it will be noted, by referring to Fig. 38, that, since R_{osc} is a constant resistance the current through which is changing from zero to twice its average value, then the power expended in R_{osc} must vary from zero to *four times its average value*. But since the power expended in R_{osc} is equal to the current multiplied by the voltage across it, it follows that not only must the current, I_{osc}, vary from zero to twice its average value, but the voltage across it must also do the same; that is, the voltage across the points Q-O, Fig. 37, must vary from zero to twice its average value.

This result would seem to be contradictory to our assumption previously made that the current I_b is constant; for if I_b is constant there can be no change whatever in the voltage across Q and O. But, as a matter of fact, I_b does vary, though the amount of this variation may be small if the inductance of the coil D (Fig. 37) is large; thus it might easily be that a variation in I_b of only 20 per cent at the modulating frequency would cause the voltage across Q and O to change from nearly zero to double its average value. In some actual radiophone sets employing this circuit we have the following:

Average value of $I_b = 0.08$ ampere.
Inductance of coil $D = 2$ henries.
Voltage of plate battery $= 300$.

If, then, a maximum variation in I_b of 20 per cent should take place at a modulating frequency of 1000 cycles per second we would have:

Maximum voltage drop over

$$D = 2\pi \times 1000 \times 2 \times 0.2 \times 0.08 = 200$$

Hence the voltage across Q and O would vary from $300-200$ to $300+200$ or from 100 to 500.

The circuit shown in Fig. 38 is not exactly equivalent to the actual tube circuit, because in this the value of R_{osc} does not remain constant, but decreases as the voltage impressed on the tube is increased. (See Fig. 61, Chapter VI.) The result of this is that the variation of the power given

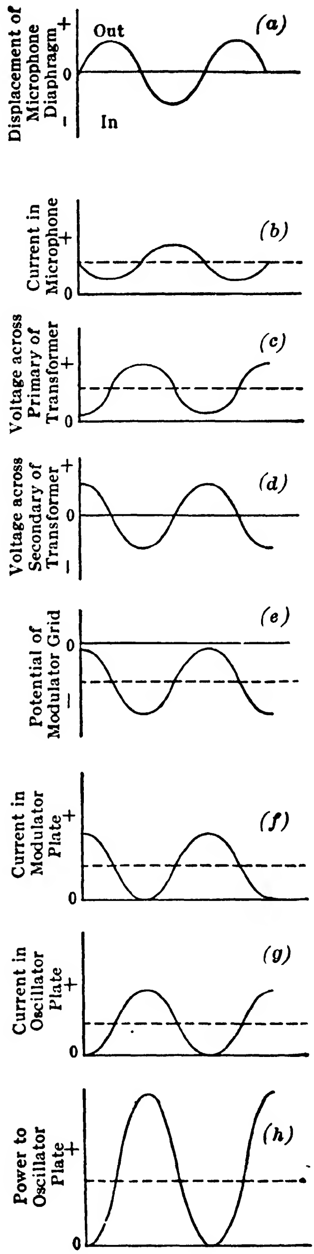


FIG. 39.—Analysis of the action of the Heising scheme of modulation.

to the oscillator is less than from zero to four times the average; but in all cases, however, the power variation is greater than from zero to twice the average.

If the power input to the antenna in the form of high-frequency current is a constant fraction of the power given to the oscillator plate, i.e., if the efficiency of the tube as a d.c.-a.c. converter is assumed constant, the power supplied to the antenna would vary about as shown in curve (*h*) of Fig. 39 and the amplitude of antenna current would vary as the square root of the amplitude of this power curve.

Actual Arrangement of Triode Transmitters.—The simple radio-phone circuits using tubes, and discussed above, are only a few of the very large

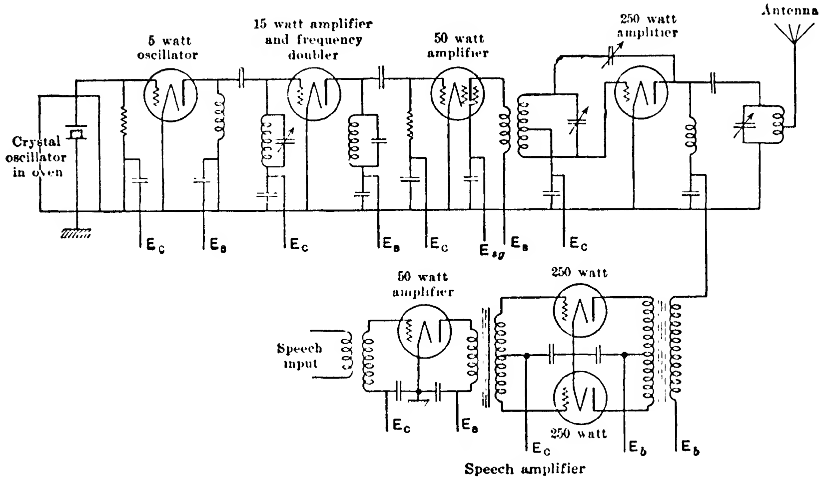


FIG. 40.—A typical 250-watt radio-phone transmitting circuit; notice that the modulator tubes have twice the power rating of the oscillator tube, that is, the 250-watt triode feeding the high frequency power to the antenna.

number of tube systems used in radio-telephony, but they are typical of such systems, and if the reader fully understands these three he will have no difficulty in grasping any other system. It must be remembered that every such system must, to begin with, have an oscillator to produce high-frequency currents in the antenna, and, in addition, it must have some means of changing the amplitude of the antenna current in accordance with sound waves of the voice; this may be done, as has already been shown, by placing the telephone transmitter directly in series with the antenna or in the grid circuit of the oscillator tube, or, again, in the grid circuit of an additional tube, known as modulator, and which amplifies the effects of the telephone transmitter.

In Fig. 40 is shown a simple 250-watt telephone transmitting set for ship use. A crystal oscillator is used to establish the frequency; this frequency is doubled and amplified by a 15-watt triode which in turn serves to excite a 50-watt tube. The output of this excites the 250-watt triode which supplies the antenna power. Connecting the plate and tuned grid circuit of this tube is a small *balancing condenser* which serves to prevent spurious oscillations in this triode circuit. The modulation of this set is *plate-circuit modulation.* The plate voltage of the 250-watt triode is made up of two components, the steady component E_b given by the plate power supply, and that given by the speech amplifier. It will be noticed that in this circuit twice as much triode power is used for modulation as is used for generating the carrier frequency power.

A 30-meter, 500-watt broadcasting transmitter is shown in Fig. 41. This outfit uses push-pull stages throughout the radio-frequency circuit.

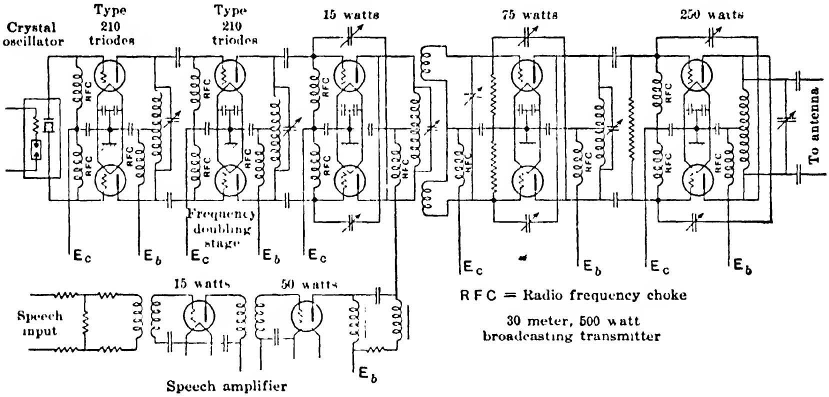


FIG. 41.—A short wave radio-phone circuit of 500-watt rating. In this, push-pull stages are used throughout; modulation is effected at the 15-watt stage, by a 50-watt triode.

The last three stages, 15-watt, 75-watt, and 250-watt triodes, all use the balanced circuit (explained in Chapter X under " Neutralization "). In this transmitter the modulation is carried out in the 15-watt stage; the two 15-watt r.f. amplifier tubes are plate modulated by one 50-watt triode. Here also it is seen that the modulator tubes are of greater power than the r.f. tubes. In both of these diagrams it will be appreciated that the terminals marked E_c and E_b do not all use the same voltage; with increasing power of triode the E_c and E_b voltages both increase. In Fig. 40 for the 5-watt tube E_c and E_b might be 20 and 200 volts, respectively; for the 250-watt tube they would probably be 500 and 2500 volts.

Analysis of Modulated Wave.—It is important at this point to note that, though the alternator, or any other source that might be used at the transmitting station, produces, when no modulation is taking place, an undamped current of constant amplitude and single frequency, except for any harmonics that might be present, yet, when modulation takes place, the current flowing through the transmitting antenna may be shown to be equivalent to a number of component harmonic currents of different amplitudes and frequencies. Thus, consider the simple case illustrated by Fig. 3 of p. 783, which represents a harmonic current, harmonically modulated.

Let I_0 = amplitude of unmodulated antenna current;
ω = angular velocity of unmodulated antenna current in radians per second;
a = instantaneous value of unmodulated antenna current.

Let the equation of this current be

$$a = I_0 \sin \omega t, \quad \ldots \ldots \quad (20)$$

when modulation takes place the amplitude of the current is varying between the maximum of $(I_0+I'_0)$ and the minimum of $(I_0-I'_0)$, and this variation takes place harmonically.

Let ω_1 = angular velocity in radians per second of modulating disturbance, or angular velocity corresponding to the cycle represented by A_1-A_2 in Fig. 3;
i = instantaneous value of modulated antenna current.

Then, the equation of i will be:

$$i = (I_0 + I'_0 \cos \omega_1 t) \sin \omega t. \quad \ldots \ldots \quad (21)$$

Eq. (21) is similar to Eq. (20) except that the amplitude of the current is now $(I_0+I'_0 \cos \omega_1 t)$, instead of just I_0.

Eq. (21) may be changed as follows:

$$i = I_0 \sin \omega t + I'_0 \sin \omega t \cos \omega_1 t$$

$$= I_0 \sin \omega t + \frac{I'_0}{2} \sin \omega t \cos \omega_1 t + \frac{I'_0}{2} \sin \omega t \cos \omega_1 t.$$

And, adding and subtracting $(I'_0/2) \cos \omega t \sin \omega_1 t$, we have:

$$i = I_0 \sin \omega t + \frac{I'_0}{2} \sin \omega t \cos \omega t_1 + \frac{I'_0}{2} \cos \omega t \sin \omega_1 t$$

$$+ \frac{I'_0}{2} \sin \omega t \cos \omega_1 t - \frac{I'_0}{2} \cos \omega t \sin \omega_1 t,$$

or

$$i = I_0 \sin \omega t + \frac{I'_0}{2} \sin (\omega + \omega_1) t + \frac{I'_0}{2} \sin (\omega - \omega_1) t. \quad \ldots \quad (22)$$

Or, letting

$$f = \text{frequency corresponding to } \omega;$$

$$f_1 = \text{frequency corresponding to } \omega_1,$$

$$i = I_0 \sin 2\pi ft + \frac{I'_0}{2} \sin 2\pi(f+f_1)t + \frac{I'_0}{2} \sin 2\pi(f-f_1)t, \quad . \quad . \quad (23)$$

which last equation shows that the harmonically modulated current of Fig. 3 is made up of three component harmonic currents of the following amplitudes and frequencies:

	Amplitude	Frequency
Component No. 1.....	I_0	f
Component No. 2.....	$I'_0/2$	$(f+f_1)$
Component No. 3.....	$I'_0/2$	$(f-f_1)$

Thus, if

$$f = 300{,}000$$

and

$$f_1 = 1000,$$

then the three frequencies will be:

300,000, 301,000, 299,000,

which means a difference between the smallest and largest frequencies of 2000 cycles or about two-thirds of 1 per cent of the frequency of the unmodulated wave (carrier wave). On the other hand if:

$$f = 20{,}000 \ (\lambda = 15{,}000 \text{ meters})$$

and

$$f_1 = 1000,$$

then the three frequencies would be:

20,000, 21,000, 19,000,

which means a difference between the smallest and largest frequencies of about 10 per cent of the frequency of unmodulated wave.

Of course, a speech-modulated current is made up of currents of a very large number of frequencies, one of which is the frequency of the carrier wave, f, and the others are

$$(f+f_1),\ (f-f_1), (f+f_2),\ (f-f_2),\ (f+f_3),\ (f-f_3)\ \ldots\ldots\ (f+f_n),\ (f-f_n),$$

where $f_1, f_2, f_3, \ldots\ldots f_n$ = frequencies included in the human voice.

The two groups of frequencies extending on either side of the carrier wave frequency are known as *side bands*. The group with a frequency higher than the carrier is called the *upper side band* and the other is called the *lower side band.* It will be noticed that the side bands extend on either side of the carrier frequency a number of cycles equal to that of the modulating microphone. Thus a voice modulated carrier would give side bands extending to 6000 cycles or more on either side of the carrier.

It is not necessary that both side bands be sent out from the broadcasting station for successful communication; one band is sufficient. Neither is it necessary to transmit the unmodulated component of the carrier current, but in this case the carrier current, in proper frequency and amplitude, must be supplied to the receiving set to make intelligible speech out of the single side band received. Single side band transmission is evidently less likely to produce interference than when both bands are sent out.

The larger the frequencies f_1 or f_2 or f_3 . . . . or f_n and the smaller the frequency f, the larger becomes the difference between the smallest and largest frequency expressed as a percentage of the carrier frequency.

This analysis leads us to the following conclusions:

Since the current in the receiving antenna is to be a reproduction of that in the transmitting antenna, it follows that the receiving antenna current must have the same frequencies as the transmitting antenna current. Thus, we at once conclude that the receiving antenna circuit must not be sharply tuned to any one frequency, to the partial or entire exclusion of all the others, but must be so designed as to be able to pick up all these various frequencies equally well. This means that, if the difference between the maximum and minimum frequencies, expressed as a percentage of the carrier frequency, is very large, the tuning of the receiving circuit must be *broad*, in order for it to respond equally well to a wide range of frequencies.

Again, in order for it to be possible to use a sharply tuned receiving circuit, such as may be obtained by the circuit discussed on p. 638, Chapter VI, the frequency of the carrier wave must be very high, that is, of the order of 500 kc. or more, corresponding to a wave length of 600 meters or less.

Furthermore, it would seem as if, for a receiving circuit having a certain degree of sharpness of tuning, a high-pitched voice would be less distinct than a low-pitched one; this effect is very noticeable if the proper adjustment of the receiver circuit is made.

The effect noted above is well illustrated when listening to radio-telephone transmission on long wave lengths, say 20,000 meters; using an amplifying circuit such as shown in Fig. 182, p. 635, the tuning characteristics of which are given in Fig. 185, p. 639, the speech is very drummy, only the low vowel sounds coming through. It is quite possible to adjust the receiving circuit to such sharp resonance that the speech is unintelligible, although very loud; decreasing the coupling of the tickler coil will decrease the sharpness of resonance of the receiving circuit, making the resistance of the circuit higher. This will, of course, decrease the strength of the received signal, but at the same time will improve the quality.

In a radio-telephone outfit, both the receiving and transmitting sets of which have been properly adjusted, the speech transmission is much better than that over the average wire line; due to the fact that all frequencies are attenuated alike (whereas in wire speech the high-frequency currents attenuate much more than the lower), the enunciation of the received signal is so distinct that the voice of the operator talking at the transmitting station may be easily recognized.

Power of Modulated Current.[1]—In case we have an unmodulated current, as represented by Eq. (20) we see that the radiation is expressed by

$$\text{Average power} = \frac{I_0^2}{2} R, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (24)$$

where R = radiation resistance of the transmitting antenna.

To obtain the average power of the modulated current, as represented by Eq. (21), we must integrate this equation over the high-frequency period $2\pi/\omega$ and also over the low-frequency period $2\pi/\omega_1$. Thus the average power of the modulated current becomes

$$P = \frac{R}{\frac{4\pi^2}{\omega\omega_1}} \int_0^{\frac{2\pi}{\omega_1}} \int_0^{\frac{2\pi}{\omega}} ((I_0 + I'_0 \cos \omega_1 t) \sin \omega t)^2 \, dt \, dt$$

$$= \left(I_0^2 + \frac{(I'_0)^2}{2} \right) \frac{R}{2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (25)$$

In case the antenna current is completely modulated, $I_0 = I'_0$ and then

$$\text{Average power} = \frac{3}{2} \frac{I_0^2 R}{2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (26)$$

Now Eq. (26) compared to Eq. (24) shows that the average power of a completely modulated wave is 50 per cent larger than that of the unmodulated wave. Also the effective value of a completely modulated current is $\sqrt{3/2}$ times as great as that of the unmodulated current.

Since the completely modulated current is made up of three components, as shown on p. 821, and since component No. 1 is the unmodulated current

[1] See "Radio Telephone Modulation," by Brown and Keener, Bull. 148, Eng. Exp. Station, Univ. of Illinois. See also "Modulation in Radio Telephony," by R. A. Heising, Proc. I.R.E., Aug., 1921, and "Relations of Carrier and Side Bands in Radio Transmission," by R. V. L. Hartley, Proc. I.R.E., Feb., 1923.

it follows that the average power as represented by Eq. (26) is distributed as follows:

	Amplitude	Frequency	Average Power
Component No. 1. . .	I_0	f	$I_0^2/2\,R$
Component No. 2. . .	$I_0/2$	$f+f_1$	$I_0^2/8\,R$
Component No. 3. . .	$I_0/2$	$f-f_1$	$I_0^2/8\,R$

Hence even in the case of a completely modulated current, components Nos. 2 and 3, which carry the modulating frequency, represent only $\frac{1}{3}$ of the total radiated power, while component No. 1, i.e., the carrier current, represents $\frac{2}{3}$ of the total radiated power. This means that in a speech modulated current, even if the modulation is complete, $\frac{2}{3}$ of the total power generated and sent out is in the carrier frequency and the balance is equally distributed in the two side bands. Actually the current is modulated, on the average, only a few per cent so it will be appreciated what a small part of the total radiated power really represents the telephone signal. The excess of power in the carrier channel must be available, however, to flow over into the side bands as the intensity of speech or music increases over its average value. In other words the great excess of generally useless power must be available to maintain the *quality* of the transmission in so far as this is determined by inflection and accentuation of the voice, etc.

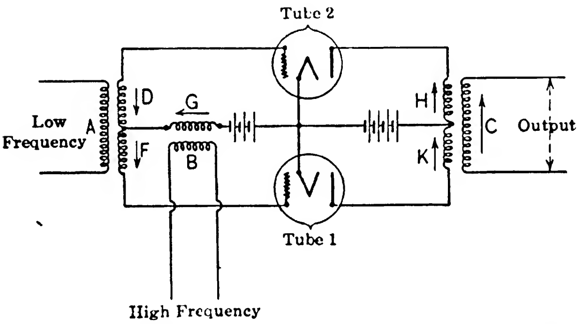


FIG. 42.—In the special arrangement shown here the high-frequency carrier current is modulated in the input and is eliminated in the output circuit; in the output appear only the two side bands.

Elimination of Carrier Frequency.—As just mentioned, the carrier frequency does not carry the signal, but simply serves as a reservoir of energy to supply the demands of the signal for side band energy. After this demand is once supplied there is no evident need of taking the carrier any farther in the transmission scheme. Furthermore one side band, by itself, carries all the characteristics of the signal; no more intelligence is carried by the other band but simply more power. Thus if advisable we might eliminate one side band also. This is done in the high-powered radio-telephone transatlantic channel and is called *single side band transmission.*[1] Its advantages are:

(a) Power saving to the extent of approximately $\frac{5}{6}$ of that required by the carrier system of transmission.
(b) Decrease of the width of the frequency band required for the channel to one-half that required by the ordinary double side band method. This makes it possible to more sharply tune the receiving set.

These advantages are especially important in transoceanic transmission because of the large saving in power and power equipment and further, because at the comparatively low frequencies which are at present being used, it is imperative that each channel shall be as narrow as possible; there are not many channels available.

As an important part in the single side band transmitter we have the *balanced modulator*[2] which is shown in Fig. 42. It will be noted that the low frequency e.m.f.s across D and F modulate the high-frequency e.m.f. across G.

Let $E'_0 \cos \omega_1 t =$ e.m.f. across D or F;
$E_0 \sin \omega t =$ e.m.f. across G.

In the grid circuit of tube 1 the e.m.f.s of G and F add to each other while in the grid circuit of tube 2 the e.m.f.s of G and D oppose each other. For this reason we have
EMF across C due to current in K

$$= S(E_0 \sin \omega t + E'_0 \cos \omega_1 t \sin \omega t). \quad (27)$$

EMF across C due to current in H

$$= S(-E_0 \sin \omega t + E'_0 \cos \omega_1 t \sin \omega t), \quad (28)$$

where S is simply a proportionality factor.

The resultant e.m.f. across C

$$= S(E'_0 \cos \omega_1 t \sin \omega t + E'_0 \cos \omega_1 t \sin \omega t). \quad (29)$$

This equation can be changed to the form

$$\text{EMF across } C = SE'_0 \left[\sin (\omega+\omega_1)t + \sin (\omega-\omega_1)t\right]. \quad (30)$$

Thus the e.m.f. across the output of this balanced modulator contains the two side bands but no carrier current.

[1] See "Transatlantic Radio Telephony," by Arnold and Espenschied in Journal A.I.E.E., Aug., 1923, and "Production of Single Side Band for Transatlantic Radio Telephony," by R. A. Heising, Proc. I.R.E., Vol. 13, No. 3, June, 1925.

[2] See "Carrier Current Telephony and Telegraphy," by Colpitts and Blackwell, Journal A.I.E.E., Feb., 1921.

A suitable band filter interposed in the transmission layout, after C, will filter out either side band leaving the other to be amplified by means of a proper power amplifier to be finally conveyed to the transmitting antenna.

The antenna current (at both the transmitting and receiving ends) has a frequency of $\omega+\omega_1/2\pi$ or $\omega-\omega_1/2\pi$ as the case may be. Let us assume that the frequency $\omega+\omega_1/2\pi$ is the one selected for transmission. In order that the low-frequency signal may be made audible at the receiving end a heterodyne receiver must be used, having in its grid circuit not only the incoming single side band signal of frequency $\omega+\omega_1/2\pi$ but another frequency of $\omega/2\pi$, which latter is obtained from a local oscillator. The telephone in the plate circuit of the detector will then respond to the difference of the two, that is, $(\omega + \omega_1/2\pi) - (\omega/2\pi) = \omega_1/2\pi$ which, of course, is the signal frequency.

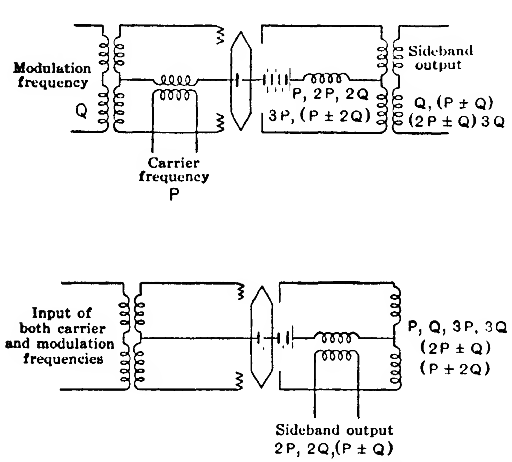


FIG. 43.—Two modulation schemes, and the frequencies they make available.

The scheme indicated in Fig. 42 evidently uses grid modulation, that is, the carrier frequency and modulation frequency first act together in the input circuit of the triode. Peterson and Keith[1] have investigated the possibilities of this type of modulation and they concluded that it is a much more effective scheme of modulation than is the plate scheme. It is a fact, however, that plate-circuit modulation is in nearly universal use at present. In Fig. 43 are shown two diagrams taken from their paper, giving the frequencies which are obtainable from the two branches of the output circuit of a push-pull amplifier.

Single Side Band Transmission.—It is evident that by the use of a push-pull amplifier properly connected, the carrier can be taken out of the output, leaving only the side bands plus any distortion frequencies that occur; band pass filters permit the selection of one side band from the other. Thus a 58.5-kc. telephone channel sends a side band including only the frequencies from 58.9 kc. to 61.1 kc.

This selection of one side band is carried out at a comparatively low power level in the train of power-amplifying tubes, so that the higher

[1] B.S.T.J., Jan., 1928.

power tubes carry one side band only. This permits the handling with given tubes, of 5 to 10 times the amount of useful power as if the ordinary method (side bands plus carrier) were used.

This scheme of side-band separation is applicable only on long-wave channels; the three short-wave present transatlantic channels therefore send carrier and both side bands.

At the receiving station the carrier frequency must be put back in the right magnitude and frequency. Its magnitude should be the same as the carrier would have had if it had been transmitted with the side band across the Atlantic.

Its frequency must be within about 20 cycles of the original carrier or disagreeable speech will result for reasons now to be analyzed. Suppose a 58.5 carrier and a 400-cycle note for modulation. The resultant frequency in the upper side band is 58.9 kc. Now suppose the 400-cycle note had second and third harmonics, of 800 and 1200 cycles. These would give side-band frequencies of 59.3 and 59.7 kc. (of course there would be others in the lower side band, which we are neglecting).

When the signal is detected (or *demodulated*, as the expression goes) the audio frequencies which appear are equal to the side-band frequencies minus the carrier frequency. Now suppose the carrier introduced into the receiver is 58.6 kc. (instead of 58.5, which it should be). The audio frequencies which come from the detector are $58.9-58.6=300$, $59.3-58.6=700$, and $59.7-58.6=1100$. But these notes, 300, 700, and 1100, have no harmonic relation, so the note will not only have the wrong pitch (300 instead of 400) but its quality will be peculiar because of the presence of the non-harmonic frequencies.

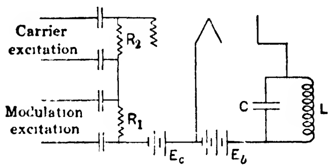


FIG. 44.—A laboratory circuit for studying modulation.

An interesting phenomenon occurs if the introduced carrier is put in with a frequency higher than the side band. After detection such a combination yields speech or music in which what were originally high notes are now low notes, and vice versa. This is called *inverted speech*.

To make the single side band radio channel successful the frequency of the transmitter, and of the oscillator used at the receiving end of the channel to supply the carrier, must be constant to within 0.015 per cent.

Experimental Analysis of Modulation.—The actual circuit arrangement by which modulation can be carried out has not yet been analyzed; the schemes of Fig. 1, p. 781, and Fig. 37, p. 814, actually do work, but nothing has been said about the circuit conditions required for modulation. To bring out the requisite conditions we first analyze the diagram of Fig. 44. Suppose the triode, with its carrier frequency impressed across R_2, is operating as a class A amplifier (see p. 983), and the circuit L-C (so called *tank circuit*) is tuned for the carrier frequency. Now we try to modulate the output by putting modulating voltage across R_1, thus impressing on the grid both carrier and modulation voltage. As long as the grid of the triode is forcing the plate current to vary only over the straight part of the E_g–I_p curve (the condition required for a class A amplifier), *the carrier frequency current in the tank circuit will not be modulated at all.* The current in the plate circuit, in the triode, will have the form shown by curve A of Fig. 45; the carrier frequency current is merely superimposed on the modulation frequency current. The current through condenser C_1 which would of course carry the modulated frequency current if there were any such, carries only the current shown in B of Fig. 45, merely the unmodulated carrier current. A class A amplifier cannot be modulated by putting the modulation voltage in either grid or plate circuit.

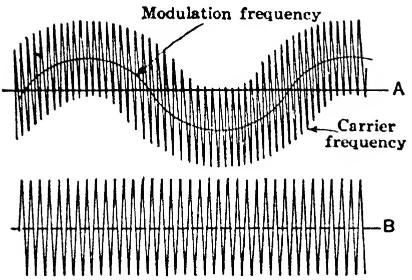


FIG. 45.—Modulation frequency and carrier frequency are both present, but the carrier frequency is *not modulated*.

Now if the grid of the triode of Fig. 44 is biased sufficiently, the plate current will fall to practically zero when the grid is unexcited. The current which flows between plate and filament when the carrier voltage is impressed on the grid is as shown in curve A of Fig. 46, and when the grid is excited by the modulation voltage in addition the current from filament to plate has the appearance of curve B of Fig. 46. This gives a r.f. current of varying amplitude entirely different from the result shown in A of Fig. 45. The current through the L–C circuit (which is tuned for the carrier) will now have the form shown in curve A of Fig. 47, in the inductance, and in the condenser, current of the form shown in curve B.

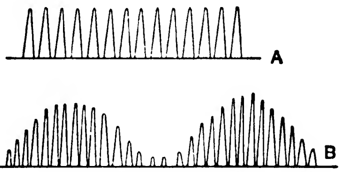


FIG. 46.—If the radio frequency triode is sufficiently biased before the modulation voltage is impressed, then the plate current will be of the form shown in curve B.

Now if another circuit, also tuned to the carrier, is supplied with power by magnetic coupling with the L–C circuit, the current in it will resemble that of curve B, Fig. 47, so far as amplitude variation is concerned, but the individual cycles of the current will approach more closely sine form. The individual cycles in condenser C (curve B of Fig. 47) are by no means sine waves, their negative loops not being of the same form as the positive loops.

In case the bias of the grid of Fig. 44 is increased still more, to make a class C amplifier, the current between plate and filament will have the form shown in A of Fig. 48, and evidently the current obtained in the condenser

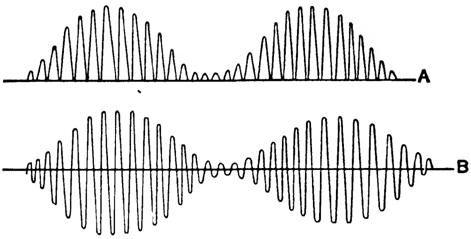


FIG. 47.—The pulsating, variable amplitude current (B of Fig. 46) will produce in a tank circuit coupled to the plate circuit, a modulated current, as shown in B.

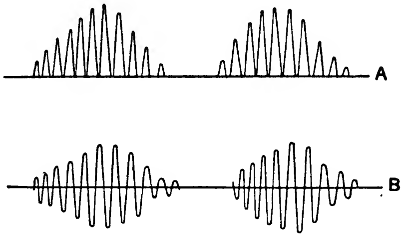


FIG. 48.—In case the grid bias (Fig. 44) is increased too much, the amplitude o the modulation voltage remaining fixed *over modulation* will result.

circuit (B of Fig. 48) does not have an envelope which resembles the sine wave of modulating voltage put on the grid of the tube. The wave is *over-modulated*; in general, this is true of class C amplifiers, and they should not be used for modulated waves as they always distort the modulation, if this is anywhere near 100 per cent.

A type 210 triode was arranged as in Fig. 44 and some of its performance is brought out by the following curve sheets. Fig. 49 shows how the current in the tank circuit (the L-C circuit) varies in amplitude as the amount of grid excitation, of carrier frequency, was varied; different bias was used for the various curves as noted in the diagram.

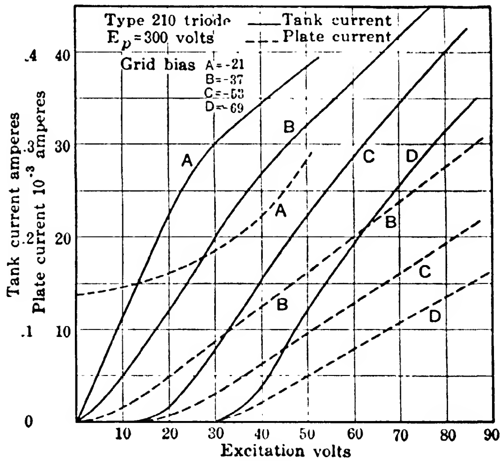


FIG. 49.—Tank circuit current of Fig. 44, as different bias and carrier excitation were used, there being no modulation voltage impressed.

Class A amplifier gives a higher output, for a given excitation, because the excitation voltage produced current in the tank circuit for both positive and negative alternations of the carrier, whereas the B and C amplifier adjustments are such that the triode carries current only during the positive alternations of grid potential.

The tank circuit, in this test, had a resistance of about 22 ohms, the carrier frequency was 15 kc., and the coil had about 4 millihenries inductance. For a tank current of 0.3 ampere the power in the circuit is $0.3^2 \times 22 = 1.98$ watts. The input to the plate circuit for curve A is 5.4 watts, 4.2 watts for curve B, and 3.6 watts for curve D. Curve A then shows a plate-circuit efficiency of 36.5 per cent (class A amplifier); curve B (class B amplifier) shows an efficiency of 47 per cent, and curves C and D (class C amplifiers) show 51 and 55 per cent respectively.

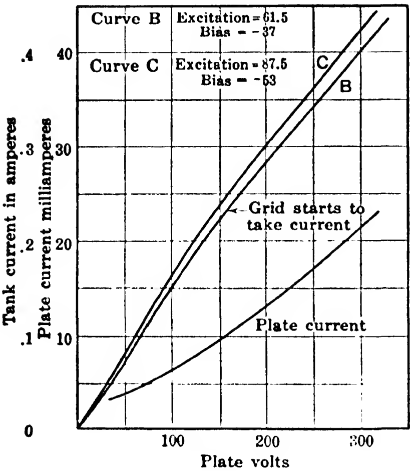


FIG. 50.—Variation of tank current (Fig. 44) as plate circuit voltage was varied; these curves show the feasibility of plate circuit modulation.

In Fig. 50 are the relations between tank current and plate voltage for a class B amplifier (grid bias $= -37$) and a class C amplifier (grid bias $= -53$). The curve of plate current shows that the resistance of this circuit, $\Delta E_p / \Delta I_p$ is about 15,000 ohms, and if a plate modulation scheme is used the transformer employed in the plate circuit to introduce the modulation voltage will be connected to a circuit of 15,000 ohms resistance. Thus if 100 per cent modulation is required, with a plate voltage (average) of 300 the modulation voltage will have to be $300/\sqrt{2} = 212$ volts (effective); and 212 volts working into a resistance of 15,000 ohms will have to supply a power of $212^2 / 15{,}000 = 3$ watts. The unmodulated power in the tank circuit (0.4 ampere in 22 ohms) is 3.6 watts (curve B), and 4 watts for curve C

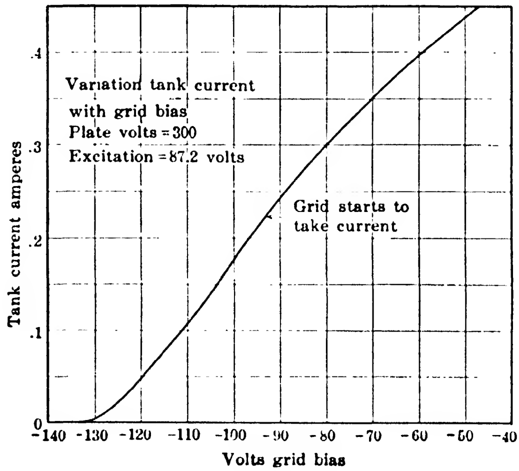


FIG. 51.—Variation of tank current (Fig. 44) as grid bias was altered, with fixed amount of R.F. grid excitation; the nearly linear proportionality shows the feasibility of grid modulation.

(Fig. 50). Thus it requires almost as much power to modulate the tank circuit power as there is power in this circuit.

Holding the plate volts at 300 and excitation voltage at 87.2 the current in the tank circuit varied with grid bias as shown in Fig. 51. This curve shows that if grid modulation is used the variation in amplitude of tank circuit will follow linearly the modulating voltage put on the grid. This shows the feasibility of using grid circuit modulation with this tube, just as the curves of Fig. 50 show the feasibility of using plate circuit modulation.

In Fig. 52 is shown a wave meter response to the tank circuit, this having unmodulated carrier current in it. With the circuit arranged as in Fig. 44, 33 volts of modulation voltage was introduced into the grid circuit. From Fig. 51 it can be seen that with −85 volts bias it requires

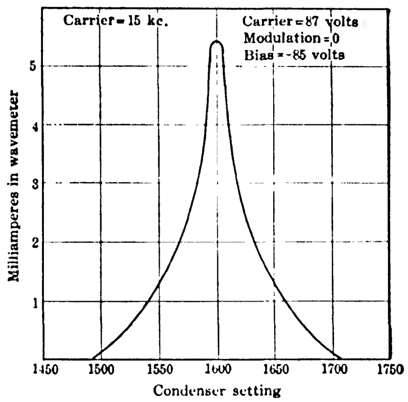


FIG. 52.—Response of a wave meter coupled to the tank circuit of Fig. 44; no modulation voltage was impressed.

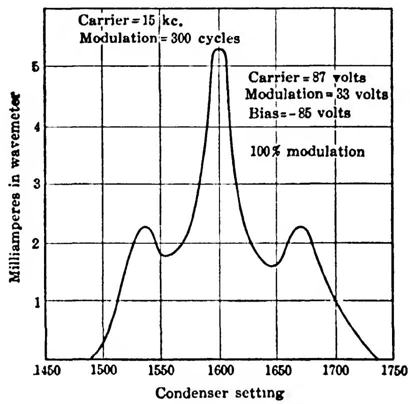


FIG. 53.—Same as Fig. 52, with 100% modulation; the two side frequencies are just 300 cycles away from the carrier frequency, as theory predicts they should be.

an added grid voltage of about 47 volts to bring the tank circuit current to zero. Hence if a modulation voltage of $47/\sqrt{2} = 33$ volts is used the circuit should have 100 per cent modulation. This was done, the modulation frequency being 300 cycles, and the wave meter response was as shown in Fig. 53. The two side bands should be at 15,300 cycles and 14,700 cycles, one 2 per cent higher and one 2 per cent lower than the carrier frequency. This means that the condenser setting of the wave meter for the side bands should be 4 per cent higher and lower than the value giving resonance for the carrier. It can be seen from Fig. 53 that these are the two settings where side band resonance occurred.

From the analysis on p. 821 it is seen that the two side bands should each have an amplitude one half that of the carrier; this is nearly so in Fig. 53

would have large currents set up in it, but the variations in amplitude of these currents would not follow the variations in the impressed e.m.f. amplitude, whereas the high-decrement circuit would have much smaller currents (same L and C supposed as for low-decrement circuit) but the variations in amplitude would more accurately follow those of the impressed modulated e.m.f. Since the voice sounds are conveyed by the *variations in the amplitude* of the current and not by the magnitude of the current itself it is evident that the high-resistance circuit would be the one to use for successful radio-telephony.

Applying this general idea to an actual case of speech transmission, we come to the conclusion that the decrements of the transmitting antenna, the receiving antenna, and closed-tuned circuit at the receiver must all be higher than the highest decrement occurring in the modulated e.m.f. Thus in Fig. 56 the e.m.f. (which is supposed accurately to represent the voice sounds) has its most rapid change in amplitude from A to B; in ten cycles its amplitude decreases in the ratio of 10 : 1, which corresponds to a decrement of 0.24. The decrement of none of the three circuits taking part in the transmission and reception should be as low as this value, if clear well-enunciated speech is expected at the receiving end.

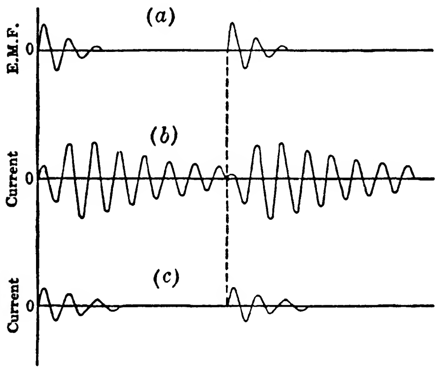


Fig. 58.—A series of damped waves of e.m.f. acting on a low-resistance receiving circuit produce a current as indicated in (*b*), evidently not of the same form as the e.m.f. a high-resistance circuit will have currents as shown in (*c*) which current closely resembles the e.m.f. causing it.

For short-wave work this idea is not of so much importance, because the permissible value of the decrement, from this standpoint, is lower than that generally attained in the construction of sets. Thus, if the time between A and B, Fig. 56, is taken as 0.0001 second and the wave length used is 300 meters, the number of cycles from A to B would be 100; a decrease in amplitude to one-tenth of its initial value in 100 cycles corresponds to a decrement of 0.023, which would be practically never obtained in either of the antenna circuits and could only be obtained in the secondary tuned circuit of the receiving set by having a tickler coupling to the plate circuit.

The same idea can be expressed by saying that the resonance curves of all the circuits receiving a radio-telephone signal should be wide enough to receive equally well all the useful frequencies. Really the proper kind of a resonance curve is one with a flat top, about 20 kc. wide, with very steep sides. Thus both side bands of a transmission will be faithfully reproduced (action of the flat top), and interference from stations in neighboring channels will be eliminated (action of the steep sides).

Multiplex Radio-telephony.—It is possible to carry on, by means of

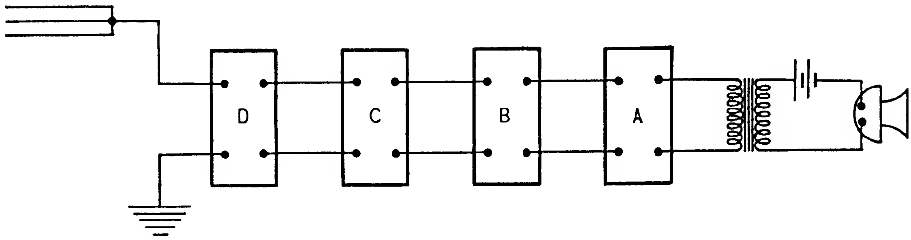


Fig. 59.—Conventional diagram of circuit for sending out a doubly modulated telephone wave.

a scheme of "double modulation," several radiophone conversations in the same area and using exactly the same high-frequency carrier wave for all stations; the extra complications of the scheme are worth while only in regions of congested communication.

The general idea of the scheme is conventionally indicated in Fig. 59, wherein A is a modulator and B is a long wave-oscillator; C is a modu-

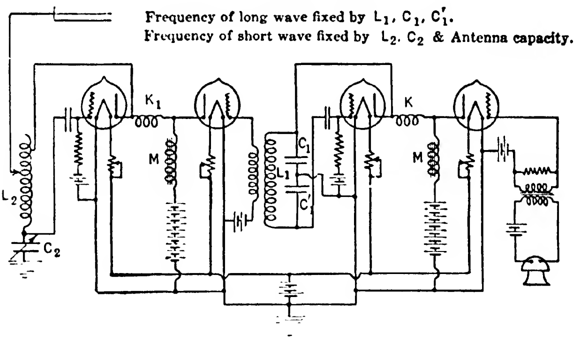


Fig. 60.—Arrangement of four tubes for carrying out the double modulation indicated in Fig. 33.

lator and D is a short wave-oscillator; the connections of a tube-transmitting set utilizing this idea are shown in Fig. 60. From these two diagrams it is evident that the antenna sends out a "doubly modulated" high-frequency wave, that is, the amplitude of the high-frequency wave follows a curve which is a voice-modulated long-wave radio-frequency. Thus generator B, Fig. 59, might generate oscillations of 25,000, and the amplitude of this 25,000-cycle current is voice-modulated by the action of A. This variable amplitude, 25,000-cycle wave, controls, through the action of modulator C, the amplitude of the high-frequency current generated by D, and sent out from the antenna.

Fig. 60 shows how the Heising modulation scheme may be made to function for multiplex transmission, and Fig. 61 shows the general reception scheme for multiplex telephony. The antenna circuit and the closed circuit, L_1–C_1, are tuned to the high frequency generated by the oscillator exciting the transmitting antenna. The action of the grid condenser and leak is to produce in the plate circuit a pulsating current, the form of which is the same as the envelope of the high-frequency wave received by L_1–C_1. This envelope is itself of inaudible frequency, it being perhaps a voice-frequency modulated, 25,000-cycle current. This 25,000-cycle current

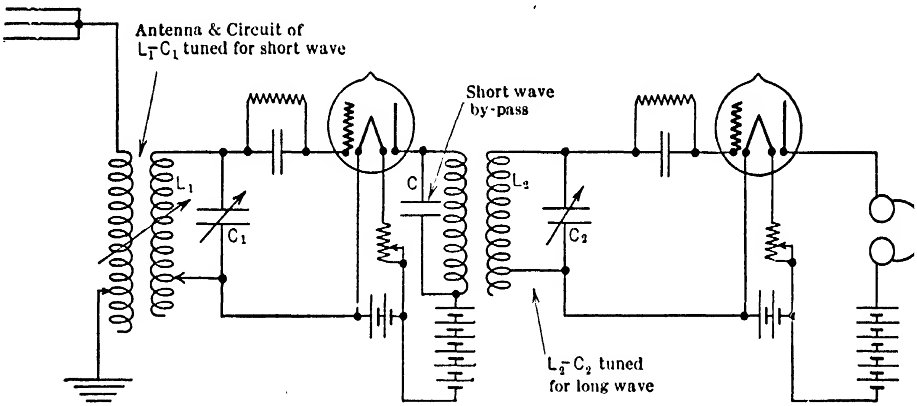


FIG. 61.—Arrangement of the receiving circuit for tuning in on a doubly modulated telephone wave.

acts on the tuned circuit L_2–C_2 coupled to the plate circuit of the first tube. The grid condenser and leak of this second detecting tube act to produce in the plate circuit of this tube (in which the telephones are connected) a pulsating current of the form of the envelope of the 25,000-cycle current. This envelope is, however, of voice frequency, and therefore makes audible the speech carried by the doubly modulated high-frequency wave.

Several stations in the same area might transmit on a carrier frequency of 3,000,000 cycles; one of the stations would send out this wave, modulated by a voice-modulated 25,000-cycle wave, another would use a voice-modulated 35,000-cycle wave, another a voice-modulated 45,000-cycle wave, etc. The selecting of the proper message at the receiving station is done by the tuning of the L_2–C_2 circuit (Fig. 61); as this is tuned to the various long-wave frequencies being used, the conversations from the several transmitting stations become audible. All receiving stations tune their respective antennas and L_1–C_1 circuits to the same high-frequency carrier current.

By using several high-frequency carrier waves, far enough apart in frequency so that no interference is encountered, and using several long wave-modulations of each of these, it might be possible to carry on, in the same area, without serious interference, perhaps fifty different conversations.

Another scheme for carrying on multiplex telephony uses an antenna tuned to several different frequencies and coupled to this antenna the same number of ordinary singly-modulated transmitting sets; it seems that this scheme may be made to work satisfactorily.[1]

Range of a Radio-telephone Transmitter.—In Chapter IV there is given a considerable amount of data on the signal distribution from a broadcast transmitter; with a good receiving set it is of course possible at night to hear a 5-kw. station, on the broadcast frequencies, 1000 miles or more, but the normal daylight range is much more limited, as indicated by the diagrams on pp. 377 and 378. The comparatively short range of a telephone transmitter, contrasted to that of a telegraph transmitter of the same power, is due to the fact that it is only the *variations in carrier amplitude* that transmit the speech, and these variations are on the average only a few per cent or less of the amplitude of the carrier.

The carrier wave of the telephone transmitter has a *nuisance range* of 5 to 10 times the useful range of the station. This carrier, if it is within a few kilocycles of the frequency of another transmitter, will produce in the receivers tuned to this distant station a whistling note which will completely spoil the program of the local station. A station which delivers a satisfactory signal only for listeners not more than 100 miles away from itself may spoil reception for listeners of another station 1000 or more miles away.

Simultaneous Radiophone Transmission and Reception.—Originally only one antenna was used to accomplish this two-way conversation; the operator had to switch the antenna from transmitting to receiving apparatus as the occasion demanded. This caused trouble because both operators might be sending or listening at the same time, thus making communication impossible. In order to overcome this difficulty encountered in the single antenna radiophone set two antennas are used at each station, one for *transmitting only* and the other for *receiving*; each operator can then talk and listen at the same time, as is done in ordinary wire-

[1] See Proc. I.R.E., Vol. VIII, No. 6, for report on the feasibility of such a scheme; article by Ryan, Tolmie, and Bach, entitled "Multiplex Radio Telegraphy and Telephony."

telephony. Attempts have been made to use a single antenna for simultaneous transmission and reception, but the results are not reported to have been very satisfactory. One possible scheme uses in the transmitting circuit two antennas of identical characteristics, one a real antenna and one a dummy; adjustments are so carried out that half the power from the transmitter goes through each antenna. The receiving coil of the receiving circuit is coupled to both antennas equally, so that when transmitting practically no voltage is induced in the receiving circuit. When the distant station is transmitting only the real antenna is excited so that the signal is received all right.[1]

An arrangement wherein two antennas are used is conventionally

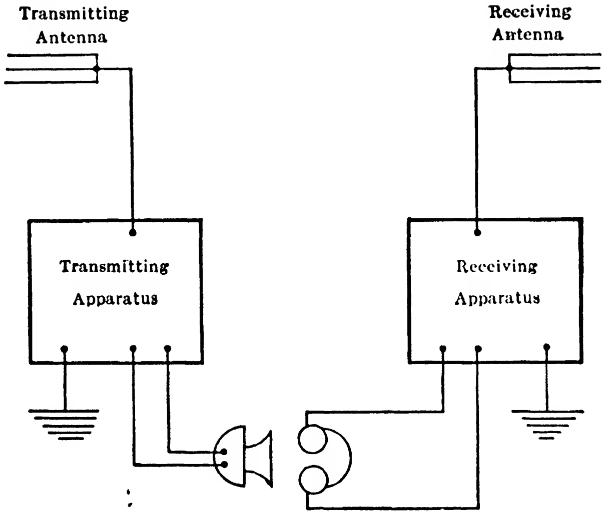


FIG. 62.—Scheme for simultaneous transmission and reception using two antennas, spaced a considerable distance from one another and tuned to different wave lengths.

shown in Fig. 62. The two antennas are put up at some distance from each other and the wave lengths of the two transmitters are made very different from each other. The reader will at once note that in such a scheme the receiving antenna has impressed upon it not only the feeble e.m.f.s due to the distant transmitting antenna, but also the far greater e.m.f.s due to the local transmitting antenna; the latter e.m.f.s are not wanted, in so far as they "swamp" the smaller e.m.f.s due to the distant transmitting antenna and make reception therefrom impossible or, at least, very difficult. In the next section will be discussed methods of avoiding this trouble.

[1] A brief description of such a set is given in the Radio Review, Vol. I, No. 15, by M. B. Sleeper.

Transatlantic Radio-telephony.—It is now possible to telephone to Europe from anywhere in the United States, at practically any time, day or night, throughout the year. There are now four radio-telephone

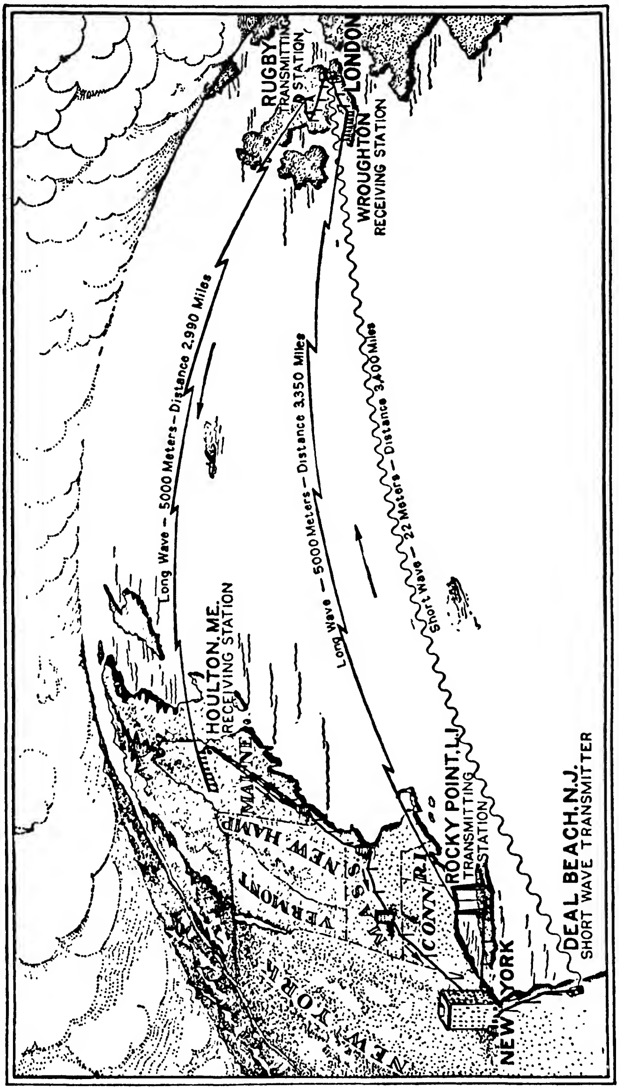


FIG. 63.—A bird's-eye view of the trans-atlantic radio-telephone channel, showing the two antennas at each end, and in addition the short-wave channel for ensuring transmission.

channels from New York to Europe, and one to South America. One of these is a long-wave channel (about 5000 meters) and the others are short-wave (15–50 meters). The long-wave channels will first be described.

In Fig. 63 is shown a bird's-eye view of this channel as it was inaugurated five years ago; since then the receiving station has been moved north from England into Scotland, where the ratio of signal to noise was found to be 5 times as great as it was near London. The short-wave channel indicated in this cut was an experimental one at the time the long-wave channel was opened; it was the forerunner of the several now in operation.

The long-wave channel uses the same frequency (58.5 kc.) for both going and return channels; evidently it must be a considerable problem to make such a system function because of the interference at the receiver in Houlton from the transmitter in Rocky Point, and the same condition in England. The power level of this channel from San Francisco to London is shown in Fig. 64; a similar figure would illustrate the return path. It is evident that the receiver at Houlton must be sensitive enough to respond to the comparatively few micro-microwatts that arrive from England; it will therefore give a tremendous response to the signal from Rocky Point, which will be thousands of times as great as that from England. It is shown in Fig. 64 that in traveling across the Atlantic the signal changes its strength by the factor 10^{14}, or 140 decibels. Starting out at 150 kw. at Rocky Point it is only 1.5×10^{-9} watt at the receiving antenna. In Fig. 65 are shown the tubes of the amplifier, supplying the 150 kw. to the antenna.

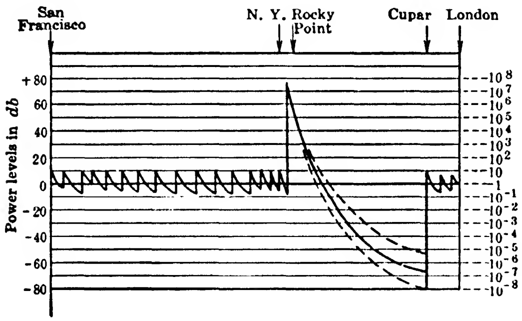


FIG. 64.—Power levels in the San Francisco-London telephone channel; at the left is given the power ratio in decibels and at the right is given the power ratio itself.

The receiving antenna is made directive for the direction of the desired signal, but even so the interference from the local transmitter would make the channel almost inoperative. An ingenious scheme to overcome this difficulty is shown in Fig. 66; this is the scheme actually used in this channel.[1] It operates in the following manner. Relay *A* is normally open so that received signals pass through to the subscriber. Relay *C* to normally closed to short-circuit the transmitting line. When the United States subscriber speaks, voice currents go into both the transmitting detector and the transmitting delay circuit. The transmitting detector is a device which amplifies and rectifies the voice currents so as to produce currents suitable for operating relays *A* and *C* which thereupon short-

[1] Bown, B.S.T.J., April, 1930, p. 258.

circuit the receiving line and clear the short circuit from the transmitting line respectively. The delay circuit is an artificial line through which the

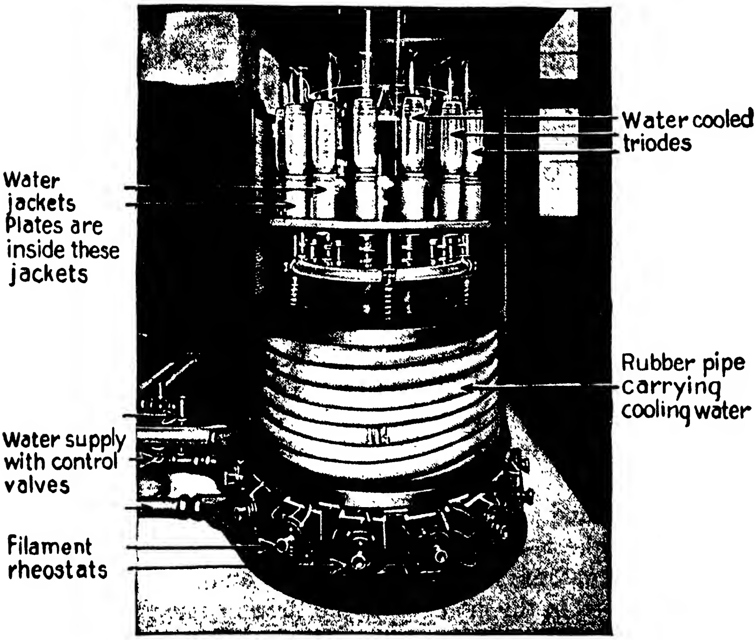


Fig. 65.—A bank of water cooled triodes for transatlantic telephony.

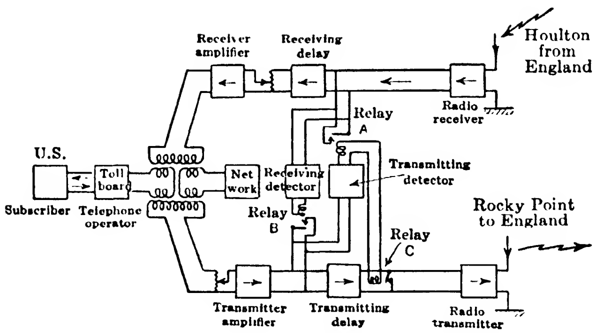


Fig. 66.—Conventional diagram of the two-way circuits of the transatlantic radio telephone channel.

voice currents require a few hundredths of a second to pass, so when they emerge the path ahead of them has been cleared by relay *C*. When the subscriber has ceased speaking the relays drop back to normal.

The function of the receiving delay circuit, the receiving detector, and relay B is to protect the transmitting detector and relays against operation by echoes of received speech currents. Such echoes arise at irregularities in the two-wire portion of the connection and are reflected back to the input of the transmitting detector where they are blocked by the relay B

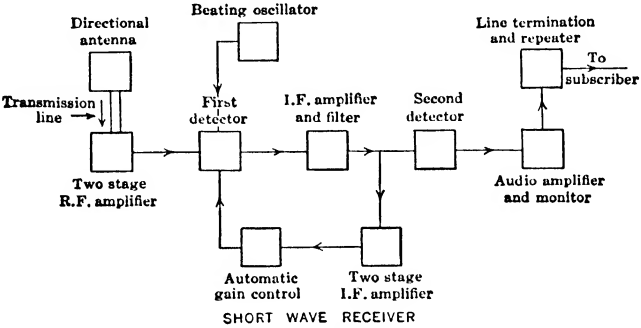


Fig. 67.—Diagram of the receiving net work of a short wave radio telephone channel.

which has closed and which hangs on for a brief interval to allow for echoes which may be considerably delayed.

The receiving antennas at Houlton, Maine, and Cupar, Scotland, are long wires stretched in the direction of the transmitting station for about three miles, perhaps 25 feet above the ground. They are of the "wave antenna" type and are described more in detail on pp. 904, 960. In a paper

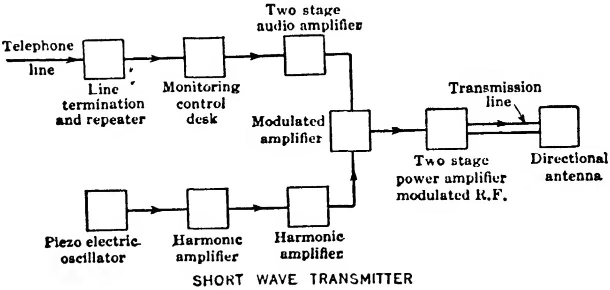


Fig. 68.—Diagram of the transmitting network of a short-wave radio-telephone channel.

dealing with this channel[1] it is stated that moving the receiving station from the vicinity of New York to Houlton increased the ratio of signal to noise by a factor of 50. Then the use of the directive antenna gives an added ratio of signal to noise of 400. If the receiving were to be done on an ordinary antenna in the vicinity of New York the power of the

[1] Bailey, Dean, and Winteringham, I.R.E., Dec., 1928, p. 1645.

transmitter in England would have to be increased by 20,000 times to get the same signal to noise ratio as is obtained at present.

Short-wave Transatlantic Channels.—For transatlantic channels 14–17 meters is good during the daytime; for either sunrise or sunset between transmitter and receiver 22 meters is best; during night over the path 33 meters is good, and for midnight in winter 50 meters may be best. Because of these conditions each short-wave transmitter is arranged to transmit on a wave length in each of three regions.

The general idea of the receiving apparatus is shown in Fig. 67, and that of the transmitting apparatus in Fig. 68. The receivers are located in Netcong, N. J., and the transmitters at Lawrenceville, N. J., as shown in Fig. 69. The transmitting antennas are in the form of curtains hung between steel towers, about 150 feet high and 250 feet apart. Each transmitting antenna occupies two of these spaces, so is in the form of a curtain 150 feet high and 500 feet long. A bird's-eye view of the transmitting antennas is shown in Fig. 70 and in Fig. 71 is given a plan of the antenna system. Each transmitting house is fitted with two transmitters, and each transmitter uses three sets of 2 bays each. Each two bays is filled with one antenna designed to transmit at the wave lengths noted on the diagram of Fig. 71. Each bay is filled with several sections of antenna made up as shown in Fig. 72; each section is higher than that shown in the figure by several more of the half wave length loops. Miniature transmission lines go to each antenna section from a common feeding point and are so adjusted in length that all vertical members have currents in the same phase. Thus an antenna consists of dozens of vertical conductors, all rigidly suspended in a vertical plane and all having currents of the same phase. This makes a sharply directive emitter.

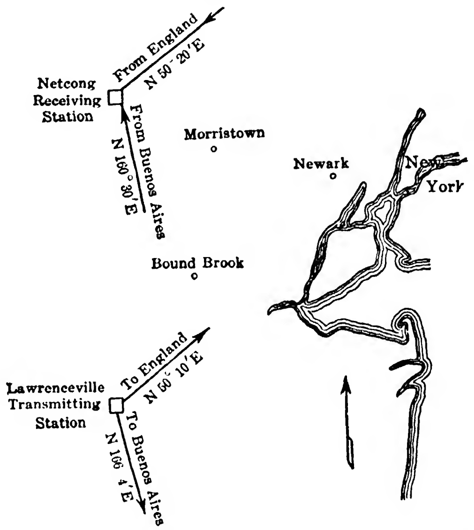


FIG. 69.—Short-wave receiving and transmitting stations of the A. T. and T. Co.

The receiving antennas are also of a sharply directive type but of different construction. The general form is shown in Fig. 73; the horizontal and vertical members are of copper pipe, rigidly supported in a wooden framework of poles and cross-braces. Each antenna is six wave lengths

FIG. 70.—Bird's-eye view of the transmitting station at Lawrenceville, N. J. Curtains of directive short-wave antennas (with reflectors) are hung between the steel towers. (Photo by Fairchild Aerial Surveys, Inc.).

long, built up of six times as many sections as shown in Fig. 73. They are built in a direction perpendicular to that from which they are to receive. The arrangement of the various antennas is shown in plan in Fig. 74; each group consists of three or four

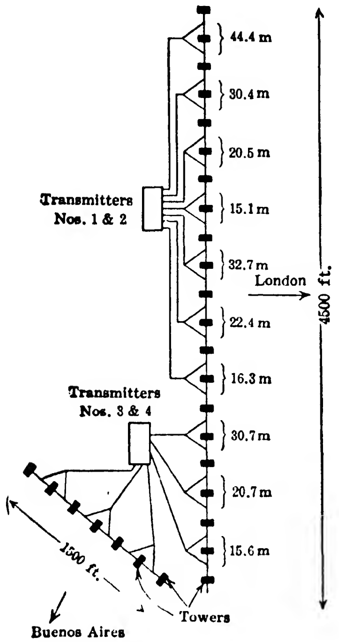


Fig. 71.—Plan of the short-wave transmitting antennas; each antenna occupies two bays of the steel tower layout.

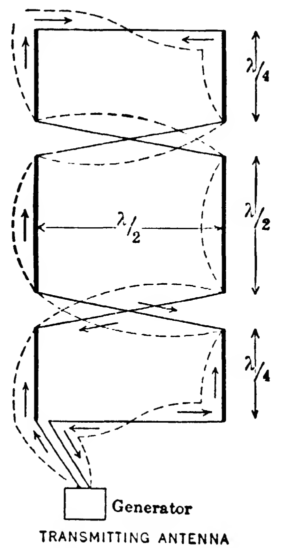


Fig. 72.—Type of antenna used for transmitting. Several of these are rigidly suspended in each bay.

antennas of such wave lengths that 24-hour communication is possible. From each group of antennas a transmission line of two concentric copper pipes (a $\frac{3}{16}$-in. pipe inside a $\frac{5}{8}$-in. pipe), these being spaced by washers of isolantite to keep the inside one central with respect to the other. The attenuation of the signal as it travels over these special transmission lines is surprisingly low, being only 2 db per thousand feet at 20 megacycles.

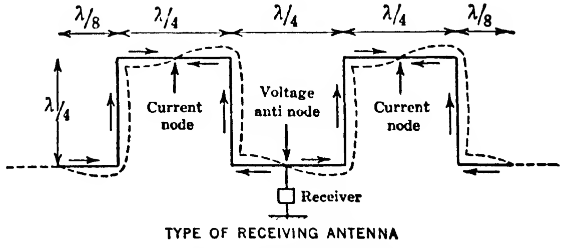


Fig. 73.—Type of antenna used for receiving short waves.

The various antennas are installed at reasonable distances from one another to prevent interaction between the used and unused antennas; all automobiles are kept a long way from the receiving antennas to prevent high-frequency waves of the ignition systems from interfering with the desired signal.

The circuit of each transmitter is crystal controlled; by proper frequency doublers and amplifying stages, the final ones using water-cooled tubes, the power supplied to the antenna by unmodulated carrier is 15 kw. At peak modulation (100 per cent) the antenna power is 60 kw.[1] The action of the directive transmitting antenna increases the signal, in the desired direction, over that which would be given by a simple vertical antenna, by 50 times; that is, the signal received in England is as great as though a simple vertical antenna was supplied with 750 kw.

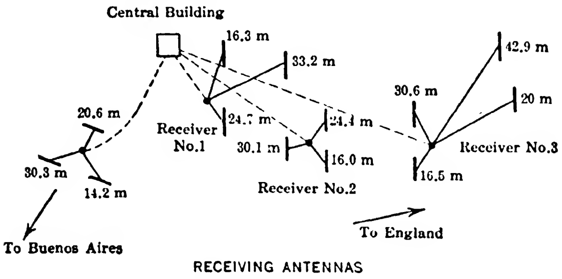


FIG. 74.—Plan of the various receiving antennas used at Netcong, N. J.

Radiophone Broadcasting. — There are today in the United States over 600 broadcasting stations, ranging in power from a few watts to 50 kw.; for experimental purposes some of the larger stations occasionally raise their power to a few hundred kilowatts. In general, they all require the same essential prices of apparatus,[2] which may be classified as below:

(*a*) The studio, microphone and speech amplifier;
(*b*) The radio-frequency system including oscillators and modulators;
(*c*) The power equipment;
(*d*) The control equipment.

The microphone transmitter has already been described; the double-button carbon type is generally employed. The amplifier for the speech currents is of the audio-frequency type (see chapter on amplifiers) and consists of three stages of amplification for a carbon transmitter and five stages for a condenser transmitter. This part of the transmitting equipment offers no features different from those mentioned in the section on microphones and the chapter on amplifiers except the "volume control." To show this, a more or less typical circuit arrangement is shown in Fig. 75. This

[1] See "Overseas Radio Extensions," Espenschied and Wilson, I.R.E., Feb., 1931, p. 282 for a good picture of one of these short wave transmitters.

[2] See "Transmitting Equipment for Radio Telephone Broadcasting," by E. L. Nelson, Proc. I.R.E., Vol. XII, No. 5, Oct., 1924; and "The Radio Telephone Broadcasting Station of the Westinghouse Elec. and Mfg. Co. at East Pittsburgh, Pa.," by D. G. Little, Vol. XII, No. 3, June, 1924.

shows the double-button microphone energized by means of the potentiometer D, connected in series with the steadying inductance F and the key K to the battery A.

The voltage across the secondary of the transformer G, due to the speech currents, is impressed upon the grid of T_1 through the potentiometer QS. Repeating from one tube into another is done by means of the iron core inductances L_1, L_2, and L_3. The final repeating into the radio transmitter is done through transformer H.

It will be understood that an amplifier of this sort is very carefully designed and built because it is most important that the output at Q_1S_1 should accurately represent the voice waves impinging on the microphone.

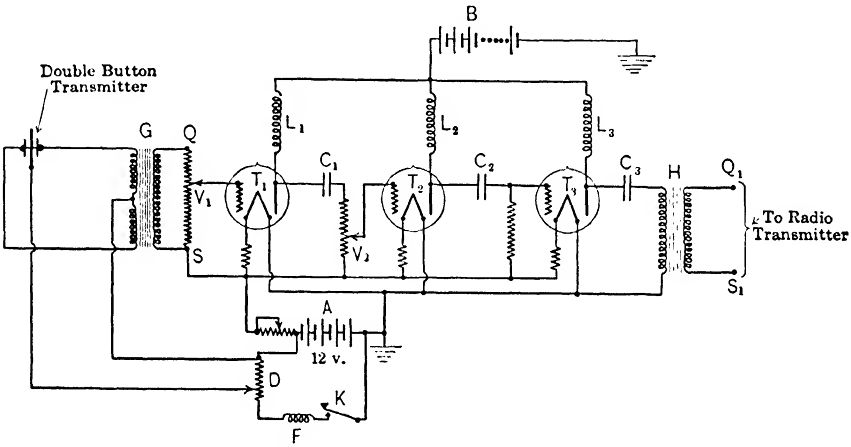


FIG. 75.—The speech amplifier circuit of a broadcast station; the two potentiometer controls V_1 and V_2 enable the operator to control the amount of amplification. The repeating inductances, L_1, L_2, and L_3, have iron cores.

The frequency characteristics of the amplifier must be flat throughout the range used in voice or music.

The variable contacts V_1 and V_2 control the amplitude of the voltages impressed on the grids of the first two tubes and so the amplitude of the voltage appearing across Q_1S_1 may be readily kept in control. This is an extremely important feature in broadcasting as it is in general necessary for the operator to decrease the range of intensity impressed on the microphone. The pianissimo and fortissimo passages of a selection are equalized to a certain extent by the radio operator. While this might be thought to distort the true character of a rendition the actual result in the receiving set is more pleasing than if the practice were not resorted to. The power range of an orchestral selection may frequently be 100,000 to 1 and the ear of the listener in the concert hall changes its sensitiveness automatically as these intensities occur. In the radio amplifier such power

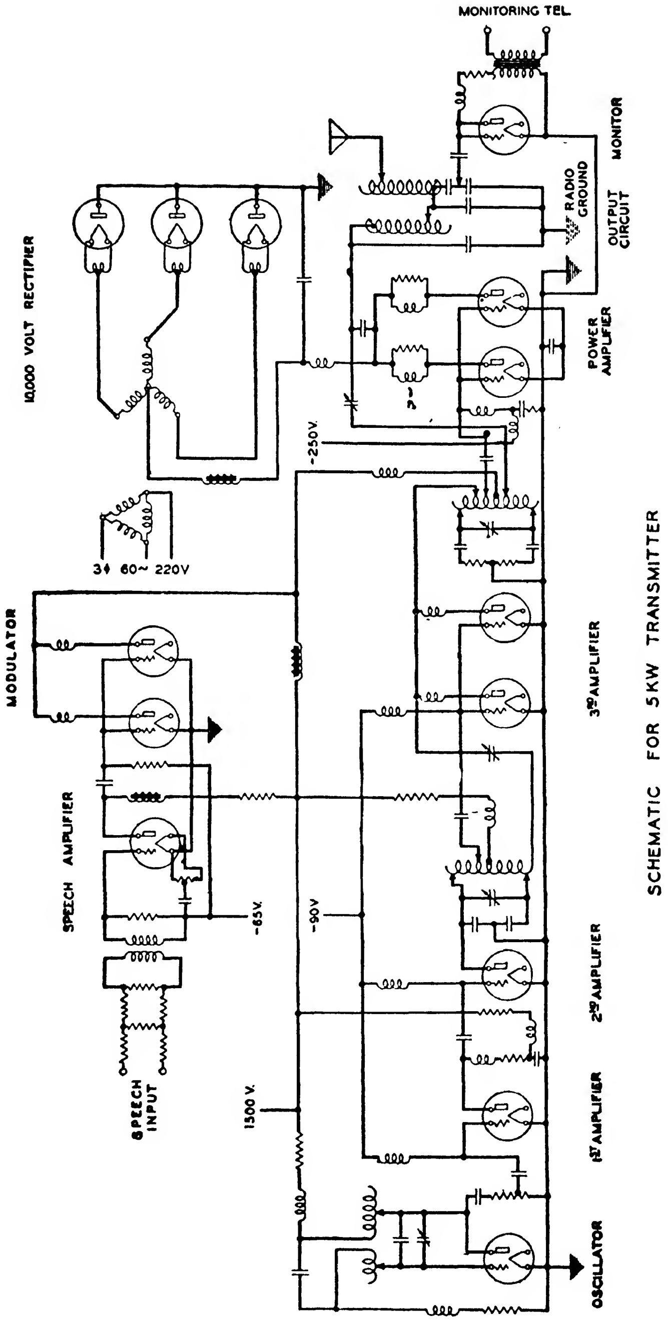


FIG. 76.—Simplified diagram of a 5 kw. broadcast transmitter.

ranges would result in unsatisfactory operation, tending to "block" some of the tubes or give so little intensity to the low passages that the radio listener gets nothing. It has been found that a power range of 1000 to 1 is plenty to give a radio reproduction reasonable similarity to the actual rendition and this is about the range used in practice.

Furthermore there is always a certain "background" of noise in a radio channel due to atmospherics, etc., and the low passages must be

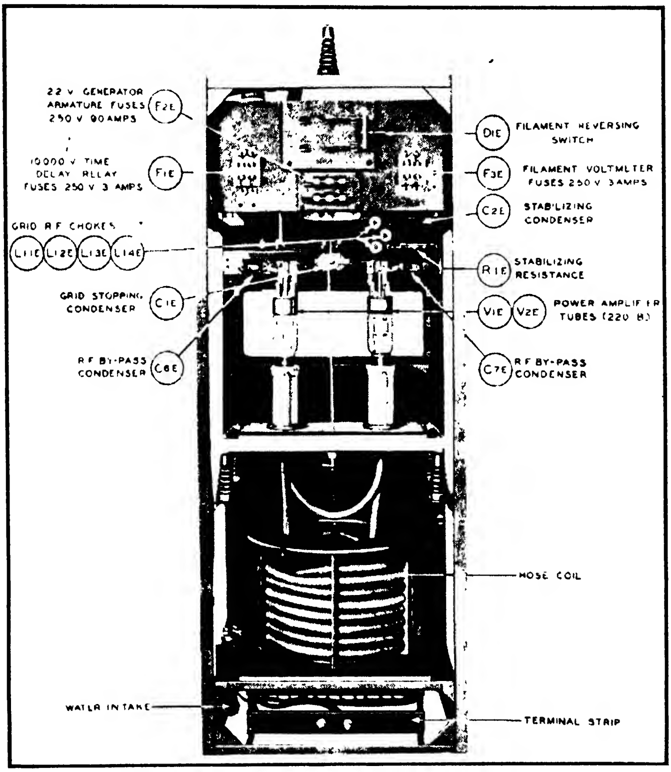


FIG. 77.—The two triodes of the "power amplifier" of Fig. 76.

kept louder than this background noise. An operator constantly keeps watch on the amplifier output and by manipulating the controls V_1 and V_2 he keeps the amplification within the proper range. The voltage output is several volts, the speech amplifier having raised the signal to this value from the few millivolts generated in the microphone.

(*b*) The radio-frequency system is no different from any of the oscillating and modulating systems already outlined. Plate-circuit modulation is the one generally used in the broadcasting stations in the United

States. Simplified circuit diagrams of low power sets are shown in Figs. 40 and 41. Fig. 76 shows the simplified diagram of a 5-kw. broadcast transmitter. In Fig. 77 are shown the two power tubes feeding the antenna, and in Fig. 78 is shown the arrangement of tuning coils and coupling

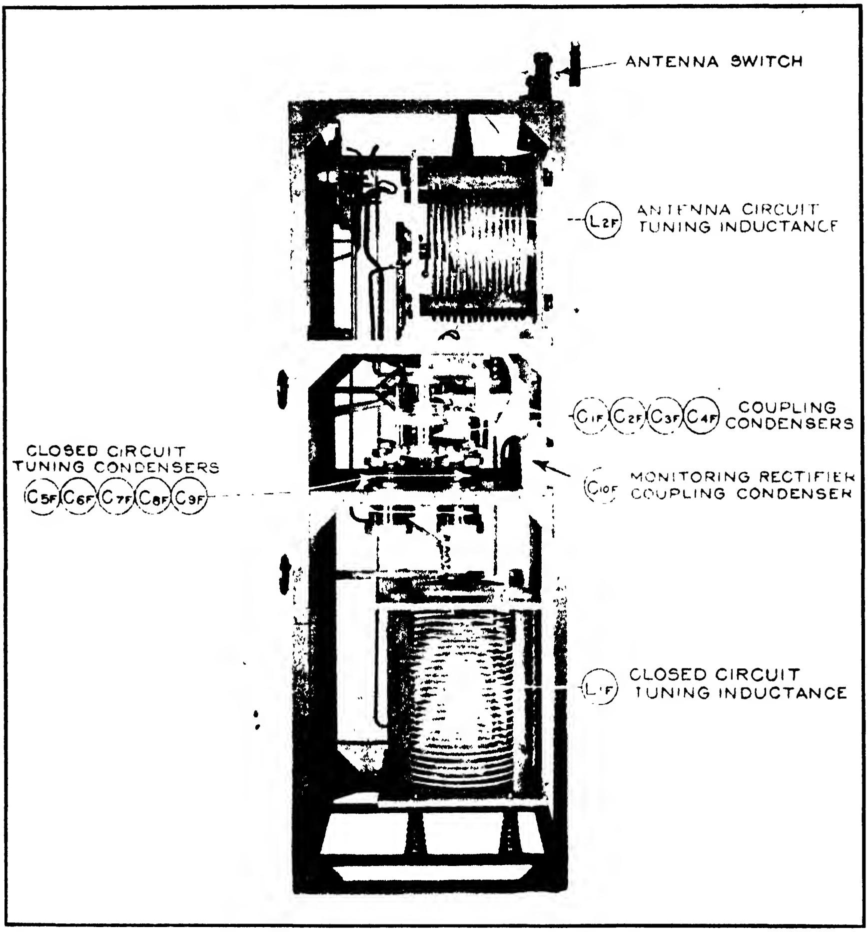


FIG. 78.—Tuning coils and coupling condensers, for getting the output of the power amplifier into the antenna.

condensers for the antenna and associated circuit. Fig. 79 is a rear view of a complete 5-kw. transmitter.[1]

(*c*) The power equipment for the smaller stations generally consists of a motor driving two c.c. generators; one of these develops 1500 to 2000

[1] These illustrations are of apparatus as manufactured by the Western Electric Co.

volts for the plate supply and the other is a 15–20 volt generator for heating the filaments. The generator voltages are controlled by suitable field rheostats. Suitable filter circuits are used to eliminate the ripples due to the commutator.

In general, batteries are used for operating the low power stages of the amplifier—this is important because any variations in power supply for the amplifier tubes is much exaggerated by the time it gets to the power tubes, and gives the familiar "background noise" of the poorer class of stations.

For stations of more than 1-kw. rating water-cooled triodes are used and the plate supply of from 10,000 to 15,000 volts is generally obtained from a bank of rectifiers. A common scheme is to arrange a three-phase

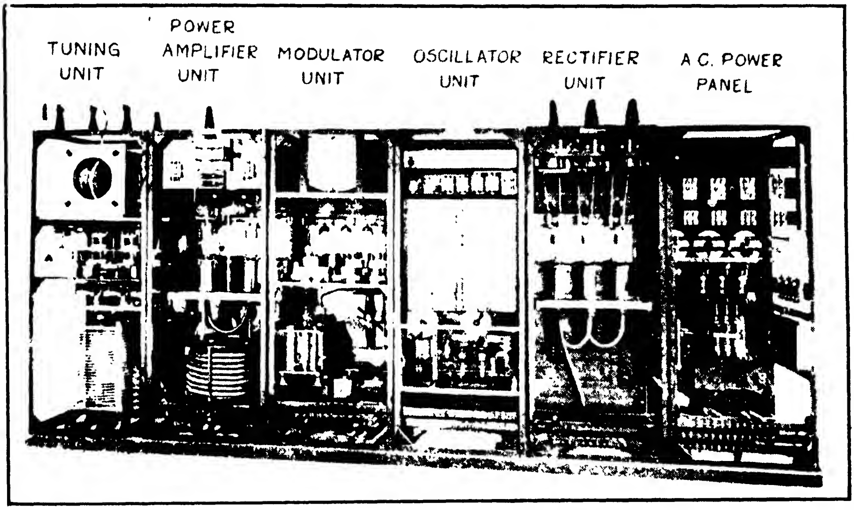


Fig. 79.—Rear view of a complete 5-kw. transmitter.

star-connected bank of transformers and rectifiers in each. The three-phase power supply eliminates much of the ripple which necessarily occurs with a single-phase power supply, thus requiring a less expensive filter system. To get the required grid bias small generators are generally used.

(*d*) The control equipment includes in addition to the field rheostats of generators, and volume control, already mentioned, a "mixing panel" from which the operator is able to pick up his signal from any one of several microphones placed in the studio or music hall. He may use several microphones at the same time in which case each microphone is potentiometer-controlled to enable the operator to "fade out" one of them as he gradually brings in the other.

The water-cooled tubes require pumps, etc., and suitable safety relays are operated by the temperature of the cooling water. Proper retard circuits enable the various operations, lighting filaments, starting water supply, putting on grid bias, putting on plate voltage, etc., to take place in proper sequence and at the proper time intervals.

Ammeters are provided for the antenna radio-frequency current; also for the plate and grid circuits of both oscillators and modulators. The ammeter in the plate circuit of the modulator responds to changes in speech volume and thus serves as a correct measure of the modulation.

In general the grid of the modulator should not swing positive and the modulator grid ammeter serves to indicate when this is so. Ordinarily this ammeter should read zero. The amount of excitation furnished to the grid of the power amplifiers[1] is ordinarily continuously variable so that the operator can readily control the antenna power.

In all of the better stations the master oscillator is operated from a piezo-electric crystal (see p. 627) thus insuring constancy in wave length to a degree finer than can be measured.

Some of the modern stations are arranged to broadcast the same program at two or more wave lengths, using two different transmitters and antennas, but the same speech amplifier. The longer wave is used for local reception (within a few hundred miles) while the short wave is intended for those a thousand or more miles away. The short-wave signal has occasionally been used to control the output of another station but such use of the ether does not seem to be good engineering. Wire lines are much more suitable for operating distant stations.

Harmonics from Transmitters.—A broadcasting station having a fundamental frequency of 600 kc. may possibly be sending out also a second harmonic of 1200 kc. and a third harmonic of 1800 kc. These upper harmonics may seriously interfere with reception from other stations assigned to operate on these frequencies. A good 50-kw. transmitter sets up a field of about 2 volts per meter at a distance of 1 mile. A harmonic of 0.1 per cent would therefore give a field strength of a 2 millivolts per meter and would completely spoil reception from another station perhaps 50 miles away, assigned to send on this frequency. As pointed out in Chapter IV, a field strength as weak as 100 microvolts per meter gives what is called satisfactory reception if there is no interference. This harmonic of 0.1 per cent from the 50-kw. transmitter would be 20 times as strong as this field strength of 100 μv per meter, and so would completely mask the other station.

It is thus evident that stringent limitation on the percentage harmonic sent out from the powerful stations must be imposed, otherwise they not

[1] See "Power Amplifiers in Transatlantic Radio Telephony," by Oswald and Schelleng in Proc. I.R.E., Vol. XIII, No. 3, June, 1925.

only appropriate their own channel but one or two others besides. This point has been discussed by Nelson,[1] and he concludes that a well-built transmitter can hold down its harmonics to 0.05 per cent, and possibly lower as designs improve.

Modulation Distortion in Transmitters.—The voice or music in the studio acts on the microphone, which itself may introduce much distortion. (See p. 795.) The speech amplifier raises the microphone output perhaps 1000 times in voltage, and this is still further raised to perhaps several hundred or a thousand volts and then applied to some convenient power stage of the transmitter, where it modulates the radio-frequency power. This modulated radio-frequency power is then further amplified by power tubes and supplied to the antenna. It will be appreciated that it requires a high grade of engineering skill to produce, in the antenna,

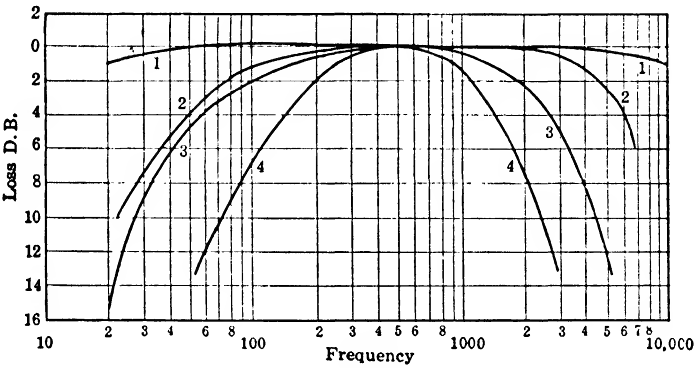


FIG. 80.—Fidelity curves of transmitters and receivers. Curve 1 is for a modern 50 kw. transmitter; curve 2 is for a 5 kw. transmitter as made eight years ago, while curves 3 and 4 are the fidelity curves of two modern radio broadcast receivers.

modulated power which varies exactly as does the few microwatts of voice power acting on the microphone. Fig. 80 shows the fidelity of modulation of a Western Electric 50-kw. transmitter; it is constant within about one decibel from 20 cycles to 10,000 cycles per second. On the same curve sheet are shown the fidelity curve of a transmitter of 1924, as well as the frequency response curves of two typical receivers. It is seen that the output of the broadcast station is of much better quality than that given off by the receiving set.

Loud Speakers.[2]—Practically all good broadcast receivers today operate a loud-speaking telephone receiver instead of the headphones

[1] I.R.E., Nov., 1929, p. 1948.

[2] See Proceedings of Physical Society of London, Vol. 36, Parts 2 and 3, Feb. and March, 1924 for discussion of loud speakers. Also see "Electrical Loud Speakers," by A. Nyman, Jour. A.I.E.E., 1923, and "Design of Telephone Receivers for Loud Speaking Purposes," by Hanna, Proc. I.R.E., Vol. 13, No. 4, Aug., 1925.

formerly used. No matter how good the transmitting station may be, or how perfectly designed the receiving set the program is spoiled for the listener if the loud speaker does not function properly.

Since speech and music contain certain frequencies ranging from 40 cycles to 15,000 cycles (see Fig. 8, p. 787), and since faithful reproduction depends upon the loud speaker giving proper response to each of the frequencies, it will be seen that a loud speaker must have two characteristics: for a given frequency the amplitude of vibration of the air result be proportional to the current input and for a given current input the amplitude must be independent of the frequency.

It is evident that vibrating diaphragms and bars can scarcely satisfy these conditions, simple as they are, for mechanically vibrating systems having mass and spring forces are sure to vibrate more readily at some frequencies than at others. The early loud speakers were remarkably resonant for certain frequencies, thus greatly accentuating these notes in an orchestral selection and thus spoiling the quality.

To eliminate the effects of mechanical resonance it is necessary to introduce damping and this damping should preferably be caused by the radiation of sound. Now it is difficult to extract much energy from a small diaphragm in the form of sound; the air next to the diaphragm " runs away " and refuses to take that compression which is necessary for the propagation of sound waves.

To hold the air close to the vibrating diaphragm and force it to extract sound energy is the function of the familiar horn. The air cannot spill sidewise from the diaphragm behind a horn as it can with the hornless diaphragm and so stays in position in front of the vibrating disc and thus receives sound energy. In the papers previously referred to mathematical relations defining the function of the horn are derived. It is shown that the opening of the horn must be reasonably large, that the cross-section of the horn must increase slowly and logarithmically and that the neck of the horn must be rigid.[1] The material of which the horn is made must be of non-resonant quality, such as papier mâché or cement. Unless the suitable precautions are taken the horn itself will act like a resonant air column and thus accentuate a certain characteristic note which depends upon the length of the horn. The lower the notes to be reproduced the longer the horn must be.

The very great effect of a horn on the characteristics of a loud speaker is well brought out by measuring the electrical constants of the loud-speaker winding, with and without horn. With the diaphragm clamped tight the measured resistance and reactance throughout the frequency

[1] See "The Function and Design of Horns for Loud Speakers," by Hanna and Slepian, Jour. A.I.E.E., 1924; also "The Performance and Theory of Loud Speaker Horns," by Goldsmith and Minton, Proc. I.R.E., Vol. 12, No. 4, Aug., 1924.

range are quite smooth, showing just as would be expected, decreasing inductance and increasing resistance with increasing frequency. With the diaphragm free to vibrate, and no horn, very pronounced effects on both resistance and reactance occur at the natural frequency of the diaphragm. This is well shown in Fig. 81, which shows the characteristic of one of the well-known loud speakers. Evidently the natural period of the diaphragm was 940 cycles per second. The increase in resistance brought about by the vibrating diaphragm at this frequency is more than 50 per cent of the total resistance, that is, the resistance introduced by the motion of the diaphragm is more than the resistance due to the winding and iron losses.

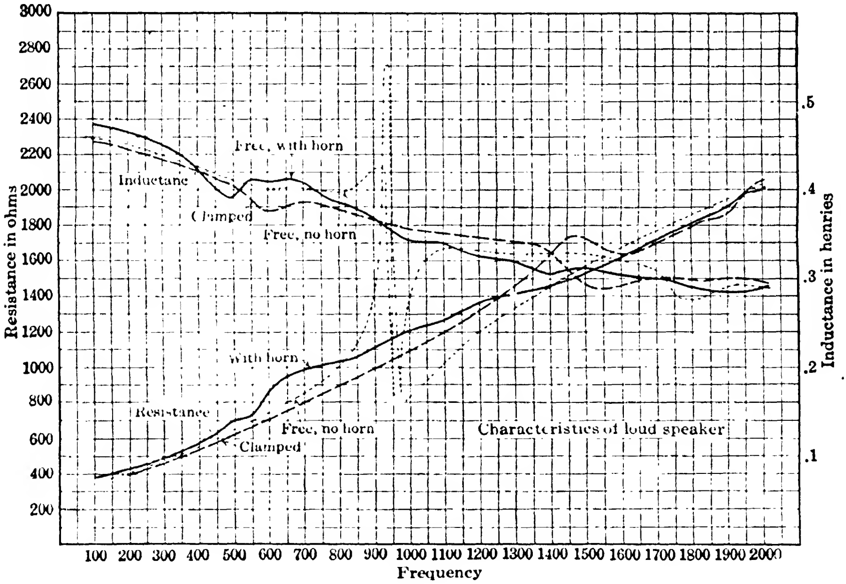


FIG. 81.—The electrical characteristics of a loud speaker under various conditions; the addition of the horn practically eliminates the resonant quality of the vibrating system.

When the horn is put in place the resonance condition disappears completely, showing how effectively the air column in the horn damps out the natural resonant vibration of the diaphragm. It will be seen that the presence of the horn quite appreciably raises the resistance throughout the low-frequency range; this electrical determination thus emphasizes the well-known fact that the addition of the horn " brings out " the low notes. The hump in the resistance curve (for the clamped diaphragm) at 1500 cycles was probably due to the fact that clamping did not hold it perfectly tight; under the clamping conditions there was probably considerable vibration at this frequency, as indicated by both resistance and inductance curves. The type of loud speaker used in getting these curves was somewhat like that shown in Fig. 82.

Provided the horn is of the right form and size the speaker shown in Fig. 82 will reproduce reasonably well the low notes of a program, but the large air spaces around the diaphragm make it inefficient for the high frequencies. The air cavities around the diaphragm must be very carefully formed if the motion of the diaphragm is to impart much sound energy to the air in proximity to it. Wente and Thomas[1] describe a moving coil type of speaker which has a very light, rigid duralumin diaphragm, and a carefully designed air chamber around it, which has remarkable efficiency. The ordinary loud speaker gives off in the form of sound but a small part of the electrical power input, probably never to exceed 5 per cent and generally much less than this. This special speaker of Wente is shown to have an efficiency as high as 50 per cent; its construction is indicated in Figs. 83 and 84, from which it will be noticed that the throat of the horn is a narrow ring opening. The field magnet develops a flux density in the air gap of 20,000 gausses. The coil is of edgewise-wound aluminum ribbon 0.015 in. wide and 0.002 in. thick with 0.0002 in.

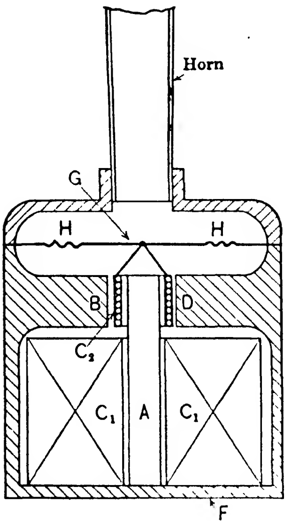


FIG. 82.—The "moving-coil" type of loud speaker; the small solenoidal coil C_2 is placed in the narrow circular air gap of the powerful electromagnet.

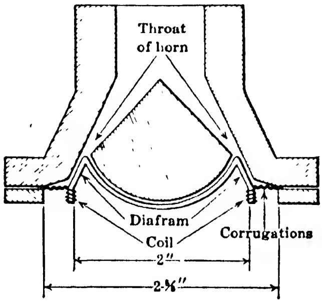


FIG. 83.—Cross-section of an efficient horn type loud speaker, designed for the low-frequency range. Mass of coil + diaphragm = 1.09 grams; area of diaphragm = 28 sq. cm.; area of throat of horn = 2.45 sq. cm.

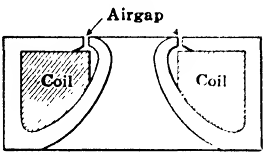


FIG. 84.—The magnetic field structure for the loud speaker shown in Fig. 83.

[1] B.S.T.J., Jan., 1928.

insulation between turns. The 70 turns are held rigidly together by the cementing action of the insulation, so the coil needs no form to hold it together. With a very long horn, having a cross-sectional area that increases exponentially with increasing distance from the diaphragm (a so-called *exponential horn*), the sound output at 60 cycles varies as shown in Fig. 85 as the electrical input is increased. The speaker will give off a large fraction of 1 watt with but little distortion. Its performance throughout the range of usual frequencies is shown in Fig. 86; in this figure " zero *db* " is taken as 100 per cent efficiency.

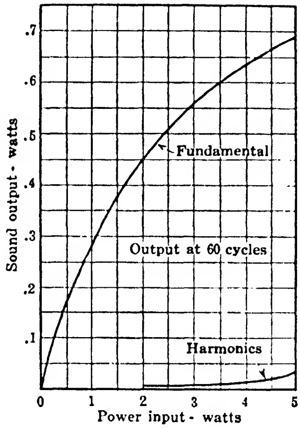


FIG. 85.—Curve showing the conversion efficiency of the speaker shown in Fig. 83.

It will be seen that even this well-designed speaker becomes inefficient at high frequencies. To reproduce the tone of cymbals, or similar percussion instruments, it is necessary for the speaker to give a good response for frequencies as high as about 15,000.

Bostwick [1] describes a speaker built in fashion somewhat similar to that shown in Fig. 83, but with more restricted air spaces in the throat, and lighter and more rigid diaphragm. The cross-section through this speaker is shown in Fig. 87; the diaphragm of 0.005-in.-thick duralumin is crowned and has a coil fastened to the edge of the crown similar to that shown in Fig. 83. The diaphragm is only 1 in. in diameter and weighs, with the coil, only 0.16 gram.

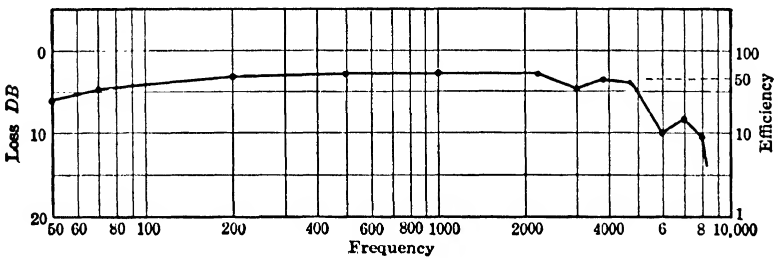


FIG. 86.—Conversion efficiency of the speaker shown in Fig. 83, throughout the useful frequency range.

The throat, close to the diaphragm, is an annular opening having an area of only 0.19 sq. cm. The horn itself is only 6 in. long, and the speaker is so small that it can be supported right inside the mouth of the horn of the larger speaker.

[1] Journal Acoustical Society of America, Oct., 1930, p. 242.

The response of this small speaker is nearly uniform from 3000 to 12,000 cycles, and the response is still appreciable at even higher frequencies. By supplying power to both the large speaker of Fig. 83 and this small one, through a filter as shown in the diagram of Fig. 88, the combined response is extremely uniform from 60 to 12,000 cycles, as shown in the curve of Fig. 88. The L_1, C_1, C_2, filter is designed to give about equal division of power at 3000 cycles; above this frequency the smaller speaker gets the larger share of the available power.

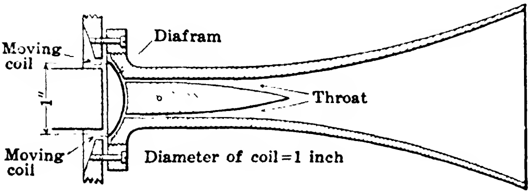


FIG. 87.—Cross-section of a speaker designed to cover the high-frequency range.

Hornless Loud Speakers.[1]—By fitting the moving coil of the speaker to a rigid, shallow cone of some light material, such as paper, the cone having an outside diameter of perhaps 10 in., a fairly good response is obtained over the ordinary frequency range. The ordinary construction is shown in Fig. 89. The outer edge of the cone is held in place by a flexible membrane of some sort (cloth or leather), and this is fastened to the edge of a circular hole in a baffle board. At low vibration frequencies the air tends to run around from the front of the moving cone to the back; the baffle board prevents this flow of air and so tends to strengthen the low-frequency response of the speaker. The coil D, fitting loosely in the annular air gap between pole piece B and ring C, is cemented to the frustrated apex of the paper cone. The magnetizing winding A, made of many turns of fine wire, is generally used as one of the chokes in the plate supply filter, so its mmf. is a pulsating one. But a pulsating flux in the air gap will induce currents in the coil and so impart to the coil a to-and-fro motion of the frequency of the magnetic field variations. This is one

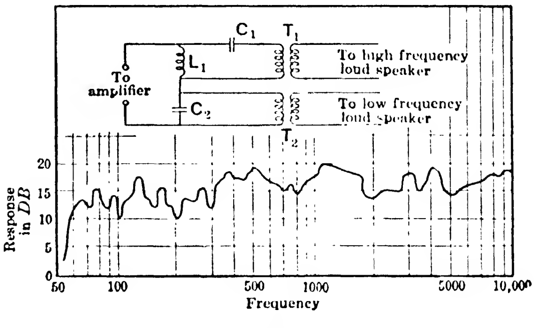


FIG. 88.—By using the speakers shown in Fig. 83 and Fig. 87 in series, through the special filter shown here, the sound output is reasonably uniform from 60 to 10,000 cycles.

[1] "Notes on the Development of a New Type of Loud Speaker," Rice and Kellogg, Jour. A.I.E.E., Sept., 1925.

reason for the familiar "hum" from the loud speaker of a poorly designed set.

A ring E composed of alternate layers of copper and iron will effectually stop pulsations of the magnetic field in the air gap. It acts to attenuate the pulsations in the magnetic field as they proceed from coil A towards the air gap. Of course these rings also act to cut down the choking action of the coil, to some extent, but this decrease is not large.

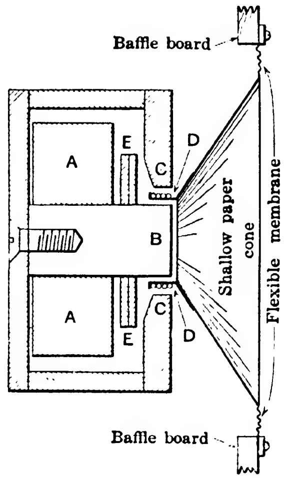


FIG. 89.—The familiar type of "dynamic" speaker; the coil D moves the paper cone back and forth.

In Fig. 90 are shown the sound response curves from a speaker of the type shown in Fig. 89, when two different sizes of baffle board are used. The moving paper cone fitted in a hole in the center of the baffle board. The sound measurements were made outdoors where reflections would be eliminated. It is quite evident that the larger baffle board materially raises the low frequency sent out from the speaker.

These paper cones, although built to act like a rigid diaphragm actually develop wave motions at the higher frequencies. The actual motion of the diaphragm is very small at the ordinary voice frequency but for low notes such as the organ gives, it may vibrate a millimeter or more. A certain speaker [1] gave an amplitude of one micron (center of the cone) with a

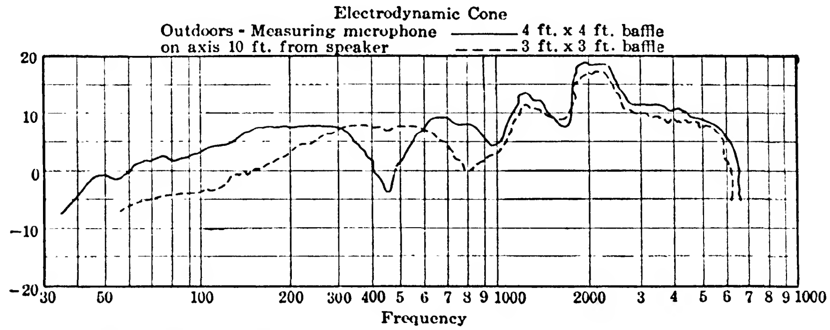


FIG. 90.—Effect of baffle board on response of speaker. Vertical scale in *db*.

500-cycle current of 0.004 ampere. With increasing distance from the center the amplitude rapidly diminished.

At 500 cycles this cone showed a reversal of motional phase along a circle 5 centimeters from the center. At 800 cycles it showed one reversal

[1] "On the Amplitude of Driven Loud Speaker Cones," Strutt, I.R.E., May, 1931, p. 839.

3 cm. from the center, a second reversal at 7 cm. from the center and a third at 10 cm. from the center. This cone showed circular nodes and antinodes which approached one another as the frequency was raised.

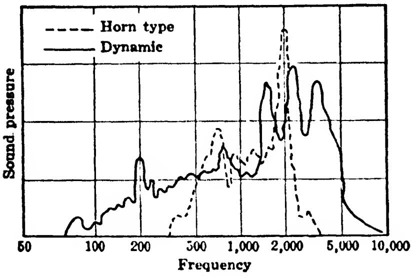


FIG. 91.—Comparative responses of a poor horn speaker and the ordinary good moving coil speaker.

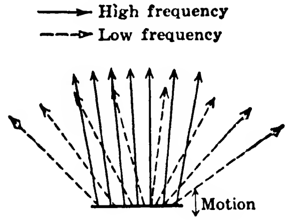


FIG. 92.—A flat, moving diagram gives off high-frequency sound nearly in the form of a narrow beam. The same diagram gives off low-frequency sound in a more nearly radial distribution.

A condenser type of loud speaker has been developed[1] but it has not come into common use, evidently being much less efficient than the types described above.

Loud Speaker Response Curves.—The loud speaker furnished with the average radio receiving set does not give a response curve as uniform as that shown in Fig. 88, by any means. In Fig. 91 are shown the response curves of a poor horn-type speaker (as furnished with some receiving sets) and a dynamic, or moving coil, speaker of the type shown in Fig. 89.

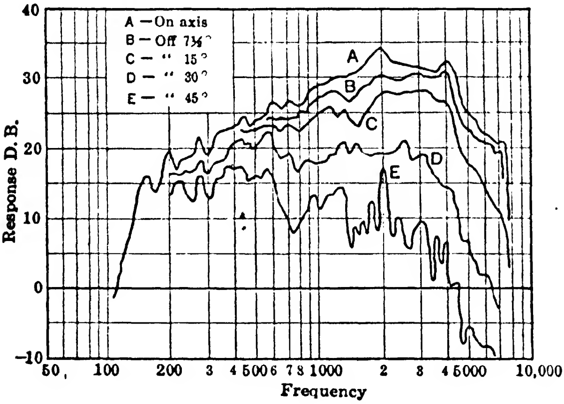


FIG. 93.—For low frequencies the speaker sends equally well inside of a solid 45° angle but at the higher frequencies points situated off the axis of the speaker receive comparatively little sound.

Distribution of Sound from a Loud Speaker.—The sound does not proceed equally well in all directions but tends to be directed along the axis of the moving diaphragm; this is especially true of the flat condenser type of speaker or of the flat eddy current speaker in which a copper plate is made to vibrate back and forth by eddy currents induced in it. Furthermore,

[1] "The Kyle Condenser Loud Speaker," I.R.E., July, 1929, p. 1142.

the higher frequencies tend to be confined to straightforward propagation more than the low ones. This idea is illustrated in Fig. 92; Fig. 93 shows the response of a microphone placed at equal distances from a loud speaker,

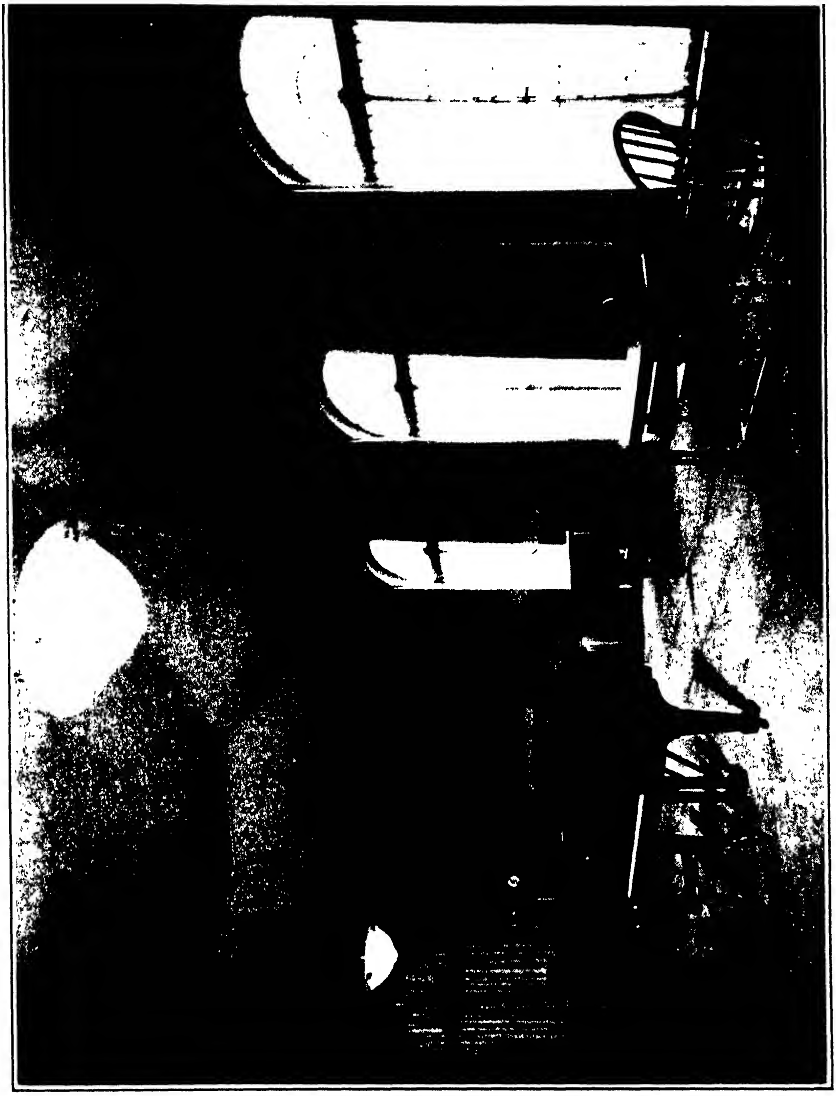


Fig. 94.—The studio of a broadcast station; the draperies, as well as the covering of floor and ceiling, are properly chosen to get the desired amount of reverberation.

in different directions. It can be seen that the low frequencies, up to say 200 cycles, were as strong 45° away from the axis of the speaker as they were on the axis, but with increasing frequency the amount of sound off the axis became less and less.

Thus at 5000 cycles the response of the microphone 45° off the axis of the speaker was 30 *db* lower than when on the axis; this means that the 5000-cycle note is only one thousandth as intense at 45° as it is directly in front of the speaker.

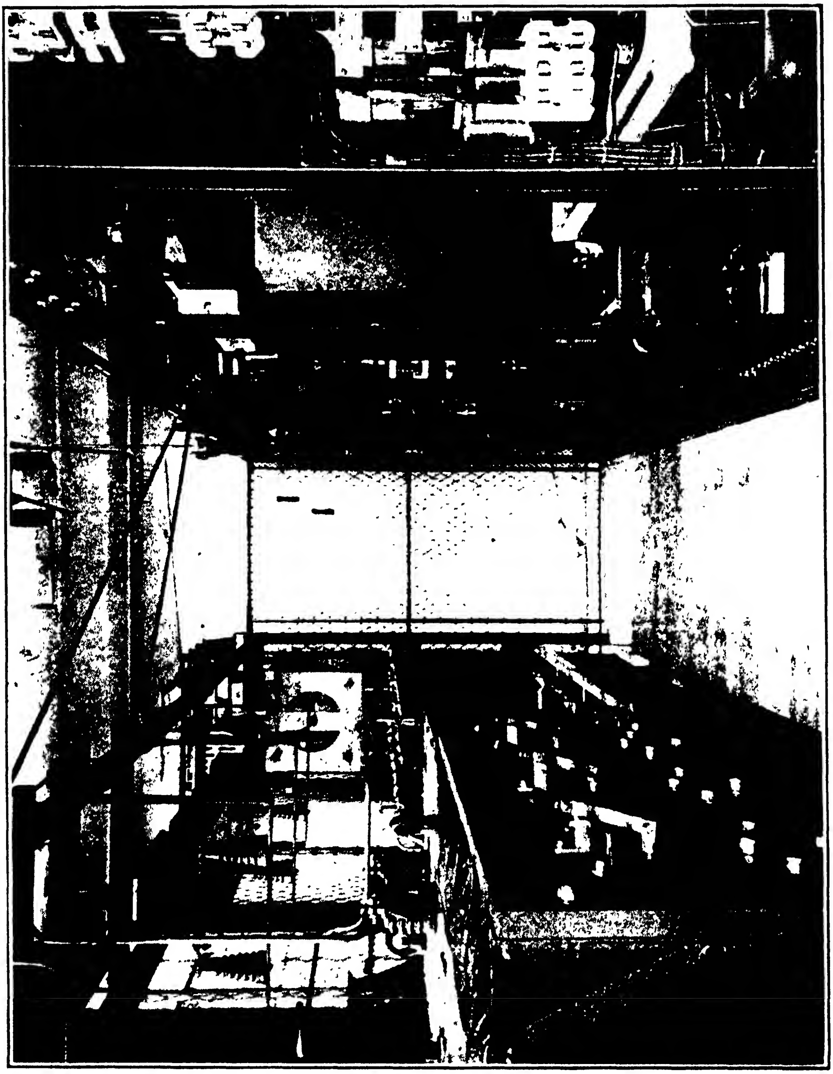


FIG. 95.—Arrangement of the apparatus behind the panel.

Typical Broadcasting Station Apparatus.—As indicative of the real engineering construction involved in modern transmitting equipment there are shown in Figs. 94 to 97 the equipment of a good 5-kw. station. The studio, with its hangings and draperies to get just the right amount of reverberation, is shown in Fig. 94 and in Figs. 95 and 96 is shown the transmitting apparatus proper. Water-cooled tubes are visible through the switchboard openings and the many meters which show how they are performing are mounted on the front of the board. In Fig. 97 is shown

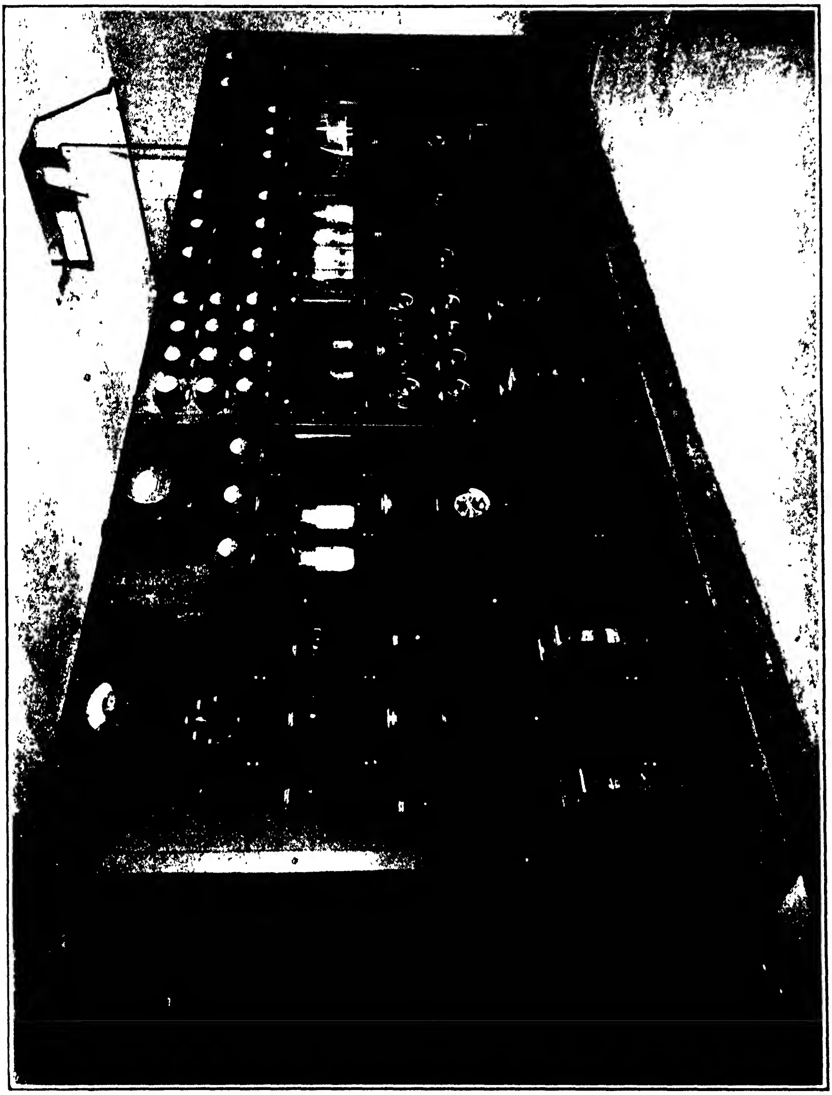


Fig. 96.—Panel view of a 5-kw. transmitter. The water-cooled tubes can be seen through the windows in the panels.

the speech amplifier, located at the studio. The output from various microphones is here "mixed" and properly controlled in amplitude before being sent to the transmitter station generally located several miles from the studio.

The antenna of this station, designed to operate at about 316 meters, has a natural wave length of about 280 meters, with no added inductance.

FIG. 97.—The speech amplifier at the studio; here the intensity of the voice current is properly adjusted before being sent over the wires to the transmitter station about twenty miles distant. Note the "monitoring" speaker overhead, for the operator's use.

The small inductance required to couple the antenna to the high-frequency power supply is sufficient to bring the wave length up to 316 meters. At this wave length the total antenna resistance is in the neighborhood of 15 ohms, of which possibly 10 ohms is radiation resistance. When transmitting with full power, the apparatus of Fig. 96 will supply 22 amperes to the antenna; the antenna power is thus between 4 and 5 kw.

The cost of such a transmitting outfit as that used in this station is between \$50,000 and \$150,000, and the cost of operating the station for one year is about the same amount.

The size of the studio naturally depends upon the type of program to be broadcast, number of performers, etc. Walls and ceiling are made of some sound-absorbing material such as hair felt, celotex, and similar materials. The reverberation is actually measured (by a form of oscillograph) and by proper construction of room, hangings, etc., is reduced to the desired amount. Generally two studios are in plain view of the station announcer so that one program may be in progress while the next is being prepared.[1]

In a broadcast station of most recent design the vacuum-tube equipment will put an average power of 50 kw. into the antenna with the possibility of delivering 200 kw. to the antenna during the modulation peaks. The input to such a station is about 350 kw. of 60-cycle 2300-volt three-phase power. Six rectifier tubes, fed through a six-phase transformer, supply 20,000 volt continuous current power for the plate circuits of the large amplifier tubes. The amplifier tubes, from the 50-watt size up through the series to the large water-cooled tubes supplying the antenna, are arranged in pairs, in the " push-pull " or balanced amplifier arrangement. (See Fig. 41, p. 819.)

Continuous-current generators are used for grid bias and filament power. Tremendous filters are used in the various circuits to eliminate the ripples, as much as 5000 microfarads (of electrolytic condensers) being used in one filter section. Some of the coils of the filter system weigh as much as four tons.

Location of Broadcasting Stations.—The first tendency was to place the broadcasting station right in the densely populated districts, but this practice should never be followed under present conditions. With many stations operating at the same time, perhaps 50 kc. apart in frequency, the listener close to one of the stations is practically cut off from the others.

All the stations should be located 20 miles or so from the city; all of the listeners in the city will then receive a field strength of a few millivolts per meter, and none of them will get signals of a hundred or more millivolts per meter as is the case when the station is located in the midst of the listeners. Furthermore when the station is located in a

[1] For detailed information, see " Design and Construction of Broadcast Studios," Hanson and Morris, I.R.E., Jan., 1931, p. 17.

district of steel construction buildings most of the radiated power is absorbed by these buildings within a mile or so of the station.

Costs of Broadcasting Stations.—The costs of constructing and maintaining a broadcasting station are much greater than would naturally be supposed, in view of the small amount of power they put out. An ordinary electric power plant, for example, costs perhaps $100 to $200 per kw. of power rating, but power in the form of modulated radio-frequency current is much more expensive.

A detailed estimate of the costs of the better class stations has been compiled,[1] from which the following brief summary is taken. The capital investment represents the average minimum, including buildings, antenna, and all actual apparatus for station and studio. The annual maintenance cost is exclusive of program talent, merely covering wages for operators, depreciation, etc. This latter item is taken as 25 per cent, in so far as actual transmitter apparatus is concerned.

Station power	1 Kw	5 Kw	50 Kw
Capital investment	$44,900	$127,000	$338,000
Annual maintenance:			
Studios and offices	24,900	52,000	83,000
Plant	24,150	73,100	213,150
Total annual cost:			
Five hours daily	49,050	125,100	
Ten hours daily	64,400	154,100	296,150
Investment and operating cost for first year	109,300	281,100	634,150

[1] "Electronics," June, 1931, p. 688.

CHAPTER IX

ANTENNAS AND RADIATION

Simple Antennas—Mechanism of Radiation.—As already described (p. 207), an antenna consists of one or more wires, suitably arranged, by means of which electromagnetic waves are radiated when high-frequency currents are sent into the wires.

The simplest type of antenna is the one shown in Fig. 1, consisting of two wires, *BC* and *DF* with an alternator, *A*, or some other source of high-frequency power, connected in the middle. In this arrangement one of the two wires may be considered as the " aerial," while the other performs the function of a " counterpoise." Both wires in this case, however, radiate electromagnetic waves, whereas in most arrangements the counterpoise is so arranged that it radiates but poorly compared to the aerial proper.

Fig. 1.—Theoretically the simplest type of antenna, the two wires *CB* and *DF* form the two plates of an *open* condenser.

The fundamental action of the alternator, as its electromotive force varies from positive to negative and vice versa, is to charge the wire *BC* positively, while, at the same time, wire *DF* is charged negatively, and, later, to reverse the charges on the two wires. It is plain that if, say, *BC* is to be charged positively, electrons must be taken from it by the alternator and transferred to some other conductor, which, in this case, is *DF*. Again, when *BC* is charged negatively, electrons must be taken away from *DF* and transferred to *BC*. Hence the obvious necessity of having electric conductors capable of storing electricity, or conductors with a reasonable amount of capacitance, connected on both sides of the alternator. Thus, it would not be advisable to use the arrangement shown in Fig. 2, for, in this case, the storage capacity of *BC* would be relatively small. As pointed out in Chapter II, the capacity of such a combination (*BC* and *DF*) depends upon the surface of each conductor; if either of them is made very small the capacity of the combination (which determines how many electrons may be transferred by the action of alternator *A*) approaches zero and the amount of radiation possible also approaches zero. On the other hand, it is common practice to connect as shown in Fig. 3, where the ground *G* forms a very good second plate of the condenser, since its surface is very large, giving a reasonable capacity to the condenser made up of *BC* for one plate and the earth for the other. Furthermore, the wire *DF* of Fig. 1 may be laid out horizontal, insulated from the earth. Also it may consist of several wires, spread out horizontally, all joining at *D*. This arrangement forms the well-known counterpoise ground, used especially where the ground resistance is high.

FIG. 2.—Without the lower wire the capacity of the condenser is so small that the alternator could not force an appreciable current to flow in in upper wire.

In order more fully to understand how energy may be radiated in the form of electromagnetic waves by means of an antenna, we will first go over some fundamental principles in connection with magnetic and electric fields.

An electric field exists in the region where electric forces are manifested, and the intensity of such a field at any point is measured by the force acting upon a unit charge of electricity placed at the point in question.

Similarly, a magnetic field exists in the region where magnetic forces are manifested, and the intensity of such a field at any point is measured by the force acting upon a unit magnetic pole placed at the point in question. The lines of action of the electric or magnetic forces are called electric or magnetic lines of force and represent, at any point, the direction of the force. It is also convenient to represent graphically the intensity of the electric or magnetic field by drawing more or less lines per unit area corresponding to a stronger or weaker field respectively; but it must be kept in mind that the force exists everywhere throughout the space in which the lines are drawn and not only at the "lines" themselves; thus the number of lines of force per unit area (electric or magnetic) which might be drawn at any point is, no matter what the intensity of the field, infinite. In other words, while it is well to visualize a field by means of lines, the significance of these lines should always be kept in mind, and

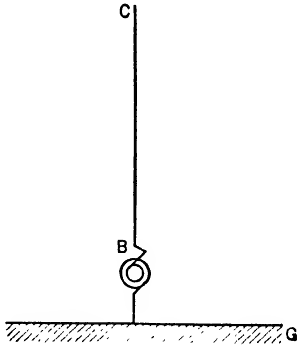


FIG. 3.—By connecting the lower end of the alternator to earth the semi-conducting surface of the earth takes the place of wire *DF* of Fig. 1 and enables the generator to send appreciable current up the wire *BC*.

it should never be forgotten that an electric field or a magnetic field is characterized by the existence of forces acting upon an electric charge or a magnetic pole respectively, and exists between the "lines of force" as much as it does at the point through which one of the lines passes.

Without attempting to go into the nature of a magnetic or an electric field we may say, however, that either field is accompanied by a strain in the material (ether or otherwise) present in the field, and that the forces manifested in the field may be considered as due to the elasticity of the material under stress, in much the same way that a stretched spring will exert a force because of the elasticity of the material tending to return the spring to its unstressed condition. Whatever the nature of the stresses and strains in an electric or a magnetic field, we may lay down certain well-known facts regarding them.

First.—An electric field or a magnetic field represents a definite amount of energy per unit volume of the field. It may be shown that this energy is, for the case of air, given by:[1]

$$W_m = \frac{H^2}{8\pi} \text{ ergs per cubic centimeter. } \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$W_e = \frac{(\mathcal{E}')^2}{8\pi} \text{ ergs per cubic centimeter}$$

$$= \frac{\mathcal{E}^2}{2.26\times 10^6} \text{ ergs per cubic centimeter, } \quad . \quad . \quad . \quad . \quad (2)$$

where W_m = energy in ergs per cubic centimeter of a magnetic field;

W_e = energy in ergs per cubic centimeter of an electric field;

H = intensity of the magnetic field in gilberts per centimeter, or in gausses;

$\mathcal{E}'$ = intensity of the electric field, in e.s.u. per centimeter;

$\mathcal{E}$ = intensity of the electric field in volts per centimeter.

Second.—A magnetic field in motion produces an electric field. This is nothing but the phenomenon of electromagnetic induction, for, the motion of the magnetic field induces an electromotive force, which must necessarily produce an electric stress or field. From Faraday's law, if:

H = intensity of magnetic field in gausses;

$\mathcal{E}$ = intensity of electric field in volts per centimeter;

V = velocity of magnetic field in centimeters per second;

$$\mathcal{E} = VH\times 10^{-8}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

[1] See J. J. Thomson, "Elements of Electricity and Magnetism," 1904, pp. 72 and 268.

Third.—An electric field in motion produces a magnetic field.

$$H = aV\mathcal{E}, \qquad (4)$$

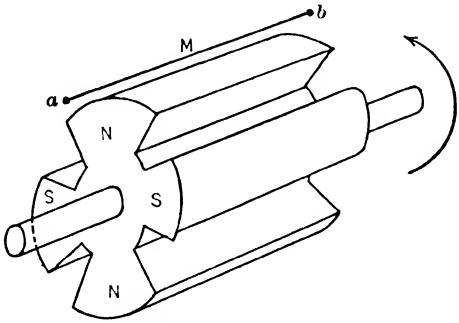


FIG. 4.—The poles of the revolving field induce an e.m.f. in the armature conductor M; it is important to note that the moving magnetic field will produce a difference of potential (hence an electric field) between points a and b *whether the conductor M is there or not.*

where a = a constant of proportionality.

This action of the moving electric field is not as easily realized as is that action by which a moving magnetic field generates an electric field. Every revolving field alternator furnished evidence of the latter effect. A revolving field (depicted in Fig. 4) generates an e.m.f. in one of the armature conductors, shown at M; one end of the conductor, a, becomes + and the other becomes −, this polarity reversing when a south pole takes the place of the north pole active in Fig. 4.

The point to be emphasized here is this—between the two points a and b (in the diagram located at the terminals of conductor M) the moving magnetic field produces a difference of electric potential or e.m.f. *whether the conductor M is there or not.*

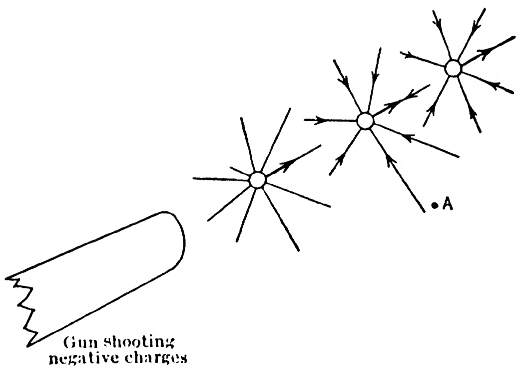


FIG. 5.—A stream of electrons shot from a gun is equivalent to an electric current and hence will produce a magnetic field at any point A; this magnetic field is really caused by the *moving electric fields of the electrons.* (Arrow heads pointing away from the electrons indicate direction of motion.)

The reciprocal relation,—a moving electric field producing a magnetic field—is not so well brought out in the action of ordinary electric machinery even though it is really the basis of every electromagnetic field. To illustrate the action let us imagine a gun shooting electrons at high speed, one following another rapidly in the same path, indicated in

Fig. 5. Each of these electrons will carry with it its electric field and so at any point in space near the stream of moving electrons (*A*—Fig. 5) there will exist a moving electrostatic field. But we know that there will

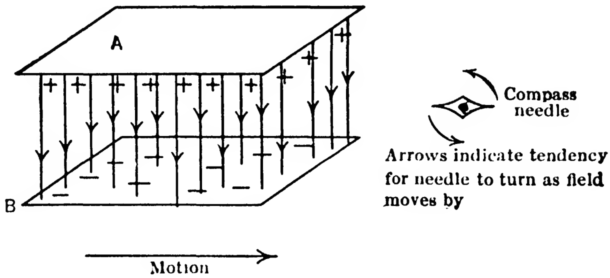


FIG. 6.—A compass needle, pivoted so that it is free to swing in the horizontal plane, will tend to set itself at right angles to the motion of the electric field as long as the electric field is moving past it, thus demonstrating the fact that a moving electric field generates a magnetic field, at right angles to itself and to its motion.

be at *A* a magnetic field (at right angles to the stream of electrons and also to the direction of the electric field) because *this stream of electrons is really an electric current*, the magnitude of current depending upon the number of electrons passing a given point per second. Thus if there were 6.28×10^{18} electrons passing a given point in 1 second, the stream of electrons would be equivalent to 1 ampere of current.

Upon exact analysis it will be found that the magnetic field at *A*, whether calculated from the well-known law of magnetic field surrounding a conductor carrying current, or from the relation given in Eq. (4), has the same value.

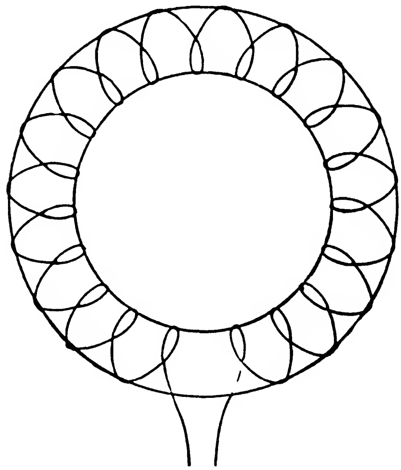


FIG. 7.—A toroidal coil is a good illustration of a *closed* magnetic field.

To illustrate this point by another simple experiment (easier to conceive than to carry out, however), we suppose two metal plates, *A* and *B*, Fig. 6, charged so that there is an electrostatic field between them as indicated. Suppose a compass needle, pivoted so as to be free to rotate in a horizontal plane and oriented in the same direction as the motion of the plates, is so placed that it is situated in the electric field as the plates move by. A magnetic force will act on the compass needle tending to make it place itself at right angles to the direction of motion of the electric field so long as the electric field is moving past, thus demonstrating the presence of a magnetic field as long as the electric field is moving past. The effect is easier to describe than to detect experimentally; the comparatively small velocities with which the electric field can be moved result in such low intensities of magnetic field that the electrostatic forces, brought into play by the induced electrostatic charges on the needle, completely mask the magnetic effects it is desired to show.

If, now, we consider a toroid such as that represented by Fig. 7 the magnetic field produced by it, when carrying a current, will be practically limited to the space within the toroid,[1] which space is not far removed from the conductors of the toroid. It is plain that if the current is reduced to zero the field collapses and in so doing it moves with respect to the conductors on the toroid and induces an electromotive force therein, thus producing an electric field. In this case, since the magnetic field is very near to the conductors, the motion of practically *all* of the magnetic field with respect to the conductors takes place at the same time, practically all of the energy given to the field is returned to the circuit, and no phenomena take place other than the well-known one of electromagnetic induction.

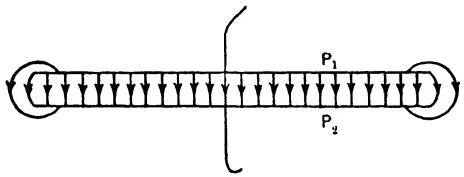


FIG. 8.—Two closely adjacent charged plates illustrate well a *closed* electric field.

Similarly Fig. 8 represents the two plates P_1 and P_2 of a condenser. The charging of the condenser produces an electric field, which is limited practically to the space between the plates. If the condenser plates are short-circuited, the electric field will collapse and here, as in the case of the toroid, since the electric field is very close to the plates, *practically all* of the energy in the field will be returned to the circuit.

If, on the other hand, we study the case of changing magnetic and electric fields which are distributed to comparatively great distances away from the seat of these fields, we meet with a new phenomenon, i.e.,

[1] This statement is not strictly true, because there is actually some magnetic field outside of the toroid as long as the current is changing. As this is an extremely small part of the total magnetic field, however, it may generally be neglected without much error. Furthermore the toroid with a continuously progressing winding as shown in Fig. 7, is magnetically equivalent to a single turn of wire, of diameter equal to the mean diameter of the toroid. This effect can be eliminated if the winding after progressing around the toroid once is then *wound back* over the toroid ending at the starting point. This means that the winding should have an even number of layers, the winding starting and ending at the same point on the periphery.

radiation of electromagnetic waves. Thus, consider the case of the two conductors of Fig. 9, to which there is connected the high-frequency alternator A. The voltage of the alternator is rapidly changing, and hence the charges on the conductors BC and DF are changing in value and in sign; the result is that a rapidly changing current is flowing through the wires, and the potential difference between the wires is also rapidly changing. In view of the above the conductors are producing a rapidly changing magnetic field, the lines of force of which are circles concentric with the wires and having planes perpendicular to the wires, and, in addition, a rapidly varying electric field, the lines of force of which are somewhat as shown in the figure.

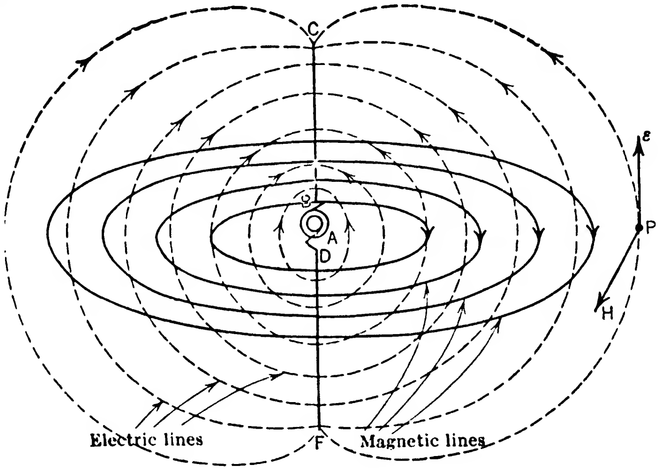


FIG. 9.—A pair of wires disposed as shown here, excited by a high-frequency alternator, illustrates what are called *open* magnetic and electric fields; these fields reach out (with appreciable strength) to distances greater than the dimensions of the circuit itself.

It is evident at first sight that this case is quite different from that of either the toroid or the two-plate condenser, for, while in the latter the field (either magnetic or electric) was existent only (at least practically) in a small space near the seat of the fields and all of it could quickly return its energy to the electric circuit, in the case of the antenna both the electric and magnetic fields extend outward in all directions and to distances as great, or greater, than the dimensions of the oscillating system. It is plain, then, that here we must consider the *time* necessary for the field to reach a certain point.

It is a matter of common knowledge that a disturbance or change of either an electric or a magnetic field travels through air or vacuum with the velocity of light. Consider then a point such as P at a distance d

from the antenna, and, for the sake of simplicity, in the equatorial plane.

Let f = frequency of alternator in cycles per second;
λ = wave length in centimeters.

We will first confine our attention to the electric field. Assume that the potential difference between the wires is on the point of starting from zero towards a maximum positive value and, therefore, the electric field is on the point of doing the same. The electric field at P will follow the variations of the potential difference between the wires, except that the variations at P will take place later, on account of the appreciable time necessary for the strain in the medium to travel the distance d. The line of action of the field at P will be vertical and represented by the line $\mathcal{E}$ in Fig. 9. We must not fail to remember at this point that an electric field means energy and therefore a certain amount of energy per cubic centimeter is present at the point P (due to the electric field) and the value of this energy is growing.

At some time, depending upon the frequency, the potential difference across the wires will reach a maximum and begin to diminish; and this will be followed, though after a definite time interval, by corresponding changes in the electric field at the point P, which will reach a maximum and then diminish. Since the electric field about the conductors is now decreasing it follows that the energy present in this field must be given back to the conductors, where it will appear as energy associated with the magnetic field set up by the current caused by the collapsing electric field. It is evident then, that the energy which had at first moved from the oscillator out towards P must now return towards the conductors. However, not all of the energy given to the electric field at the point P and beyond will reach the conductors before the potential difference across them begins to build up in the opposite direction, thus again sending out energy, in the form of an electric field in the opposite direction. There is then left [1] at the point P a certain amount of energy in the form of an electric field in the direction indicated by $\mathcal{E}$, Fig. 9, and this energy is unable to return to the conductors since they are already *sending out* more energy in the form of an electric field in the opposite direction to that of $\mathcal{E}$ Fig. 9.

The energy left at P or at any other point in the field cannot remain stationary, but must travel outward. This, however, could not happen were it not that, at the same time and for the same reason that a certain amount of energy is left detached at any point in the form of an electric

[1] In trying to picture radiation in this elementary fashion, statements are necessarily made which will appear, to the mathematical physicist, rather crude and artificial.

field, an equal amount of energy in the form of a magnetic field, acting in a horizontal direction as shown by H, Fig. 9, also remains at each point. These two energies, moving outward with the velocity of light, *can now sustain each other and are completely independent of the conductors wherefrom they issued.* For, it must be here remembered that, as pointed out on p. 870, a moving electric field produces a magnetic field and vice versa. That the energies of the two fields must be equal at all points and times follows from the fact that, if one were larger than the other, the difference could not exist by itself while moving in space;[1] for, in so doing, it would produce the other type of energy, hence it would either

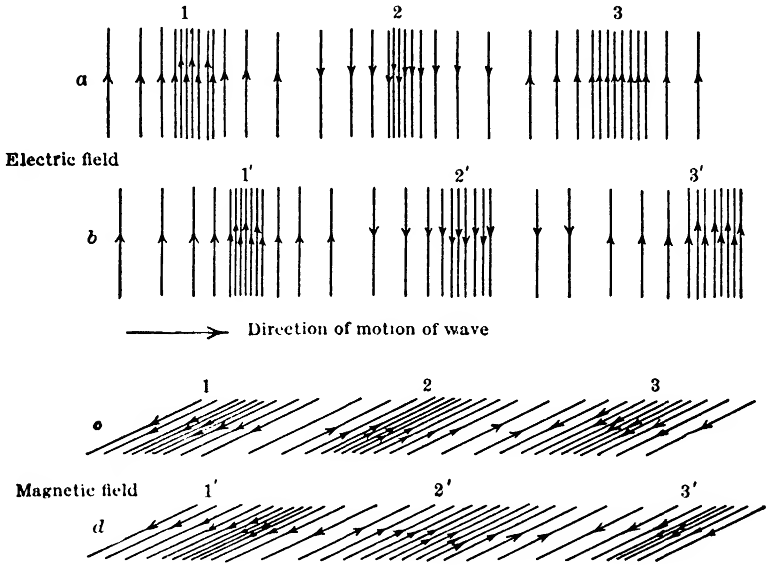


FIG. 10. Electric and magnetic fields associated with a wave of radiation at two successive instants of time; magnetic field c occurs with electric field a, dense magnetic field occurring where dense electric field is and vice versa. The magnetic and electric fields at any point are in *time phase* and *space quadrature.*

have one-half of itself transformed into the other type of energy, both of which would continue to move together, or else it would be absorbed by the medium or some conductor in the path.

In the brief discussion given above we have considered energy, in the form of a varying electric field acting in a *certain direction*, to be detached from the antenna; but, of course, in a similar manner energy is also detached in the form of an electric field acting in the opposite direction, so that the electric field, equivalent to the energy which is detached from

[1] This same idea holds good for water waves also; when the two types of energy associated with the wave become unequal the wave "breaks."

the antenna, is, at any point, varying continually in value and direction similarly to the antenna current. If this is harmonic the variation of the detached field will at any point be harmonic. Furthermore, since it takes time for the field to travel any distance, it follows that the phase of the field will be different at each point; in other words we shall, as already outlined in Chapter IV, have a wave constituting an electromagnetic disturbance in the medium, so that while at a certain instant of time the electric field in a certain portion of the space may be represented by (*a*) Fig. 10, the maximum intensities occurring at 1, 2, 3, a little later the electric field will appear as at (*b*), the maximum intensity now occurring at 1′, 2′, 3′, and the wave of electric disturbance having traveled the distance from 1 to 1′. The above also applies to the magnetic field, the latter acting in a direction perpendicular to the electric field, and both moving together in a direction perpendicular to both. Thus, at a certain instant the magnetic field, in the portion of the space for which the electric field is given in Fig. 10 (*a*) and (*b*), will be represented by (*c*) and (*d*) Fig. 10, which will correspond to (*a*) and (*b*), respectively. Since, as already stated, a moving electric field produces a magnetic field proportional to its own intensity, and vice versa, it follows that the intensities *of the two fields are in time phase,* though in space quadrature.

From p. 869 we have

$$\mathcal{E}=VN\times10^{-8}. \qquad (3)$$

$$H=aV\mathcal{E}. \qquad (4)$$

Since in the case under discussion the electric field is produced by the motion of the magnetic field and the latter is produced by the motion of the electric field, it follows that the H and $\mathcal{E}$ of Eq. (3) are the same as the H and $\mathcal{E}$ of Eq. (4), and may be substituted therein. Thus, from (3)

$$H=\frac{\mathcal{E}}{V}\times10^{8},$$

and substituting in (4)

$$10^{8}\times\frac{\mathcal{E}}{V}=aV,$$

or

$$a=\frac{10^{8}}{V^{2}},$$

and of course the second equation becomes the same as the first, i.e.,

$$H=\frac{10^{8}}{V^{2}}\times V\mathcal{E},$$

or

$$\mathcal{E}=VH\times10^{-8}.$$

In our case V, the velocity of magnetic field and of the electric field, is the velocity of light; since the velocity of light is 3×10^{10} cms. per sec. we may substitute this in Eq. (3) and thus obtain

$$\mathcal{E}=300\,H. \qquad (5)$$

From this relation we conclude that a magnetic field, of intensity represented by one gauss, when moving with the velocity of light, generates an electric field, at right angles to itself and to the motion, of the intensity of 300 volts per centimeter.

From our brief qualitative consideration of the phenomena around an antenna carrying an alternating current it follows that we may consider the space about an antenna as occupied by two components of electric and magnetic fields. One of these is continually moving backwards and forwards from the antenna, so that energy is alternately given to it by the antenna and returned by it to the antenna. Because of this backwards and forwards motion the average displacement of this component of either field is zero, and may therefore be known as the " stationary " component, also known as the " induction " field; it is this component with which students of electrical engineering are more familiar, in so far as it is this which produces the well-known phenomena of induction (either magnetic or electrostatic).

The other component of either field is the one which, once having left the antenna, is prevented from returning to it and is thereafter urged away from the antenna and continually travels outward from this with the velocity of light. This component, while fundamentally of the same nature as the stationary component, it is yet very different in so far as it is completely detached from the antenna. It is known as the " radiation " field and represents energy which is transferred by the antenna to the medium around it, which energy is never again returned to the antenna. At any given point in space the induction fields (magnetic and electric) are out of time phase by 90°; at the instant one of them is a maximum the other is zero. The two components of the radiation field, on the other hand, are *in time phase* with one another; at a given point in space the two components rise and fall simultaneously.

Both of the above types of the fields, i.e., induction and radiation, exist at any point at any distance from the antenna; but at points near it the induction field is much greater than the radiation field, while at points far away from the antenna the radiation field is so much greater than the induction field that the latter may be said not to exist. The reason for this is that the amplitude of the induction field at any point varies inversely as the square of the distance while that of the radiation field varies inversely as the first power of the distance.[1] Thus any effects of the field near the antenna are mostly due to the induction field, while at great distances from the antenna they are mostly, and practically wholly, due to the radiated field. Hereafter when speaking of the field about an antenna we will, unless otherwise specified, mean to refer to the radiation field, since this is the one by means of which intelligence is transmitted to great distances without wires.

The radiation component of the field is most important when the currents in the antenna are of high frequency; but it must not be understood that no radiation component exists at low frequencies; for a radiation component exists at any and all frequencies. Since, however, the very reason for the existence of such a component is to be found in the inability of the energy given to a rapidly changing field to return in its entirety to the circuit giving out the energy, it follows that, for slowly changing fields, this effect is negligible, and hence the radiation field is practically non-existent and is never considered in low-frequency circuits.

It must not be concluded, as a result of the foregoing elementary analysis, that there are actually two different fields to be considered, one induction and one radiation. At any point in space in the neighborhood of a radiating system, the magnetic and electric fields both go through harmonic variations. Close to the radiator these two fields are both of intense amplitude (comparatively) and they are very nearly 90° out of time phase; as the distance from the oscillator increases both of these fields fall off in intensity and with increasing distance the phase difference is diminished until at very great distances (perhaps a wave length from the radiator) the electric and magnetic field are in phase.

This point is illustrated in Fig. 11; in (*a*) are shown the magnitudes of the actual electric and magnetic fields at various distances from the radiator (points supposed in the equatorial plane) and in (*b*) and (*c*) are shown the induction and radiation components of the actual field. The electric and magnetic fields are, for all conditions, in *space quadrature* (i.e., at right angles with one another) but the time phase between the two fields varies as indicated in the diagram.

The above discussion has been given on the basis of the antenna and counterpoise represented by Fig. 1, but it applies equally well no matter what the counterpoise and no matter what the nature of the source which produces alternating currents in the antenna.

A Simple Analogy to an Antenna.—A very simple picture, which contains much of the characteristic behavior of the radiating antenna, supposes a large sheet of rubber held tight at its distant boundaries by clamps of some sort. We imagine this stretched rubber sheet held horizontal and

[1] See "Principles of Radio Transmission and Reception with Antenna and Coil Aerials," by J. H. Dellinger, Proc. A.I.E.E., Oct., 1919.

to its middle point a vertically held stick is attached, by a tack. The stick is now made to execute simple harmonic motion in the vertical direction, of course, carrying along with it the central point of the rubber sheet.

Any point on the rubber sheet close to the place where the stick is attached will be lifted up and down following the motion of the stick, but distant points on the rubber sheet will experience no motion if the motion

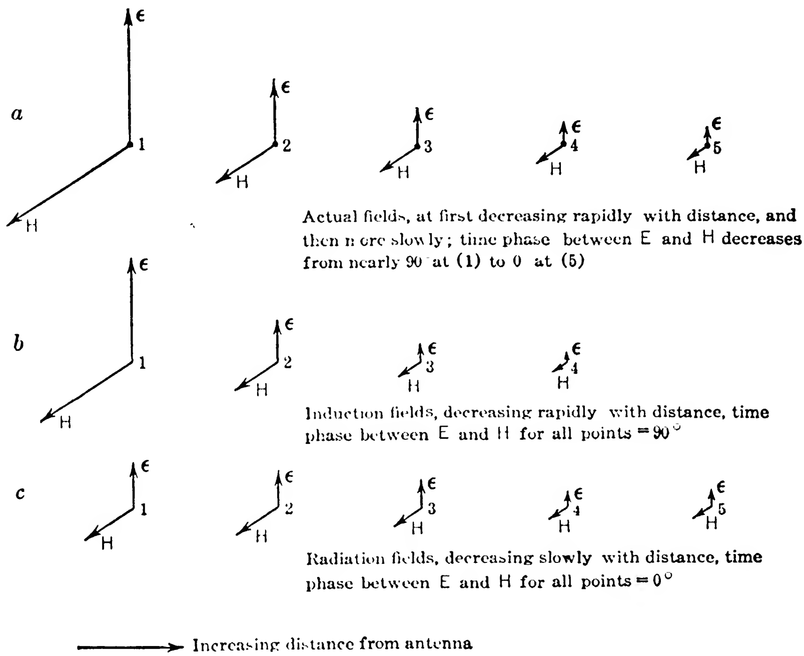


FIG. 11.—Actual electric and magnetic fields at different points in the vicinity of an antenna shown in *a*; these actual fields decrease in magnitude with distance from the antenna and at the same time come more nearly into time phase. The components of the fields which are 90° out of phase (in time) are called the induction fields, shown at *b*, while the components which are in time phase with each other constitute the radiation fields; the latter decrease with the first power of the distance while the former decrease with the second power of the distance.

of the stick is reasonably small. However, when the oscillatory motion of the stick is increased in frequency it will be seen that waves are sent out over the rubber sheet and these waves will set into motion the more distant parts of the rubber. That is, points on the sheet so far distant from the stick that steady displacements of the stick produces no perceptible displacement of the rubber sheet will be affected, and move up and down, if the stick is moved up and down with sufficient rapidity.

The distortion and motion of the rubber sheet, near the stick when the stick is moved up and down slowly, correspond to the induction fields of the antenna whereas thè distortion and motion of the rubber sheet produced by the waves, correspond to the radiation fields.

Radiation from the Electron Viewpoint.—The foregoing simple explanation of radiation has been developed along the well-known lines used in ordinary engineering texts; it is, in a sense, unsatisfactory in that it suggests two sets of fields around an antenna, "induction" and "radiation" fields. It is, however, possible to construct a satisfactory picture of radiation from the electron viewpoint, using the conception of Faraday that when a charge moves it carries along its electric field and when the position of an electron is changed a disturbance travels out over its field, with the velocity of light.[1]

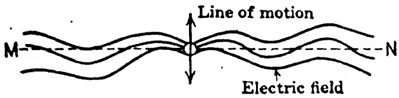


FIG. 12.—We may get a reasonable picture of self and mutual induction as well as radiation by imagining the radial electric field of the electron to actually carry waves, as the electron is oscillated back and forth.

The electric field surrounding a stationary electron varies in strength in accordance with Coulomb's law, that is, inversely with the square of the distance. Suppose in Fig. 12 the electron oscillates with simple harmonic motion in a vertical line, carrying its electric field with it. Considering then the electric field in the equatorial plane, M–N it is evident that sinusoidal waves will travel out from the electron with the velocity of light. They have the same velocity as light because they travel in the same medium as does light, namely the electron's field.

These transverse waves are sent out in all directions in the equatorial plane so that if the eye could see the electron's field it would resemble the rubber sheet of the previous section, when the stick was being moved up and down rapidly. Waves are, of course, sent out in other directions than those embraced by the equatorial plane, but we shall confine our attention to this plane as it is here that the maximum radiation is sent out.

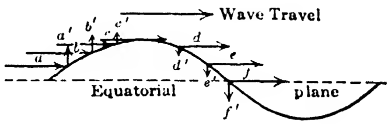


FIG. 13.—The waves set up according to the idea of Fig. 12 may be split up into their horizontal and vertical components. The horizontal component is balanced out by the field of the positive electrons, but as these do not move, the vertical component, of the waves of the electron's field, is left unbalanced.

Analyzing now this wavy electric field we picture a portion of it in Fig. 13. The actual direction of the field varies from point to point at a

[1] J. J. Thomson uses this conception in the explanation of X-rays. See his "Conduction of Electricity through Gases," pp. 657 et seq.

given instant, or at a given point the field varies in direction from instant to instant. This field can always be imagined as made up of two components, one in the equatorial plane and the other perpendicular to it; as the wave passes a given point the equatorial component of the field remains constant but the perpendicular component varies harmonically as the wave passes by.

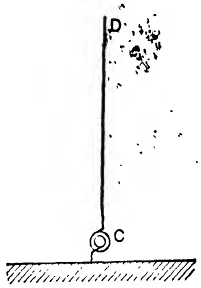


FIG. 14.—A simple vertical wire grounded antenna.

We have previously shown that an electric field moving transversely to its direction sets up what we call a magnetic field perpendicular to itself and to the direction of its motion and we see then that the transversely vibrating electric field of Fig. 13 will set up a magnetic field perpendicular to the plane of the figure. This magnetic field depends upon the strength of the electric field and upon the rapidity of its motion. The magnetic field associated with the transversely moving equatorial electric field is the magnetic induction field and the magnetic field associated with the outwardly traveling vertical component, a', b', c', d', etc. of the electric field is the magnetic radiation field.

The transverse velocity of the equatorial electric field is very slow (fraction of 1 centimeter per second) whereas the velocity of the vertical electric field is equal to that of light. The strength of the equatorial field, however, varies inversely with the square of the distance so that at a short distance from the antenna the induction field becomes small compared to the radiation field. The transverse amplitude of the wave of Fig. 13 is only about 1×10^{-12} as great as its length, so that the electric field actually has no such evident "kink" as that given in Fig. 13.

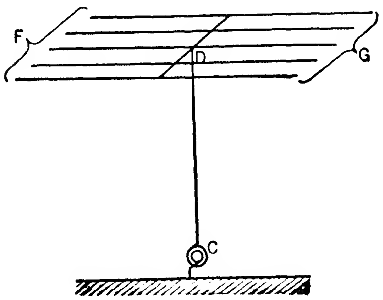


FIG. 15.—If the antenna has a considerable network of wires above, the current in wire CD will be nearly the same in amplitude at all points of the wire.

Radiated Field at any Distance from Antenna.—Before taking this up we will discuss very briefly the distribution of the current in an aerial. In the case of the aerial shown in Fig. 14 it is plain that, since the current in the wire CD flows only to charge the capacity of the wire, the effective value of the current at C will be a maximum and at D it will be zero, for the current at C represents the electricity flowing through that point which goes to charge the rest of the wire, while at the point D no electricity whatever flows, since there is nothing to which it can flow.

On the other hand, if a metallic plate or a system of conductors be arranged at the end of the wire CD, as at FG, Fig. 15, and if FG has a very large surface as compared with that of CD, it is plain that the effective value of the current at D will then be only slightly smaller than that at C, since the current at D must be such as to charge the large capacity FG. Under such conditions the effective values of the current in all parts of the vertical wire of the antenna will be sensibly equal and will be considered as such in the following discussion. Consider, then, the aerial as represented in Fig. 16, where the counterpoise is represented by a horizontal system of conductors, $F'G'$, laid near the ground but insulated therefrom, being in every way similar to the system of conductors at the top of aerial FG. The current in the vertical part of the

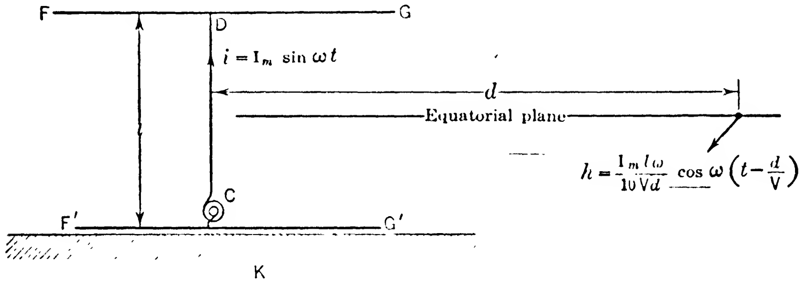


FIG. 16.—With uniform current in wire CD the magnetic field due to this current, is given as above.

aerial CD will be assumed to have the same effective value throughout, so that at every point of CD we will have for the equation of the current:

$$i = I_m \sin \omega t,$$

where i = instantaneous value of aerial current in amperes;
I_m = maximum value of aerial current in amperes;
ω = angular velocity of current vector in radians per second;
t = time in seconds.

Under these conditions it may be shown that the radiation component of the magnetic field, at any point in the equatorial plane of the aerial, is given by:[1]

$$h = -\frac{l\omega I_m}{10Vd} \cos \omega \left(t - \frac{d}{V}\right), \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

[1] The normal development of the equation of radiation field requires more mathematical background than the average radio engineer possesses and it is not thought well to introduce it here; a short analysis of the problem is given in Berg's "Electrical Engineering, Advanced Course," pp. 278 et seq. Eq. (20), p. 289, of that volume is

where h = instantaneous value of magnetic field in gausses;
l = height of antenna in centimeters;
V = velocity of light in centimeters per second;
d = distance of point in question from antenna in centimeters.

The above equation shows that the radiation magnetic field is a function similar to the antenna current (in this case a harmonic function), and that the phase angle is different for points at different distances since this angle is equal to $\omega d/V$. Substituting $\omega = 2\pi f$ and $V = \lambda f$ we have:

$$\text{phase angle} = \frac{2\pi d}{\lambda} \text{ (}\lambda\text{ being measured in centimeters),}$$

whence, Eq. (6) becomes:

$$h = -\frac{2\pi l I_m}{10\lambda d} \cos\left(\omega t - \frac{2\pi d}{\lambda}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Since the "radiation" component of the electric field bears a fixed relation to the "radiation" component of the magnetic field as given by Eq. (5), we may write:

$$\epsilon = 300h = -\frac{600\pi l I_m}{10\lambda d} \cos\left(\omega t - \frac{2\pi d}{\lambda}\right), \quad . \quad . \quad . \quad . \quad (8)$$

where ϵ = instantaneous value of electric field in volts per centimeter.

From Eqs. (7) and (8) we obtain the effective values of the radiation components of the two fields. Thus, if:

H = effective value of magnetic field in gausses;
$\mathcal{E}$ = effective value of electric field in volts per centimeter,

$$H = \frac{2\pi l I}{10\lambda d}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

$$\mathcal{E} = \frac{600\pi l I}{10\lambda d}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

where I = effective value of the current in aerial, in amperes.

Eqs. (9) and (10) show that the effective value of either field varies directly with the effective value of current in the aerial and with the height of the aerial and inversely as the wave length and distance from the aerial.

Now consider the case represented by a loop of wire as shown in Fig. 17. Assume, similarly to the previous case, that the capacity of the condenser, P_1 P_2, is so large as compared with the distributed capacity

the same as the Eq. (6) given above, it being noted that Berg has used h to signify one-half the length of the oscillator.

of the loop *CDGF* that the latter has the same effective value of current throughout its length.

Consider the magnetic field at a point P at a distance d from vertical wire CD and a distance $s+d$ from vertical wire GF. Assume the positive direction of current to be as shown by the arrows. Then the field at P must be equal to the difference of the field due to CD and that due to FG.

Let h_1 = instantaneous value of magnetic field at P due to CD;
h_2 = instantaneous value of magnetic field at P due to FG

Then, from Eq. (7) we have

$$h_1 = -\frac{2\pi l I_m}{10\lambda d}\cos\left(\omega t - \frac{2\pi d}{\lambda}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

$$h_2 = +\frac{2\pi l I_m}{10\lambda(d+s)}\cos\left(\omega t - \frac{2\pi(d+s)}{\lambda}\right). \quad . \quad . \quad . \quad . \quad (12)$$

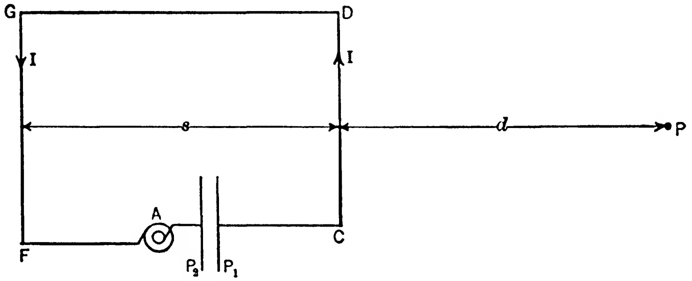


FIG. 17.—In the case of a coil antenna the magnetic field at P is calculated by adding the two fields due to CD and FG, it being noted that the currents are opposite in direction.

It will be noted that the amplitude of these two fields is practically the same, since, for great distances, d is practically equal to $d+s$, but the phases of the fields are different by the amount

$$\pi - \frac{2\pi s}{\lambda} \text{ radians.}$$

The resultant field (h) is given by:

$$h = h_1 + h_2 = -\frac{2\pi l I_m}{10\lambda d}\cos\left(\omega t - \frac{2\pi d}{\lambda}\right) + \frac{2\pi l I_m}{10\lambda(d+s)}\cos\left(\omega t - \frac{2\pi(d+s)}{\lambda}\right)$$

$$= -\frac{2\pi l I_m}{10\lambda d}\left\{\cos\left(\omega t - \frac{2\pi d}{\lambda}\right) - \cos\left(\omega t - \frac{2\pi(d+s)}{\lambda}\right)\right\}$$

$$= -\left(\frac{4\pi l I_m}{10\lambda d}\sin\frac{\pi s}{\lambda}\right)\sin\left(\omega t - \frac{2\pi}{\lambda}\left(d + \frac{s}{2}\right)\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

From which the effective values of the resultant magnetic and electric fields are given by

$$H = \frac{4\pi l I}{10\lambda d} \sin \frac{\pi s}{\lambda} \quad (14)$$

$$\mathcal{E} = \frac{1200\pi l I}{10\lambda d} \sin \frac{\pi s}{\lambda} \quad (15)$$

The vector addition of H_1 and H_2 by which Eq. (14) is obtained is shown in Fig. 18. These equations show that the effective value of the resultant field is equal to twice that due to either wire multiplied by the sine of an angle which varies with the distance between the two wires.

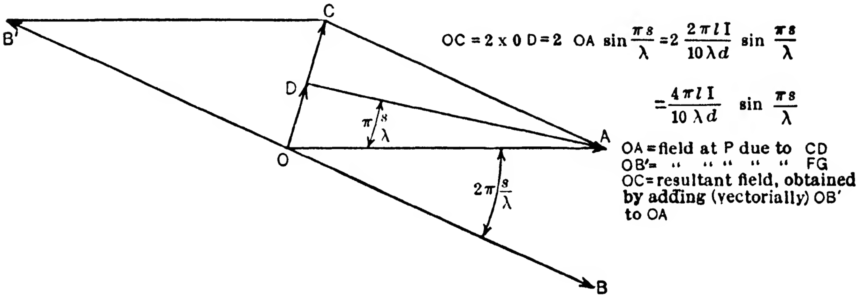


FIG. 18.—The field due to wire CD is shown by vector OA; that due to FG is shown by OB' nearly 180° out of phase with OA. The actual field is obtained by adding vectorially OB' and OA, it being noted that they differ in phase by $(\pi - 2\pi s/\lambda)$.

Thus if $s = \lambda$

$$\sin \frac{\pi s}{\lambda} = \sin \pi = 0$$

and if $s = \lambda/2$,

$$\sin \frac{\pi s}{\lambda} = \sin \frac{\pi}{2} = 1.$$

It may then be seen that if the distance between the two wires of the loop is exactly equal to one wave length, the resultant field at all points in the plane of the loop is zero, while if the distance between the two wires is one-half a wave length the resultant field in the plane of the loop is equal to twice that of one wire. In other words the resultant at any one point is due to fields of the same amplitude but different phase, the latter depending upon the distance between the wires, since in one case the field has to travel a greater distance than in the case of the other wire. Thus, if the two wires were close together the resultant field at any point would be zero.

Again, if a point be chosen such as Y, Fig. 19, in a plane perpendicular to the plane of the loop and equidistant from both wires, it is plain that the fields at Y due to wires CD and FG must be 180° out of phase, since they have to travel the same distance, and the result is that the resultant field at Y is zero. For points other than those such as point Y of Fig. 19 and point P of Fig. 17 the maximum value of the field for a certain distance from the aerial varies from zero at Y to a maximum at P.

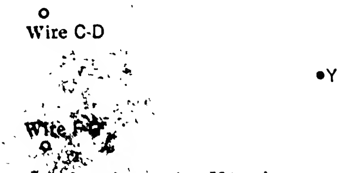


Fig. 19.—At a point Y in the equatorial plane of the coil, equidistant from both wires CD and FG the radiation field is zero.

If a curve were plotted to polar coordinates, showing the effective values of the magnetic field intensity at all the points around the circumference of a circle having the loop as a center, we would obtain a diagram as shown in Fig. 20, the intensity

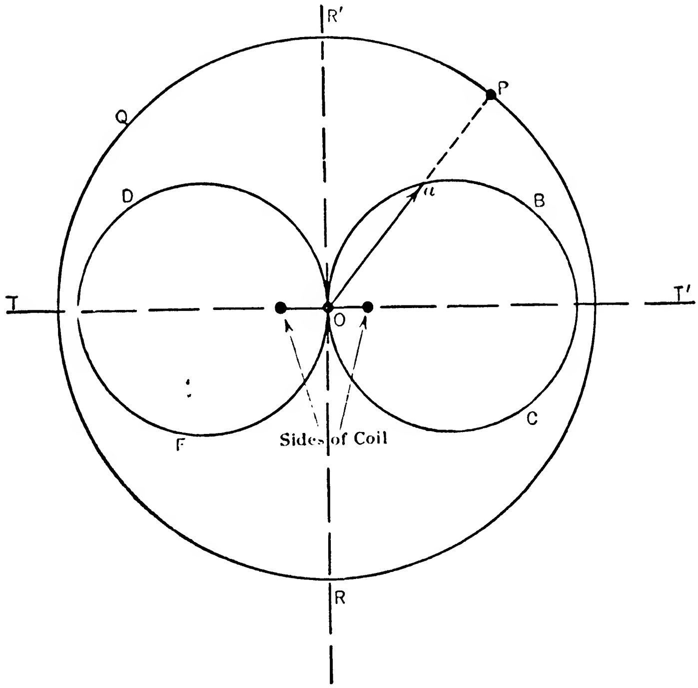


Fig. 20.—The distribution of radiation field in the equatorial plane of a coil antenna.

of the field at any point P along the circumference PQR being represented by the line Oa. It may be easily shown that the intensity of the field varies harmonically from zero at points R and R' to maxima at points T

and T', and, therefore, the curves OBC and ODF should be circles with a diameter equal to the intensity of the field in the direction TT'. Such a loop will, then, radiate most energy in the direction TT' in the plane of the coil and practically none in the direction RR'.

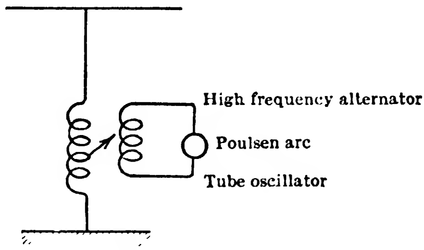


FIG. 21.—Excitation of antenna by magnetic coupling to generator.

Methods of Producing Current in the Antenna.—So far we have discussed simple antennas energized by means of an alternator placed directly in series with the aerial; but it has already been stated that an antenna may be energized by means other than this one. Thus the diagrams of Figs. 21, 22, and 23 give various methods of energizing the antenna, all of which methods have already been studied. Fig. 21 shows the alternator inductively coupled to the antenna circuit, instead of having the alternator directly in the antenna circuit. This has the advantage of eliminating some of the harmonics of the alternator, so that the current in the antenna is nearly sinusoidal. It is to be noted that instead of a high-frequency alternator, a tube generator or a Poulsen arc may be used, and, in every case the antenna current will be nearly harmonic and undamped.

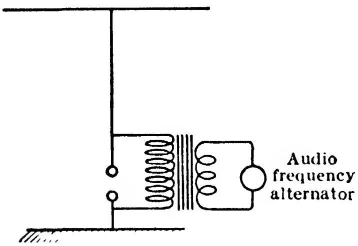


FIG. 22.—Simplest scheme for spark telegraphy excitation.

On the other hand, the arrangements of Figs. 22 and 23 are meant to produce trains of damped currents in the antenna. In Fig. 22 the spark gap is directly in the antenna, while in Fig. 23 the spark gap is placed in the so-called closed oscillating circuit. The disadvantage of placing the spark gap directly in the antenna is due to the fact that such a gap has considerable resistance, especially in the case of high-powered, high-voltage sets where the gap distance must be large, and when so used will make the decrement of the antenna proper very high, which is objectionable. Hence, with very few exceptions, i.e., low-power sets, all modern

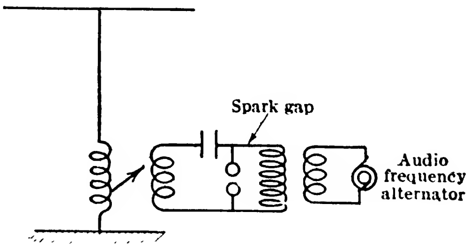
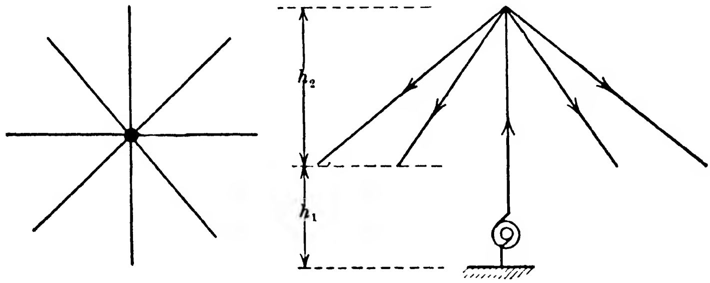


FIG. 23.—Ordinary scheme of excitation for spark telegraphy.

sets place the spark gap in the closed oscillating circuit, instead of in the antenna.

The methods outlined above are only typical, and there are several other ways of energizing the antenna, which have already been taken up in Chapters IV, VI and VIII.

Various Types of Antennas.—It was stated on p. 881 that if a single vertical wire be used for an antenna the effective value of current at the base of the wire will be maximum, while at the top it will be zero. Since, the intensity of the field radiated by an antenna is directly proportional to the current therein (on the basis of a constant current throughout the antenna) it is plain that a single vertical wire with non-uniform current will not radiate as well as if it had a capacity at the top end, when the current would be more nearly uniform, and also larger, for a given voltage impressed by the power source. Such a capacity is used at the top end of an antenna in actual practice, the capacity being in the form of wires stretching outward from the antenna proper. Depending on how these wires are arranged we have several types of antennas, known, as: umbrella,



Umbrella type

FIG. 24.—Umbrella antenna.

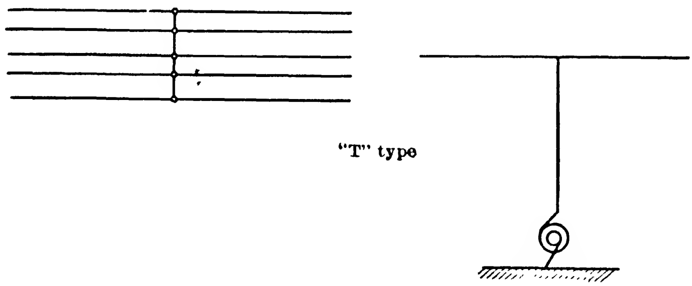


FIG. 25.—Antenna of the T type.

T-type, inverted L-type, "Fan or Harp" type, "Multiple-tuned" type, "Coil" type, and Wave-antenna type.

These various types are shown in the conventional diagrams of Figs. 24–29, respectively.

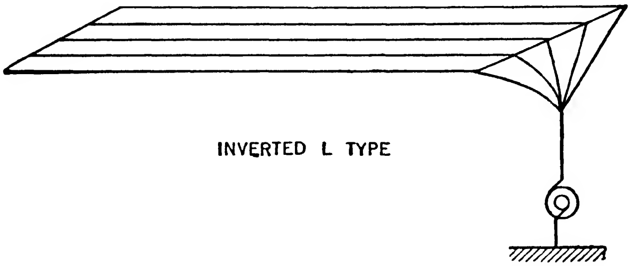


FIG. 26.—Antenna of the inverted L type.

The characteristics of these various types of antennas will now be discussed.

Umbrella Type.—Since the top wires are symmetrically arranged all around the central radiator it is easily inferred that at a given distance from the aerial the intensity of the field all around the radiator is the same, that is, the curve of distribution of field intensity around the radiator should be a circle. It must be noted that, while in the case of a single

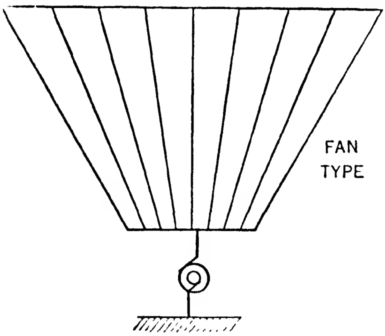


FIG. 27.—Fan or harp antenna.

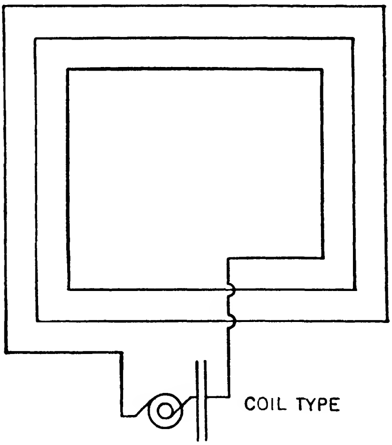


FIG. 28.—Coil antenna.

wire or of a Hertzian double (Fig. 1, p. 867) for a radiator, vertical wires only are used to radiate energy, in the inverted L-type aerial the horizontal top wires radiate a certain amount of energy in the direction perpendicular to the wires themselves. Thus, while in the former case we would likely find the intensity of the field directly over the top of the antenna practically *nil*, in the latter case (the inverted L-antenna) the field directly over the top might be of considerable strength and is successfully used to signal to aeroplanes, even though they be directly over the antenna.

On the other hand, the top spreaders of an umbrella-type antenna subtract, to a certain extent, from the radiating ability of the central

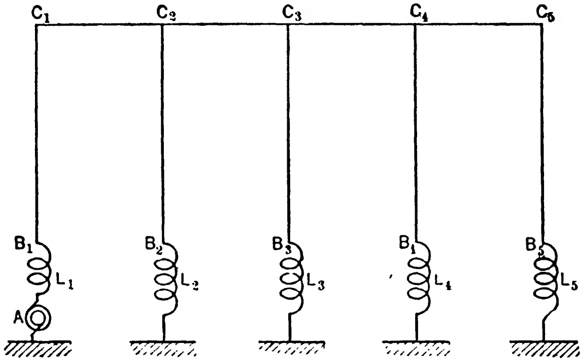


FIG. 29.—Multiple-tuned antenna.

vertical wire, for the following reason. We have already stated that the ability of the vertical wire or wires as a radiator of energy depends upon the fact that in view of its very configuration it is capable of setting up a field, magnetic and electric, which extends to very great distances from the wire and is not mainly confined to a space near the wire; thus we have seen that in the case of the two-plate condenser of Fig. 8 the energy stored in the electric field is mainly in the space between the plates, which constitute a "closed electric circuit." If we were to imagine an umbrella aerial with a very large number of spreaders reaching nearly to ground, as shown in Fig. 30, it is plain that these spreaders would act like one plate, and the ground like the other plate, of a closed electric circuit, and practically no energy could then be radiated because the electric field of the antenna would, for the most part, be confined in the space under the spreaders, and there would be little likelihood of any energy being detached from the antenna. The radiation from such an arrangement would of course, be very small. In an actual umbrella-type antenna the spreaders

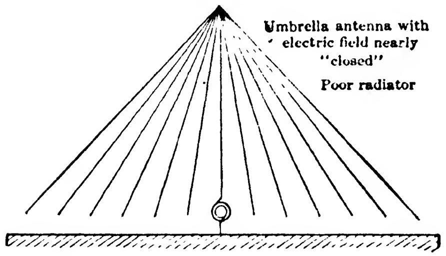


FIG. 30.—Umbrella antenna of this form is a poor radiator; the spreaders come too low.

do not reach to anywhere near ground, hence they do not seriously interfere with the radiation, though they do so to a certain extent.

Another reason for the spreaders interfering with radiation is to be found in the fact that, at any time, the direction of the current flowing through the vertical wire is opposite to that flowing in the spreaders; that is, if the current in the vertical wire is upward that in the spreaders is downward. In the extreme case where the spreaders might be considered as being close to the vertical wire, as in Fig. 31, the portion of the vertical wire AB would be seriously limited in its radiating action, since the action of the current in the vertical wire is opposed by that of the spreaders.

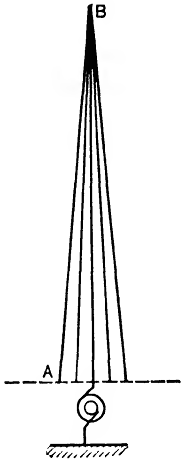


FIG. 31.—If the spreader wires are brought down very close to the antenna proper the radiation is practically zero.

However, with spreaders as generally arranged, the total interference with radiation from the vertical wire is less than what the spreaders contribute towards increasing the radiation through causing a more uniform current and, for the same voltage, a larger current, to flow through the vertical wire. Several large antennas of this type have been used for long-distance transmission. In the smaller sizes they are very convenient for portable sets where the spreaders, anchored through insulating clamps to the ground, serve the purpose of holding the central support, in addition to increasing the capacity at the top end of the vertical wire.

The effect of the spreaders may be looked upon as if the height of the vertical wire had been diminished and it may be shown that the "effective height" of an umbrella antenna is approximately given by

$$h=h_1+\frac{h_2}{3},$$

where h = effective height;

h_1 and h_2 = the heights as represented in Fig. 24.

Fig. 32 shows the arrangement of spreaders, vertical wire and vertical support, insulation, etc., for a small umbrella aerial, where *aaabbb* represent insulators and *cd* the vertical radiating wire. The spreaders are generally made long enough to extend about two-thirds the length of the mast. A large piece of wire netting (called a ground mat) may serve as a counterpoise for the oscillating system.

" T " Type.—Since the top wires are on this type unsymmetrically arranged, i.e., extending outward from the vertical wire in two directions only, it would seem at first as if the field produced by such an aerial would not be quite the same all around the antenna. This is probably the case at comparatively short distances from the aerial, but it is not found to be so at large distances away, in view of the tendency of the field to become uniform as it spreads out in all directions away from the aerial.

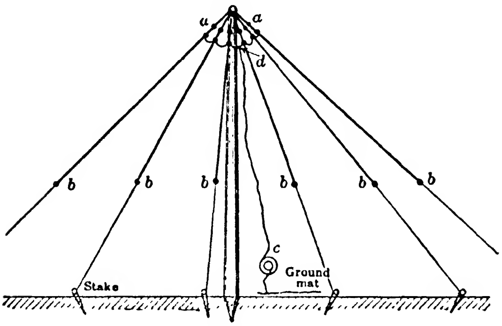


FIG. 32.—Showing construction of a small portable umbrella antenna; *a-a-a-*, *b-b-b-*, etc., are insulators. To get a fair "ground" a net of copper wires is generally spread out underneath the antenna, the lower side of the generator being connected to this.

Here, as in the case of the umbrella type, some energy is also radiated in a direction directly above the antenna. Antennas of this type are very widely used on shipboard where the flat top is easily suspended between two masts; also for portable sets an aerial of this type is easily suspended between two trees.

Inverted " L " Type.—The main difference between this type and the " T " type is that the " L " type has a more pronounced directional effect,

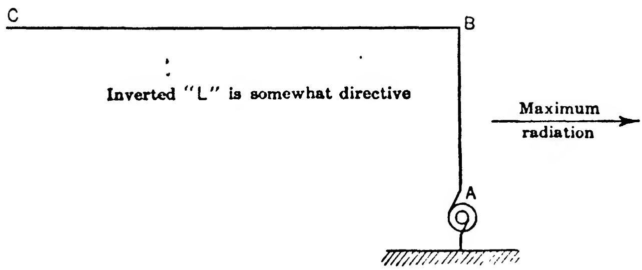


FIG. 33.—An inverted L antenna is somewhat directional giving maximum radiation in the direction shown above.

that is, it is capable of producing a greater intensity of radiation in one direction than in any other. This action, in the case of the " L " antenna, is not yet very fully understood and it is by some stated to be too small to actually claim for this type of antenna directional ability. However, this type of antenna has been used by the Marconi Co. for the large transatlantic stations and has actually been found to develop, even at considerable distance from it, a field stronger in the direction of the arrow, Fig 33, than in any other. This effect depends especially upon the length of the flat top, *BC*, as compared with the vertical wire *AB*. The longer *BC* is made relative to *AB* the greater seems to be the directional effect of the antenna. It is probable that this is due to an interfering action of some sort, between the currents in the vertical and horizontal portions of the antenna, which occurs on one side of the antenna to a much greater extent than on the other. This would, of course, take place to a greater extent the larger the horizontal portion of the aerial relative to the vertical portion. The Clifden station of the Marconi Co. has a vertical portion about 60 meters high and a horizontal portion about 2000 meters long; it is said to have a large directional effect. In " L " type aerials as used on board ships, however, the horizontal portion is never very much longer than the vertical portion and it is doubtful if in this case any appreciable directional effect is present, even at short distances from the aerial.

A directional effect is noted in the case of aeroplanes carrying a long vertical wire weighted at one end and dangling beneath the aeroplane proper; this wire, when the aeroplane is in flight, bends somewhat as shown in Fig. 34 and very much in the form of an inverted " L" aerial. The greatest field intensity is in the direction of flight or away from the horizontal portion of the aerial; in this case the framework of the plane is the counterpoise.

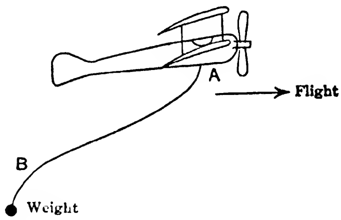


Fig. 34.—An aeroplane antenna is directional, sending out most power in the direction of flight.

Not only are inverted " L " aerials used in large transatlantic stations but they are also favorites on board ships, where they are as easily installed as the " T " type. They are also widely used for small stations and by amateurs. When used on board ships it is customary to install them where the distance between the masts does not exceed about 30 meters; for over 30 meters the " T " type is used.

" *Fan Type* " *Aerial*.—In this case a large number of vertical or nearly vertical wires in multiple are used, the top ends of these wires being perhaps free, that is, not connected electrically to any other wires. In such an arrangement the effective value of the current at the base of each wire is a maximum while it is zero at the top, or, in other words the current distribution is very far from uniform, and in this respect the arrangement is objectionable. On the other hand, the capacitance of the aerial is very large, in view of the capacitance of so many wires connected in multiple; in fact the whole arrangement may be thought of as a single wire having a capacity equal to that of all the wires. The current through the combination of all the wires may, because of the large capacity, be made very large without excessively high voltages. Another advantage of this type of aerial as compared with the others previously discussed is that there are no horizontal or inclined wires to interfere with the radiation from the vertical wire. As a matter of fact such an arrangement is considered one of the best and most efficient radiators. In spite of this, however, the fan type is not very widely used because of the difficulty of installing it, especially in the case of ships where such a multitude of vertical wires would be in the way of some of the projecting parts of the ship.

"*Multiple-tuned*" *Type.*—This type of antenna is of great value when the location of the radio station has been determined by reasons other than the electrical behavior of the antenna. Thus if the antenna is in swampy country no trouble is experienced in getting a good ground, but when the station is located on the top of a huge sand bank (as is the case with one of America's large stations) then the ground resistance of the antenna becomes of paramount importance.

Sand permits water to run off immediately after a rain and this action carries off all soluble matter. Thus the top of the sand bank is dry and highly insulating, in fact it may be 500 times the resistance of a salt marsh ground, such as is available near the seashore. In such a location most of the station power will be used in warming up the sand unless special precautions are observed. By grounding the antenna at many points the ground resistance is diminished according to the number used and thus the multiple-grounded antenna tends to nullify the bad effects of a poor location. Such an antenna consists, as shown in the diagram, Fig. 29, of a horizontal top similar to the top of "T" antenna, fed at one end by means of the alternator A connected in series with the tuning inductance L_1, and the vertical wire B_1C_1, and in addition, of a number of vertical wires attached to the horizontal top at suitable points and each separately connected to ground through a tuning inductance. The result of this is that each of the vertical wires acts as a vertical antenna, the whole arrangement constituting a number of vertical antennas connected in multiple, and hence radiating as if they were a single antenna. The advantage lies in the fact that, since each vertical wire is independently connected to ground, it follows that all the ground resistances are connected in multiple, and hence the total ground resistance is very much less than would be found to be the case with any other type of antenna of the same power capacity, thus giving a very high efficiency.[1]

[1] It must be pointed out here that the radiation resistance of each vertical wire of the multiple-tuned antenna cannot be calculated as though the wire stood alone, using e.g. Eq. (21), p. 912. The presence of the other vertical wires, also carrying current, will affect this radiation resistance, the amount of this effect depending upon the proximity of the various vertical wires, and upon the relative phases of their currents.

Of course it is hardly necessary to mention that the phases of the currents must be adjusted so that all the vertical wires will be radiating in phase with one another in order to obtain maximum radiation; the tuning coils L_1, L_2, . . . L_5 are used for the purpose of making this adjustment.

On the other hand, if the vertical wires be suitably spaced and if, in addition, the phases of their currents be suitably adjusted, it is possible by means of this type of antenna to obtain greater radiation in one direction than in another, thus producing directional transmission. Thus, in the case of the multiple-tuned antenna the intensity of the radiated field at any point is the resultant of the fields due to each of the vertical wires and, if suitably designed and adjusted, the resultant field in certain directions may be made a minimum and in others a maximum, thus producing a directional effect.[1]

An elementary analysis shows the normal operation of this antenna to be slightly directive, the maximum radiation taking place at right angles to the length of the antenna. If directive radiation is obtained by phase shifting in the different vertical wires, the radiation resistance of the antenna as a whole falls to a small fraction of its normal value.

"*Coil Antenna.*"—This has already been discussed on p. 886, where it was shown that such an aerial has a very decided directional effect, and that the intensity of the field in the plane of the coil, where it is a maximum, is a function of the distance between the two vertical sides of the coil and is greatest when this distance is equal to one-half a wave length. A comparison may here be made of the single vertical wire with uniform current throughout and of the coil antenna with uniform current throughout. Thus, from Eqs. (9) and (14) on pp. 883–885 for the effective values of the intensity of the magnetic field at any distance from antenna we have:

$$H = \frac{2\pi l I}{10\lambda d}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

for single wire

and

$$H = \frac{4\pi l I}{10\lambda d} \sin \frac{\pi s}{\lambda}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

for coil of one turn. Of course, if the coil aerial has more turns than one the intensity of the field is directly proportional to the number of turns, provided that the current is uniform throughout.

[1] See paper by E. F. W. Alexanderson, "Transatlantic Radio Communication," Proc. A.I.E.E., Oct., 1919. In reading this paper the student should bear in mind that the quantitative results predicted (magnitudes of currents, etc.) do not hold good for the transient state which, in an antenna of this kind, may be a large fraction of the duration of a "dot."

If N = number of turns

$$H=\frac{4\pi NlI}{10\lambda d}\sin\frac{\pi s}{\lambda}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14a)$$

for coil of N turns.

If I, l, λ, and d are the same for the two types of antennas we may obtain the ratio of the magnetic fields due to the coil and to the single wire by dividing (14a) by (9). Thus

Ratio of field of coil to that of single wire

$$=2N\sin\frac{\pi s}{\lambda}.$$

In order to make the two fields the same we must have

$$\sin\frac{\pi s}{\lambda}=\frac{1}{2N},$$

or

$$s=\frac{\lambda}{\pi}\sin^{-1}\frac{1}{2N}.$$

Below is given a table shown the value of s for different values of N.

TABLE I

Distance between sides of a coil aerial of the same height as a corresponding single vertical wire aerial necessary to make the fields from the two aerials alike.

N	s	N	s
1	0 17λ	5	0 032λ
2	0.08λ	10	0 016λ
3	0 053λ	100	0.0016λ

It is understood that the coil aerial field has, in the above discussion, been considered, which exists in the plane of the coil, i.e., the plane of maximum field intensity. The above table shows that for a single-turn coil aerial the width must be as large as 0.17λ in order for it to have an effect equivalent to that of a single wire of the same height. But with a larger N the width may be made much smaller, so that with 100 turns the width need only be a few meters, even with large wave lengths. However, with a large number of turns the question of the capacity between turns and the effective resistance of the coil plays an important part.

If the capacity from turn to turn is large (i.e., the turns close together) the current will not be uniform throughout, and, furthermore, the phase of the current at every point will be different, a condition which is not conducive to best results as regards radiation. Hence the turns should be separated by a considerable distance from one another. This may be stated by saying that the capacity of the coil itself should be such as to make the fundamental wave length of the coil no larger than about one-third of the wave length to be used. The effective resistance of the coil antenna is taken up on p. 910 of this chapter.

As examples of coil antennas which seem satisfactory for receiving purposes it may be noted that for a 600-meter wave a square coil, 120 cm. on a side, of 10 turns, spaced about 0.5 cm. from each other, requires a tuning condenser somewhat less than 0.001 μf.

By installing the coil (or "loop" as it is more frequently called) as indicated in Fig. 35, it may be used with the *D. P. D. T.* switch down, for general reception, the loop merely acting as a low antenna, tuning being accomplished by the variometer, *L*. When the desired signal is received the switch may be thrown upwards and the directive effect of the coil thus be obtained.

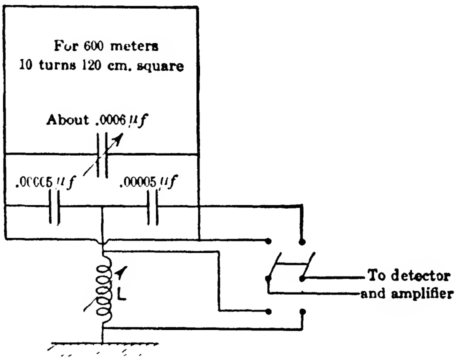


FIG. 35.—Use of a coil receiving antenna; by throwing the switch down the coil acts as a simple antenna, the coil *L* being used for tuning. When it is desired to get the directional effect of the coil the switch is thrown up.

For wave lengths from 10,000–20,000 meters a square coil, about 6 meters on a side with 50 turns spaced 4 cm. apart, is suitable.

Because of the comparatively low receptive power of loop antennas the receiver used must be the most sensitive available; the use of such a detector with a good amplifier is possible because of the comparatively low intensity of the " strays " picked up by a loop.

Aeroplane and Airship Antennas.—The aerial system of aircraft comes nearest to approximating the conditions represented by the simple aerial system of the Hertzian double (see Fig. 1), in so far as the counterpoise is not the ground, and furthermore the antenna and counterpoise are at considerable distance from the ground, so that the electromagnetic waves generated by such a radiating system travel outward in space without coming in contact with the ground except at considerable distance from the radiating system.

The various types of radiating systems used may be classified into two general headings:

(1) Those which may be used only when the ship is in flight.
(2) Those which may be used at any time whether the ship is in flight or not.

The first class includes by far the most effective type of aircraft aerial; in this case the aerial is a trailing wire dangling from the aircraft while the counterpoise consists of all the metal parts of the craft electrically connected together. The trailing wire is made up of a length of phosphor bronze or silver bronze wire ranging between 150 and 300 feet with a weight attached at its free end dangling from the aircraft somewhat as shown in Fig. 34. The transmitting or receiving apparatus is connected between the trailing antenna wire and the metal parts of the craft which, as already stated, form the counterpoise; this is schematically shown in Fig. 36; when the aircraft approaches ground the aerial wire is reeled in, the reeling-in apparatus being operated by hand or by a small electric motor.

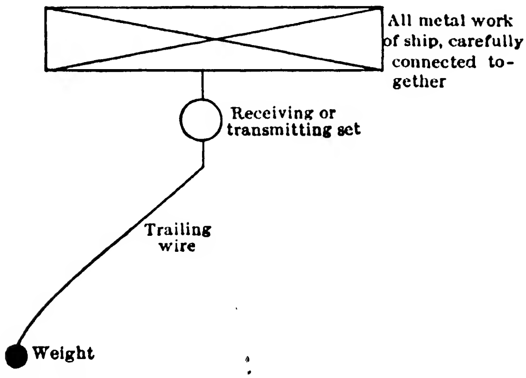


FIG. 36.—Arrangement of apparatus on aeroplane antennas.

Such an arrangement as the one above described has been used with success on practically all types of aircraft, including lighter-than-air ships. Its only disadvantage seems to lie in the fact that in the case of a forced landing, and, more especially, in the case of an aeroplane being compelled to dive or to "loop-the-loop" the presence of the trailing antenna wire might prove disastrous unless it were reeled in very quickly. Again, it may be easily understood that such an arrangement cannot be used unless the aeroplane is in flight.

The second class of aircraft aerials comprises various types which enable signals to be sent out or received even while the craft is on the ground. The following types have been used:

(*a*) Skid-fin aerials for aeroplanes.
(*b*) Coil aerials for aeroplanes.
(*c*) T-antenna for airships.

(*a*) The skid-fin antenna is nothing more than an inverted "L-antenna" the top of which is mounted a few feet above the uppermost plane and covers in length and width practically the entire wing, somewhat as shown by *ABCD* in Fig. 37, where the wire *DF* is the leading-in wire and connects directly to the transmitter or receiver; the counterpoise consists as usual of all the metal parts electrically connected together. Such an antenna has been extensively used by U. S. Navy aeroplanes. It must be understood that because neither the length of the leading-in wire nor that of the top wires can be made very large, as also because of the small separation between the antenna proper and the counterpoise the aerial is not a very good radiator, and, in general, aircraft carrying a skid-fin antenna also carry a trailing wire antenna. It may be said, in a general way, that the transmitting range of a skid-fin antenna is about one-half that of a dangling wire antenna for the same aircraft and transmitting apparatus.

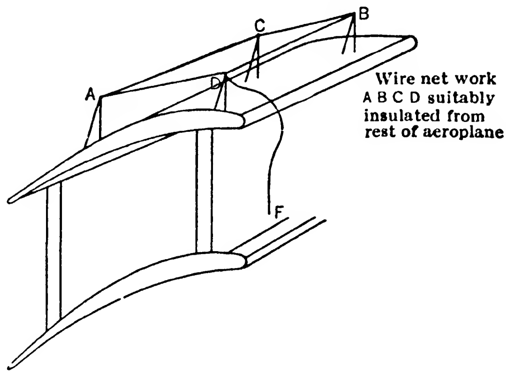


Fig. 37.—Aeroplane antenna of the skid-fin type.

When the metal work of a ship is used for counterpoise it must be all very carefully bonded together, otherwise sparks may occur, when transmitting, which are, of course, an unnecessary fire risk.

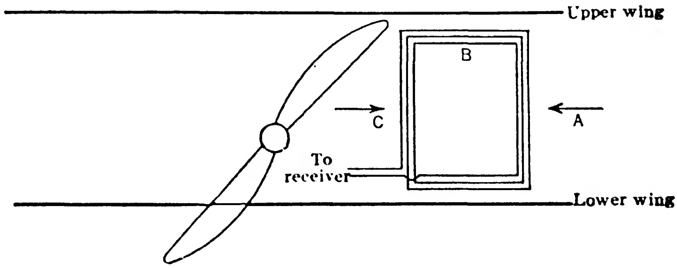


Fig. 38.—Coil-type antenna installed between the wings of an aeroplane: the coil sides are placed behind the struts between the wings.

In another type the antenna wires are strung along the back of the wing and from wing tips to tail.

(*b*) Coil aerials have been used more especially for receiving purposes, in view of their ability to detect the direction from which the waves may be coming. They are made up of several turns and of such dimensions as will enable them to fit in between the two wings of a biplane, somewhat as shown diagrammatically by B, Fig. 38. In this case no counterpoise is necessary. When the coil is used as a transmitter the greatest radiated field will be in the plane of the coil, similarly if the coil is used for receiving it will respond most vigorously to signals coming from the direction of A or C. In order to either send or receive in certain directions the coil may be rotated or else the aeroplane itself may be veered

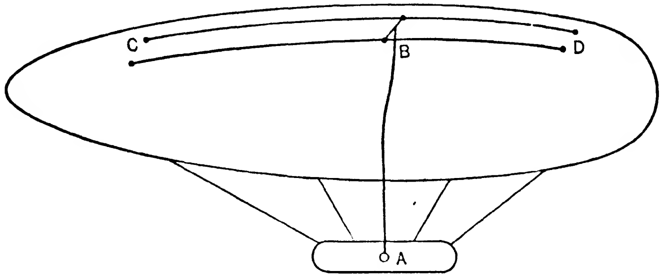


FIG. 39.—In a dirigible balloon a T-type antenna is used, the counterpoise consisting of all the metal work around the engines, etc.

around until the plane of the coil points in the desired direction. In order to avoid either one or the other of these operations another coil may be used with its plane at right angles to the first, in which case the operator need do no more than move small coils within his easy reach; this will be more fully explained later, in the section on direction finders, p. 944.

The range of transmission of coil antennas is small, but they are used for receiving from very great distances. Some aeroplanes carry a trailing wire for long-distance transmission while in flight, a skid-fin antenna while stationary, and a coil aerial for directional reception.

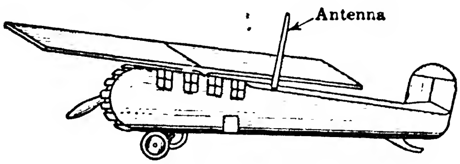


FIG. 40.—More recently aeroplanes use a short stream-lined rod, about 8 ft. high, as their antenna.

(c) The T-aerial for airships is schematically illustrated in Fig. 39, where AB is the leading-in wire and CD the top of the "T." The counterpoise consists of the metal parts of the suspended car, including engine, etc. Such an antenna has practically the same transmitting characteristics as a "T" antenna of the same dimensions used on the ground; and because the wire AB is quite long and the wires CD may be made very long as well, the range of the antenna is comparatively large. It need hardly be stated that the construction of such an aerial is such as to permit it to be used with equal effectiveness whether the airship is in flight or not, and is a great improvement over the trailing-wire antenna at first used on such ships. Care must, of course, be observed regarding the fire risk of the installation.

(*d*) More recently a short, rigid metal pole (stream lined) has been used on aeroplanes, especially when the radio receiver is to be used for direction finding, by means of radio beacons. Such an antenna is shown in Fig. 40.

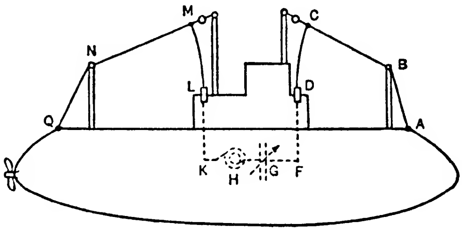


Fig. 41.—Arrangement of loop antenna in a submarine.

Still more recently it has been found that a low T antenna, of symmetrical design, is as satisfactory as the pole in so far as signal reception is concerned, and better mechanically. A symmetrical T antenna having a top wire extending 15 ft. either side of the fuselage, to the wing tips, and 12 inches above the wing surface, with a 12-inch vertical lead-in, gave slightly better reception than a 5-foot vertical pole.

A symmetrical T arranged above and along the fuselage, having the lead-in half way between front and back of the aeroplane, is just as good as the transverse one and offers less wind resistance. The radio receiver in this case, however, must be located midway along the fuselage. To keep these T antennas free from directional errors, the lead-in wire must be at the center, and vertical.[1]

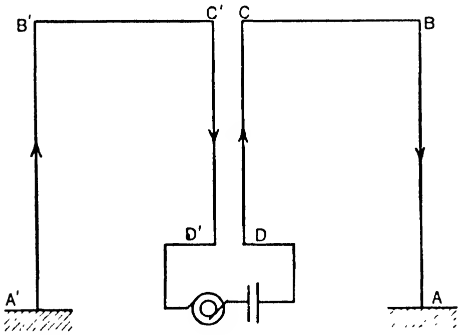


Fig. 42.—Electrical circuit of the installation of Fig. 41.

Underwater Antennas.—The problem of underwater antennas is especially important in connection with submarines. Up to a few years ago communication by radio with a submarine, while submerged, was considered very unsatisfactory, because use was being made of antennas similar to ground antennas such as the "T" type or inverted "L." T antennas, even if made of heavily insulated wire, are more or less likely to be short-circuited by the water (particularly salt water) more especially

[1] "Characteristics of Airplane Antennas for Radio Range Beacon Reception," Diamond and Davies, I.R.E., Feb., 1932, p. 346.

because, as will be more fully discussed on p. 931, the highest potential is, when transmitting with such antennas, present at the very end of the wires, where it is most difficult to guard against the short-circuiting effect of the water. Aside from these considerations which are not, however, so very serious when using the antenna for receiving purposes, the more serious handicap was the fact that such an antenna projects too far above the topmost part of a submarine, even above the periscope and made it necessary for the submarines to submerge more deeply than would otherwise have been the case or else to use a short, ineffective antenna.

Real progress was made in submarine radio transmission by the introduction of the loop antenna; in the application to submarine work the loop is made up somewhat as shown in Fig. 41. The wires $ABCD$ and $QNML$ are grounded at A and Q, and insulated from the boat everywhere else. Thus the loop may be diagrammatically represented as in Fig. 42, which should be compared with the diagram of the simple loop discussed on p. 884, and reproduced in Fig. 43 for the sake of convenience.

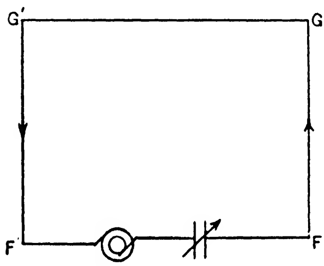


Fig. 43.—As wires CD and $C'D'$ of Fig. 42 radiate no appreciable power this arrangement is equivalent to the single-turn coil here shown; wire AB corresponding to FG and $A'B'$ to $F'G'$.

In the simple loop the wires FG and $F'G'$ radiate most effectively when the distance between them is one-half a wave length and the strongest field is produced by the loop in its own plane. Similarly in the case of the submarine loop the wires AB and $A'B'$ are the radiators while CD and $C'D'$ radiate very little energy since they are very close together and the fields created by them practically neutralize each other; of course the best distance between AB and $A'B'$ is one-half a wave length, and, again, as in the simple loop, the submarine loop will radiate best in its own plane.

Another arrangement used for submarines is a coil antenna consisting of a large number of turns and enclosed in water-tight container which is supported above the deck of the submarine. The dimensions of such a coil are necessarily small (perhaps one meter square), and its effectiveness as a transmitter is consequently low, but it has been used to receive from very great distances with considerable success.

A word should here be said regarding the transmission of electromagnetic waves in water. It has already been pointed out in Chapter II that electromagnetic waves may be transmitted through any medium to more or less extent depending more especially upon the electrical conductivity of the medium. An electromagnetic wave will, on striking a wall of ordinary conducting material, be partly reflected and partly absorbed in the production of currents in the material, so that practically no electromagnetic field would be found at even a small depth below the surface of the material. On the other hand, if the material is, as in the case of salt water, only a partial conductor, the electromagnetic waves are able to penetrate into it for considerable distance before the energy represented by them is completely absorbed by currents produced in the water. It is a well-known fact that the greater the frequency (the smaller the wave length) of a magnetic or electric field the smaller is the depth to which it will penetrate into a conducting or semi-conducting medium; therefore, in the case of electromagnetic waves in water, the extent to which they penetrate below the surface is very much dependent upon the wave length.

The equation for penetration of an electromagnetic wave into a conducting medium was given on p. 168. Although there given as the penetration of a *current* the same formula holds if written to express either electric or magnetic fields. Thus we may write

$$H_x = H_0 \epsilon^{-\left(\sqrt{\frac{2\pi\omega\mu}{\rho}}\right)x}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

in which H_0 = intensity of magnetic field, of the electromagnetic wave, just at the surface of the ocean;

H_x = intensity of magnetic field x centimeters below surface;

$\omega = 2\pi \times$ frequency;

μ = permeability of sea water = unity;

ρ = resistivity of sea water in abohms per cubic centimeter = approximately 10^{11}.

If we assume a signal detectable if H_x is only 1 per cent of H_0, then the depth at which the signal should be detectable is obtained from

$$\epsilon^{-2\pi\left(\sqrt{\frac{f}{\rho}}\right)x} = 0.01.$$

For a wave length of 10,000 meters the value of x calculated from this relation is about 1500 cm. or 15 meters.

As an example of the effect of wave length it has been stated that signals have been received by submarines with loop antennas with the top of the loop 16 feet below the surface of the water at a wave length of 6000 meters and 200 miles from the transmitting station, while for a wave length of 2500 meters and the same distance signals could only be heard with the top of the loop 8 feet below the surface of water.

If we assume that the loop was such that the "mean depth" was 5 feet lower than the top of the loop, so that in one case the effective depth was 21 feet in the first case and 13 feet in the other the experimental results agree very well with those predicted from Eq. (15). Thus we have

$$\sqrt{\frac{6000}{2500}}\times\frac{13}{21}=0.96.$$

Again, in case the submarine is transmitting while submerged, the transmitting range is very small because the electromagnetic waves are practically entirely absorbed in their passage through the water and issue therefrom with very feeble strength. Thus, it has been found that a submarine when submerged so that its loop antenna was only a few inches below the surface could only transmit to a distance of about 9 miles with a wave length of about 300 meters and an antenna current of 6 amperes while it could transmit 50 miles or more when on the surface. Probably better transmission through the water would be expected if the wave length were much larger (10,000 or more méters); but a large wave length implies an antenna of dimensions too large to be carried by a submarine.

It is to be remembered that the question of *reflection* at the surface of the water is to be considered when analyzing communication possibilities from a surface station to a submerged boat, or vice versa; this has not been attempted here. It may generally be stated that the present state of the art does not permit a submerged submarine to transmit to any greater distances than about 10 to 20 miles, while, on the other hand, enabling it to receive from almost any distance provided it is not too deeply submerged.

It has been found possible to send radio signals to trains when they were in long tunnels, a hundred feet underground.

Wave Antenna.—It has been found that for the reception of long-wave telegraph signals a long, low wire is more advantageous than a high one. The ratio of signal to disturbance is greater with the low one and this ratio is a true measure, in general, of the utility of an antenna.

For transoceanic signals an antenna several miles long is used, supported on insulators a few feet from the ground. It is laid in line between the receiving and transmitting stations and the receiving apparatus is placed between the antenna and ground at the end more distant from the transmitter. The other end is grounded through a resistance equal to the "surge resistance" of the line. It will be evident at once that these long horizontal antennas abstract energy from the radio wave only as a result of the inclined wave front. If the electric field component of the wave were perpendicular to the earth's surface (*i.e.*: vertical) the wave as it traveled along the horizontal antenna would induce no voltage in it. The theory of behavior of this antenna has been thoroughly analyzed by Beverage, Rice, and Kellogg.[1]

[1] See Journal A.I.E.E., March, 1923, and subsequent issues.

Antenna Design.—An antenna should be designed and built according to the purpose intended. In general the antenna of a broadcasting station for example, should not be in, or too close to, a large city but perhaps 10–25 miles distant. An antenna for transoceanic communication should generally be in a swampy location near the ocean. For short-wave transmission the ground conditions are of little consequence, as a counterpoise rather than a ground connection is generally used. The height is determined primarily from economic considerations, the cost increasing greatly above a certain height.[1]

Law of Radiation of Power from an Antenna.[2]—Upon consulting the literature there will be found many formulas which are supposed to give the power radiated from an antenna, in terms of the height, wave length, etc., but in general they do not agree, and it is difficult to appreciate the derivation of some of them. The simple derivation given below yields a result different from those given by accepted authorities, but it probably represents the true state of affairs as well as any of them.

Practically all analyses start from the theory of the Hertzian doublet, supposedly modifying it properly to make it apply to the grounded antenna. In some derivations the amplitude of the current in the antenna is supposed constant (i.e., the effective value of the current the same at the top of the antenna as at the grounded end), and in others the amplitude is supposed to vary in some prescribed manner. Some formulas use as the height of the antenna the actual height and others use a certain "effective height," measured to the "center of gravity" of the capacity of the antenna.

We shall consider the energy per cubic centimeter at a point P (Fig. 44), in the equatorial plane on the oscillator and distant from it several wave lengths, so far that the induction field is negligible. Our first assumption is that the effective value of the amplitude of the current in the vertical part of the antenna is at all points the same; this is nearly true for the ordinary antenna, in which the capacity of the vertical wire is small compared to the capacity of the network of wires generally used for the top of the antenna. This assumption will give us a radiation somewhat greater than the true value. The next assumption we make is that the actual height of the antenna, l, represents the distance between the positive and negative charges of the antenna, the flow of which causes the antenna current I. In case of a ship antenna the height l is from the water to

[1] For a general review of antenna design and construction the reader is referred to an article by Lindenblad and Brown, Proc. I.R.E. for June, 1926. For a general treatment of the electrical behavior of an antenna see article by Bennett in Jour. A.I.E.E., Nov. and Dec., 1920.

[2] For a thorough mathematical discussion, see Pierce, "Electric Oscillations and Electric Waves."

the top of the antenna. In case of a land antenna, with possibly a poor ground, it is likely that the average distance between the charges of the antenna is greater than the distance from the top of the antenna to the ground, so that it might seem that in this case we should take a distance

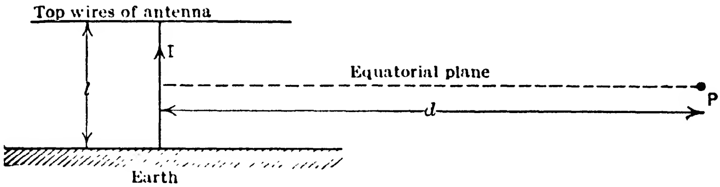


FIG. 44.—The energy radiated from an antenna is to be calculated from the law giving the strength of magnetic field at P, in terms of the antenna constants.

greater than the actual height, if the theory of the doublet is to be applicable.

This is indicated in Fig. 45; it may be, for such a ground condition, that the distance l' (average distance between charges) is considerably greater than l. We shall neglect this extra height $(l'-l)$, however, as it is not only indeterminable, but it contributes but little to the radiation reaching the distant point, P; the electromagnetic energy sent off from this subterranean part of the antenna could only reach P by traveling through the earth's crust, in which case the attenuation is so rapid that the amount of energy arriving at P by this path will be negligible compared to that reaching P from that part of the antenna specified by the height l.

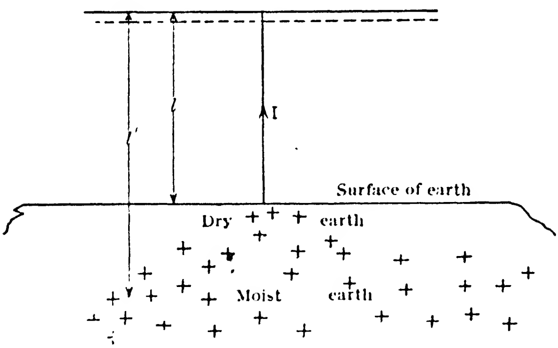


FIG. 45.—In an actual antenna there is undoubtedly a vertical motion of the charges in the earth under the antenna; this subterranean current will contribute practically no radiation at distance points because of absorption in the earth's surface.

We shall therefore assume that Eq. (9) p. 883 represents accurately the radiation field at point P, the symbols having the definite meaning given below.

$$H_m=\frac{2\pi l I_m}{10\lambda d},$$

in which H_m = maximum value of magnetic field at P, in gilberts per centimeter;

l = actual height of antenna, in centimeters, from ground to top, for flat-topped antenna;

I_m = maximum value of current (in amperes) in antenna, this value being assumed the same throughout the height of the antenna;

λ = wave length radiated, in centimeters;

d = distance from antenna to point P, in centimeters.

Now the energy per cubic centimeter at P, due to this magnetic field, is equal to $H_m^2/8\pi$, and as the electric field set up at P by this moving magnetic field must be of such magnitude that it represents the same energy per cubic centimeter as that possessed by the magnetic field, the total energy per cubic centimeter (maximum value) must be $H_m^2/4\pi$. As the electromagnetic wave travels past point P with the velocity of light, the electric and magnetic fields at this point both go through sinusoidal variations, so that the average value of the energy per cubic centimeter, in terms of the maximum value of magnetic intensity, must be equal to one-half of the maximum energy, or $H_m^2/8\pi$.

If we now consider the effective value of the magnetic field at the point P, we have (as $H_m^2 = 2H^2$, H being the effective value) the average energy of the radiation field at P equal to $H^2/4\pi$, the energy being in ergs per cubic centimeter.

This energy of radiation travels past point P with the velocity of light, V, so that the energy streaming past P per square centimeter (plane of the square centimeter being perpendicular to distance d) per second is equal to $H^2V/4\pi$. Using now Eq. (9) to express H and substituting I (effective current) for I_m, we get

Energy, in ergs, per square centimeter per second

$$=\left(\frac{2\pi lI}{10\lambda d}\right)^2\frac{V}{4\pi}=\frac{\pi I^2l^2V}{10^2\lambda^2d^2}. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (16)$$

In calculating the total radiation from the antenna we must assume some law of variation in the value of H, as the point P is moved over the surface of a sphere of radius, d. In the ideal case the distribution of H over the surface follows a cosine law as indicated in Fig. 46; it has a maximum value in the equatorial plane of the oscillator and zero directly above or below the antenna.[1] As the power per square centimeter varies

[1] This statement neglects the radiation from the horizontal currents in the upper wires of the antenna and in the earth. The amount of this radiation may be considerable and should be calculated in getting the total radiation from the antenna. As the problem lends itself at best to approximate treatment only, due to earth conditions,

with the second power of H, and as H has a sinusoidal variation with respect to θ, Fig. 46, it is evident that the average power per square centimeter over the whole imaginary sphere is $\frac{2}{3}$ times as great as at the equatorial plane, and this gives

$$\text{Average power per square centimeter} = \frac{2}{3}\left(\frac{\pi I^2 l^2 V}{10^2 \lambda^2 d^2}\right)$$

Now as the current in the upper part of the antenna must be smaller than that at the base where the current is measured the radiation must be less than that calculated on the assumption of uniform current distribution. In the ordinary flat-top antenna it is quite reasonable to put $\frac{1}{2}$ in place of $\frac{2}{3}$ in the foregoing formula, to take care of this non-uniform current distribution, and then, as the total area of the sphere is $4\pi d^2$, we have, for the total radiation from the oscillator

Total radiation, in ergs per second

$$= \left(\frac{\pi I^2 l^2 V}{10^2 \lambda^2 d^2}\right) 2\pi d^2$$

$$= \frac{2\pi^2 I^2 l^2 V}{10^2 \lambda^2}.$$

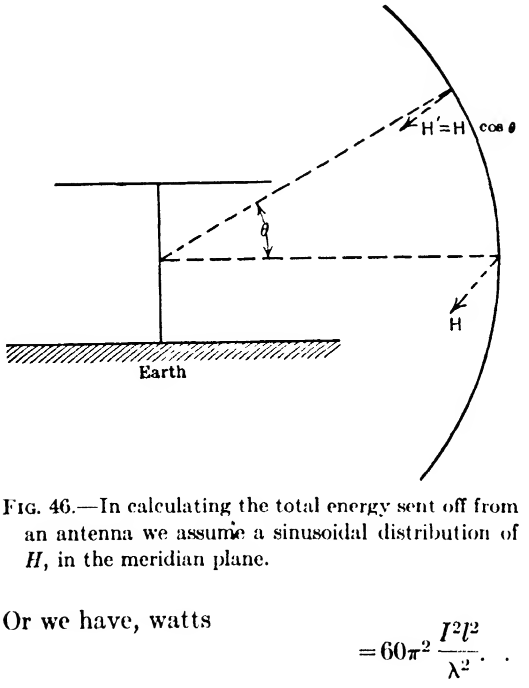


FIG. 46.—In calculating the total energy sent off from an antenna we assume a sinusoidal distribution of H, in the meridian plane.

Or we have, watts

$$= 60\pi^2 \frac{I^2 l^2}{\lambda^2}. \qquad (17)$$

In this formula I is measured in amperes (effective) and l and λ are measured in any convenient unit, providing it is the same for both.

It will be noticed that in this derivation the treatment does not agree with that ordinarily given [1] in that the radiation is considered as occurring over a *whole sphere* instead of only a hemisphere. It will be appreciated that this way of looking at the question is correct if any analogous problem

etc., it is not thought worth while to include the calculation of this up-and-down radiation.

[1] See Berg, "Electrical Engineering," advanced course, p. 292.

in radiation is considered. Thus imagine an upright incandescent filament sending out light as indicated in Fig. 47. The filament is supposed to have its lower end resting on a surface which absorbs part of the incident light and reflects the rest.

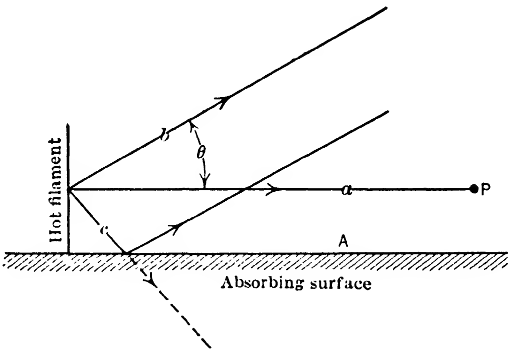


FIG. 47.—The radiation of light from an incandescent filament standing in a partially reflecting surface is exactly analogous to the radiation of radio waves from an antenna.

Let us suppose that, by use of accepted formulas, we have obtained the intensity of illumination at P, due to light traveling from the filament directly to P. (This excludes light arriving at P after being reflected from surface A.) Suppose further that we know the law for the distribution of radiation, with respect to the angle, θ, this law representing the distribution in a homogeneous medium, i.e., exclusive of any such reflecting surface as we have at A. From this law we can obtain the average lumens per square centimeter which would exist over the surface of a sphere through P if the reflecting surface A were not present. To get the total radiation it is evident that we must multiply this average illumination by the whole surface of the supposed sphere if we are to get the total radiation from the filament. To be sure, the lower half of the sphere (below the surface A) actually gets inappreciable illumination, due to reflection at the surface and to absorption in the material below A, but this fact in no way alters the radiation from the filament, it merely redistributes the lumens *after they have left the filament,* and increases to some extent the illumination in the upper hemisphere. The surface of the earth acts in the same way towards the radio waves as does the surface A to the light rays striking upon it.

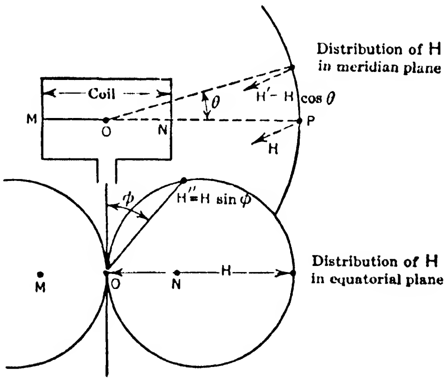


FIG. 48.—In calculating the power radiated from a coil we assume a sinusoidal distribution of H in both equatorial and meridian planes.

Radiation from a Coil.—In the case of a coil the formula for radiation may be at once obtained by using the proper value for H in the previous deduction for the ordinary antenna. We suppose a coil of one turn the length of whose vertical sides is l, and the width between these sides is s; the value of H in the equatorial plane is

$$\left(\frac{4\pi lI}{10\lambda d}\right)\sin\frac{\pi s}{\lambda}.$$

If the coil has N turns of course this value of H must be multiplied by N.

The formulation of the total radiation for the coil requires the knowledge of the distribution in the meridian plane as well as the equatorial plane. Assuming both these distributions sinusoidal,[1] as indicated in Fig. 48, we find the average value of H^2 and thus we get the total radiation from the coil

$$\text{Watts} = 120\pi^2\frac{I^2l^2}{\lambda^2}\sin^2\frac{\pi s}{\lambda}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

In case the coil is so narrow that $\sin \pi s/\lambda = \pi s/\lambda$ we have

$$\text{Watts} = 120\pi^4\frac{I^2l^2s^2}{\lambda^4}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

and if the coil is square so that $s = l$ we have

$$\text{Watts} = 120\pi^4 I^2\left(\frac{l}{\lambda}\right)^4. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

It was mentioned when calculating the radiation from an ordinary antenna that the horizontal parts of the antenna give off considerable radiation, which was neglected in getting the total radiation. It must be noticed that in the case of the coil antenna this omission causes a very large error, because by its very form, the coil radiates as much from its horizontal sides as it does from its vertical sides, if the coil is a square. In case the coil is not square its radiation due to the horizontal sides may be obtained at once by interchanging the symbols s and l in Eq. (14). Taking this extra radiation into account it would seem that the total power radiated from a square coil is twice the value given by Eq. (20).

All of the foregoing formulas for radiation have been obtained on the assumption that the current was uniform in amplitude throughout the

[1] As noted before, the treatment of radiation given here is elementary and approximate only; the student is referred to Chapter IX of Pierce's "Electric Oscillations and Electric Waves" for a full treatment of the subject.

length of the radiating portion of the antenna. It is evident that when such is not the case (as, for example, a straight vertical grounded wire) the average value of the current must be approximated and this value used in the proper formula. Thus, for the single wire just referred to, if considerable loading is used, the average current is one-half the value of current at the ground end of the antenna and the radiated power would be one-quarter of the value given by Eq. (17). If, on the other hand, the wire was oscillating at its fundamental ($l=\lambda/4$) the average current would be $2/\pi$ of the current at the base and the power would be $(2/\pi)^2$ or 41 per cent of the value given by Eq. (17).

Both Eqs. (17) and (18) show that the power radiated by either a coil or a simple antenna is a direct function of the square of the height and the square of the current, and an inverse function of the square of the wave length.

We will illustrate the influence of the wave length upon the power radiated by means of an example. Assume a simple antenna for which

$$l=10{,}000 \text{ cms.}=100 \text{ meters}$$

$$I=20 \text{ amperes}$$

then if

$$\lambda=1000 \text{ meters } (f=300{,}000 \text{ cycles per second})$$

$$\text{Power}=60\pi^2\times\frac{100^2\times 20^2}{1000^2}\cong 2400 \text{ watts,}$$

while, if

$$\lambda=100{,}000 \text{ meters } (f=3000 \text{ cycles per second})$$

$$\text{Power}=60\pi^2\times\frac{102^2\times 20^2}{10^{10}}\cong\frac{24}{10^2}=0.24 \text{ watt.}$$

Thus it may be seen that it is impossible to radiate power to any great extent at low frequencies and it must also be remembered that this hypothetical case of 20 amperes supplied to an antenna at 3000 cycles is impossible of realization.

Variation in Induction Field and Radiation Field.—In the case of a coil antenna the distribution of radiation field is entirely different from that of the induction field so it might seem possible to get measurements to separate them. Ramsey [1] carried out tests of this kind, and in Fig. 49 are shown three of his diagrams. In diagram *A* are plotted in polar diagram the induction field, and in dashed line is shown the locus of his experimental points, for a distance from the coil center of 1/20 wave

[1] I.R.E., Aug., 1928, p. 1118.

length (which was 16 meters). At a distance $\lambda/2\pi$, the radiation field and induction field are equal, in the plane of the coil, as shown in diagram *B*. In diagram *C* are shown the relations of the two fields at a distance of $\lambda/2$ meters. The three diagrams are not to the same scale; this radiation field of course actually falls off uniformly with increasing distance, and the induction field at a much faster rate; the diagrams are correct only for *relative* magnitudes of the two fields compared one with the other.

Radiation Resistance.—Radiation resistance is a fictitious resistance the value of which is such as will absorb the same power as is radiated for the same current as flows in the antenna.

From the definition the radiation resistance may be found by divid-

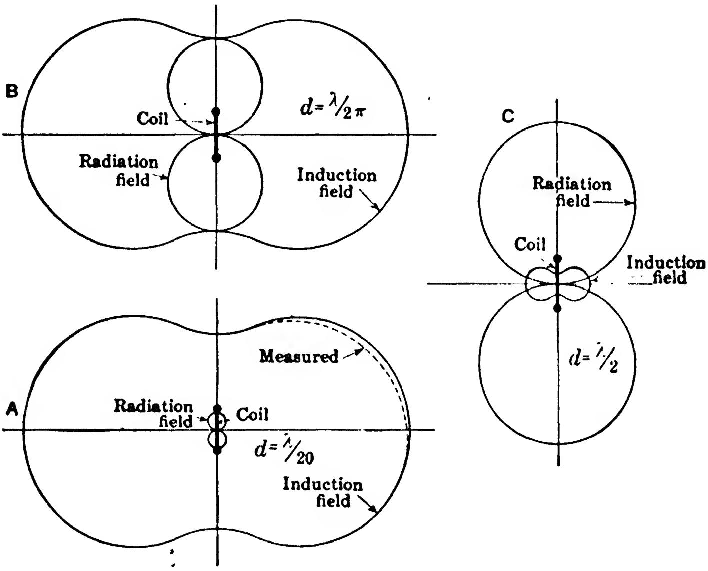


FIG. 49.—Experimental determination of the induction field and radiation field around a coil antenna, for three different distances from the coil.

ing the power radiated by the square of the antenna current. Thus, from Eqs. (17) and (19) we find

Radiation Resistance for simple antenna

$$=60\pi^2 \frac{l^2}{\lambda^2}. \qquad (21)$$

Radiation Resistance for single turn coil having a width equal to s, small compared to one-half a wave length

$$=120\pi^4 \frac{l^2 s^2}{\lambda^4}. \qquad (22)$$

The radiation resistance is used as a measure of the ability of an antenna to radiate power. An antenna with a high radiation resistance is a good radiator, and vice versa.

As previously pointed out the values of resistance obtained from Eqs. (17) and (19) may be far from correct for an actual antenna. A single vertical wire (no top wires) will have a resistance only 41 per cent of the value given by Eq. (17) when oscillating at its natural period and if much loading is used, so that the amplitude of current decreases uniformly from base to top of antenna the radiation resistance will be but 25 per cent of the value calculated from Eq. (17).

In the case of the coil antenna, radiating up and down, as well as horizontally, the radiation resistance is probably much greater than the value given by Eq. (18); for a square coil perhaps twice as much.

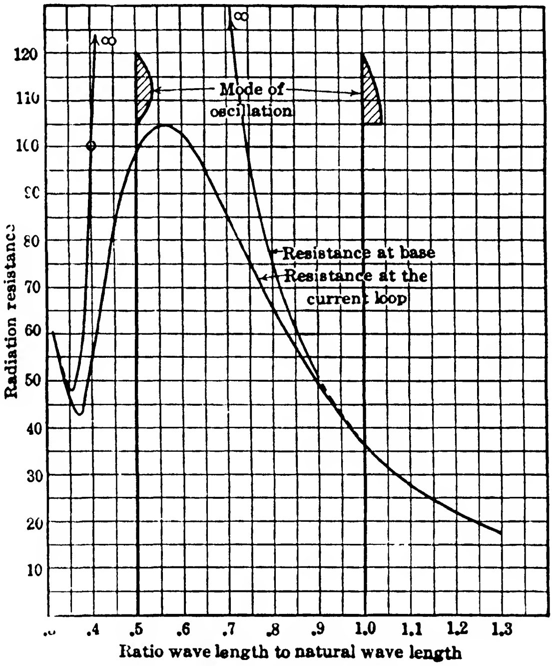


Fig. 50.—The calculated radiation resistance of a simple vertical antenna operated at wave lengths shorter than its natural wave length (antenna length one-quarter wave length). Curve marked "Resistance at base" calculated on assumption that ammeter is in the ground end of the antenna; the other assumes the ammeter is at the point of maximum current.

Antennas Operated Below Their Natural Wave Length.—With the increasing utility of the very short waves it has become customary to operate antennas at less than their natural wave length; this requires the insertion of a series condenser.

The resistance of such an antenna will rapidly rise as the wave length is decreased below the natural wave length due to two effects. The radiated power actually increases with the increase in frequency and the current at the ground connection decreases. The loop of current (maximum value) does not occur at the ground connection as it does on the ordinary antenna, but occurs part way up the antenna.

Ballantine[1] has calculated the radiation resistance for an antenna used in this fashion and gets results as shown in Fig. 50. The natural wave length is four times the height of the vertical wire and from Ballantine's results it is seen that this radiation resistance for this condition is 36.5 ohms. This result is obtained after very arduous work calculating the summation of complicated series, etc. The very simply derived result of Eq. (21) gives a value of 37 ohms.

In Fig. 50 one of the curves gives the resistance in terms of the ammeter reading at the current loop while the other gives the resistance in terms of the ammeter reading at the ground connection.

Current in Receiving Antenna.—It is important to be able to calculate the current in the receiving antenna, because the value of this current determines whether or not it is possible to hear the signals which cause such a current to flow in the receiving antenna. It must be here stated that were it not for the interference of the so-called "strays" (see p. 388) it would be possible, due to the extreme refinement and sensitiveness of modern receiving apparatus, to hear signals, no matter how small the currents in the receiving antenna. In view of the "strays," however, which also produce currents in the receiving antenna, the signal currents must be larger than would otherwise be necessary, so that the "strays" may interfere with the signals as little as possible; since the "strays" currents have considerable magnitude it follows that attention must be paid to making the signal currents large.[2] Hence the importance of knowing the factors affecting the signal current in the receiving antenna. We will determine this for a simple antenna and for a coil antenna.

Received Current in Simple Antenna.[3]—Consider the antenna represented by Fig. 51 in the path of electromagnetic waves moving, as shown, in a direction perpendicular to the antenna lead-in wire $A-B$. The electric field will act in a direction parallel to AB, hence there will exist a difference of potential across AB which will be equal to its length multiplied by the intensity of the electric field. Thus, if:

$\mathcal{E}$ = effective value of intensity of electric field at AB, in volts per cm.;
l_r = height of receiving antenna, in centimeters;

[1] "On the radiation resistance of a simple vertical antenna at wave lengths below the fundamental," by Stuart Ballantine, Proc. I.R.E., Dec., 1924.

[2] For example in England during the winter time the average atmospheric disturbance amounts to from one to ten microvolts per meter, when the receiving apparatus is tuned for 50 kc.

[3] In an article by Bennett, in the Journal of the A.I.E.E., for Nov. and Dec., 1920, various properties of antennas are analyzed and exact expressions for them derived. Among other things, he shows that an antenna having negligible resistance (other than radiation), the amount of power which can be abstracted by a receiving antenna is equal to about 6 per cent of the amount flowing through an area (parallel to the wave front) equal to $(\lambda)^2$ square meters, λ being in meters.

I_r = effective value of current in receiving antenna, in amperes;
R = effective resistance of the antenna, this of course depending among other things upon the coupling and adjustments in the closed receiving circuit, type of detector used, etc.

Then, since the receiving circuit is always adjusted to resonate to the frequency of the incoming waves, it follows that the reactance will be zero and current in this circuit will be given by:[1]

$$I_r = \frac{\mathcal{E} l_r}{R}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

It now becomes necessary to substitute for $\mathcal{E}$ its value in terms of the transmitting antenna constants.

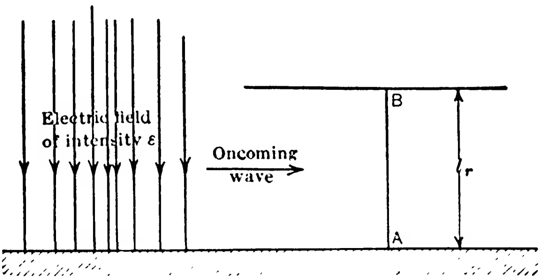


FIG. 51.—Wave with electric gradient, $\mathcal{E}$, approaching a receiving antenna.

From Eqs. (10) and (15) we have:

$$\mathcal{E} = \frac{60\pi l I}{\lambda d},$$

volts per centimeter for a simple antenna;

$$\mathcal{E} = \frac{120\pi N l I}{\lambda d} \sin \frac{\pi s}{\lambda},$$

for a coil of N turns and width s.

Substituting in Eq. (23), we have:

$$I_r = \frac{60\pi l l_r I}{\lambda d R} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

for a simple transmitting antenna;

$$I_r = \frac{120\pi N l l_r I}{\lambda d R} \sin \frac{\pi s}{\lambda}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

[1] These solutions hold only for the steady state; they are not good until the transient condition is past.

for a coil transmitter of N turns and a width s with the receiving antenna in the plane of the coil.

Of course these formulas are only approximate; the effect of the curvature of the earth's surface, effect of sky waves, etc., are left out of the question, and they may have a very important bearing on the amount of received current, especially when the distance between transmitter and receiver is large.

Received Current in a Coil Antenna.—Assume the single-turn coil of Fig. 52 placed in the path of incoming electromagnetic waves, the wave front and plane of the coil being perpendicular to each other and the electric field of the wave being parallel to conductors AB and $A'B'$. Then,

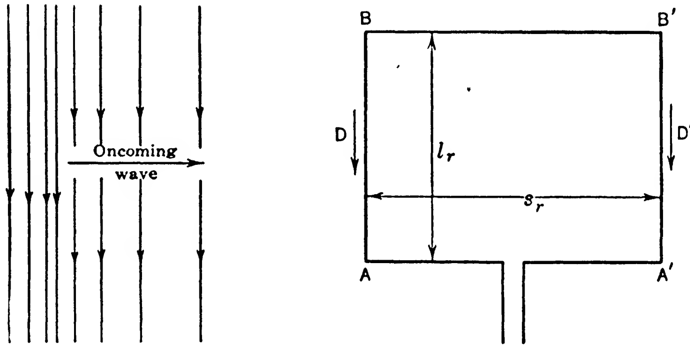


FIG. 52.—Wave approaching a coil antenna.

if D and D' represent the assumed positive direction of the potential difference established across AB and $A'B'$ and, if:

$\mathcal{E}$ = effective value of intensity of electric field at AB, in volts per centimeter;

$\mathcal{E}_1$ = effective value of intensity of electric field at $A'B'$, in volts per centimeter;

s_r = width of coil in centimeters;

we have:

$l_r\mathcal{E}$ = effective value of potential difference across AB;

$l_r\mathcal{E}_1$ = effective value of potential difference across $A'B'$;

$\dfrac{2\pi s_r}{\lambda}$ = phase difference between $l_r\mathcal{E}$ and $l_r\mathcal{E}_1$ in radians.

The total electromotive force within the entire coil is equal to the vector difference of $l_r\mathcal{E}$ and $l_r\mathcal{E}_1$. Since AB and $A'B'$ are at practically the same distance from the radiating antenna, the magnitudes of $l_r\mathcal{E}$ and $l_r\mathcal{E}_1$ will be practically the same, but their phases will differ. Hence, total electromotive force within the coil is obtained by taking the vector difference of the e.m.f.s in the two sides of the coil, as shown in Fig. 53.

We then have

$$2l_r\mathcal{E} \sin \frac{\pi s_r}{\lambda} = \text{electromotive force in coil}$$

and effective value of current on the assumption of a tuned receiving circuit is given by:

$$I_r = \frac{2l_r\mathcal{E}}{R} \sin \frac{\pi s_r}{\lambda}, \qquad (26)$$

for a single-turn coil.

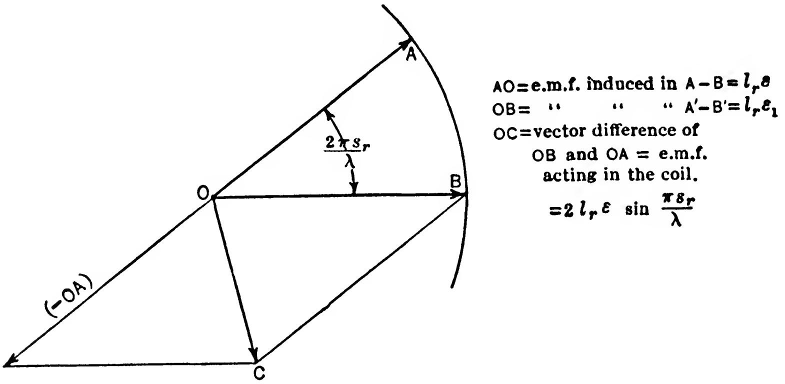


FIG. 53.—The effective induced e.m.f. of a receiving coil antenna is obtained by subtracting vectorially the e.m.f.s. induced in the two sides.

In case the receiving coil has N_r turns Eq. (26) becomes

$$I_r = \frac{2l_r\mathcal{E}N_r}{R} \sin \frac{\pi s_r}{\lambda}, \qquad (27)$$

for a coil of N_r turns.

And substituting for $\mathcal{E}$ the expressions obtained from Eqs. (10) and (15) we finally have the following:

$$I_r = \frac{120\pi N_r l l_r I}{\lambda d R} \sin \frac{\pi s_r}{\lambda}, \qquad (28)$$

for current received in a coil antenna from a simple transmitting antenna lying in the plane of the receiving coil; and

$$I_r = \frac{240\pi N N_r l l_r I}{\lambda d R} \sin \frac{\pi s}{\lambda} \cdot \sin \frac{\pi s_r}{\lambda}, \qquad (29)$$

for a transmitter coil of width s and placed with its plane in that of the receiving coil.

For the sake of convenience all receiving formulas are collected below; it is assumed that: the vertical wires of the receiving antenna are parallel to the electric field of the oncoming wave, that the transmitting antenna current is undamped and of uniform amplitude throughout, that there is no energy absorption by the medium, that the receiving circuits are tuned to the transmitting frequency, that the curvature of the earth has no effect, that the sky wave is of negligible importance, and that the planes of the coils (either transmitters or receivers) are directed towards the other antenna or coil.

Antenna to antenna

$$I_r = \frac{188 l l_r I}{\lambda d R}. \qquad (30)$$

Coil to antenna

$$I_r = \frac{376 N l l_r I}{\lambda d R} \sin \frac{\pi s}{\lambda}. \qquad (31)$$

Antenna to coil

$$I_r = \frac{376 N_r l l_r I}{\lambda d R} \sin \frac{\pi s_r}{\lambda}. \qquad (32)$$

Coil to coil

$$I_r = \frac{752 N N_r l l_r I}{\lambda d R} \sin \frac{\pi s_r}{\lambda} \sin \frac{\pi s}{\lambda}. \qquad (33)$$

In case the angle $\pi s/\lambda$ is sufficiently small that the angle may be substituted for its sine these formulas become somewhat simpler in form. We have

Coil to antenna or antenna to coil[1]

$$I_r = \frac{1180 N l l_r s I}{\lambda^2 d R}. \qquad (34)$$

Coil to coil

$$I_r = \frac{7450 N N_r l l_r s s_r I}{\lambda^3 d R}. \qquad (35)$$

In every one of the preceding formulas it is to be noted that the received current is:

> A direct function of receiving and transmitting antenna heights, and the transmitting antenna current.

[1] It is to be pointed out that whereas the constants in these formulas are given to the third significant figure, the actual received current may differ from the predicted value by a factor of three or more; refraction and reflection of the waves play an important role in transmission.

An inverse function of the wave length, to the first, second, or third power, the distance and the resistance of the receiving antenna circuit.

It will be seen that for a fixed frequency the received current is directly proportional to the height of the transmitting antenna and to its current. This has led to the suggestion that the unit for defining the ability of an antenna to radiate be taken as the *meter-ampere*. It makes no difference how the product of meter-amperes is obtained, by a large current with low antenna or vice versa.

Returning to the matter of the effect of " strays " it is apparent that if the received signal current is made large by suitably arranging the receiving antenna constants, then the " strays " current will at the same time be made large, and thus reception may be poor. On the other hand, if the receiving antenna constants are poor and the transmitting antenna constants very good, then the received signal current will be large while the " strays " current will be small, with consequent improvement in reception. This explains the modern tendency towards radiating systems of very large dimensions and receiving systems of small dimensions.

In using a coil antenna for reception of signals, a regenerative, or " feed-back " connection of some sort should be used, to reduce the resistance of the coil as much as possible. Such a scheme involves a connection similar to the connection of the closed circuit of Fig. 182, p. 635, where it must be supposed that the coil L_2 of the diagram represents the coil antenna used for receiving.

The reception Eqs. (30) to (35) should be modified by multiplying by suitable factors when the transmitting antenna current is damped, when there is absorption of energy by the medium and when the plane of the coil is not parallel to the direction in which the waves are being propagated. The factors are as follows:

When the transmitting antenna current is damped [1]

$$\text{Factor is } \sqrt{\frac{\delta}{\delta+\delta'}}.$$

When there is absorption [2]

$$\text{Factor is } \epsilon^{-0.000047\frac{d}{\sqrt{\lambda}}}$$

When the plane of coil is not parallel to the direction in which the waves are being propagated

$$\text{Factor is } \cos\alpha,$$

[1] See Dellinger's paper on "Radio Transmission," Proc. A.I.E.E., Oct., 1919.

[2] See Scientific Paper No. 226 of the Bureau of Standards. This absorption coefficient holds good only over the ocean, in daylight. Over land, and over either land or ocean at night time, the transmission is too erratic to make a formula worth while.

where δ = decrement of receiving antenna circuit;
δ' = decrement of transmitting current;
d = distance between station in meters;
λ = wave length in meters;
α = angle made by plane of coil with the direction of propagation of the waves;
ϵ = base of natural logarithms.

Comparative Merits of Different Types of Antennas.—At the transmitting station it will probably always be necessary to use a high antenna, directive or not as desired, but for receiving a signal it is scarcely ever advantageous to use the same high antenna as used for transmitting.

The readability of a signal depends not upon the actual strength of the signal, but upon the ratio of signal strength to that of the disturbing noises also present. If static interference comes from all directions the ratio of signal to static may evidently be increased by using a directional receiving antenna; also, as it seems probable that most of the energy of strays may be considered to exist in the form of highly damped, long-wave signals, the best antenna will be one that absorbs but little energy from waves greater than that for which it is tuned. A coil antenna satisfies both of these requirements better than the ordinary high antenna, it being directional and having induced in it a voltage inversely proportional to the wave length, for an oncoming wave of fixed value of electric field, $\mathcal{E}$; the induced voltage in the ordinary antenna under the same conditions is independent of wave length. (The statement regarding the coil antenna presupposes a coil width s, small compared to the wave length, practically always the case.)

It is, of course, true that the intensity of signal received by the coil antenna will be only a small fraction of what it would be with the other antenna but *the static interference will be even a smaller fraction.* Hence by a good amplifier the signal may be brought up to readable intensity and (if the amplifier increases static and signal equally) the amplified weak signal from the coil will be more easily read than an equally loud, unamplified, signal from the high antenna.

The validity of the above argument depends to some extent upon the actual ratio of radiation resistances of the two antennas; if, e.g., the coil has an induced signal current only 0.0001 as much as that of the high antenna, then it will be necessary to use an amplification of 10,000 times (in volts) to make the coil signal as loud as that from the other antenna. In the present state of the art it is likely that such an amplifier would generate in itself sufficient noises (due to microphonic resistances, "dirt" on hot filament, thermal voltage, etc., see Chapter X) to make the signal unreadable.

Counterpoises.—It has already been stated that an antenna, other than a loop or coil antenna, must necessarily consist of a so-called aerial, which radiates, and a counterpoise, which may or may not radiate. In the simple Hertzian double the counterpoise radiates, and this is also true to a certain extent of the counterpoise used in aircraft, made up, as it is, of the metal parts of the craft. In most cases, however, the counterpoise is the ground itself.

Sometimes when the ground is dry and therefore a poor conductor the counterpoise consists of a network of wires laid on the ground (in some cases insulated from it) directly underneath the top of the antenna proper. In every case it must be understood that the purpose of the counterpoise is to enable charges of electricity to be transferred to and fro between itself and the aerial with as little loss (due to heat development) as possible, and for this reason it must have low resistance and it must also have sufficient capacity. The metallic surface of the counterpoise should be at least equal to that of the antenna and is in most cases much larger.

A counterpoise having a small surface has the same effect as a small capacity connected in series with the aerial; that is, it makes the antenna capacity very small. In order to make such a low capacity aerial resonate at desirable frequencies it will probably be necessary to use a large loading inductance, which is generally accompanied by a large resistance; hence the power lost and the decrement of such an aerial are very large. As a matter of fact, it is generally attempted to make the counterpoise of as large a surface and as low a resistance as possible. As a rule, when the ground underneath the antenna is a good conductor (wet, soft earth) the ground itself is used as a counterpoise and connection is made with it by means of copper plates or network of wires sunk into the ground at various places within the area underneath the antenna. These buried conductors should be put deep enough so that the earth around them is permanently moist.

It may sometimes be necessary to insert radio-frequency choke coils in series with various connecting wires of the buried network of wires to make the current divide properly between various parts of the counterpoise. Strange as it may seem, adding these coils to the feed wires of the counterpoise actually decreases its resistance.

Antenna Resistance.—An antenna or coil transmitter absorbs power when supplied with high-frequency currents by an alternator or some other generator of such currents. Some of this power is, as has already been pointed out, radiated in the form of an electromagnetic field, and represents useful power, while the rest is consumed in various ways and represents a complete loss, in so far as it contributes nothing towards radiation.

In dealing with the power absorbed by a circuit such power is, for the sake of simplicity, looked upon as if it were expended in a resistance of such a value as would consume the actual power expended in the circuit for the same current as flows therein. This fictitious resistance is known as "effective resistance."[1]

Since the total power expended in an antenna is partly radiated and partly "lost" due to various causes we may divide the "effective resistance" of an antenna in two parts, i.e.:

(*a*) Radiation resistance.
(*b*) Loss resistance.

The radiation resistance has already been defined on p. 912, and the expressions therefor have been derived for a simple antenna and for a coil transmitter; these expressions are, for convenience, rewritten below:[2]

$$R=60\pi^2\frac{l^2}{\lambda^2}, \qquad (21)$$

for simple antenna

and

$$R=120\pi^4\frac{l^2s^2}{\lambda^4}, \qquad (22)$$

for single-turn coil having a width s small compared to λ.

The loss resistance is due to a number of losses which are enumerated and discussed below:

(1) Loss in poor dielectrics in the neighborhood of the aerial.
(2) Loss in the resistance of the aerial.
(3) Loss in the resistance of the counterpoise, generally the ground.
(4) Loss due to eddy currents in neighboring conductors.
(5) Loss due to leakage over insulators, etc.
(6) Loss due to corona.

(1) The loss in poor dielectrics is due to the hysteresis[3] phenomenon taking place in all dielectrics, and most especially in poor dielectrics such as moist wood, concrete, masonry, trees, etc., which may happen to be in the vicinity of the aerial and hence acted upon by the electrostatic field about the aerial. This loss resistance is analogous to that due to magnetic hysteresis in iron and is an inverse function of the frequency or a direct function of the wave length as discussed on p. 255. The effective resis-

[1] See pp. 163 et seq.

[2] Dellinger gives for the resistance of a vertical antenna $R=(39.7\ l/\lambda)^2$ and for a coil antenna (having $l=s$) $R=(13.3\ l/\lambda)^4N^2$.

[3] See pp. 252 et seq.

tance due to dielectric loss must, therefore, increase as the wave length increases or as the frequency diminishes. This loss is one of the most important [1] taking place in a radiating system and should be reduced to a minimum by keeping the field of the antenna free from unnecessary obstructions wherein a dielectric loss is likely to take place. As the highest electric gradient occurs near the end of an ⅂ or T antenna, especial care must be taken to keep poor dielectrics away from this part of the antenna.

In the case of a ship's antenna much loss may occur in the "lead-in" insulator where it enters the radio room; in case the radio room is wood, no metal (except the wire itself) should be used in this insulator. When the radio room is metal or where the "lead-in" wire has to go through metallic bulkheads, a considerable unavoidable power loss occurs in the insulator.

(2) The loss in the ohmic resistance of the aerial wire should be kept low by making the wire of large cross-section and good conducting material. The large useful cross-section may be obtained by using a large number of very fine wires which are insulated from one another in order to prevent the skin effect increasing the resistance.[2] The material is generally some bronze (phosphor or silicon bronze), since this combines fair conductivity with great tensile strength; mechanical considerations generally determine the kind of cable to use, in that many times a seven-strand cable is used which has practically as much skin effect as solid wire.

(3) The loss in the resistance of the counterpoise necessarily occurs because there are currents flowing therein which must produce a Joulean loss of power as they encounter a resistance. A counterpoise should be made of the smallest possible resistance. Where the ground is the counterpoise it is important that connection be made thereto by means of a large number of copper plates buried all around the antenna in soft, moist soil. It was previously pointed out how the multiple-tuned antenna may be employed in order to diminish as much as possible the ground resistance.

(4) Loss due to eddy currents in neighboring conductors may be diminished by eliminating as much as possible all metal masses from the neighborhood of the antenna. Of course this is quite impossible in so far as metallic masts are generally used to support the antenna, and, besides, if these masts were replaced by wooden or concrete masts the latter might suffer considerable dielectric loss.

Since eddy currents and the loss due thereto increases with an increase of the frequency (for a fixed value of current) it follows that the effective resistance representing this loss increases with the frequency or decreases with an increase of the wave length.

[1] See Bureau of Standards Scientific paper No. 269, by J. M. Miller.

[2] See pp. 174 et seq.

(5) The loss due to the leakage currents flowing between the aerial and the counterpoise should be kept down by using suitable insulators between the antenna wires and the supports and also between the lead-in wire and any walls through which it passes so that the resistance of the leakage paths may be made as high as possible. The resistance of the leakage paths is, of course, very much diminished in wet weather and, especially, where sprays from a rough sea reach the aerial. It has already been pointed out that in the case of submarines the ordinary antenna is very inefficient, except on a smooth sea, because of the salt-water sprays producing large leakage currents to ground and thus absorbing the largest part of the energy given to the antenna.

Since the loss due to leakage is a direct function of the $(\text{voltage})^2$ and the voltage is inversely proportional to the frequency (for a given current) it follows that the effective resistance corresponding to leakage loss varies inversely as the square of the frequency and directly as the square of the wave length.

(6) The loss due to corona takes place at high voltages and is due to the partial ionization of the air about the antenna wires, which causes the air to become a partial conductor and carry a current. At night the corona effect is visible through the glow which accompanies it. The corona does not begin to take place except at a certain definite voltage, which, however, varies with the shape and size of the conductors; this critical voltage is smallest where the conductors are small and at points and corners. Once the critical voltage has passed, a large amount of energy loss may take place due to corona. As a matter of fact this phenomenon is to a certain extent a limitation upon the amount of power which may be radiated by an antenna in so far as, for an antenna of certain dimensions, the greater the power given thereto the greater must be the voltage and hence the greater the corona loss; thus, for a certain antenna there is a limit to the power input, beyond which it is inadvisable to go because a large amount of power is wasted due to corona loss, and little is gained as far as power radiated is concerned.

This limit is reached when the voltage at the ends of the antenna is in the neighborhood of 150,000 volts.[1] This is one reason why the use of very large radiating systems for large stations is imperative in order that the large capacity resulting therefrom may keep the voltage below the limit of corona loss even for large amounts of power input. The effective resistance representing this loss is for a fixed current an inverse function of the frequency and a direct function of the wave length, for voltages above the critical value.

[1] As mentioned before this limit depends upon how well the antenna conductors are kept free from sharp points and edges.

From the above we have, then, that for a certain antenna and for a fixed current therein:

Radiation resistance is an inverse function of $(\lambda)^2$;

Resistance corresponding to (2), (3) and (4) (eddy currents and skin effect) decreases as λ increases;

Resistance corresponding to (1), (5) and (6) (dielectric loss, leakage, corona), increase with increase in λ.

The above relations are roughly indicated in Fig. 54, where the various components of the antenna resistance have been plotted, together with curves showing the total loss resistance and the total antenna resistance.

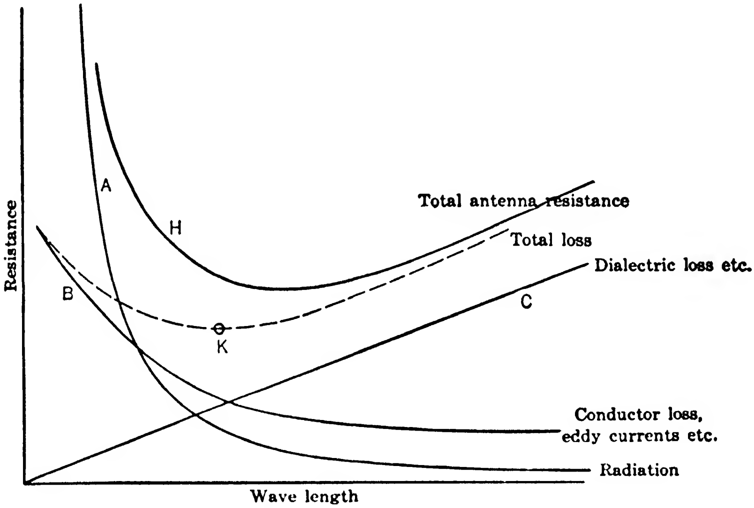


FIG. 54.—Various components of antenna resistance, showing approximately how they vary with wave length.

From the component curves B and C, we have obtained the total loss resistance curve, by adding the ordinates of these two curves and finally, the total antenna resistance, curve H, by adding the ordinates of the curves A, B, and C. The important point brought out by the curves is that, because some of the loss resistance components are a direct function, and others an inverse function, of the wave length, it follows that the total loss resistance has a minimum value, as represented by the point K on the " total loss " curve. It would seem, then, as if from the point of view of the losses the best wave length at which to use an antenna should be that corresponding to point K, and in practice this is approximately the most efficient wave length at which to operate an antenna.

The curves given above are purely of a theoretical nature, because the components of the antenna resistance cannot be satisfactorily measured by the methods at present available. However, total resistance curves of actual antennas, which are easily obtained, are all found to have the shape of curve H, and the point of minimum resistance is always found to be at wave lengths considerably greater than the fundamental wave length, perhaps twice as great.

Some typical resistance curves of actual antennas are given herewith; Fig. 55 shows the resistance for a ship's antenna for which the minimum resistance takes place at a wave length of 3.5 times the fundamental.

Fig. 56 shows the resistance of another ship antenna, plotted against frequency instead of wave length. The irregular character of this resistance variation indicates losses in the ship's rigging, masts, etc. In this same diagram is shown the change in apparent capacity of the antenna as the frequency increased; this is due to the change in voltage distribution along the antenna, an effect analyzed later in this chapter.

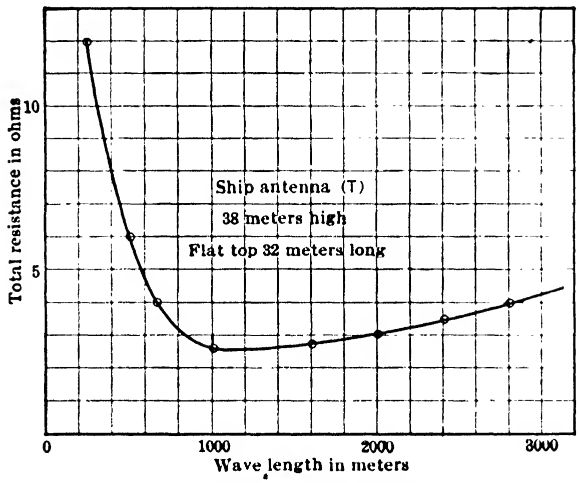


Fig. 55.—Total resistance curve for a ship's antenna.

The ordinary land station antenna resistance resembles these curves in form, but generally the resistance increases with the longer wave lengths more rapidly than does that given in Figs. 55 and 56. Fig. 57 gives the total resistance for an aeroplane antenna of the trailing-wire type and one of the skid-fin type; both of these antennas have about the same fundamental wave length. It is to be noted that the trailing-wire antenna curve shows large resistance to the left of the minimum value, while the other curve shows large resistance to the right of the minimum value. This is accounted for as follows: a trailing-wire antenna is a much better radiator than a skid-fin antenna, hence the radiation resistance should be larger in the former and therefore that part of the curve to the left of the minimum, which is very much affected by the radiation resistance, should have the larger ordinates in the trailing antenna than in the skid-fin antenna curve. On the other hand, since the skid-fin antenna is very close to the aeroplane structure, the dielectric and leakage-loss resistance should be very much greater than in the trailing-wire antenna, and hence the ordinates of the resistance curve to the right of the minimum value should be much larger. The minimum resistance for the skid-fin antenna is seen to be less than for the other in view of the shorter length of wire used and hence less ohmic resistance.[1]

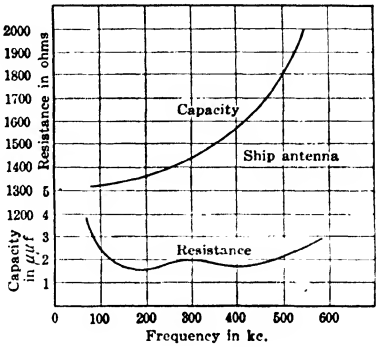


FIG. 56.—Variation in effective capacity, and peculiar resistance variation, of a ship's antenna.

In Figs. 58 and 59 are shown the resistances of trailing-wire antennas of different lengths, suspended from aeroplanes and dirigibles. The 250-foot wire hung from the plane gave consistent results (as did the others also) showing a resistance increasing rather rapidly as the frequency increased past 1000 kc.; this is to be expected, as the quarter wave length oscillation of such an antenna occurs at about this frequency, giving a high radiation resistance. In fact, the rapid rise in resistance of all these antennas is due to radiation, indicating efficient emitters. In the case of the 250-foot wire from the dirigible, however, the resistance increased to excessive values at 600 kc., much less than its natural frequency. This is due to resonance effects in some other conductor (quite possibly the dirigible itself) which is electrically coupled to the trailing wire.

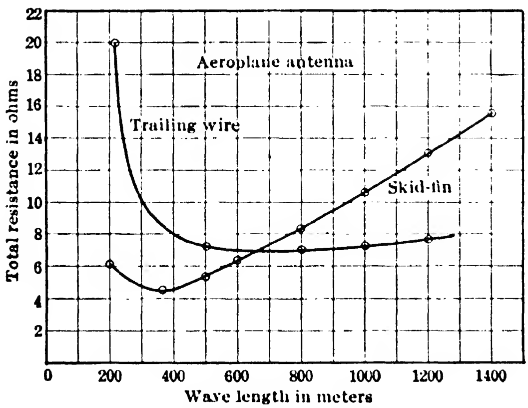


FIG. 57.—Resistance curves for two types of aeroplane antenna.

The very large land stations have a minimum antenna resistance between 1 and 2 ohms; the minimum resistance for the antenna of a 5-kw. set is generally between 5 and 10 ohms. Portable field antennas sometimes have a resistance as high as 50 ohms.

[1] For a number of curves of aircraft antenna resistance see Johnson's paper in I.R.E., Vol. 8, Nos. 1 and 2.
Also see Scientific paper No. 341 of Bureau of Standards, by J. M. Cork.

What has been said of the resistance of an ordinary antenna applies to a coil radiator as well, except, of course, that the components of the total resistance are related to one another in a somewhat different way;

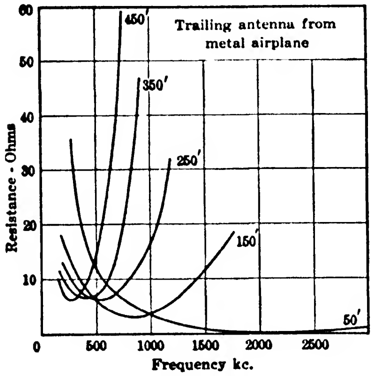


FIG. 58.—Total resistance curves of wires trailing from a metal airplane; radiation from the 50-foot wire was extremely small.

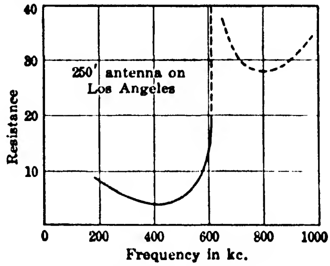


FIG. 59.—A 250-foot wire trailing from a metal frame dirigible showed excessive resistance at a certain frequency; this was probably due to resonance in some conducting part of the dirigible, being excited by coupling of some sort (electric or magnetic) with the trailing wire.

and this is true, to a certain extent, of any one type of antenna relative to any other. The most important thing about a coil radiator is that its counterpoise or ground resistance is practically eliminated, and hence a much less total resistance is obtained. Therefore, a certain voltage will, when impressed upon a coil radiator, produce a much larger current than in a simple antenna having the same radiation resistance as the coil; hence it is possible to radiate larger amounts of power by means of the coil than one might at first think; for ordinary-sized coils, however, the frequency must be very high if appreciable power is to be radiated.

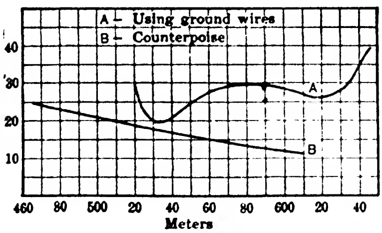


FIG. 60.—This pair of curves shows the diminution in resistance possible, when a counterpoise is substituted for earth connection; it seems that the radiation resistance, as well as the loss resistance, is cut down by the change.

It will be recalled, however, that the submarine antenna (p. 901) is nothing but a single-turn coil. Furthermore, the radio beacons used in airway navigation are also nothing but one-turn coils, though several hundred feet long and of considerable height.

Reduction of Antenna Resistance by Counterpoise.—If the ground below an antenna is a poor conductor there will be much loss there, giving a high antenna resistance. In this case the use of a network of wires (suspended perhaps 10 feet above the ground on insulators mounted on short poles) in place of the ground connection will appreciably reduce the antenna resistance. Such an effect is shown in Fig. 60;[1] this decrease in resistance is accompanied by some decrease in radiation, because the effective radiating height has been reduced. However, in this special case most of the decrease in resistance was due directly to decrease in ground losses.

Natural Wave Length of Antenna.—Consider an antenna in its simplest form, i.e., a long vertical wire connected to the alternator as shown in Fig. 61. The antenna wire has:

(1) Distributed inductance.
(2) Distributed capacity.
(3) Distributed resistance.

(1) The distributed inductance is due to the ability of every part of the antenna to develop magnetic lines of force. Assuming the absence of magnetic material near the antenna, its inductance per unit length should be practically uniform throughout its height.

(2) The distributed capacity consists of the capacity between the wire and the counterpoise, or earth, and is, in general, different for different parts of the antenna.

(3) The distributed resistance of the antenna is due to radiation and all the losses taking place. This total resistance per unit length of wire may be considered to be about uniform throughout the entire antenna height.

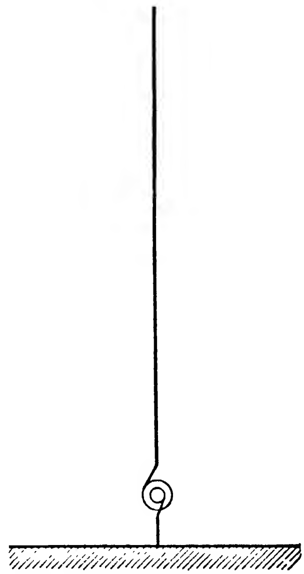


Fig. 61.—In an actual antenna the capacity, inductance, and resistance, are distributed and must be so considered when accurate equations for current and voltage are desired.

The antenna as a whole has a certain value of effective inductance, effective capacity and effective resistance, all of which, when defined in terms of current at the base of the antenna, change with the frequency of the currents flowing through the wire. It has already been shown how the effective resistance changes with the frequency or wave length. The effective inductance and capacity change with the frequency because, as will be presently demonstrated, the distribution of the voltage and current over the antenna changes with the wave length or frequency. Thus, if the antenna inductance, say, is measured for a certain value of λ and some effective value of current, I_o, at the alternator, and if, then, the

[1] From an Article by O'Neill, I.R.E., July, 1928, p. 872.

wave length is changed while the effective value of the current I_o is maintained the same, the total magnetic flux emanating from the antenna will be different on account of the different distribution of current, and the inductance will necessarily be different. If, for instance, the effective value of the current over the antenna were as in *a*, Fig. 62 (an impossible condition) every part of the antenna would be nearly as effective in producing magnetic flux, while if the current distribution were as in *b*, with the same current, I_o, at the alternator, the parts of the antenna farthest removed from the alternator would not be very effective in producing magnetic flux, thus making the inductance smaller as compared with *a*.

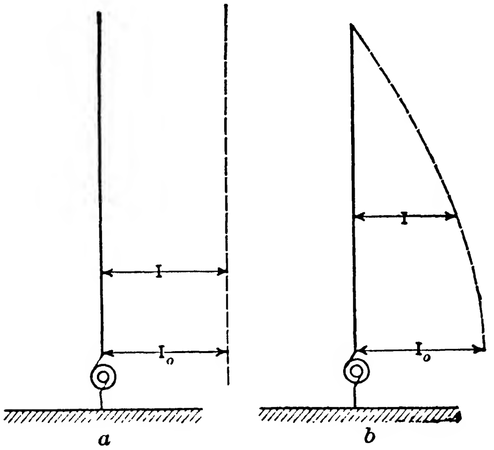


Fig. 62.—If the current at the base of an antenna is held constant while frequency is changed the distribution of current along the antenna will change; this will change the amount of magnetic energy associated with the antenna and hence will change its effective self-induction.

A similar thing happens in the case of the capacity, because the voltage distribution along the antenna varies with the wave length, and hence the ability of the different parts of the antenna to produce electrostatic lines of force varies with the wave length.

Effect of Ground on Natural Wave Length of an Antenna.—An antenna of given height above dry ground will have actually a greater electrical height, and have a larger natural wave length, than one on wet ground, because the ground of the aerial must be the semi-conducting layer of earth, as illustrated in Fig. 45, p. 906. An interesting illustration of this question is shown in Fig. 63; a wire of given length, trailed from

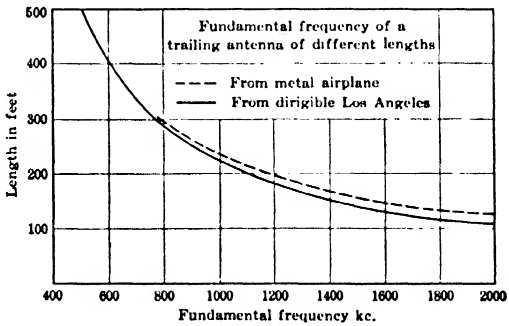


Fig. 63.—A trailing wire, of given length, has a lower natural frequency when suspended from a dirigible than when suspended from an airplane; this is due to the larger counterpoise (or earth) supplied by the extensive metal framework of the dirigible.

an aeroplane, will have a shorter wave length than when trailed from a dirigible, because the " earth " is so much larger in the latter case. In the first, the metal work of the aeroplane is the " earth," and in the other, the metal work of the dirigible. The natural frequency is nearly 10 per cent higher in the case of the aeroplane. The values given in this curve check pretty well with the idea that the natural wave length is four times the antenna length; thus with an antenna 400 feet long we calculate the natural frequency as 615 kc., which is what the curves show.

Distribution of Current and Voltage along the Antenna Wire.—It has already been pointed out that the current and voltage cannot be the same throughout the antenna wire of Fig. 61. The current in such a wire exists because electrons are being made to flow alternately *into* and *out* of a capacity; at the very end of the wire past which there is no capacity the current must be zero and will grow in value for points farther away from the end. The flow of this rapidly alternating capacity current (leading current) through the inductance of the antenna wire produces an increasing voltage as we proceed towards the end of the wire, a phenomenon which is well known to the electrical engineer in the case of long-distance transmission lines. In order more fully to understand the distribution of current and voltage we are giving below the expression for the current and voltage in a simplified antenna or, more definitely, an antenna having *uniformly distributed inductance and capacity* and no resistance whatever. Thus, let AB, Fig. 64, represent such an antenna.

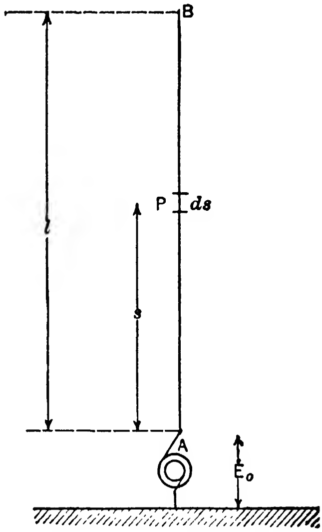


FIG. 64.—By considering the leakage, capacitance, inductance, and resistance of a small element ds, the equations for current and voltage may be obtained. Even if the resistance and leakage are neglected fairly accurate expressions will be obtained.

Let $\dot{E}$ = e.m.f. vector at any point P, the effective value of the e.m.f. being E;

$\dot{I}$ = current vector at any point P, the effective value of the current being I;

s = distance from A to P, in centimeters;

l = length of antenna in centimeters;

L_1 = inductance of antenna in henries per centimeter;

C_1 = capacity of antenna in farads per centimeter;
ω = angular velocity of alternator e.m.f. in radians per second;
x = inductive reactance in ohms per centimeter $= \omega L_1$;
b = capacity susceptance in ohms per centimeter $= \omega C_1$;
$\dot{E}_o$ = e.m.f. vector at alternator end of antenna;
$\dot{I}_o$ = current vector at alternator end of antenna.

By suitable mathematical analysis[1] it may be shown that, if the antenna resistance is neglected,

$$\dot{I}_o = j\frac{b}{\sqrt{bx}}\dot{E}_o \tan\sqrt{bx}\,l \quad . \quad . \quad . \quad . \quad . \quad . \quad (36)$$

$$\dot{E} = \frac{\dot{E}_o}{\cos\sqrt{bx}\,l}\cos\sqrt{bx}(l-s). \quad . \quad . \quad . \quad . \quad (37)$$

$$\dot{I} = -\frac{\dot{I}_o}{\sin\sqrt{bx}\,l}\sin\sqrt{bx}(l-s). \quad . \quad . \quad . \quad (38)$$

If we let d = distance from the upper end of the antenna in centimeters, then

$$d = l - s, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (39)$$

call

$$a = \sqrt{bx}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (40)$$

Substitute (39) and (40) in Eqs. (36), (37), (38) and we have the simple equations:

$$\dot{I}_o = j\frac{b}{a}\dot{E}_o \tan al. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (41)$$

$$\dot{E} = \frac{\dot{E}_o}{\cos al}\cos ad. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (42)$$

$$\dot{I} = \frac{-\dot{I}_o}{\sin al}\sin ad. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (43)$$

From Eqs. (42) and (43) we note that both $\dot{E}$ and $\dot{I}$ are trigonometric functions of the distance from the upper end of the antenna d. $\dot{E}$ varies with cos ad and $\dot{I}$ varies with $-\sin ad$, therefore $\dot{E}$ and $\dot{I}$ must differ by 90° in "space phase." In Fig. 65 the abscissas of curves $\dot{E}$ and $\dot{I}$ represent the values of the e.m.f. vector and current vector, respectively, at

[1] See John M. Miller, "Electrical Oscillations in Antennas and Inductance Coils," Proc. I.R.E., Vol. 7, No. 3.

various points along the antenna, obtained by the application of Eqs. (42) and (43). At the end of the antenna the current vector is zero while the voltage vector is a maximum. At the alternator end, $\dot{E}$ and $\dot{I}$ may have any value, depending upon the value of $\sqrt{bx}$ and the height of antenna. The example represented by the curves is not one which is generally attempted in practice, since the antenna would, in this case, produce a peculiar distribution of radiation, being a maximum in some zones, and zero in others, due to the fact that the current and voltage are positive over certain portions of the antenna and negative over others; hence in some zones the effect of certain parts of the antenna would be partly or fully neutralized by other parts. The curves, however, show a more or less extreme possibility. The more usual case is that represented by Fig. 66; this case will be discussed more fully a little later.

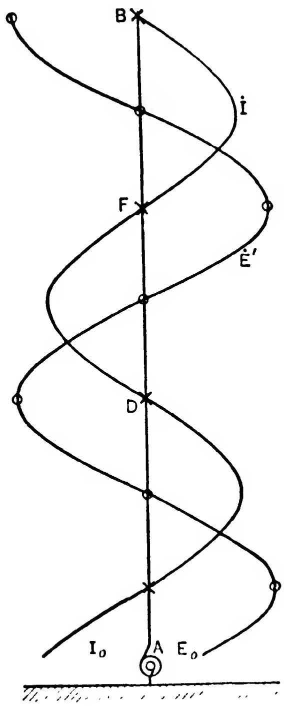


FIG. 65.—A possible form of excitation of an antenna, at a frequency much higher than its natural frequency.

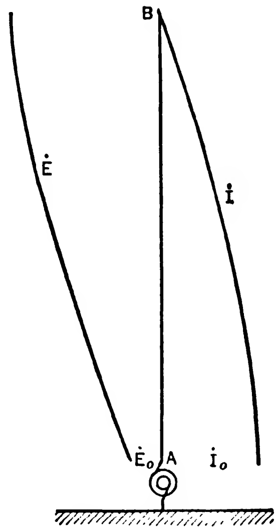


FIG. 66.—The ordinary form of voltage and current distribution on an unloaded antenna, excited at its natural wave length.

Again, we note from Eqs. (42) and (43) that, since $\dot{E}$ and $\dot{I}$ are trigonometric functions of ad, this quantity must represent a *space rate of change of angle*, as distinguished from ω, which represents a *time rate of change of angle.* Now, looking at the curves of Fig. 65, which represent nothing but so-called *stationary waves* of e.m.f. and current, the distance between such points as B and D must be the length of the stationary wave over the antenna, and this distance must be such as to make

$$a\lambda_1 = 2\pi,$$

where

λ_1 = wave length of stationary waves in centimeters,

or

$$\lambda_1 = \frac{2\pi}{a}. \qquad (44)$$

Since

$$a = \sqrt{bx} \text{ and } b = \omega C_1,\ x = \omega L_1,$$

$$a = \omega\sqrt{L_1C_1}. \qquad (45)$$

If

f = frequency of alternator in cycles per second,

$f = \frac{\omega}{2\pi}$ and substituting in (45)

$$a = 2\pi f\sqrt{L_1C_1}. \qquad (46)$$

The quantity $1/\sqrt{L_1C_1}$ is shown in electrical engineering texts to be very nearly equal to the velocity of light or to the velocity of propagation of electromagnetic waves emanating from an antenna through the air.

If

V = velocity of propagation of electromagnetic waves in centimeters per second,

$$\frac{1}{\sqrt{L_1C_1}} = V,$$

and substituting in (46)

$$a = \frac{2\pi f}{V},$$

and finally substituting this expression for a in Eq. (44) we have:

$$\lambda_1 = \frac{V}{f}. \qquad (47)$$

Thus, the length of the antenna stationary waves, λ_1, is equal to the velocity of propagation of the waves divided by the frequency; but this quotient represents the length of the electromagnetic waves, therefore, the *wave length of the stationary antenna waves is equal to the wave length of the electromagnetic waves in free space.*

Now, going back to Eq. (41) on p. 932, we may solve for the value of $\dot{E}_o/\dot{I}_o$, thus:

$$\frac{\dot{E}_o}{\dot{I}_o}=-j\frac{a}{b}\cot al. \qquad (48)$$

Since $\dot{E}_o$ and $\dot{I}_o$ are the e.m.f. and current at the alternator, their ratio must be the effective impedance of the antenna at the point where the alternator is connected. In our case the expression for this impedance is always imaginary, and therefore, represents the value of the reactance. This result was to be expected, since the resistance has been omitted in our simplified discussion.

Let X_o = antenna reactance at the alternator in ohms.
Then

$$X_o=-\frac{a}{b}\cot al. \qquad (49)$$

Substituting for

$$a=\frac{2\pi f}{V}=\omega\sqrt{L_1C_1}, \quad b=2\pi fC_1=\omega C_1,$$

we have

$$X_o=-\sqrt{\frac{L_1}{C_1}}\cot 2\pi\frac{f}{V}l, \qquad (50)$$

or

$$X_o=-\sqrt{\frac{L_1}{C_1}}\cot\frac{2\pi l}{\lambda}, \qquad (51)$$

where λ = wave length of electromagnetic waves, in centimeters.

The value of this reactance will apparently vary, for a fixed λ, as we vary the antenna height l. When l is such as to make $\cot 2\pi l/\lambda=0$, then the reactance is zero and the antenna will resonate to the alternator frequency. This will happen when

$$\frac{2\pi l}{\lambda}=\frac{\pi}{2} \text{ or } \frac{3\pi}{2} \text{ or } (2n+1)\frac{\pi}{2}, \qquad (52)$$

or when

$$l=\frac{\lambda}{4} \text{ or } l=\frac{3}{4}\lambda \text{ or } l=\left(\frac{2n+1}{4}\right)\lambda.$$

And since λ is also equal to the wave length of the stationary antenna waves, it follows that the antenna will resonate to the frequency of the alternator when the antenna height is such that there will result a distribution of e.m.f. and current vectors which wlll produce either $\frac{1}{4}$ or $\frac{3}{4}$ of $\frac{5}{4}$, etc., of a stationary wave.

If the expression for X_0 as given by Eq. (50) be plotted against values of alternator frequency,[1] everything else remaining the same, we would have the curve shown in Fig. 67. At the points 1, 3, 5 the antenna reactance is zero, and the frequencies at 1, 3, 5, etc., are in ratios 1 : 3 : 5 : 7, etc. Hence the antenna can be made to resonate at the frequency f_1 and at frequencies three times, five times, seven times, etc., f_1. On the other hand, a little to either side of the points 2, 4, 6, etc., the reactance is given by the formula as infinite and directly at the points 2, 4, 6, etc., the resistance of the antenna, as measured at the base, is found by the formula to be infinite so that practically no current can be caused to pass into the antenna at the frequencies f_2, f_4, f_6, etc., which are two times, four times,

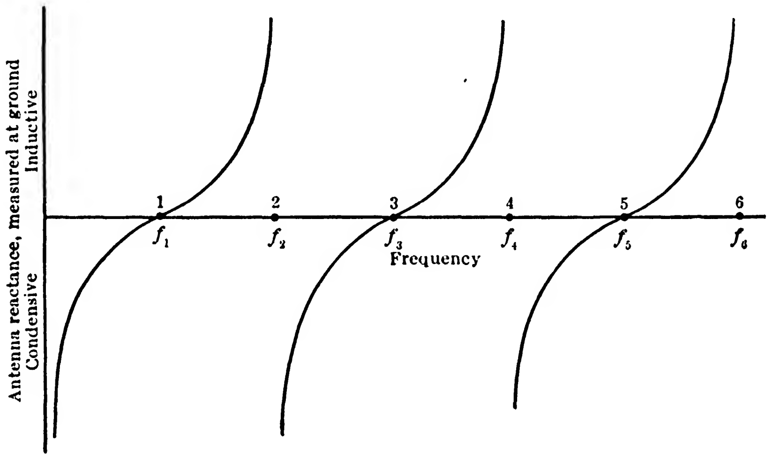


FIG. 67.—As the frequency impressed on an antenna is varied the reactance (as measured at the base) goes through the changes indicated here; in case an antenna with appreciable resistance had been considered the reactance changes from its high positive value to high negative value by going through zero values at 2, 4 and 6.

six times, etc., the first resonating frequency f_1. It is not out of place to point out here that the first resonating frequency f_1 is such as to produce one-quarter of a stationary wave over the antenna, as may be easily seen from the preceding discussion. This frequency and the wave length corresponding to it are known as the *fundamental or natural frequency* and *wave length of antenna,* respectively.

An antenna if excited by means of a spark gap will naturally have currents produced in it of the frequency corresponding to zero reactance, and therefore of the fundamental frequency and wave length; it is possible by putting proper discontinuities in the antenna, to cause this frequency to be three times and even five times the fundamental. In general, it

[1] For experimental curves showing this effect see Fig. 166, p. 160.

may be said that, whenever the simple antenna oscillates freely, no matter how excited, it does so at the natural or fundamental wave length.

The curves of Fig. 67 also show that the reactance of the antenna may be negative (condensive) or positive (inductive), depending entirely upon the frequency at which it is used.

In the above discussion we have assumed an antenna consisting of a vertical wire and having distributed inductance and capacity, and no resistance. The presence of the resistance makes the results regarding frequencies only slightly different, and so does the fact that the capacity is not quite uniformly distributed. The effect of properly introducing a resistance term into the development of Eqs. 50–51, etc., is to yield finite values of reactance for all values of frequency, a result we know is required by the physical facts.

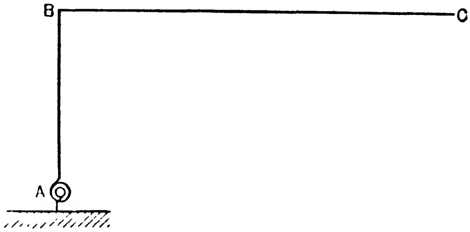


FIG. 68.—In the case of an inverted L antenna the natural wave length is slightly more than four times the extreme length A–B–C.

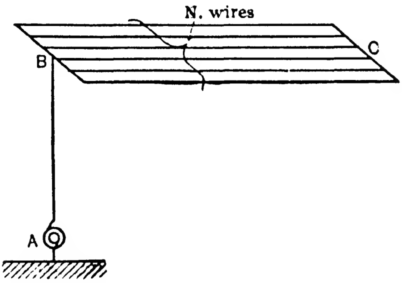


FIG. 69.—An antenna with a wide top has a natural wave length considerably greater than four times the extreme length A–B–C; by spreading out the wire A–B (separating the different strands sufficiently and bringing them down in a cylindrical form) the natural wave length may be brought down to very nearly four times the length A–B–C.

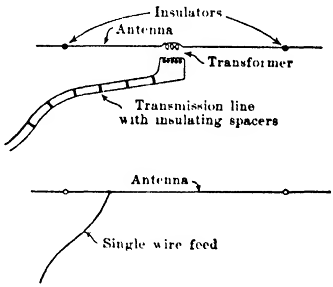


FIG. 70.—A half wave length antenna (frequently called a Hertz antenna) may be supplied with power by means of a transmission line (2 wire) and suitable transformer connected at its midpoint, or it may be supplied by a single wire transmission line, directly connected to the antenna at a proper point.

Wave Length of an Antenna.—Consider now an actual antenna with a flat top. If the top consists of a single horizontal wire of the same size as the vertical wire, then it may be shown in the manner already illustrated for the simplest antenna that, assuming uniformly distributed capacity and inductance throughout, the total length ABC, Fig. 68, represents one-quarter of the fundamental wave length of the antenna.

If, on the other hand, the top consists of a number of horizontal wires, as in Fig. 69, then the problem is somewhat complicated, because the capacity and inductance per unit length of the part BC are different from those for part AB. However, in view of the fact that for the part BC the capacity per unit length is kn times [1] that of a single wire, while the inductance per unit length is $1/kn$ times that of a single wire, the product of these two quantities remains the same, and it is safe to take the distance ABC as again being approximately one-quarter of the fundamental wave length of the antenna.

The inaccuracy of this simple rule increases as the form of the aerial departs from the simple one given in Fig. 64. It has been found experimentally that the natural wave length is connected to the extreme length of the antenna (from ground, up lead wire to farthest point of aerial) about as given here.

Vertical wire	4–4.1l
T aerial with small tops	4.3– 5l
T aerial with broad tops	5– 6l
Umbrella aerial	6–10l
Horizontal wire, 1 meter from ground	5l

For a horizontal wire, not connected to earth, and above the earth a distance comparable with its length, Englund found [2] that it oscillates at a wave length equal to 2.12 its length, within 1 per cent. His oscillators were between 200 cm. and 300 cm. long, of $\frac{1}{2}$-inch copper tube.

This type of antenna must evidently have a current node at either end, it must then oscillate in the half wave length mode. A conductor 300 cm. long resonates to a wave length of $3.00\times2.12=6.36$ meters. These horizontal half wave length antennas are frequently called Hertz antennas; they are frequently supplied with power from a single wire, as will be discussed later in this chapter, but more often a small transmission line with transformer, in the middle of the antenna is used for excitation. These two methods are illustrated in Fig. 70.

The capacities of antennas vary from perhaps 0.2×10^{-9} farad to 40×10^{-9} farad; the lower value being for small portable field antennas and the higher value for large high power stations. The ordinary ship antenna has a capacity between 1×10^{-9} and 2×10^{-9} farad.

[1] k is a constant less than unity; it approaches unity as the different wires of the antenna are spaced farther apart.

[2] B.S.T.J., July, 1928, p. 404.

In the case of a coil radiator the capacity of the condenser C, Fig. 71, is generally very large as compared with the distributed capacity over the conductor $ABCD$, hence the latter may be neglected and the fundamental wave length of the circuit may be obtained from the inductance of the coil and the capacity C.

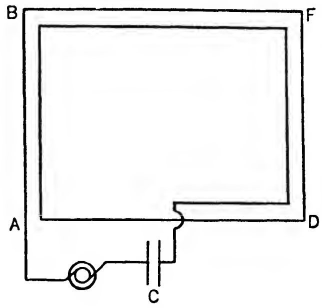


FIG. 71.—The natural wave length of a coil antenna is seldom used; the wave length is calculated from the value of L of the coil and the amount of capacity in C.

Current and Voltage Distribution in Antenna for Various Loadings.—The expression "Loading of an antenna" applies to the insertion of an inductance or a condenser in series therewith for the purpose of changing the fundamental wave length of the antenna circuit. This is best understood by referring to the curves of Fig. 72, which give the reactance, at different frequencies, for an antenna, for a coil, and for a condenser. The antenna reactance curve A is the same as the first section of the curve of Fig. 67, the curve for the inductance is a straight line, since inductive reactance varies directly with the frequency, and the curve for the condenser is an equilateral hyperbola, since condensive reactance varies inversely as the frequency.

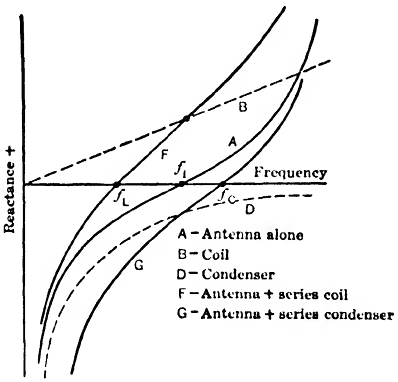


FIG. 72.—The diagram of reactance of an antenna (A), a coil (B), and a condenser (D), shows how the natural wave length of an antenna circuit is changed by adding loading coil or shortening condenser in the base of the antenna.

If the antenna and the coil are connected in series it is plain that the total reactance will, for any frequency, be the algebraic sum of the two individual reactances; a similar thing applies to the case where a condenser is connected in series with the antenna. The resultant reactance curves are shown as F and G. Now, considering the three curves A, F, G, it will be seen that the antenna alone has a natural frequency of f_1, the antenna with the coil in series has a natural frequency of f_L, and the antenna with the condenser in series has a natural frequency of f_C. Thus, the effect of the series inductance is to make the natural frequency of the entire antenna circuit smaller (larger wave length) than that of the antenna alone, and vice versa for the case of the series condenser.

It will be noted that by making the slope of the curve B very great (large inductance) the antenna circuit may be caused to have a very much lower fundamental frequency than that of the antenna alone, the limit being zero. In the case of the series condenser it will be observed that no matter how large we make its reactance (how small its capacity) the maximum frequency obtainable is twice that of the fundamental frequency of the antenna proper. Thus, if an antenna has a natural wave length of, say, 500 meters, it is impossible to change this to anything less than 250 meters by placing a condenser in series with the antenna.

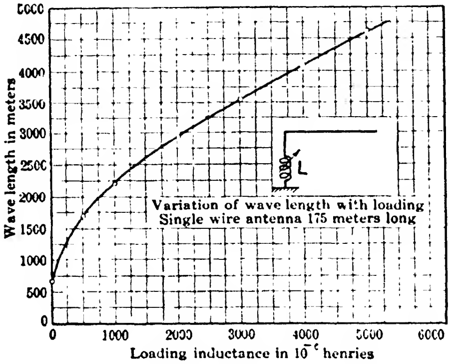


FIG. 73.—Effect of loading coil on the wave length of a single-wire antenna.

The changes which take place in the natural wave length of an antenna, as various coils or condensers are used in series with it, are shown in Figs. 73, 74, and 75. A single-wire antenna was used in the test, about 175

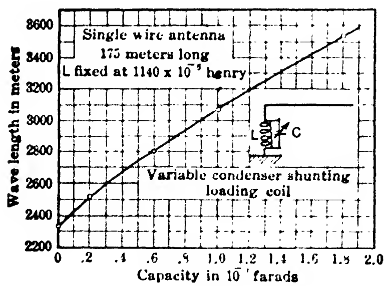


FIG. 74.—Effect of shunting the loading coil of Fig. 73 by a variable condenser.

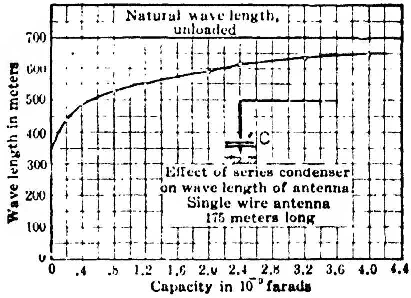


FIG. 75.—Effect of putting a variable condenser in series with the base of the antenna.

meters long, having (unloaded) a natural wave length of 700 meters. As a variable inductance, in series with the ground connection of the antenna, was changed, the natural wave length of the loaded antenna increased as shown in Fig. 73. Then keeping the value of the loading inductance fixed at 1140 μh a variable condenser shunted around this load coil brought about the changes in wave length shown in Fig. 74.

The effect of putting a "short-wave" condenser in series with the base of the antenna is shown in Fig. 75; it will be seen that with no capacity in series with the base of the antenna (that is, the lower end of the antenna merely left free, connected to nothing) the natural wave length decreased (by extrapolation of the curve) to half the natural wave length of the grounded antenna.

It has already been stated that an antenna is generally used at frequencies lower than its fundamental, and therefore antennas have generally a loading inductance inserted in series.

Current and Potential Distribution over Antenna.—We will consider the following cases:

(1) Simple antenna (single vertical wire) with no loading inductance or series condenser.
(2) Simple antenna with loading inductance in series.
(3) Simple antenna with condenser in series.
(4) Commercial antenna with large top and no loading inductance or series condenser.
(5) Commercial antenna with loading inductance.
(6) Commercial antenna with condenser in series.

It is understood that in every case the antenna circuit is operated at the fundamental frequency of the circuit, for at this frequency the reactance is zero, the resistance is a minimum, and the current a maximum.

Case (1). Eqs. (42) and (43), of p. 932, give

$$\dot{E} = \frac{\dot{E}_o}{\cos al} \cos ad$$

and

$$\dot{I} = -\frac{\dot{I}_o}{\sin al} \sin ad$$

and indicate that, since the antenna height (l) is equal to one-quarter of a wave length, the voltage and current curves will be as shown in Fig. 66, the curves being sinusoidal; the current curve will be one-quarter of a complete sine wave.

Case (2), Fig. 76. Here the λ of the entire circuit will be larger than that of the antenna alone, therefore the antenna height will represent less than one-quarter of the wave length. Furthermore the current through the inductance will be constant, but the voltage over it will vary from DK to AH. Hence the voltage E'_o at the beginning of the antenna wire will be much larger than in case (1), and the insulators at the point A will need to be such as to stand a large voltage.

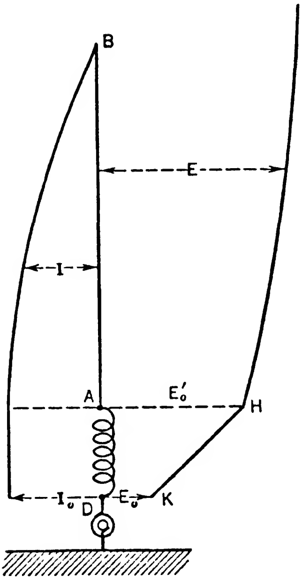


FIG. 76.—Voltage and current distribution in simple antenna with loading coil.

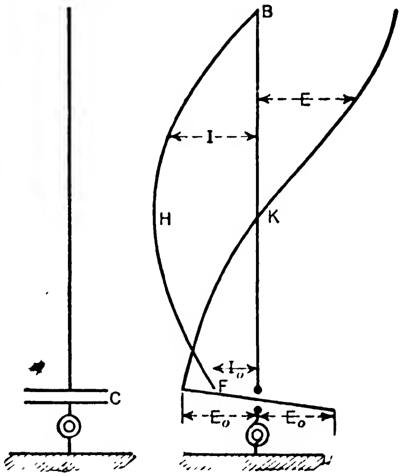


FIG. 77.—Voltage and current distribution in simple antenna with shortening condenser.

Case (3), Fig. 77. Here the λ of the entire circuit is less than that of the antenna alone; hence the antenna height will represent more than one-quarter of the wave length. The current curve, therefore, has its zero at B, its maximum at H and decreases to I_o at F; it has the same value on both sides of the condenser. The voltage curve is a maximum at B, zero at K and becomes negative thereafter; however, it again changes sign over the condenser.

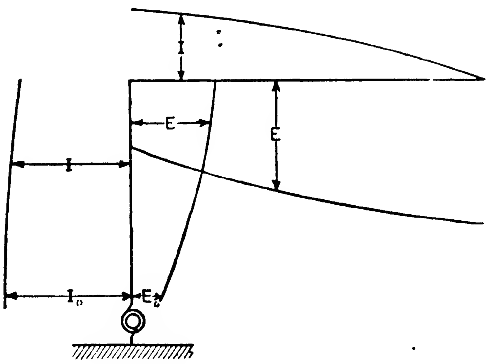


FIG. 78.—Voltage and current in unloaded inverted L antenna.

Cases (4), (5) and (6), illustrated in Figs. 78, 79, and 80, are analogous to cases (1), (2) and (3), respectively, except that the distribution of voltage and current takes place over the entire antenna length and not over the vertical part alone. The result of this is that the vertical part has a current of more nearly constant effective value over the entire height; of course this result is especially desirable in view of the better radiation produced by a uniform current over the vertical wire.

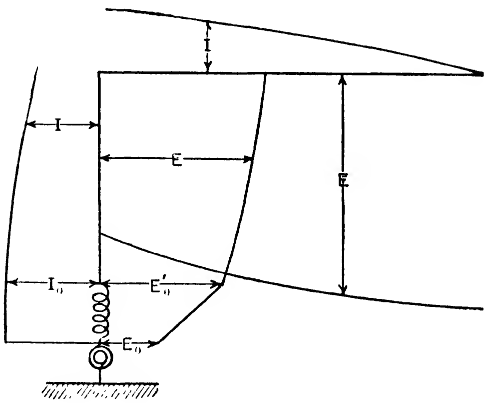


FIG. 79.—Current and voltage in inverted L antenna having a loading coil.

Experimental curves of voltage and current distribution for a low-frequency circuit, representing at low frequency what an antenna does at high frequency, bear out the theoretical predictions already discussed, except that, whereas in the theoretical curves of Figs. 77 and 80 we have shown the effective value of the voltage to actually become zero at points marked K, this does not happen in the experimental curves.[1] The reason for this lies in the fact that the theoretical curves have been plotted on the basis of Eq. (37) which takes no account of the resistance of circuit, while actually there is resistance. The effect of the resistance upon the effective value of the voltage along the antenna is, generally, to make it impossible for it to become zero for, at the nodal point, where the voltage should be zero, there is energy flowing past the nodal point to supply the losses for the rest of the antenna, and in order for this to take place the voltage must be greater than zero. In the case of no resistance the voltage along the antenna has different effective values, but the same phase for the same half-wave and changes in phase through

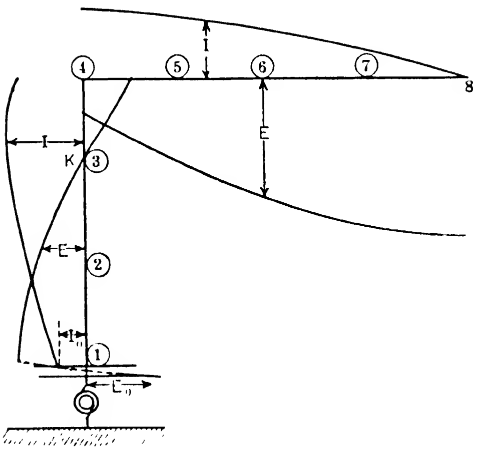


FIG. 80.—Current and voltage in inverted L antenna having shortening condenser.

[1] See Morecroft, "Experiments with Long Electrical Conductors," Proc. I.R.E., Vol. 5, No. 6, Dec. 1917.

180° at the point where it passes through its zero value. On the other hand, in the actual case the voltage all along the antenna has not only different effective values, but different phases as well, as may be shown by the vector diagram of Fig. 81 where the vectors represent voltages at different points of the antenna of Fig. 80, the numbered vectors corresponding with the numbered positions on the antenna. At nodal points the voltage would be very small as shown at E_3; its magnitude (for a given impressed voltage) becomes smaller as the resistance of the upper part of the antenna is decreased.

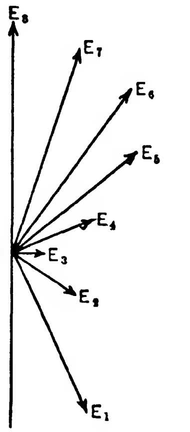


FIG. 81. — Voltage magnitudes and phases of the antenna shown in Fig. 80; at the nodal points of such an antenna the voltage is not zero as the curves of Figs. 77 and 80 would indicate, but a certain small value depending upon the resistance of the antenna.

Antenna Operated at Less Than Quarter Wave Length. — Meissner[1] has advocated the operation of antennas at a wave length considerably less than the natural wave length, that is, less than the wave length it shows when grounded, with no loading coil. He is of the opinion that the signals sent off from such an antenna experience much less fading than do those operated at the quarter wave length frequency or less.

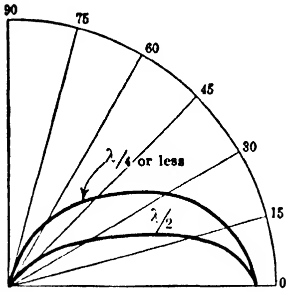


FIG. 82.—Polar diagram of radiation from a quarter wave length antenna and a half wave length antenna (Meissner).

In Fig. 82 is shown the form of the curve of energy distribution for a quarter wave length antenna and a half wave length antenna, according to Meissner. The greater amount of energy travels close along the earth's surface, with the half wave length antenna, thus not only increasing the strength of the direct earth wave, for a given power, but also diminishing the intensity of the sky wave and thus diminishing the interference between the two. It is this interference which produces fading.

In Fig. 83 is shown an effective antenna for broadcast frequencies; it is operated at a wave length of 545 meters, although when grounded its natural wave length is 930 meters. The length from the ground to the spreader is 150 + 22 = 172 meters. The ratio of natural wave length to this length is 930/172 = 5.4. This factor, for antennas of this type, is given on p. 938 as between 5 and 6.

[1] I.R.E., July, 1929, p. 1178.

The field strength distribution of this antenna, with 15-kw. input, is shown in Fig. 84; according to Meissner this shows a considerably better average field strength than any other broadcast antenna of similar general type; two others which he uses for comparision give 12 mv. per meter and 13 mv. per meter at 50-km. distance with 15-kw. input, whereas this one shows 19 mv. per meter. At 100-km. distance this one shows 5 mv. per meter, whereas the other two showed 3.2 and 4.2 respectively. In another of these high-efficiency antennas the height is 150 meters and the operating wave length is 497 meters; this approaches the half wave length condition more closely than the one described above.

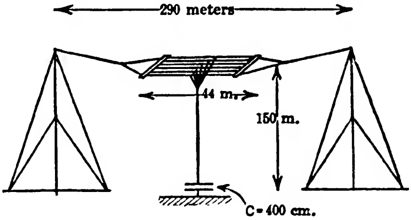
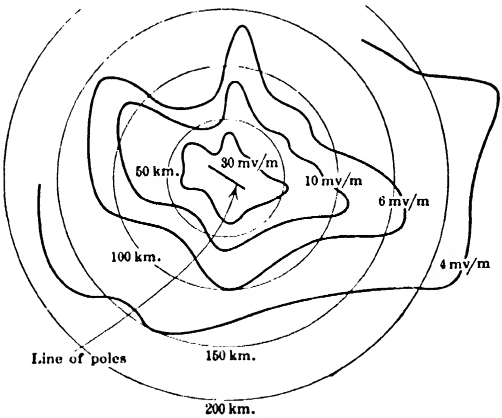


FIG. 83.—An efficient design of antenna for a broadcast station; it is operated at a wave length much less than the natural (quarter) wave length of the unloaded antenna.

Half Wave Length Antenna Not Dependent on Ground Condition.—The maximum current in these half wave length antennas occurs part way up the vertical lead. At the base the current is much less than its maximum value. (It will be recalled that the quarter wave length antenna has a maximum current at the base.) This small current at the base means that the loss in the earth connection will be comparatively small, so the ground network, or counterpoise, if such is used, need not be as extensive or expensive as it is for the quarter wave length condition.



15 kw. input

FIG. 84.—Field strength distribution map of the antenna of Fig. 83, when 15 kw. are being supplied to it.

Effect of Masts.—The steel masts from which the antenna is suspended should always be as far away from the antenna as feasible, otherwise much of the radiated power will be at once absorbed by them. They exert a large influence on the distribution of radiated energy even if they do not absorb much power. It is frequently possible to get much more current into the antenna, with a given power supply, by grounding the towers, if they have been built on insulating pedestals as is frequently the case. However, it is often found that, even with this increased current in the antenna, the radiated field is actually less. This is caused by the currents flowing in the towers neutralizing to some extent the effect of the current in the antenna proper.

If the grounded tower has a natural frequency near to that of the antenna current, a very great shift in radiation distribution will occur with small changes in antenna frequency. When this is lower than the tower frequency the phase of the tower current will be ahead of that of the voltage induced in the tower, and vice versa. The rapid change in phase, nearly 180°, will greatly change the radiation pattern.

In Fig. 85 is shown the shift in pattern of radiation field when the two

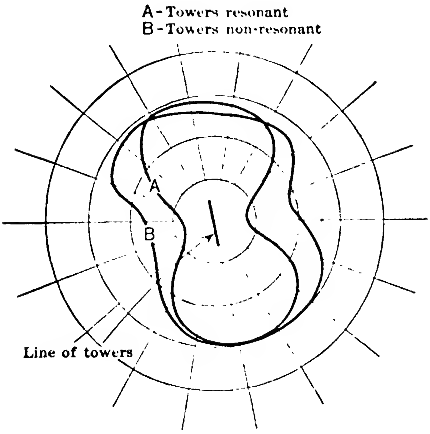


FIG. 85.—Effect of the two steel supporting towers upon the field distribution sent out from the antenna.

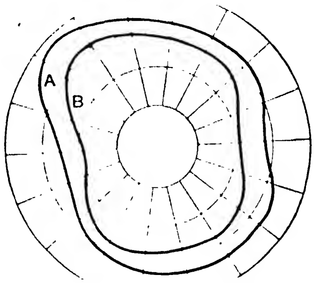


FIG. 86.—Radiation pattern B was obtained from a T antenna, and pattern A was obtained from a single vertical wire, strung from the same masts. This indicates that the steel supporting masts have a much greater influence upon the radiation pattern than does the form of the antenna.

towers between which the antenna was strung were tuned, and when not tuned. In Fig. 84 it will be noticed that the field is asymmetrical, being greater in the line of the towers. By properly adjusting the resonant frequency of these towers, and proper spacing from the antenna in terms of wave length, the pattern may be greatly altered.

Directivity of the T Antenna.—It has been thought that a T antenna shows directive radiation, but results obtained by O'Neill[1] rather disprove this. In Fig. 86 are shown the radiation from two antennas of entirely different shape; A was a vertical wire 240 feet long and B was a T antenna having a vertical lead wire 80 feet long and a T top 240 feet long. The radiation patterns of both antennas are practically the same;

[1] I.R.E., July, 1928, p. 872.

its asymmetry in both cases is quite evidently caused by the action of the towers between which the antennas were suspended, and not by the antenna shape.

Direction Finders.—This is the name given to receiving antennas so constructed as to indicate the direction from which the signals are coming. The simplest direction finder is a receiving coil antenna; it has already been pointed out on p. 886, that such a coil when used as a transmitter will produce the maximum intensity of field in its plane and the minimum at right angles thereto; in a similar manner the coil will, when receiving, have the greatest current produced in its circuit when its plane is in the plane of propagation of the waves and the minimum when its plane is perpendicular to the plane of propagation of the waves. Thus, if the coil be arranged so that it may be made to rotate with respect to its vertical axis while signals are being received, then when the coil is placed into a position of minimum or zero strength of signals the normal to its plane indicates the direction from which the waves are coming.

It has already been stated that in the case of aeroplanes a coil antenna is sometimes used for receiving, which is kept fixed in position with respect to the aeroplane while the aeroplane is maneuvered until maximum or minimum strength of signals is obtained.

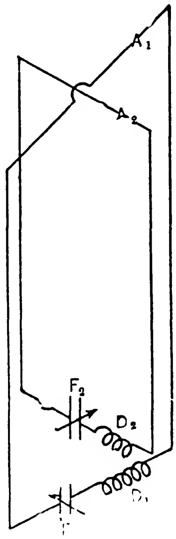


Fig. 87.—Pair of similar coil antennas placed at right angles to each other constitute the Bellini-Tosi direction finder.

In order to obviate the necessity of moving the coil while obtaining bearings Bellini and Tosi invented the so-called goniometer which bears their name. It consists of two similar coil antennas, Fig. 87, at right angles to each other, the antennas being kept stationary.[1] Each of the antennas is connected in series with similar coils D_1 and D_2 and variable condensers F_1, F_2, such as to enable the operator to tune to the incoming waves. The condensers are constructed so that they may both be varied at the same time and by the same amount, in order for both antennas to be simultaneously tuned to the incoming waves. The coils D_1 and D_2 are constructed in two parts as shown in Fig. 88, leaving a space in the middle for a coil K which may be rotated with respect to a line through O as an axis. The coil K is connected to a tuning condenser to which there is attached the detecting circuit.

The signal strength will vary as the coil K is rotated. This may

[1] For another type of directive receiving system see "A New Directional Receiving System," by H. T. Friis, Proc. I.R.E., Vol. 13, No. 6, Dec., 1925.

be shown as follows: Let, in Fig. 89, D_1, D_2 and K, represent the planes of the stationary coils D_1 and D_2 and of the movable coil K, respectively; also assume, for the sake of simplicity, that the coil antennas A_1 and A_2 are placed so that the plane of A_1 is parallel to that of D_1 and the plane of A_2 parallel to that of D_2. It is understood that the coils D_1 and D_2 together with the respective antennas A_1 and A_2 and the condensers F_1 and F_2 (see Fig. 87) are so adjusted that each circuit has a natural wave length equal to that of the incoming waves and the same value of resistance as the other circuit; this means that the circuits

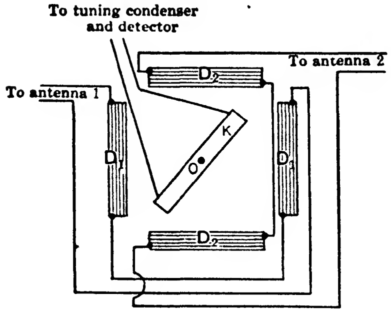


FIG. 88.—Arrangement of coils in the base of the two antennas; coil K may be rotated and by the magnitude of the signal strength induced in it the direction of the sending station ($\pm 180°$) can be obtained.

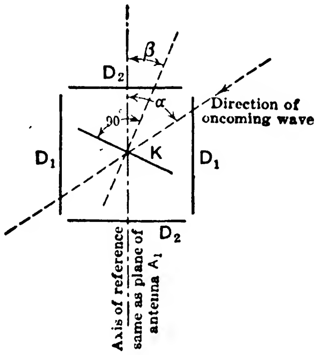


FIG. 89.—Diagram for analysis of the action of the direction finder.

of the two antennas must be exactly similar. Assume that the incoming electromagnetic waves are harmonic and that, therefore, harmonic e.m.f.s, will be induced in A_1 and A_2 which will, in turn, produce harmonic currents in their respective circuits.

Let α = the angle made by the direction of incoming waves with the plane of the A_1 antenna;

β = angle made by the normal to the plane of the revolving coil K with the plane of the A_1 antenna;

i_1 = instantaneous value of current in circuit A_1-D_1-F_1 (Fig. 87);

i_2 = instantaneous value of current in circuit A_2-D_2-F_2 (Fig. 87);

e_1 = instantaneous value of e.m.f. induced in K by current in D_1;

e_2 = instantaneous value of e.m.f. induced in K by current in D_2;

e = instantaneous value of total e.m.f. induced in K by the simultaneous action of D_1 and D_2;

I_m = maximum value of the current which would flow in A_1-D_1-F_1 or A_2-D_2-F_2 if either were placed with its plane parallel to the direction of the incoming waves;

ω = angular velocity of currents flowing in A_1–D_1–F_1 and A_2–D_2–F_2;
M = coefficient of mutual induction between K and either D_1 or D_2 when the plane of K is parallel to either D_1 or D_2.

It was stated on p. 919 that the effective value (the same applies to the maximum value) of the current flowing in a receiving coil antenna, whose plane is inclined to the direction of the incoming waves, is equal to that which would flow, were its plane parallel to the direction of the waves, multiplied by the cosine of the angle which the direction of the waves makes with the plane of the coil. In our case, therefore, we have:

$$i_1 = I_m \cos\alpha \sin\omega t. \qquad (53)$$

By imagining that D_1 is rotated (counter-clockwise) until it coincides with position shown for D_2, we see that the equation for current in coil D_2 must be

$$i_2 = I_m \cos(\alpha + 90) \sin\omega t = -I_m \sin\alpha \sin\omega t \qquad (54)$$

From the well-known law of electromagnetic induction

$$e_1 = -M \sin\beta \frac{di_1}{dt}. \qquad (55)$$

$$e_2 = -M \cos\beta \frac{di_2}{dt}. \qquad (56)$$

Substituting in (55) and (56) the values of i_1 and i_2 of (53) and (54) we have:

$$e_1 = -\omega M I_m \cos\alpha \sin\beta \cos\omega t, \qquad (57)$$

$$e_2 = \omega M I_m \sin\alpha \cos\beta \cos\omega t, \qquad (58)$$

and

$$e = e_1 + e_2 = -\omega M I_m \cos\omega t(\cos\alpha \sin\beta - \sin\alpha \cos\beta). \qquad (59)$$

The maximum value of e for a given value of α and β evidently occurs when $\cos\omega t = 1$ or

$$\text{Max. value of } e = \omega M I_m (\cos\alpha \sin\beta - \sin\alpha \cos\beta). \qquad (60)$$

Since the maximum value of the current flowing in the coil K is directly proportional to the maximum value of e, and since this latter changes as the angle β is changed, i.e., as the position of K changes, it follows that the signal strength will vary as K is rotated about its axis.

We may now find the values of β which will make the signal strength zero or a maximum respectively; this will occur when the value of the parenthesis of Eq. (60) is zero or a maximum.

We can put

$$\cos \alpha \sin \beta - \sin \alpha \cos \beta = \sin (\beta - \alpha)$$

and then get

$$\sin (\beta - \alpha) = 0 \text{ when } \beta - \alpha = 0^\circ \text{ or } 180^\circ$$

$$\text{from which } \beta = \alpha \text{ or} = 180^\circ + \alpha. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (61)$$

$$\sin (\beta - \alpha) = \text{maximum when } \beta - \alpha = 90^\circ \text{ or } 270^\circ$$

$$\text{from which } \beta = 90^\circ + \alpha \text{ or } = 270^\circ + \alpha. \quad . \quad . \quad . \qquad (62)$$

We may therefore state that extinction of the signals will take place when the normal to the plane of the coil K is parallel to the direction of the incoming waves, and that maximum strength of signals will result when the normal to the plane of K is at right angles to the direction of the incoming waves. It will be noted that, in this particular case, where D_1 and D_2 are parallel to A_1 and A_2, respectively, the results are the same as if the whole system of coils were reduced to the coil K alone used as a coil antenna; for, when the plane of K is perpendicular to the direction of the waves, the strength of signals is a minimum, and when the plane of K points towards the direction of the waves, the strength of signals is a maximum.

A discussion similar to the one given above may be applied in a similar manner and with similar results to the case of damped waves. Of course it is plain that the results expressed by Eqs. (61) and (62) are vitiated by the existence of any dissimilarity between the circuits A_1-D_1-F_1 and A_2-D_2-F_2. In order to avoid any dissimilarity as much as possible, even at the expense of sensitiveness, the condensers F_1 and F_2 are often dispensed with, and the circuits are thus made aperiodic.

Direction Finders in Navigation.—By fitting coil K with a suitably calibrated dial and rotating the coil until weakest signals are obtained, the direction of the incoming waves may be determined with a comparatively small percentage of error. Use is continually being made of direction finders for determining the position of a ship or aircraft of some kind. Thus, a ship S which is nearing the port may get her bearings quite accurately in one or two ways, as indicated below. (See Fig. 90.)

(*a*) The ship may be fitted with a directional receiver, and the stations A, B, C, D may be fitted with non-directional transmitters continually sending out different identifying letters. The operator on board the ship is assumed to know the positions of the stations A, B, C and D on his chart. He would obtain the angles α, β, γ (see Fig. 90) by manipulating his directional receiver. By plotting the points A, B, C, D and the angles α, β, γ the position of the ship may be obtained.

(*b*) The ship may be fitted with a nondirectional transmitter continually sending out some identifying letter, and the stations *A*, *B*, *C*, *D* may be fitted with directional receivers. The operators at *A*, *B*, *C*, *D* would, by manipulating their directional receivers, obtain the angles which the lines *SA*, *SB*, *SC*, *SD* make with the north and south line and report these angles by telephone to a central station *F*, where the angles are plotted and the position of the ship is determined. Station *F* will then transmit the position of the ship by radio to the operator on board the ship.

This latter method is the one used in the port of New York and seems to be preferable to the former, in so far as this requires the presence of a skillful operator, capable of plotting the ship's position, on board each ship, whereas in the other case all the plotting is done in one single central station, where much greater accuracy may be obtained.

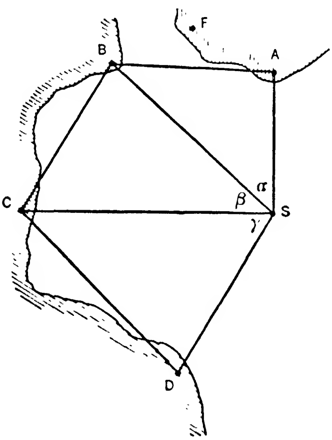


Fig. 90. Arrangement of shore stations around a port to furnish radio compass service to incoming ships.

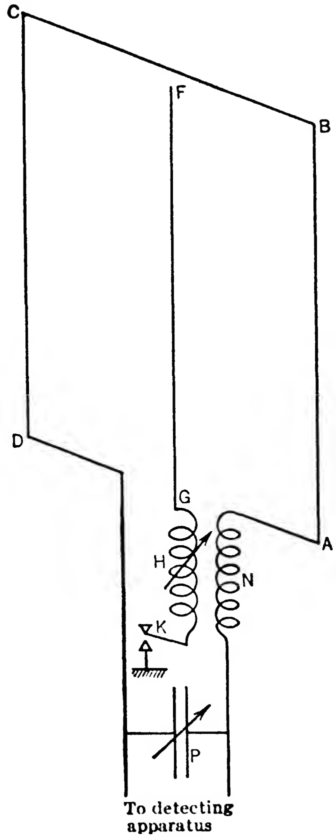


Fig. 91.—To eliminate the 180° uncertainty it is necessary to use a simple antenna in connection with the coil antenna.

So far we have shown how, by means of the single-coil antenna or by means of a *goniometer*, we may be able to determine the plane parallel to which the electromagnetic waves are acting; but we have not yet determined the exact direction of the incoming waves. Thus, we have been able to find that the waves may be acting along the line AB, but not whether they are coming from A or from B; this determination is technically known as the "elimination of the 180° uncertainty." In most instances the direction from which the waves are coming is known, especially in communication between ship and shore and vice versa; but sometimes this is not the case.

In order to eliminate the 180° uncertainty the single-coil antenna or the double-coil antenna of a goniometer is accompanied by a vertical-wire antenna located in the axis of the coil or coils, as shown for the case of the single-coil antenna of Fig. 91, where $ABCD$ is the coil antenna, FG the vertical-wire antenna, connected to ground in series with the tuning inductance H and the key K. The inductance H is loosely coupled to the coil N inserted in series with the coil antenna. The operation of obtaining the direction of the incoming waves would be as follows:

(1) With key K open and the coil antenna turned into some position where the signals may be easily heard, tune the coil antenna circuit to the incoming wave-frequency by means of condenser P.

(2) Close K, and, without changing condenser P, adjust H until the circuit of the vertical wire antenna is tuned to the frequency of the incoming waves, which will be denoted by maximum noise in the receivers connected in the detecting apparatus.

(3) Again open key K. Turn the coil antenna until the signals disappear or become a minimum. The normal to the plane of the coil when in this position represents a line parallel to the direction of the incoming waves.

(4) With key K still open turn the coil antenna 90° from position of (3). Maximum signal strength will then be obtained.

(5) With the coil antenna in the position of (4) depress key K. The signal strength will either increase or decrease relative to that of (4), depending upon the exact direction from which the waves are coming. If the signal strength decreases upon closing K the waves are coming from a certain direction, and if it increases the waves are coming from the opposite direction. Whether it is one direction or the other may be told by previously calibrating the entire apparatus. Waves are used for this calibration which are known to come from a definite direction.

The reason for the behavior of the vertical-wire antenna together with the coil antenna is as follows: Consider Fig. 92 and let the arrows represent the assumed positive directions of the electromotive forces in the wires AB, FG, CD. Let the direction of the incoming waves be as represented by W, and let the plane of the coil be parallel to the direction of the waves.

Call E_1 = effective value of e.m.f. produced in wire AB due to waves W;
E_2 = effective value of e.m.f. produced in wire FG due to waves W;
E_3 = effective value of e.m.f. produced in wire CD due to waves W;
α = angle equivalent to distance S_1 between AB and FG, and between FG and CD;
E = effective value of total e.m.f. in coil antenna due to waves W;
I_2 = effective value of current produced in the vertical wire antenna;
E_n = effective value of e.m.f. induced into N by the current in H.

Since the waves strike wire AB first it is plain that the e.m.f. produced therein will be ahead of that of FG and CD and, therefore, the various e.m.f.s will be as shown in Fig. 93 below, where:

$$\dot{E} = \dot{E}_1 - \dot{E}_3$$

It is plain that no matter what the angle α the vector E will always be at right angles to E_2. The current I_2 will, since the wire antenna is tuned to the incoming waves, be in phase with the e.m.f. E_2. The e.m.f. E_n induced in N will be 90° behind the current I_2 or 180° from the e.m.f. E. Since the total e.m.f. producing the current in the coil antenna is $E - E_n$, this e.m.f. will, in this case, be OA, less than if the coil antenna alone were acting, when the total e.m.f. would be E.

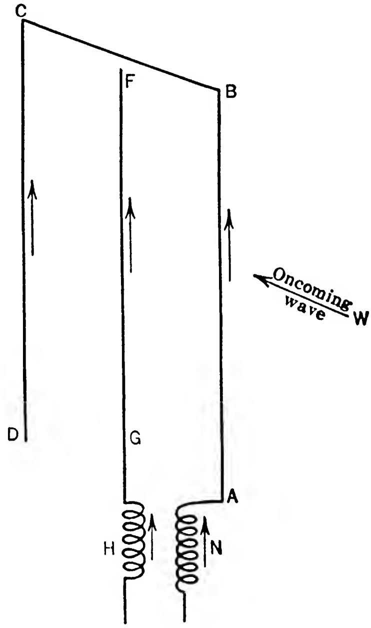


Fig. 92.—Direction of assumed positive e.m.f. induced in the conductors of the two antennas of Fig. 88.

Now consider the case when the waves are coming from the opposite direction to W. Let the symbols: E'_1, E'_2, E'_3, E', I'_2, E'_n represent quantities corresponding to E_1, E_2, E_3, E, I_2, E_n, with the waves from the direction opposite to W. In this case the waves will strike conductor CD first, and hence the e.m.f. produced therein will lead the e.m.f.s of FG and AB. The vector diagram will then be as shown in Fig. 94. As before $\dot{E}' = \dot{E}'_1 - E'_3$ and will be always perpendicular to E'_2. The e.m.f. E'_n will now be in phase with E' and the total e.m.f. $(\dot{E}' + \dot{E}'_n)$ producing the current in the coil antenna will, in this case, be OA, larger than if the coil antenna alone were acting, when the total e.m.f. would be E'.

Thus it has been shown that if the waves are coming from W, Fig. 92, the action of the current in the vertical-wire antenna is to diminish the current in the coil antenna (and hence the strength of signals), while if the waves are coming from the opposite direction the action of the vertical-wire antenna is to increase the strength of the signals. It will be understood that whether the signal strength is increased or decreased by the action of the vertical-wire antenna will depend not only upon the direction of increasing waves, but also upon the direction of the winding on the coils H and N and the position of these coils relative to each other. This is the reason why the entire apparatus has to be calibrated beforehand. In the case of a goniometer the vertical wire antenna is coupled to both of the coil antennas, and the manipulation of the apparatus is similar to that for the single-coil antenna.

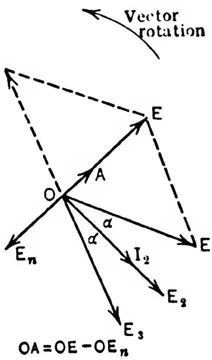


FIG. 93.—The e.m.f. acting in the coil antenna (Fig. 92) is the vector difference of the e.m.f. in its two sides and is shown at OE; current flowing in the simple antenna is shown at OI_2 and this induces a voltage in the coil antenna equal to OE_n.

Incomplete Extinction of Signals. — Unless special precautions have been taken coil antennas do not give zero signal, in any position; the signal goes to a minimum, but is not extinguished. This effect is produced by the coil acting to some extent like a simple antenna. The two wires leading from the coil to the detecting apparatus unbalance the coil electrically, one of them going directly to ground (filament circuit of detecting tube) and the other connecting to ground only through a very high impedance. This asymmetry is sufficient to prevent a " silent " setting to be made with the coil, because the antenna effect gives an e.m.f. 90° out of phase with the coil effect. By a suitable auxiliary circuit it is possible to eliminate this antenna effect, thus getting a more accurate setting, if necessary.

The lack of complete extinction may also be due to the fact that there are two waves acting on the coil at the same time, one that came directly along the ground and the other having reflected from the sky. As these two waves may not come from the same direction it is quite likely that complete extinction is found impossible. In Chapter IV, Fig. 28, it was shown that when the compass bearing went through large errors the minimum signal setting was broad, indicating that the sky wave and direct wave were producing interfering effects.

Rotating Beacon on Shore.—In one type of beacon suggested a loop antenna on shore rotating exactly once a minute sends out a characteristic signal (such as N) when the axis of the coil is pointing north. It then sends out its specified code signals as it revolves, and of course any ship listening in on its ordinary non-directive receiver hears the signal wax and wane. By taking the time required for the signal to go to zero, after the N has been sent, the direction of the beacon loop at this time is at once calculated. Thus, if the ship hears a minimum signal 10 seconds after the N signal has been sent, the beacon is N 60° E (or S 60° W) of the ship. The beacon turns east from north, so 10 seconds after it sends the N signal the axis of its coil must be 60° east of north. Which of these two opposite bearings is correct will presumably be known by the navigator.

It might so happen that the listening ship is due south (or north) when the N signal is sent, and so the signal is not heard. To offset this difficulty another characteristic signal (such as E) is sent when the axis of the beacon coil is pointing east.

Reliability of Direction Finders.—The precision with which a direction-finding receiving coil can be set (under laboratory conditions) is probably less than 1°; in general an operator can set more precisely for minimum signal strength than for maximum unless two coils, at right angles to each other, are used and one of them arranged for commutation. In this scheme the combination of coils is so placed that one coil (the one without the commutator) lies approximately in the direction of the signal, thus being set for maximum reception. The other coil (evidently set for minimum) is connected in series with the first by means of the commutator. The operator then orients the apparatus until the commutation of the one coil makes no difference in the signal strength. The precision of setting with this apparatus is probably much better than 1°.

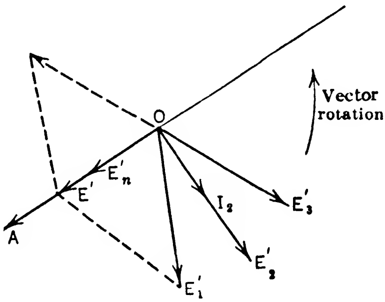


Fig. 94.—This diagram shows how the phase relations of the various e.m.f.s of Fig. 93 change if it is assumed that the signal waves are coming from the opposite direction to that assumed in Fig. 93.

It would seem that it is not worth while to increase the precision of direction finders beyond that now attainable, because of the non-linear propagation of radio waves. With short waves there is not much deviation from straight line propagation, under ordinary conditions; with the long-wave signals, however, the propagation seems to be rather erratic.[1] With signals from 10,000–20,000 meters long, an apparent change in direction of a transmitting station of as much as 90° may occur, the change occurring quite rapidly (as much as several degrees per minute). This variation occurred when the two stations were less than 200 miles apart and might, of course, have been greater if the distance had been greater.

[1] See Bureau of Standards Scientific Paper No. 353, reporting experiments by A. H. Taylor.

In view of Taylor's experiments it seems hardly advisable to use highly directional receiving antennas for communication between long-wave stations. It would seem as though many experiments on attenuation measurement with long waves must be of extremely doubtful value, if the receiving antenna was at all directional.

The charts giving the location of radio compass bearings give the angle in which the observed bearing is reasonably accurate; this angle will generally not include any path in which the signal travels closely parallel to a shore line.

The coils used for direction finding on board ship have certain errors due to the iron work of the ship, just as the old magnetic compass had. These errors in the radio compass must be compensated or at least known, by "swinging ship" as was done with the compass.[1]

A direction-finding coil antenna is of no service at all in the rooms of a modern steel frame work building; no direction can ordinarily be found

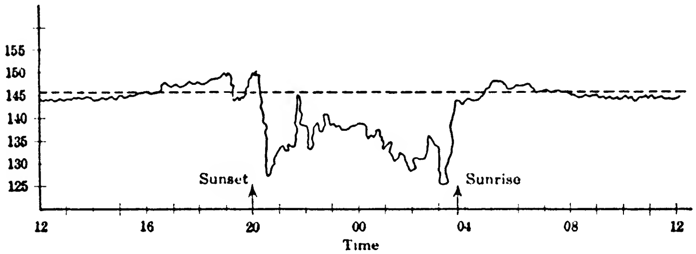


FIG. 95.—Change in radio compass bearing of a long wave station, as measured a few hundred miles away.

in which the signal is extinguished and the maximum signal will generally be received when the coil is closely coupled, magnetically, with one of the steel columns of the building.

In a comprehensive paper, Smith-Rose[2] summarizes the situation with respect to reliability of direction finders, as aids to navigation. He concludes that bearings taken in the daytime, with water only intervening between transmitter and receiver, the error is seldom 2° for distances up to 100 miles. At night time the variation is somewhat greater, so that perhaps not more than 30 miles is allowable if the error is not to be greater than 2°.

Fog seems to have no effect whatever in the accuracy of a direction finder. If the wave travels mostly over land, and the distance is a few hundred miles, great variation in apparent bearing occurs at night, especially on long waves. Fig. 95 shows the apparent direction of St. Assise

[1] For latest type of direction finder see Bureau of Standards Scientific Paper No. 525. "A Unicontrol High-frequency Radio Direction Finder," by F. W. Dunmore.

[2] I.R.E., March, 1929, p. 425.

(14.3-km. wave), France, as measured in England. During the daytime the apparent bearing and true bearing agree within 2°. Shortly before sunset, however, irregularities occur, and large errors, present through the night, persist until an hour or more after sunrise.

A station thousands of miles distant, over the ocean, shows scarcely any deviation, even at night. Thus an American station (16.5-km. wave length) did not change its apparent bearing, in England, by more than 2° throughout the 24 hours.

Cause of Change in Apparent Direction.—It is now well known that fading is generally caused by interference of waves reaching the observer by different routes; they thus have a difference in phase and so may assist or oppose one another. In direction finding, the loop used for picking up the signal receiver waves not only with its two vertical sides but also with its two horizontal sides. Now the vertical sides (which are the only ones generally considered) receive the ground wave, the one that comes directly from the transmitter. The top and bottom of the coil receive a wave that has traveled up to the sky (the Heaviside layer) and been reflected down again. In this reflecting action the direction of its electric field may have been greatly changed, depending upon the character of the ionized region where the reflection occurs. The direction from which the sky wave comes to the receiver loop may have no relation at all to the geographical bearing of the transmitter.

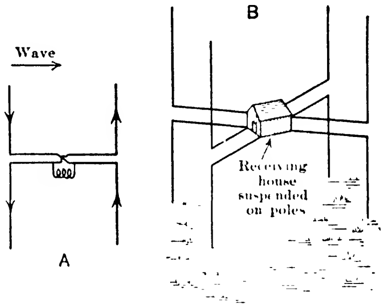


FIG. 96.—A special type of compass station not affected by sky waves; its bearing indications are not subject to errors as are the ordinary coil antennas used in direction finders.

Now the sky wave is greater at night than in daytime, and so interference phenomena and irregularities in direction finding are correspondingly pronounced. Close by a transmitter the direct wave is very strong compared to any sky wave that arrives, so the error produced is small. This is the condition for 100 miles in the daytime, and for 30–40 miles at night, distances within which we have stated the loop gives direction correct to within 2°.

Elimination of Direction Error Due to Sky Wave.—Evidently if the loop had no top or bottom parts to pick up the sky wave this source of error would be eliminated. This is possible by the use of a special antenna system due to Adcock, shown in Fig. 96. He uses two specially formed antennas, crossing one another as shown at *A*. The coil joins the two separate antennas in such a way that current flowing up the front of the antenna flows through the coil in the same direction as does current flowing down the back. And as the wave passes, and directions of current in the front and back reverse, the current through the coil reverses.

But any wave proceeding downward (such as a wave from the sky) will produce no effect on the coil because of the method of connecting the two antennas. The two horizontal conductors will have voltages introduced in the same direction, thus producing no net effect to send current through the coil.

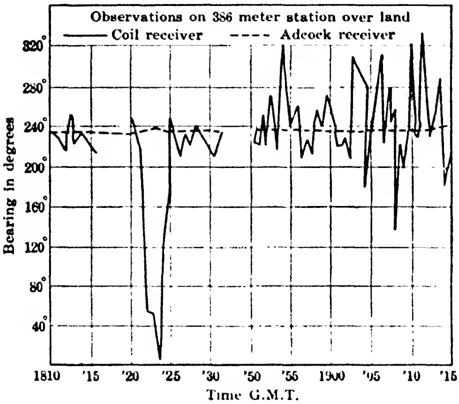


FIG. 97.—Comparative performances of an ordinary coil direction finder and the special one shown in Fig. 96.

By using two such antenna systems, and coils, at right angles, and using the coils in the scheme shown in Fig. 88, a goniometer is obtained the readings of which are independent of sky wave action. In sketch *B* of Fig. 96 is shown such an antenna system. It is installed between four heavy wooden poles, and cross-members on the pole tops serve to suspend the observer's cabin.

The remarkable performance of this device is shown in Fig. 97, taken from the paper of Smith-Rose referred to above. The full line shows the bearing of a station taken with the ordinary radio compass, the loop aerial. It shows variations in the apparent bearing of 300°! At the same time bearings were taken by the Adcock goniometer; these bearings are plotted by the dashed line, and it is seen that they did not change by more than one or two degrees. This test shows conclusively that direction errors with the ordinary radio compass are due to the wave coming down from the sky.

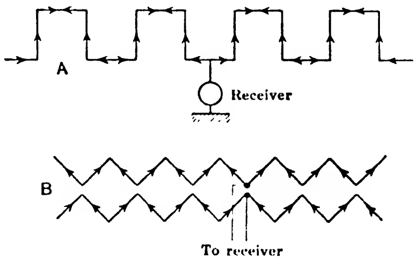


FIG. 98.—Two special forms of antennas showing marked directional properties.

It is of course to be appreciated that the voltage picked up by this Adcock system is very small compared to that picked up by a coil of many turns; it is therefore necessary to use an amplifier with high gain if weak signals are to be picked up. Of course this is not really a serious drawback because radio compass bearings are never required unless the ship is within reasonable distance of the station, so that the signals will be reasonably strong. The one case when the inefficiency of this antenna might work harm is in trying to locate a distress call; this might be from a ship hundreds of miles away and the signal thus weak. Even this, however, does not seem to be enough of a serious drawback to prevent the adoption of this type of compass station.

Special Directional Antennas of Chireix and Mesny.—Two antennas of unusual design have recently come into great prominence; they are due, according to Ballantine,[1] to the French engineers Chireix and Mesny. The first one shown at A, Fig. 98, uses a series of horizontal and vertical members, each a half wave length long, except the end sections, which are only a quarter of a wave length long. The voltages set up in the vertical members all alternate in the same phase for waves coming at right angles to the plane of the antenna, and hence standing waves (with relative current directions as shown in the diagram) are set up, and the receiver is connected from a voltage loop to ground. The standing waves will be set up only if the vertical members all receive their impressed voltages in the same phase, and this can be so only if the wave is coming broadside on the antenna. Furthermore, only those waves having a length equal to two of the antenna members will be able to set up the resonance condition. This then is a highly directive and selective antenna and is the one used to receive transatlantic radio telephone waves. (See p. 845.)

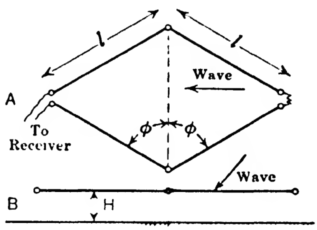


Fig. 99.—An antenna (credited to Bruce) which is similar to one section of the antenna shown in Fig. 98-B.

The other type of antenna, shown at B of Fig. 97, uses two wires made up of right-angled sections, each one-half wave length long. It will receive equally well waves coming from the sky, or waves coming horizontally, according as its plane is vertical or horizontal.

An antenna made up of one section of this type of antenna is described by Bruce,[2] who shows it to be highly directive and selective. A plan and elevation of the antenna are shown in Fig. 99 A and B. Fig. 100 shows a perspective view of one of these antennas, and Figs. 101 and 102 show the directivity of the antenna in vertical and horizontal planes. For these figures the "tilt angle" was 65°, the length of each section of the diamond was four wave lengths, and the height above ground was one wave length. All these factors enter into the type of directivity obtained.

The two wires are connected together at one end through a proper

[1] I.R.E., Sept., 1928, p. 126.
[2] B.S.T.J., Oct., 1931, p. 656.

resistance (a so-called terminating impedance), and the receiver is connected between them at the other end. The advantages of this type of antenna are, according to Bruce:

(1) The directivity of the horizontal diamond antenna can be aimed to some extent, at the most desirable vertical angle, by selecting the proper tilt angle, ϕ.

(2) The high angle directivity tends to discriminate against local disturbances, such as automobile ignition systems, etc.

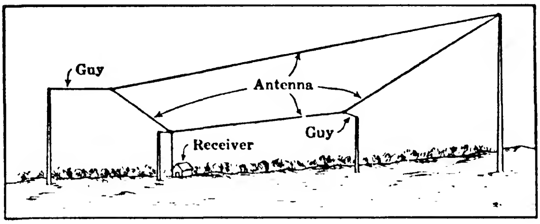


FIG. 100.—Perspective view of a Bruce antenna. At the end opposite to the receiver the two antenna wires are joined by an absorbing resistance, instead of being connected directly together.

(3) The horizontal antenna is stable against changing weather conditions, as horizontally polarized waves, coming from above, are less affected by ground condition.

Of course the use of an antenna of this type, being directive for a wave coming from the sky and having a horizontal electric field (instead of vertical as required for the ordinary antenna or that shown in A, Fig. 97) assumes that the horizontally polarized wave of the signal is as intense as the wave with vertical polarization. This seems to be generally true.

(4) It is possible to use the same antenna (and it is being used) to send and receive on different wave lengths, not possible with other types of directive arrays.

Arrays of Wave Antennas.—At Houlton, Maine, the American end of the 5000-meter transatlantic radio-telephone channel, great ingenuity has been expended in designing an antenna system which shall pick up sufficient of the minute signal energy which reaches this side of the Atlantic and shall at the same time pick up a minimum amount of atmospheric and other disturbances. In other words, a highly directive antenna of large dimensions was required. Because of the 5000-meter wave length the type of directive antenna shown in Fig. 98 cannot be used.

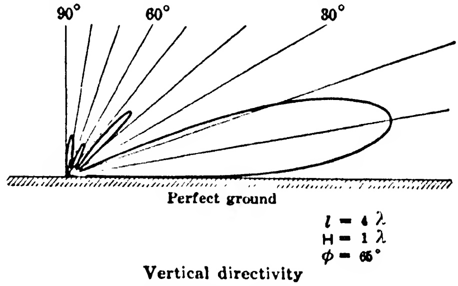


FIG. 101.—Vertical directivity of a Bruce antenna.

So-called wave antennas are used; long wires are strung in lines of poles parallel to the direction from which the signal waves are coming. One such wire alone gives considerable directivity, but by using four of them properly spaced and connected together by suitable phase-changing circuits, etc., the undesired signals and atmospheric disturbances have been greatly reduced and the signal intensity raised. The combination of antennas used is shown in Fig. 103; from this dimensioned plan it will be realized that a directive antenna for long waves is large and expensive. In Fig. 104 are shown the directive properties of one of these wave antennas, and of two of them, A and C of Fig. 103, acting together.

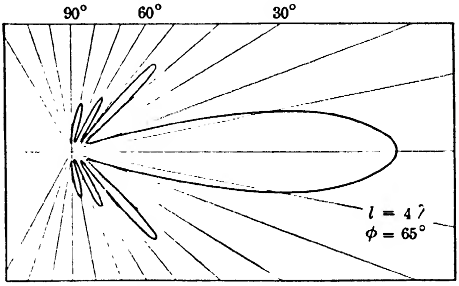


FIG. 102.—Horizontal directivity of a Bruce antenna.

In describing this system[1] the authors state that in moving the receiver from New York to Maine the signal to noise ratio improved 50 times and that the use of the array of antennas gives an improvement over the ordinary antenna, signal to noise ratio, of 100 times. They say that using the wave antenna array in Maine, compared to a coil antenna used in New York, is as effective in improving the signal to noise ratio, as the increasing of the power of the transmitter in England 20,000 times. The receiving system still gives a readable signal when the received field strength is as low as 0.4 microvolt per meter.

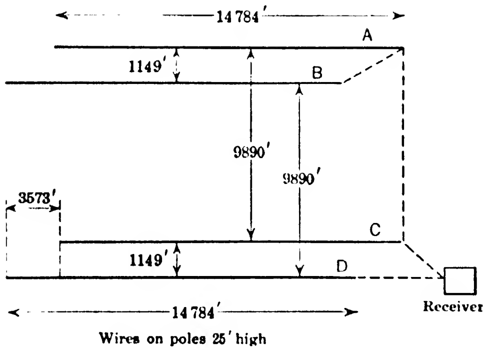


FIG. 103.—Array of antennas used in U. S. for receiving a transatlantic radio telephone signal, employing a wave about 5000 meters long.

In Cupar, Scotland, the other end of the transatlantic channel, a somewhat similar array of wave antennas is used.

Diversity Factor in Radio Reception.—It was shown in Chapter VIII (Fig. 73), that highly directional antennas are used in receiving transatlantic radio telephone messages, but in addition it is now found advisable to use several of these directional antennas, located some distance from each other, all contributing their energy to produce the

[1] Bailey, Dean, and Wintringham, I.R.E., Dec., 1928, p. 1645.

desired signal. To reduce fading and increase the ratio of signal to noise R. C. A. Communications, Inc., now employs three spaced antennas for short wave reception, each connected to separate receivers. Each receiver output is rectified. The rectified output of all three is sent into a common resistor, the drop in which is used to control a locally generated tone (this is for telegraphy).

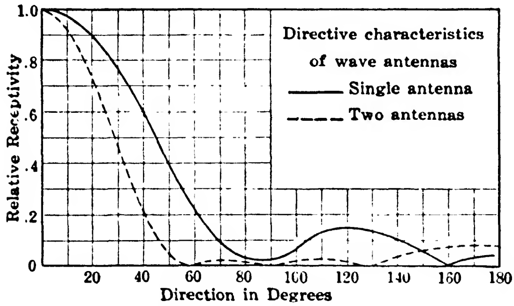


FIG. 104.—Directive characteristic of one of the antennas of Fig. 103, and of two of them properly connected together. By using all four the directional properties are still further improved.

Each receiving antenna is aperiodic and directive. It consists of a two-wire transmission line stretched in the direction from which the signal comes. The wires are close together, and about 15 meters high. A great many horizontal doublets are arranged perpendicular to the direction of this line, and in its plane. Each couples to the line by a pair of condenser plates. The doublets are untuned, the combination of doublets and transmission line receiving equally well waves from 14 to 25 meters. Two of these peculiar antennas (suspended on wooden poles), close together and extending in the same direction, are connected in parallel. Three of these double, untuned, directive transmission line antennas properly spaced from one another act to give the desired freedom from fading and increased ratio of signal to noise. The combination increases the ratio of signal to noise, for European stations, about 32 db. over the ratio given by a single horizontal doublet.[1]

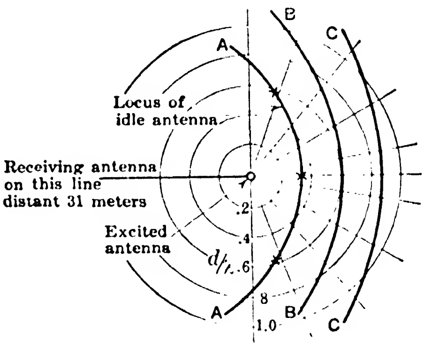


FIG. 105.—This diagram shows the locus of an idle, reflecting, antenna for maximum signal reception at the receiving antenna. Locus B–B gave minimum signal, while A–A and C–C gave maxima.

Effect of Reflections from Antennas.—In short-wave radio much work has been done in guiding the radiated energy along certain channels; this is readily accomplished

[1] "Diversity Receiving System of R. C. A. Communications, Inc." Beverage and Peterson, I.R.E., April, 1931, p. 531.

when the wave length is 15–40 meters where it would be impossible with waves thousands of meters long.

Hertz, the discoverer of radio communication, explains in his book "Electric Waves" how he found it possible to focus the waves generated in his laboratory by either reflecting mirrors, or lenses. Some of the commercial schemes, first advocated apparently by Marconi, are almost exact replicas of the devices Hertz himself used.

The simplest device which acts somewhat like a reflector is an unexcited antenna, similar to the one being used; it must be placed parallel to the one being used and a distance of about $\lambda/4$ behind it.

Englund and Crawford [1] investigated the effect of idle antennas in the neighborhood of an excited one and so studied the reflections accomplished for various positions of the idle antenna. Fig. 105 gives some results from their tests; a receiving antenna was held vertical and fixed in position, 31 meters away from a vertical transmitting antenna; an idle antenna, similar to the transmitter, held vertical, was moved into different positions behind the transmitting antenna. In some positions it increased the response of the receiving antenna and in others it decreased it.

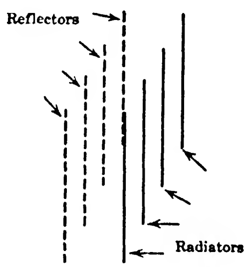


FIG. 106.—Showing an array of reflectors behind a curtain of transmitting antennas.

Line A–A of Fig. 105 shows the locus of the idle antenna to give maximum reading of the receiver, this showing a maximum reflecting action. It will be noticed that when the reflector was directly behind the transmitter its best position placed it about $\lambda/3$ behind the transmitter. The best distance is generally taken as $\lambda/4$, and commercial reflecting antennas are so built.

Line B–B shows the locus of reflector position to give minimum reading on the receiving antenna; here it acted to diminish the forward-going signal wave. Line C–C shows a locus of reflector positions for another maximum reading on the receiver; the effect of the reflector was not as marked as it was when on the locus A–A, showing that the reflector should be as close to the transmitter as possible.

The three points marked x, on locus A–A, show the proper position to put three reflectors to give reasonably selective radiation along the transmitter-receiver line.

Radiation from an Array of Short-wave Antennas.—Shelling [2] has investigated the radiation pattern of a group of vertical, half wave antennas, backed up by a corresponding set of idle reflectors, of similar

[1] I.R.E., Aug., 1929, p. 1277.
[2] I. R. E., June, 1930, p. 913.

spacing and dimensions. The idler is shown in Fig. 106; here four couples are shown. Each antenna (and reflector) is half wave length long and the antennas are placed in a plane, a distance $\lambda/2$ apart. The corresponding curtain of reflectors is a distance $\lambda/4$ behind the antennas.

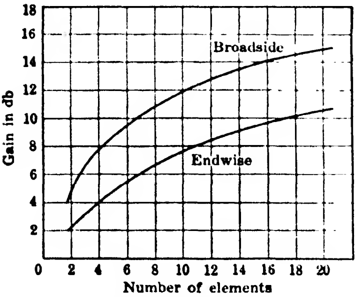


FIG. 107.—Showing how the radiation in front of such an arrangement as given in Fig. 106, varies as the number of elements is increased. The transmitting antennas all excited in the same phase for the "broadside" curve, and every other one 180° out of phase for the "endwise" curve.

Such a scheme tends to concentrate its power in front of the antennas, and the gain in power picked up by a receiver in front of the antenna array (a few wave lengths away) varies with the number of couples used as shown by the curve of Fig. 107. If all the antennas are excited in the same phase the radiation is perpendicular to the plane of the array and the gain is given by the "broadside" curve. If every other antenna is excited 180° out of place, maximum radiation takes place in the plane of the array and the gain to be expected in this direction, as the number of antennas is increased, is given by the "endwise" curve.

These gains represent the ratio of power picked up by a receiving antenna, when a given power is supplied to the array, to the power picked up by the same antenna in the same position, when the same amount of power is supplied to a single vertical antenna in the same position as the array.

In Fig. 108 is shown a horizontal plan of the radiation distribution from an array of 24 half wave-length vertical antennas, arranged in a curtain, with $\lambda/4$ separation between adjacent antennas. The curtain of reflectors was a distance $\lambda/4$ behind the antenna array. This diagram is from an article by Southworth.[1]

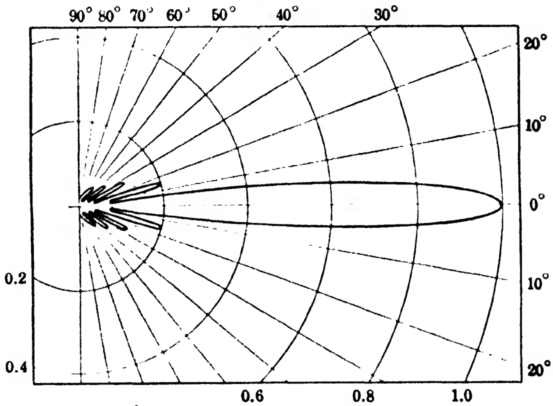


FIG. 108.—Directivity of radiation from an array similar to that given in Fig. 106; there were 24 antennas and an equal number of reflectors.

[1] B.S.T.J., Jan., 1931, p. 63.

A commercial arrangement of antenna arrays was shown in Fig. 70 of Chapter VIII; in each bay (distance between masts) there are about 50 half wave-length antennas, and behind, at a distance $\lambda/4$, an equal number of reflector antennas.

Feeding a Short-wave Antenna.—A Hertz antenna may be excited in the manner shown in Fig. 72, p. 845, by being used as one member of a group all connected together in some way, or it may be fed by an individual feeder. Both of these methods have been indicated in Fig. 70, p. 937. The scheme in which the power is fed into the middle of the antenna, using a two-wire transmission line, has often been called "current feed"; the other scheme shown in Fig. 70 is added a "voltage feed." There was some justification, possibly, for these ambiguous terms owing to the fact that it was thought that the single wire feed required a high voltage and the center feed required much more current than the other. There is not actually much difference between the two methods, when each is properly adjusted. The single feed line is really a double one; the earth serves as the other side.

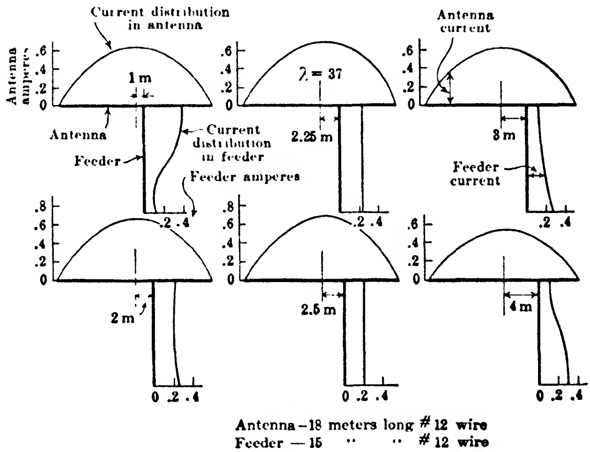


FIG. 109.—Current distribution in antenna and feeder of a single wire fed Hertz antenna.

The surge impedance of a two-wire transmission line is around 700 ohms, and the resistance of a Hertz antenna measured at its middle point, is about 70 ohms. As the load on the end of a high-frequency transmission line should match the surge impedance of the line, it is necessary to use a step-down transformer as shown in Fig. 70; it should have a ratio of about 3 to 1.

The ratio of voltage to current (at a given point) in a Hertz oscillator varies from a small value at the middle (where voltage is low and current high) to a high value near the end. Now the ratio of voltage to current in a circuit is its impedance at that point, so we may say that the input resistance of the Hertz antenna varies from a small value near the middle to a high value near the end. A single-wire feeder should then attach to the antenna at such a point that the apparent resistance of the antenna is equal to the surge impedance of the feed line, which is a few hundred ohms, depending somewhat on the size of the wire.

When the feeder is connected to the proper point of the antenna there will be no reflection of current at the junction point, which means that there are no standing waves on the feeder. Everitt and Byrne[1] have investigated this point, and Fig. 109 shows some of their results.

To know when resonant frequency is being impressed on the antenna they advise measuring the current in the antenna at either side of the feed point, with the feeder connected somewhere near the center of the antenna; the value of these currents should be the same. To find the right point of attachment the feeder should be moved back and forth until there is no standing wave of current in the feeder. When the ammeter reads the same current all along the feeder there are no standing waves, and the feeder is properly connected. They show an ingenious scheme, using a trolley, by which the feeder current (or a definite fraction of it) is easily measured without opening the feeder.

In Fig. 109 there are given six diagrams of current distribution in both feeder and antenna, as the point of feeder attachment was moved along. The frequency impressed on the feeder gave a 37-meter wave; this is the proper wave for an 18-meter antenna oscillating at half wave length. It will be noticed that the current distribution in the antenna is a half sine wave for all points of feeder connection, but that it is a maximum when the feeder is attached 2.5 meters from the center; here the antenna current (at the center of the antenna) is 0.75 ampere while the feeder current is only 0.2 ampere. This value of 0.2 ampere is the same all along the feeder for this connection point; for all other points of connection the magnitude of current in the feeder was different at different points. In Fig. 110 is shown the variation in maximum antenna current as the point of feeder attachment was moved along. According to the results of tests reported by these experimenters, the feeder can have a length many times the wave length of the radiated power, without serious loss of power in the feeder itself.

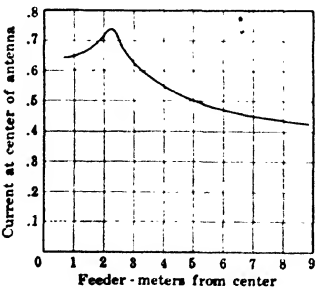


FIG. 110.—Variation of maximum antenna current with position of feeder connection.

Radio Beacons.—The airways of the world today are using radio as an important aid in navigation. The transmitter, on the ground, consists of two single-turn loops, at right angles to each other, as shown in Fig. 111. These loops may be several hundred feet long and about 50 feet high. The vacuum-tube power-supply circuit is connected first to one antenna

[1] I.R.E., Oct., 1929, p. 1840.

and then to the other, by a motor-driven key, using designated letters of the alphabet, one for each loop. The radiation from each loop is of the figure eight form as shown in Fig. 20, p. 886 and indicated again in Fig. 113. In the shaded zones both loops give about the same power.

These beacon stations are about 200 miles apart, and their radiation patterns are arranged to fit into one another as shown in Fig. 114, which shows one airway in the eastern part of the United States. An airplane leaving one zone is presumably able to pick up the next beacon.

It is possible to use two vertical antennas, instead of the loops of Fig. 111, to obtain directional radiation. It has been shown (p. 886) that the loop gives maximum radiation in its plane, due to the interfering action of the currents in its two sides, these currents being 180° apart in phase. Of course, considering the loop as a whole the two currents are *in phase* but considering each side of the loop as a vertical antenna the two currents are *180° out of phase*, one being *up* while the other is *down*.

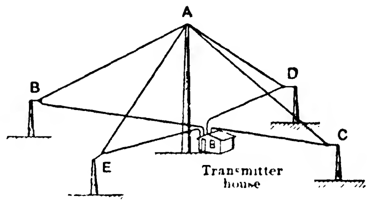


Fig. 111.—Arrangement of transmitting loop antennas for an airway beacon station.

Fig. 112 shows how it is possible to get this effect without constructing an actual loop. The two self-supporting steel towers, *A* and *B*, may be 150 feet high and several hundred feet apart. They are insulated from ground, and are connected together at their bases by the wire, *a*, *a*, shown as an open wire in this diagram, but which may be run in a suitable conduit of copper pipe. The coil *C* permits coupling the radiating system to the power supply, preferably located in a house midway between the towers. This connection scheme evidently results in the currents in towers *A* and *B* having 180° phase separation; by tuning the system *A*, *a*, *C*, *a*, *B* to the desired frequency, an efficient directive radiator is obtained. One pair of towers takes the place of each loop of the arrangement of Fig. 111.

We refer again to Fig. 113 to show the scheme of signaling. The motor-driven key sends the dash of the letter "n" on loop *ED* and then the dot of the letter "a" on loop *BC*; next it sends the dot of the letter "n" on loop *ED* and lastly the dash of the letter "a" on loop *BC*. The result is that in the positions marked "dash zone" where the "a" and the "n" are equally loud, and interspersed as described above, they are not heard as separate letters but as one long dash. If the plane veers to one side or the other of its proper course (the dash zone) either the "a" or the "n" becomes readable out of the dash, and the pilot knows which way to bear to get back on the course.[1] This sequence of "a" and "n" is continually repeated.

[1] Walls, I.R.E. Dec. 1929, p. 2141.

The latest type of airway beacon service[1] uses a goniometer transmitter of the same type as that shown in Fig. 113 but sends out on both loops all the time, one loop with a modulation of 65 cycles and the other modulated at 86.7 cycles. The receiver on the plane is equipped with two reeds excited by the output tube, one reed being mechanically resonant at 65 cycles and the other at 86.7 cycles.

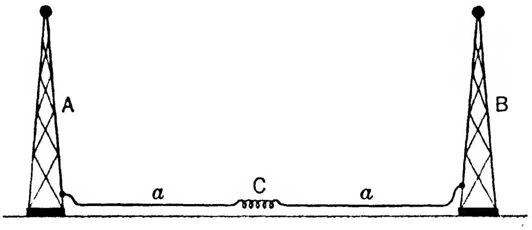


Fig. 112.—A possible directive radiating system to take the place of one loop of Fig. 111. Power is supplied by coupling the power set to coil C.

The system operates as shown in Fig. 115. When the plane is on its course the two reeds vibrate equally; but as it veers to one side or the other, one reed increases its vibration and the other decreases. The difference in vibration amplitudes indicates how far the ship is off the course.

In a further development of this modulation scheme a 12-course beacon has been developed, which uses three goniometer loops at the transmitter, modulated at different frequencies, and a receiver on the plane which has three reeds, instead of two. By the use of special windows and shutters in front of these reeds the operator can tell whether he is on one of these 12 courses within 2°, and which way he is flying on it.[2]

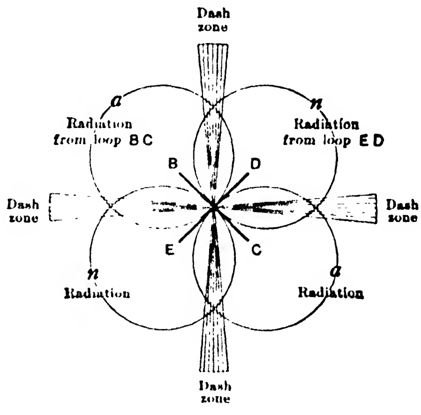


Fig. 113.—Distribution of radiation from the system shown in Fig. 111; one loop is modulated to send out the letter *a* and the other to send the letter *n*.

Dellinger states that up to 75 miles this visual airway beacon navigation is accurate; from 75 to 125 miles variations up to 10°–20° occur, first in one direction and then in the other, but that a skilled operator can allow for this variation and still stay on the course, and that even up to 200 miles distance the beacon is quite good and can be used if reasonable allowance for variations is made.

[1] Dellinger, Diamond, and Dunmore, I.R.E., May, 1930, p. 796.

[2] Dunmore, I.R.E., June, 1930, p. 963; and Diamond and Kear, I.R.E., June, 1930, p. 939.

A radio system to permit a flier to make a "blind landing" at an aviation field has now passed the experimental stage.[1] By the use of three

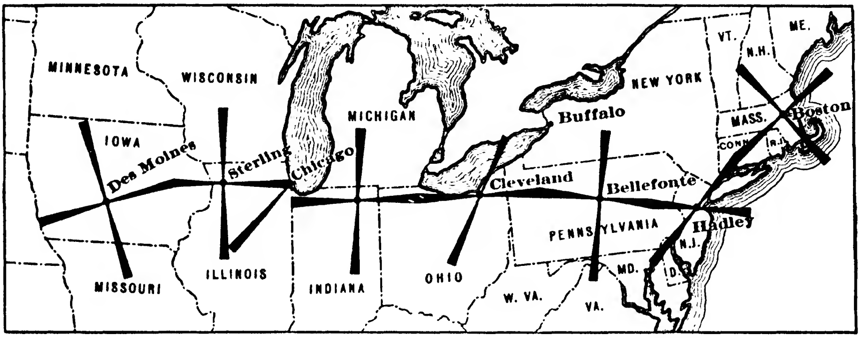


FIG. 114.—Showing how some of the airway beacons in the U. S. interlock.

special transmitters properly located on the field, by using the regular range receiver both visually and orally, and by the additional use on the plane of a special short wave receiver, it has been possible to land a plane successfully without the pilot seeing the ground.

Setting up the Steady State in an Antenna.—It has been noted previously that after the sending key is depressed it may be an appreciable time before the current reaches the value predicted by the steady state equations; some of the effects obtained are shown in Figs. 116, 117, and 118. These oscillograms were obtained on an artificial antenna having inductance and capacity great enough to reduce the natural frequency to such a low value that the ordinary oscillograph could easily

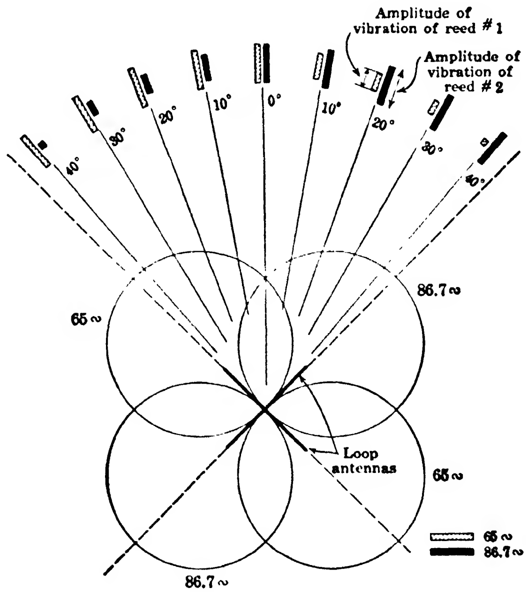


FIG. 115.—Type of radiation and showing of the vibrating reed indicator for the "visual" airways beacon.

[1] A radio beacon and receiving system for blind landing of aircraft.—Diamond and Dunmore, I.R.E., April, 1931, p. 585.

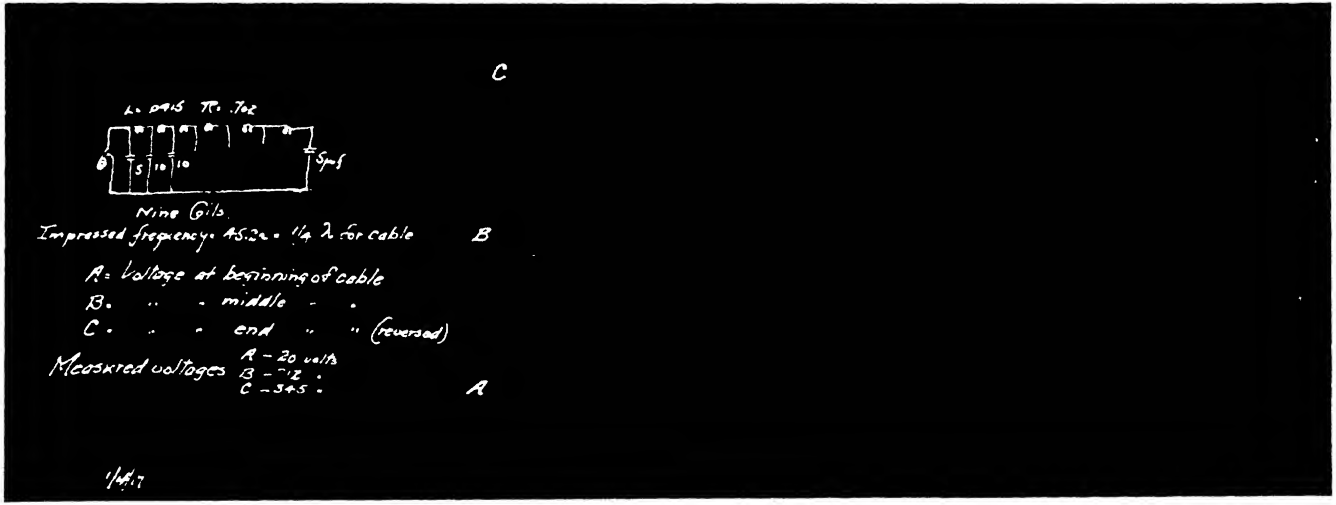


FIG. 116.—The film shows how the "steady state" is set up when an antenna is first excited; curve A shows the voltage impressed, B the voltage at the middle and C the voltage at the open end. Voltage impressed = 20, voltage at B (steady state) = 212, voltage at the end of the artificial antenna = 345.

record the voltages and currents. In getting the film shown in Fig. 116 the impressed frequency was such as to set the artificial antenna into quarter wave-length oscillation; the three curves on the film show the

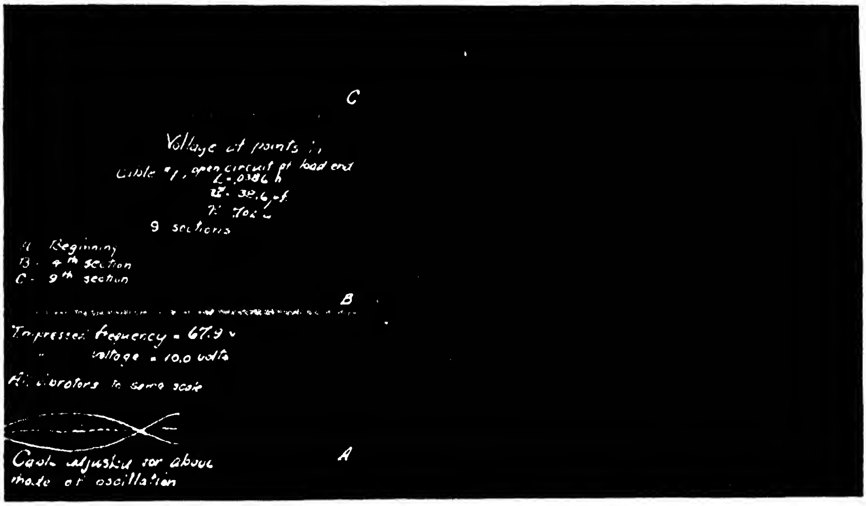


FIG. 117.—Here the artificial antenna was forced to vibrate at three times its fundamental frequency; it will now be noted that the voltages at B and C are in opposite phase in the steady state. From the film it can be seen that the original pulse arrives at C one-half a cycle after passing point B.

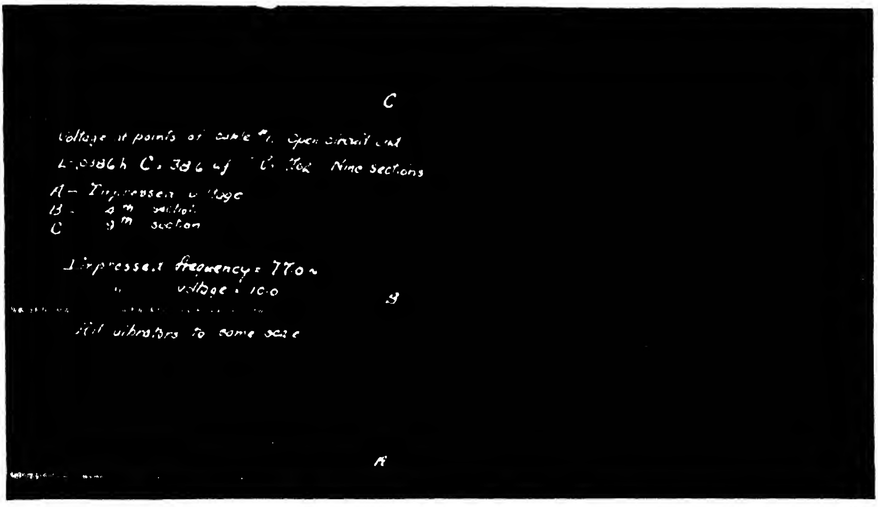


FIG. 118.—While the steady state is being set up some sections of the antenna may have voltages greater than the steady state values.

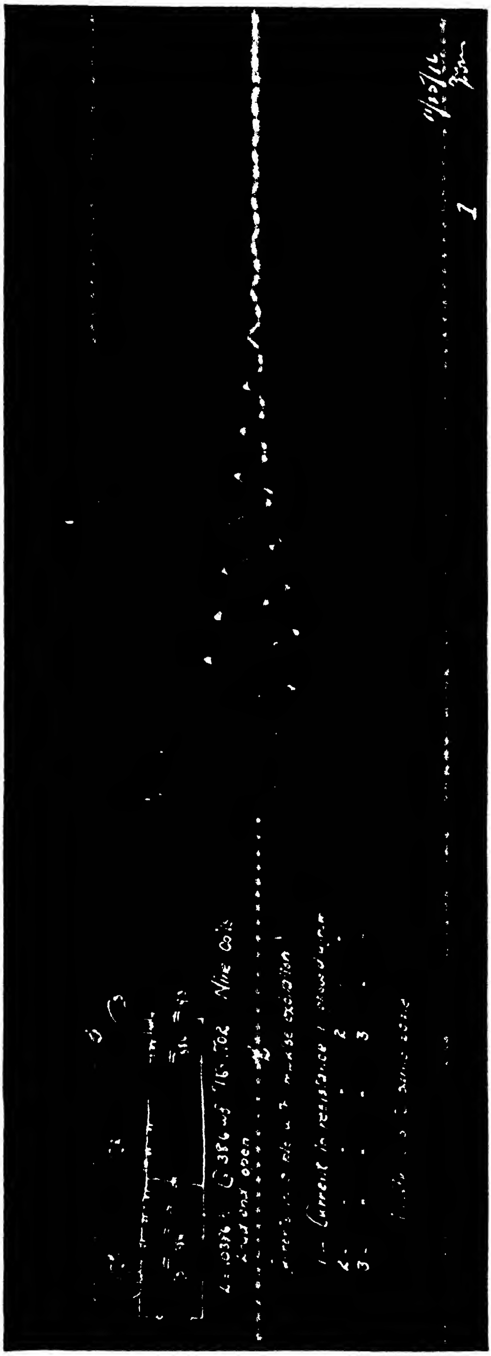


FIG. 119.—This film shows how an antenna is affected by a pulse; a unidirectional pulse of current, shown in the lower curve, was impressed on the antenna, and the current set up in the antenna is oscillatory.

voltage impressed, voltage half-way along the antenna, and voltage at the open end. They are not shown in the film to the same scale, as the voltage at the open end measured 345 volts, that at the middle 212 volts while the voltage impressed was only 20 volts.

It took this artificial antenna about 20 cycles to obtain its steady state values; in an actual antenna it may take 100 cycles or more before the steady state is reached, i.e., before normal radiation is established.

By examination of the film it may be seen that the voltage at the base (a nodal point) is 90° out of phase with the voltage at the end of the antenna; this is in accordance with the ideas brought out in discussing Fig. 81.

In getting the film of Fig. 117 the frequency was increased to three times the value for quarter wave-length oscillation. It may be found from measurement of the film that it took the first pulse three-quarters of a cycle to travel from the beginning of the antenna to the end. Furthermore, it may be noted that in the steady state the voltage at C (end of antenna) is 180° out of phase with the voltage at B, as predicted in Fig. 81, and voltage at A is 90° out of phase with the voltage at B.

In establishing the steady state it may happen that certain sections of the antenna built up to a voltage higher than the steady state value; this is indicated in Fig. 118 in which the impressed frequency had no particular relation to the fundamental frequency of the antenna.

Effect of Pulse Excitation of an Antenna.—In Fig. 119 is shown the effect of putting a square pulse of current into the antenna and then disconnecting the antenna from earth; an oscillatory current is set up in the antenna (as shown by the middle curve) the frequency of which for the conditions used is that of the half wave-length oscillation of the antenna. Thus pulses of "static" always excite an antenna to oscillate at its natural period.

Effect of High-speed Signal Transmission.—As has been previously noted one of the chief advantages of continuous wave sending, as compared with spark, or damped-wave excitation, is the greater selectivity available, and consequent diminution in interference. But as automatic sending and receiving is developed the tendency is to increase the speed and it is to be noticed that continuous wave sending approaches spark wave sending in character as the speed is increased. The width of the resonance curve at a distant receiving station increases as the dashes and dots are made shorter, and follow one another in more rapid succession.

CHAPTER X

AMPLIFIERS

Amplification and Its Measurement.—If we put into a device a certain amount of power P_1 and, by the operation of the device itself, get out of that device a different amount of power P_2 of corresponding variation, the device must be either a source of energy, or an energy sink. If P_2 is greater than P_1 there must be some local energy supply, the output of which is controlled by the amount of power P_1 supplied to the device. If there is no local supply of energy (which is or can be transformed into electrical energy), then P_2 must be smaller than P_1 and this means that part of the input power P_1 has been used up, that is, transformed into some other type of energy, in passing through the device. The ordinary *attenuator*, or circuit for using up a definite known proportion of the power supplied to it, is of this sort.

The ordinary vacuum-tube circuit evidently constitutes an amplifier; the input to the grid circuit is only a small fraction of the amount of power sent off from the plate circuit. A microphone also is an amplifier, although of a different sort. This device takes in power in the form of sound waves and sends off a much greater amount of corresponding power in the form of voltage and current and is thus a converter as well as an amplifier.

It is to be noticed that a transformer which, by virtue of the different number of turns in the two windings, raises the voltage of a power supply is not an amplifier. The amount of power which can be taken from the secondary of a transformer is less than the amount of power supplied to its primary, hence it is really an attenuator, even if it does increase the voltage many times.

The scheme of amplification measurement which is now universally used in electrical communication, and sound measurement, is based upon a logarithmic series, and the unit of amplification is taken as the *bel*. A device gives an amplification of 1 bel when the two powers P_2 and P_1 have a ratio of *ten*. The number of bels of amplification is given by the logarithm to the base ten, of the power ratio. Thus:

$$\text{Number of bels} = \log_{10}\frac{P_2}{P_1}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Formerly the communication engineer used the *mile* as his unit of attenuation or amplification. A certain type of telephone cable, made exactly according to fixed specifications, was used to attenuate P_2 to bring it down to the P_1 level. The number of miles of the cable required to thus equalize the two powers was given as the measure of P_2 in terms of P_1.

Now the bel is too large a unit for many purposes, so the *decibel* (db) is now always used as a measure of amplification. As this unit is quite evidently *one tenth of a bel*, we have, as the fundamental definition of the present amplification scale

$$\text{Number of decibels} = 10 \log_{10} \frac{P_2}{P_1}. \quad \text{. . . . . . .} \quad (2)$$

The decibel is practically the same as the formerly used *mile* of attenuation. It so happens that this unit, the decibel, has a physiological significance; if one sound is 1 db more powerful than another the average ear can just about detect the difference. The louder sound has about 25 per cent more power than the other.

This rather large change in power required to affect the ear (and a similar effect exists with the eye) is probably due to the fact that the ear and eye are very flexible measuring instruments; as the stimulus exciting them increases they automatically cut down their sensitivity by some physiological adjustment. If it were not for this characteristic, it would be impossible for them to respond, without discomfort, to such a great range of stimulus. Thus the ear can accommodate itself to weak sounds (and hear them satisfactorily) which have one-millionth as much pressure as the loudest sound that can be listened to without discomfort. Any one familiar with ordinary measuring instruments knows that it is impossible to read on the scale any effect smaller than one-thousandth of full-scale reading; it thus appears that the ear and eye have a range in intensity response which is one thousand times as great as that of any ordinary measuring instrument.

In the accompanying table are given the power ratio equivalents for a wide range of decibels, abbreviated db. If one power is greater than another it is said to be *up* so many decibels or plus so many decibels; if it is smaller it is *down* or minus so many decibels. Thus 5 watts, compared to 1 watt, is 7 db up or +7db, and the 1 watt, compared to the 5 watts is down 7 db or −7 db.

It will be noticed that this scale of amplification is based on *power* ratio, but it is often used also for voltage and current ratios, with the understanding, however, that the two quantities compared are acting on (or through) the same resistance. Acting on the same resistance the power used varies as the square of the voltage so that if one voltage is 10 times another the power used will be 100 times as great. Hence, if an output voltage is ten times the input voltage, the amplifier which is

Number of db	Power Ratio		Number of db	Power Ratio		Number of db	Power Ratio	
	Gain	Loss		Gain	Loss		Gain	Loss
0.1	1.023	0.977	3.6	2.29	0.437	7.1	5.13	0.195
0.2	1.047	.955	3.7	2.34	.427	7.2	5.25	.191
0.3	1.072	.933	3.8	2.40	.417	7.3	5.37	.186
0.4	1.096	.912	3.9	2.45	.407	7.4	5.50	.182
0.5	1.122	.891	4.0	2.51	.398	7.5	5.62	.178
0.6	1.148	.871	4.1	2.57	.389	7.6	5.75	.174
0.7	1.175	.851	4.2	2.63	.380	7.7	5.89	.170
0.8	1.202	.832	4.3	2.69	.372	7.8	6.03	.166
0.9	1.230	.813	4.4	2.75	.363	7.9	6.17	.162
1.0	1.259	.794	4.5	2.82	.355	8.0	6.31	.158
1.1	1.288	.776	4.6	2.88	.347	8.1	6.45	.155
1.2	1.318	.759	4.7	2.95	.339	8.2	6.61	.151
1.3	1.349	.741	4.8	3.02	.331	8.3	6.76	.148
1.4	1.380	.724	4.9	3.09	.324	8.4	6.92	.144
1.5	1.413	.708	5.0	3.16	.316	8.5	7.08	.141
1.6	1.445	.692	5.1	3.24	.309	8.6	7.24	.138
1.7	1.479	.676	5.2	3.31	.302	8.7	7.41	.135
1.8	1.514	.661	5.3	3.39	.295	8.8	7.59	.132
1.9	1.549	.645	5.4	3.47	.288	8.9	7.76	.129
2.0	1.585	.631	5.5	3.55	.282	9.0	7.94	.126
2.1	1.622	.617	5.6	3.63	.275	9.1	8.13	.123
2.2	1.660	.603	5.7	3.72	.269	9.2	8.32	.120
2.3	1.698	.589	5.8	3.80	.263	9.3	8.51	.118
2.4	1.738	.575	5.9	3.89	.257	9.4	8.71	.115
2.5	1.778	.562	6.0	3.98	.251	9.5	8.91	.112
2.6	1.820	.550	6.1	4.07	.245	9.6	9.12	.110
2.7	1.862	.537	6.2	4.17	.240	9.7	9.33	.107
2.8	1.906	.525	6.3	4.27	.234	9.8	9.55	.105
2.9	1.950	.513	6.4	4.37	.229	9.9	9.77	.102
3.0	1.995	.501	6.5	4.47	.224	10.0	10.00	.100
3.1	2.04	.490	6.6	4.57	.219	20.0	100	.01
3.2	2.09	.479	6.7	4.68	.214	30.0	1,000	.001
3.3	2.14	.468	6.8	4.79	.209	40.0	10,000	.0001
3.4	2.19	.457	6.9	4.90	.204	50.0	100,000	.00001
3.5	2.24	.447	7.0	5.01	.200	60.0	1,000,000	.00000

raising the voltage this amount is delivering 100 times as much power as it is taking in and is therefore giving a power amplification of 20 db.

For power ratios not directly readable from the table the values there given are used just as are ordinary logarithms. If one power is 162 times another the decibel ratio is not directly readable. However, 162 is equal to 100×1.62; now the decibel corresponding to a power amplication of 100 is 20, and the decibel corresponding to a power ratio of 1.62 is 2.1, so the decibel number corresponding to a power amplification of 162 is equal to $20+2.1$ or 22.1 db. Again, if a power of 1 watt is attenuated to 2 milliwatts the power ratio is 0.01×0.2, and each of these can be taken from the table. In this case it is $20+7=27$ db down, or -27 db.

Reference Power for Amplification Table.—The decibel table given above rates one power in terms of another but of course tells nothing about the amount of power involved. "Zero level" has been taken in communication measurements as 0.006 watt, 0.010 watt and 0.012 watt (even as late as 1930), but is now taken as 0.001 watt.[1] Thus if a telephone channel is 10 db up its power is 0.01 watt; if it is 15 db down it is 0.0000316 watt.

In rating microphones, zero level is taken as the output of a microphone which delivers 1 volt (open circuit) when the sound wave has a pressure of 1 dyne per square centimeter (1 bar); ratings of microphones given in this way are shown in Figs. 15, 20, and 21 of Chapter VIII. In Fig. 15, e.g., at 5000 cycles the microphone is rated as -48 db, and in the same curve this is given as 0.004 volt. Now 4 millivots, on a certain resistance, will develop 0.000016 as much power as will 1 volt. This power ratio is then $10^{-4} \times 0.16$. The first factor corresponds to 40 db and the second to about 8 db, so that 0.004 volt is $40+8=48$ db down, and this is what the decibel scale of the curve sheet reads.

One antenna is said to be 17 db above another; the 10 db gives a power ratio of 10 and the 7 db gives a power ratio of 5, so the power ratio is $10 \times 5 = 50$. The power picked up by a receiving antenna varies with the square of the field strength. The value of $\sqrt{50}$ is 7.07, hence we find that the 17-db antenna gives a field strength 7.07 times as much as the other.

Amplifiers in General.—An amplifier is, as the name implies, an apparatus for increasing the strength of incoming signals. It performs, in modern radio communication, and also in ordinary wire communication, a very important function, in so far as it makes possible the detection of very feeble signals and thus increases the practical range of transmission.

The reader is already familiar with the fact that the signals received in radio transmission consist of very high-frequency currents and voltages, which may be of constant or varying amplitude, depending upon the system used. These signals are generally "heard" in telephone receivers by first reducing the frequency of the incoming currents and voltages from a very high value to an "audible value," and thereafter causing the "audio frequency" currents to flow through the receivers. In using an amplifier either of the following two schemes may be resorted to:

(*a*) The amplifier may be so connected that the incoming high-frequency currents and voltages are first strengthened and thereafter reduced in frequency.

(*b*) The amplifier may be so connected as to strengthen the currents and voltages *after* they have been reduced in frequency.

[1] Everitt, "Communication Engineering," McGraw-Hill Co., 1932.

(c) In another scheme the high-frequency currents from the antenna are amplified to some extent and then reduced in frequency about 10 to 1, and are again amplified. This frequency is still far above audibility and is called an intermediate frequency. After one or more steps of amplification at this intermediate frequency the signal is reduced in frequency to audibility and again goes through one or more steps of amplification.

The above forms the basis of the division of amplifiers into three general classes, i.e., "high-frequency," "intermediate frequency," and "low-frequency."

While these three general types of amplifiers are fundamentally the same, yet the constants of the apparatus used in their construction are often so different that they cannot, in general, be used interchangeably.

The amplifiers used in radio communication consist invariably of one or more multi-electrode vacuum-tubes with other suitable apparatus. As a matter of fact, it was not until the advent of the vacuum-tube that suitable amplifiers could be constructed and operated. The characteristics of a good amplifier should be such that the signal currents are strengthened without any distortion; the vacuum-tube can be made to fulfill these conditions admirably, and it is practically the only apparatus which can. It will be noted from this brief outline that an amplifier must be akin to a trigger which, actuated by the very weak voltages impressed by the antenna, releases from a local energy supply an amount of energy much greater than that actuating the antenna. The suitability of the three-electrode tube for this purpose is at once evident from the analysis of its action given in Chapter VI.

The General Characteristics of Triodes.—These have been quite thoroughly discussed in Chapter VI; on pp. 703 et seq. the possibility of using a tube as an amplifier was pointed out and an elementary analysis given.

The "static" relation between the plate current and grid and plate potentials was shown to be expressible by

$$I_b = A(E_b + \mu E_c)^x, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

and it was also pointed out that, for small variations in the tube potentials, the exponent might be treated as unity.

It was further shown that, if a sufficiently small sine wave e.m.f. was impressed on the grid, the pulsations in the plate current would be sinusoidal in form, and the constant $(1/A)$ acquires the significance of "alternating current plate circuit resistance.".

We then have the equation which was used throughout Chapter VI in analyzing tube action, i.e.:

$$I_p = \frac{1}{R_p}(E_p + \mu E_g), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where I_p = effective value of alternating component of plate current;
R_p = a.c. plate circuit resistance;
E_g = effective voltage of alternating component of grid voltage;
E_p = effective value of alternating component of plate voltage.

We must point out again the limitations of the applications of this relation. The steady values (c.c. components) of grid and plate potentials must be so chosen that, for the value of E_g impressed the linear relation of Eq. (4) holds good. This requires in general that E_g and E_b of Eq. (2) above be properly related.

As pointed out on p. 705 et seq. and illustrated by the curves of Fig. 263, p. 710, when there is considerable outside impedance in the plate circuit the plate current changes linearly with respect to E_g over much wider ranges than might be judged from the static characteristic. This is conventionally illustrated in Fig. 1. With no external resistance in the plate circuit the static characteristic of a tube might be as shown by curve A, whereas if a resistance is put in series with the plate (about equal to R_p) and the plate voltage E_b be increased sufficiently to make I_b (for $E_g = 0$) the same as for curve A, then curve B will be obtained, which is evidently of such a shape as to satisfy Eq. (4) over a change in E_g from perhaps -4 volts to zero.

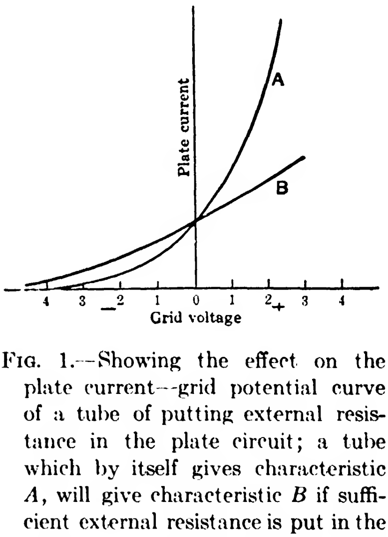


FIG. 1.—Showing the effect on the plate current—grid potential curve of a tube of putting external resistance in the plate circuit; a tube which by itself gives characteristic A, will give characteristic B if sufficient external resistance is put in the plate circuit and the voltage in this circuit suitably increased.

Hence if a constant biasing potential of -2 volts is applied to the grid of a tube having the characteristics of curve B, Fig. 1, would operate satisfactorily with an impressed alternating grid signal of 2 volts maximum value.

It will be noticed that, even with the grid potential positive, curve B is still nearly straight so that it might seem possible to operate the tube satisfactorily with signals sufficiently intense to make the grid swing positive. Such is not generally the case, however; if the grid is allowed to become positive it takes current (it takes negligible current as long as it is negative), and, as will be explained later, this may seriously interfere with proper amplification.

In order to keep the grid of an amplifier suitably negative either a small dry battery may be inserted in the grid-leak resistance circuit, or the leak resistance may be attached to a point in the filament circuit which is sufficiently negative with respect to the filament. These two schemes are indicated in Fig. 2, *a* and *b*; in scheme *b* an extra resistance R is put in series with the filament having such a resistance that when the normal filament current flows through it the IR drop is the required amount. In some multi-stage amplifiers (several tubes repeating one into the other) the filaments are all connected in series, and in this case, the filament of the preceding tube may serve as the resistance R, as indicated in Fig. 3.

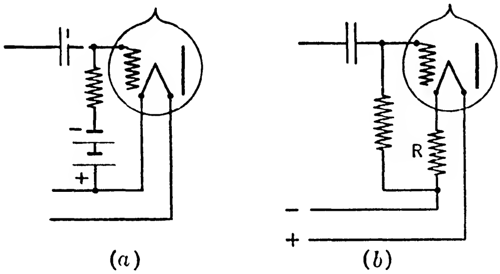


FIG. 2.—To keep the grid of an amplifier tube negative either a small battery of dry cells may be used (*a*) or a resistance inserted in the negative leg of the filament may be employed (*b*).

In the operation of tubes as amplifiers the following quantities play a very important part:

(*a*) A.C. resistance of plate to filament, or output circuit of tube.

(*b*) A.C. resistance of grid to filament, or input circuit of tube.

(*c*) Capacity of grid to filament under static conditions and under actual operating conditions.

(*d*) Capacity between grid and plate.

(*e*) In a few circuit arrangements capacity from plate to ground.

All of the above quantities have been fully discussed in Chapter VI, and the reader will do well, before proceeding with the study of this chapter, to go over the fundamental principles of three-electrode tubes as outlined in the beginning of Chapter VI. The fact should here be emphasized that in the case of a triode the capacity of the grid to filament, while small under static conditions, may attain comparatively large values under actual operating conditions. Again the circuit from plate to filament or grid to filament is made up of a resistance in multiple with a capacity, and, while ordinarily

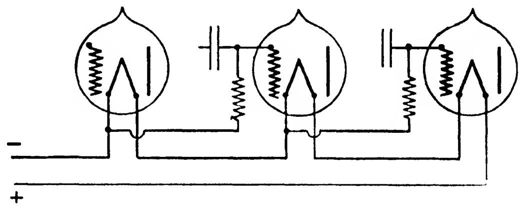


FIG. 3.—In case several tubes are used in cascade it is possible to connect all filaments in series and connect the leak resistances behind the filament of the preceding tube. This makes the grid of each tube negative with respect to its filament by an amount equal to the IR drop of the filaments.

the impedance of either circuit is practically equal to its resistance, there are cases when the frequency is high enough to make the impedance much less than the resistance. That is, the capacity reactance of the circuit, shunting the resistance, may be low enough to determine the impedance of the path.

Effect of External Resistance in the Plate Circuit.—As pointed out in Chapter VI the function of a triode when used as amplifier is to make available in the external plate circuit a voltage similar to that impressed on the grid, and as much larger as feasible. The amount of increase depends upon the μ of the tube used, and on the impedance introduced in the plate circuit.

If a resistance R is put in the external plate circuit the total impedance of the plate circuit is R_p+R, on the assumption that the effect of the tube capacity is negligible. The magnitude of alternating current set up in the plate circuit by a sine voltage E_g acting between grid and filament is given by

$$I_p=\frac{\mu E_g}{R_p+R}, \quad \ldots \ldots \ldots \ldots \quad (5)$$

and this alternating current flowing through the resistance R gives an available voltage in the plate circuit of

$$I_pR=E_g\mu\frac{R}{R_p+R}. \quad \ldots \ldots \ldots \ldots \quad (6)$$

This is indicated in Fig. 4, and experimental curves showing how the amplifying power of a tube varies with the value of R used are given in Fig. 260 of Chapter VI.

It is evident that if resistance is used in the plate circuit, more voltage must be supplied by the B battery to maintain the *plate voltage* at its proper value. Unless this is done the expected amplification $\mu\, R/(R_p+R)$ does not increase with R as rapidly as might be expected because as R is increased the plate voltage (which is equal to E_b-I_bR) decreases and, as pointed out on p. 518, *this gives an increase in* R_p. Hence if resistance is used in the plate circuit of an amplifying tube the B battery e.m.f. must be considerably greater than rated plate voltage of the tube, ordinarily two or three times as much. If the external resistance R is taken equal to the tube resistance R_p the B battery must have a voltage about

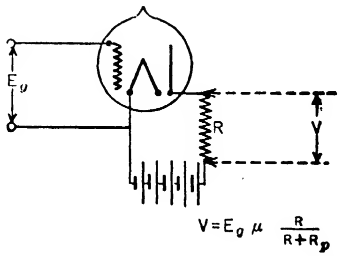


FIG. 4.—Amount of amplified voltage with resistance in plate circuit.

50 per cent greater than the rated plate voltage of the tube, and the amount of voltage amplification obtainable is $\mu/2$.

The effect on the amplifying power of a tube of having the grid at different potentials, E_c, is well brought out by the curves of Fig. 262, p. 707. It is there seen that not only must a proper plate resistance be used, but also the grid must be at a proper average potential if the maximum possible amplification is to be obtained.

Effect of Reactance in the Plate Circuit.—If we use in the plate circuit, a low-resistance reactance, instead of a resistance, the amplifying qualities of the tube are much better. Thus if we put in series with the plate, $Z=R+j\omega L$, we shall have

$$I_p=\frac{\mu E_g}{\sqrt{(R_p+R)^2+\overline{\omega L}^2}},$$

and hence the available drop in the external circuit is

$$V=E_g\mu\frac{Z}{\sqrt{(R_p+R)^2+\overline{\omega L}^2}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

It is to be pointed out here that R and L are the a.c. constants of coil Z, measured under the conditions which obtain in the actual use of the coil; i.e., R and L must be measured in an a.c. bridge (or similar scheme) with the frequency and magnitude of voltage to which the coil is subjected when used in the tube circuit. Also when these measurements are made there must be flowing through the coil a continuous current equal to the average plate current, I_b. These precautions in determining Z are not necessary if an air-core coil is used, but this is seldom the case; generally an iron-core coil is used.

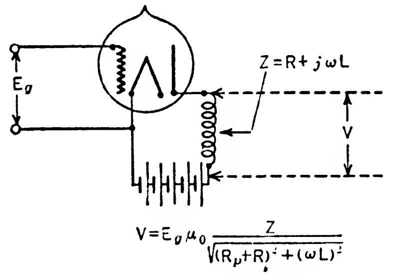


FIG. 5.—Amount of amplified voltage with inductance in plate circuit.

If the resistance R is small compared to R_p and ωL we can write the voltage across the external circuit as

$$V=I_p\omega L=E_g\mu\frac{\omega L}{\sqrt{R_p^2+\overline{\omega L}^2}}. \quad . \quad . \quad . \quad . \quad . \quad (8)$$

The voltage amplification

$$\frac{V}{E_g}=\mu\frac{\omega L}{\sqrt{R_p^2+\overline{\omega L}^2}},$$

may be made nearly equal to μ, by making ωL sufficiently large; at the same time the B battery need have a voltage only slightly greater than the actual voltage of the tube since the resistance in series with the plate circuit is assumed small.

Types of Amplifiers.—An amplifier generally consists of two or more vacuum tubes so arranged that the varying signal voltage is impressed upon the grid of the first tube, thus producing a variation of the plate current in this tube; this varying plate current is then made to produce a varying voltage between the grid and filament of the second tube, and, similarily, the varying voltage is relayed from the second to the third tube, etc., until the plate circuit of the last tube is reached, wherein are placed the telephone receivers or any other device used for making the signals readable. From this brief description it is plain that the signals must be " repeated " from one tube into the next. Amplifiers, either for low-frequency or for high-frequency, are divided into the following classes, according to the arrangement used for " repeating."

(1) Transformer-repeating amplifiers.
(2) Resistance-repeating amplifiers.
(3) Inductance-repeating amplifiers.

A tube, together with all co-acting apparatus used for amplifying purposes, is known as a " stage of amplification "; an amplifier consisting of n such tubes is known as an n-stage amplifier.

The two terminals of the amplifier upon which the incoming signal voltages are impressed are known as the " input " terminals, while the two terminals across which exist the amplified signal voltages are known as the " output " terminals.

Classification of Amplifiers.—Amplifiers may be of various types—radio frequency, intermediate frequency, or audio frequency—but all of them function as a result of the change in plate current brought about by change in grid potential. According to the range of change in plate current taking place during the normal operation of the amplifier, these have been classified as types A, B, or C. The Standardization Rules of the Institute of Radio Engineers define these three types as follows.

Class A Amplifier.—A class A amplifier is one which operates in such a manner that the plate output wave form is essentially the same as that of the exciting grid voltage.

This is accomplished by operating with a negative grid bias such that some plate current flows at all times, and by applying such an alternating voltage to the grid that the dynamic operating characteristics are essentially linear. The grid must usually not go positive on excitation peaks, and the plate current must not fall low enough at its minimum to cause distortion due to curvature of the characteristic. The amount of second harmonic present in the output wave which was not in the input wave is generally taken as a measure of distortion, the usual limit being taken as 5 per cent.

The characteristics of a class A amplifier are low efficiency and output, with a large ratio of power amplification.

Class B Amplifier.—A class B amplifier is one which operates in such a manner that the power output is proportional to the square of the grid excitation voltage. This is accomplished by operating with a negative grid bias such that the plate current is reduced to a relatively low value with no grid excitation voltage, and by applying such excitation that pulses of plate current are produced on the positive half cycle of the grid voltage variations. The grid may usually go positive on excitation peaks, the harmonics being removed from the output by suitable means.

The characteristics of a class B amplifier are medium efficiency and output with a relatively low ratio of power amplification.

Class C Amplifier.—A class C amplifier is one which operates in such a manner that the output varies as the square of the plate voltage within limits.

This is accomplished by operating with a negative grid bias more than sufficient to reduce the plate current to zero with no excitation. An alternating grid excitation voltage is applied such that large amplitudes of plate current are passed during a fraction of the positive half cycle of the grid excitation voltage variation. The grid voltage usually swings sufficiently positive to allow saturation plate current to flow through the tube. Thus the plate output waves are not free from harmonics, and suitable means are usually provided to remove harmonics from the output.

The characteristics of a class C amplifier are high plate-circuit efficiency and output with a relatively low ratio of power amplification.

Uses of the Three Types of Amplifiers.—The ordinary audio-frequency amplifier of a radio receiver, having single triodes in each stage, is a class A amplifier; its useful output is only a small part of the safe plate dissipation, as can be seen from the table of triode ratings on p. 708.

Type B amplifier is used in the special push-pull audio amplifier with type 46 triodes, as described on p. 1007. Also triodes used as grid circuit modulators are operated as class B amplifiers; also the rest of the r.f. amplifiers, after the modulator, is made up of class B amplifiers.

Type C amplifiers are used as plate-circuit modulators.

Transformer-repeating Amplifiers.—These are used for amplifying both radio-frequency and audio-frequency signals; we will first discuss the operation of an audio-frequency transformer-repeating amplifier. The radio-frequency transformer practically always uses a tuning condenser across its secondary coil and thus requires a somewhat different kind of analysis; it will be taken up in a later section.

Referring to Fig. 6, the audio-frequency varying voltage is connected at D and stepped up by means of the transformer T, after which it is applied between the grid and filament of the first tube; this produces a corresponding variation of the plate current of Tube 1. The varying current flowing through the primary P_1 of the transformer T_1 induces an e.m.f. in the secondary S_1. This e.m.f. is applied to the grid and filament of the second tube, and thus the varying signal voltage is "repeated" from the first into the second tube and finally from the second into the third tube, the varying plate current of which is caused to affect the telephone receivers.

It will be at once apparent that in an arrangement of this kind, while each tube itself is always amplifying, the advantage of this may be lost by a poor repeating device. The object to be gained is, of course, to make the varying voltage between the grid and filament of each tube greater than for the preceding tube. This requires correct proportioning of the

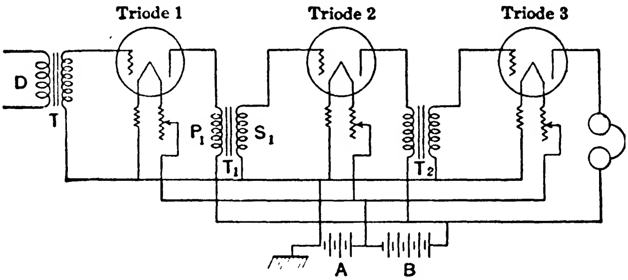


FIG. 6.—A normal, three-stage, transformer-repeating amplifier for audio frequency currents.

primary and secondary of the repeating transformers T_1, T_2; otherwise the grid-filament voltage of the second tube may be but slightly larger, or even smaller, than for the first tube.

We will study the repeating action from the first into the second tube. The capacity between the plate and filament, as well as that between grid and filament, will be neglected at first because its effect is nearly negligible, except at the very highest voice frequencies. For the sake of simplicity we may assume that the repeating transformer has neither leakage inductance nor resistance and also that the magnetizing current is zero; this is equivalent to saying that the transformer is ideal. In so far as the a.c. relations of the circuit are concerned, such a transformer may, if the secondary is loaded by means of non-inductive resistance, be replaced by a fictitious resistance placed in the primary and equal to the secondary circuit resistance divided by the square of the ratio of transformation.

Let E_{g_1} = effective value of alternating voltage between grid and filament of Tube 1;

E_{g_2} = effective value of alternating voltage between grid and filament of Tube 2;
R_{p_1} = plate-filament a.c. resistance of Tube 1;
R_{g_2} = grid-filament a.c. resistance of Tube 2;
μ = amplifying constant of Tube 1;
V = effective value of alternating voltage across primary of repeating transformer, T_1;
n = repeating transformer ratio expressed as the ratio of secondary to primary voltage.

The above quantities are illustrated in Fig. 7. The action of E_{g1} upon the plate current of Tube 1 is the same as if an alternating voltage equal to μE_{g_1} were acting in the plate circuit, in addition to the battery e.m.f. This alternating voltage μE_{g1} is impressed upon a circuit which may be simplified as shown in Fig. 8 and consisting of the plate resistance of the first tube in series with the equivalent resistance of the repeating transformer transferred to the primary. This is probably the simplest way to treat the problem when the coupling between the primary and secondary of the transformer is tight and the load circuit of the transformer is resistive only. For the more general case, i.e., leaky transformer and reactive secondary load, the action of the tube is best analyzed by using for the external impedance in the plate circuit the resistance and reactance of the primary of the transformer as calculated from the general equations given on pp. 118–119.

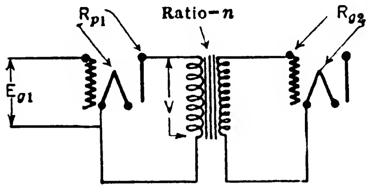


FIG. 7.—Circuit detail of the amplifier shown in Fig. 6.

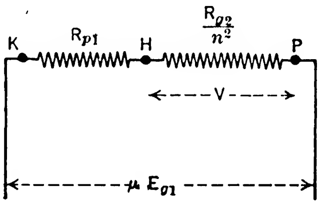


FIG. 8.—Under ideal conditions (transformer requiring no magnetizing current, having zero internal impedance, and secondary load resistive only) the circuit of Fig. 7 may be replaced by the one of Fig. 8.

From Fig. 8 the following equation is easily derived:

$$V=\frac{\frac{R_{g_2}}{n^2}}{R_{p_1}+\frac{R_{g_2}}{n^2}}\cdot\mu E_{g_1}=\frac{\mu R_{g_2}}{n^2R_{p_1}+R_{g_2}}E_{g_1}. \quad . \quad . \quad . \quad (9)$$

The voltage between grid and filament of the second tube is equal to the voltage across the transformer primary multiplied by the ratio of transformation; thus:

$$E_{g_2}=nV=\frac{\mu n R_{g_2}}{n^2R_{p_1}+R_{g_2}}E_{g_1}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

and

$$\frac{E_{g_2}}{E_{g_1}}=\frac{\mu n R_{g_2}}{n^2R_{p_1}+R_{g_2}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Eq. (11) may be written as:

$$\frac{E_{g_2}}{E_{g_1}}=\frac{\mu n a}{n^2+a}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

where

$$a=\frac{R_{g_2}}{R_{p_1}}.$$

It will be noted from Eq. (12) that the ratio E_{g_2}/E_{g_1} varies directly with the amplifying constant of the first tube and it also varies in a complex manner with R_{g_2}/R_{p_1} and with n. It will further be noted that:

First. If μ and n are kept constant and the ratio R_{g_2}/R_{p_1} increased from a low value, then E_{g_2}/E_{g_1} will constantly increase towards the limiting value μn which will be theoretically reached when $R_{g_2}/R_{p_1}=\infty$.

Second. If μ and R_{g_2}/R_{p_1} are kept constant and the value of n changed then E_{g_2}/E_{g_1} may be shown to have a maximum when

$$n^2=\frac{R_{g_2}}{R_{p_1}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

It follows that the resistance R_{g_2} should be made as high as possible, and that, once this has been done, a transformer should be chosen with a transformation ratio about equal to $\sqrt{R_{g_2}/R_{p_1}}$. It is not always possible adequately to satisfy this latter condition, as will be more fully explained later. The resistance R_{g_2} is made high by preventing the potential of the grid of Tube 2 from ever becoming positive, for, in this case, the grid-filament resistance is theoretically infinite; this is accomplished by keeping the grid at a negative potential by a suitably connected battery, or by any of the circuit arrangements already explained on p. 980. Practically, on account of gas in the tube and the leakage from grid to filament outside of the tube, the grid-filament resistance, while very large, is at the most of the order of one million ohms, and may in many cases be as low as one hundred thousand ohms.

For the ideal value of n^2 the ratio E_{g_2}/E_{g_1} is found by substituting (13) in (12), thus:

$$\frac{E_{g_2}}{E_{g_1}}=\mu\frac{n}{2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

If the tubes used for the various stages of amplification are similar, the transformers may have the same ratio throughout.

The results indicated by Eqs. (13) and (14) have been obtained on the basis of ideal transformers having neither leakage inductance nor coil resistance and requiring no magnetizing current. The effect of all of these in an actual transformer would be such as to alter the best value of the transformation ratio, and, more than this, to diminish the ideal ratio E_{g_2}/E_{g_1} as given by Eq. (14). The leakage inductance and coil resistance of the transformer can be made quite small and negligible as compared with the resistance R_{g_2} and their effect will, therefore, be but small. On the other hand, it is very important to make the magnetizing current very small, or, in other words, to make the no-load reactance of the transformer primary very high. This will be made clearer by a study of the diagram Fig. 9, which is similar to Fig. 8, with the exception of the introduction of X_o in multiple with R_{g_2}/n^2, where X_o = reactance of transformer primary at no load.

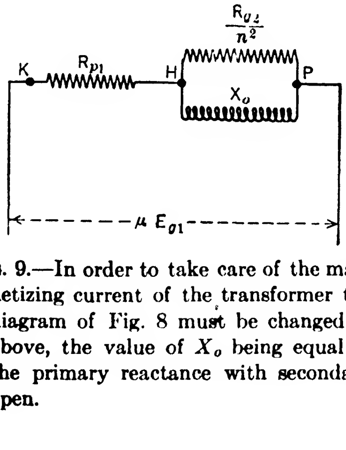


FIG. 9.—In order to take care of the magnetizing current of the transformer the diagram of Fig. 8 must be changed as above, the value of X_o being equal to the primary reactance with secondary open.

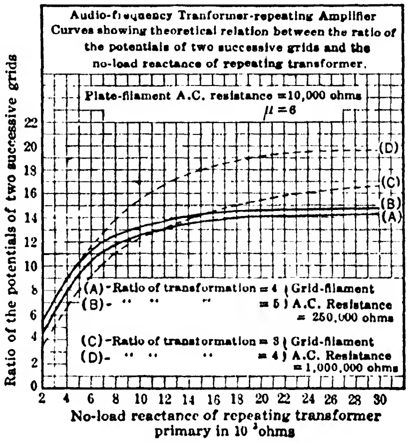


FIG. 10.—Calculated values of voltage amplification using transformers of different ratios and two different values of input circuit resistance of the second tube. The curves show the effect of varying the no-load reactance of the primary of the transformer, abscissas being no-load reactance in thousand ohms.

A resistance should, in the above diagram, be inserted in series or parallel with X_o to represent the core losses, but we have omitted it for the sake of simplicity.

The diagram shows that X_o is in multiple with R_{g_2}/n^2 and therefore diminishes the equivalent impedance of the circuit H–P; if, then, X_o were very low the voltage drop across H–P and, therefore the secondary voltage (E_{g_2}) would be small. It is important, then, to make X_o as high as possible, or, in other words the primary must have a very large number of turns. There is a point, however, beyond which it is uneconomical to increase the value of X_o, since the gain in amplification is too small to make it worth while. To show this we have worked out the theoretical curves of Fig. 10, after having assumed the following:

$$\mu = 6,$$
$$R_p = 10{,}000.$$

For R_{g_2} two different values were chosen, i.e.:

$$R_{g_2} = 250{,}000 \text{ ohms and } R_{g_2} = 1{,}000{,}000 \text{ ohms.}$$

For R_{g_2} of 250,000 ohms the best transformer ratio for an ideal transformer is found, from Formula (13), to be: $\sqrt{250{,}000/10{,}000} = 5$, and similarly for R_{g_2} of 1,000,000 the best transformer ratio would be $\sqrt{1{,}000{,}000/10{,}000} = 10$.

From Formula (14) we have:

Maximum possible value of E_{g_2}/E_{g_1} for R_{g_2} of 250,000 $= 6 \times \frac{5}{2} = 15$.

Maximum possible value of E_{g_2}/E_{g_1} for R_{g_2} of 1,000,000 $= 6 \times \frac{10}{2} = 30$.

The points on the curves have been plotted by assuming different values of X_o and then obtaining the voltage across H–P and also the secondary voltage E_{g_2}, after which E_{g_2}/E_{g_1} was computed. The assumption was made that the transformer had no leakage inductance, no coil resistances, and no core losses. Curves were drawn for two different ratios of transformation, i.e., 4 and 5 for R_{g2} of 250,000 and 3 and 4 for R_{g_2} of 1,000,000 ohms. They show:

First.—That for low values of X_o the ratio E_{g_2}/E_{g_1} may be very small, even smaller than the amplifying factor of the tube, which is in this case 6. Thus, a transformer with low no-load reactance might make the result of the two tubes no better, *or even worse,* than for one tube alone.

Second.—That for R_{g_2} of 10^6 ohms the ratio of E_{g_2}/E_{g_1} is larger than for R_{g_2} of 250,000 ohms, for the same transformer ratio, even though the value of n used for R_{g_2} of 250,000 ohms is much nearer the ideal value than for R_{g_2} of 10^6 ohms.

Third.—Beyond certain values of X_o the ratio E_{g_2}/E_{g_1} does not increase much with increase of X_o. Not only is it of no advantage to increase the reactance X_o above a certain amount, but it is actually disadvantageous. The X_o is, of course, increased by increasing the cross-section of the core or the number of turns in the primary winding. The former of these expedients is objectionable because it increases the space requirements. As regards increasing the number of primary turns, it must be noticed that, if this is done, the secondary turns must be proportionately increased if the ratio of transformation, n, is to be constant. Now, the higher the number of transformer turns the higher become the internal resistance and leakage reactance of the transformer, which have so far been neglected in our discussion.

A high internal transformer impedance may produce a large internal drop due to the "load" attached to secondary, i.e., the grid-filament resistance and reactance of the tube into which the transformer is repeating, and also the *internal distributed capacity of the secondary winding itself;* the final result would be that the voltage applied to the grid-filament might be far less than that calculated on the basis of negligible internal transformer drop. This may be summed up by stating that, for a given frequency, the higher the number of turns used the more does the ratio of terminal voltages (secondary to primary) depart from the turn ratio n, being only a fractional part of n. In fact, it is possible to increase the transformer turns to such an extent (more especially if the ratio be high, say: 10 to 1), that the terminal voltage of the secondary (when used in the tube circuit) is *less than the voltage impressed upon the primary* winding. The above phenomenon may take place if the number of turns is kept constant and the frequency raised. In practice the value of X_o, for the lowest frequency it is desired to repeat with the amplifier, is made equal to about once or twice the value of the plate-filament a.c. resistance.

Fourth.—The higher the ratio of transformation the greater the amplification. In connection with this it will be noted, however, that a point may be reached beyond which it is uneconomical to increase the ratio, since the gain in amplification is too small, as for example in the case of curves A and B for transformer ratios of 4 and 5, respectively. As a matter of fact if we consider that, for a constant number of primary turns, the increase in ratio is obtained by increasing the number of secondary turns and that simultaneously the internal impedance of the transformer and the effect of the distributed capacity of the secondary are increased, it will be apparent that, due to the large internal drop, the voltage across the secondary may be smaller for a high than for a low ratio of transformation. This effect has already been pointed out in connection with the value of X_o and plays such an important part in connection with the transformation ratio that it has been found advisable, in practice, to keep the value of this ratio below about 4 or 5.

In the case of the first transformer T (see Fig. 6) it may be shown by a method similar to that used for the other transformers that the ideal ratio of transformation is given by:

$$S=\sqrt{\frac{R_{o_1}}{R}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

where S = ratio of secondary to primary voltage for transformer T;
R = resistance connected in series with the primary of transformer T.

The resistance R may be that of the plate-filament circuit of some other tube or of a telephone line or anything else which may be in series with the transformer primary.

Again, as in the case of the other transformers, the no-load inductance of the primary should be high, and the ratio S should not be made so high as to permit the internal capacity of the transformer to have much effect.

Effect of Resistance across Secondary Terminals.—Obviously if a low resistance is put across the secondary terminals of any transformer the terminal voltage must drop. In the case of the transformers being discussed here, having windings of very fine wire and extremely large inductance, even high values of load resistance may seriously diminish the terminal voltage.

In testing transformers it is customary to use a resistance, instead of a triode circuit itself, in the primary circuit, the value of resistance being chosen equal to the a.c. resistance of the plate circuit of the tube with which the transformer is to be used. This is shown in Fig. 11; the value of R_1 is taken equal to the tube plate circuit; for the average amplifier tube (excluding tetrodes and output triodes) 10,000 ohms is representative.

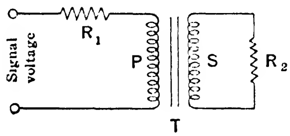


FIG. 11.—In testing the characteristics of audio frequency transformers, it is customary to use in series with the primary, a resistance R, equal to the a.c. resistance of the plate circuit of the tube with which the transformer is to be used.

Now as R_2 is varied the voltage across S will also change; typical experimental results are shown in Fig. 12. This was a representative transformer, having a turn ratio of 4.5. It is seen that a resistance of even one million ohms across the secondary will appreciably cut down the terminal voltage. But most input circuits of triodes have about this resistance, when sufficiently biased. Fig. 75 of Chapter VI shows the input circuit resistance of a typical amplifier tube, and this is seen to vary from about 100,000 ohms up, as the bias is increased from zero.

Effect of Capacity across Secondary.—The input circuit of a tube has capacity, generally measured in a few micro-microfarads; with the capacity of the tube socket, and wiring, in parallel with the grid-cathode circuit the total capacity connected to the secondary terminals might be as as much as 20 $\mu\mu f$. There is an additional capacity in the secondary winding itself and also between the primary and secondary windings so that the actual circuit is somewhat complex.

Simply the circuit of one of these transformers is as shown in Fig. 13. The primary has been assumed to have 50 henries of inductance (which is a reasonable value for a good transformer), and the turn ratio as 1 to 3, so the secondary has 450 henries inductance. The capacity C is taken as 20 $\mu\mu f$.

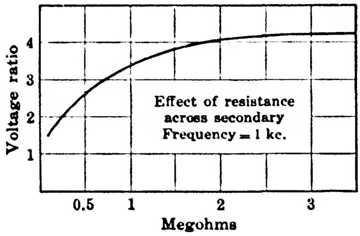


FIG. 12.—The effect of resistance across the secondary of a modern audio frequency transformer, when tested at a frequency of one kilocycle.

At low audio frequencies the amount of charging current taken by this small condenser is so small that no appreciable effect is produced on the transformer action, but at high frequencies this is not so. Fig. 5 of Chapter V shows how a condenser load on the secondary of a transformer may very greatly increase the secondary terminal voltage at some frequencies and then decrease it at higher frequencies. As a matter of fact, the curve of Fig. 5 was obtained at fixed frequency and variable capacity, but from the action there shown we deduce the fact that a certain capacity will show resonance at some frequency, giving high secondary terminal voltage, and at frequencies much higher than this the secondary voltage will fall rapidly.

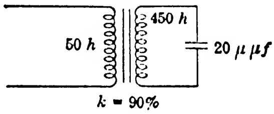


FIG. 13.—A 3 to 1 transformer, having a capacity of 20 $\mu\mu f$ across its secondary. This might be the capacity of the next tube, plus the wiring, and the internal capacity of the secondary winding itself.

The coefficient of magnetic coupling of an ordinary audio-frequency transformer is not very high (as transformers go) probably being seldom in excess of 90 per cent, thus giving 10 per cent so-called leakage reactance. Under this assumption, and under the further assumption that the losses in the transformer can be neglected, we can replace the circuit of Fig. 13 by that of Fig. 14. Here the transformer has been changed in A to a 1 : 1 ratio, permissible if the capacity is changed by the (turn ratio) ^{2}; this makes it 180 $\mu\mu f$.

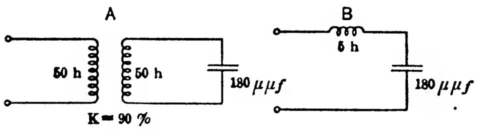


FIG. 14.—Simplifications of the circuit of Fig. 13.

Diagram B of this figure is the equivalent of A, the two 50-henry coils, coupled 90 per cent being replaced by one coil of $50 \times (1-0.90) = 5$ henries. Now the circuit shown in B is resonant at a frequency of $1/2\pi\sqrt{LC} = 10^6/2\pi\sqrt{5 \times 180} = 5300$ cycles. Thus we might well expect that the transformer shown in Fig. 13 might show a resonant "hump" at this frequency. This resonance will not be very sharp because the coils, of fine wire, and the iron core, both result in high effective resistance and so limit the resonant rise in voltage.

Internal Capacity of Transformer Itself.—Most transformers have layer wound coils, with paper between layers. Sometimes one layer of paper is used and sometimes two; in some cases the paper is wax impregnated and in others it is not so treated. Also there is capacity between the primary and secondary coils, because one is wound right on top of the other. The effect of this inter-coil capacity will depend upon the relative connections of primary and secondary. This point is illustrated in Fig. 15. In A the relative placing of the coils on the transformer is shown; p_1 and p_2 represent the inner and outer ends of the primary winding, and s_1 and s_2 show the corresponding terminals of the secondary. In B is shown one way of connecting the transformer; it will be seen that the inner end of the secondary s_1 and inner end of the primary p_1 are grounded in so far as alternating voltages are concerned. Thus the capacity between the p_2 end of the primary and ground is relatively high, the inner layer of the secondary winding being at ground potential.

It might be that the secondary winding is connected in the opposite direction to that shown in Fig. 15; then the p_2 and s_1 terminals would connect to the plate and grid respectively, and the effective capacity between windings will be much different. As a matter of fact, this simple changing of the polarity of the secondary winding connections has a very marked effect on the amplifying characteristic of the transformer.

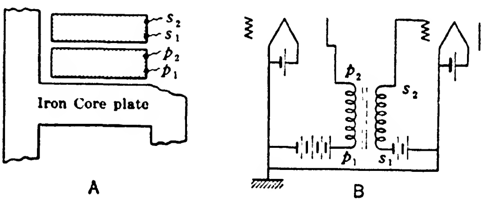


Fig. 15.—Arrangement of windings of a transformer, and connection in an amplifier.

In Fig. 16 are shown the amplification characteristics of a well-designed transformer for various circuit conditions; for comparison is shown the amplification curve of a transformer of poorer design (less iron and less copper). It will be noticed that with normal connection (grid connected to outer end of secondary coil) there is a high resonance peak in amplification at 5500 cycles. This transformer had a high value of primary self-induction, about 100 henries, and a secondary inductance of about 2500 henries. The vacuum-tube voltmeter used to measure the secondary voltage had about 8 $\mu\mu f$, and the internal capacity of the transformer was about 7 $\mu\mu f$, giving 15 $\mu\mu f$ together. The coefficient of coupling was about 90 per cent, so the resonance frequency calculates to be 5500 cycles. The curve B shows this to be the actual resonance frequency.

When the secondary winding was reversed the effective internal capacity of the transformer was made greater, and the internal losses were evidently much increased; the resonance peak is now much less marked and occurs at lower frequency, 2500 cycles. The effect of putting 1 megohm across the secondary terminals is shown by curve D and of putting 0.5 megohm across by curve E. As seen by the curve of Fig. 75 of Chapter VI this change in resistance of the input circuit of a triode can be brought about by changing the grid bias from 4.5 volts to 3 volts.

The transformer used in getting the results of Fig. 16 had two layers of unimpregnated paper between layers of secondary winding. Using just the same winding, with only one layer of paper between layers, and having

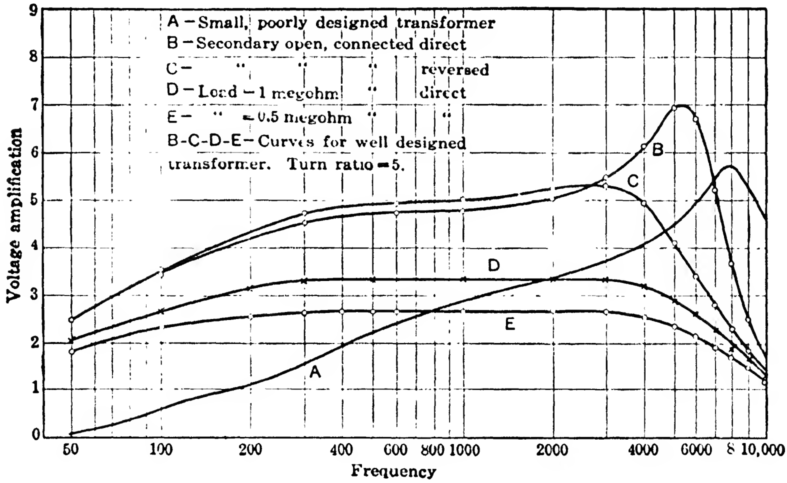


FIG. 16.—Performance of a small cheap transformer, A, and that of a heavier, bigger one of better design. For different connections (see Fig. 15) and different resistances across the secondary the performance changes as shown here.

the paper impregnated with wax, changed the characteristics as shown by curves B and C of Fig. 17. Curve A is reproduced from Fig. 16 for comparison. It can be seen by comparison of curves A and B that the effect of using thinner insulation, wax-impregnated, was to produce resonance at a much lower frequency, 3500 cycles compared to 5500 cycles. This change in frequency means that the capacity in the secondary circuit had increased about 2.5 times.

Sectionalized Windings.—It is possible to wind the two coils in sections, on forms, so that the primary and secondary turns are more interspersed. This will reduce the magnetic leakage between windings, and may also decrease the internal capacity. Better iron may be used, thus raising the self-induction of the primary for a given winding. Thus the amplification may be made nearly flat from about 50 cycles to perhaps 8000 cycles.

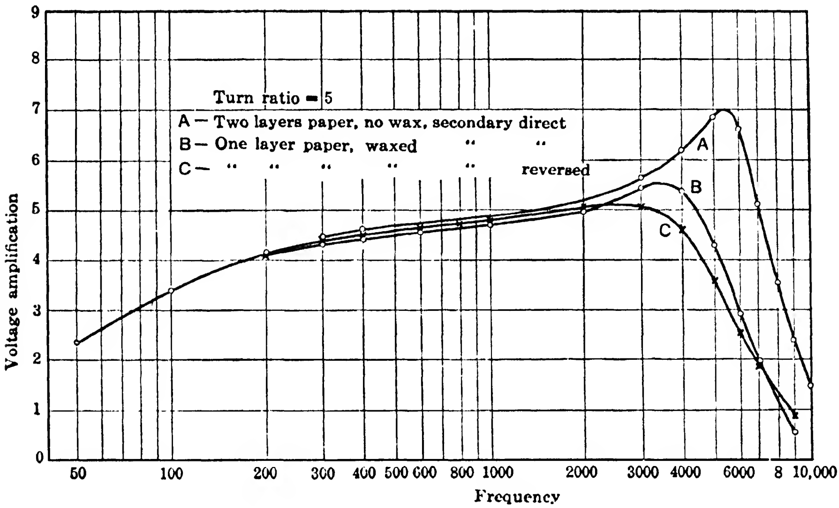


FIG. 17.—These curves show the effect of the internal capacity of the transformer, with one and two paper insulation.

One scheme for making these low-capacity, low-magnetic leakage windings, is shown in Figs. 18 and 19. Fig. 18 shows a cross-section through one side of the winding, and Fig. 19 shows one of the spools (molded from some insulation material) on which the turns are wound. Three spools fit snugly

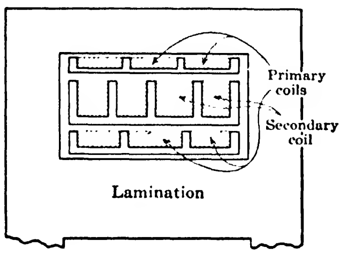


FIG. 18.—To keep the internal capacity low, some transformers use the idea incorporated in all power transformers, namely winding primary and secondary in sections, and properly interspersing them.

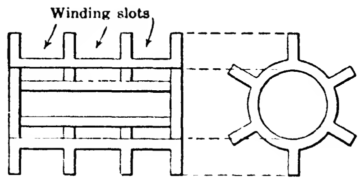


FIG. 19.—The coils of Fig. 18 are wound upon molded forms like these, which fit inside one another.

into one another, and their combined depth is just equal to the width of the " window " in the core lamination. By winding the primary in two parts one inside and one outside the secondary, the leakage reactance is made low; this, in combination with the low internal capacity of the windings, advances the resonance hump well into the higher-frequency range.

Effect of " Free " Secondary Winding.—It will be noticed in Fig. 15

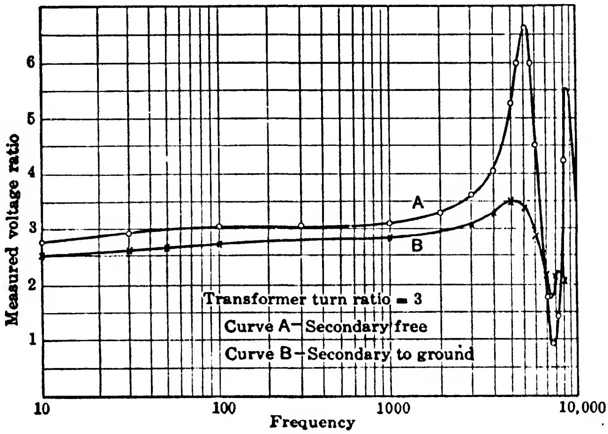


FIG. 20.—If, in testing, the secondary winding is entirely free from the primary, it is quite likely that two resonance humps may be found as was the case here.

that one side of the secondary winding is connected to one side of the primary and both are connected to ground. The B battery between the lower end of the primary and ground has a negligible impedance for alternating currents, and therefore, so far as alternating current is concerned, the lower end of the primary winding is at ground potential. With this normal connection the amplification curve shows forms as given in Fig. 16.

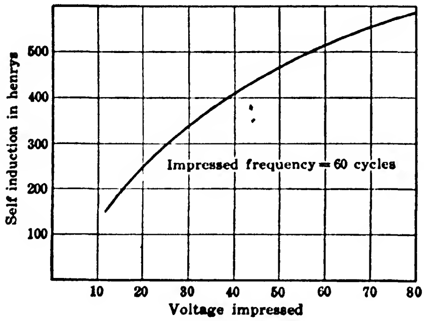


FIG. 21.—Variation of self-induction of the primary of a good, modern transformer, as the test voltage was varied.

Now if the secondary winding is left free of the ground, the voltage across its terminals being measured by a vacuum-tube voltmeter which also is not grounded, peculiar results may be obtained. The internal capacity of the two windings, as well as the capacity between windings, act to give a double resonance hump, as shown in Fig. 20. There is also presented the curve for normal connection of the secondary winding; it also shows a small second resonance peak.

In obtaining these curves the action of the transformer only was being studied; there was no resistance in series with the primary winding to represent the plate-circuit resistance of a triode, so that the signal voltage was impressed directly in the primary winding.

It can be seen that the transformer itself will act efficiently even as low as 10 cycles; of course if it were used in the plate circuit of a triode it would not give the flat curve at the lower frequencies because the primary reactance would be low compared to the tube resistance, so but a small part of the tube voltage would appear across the transformer primary.

Variation in Self-induction of the Primary Winding.—In Fig. 21 is shown the variation in self-induction of the primary winding (of one of the best obtainable transformers) as the value of the voltage impressed on the primary is varied; the measurements were made at 60 cycles. This curve shows that as the impressed signal is increased in amplitude the self-induction increased very much. Now we have shown, Fig. 10, that at low frequencies the voltage amplification increases as the self-induction of the primary is increased. It therefore follows that this transformer would amplify strong signals, of low frequency, more than it would weak ones. However, there is another factor, the resistance across the secondary terminals (input circuit of the next triode) which also varies with signal strength, and in such a way as to offset this increase in self-induction of the primary.

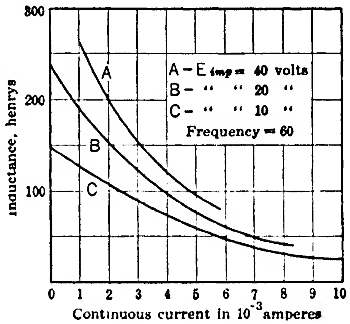


FIG. 22.—Effect of plate circuit current upon the a.c. self-induction of the transformer of Fig. 21.

Effect of Plate Current on Primary Inductance.—As ordinarily used, an audio-frequency transformer has a continuous m.m.f. acting on its core, in addition to that of the alternating current which is produced by the signal voltage. This continuous m.m.f. is due to the average value of the plate current of the triode.

In Fig. 22 are given curves showing how the plate current affects the self-induction of the primary winding. These results on actual transformers bear out the ideas formulated in Chapter II in which the various factors affecting the permeability of iron were discussed.

Reason for Negative Grid Bias.—Except in special circuits the grids of audio-frequency amplifier tubes should always be biased to such an extent that the strongest signal which will be handled is not sufficient to force the grids to become positive with respect to the cathode. From

Fig. 75, Chapter VI, it is seen that if the grid goes positive the input circuit resistance falls to very low values, and Figs. 12 and 16 show that low input circuit resistance results in low amplification.

A more important reason has to do with the action of the grid current, which flows when the grid is positive. If a sine wave signal is being repeated by the amplifier its form will be greatly changed if the grid of one of the amplifying tubes is allowed to draw current. Inspection of Fig. 15 *B* shows that any current taken by the grid must flow through the resistance and reactance of the secondary winding. This results in taking the "tops" off the positive alternations of the voltage wave, as suggested in Fig. 23. This peculiar distortion results in a disagreeable change of quality in signal reproduction, due to the even harmonics, always present in a wave of this form.

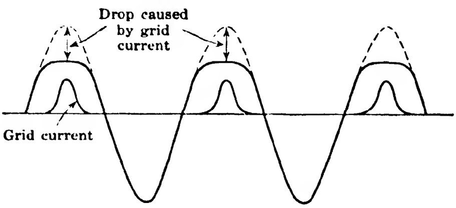


FIG. 23.—If the grid of an A.F. amplifier draws current during the positive alternation, the internal drop in the secondary winding results in a terminal voltage having flat tops on the positive parts of the wave—this gives "second harmonic" distortion.

It then follows that as a signal builds up in amplitude from stage to stage the successive grids should have an increasingly greater negative bias, so that no grid draws an appreciable current at any time.

Construction Details of a Transformer.—The transformers used in getting the curves of Figs. 16 and 17 had a cross-section of iron of 0.52 sq. in., made up of 46 plates each 14 mils thick. The laminations were 0.8 in. wide, for the center leg of the shell type core. The length of magnetic path was 7 in., and there were 6500 turns in the primary coil. A reasonable secondary coil has 18,000 turns. At 60 cycles, with 14 volts impressed, the inductance was 101 henries. The maximum flux density with 14 volts of 60-cycle signal was 1800 lines per square centimeter, and the effective permeability was 1020. With continuous current

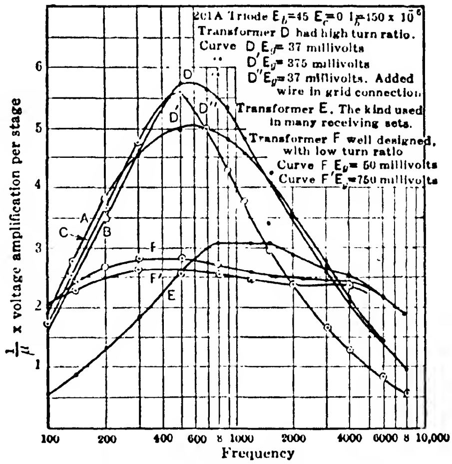


FIG. 24.—Performance of typical transformers.

through the winding the self-inductance and permeability both decrease very much.

Effect of Plate Voltage upon Amplification.—Leaving other factors the same it would be thought that raising the plate voltage would increase the amplification, especially at the lower frequencies, because the plate-circuit resistance is lowered. This will raise the ratio of the transformer reactance to the tube resistance and hence, by Eq. (8), p. 982, the amplification. However frequently the opposite effect takes place, the amplification drops. The increased plate voltage draws a greater plate current, and this, according to Fig. 22, drops the self-induction of the primary of the transformer and hence its reactance; this results in lowered amplification.

Transformers as Built To-day.—In Fig. 24 are shown a few curves typical of the transformers used today in many amplifiers; these are all well-known makes. Transformer *D* is quite evidently built with too high a turn ratio to give satisfactory amplification throughout the voice range, although it would be suitable for the reception of telegraph signals. With the beat-note method of reception, for example, the operator could adjust the note for 600 cycles per second and this transformer would give four times as loud a signal as would transformer *F*, yet *F* is a much better transformer for a radio-telephone amplifier than *D*.

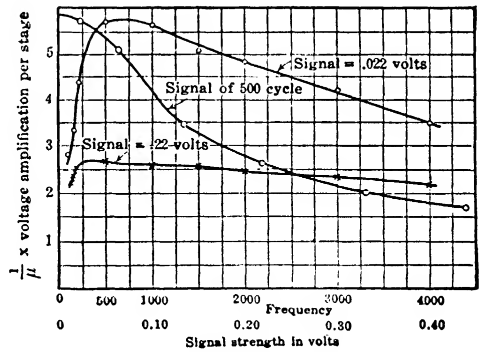


FIG. 25.—Effect of signal strength on amplifying power of a transformer having a turn ratio of six. No grid bias was used, the grid thus taking current on the positive alternations. With fixed frequency the amplification changed with signal strength as shown by the curve so marked.

In general the louder the signal the less does the transformer amplify; this is primarily because of the greater grid current taken by the stronger signal. The grid current, flowing through the secondary winding, lowers the terminal voltage of the secondary. Now the grid current is greatest when the signal voltage is amplified the most and hence it can be seen that this effect tends to flatten the amplification curves of Fig. 24. By using a very loud signal when testing a transformer it is possible to make such a transformer as *D*, which really gives very uneven amplification throughout the frequency range, give a curve which is quite flat. This effect is shown by curves *D* and *D′* of Fig. 24 as well as by curves *F* and *F′*. The smoothing out of the characteristic is more pronounced as the amplification, on weak signals, is the more uneven.

In curve D'' is shown the effect of having too much capacity attached to the terminals of the transformer secondary. The signal strength was the same as for curve D but the wire leading from the transformer secondary to the grid of the next tube was made much longer than necessary, just to bring out this effect. In getting curve D a short wire a few inches long was used and in getting D'' the connecting wire was 4 feet long. Whereas such a condition could never arise in an actual, wired amplifier set, it might occur in a laboratory test, where due care in arranging the wiring had not been observed.

Curve E is quite typical of the amplifier performance in hundreds of thousands of radio receivers. The transformer is under size, has too little iron and too little copper. It is the performance of such transformers which makes the reproduction of radio programs no better than the old-fashioned talking machine reproduction. Only a few octaves of the audible scale are reproduced, resulting in that unnatural quality which gave rise to the term " canned music."

Effect of Signal Intensity on Amplification.—In the foregoing section it was pointed out that the amplifying power of a transformer decreases as signal strength is increased and Fig. 24 shows some curves of the effect. The effect is greater in transformers with large magnetic leakage between the windings and Fig. 25 is given to illustrate this effect. The transformer was one of the commercial types and was used in a 201A tube with 45 volts in the plate circuit and no grid bias. It is seen that with the strong signal the amplification curve is almost flat, yet with weak signal great variations occur. With the strong signal, the grid was going positive part of the time, hence drawing current and greatly distorting the wave form. In Fig. 25 is also shown the curve of amplification for a signal of fixed frequency (500 cycles) and varying amplitude.

Impedance of Telephone Receivers.—The receivers (or loud speaker) used in the plate circuit of the last tube of the amplifier should be suitably chosen. The impedance of a telephone receiver is made up of the following four components:

First. The static reactance.
Second. The static effective resistance.
Third. The motional reactance.
Fourth. The motional resistance.

The first two components are due to the constants of the electric and magnetic circuits of the receiver and the losses taking place therein and are the effective reactance and resistance measured with the diaphragm " locked."

The motional reactance and resistance are produced by the motion of the diaphragm and are to be added to the static reactance and resistance respectively. It is plain that the motional resistance is the resistance equivalent of the power expended in moving the diaphragm to and fro, part of which is useful in producing sound; in other words, for a certain receiver, the greater the motional resistance the greater will be the receiver response to a certain value of incoming alternating current. The value of this resistance varies with the frequency and is a maximum at about 900 to 1000 cycles per second. Curves are given in Fig. 26, showing the relation between frequency and resistance and reactance of a receiver with the diaphragm locked and with the diaphragm vibrating.[1] A telephone receiver may be thought of as a motor receiving electrical energy and transforming part of this into motional energy of the diaphragm; again part of the motional energy of the diaphragm is transformed into acoustical energy. The motional resistance, when multiplied by the square of the current gives a measure of the power supplied to the diaphragm, a small percentage of which is converted into sound.[2]

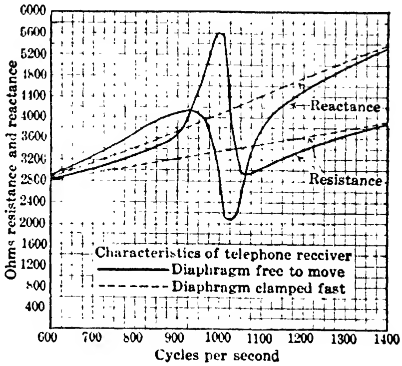


FIG. 26.—Resistance and reactance of such a receiver as is used in radio work as a function of the frequency; one set of curves gives the characteristics with diaphragm free to move and the other with it clamped tight.

The overall efficiency of a receiver is the ratio of the power given off as sound to the power supplied electrically and this varies of course, for various frequencies. It is a very difficult task to make accurate measurements of sound energy, but from what data are available it appears that the average efficiency of a telephone receiver throughout the voice range of frequency is less than 1 per cent. Most of the power supply is used up in the windings, as eddy currents and hysteresis in pole pieces and diaphragm, as frictional losses in the iron of the diaphragm itself, as frictional losses at the rim where it is clamped, and in air eddies in the cavity above the diaphragm.

In view of the many components of the impedance of a receiver, and

[1] It must be pointed out that where the motional resistance is negative, the diaphragm is not acting like a generator, giving off electric power due to its motion; the negative motional resistance merely signifies that the diaphragm in motion (in the right motional phase) absorbs less power as eddy currents and hysteresis than if it is locked and so unable to move.

[2] See "The Mechanics of Telephone Receiver Diaphragms," by Kennelly and Affel, Proc. Am. Acad. of Arts & Sciences, Nov., 1915.

their variation with the frequency, it is difficult to lay down any set rules for the choice of a receiver to be used in connection with an amplifier. The following analysis, however, will be of value:

Let R_s = static resistance of receiver;
R_m = motional resistance of receiver at a definite frequency;
X = total reactance of receiver when in motion, at the same frequency as assumed for R_m;
R_p = plate-filament a.c. resistance of tube;
E_g = effective value of voltage impressed on grid;
I_p = effective value of plate current produced by E_g;
P_m = power expended in R_m.

From p. 982 we have

$$I_p = \frac{\mu E_g}{\sqrt{(R_p+R_s+R_m)^2+X^2}}$$

and

$$P_m = I_p^2 R_m = \frac{\mu^2 E_g^2 R_m}{(R_p+R_s+R_m)^2+X^2}. \quad . \quad . \quad . \quad . \quad . \qquad (16)$$

For maximum response in the receiver, the power (P_m) expended in the motional resistance should be a maximum. Eq. (16) shows that if all the other quantities are kept constant while R_m is varied, P_m will vary directly with R_m while R_m is small, but when R_m becomes sufficiently large as compared with R_p, R_s, and X then P_m will decrease with increase in R_m. R_m may be increased by suitable design of the diaphragm and the air cavity above it, but it is hardly possible to make R_m too large.

Eq. (16) also shows that for a receiver of fixed constants and for a tube of definite μ the power P_m increases as R_p is decreased, therefore the tube resistance should be small as compared with the impedance of the receiver in motion. As it is impractical to reduce the plate-filament resistance very much it is necessary to use a receiver with as high an impedance as practical. Ordinarily the telephone impedance is at least equal to the tube resistance.

We will calculate the value of P_m for a definite case. Assume a receiver having the characteristics of Fig. 26 from which, at 900 cycles.

$$R_s + R_m = 3700$$
$$R_s = 3200$$
$$R_m = 500$$
$$X = 4100$$

Assume this receiver used with a tube of $R_p = 10{,}000$ and $\mu = 6$ and let the signal voltage $= 0.03$ volt

$$P_m = \frac{6^2 \times 0.03^2 \times 500}{(10{,}000+3700)^2+4100^2} = 0.081 \text{ microwatt.}$$

Of this amount of power supplied to the diaphragm possibly 0.001 microwatt would change to sound.

It has been reported [1] that with a good loud speaker an electrical input of 2 milliwatts gives the equivalent of normal speech. Remembering that normal speech represents about 10 microwatts of power we find the efficiency of this loud speaker to be ½ per cent, about the same as was mentioned above for a head receiver.

It might be thought that the impedance of the receivers would change appreciably with signal strength, because of the iron used in the magnetic path. In general, with increasing current the impedance of an iron-core coil diminishes, but it happens that the currents occurring in telephone receivers are so small that the opposite effect occurs, to a slight extent.

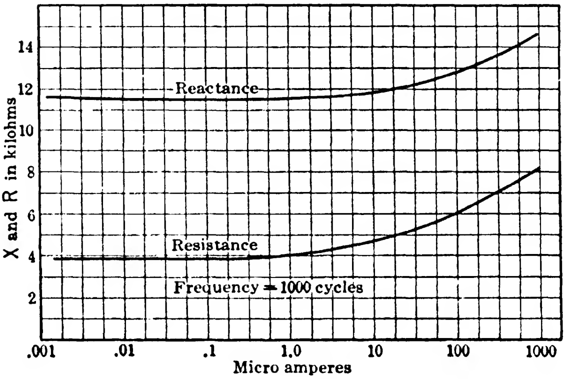


FIG. 27.—The resistance and reactance of a pair of head phones for very weak currents.

Using a well-designed set of head receivers, having 11,000 turns per receiver, Curtis [2] found that for currents from 10^{-9} to 10^{-6} amperes the impedance was practically constant. At higher currents the impedance increased somewhat. One set of his results is given in Fig. 27. Incidentally this figure shows what the power factor of a good receiver should be. This receiver is a much more efficient one than that used in getting the results of Fig. 26, the motional resistance at its resonant frequency being about four times the resistance with diaphragm locked.

When telephone receivers are held tight against the ears the abrupt change in resistance and reactance (shown at 1000 cycles in Fig. 26) does

[1] See "Loud Speakers," by E. K. Sanderman, Wireless World and Radio Review, Jan. 23 and 30, 1924.

[2] "The Vibratory Characteristics and Impedance of Telephone Receivers at Low-Power Inputs," by A. S. Curtis, B.S.T.J., Vol. IV, No. 3, July, 1925.

not occur; the air column in the ear cavity acts like a horn and loads the diaphragm as was shown for the loud speaker on p. 855.

Loud speakers were analyzed somewhat in detail in Chapter VIII, pages 853, et seq. They are in general more efficient sound producers than head phones, and give a more uniform response over the frequency scale. As the loud speaker, or telephone receiver, should "match" the resistance of the plate circuit of the output tube, it is generally the practice to use an "output transformer" of suitable ratio.

The "Push-pull" Type of Amplifier.—Because of the fact that the form of the E_g–I_p curve of the triode is not a straight line, distortion in the amplified signal necessarily occurs. One arrangement which has been much used to eliminate some of this distortion resembles the double-button microphone discussed on p. 791. It requires two tubes per stage and

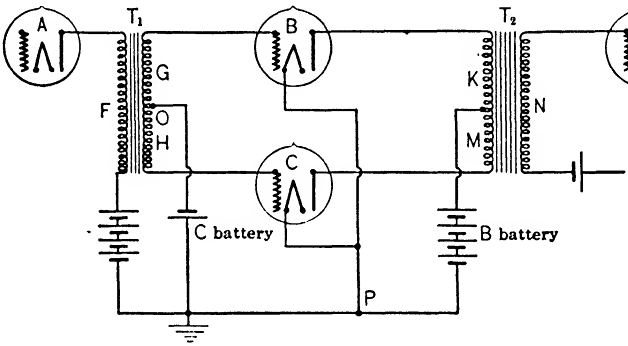


FIG. 28.—The circuit arrangement for a "push-pull" stage of amplification. It is not customary to put the output of a push-pull stage into a single triode as shown here because the distortion, which is prevented by the push-pull stage, is likely now to occur in tube N. Succeeding stages are also generally of the push-pull type.

so is not as efficient a use of apparatus as is ordinarily employed, but in transmitting stations the cost of a few extra tubes is of no consequence and here we find the balanced amplifier nearly always used. Nearly all modern broadcast receivers use the push-pull arrangement in the last stage, the one that supplies power to the loud speaker. Some transmitting sets use push-pull stages throughout. (See Fig. 41, p. 819.)

The two tubes are arranged as shown in Fig. 28. In this diagram A is the tube of one of the amplifier stages, F is the primary of repeating transformer T_1, by means of which the signal is repeated from tube A to the combination of tubes B and C. This combination forms, together with G–H and K–M, the push-pull circuit.

It will be noted that any change of potential taking place between points O and P will affect the grids of B and C similarly and will produce equal and opposite changes of ampere turns in K and M, hence a disturbance will not be repeated into N and the tube D, to which there may be connected a loud speaker. Tube D would be of much greater capacity than B and C, if such an arrangement were used, to prevent D introducing the distortion which the combination of B and C have avoided. In the ordinary receiving set amplifier, if a push-pull stage is used, the output of the transformer T_2 goes directly into the loud speaker. A signal coming through from F will produce equal and opposite potential changes in the grids of B and C and hence equal and co-directional changes in the ampere turns of K–M. This will repeat the signal into the secondary N and tube D.

Such an arrangement eliminates the distortion which would be produced by the asymmetry of the E_g–I_p curve on the two sides of the point on the curve determined by the C battery. Saying the same thing in other words, the push-pull arrangement eliminates distortion in so far as these are produced by *even harmonics* of the signal voltage.

It will generally be found advisable to " match " the tubes used in a push-pull amplifier if the results are to be appreciably better than those of a single-tube amplifier.

It will be noticed that transformer T_2 is not subject to any continuous m.m.f.; the plate current of tube B tends to magnetize the core in one direction and that of tube C tends to magnetize it in the opposite direction. If the two tubes are similar, so that they draw equal plate currents, *there is no flux in this transformer core when no signal is coming in.* Hence the self-induction of the primary windings is maintained at its highest value. (See Fig. 22.) The balance in m.m.f.s of parts K and M is destroyed when a signal comes in, because one plate current increases and the other decreases; this is the result of the grid connections; when that of B goes up in potential that of C goes down, and vice versa. Hence flux is set up in the core of T_2 which is proportional to the signal coming from tube A, and so the signal is repeated to tube D.

Effect of Push-pull Circuit on Distortion.—The arrangement of two tubes as in Fig. 28, because of its nullifying the effects of a certain part of the asymmetry of the E_g-I_p curve, permits each of the two tubes B and C to be excited by a much higher signal strength than would otherwise be advisable. In Fig. 29 are shown the percentages of harmonics in the output of a single 5-watt telephone repeater triode, as the grid was excited with a sine wave signal of increasing intensity.[1] It has been found by tests that even harmonics produce on the average listener a much more disagreeable quality of distortion than the odd ones, hence it is important to avoid even harmonic distortion. In Fig. 30 is shown the analysis of the output of two 5-watt tubes (each similar to the one used in getting the

[1] This and the succeeding figure are from results given in Bell Laboratories Record for April, 1927.

results of Fig. 29); it will be seen that a much greater output is obtainable from the two in push-pull connection than twice than obtainable from one tube, for a given amount of distortion. Thus with 2-watt output from the single tube there is 10 per cent of second harmonic, whereas with 4-watt output for the push-pull circuit the second harmonic content is only 1 per cent. Similarly with 3 watt-output from the single tube the fourth harmonic is 15 per cent of the fundamental, whereas with 6-watt output from the push-pull circuit the fourth harmonic is only a fraction of 1 per cent.

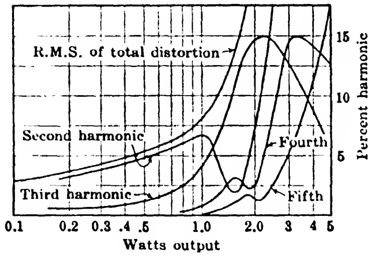


FIG. 29.—Distortion in the output of a 5-watt triode, as grid excitation is increased.

A High Output Push-pull Amplifier.—It is almost an axiom that amplifier tubes used for audio-frequency signals should be operated at that part of the E_g-I_p curve where the relation of plate current and grid voltage is linear for considerable variations of grid voltage either way from its " no signal " value. A tube so used gives a small output reasonably free from distortion; like the microphone, or any other device of this nature, it must be used inefficiently if its output is to be reasonably similar to the input. Furthermore, the grid should not be allowed to draw current.

A new tube just become available is recommended for use as a high power amplifier under conditions far from the conventional ones stated above. This tube, the type 46, has two helical grids one under the other. Both are brought out to pins on the tube base. This tetrode can be used as a triode as indicated in Fig. 31 *A*, with one grid tied to the plate, or as a triode as shown in sketch *B* of this figure, with both grids connected directly together to act as a control grid. Connected as in *A* its characteristics are those of an ordinary triode. $E_p=250$, $E_c=-33$, $\mu=5.6$, $R_p=2380$, $S_m=2350$, and $I_p=0.022$. The proper load resistance for maximum undistorted power is 6400 ohms, and the maximum undistorted power is 1.25 watts. (Power is said to be "undistorted" when the "total of the generated harmonics does not exceed 5 per cent of the fundamental," from the I.R.E. rules.)

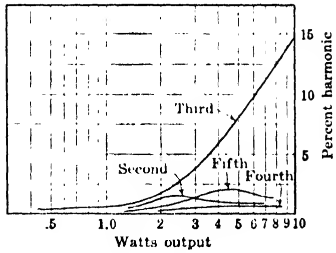


FIG. 30.—Using two 5-watt triodes, similar to that used in getting the results of Fig. 29, more than twice as much power is available with less distortion.

When connected as in B of Fig. 31, the tube has a considerably higher amplification factor, so much so that when the grid bias is zero the plate current is also essentially zero. (Contrasted with this is the fact that when connected as in A of Fig. 31 with the control grid at zero potential the plate current is about 50 milliamperes.)

With $E_p = 400$ volts, $E_c = 0$, $I_p = 0.006$ ampere. If the grid is allowed to go about 55 volts positive the plate current increases to 0.200 ampere.

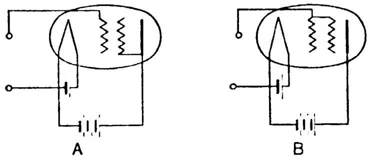


FIG. 31.—A modern tetrode can be connected as at A to give a high output tube, or as at B to give a high voltage amplification.

As this double grid triode is used, the grid fluctuates, with signal excitation, from −55 volts to +55 volts; during nearly all of the negative alternation of grid swing the plate current is zero whereas on the positive alternation the plate current varies almost linearly with the grid voltage, from 0.006 to 0.200 ampere.

The grid draws a large current on its positive swing, so that a considerable amount of power is required to excite it. When the signal is 40 volts (effective) the power used in the grid circuit is about 0.5 watt.

In Fig. 32 is shown the use of three of these type '46 tetrodes to take in a signal of a few volts and give a reasonably good output of 20 watts. Connected as a single-grid triode tube No. 3 serves as an amplifier to give about 1-watt output for exciting the other two tubes connected as double-grid tubes, connected in push-pull arrangement. It will be noticed that no grid bias is used in this stage.

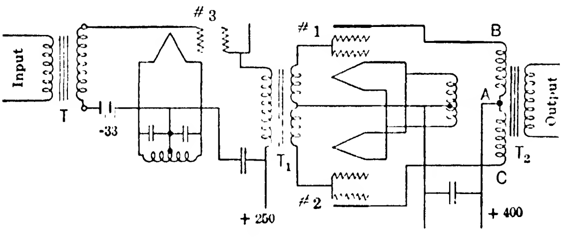


FIG. 32.—Three of the tetrodes of Fig. 31, thus connected, constitute a very efficient audio frequency amplifier, of large output. Tubes 1 and 2 constitute a type B amplifier.

The operation of the push-pull circuit can be understood with the help of Fig. 33, which shows how the currents of tubes 2 and 1 change as the signal goes positive and negative. As the signal goes positive the plate current of tube 1 increases linearly from 6 milliamperes up and the current of tube 2 (on which the actual voltage is going negative) decreases from 6 milliamperes to zero. Hence except for the low values of instantaneous signal voltage there is no current in part A–C of transformer T_2 primary, whereas part A–B is carrying a current varying linearly with the voltage impressed on the grid of tube 1. When the signal polarity reverses, the current of tube 1 is brought zero, and tube 2 has a plate current varying linearly with the voltage on the grid of this tube. As the magnetizing actions of the two plate currents in the primary winding of transformer T_2 (when two currents *do* flow, i.e., for low signal voltages) are in opposite directions it can be seen from Fig. 33, that as the signal voltage varies a small amount in either direction from zero the net magnetizing action of the primary coil will follow the dashed line shown in this diagram, and of course for larger signal voltages the magnetizing action follows the straight part of the E_g-I_p curve of one tube or the other. Referring again to Fig. 33 we see that as the signal voltage flucutates about zero as its average value, the change in magnetizing action in transformer T_2 follows the straight line A B. Thus the output of transformer T_2 is similar to the input of T_1.

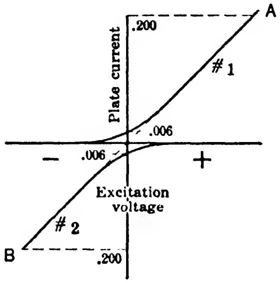


FIG. 33.—Plate current characteristics of the push-pull stage of Fig. 32.

As only one tube is working at a time, and the signal voltage is able to *increase* the plate current only, it has been suggested that the device be called a "push-push" amplifier.

To get full output of this amplifier it is necessary to supply a signal to the grid of about 40 volts (effective); this means that the secondary of transformer T_1 must give about 80 volts. As the tube 3 can generate in its plate circuit about 120 volts, it is seen that transformer T_1 can be built with a step-down turn ratio about 1.5 to 1.

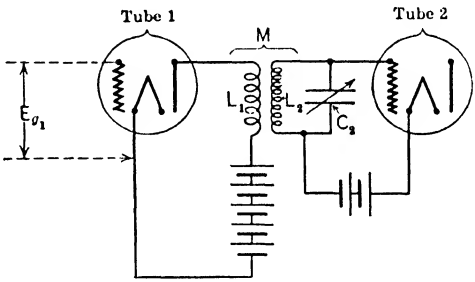


FIG. 34.—A stage of radio-frequency amplification on modern sets generally uses a transformer with tuned secondary. The step-up ratio may be as much as six to one, and the coupling between primary and secondary is generally between 40 per cent and 70 per cent.

Without danger of overheating, this amplifier can safely deliver 20 watts of output. The loss on each plate will be 10 watts when giving this output, so the plate-circuit efficiency is 50 per cent. Tube 3, acting as exciter, is operating at a plate-circuit efficiency of about 25 per cent.[1]

[1] For further discussion of these small, high-powered tubes see "High Audio Power from Relatively Small Tubes," by Barton, I.R.E., July, 1931, p. 1131.

Need for Tuned Radio-frequency Amplifiers.—There are two factors which make it advisable to have one or more stages of radio-frequency amplification, tuned, ahead of the detector. As was shown in Chapter VI the efficiency of any ordinary kind of rectifier, acting as detector, varies nearly with the square of the signal voltage. Thus if the strength of a signal, while it is still in radio-frequency form, can be amplified ten times the strength of the signal after it has been changed to audio frequency by the detecting tube will be one hundred times as strong. If we can use two stages of radio-frequency amplification ahead of the detector, each giving a voltage amplification of ten times the voltage of the signal at the detector, input will be one hundred times as great as though no radio-frequency amplification had been employed. And the current in the loud speaker will be *ten thousand times as great.* It is really the peculiarity of the law of detector action which makes radio-frequency amplification so useful from the standpoint of signal strength.

With the ever-increasing number of broadcasting stations crowding close to each other in the available frequency range, the advantage of the tuning available in two or more radio-frequency circuits spells the difference between success and failure of radio broadcasting. An untuned receiver today would be practically worthless.

Radio-frequency Transformer with Tuned Secondary.—In Fig. 34 is shown the simple circuit arrangement of one stage of a tuned radio-frequency amplifier; for broadcast frequencies the two coils L_1 and L_2 are generally single-layer solenoids, but for lower frequencies the primary coil L_1 is sometimes made of two or three sections wound in shallow slots in a molded form of some insulation material.

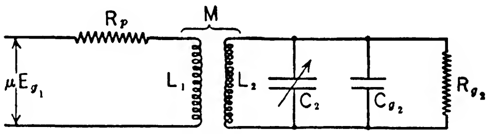


FIG. 35.—The circuit equivalent of Fig. 25. In this diagram the capacity between grid and plate of the triode has been neglected.

The circuit of Fig. 34 may be replaced by that of Fig. 35 in which C_{g_2} and R_{g_2} are the capacity and resistance of the input circuit of tube 2, and the signal voltage E_g actually impressed on the grid of tube 1 is replaced by the voltage μE_{g_1} acting directly in the plate circuit of tube 1.

We first replace R_{g_2} by an equivalent series resistance in accordance with Eq. (38) of Chapter II; for a resonant circuit this may be written

$$r=\frac{1}{\omega^2 C_2{}^2 R}=\frac{\omega^2 L_2{}^2}{R}.$$

Let this give for the total series resistance of the L_2–C_2 circuit a value R_2.

Now applying formulas (108) and (109) of Chapter I and remembering that we adjust C_2 for resonance with the signal frequency,

$$R'_1 = R_p + \frac{\overline{\omega M}^2}{R_2}$$

$$L'_1 = L_1.$$

Total effective primary impedance $Z'_1 = \sqrt{\overline{R'_1}^2 + \overline{\omega L_1}^2}$.

Primary circuit current $I_1 = \dfrac{\mu E_{g1}}{\sqrt{\overline{R'_1}^2 + \overline{\omega L_1}^2}}$.

Voltage induced in secondary coil, $E_2 = \omega M I_1 = \dfrac{\omega M \mu E_{g1}}{\sqrt{\overline{R'_1}^2 + \overline{\omega L_1}^2}}$.

Secondary current $I_2 = \dfrac{E_2}{R_2} = \dfrac{\omega M \mu E_{g_1}}{R_2\sqrt{\overline{R'_1}^2 + \overline{\omega L_1}^2}}$.

The voltage across the input circuit of the second tube, E_{g_2}, is evidently the same as the voltage across L_2 so

$$E_{g_2} = \omega L_2 I_2 = \frac{\omega^2 M L_2 \mu E_{g_1}}{R_2\sqrt{\overline{R'_1}^2 + \overline{\omega L_1}^2}}.$$

The quantity $\overline{\omega L_1}^2$ is ordinarily very small compared to $\overline{R'_1}^2$, so we may simplify the above to the approximate form

$$E_{g_2} = \frac{\omega^2 M L_2 \mu E_{g_1}}{R_2 R'_1} = \frac{\omega^2 M L_2 \mu E_{g_1}}{R_2 R_p + \overline{\omega M}^2},$$

and

$$\frac{E_{g_2}}{E_{g_1}} = \frac{\omega^2 M L_2 \mu}{R_2 R_p + \overline{\omega M}^2}. \qquad (17)$$

Evidently the voltage amplification depends upon the value of M. By differentiation the optimum value of M, so far as voltage amplification is concerned, is given by

$$M^2 = \frac{R_p R_2}{\omega^2}. \qquad (18)$$

Substituting this value of M in (17) gives as the highest amplification obtainable

$$\frac{E_{g_2}}{E_{g_1}} = \frac{\omega L_2 \mu \sqrt{R_p R_2}}{2R_p R_2} = \frac{\mu}{2\sqrt{R_p}} \times \frac{\omega L_2}{\sqrt{R_2}}. \qquad (19)$$

It appears then that to get the highest amplification we must use a good coil in the tuned circuit (high $\omega L_2/R_2$ ratio) and a tube should be used which has a high ratio of μ/R_p, that is, high transconductance. Selecting as typical values $\mu = 6$, $R_p = 10{,}000$ ohms, $L_1 = 36\ \mu h$, $L_2 = 256\ \mu h$, $\omega = 3 \times 10^6$, $R_{g_2} = 10^6$, $R_{2(\text{coil})} = 3.4$ ohms. Changing R_{g_2} into its equivalent series resistance gives $r = \overline{\omega L_2}^2/R_{g_2} = 0.59$ ohm. Hence the total equivalent series resistance of the secondary circuit is practically 4 ohms.

Now by Eq. (18) the optimum value of M is equal to $\sqrt{4 \times 10^4/(3 \times 10^6)}$ $= 66.6\ \mu h$. Using this value of M to get maximum amplification gives a ratio of $E_{g_2}/E_{g_1} = 11.5$. Thus one stage of this amplifier would raise the voltage for which it is tuned 11.5 times.

In an actual set using the constants stated above the manufacturer used a mutual induction of only 29 μh. This value of M, substituted in Eq. (17) gives

$$\frac{E_{g_2}}{E_{g_1}} = \frac{(9 \times 10^{12}) \times (29 \times 10^{-6}) \times (256 \times 10^{-6}) \times 6}{(4 \times 10^4) + (9 \times 10^{12})(29 \times 10^{-6})^2} = 8.45.$$

Thus dropping the mutual induction from 66.6 μh to 29 μh decreases the voltage amplification from 11.5 to 8.45.

In another set as manufactured $L_1 = 5\ \mu h$, $L_2 = 250\ \mu h$; $K = 0.5$; $M = 17\ \mu h$; and, for $\omega = 6 \times 10^6$, R_2 was 15 ohms. This gives a voltage amplification per stage (using a tube having $\mu = 8$) of 6.1.

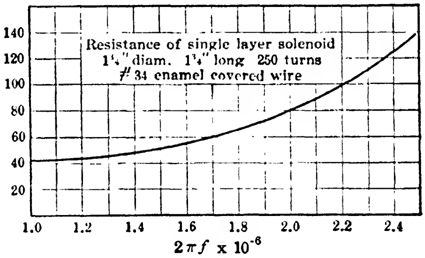


FIG. 36.—Resistance variation of a single layer solenoid designed for use at about 400 kc.

Effect of Coupling upon Selectivity.—From the above calculation it would seem that tight coupling should always be used between the primary and secondary coils of the transformer, to get the high value of voltage-amplification, yet in actual sets it is never done. The reason is that the selectivity of the tuned circuit rapidly decreases as the primary and secondary coupling is increased, and with the optimum coupling given in Eq. (18) the effective resistance of the secondary circuit is just doubled by the effect of the triode plate circuit. The selectivity of the circuit is therefore halved, and it is because of this effect that ordinarily a tuned radio-frequency amplifier uses less than the optimum value of coupling. Thus, in the set specified at the end of the above section, the mutual induction to get maximum amplification would be 45.6 μh, a value unattainable with the values of L_1 and L_2. This value of M would have increased the effective resistance of the tuned secondary circuit from 15 ohms to 30 ohms; with the value of M actually used the circuit resistance was increased from 15 ohms to only 17 ohms by the effect of the plate circuit of the triode.

Experimental Proof of Above Theory.—Pratt and Diamond [1] report the result of tests on radio-frequency amplifiers, used in aeroplane receivers; they show the effect of different construction of coils, and various couplings, on amplification and selectivity. In all their tests the secondary coil L_2 was a single-layer solenoid, $1\frac{3}{4}$ in. in diameter, winding $1\frac{3}{4}$ in. long, consisting of 250 turns of No. 34 enamel-covered wire. This coil showed a resistance with frequency variation, as shown in Fig. 36. Its inductance was nearly 0.002 henry.

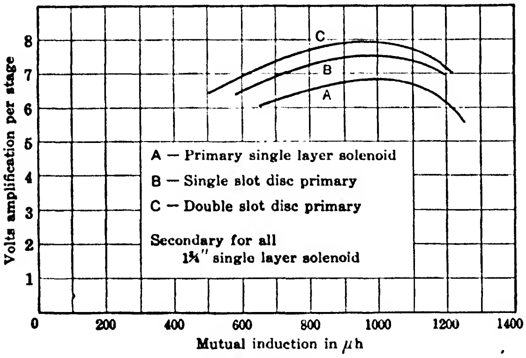


FIG. 37.—Variation of voltage amplification (R.F.) per stage as coupling is varied.

They tried the effect of using primaries of the single-layer solenoid type and of windings concentrated in one or more slots. The coils all had a diameter of $1\frac{1}{4}$ in.

By using more or less turns in the primary winding the mutual induction was changed and the amplification per stage was measured. Fig. 37 shows some of their results. Evidently the double-slotted primary was the best coil to use, and it is also evident from the curves that the value of mutual induction is not a very critical factor in determining maximum amplification. For instance, when M was changed from 1000 μh to 600 μh the voltage amplification changed only about 10 per cent. From the standpoint of selectivity the value of 600 μh would be much better than 1000 as the resistance introduced into the tuned circuit, by the plate circuit of the triode, would be only $(600/1000)^2$ or 0.36 as much.

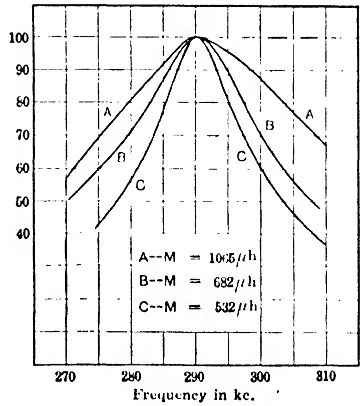


FIG. 38.—Effect of coupling upon selectivity of secondary circuit.

In Fig. 38 is shown the effect on selectivity, of the tuned circuit, of varying the coupling between primary and secondary coils, the primary being the double-slotted one of Fig. 37. It is evident that the greater selectivity is obtained when smaller coupling is used. The coupling of

[1] I.R.E., Feb., 1929, p. 283.

532 μh gives more than twice the selectivity had with $M = 1065$ μh, and inspection of Fig. 37 shows that the change in amplification as M is decreased from 1065 μh to 532 μh is only about 12 per cent.

Effect of Capacity Coupling between Transformer Coils.—The lower part of the primary and secondary coils are ordinarily connected together at the bottom; not actually connected together by a wire but by the use of bypass condensers they act as though they were connected together so far as alternating potentials are concerned. Then the upper part of the primary coil will act as toward the secondary as one side of a condenser, giving capacity coupling between the two coils as suggested by the small condensers C, C, of Fig. 39. This capacity coupling, in addition to the magnetic coupling between the coils, makes the solution of best mutual induction, etc., difficult.

The circuit is outlined in Fig. 40, the condenser C_0 representing the condensers C, C of Fig. 39. Diamond and Stowell[1] investigated this problem experimentally, and some of their results are given here. The secondary coil was the solenoid previously referred to, 1¾ in. in diameter and 1¾ in. long, of 250 turns of No. 34 enameled wire, having 1900 μh inductance. Their tests were carried out at 290 kc. where the coil had 70 ohms resistance. The triode they used had $\mu = 6$ and $R_p = 18{,}000$ ohms. The primary was wound as a single-layer solenoid 1¼ in. in diameter, of various lengths up to 1¾ in. The capacity coupling (C_0 of Fig. 40) naturally increased as the primary coil was made longer, i.e., wound with more turns. Their results were as follows:

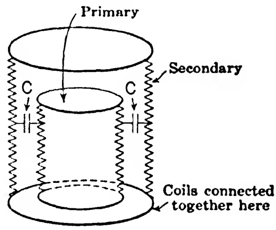


FIG. 39.—A radio frequency transformer has capacitive coupling as well as magnetic coupling.

$M\ \mu h$	$L_1\ \mu h$	$C_0\ \mu\mu f$	*Amplification*
200	103	8 1	5.05
400	289	11.5	7.90
600	514	14.2	8 65
800	748	16.3	8 90
1000	1000	18.2	8.80
1092	1265	20.0	8.45

By using the conditions given in Eqs. (18) and (19), it seems that the optimum coupling demands a mutual induction of 630 μh, and that with this amount the amplification per stage should be 9.5. The greater value of M required and the smaller amplification obtained may be due to the

[1] I.R.E., Sept., 1928, p. 1194.

resistance of the input circuit of the next triode, connected across the tuned secondary circuit, or it may have been due to the effect of the capacity coupling, which was not considered in deriving Eq. (18).

Effect of the Input Circuit upon Amplification.—If the resistance of the grid-filament circuit of the second tube is 10^6 ohms its equivalent series value is 2.25 ohms (when $L_2=250$ μh and $\omega=6\times10^6$) and so as the coil itself will have about 15 ohms resistance, the grid-filament circuit does not have a serious effect on the value of R_2. If, however, the value of R_{g_2} falls as low as 10^5 ohms (as is possible) then the resistance of the L_2–C_2 circuit is more than doubled as a result of the grid leakage. By holding the grid of the tube sufficiently negative, the resistance of the grid-filament circuit is maintained at a reasonably high value.

The capacity of the input circuit proves most troublesome in limiting the frequency range to which the circuit can be tuned. If the effective input capacity of the tube is 20 $\mu\mu f$ and that of the wiring, distributed capacity of L_2, etc., is 10 $\mu\mu f$ then the minimum circuit capacity is 30 $\mu\mu f$. If the maximum value of the condenser C_2 is 300 $\mu\mu f$ (a common value) the frequency range is only $\sqrt{330/30}=3.3$. This is about what is available in a well-designed set. If the wires have not been carefully placed, and stray capacities not reduced as much as possible, less frequency range than this is available.

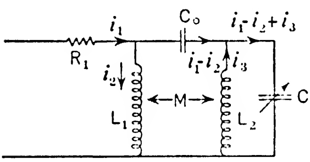


FIG. 40.—Circuit equivalent to that of Fig. 39.

It will nearly always be found that the tuning of the L_2–C_2 circuit is very broad at the higher frequencies. There are two reasons why this is so, resistance, and effect of the plate circuit. The effective resistance of the L_2–C_2 circuit is much higher than at the larger condenser settings, due primarily to the grid-filament resistance of the second tube. If the frequency range is 3.3 to 1, as above, then the equivalent series resistance of the grid-filament shunt resistance is ten times as great on the high frequency as on the low.

It is possible that this very great increase in resistance of the L_2–C_2 circuit is not noticed in an actual set because the regenerative action of the grid-plate capacity of the tube is more effective at high than at low frequencies. There are many actions taking place simultaneously in these circuits so it is difficult to make measurements showing the effect of only one of them.

Difficulties in Tuned Radio-frequency Amplifiers.—As was mentioned in the discussion of Fig. 81 of Chapter VI the effect of the capacity coupling, inside the triode, between the grid and plate circuits is very likely to result in oscillations in the tuned circuits. It is not necessary to have both circuits tuned; a tuned grid circuit, with sufficient inductive reactance in the plate circuit, may set up oscillations with no other coupling than that inside the tube. This becomes increasingly likely as the oscillatory circuit is made more efficient (less resistance) and as the circuit is adjusted for higher frequencies. Thus a radio-frequency, transformer-coupled, tuned amplifier may operate well at the larger values of the tuning condensers; as these are diminished the tuning becomes sharper, denoting to the skilled engineer the presence of a regenerative action, and as they are made still smaller, oscillations are set up in the tuned circuits and for reception of radiophone signals the amplifier is useless.

Use of the Loss Method to Prevent Oscillation.—Various schemes have been devised to control these oscillations in tuned amplifiers, in which the resistance of the oscillatory circuit is sufficiently increased at the higher frequencies to offset the tendency of the tube to generate oscillations.

Some attempts have been made to cause the required increase in resistance to take place automatically as the tuning condenser is turned to its smaller values, but this is a feat difficult of accomplishment because the regenerative action depends on several other features besides the value of the tuning condenser. Thus the grid bias, the voltage of the B battery, and filament temperature affect the regenerative action, and of course, the regenerative action may change by 25 to 50 per cent when a new tube is substituted for an old one. In general then the loss methods must be under the control of the operator.

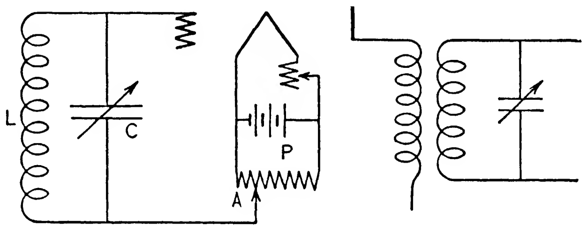


Fig. 41.—One scheme for controlling regeneration. The average potential of the grid is determined by the setting of A, on potentiometer P, and thus the effective resistance of the tuned input circuit is controlled.

Grid Bias Method of Control.—In Fig. 41 is shown the scheme which has frequently been used to stabilize high-frequency tuned amplifiers. A high resistance, P, is connected directly across the A battery and the tuned input circuit, L–C, is connected to the filament circuit at the sliding contact A, on this potentiometer. Thus the average grid potential (grid bias) can be given any value between that of the negative end and that of the positive end of the A battery. By reference to Fig. 75 of Chapter VI it will be seen that the resistance of the input circuit of the average tube can be changed from about 1 megohm down to a few thousand ohms by such a scheme. If higher values of input resistance are desired the filament rheostat may be put in the negative lead to the filament, instead of the positive as shown in Fig. 2, p. 980.

As has been previously pointed out, the resistance of the input circuit of the tube is a shunt resistance for the tuned circuit and may be changed to its equivalent series resistance by the relation

$$r_{\text{series}} = \frac{\omega^2 L^2}{R_{\text{shunt}}}.$$

It is thus possible to change the equivalent series resistance of the tuned circuit from its normal value of a few ohms to several hundred ohms, and this is generally sufficient to control the tendency of the tube to generate oscillations.

Another scheme for preventing undesired oscillations due to grid-plate capacity is to use a reasonably high resistance in each grid lead, right next to the grid. Several hundred ohms are sometimes necessary. (See p. 625.)

Balanced or Neutralized Radio-frequency Tuned Receivers.—The tendency of the tuned radio-frequency amplifier to generate oscillations is caused by the feeding of energy from the plate circuit to the grid circuit. The action must, of course, be a reciprocal one in the production of oscillations, because a condenser permits current to flow both ways. The grid potential first affects the plate current (for a disturbance starting in the grid circuit) and the change in plate current, acting through the inductive reactance in the plate circuit, produces a change in plate voltage. This change in plate voltage acts through the grid-plate capacity to cause a voltage in the grid circuit to maintain and augment the disturbance we have imagined there. From this viewpoint it is evident that if the grid potential is not allowed to change, due to the change in plate potential, this feed-back, or regeneration, action could not occur. Another, and opposing, voltage can be introduced into the grid circuit by an electromagnetic coupling between the grid and plate circuits. Such an expedient can be expected to work over a comparatively narrow frequency band, however, as it is not possible to just balance a capacitive feed-back by a magnetic feed-back throughout a wide range of frequencies. The magnetic feed-back must be made adjustable if such a scheme is to be most effective, and the operator will have to change the magnetic coupling as he changes the tuning condenser.

The better scheme is to utilize another capacity feed-back between the plate and grid circuits and to so arrange the circuit that the voltage impressed on the grid through this added condenser is just equal and opposite to that impressed on the grid through the plate-grid capacity. It will be seen that this scheme involves the selection of a suitable point in the grid circuit and connecting this point to the plate through the added balancing condenser, or the selection of a suitable point in the plate circuit and connecting this point to the grid through the balancing condenser. In case the circuit arrangement is such that the suitable points called for cannot be located it may be necessary to add a coil in one or the other circuit, or to seek a point in a circuit coupled to the grid or plate circuits.

In either case the requirement is very simple. Another feed-back must be introduced between the grid and plate circuits, utilizing a condenser through which to impress the additional desired voltage and arrange this added feed-back so that the voltage it supplies is just equal and opposite to that supplied through the grid-plate capacity. Many sets have been put on the market in which this balancing scheme is employed; some of them neutralize from grid circuit to plate and others neutralize from plate circuit to grid.

In Fig. 42 is shown one method of applying neutralization, this being of the so-called grid form. The filament is connected, not to the end of the coil, *AC*, of the input circuit, but at an intermediate point *B*, which may be the middle. Then condenser C_2 is connected as shown, and, if point *B* is the midpoint of coil *AC* then condenser C_2 is given a capacity equal to the grid-plate capacity of the tube, indicated in Fig. 42 by the condenser C_1. Merely by inspection it can be appreciated that any tendency to make the *L–C* circuit oscillate, due to voltage from the plate being impressed on point *A* through condenser C_1 will be nullified by the equal and opposite voltage impressed on point *C* through C_2. Thus a change in plate voltage cannot start oscillations in the *L–C* circuit. It will be further appreciated that any disturbance set up in the *L–C* circuit cannot affect the plate voltage as condensers C_1 and C_2 will produce equal and opposite effects on the plate for any current circulating in the *L–C* circuit.

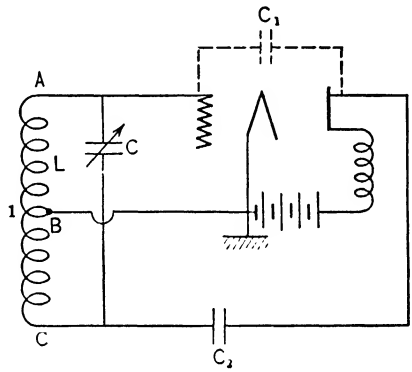


FIG. 42. In this scheme of neutralization the condenser C_2 is added to the circuit. If *B* is the mid-point of the coil, condenser C_2 is made equal to condenser C_1.

It is not necessary to have point *B* in the middle of coil *L* at all. It might be anywhere in the coil and the scheme will still work provided the capacity of condenser C_2 is properly chosen. It will be appreciated that, to get equal and opposite voltages on the plate the ratio of condensers C_1 and C_2 must be fixed by the ratio of voltages across *BC* and *BA*. But condenser C_1 cannot be altered as it is inherent in the tube, so condenser C_2 is the one which must be changed to effect neutralization. This is so adjusted that the ratio of C_1 to C_2 is the same as the ratio of turns from

CB to BA. The fewer turns there are in the coil L from C to B the larger must C_2 be to balance the effect of C_1.

To make the scheme work satisfactorily it is advisable to have reasonably tight magnetic coupling between the two parts of coil L, AB and BC.

In Fig. 43 is shown one arrangement for effecting neutralization from the plate side of the triode. As it is normally impossible to find a point directly in the plate circuit which has opposite potential to that of the plate we have recourse to the same expedient as used in Fig. 42. Instead of making the filament connection (through the B battery) to the end of the plate coil AC we connect it at a midpoint B. Then point C will go up and down in potential in phase opposite to the potential of the plate; just the same as in Fig. 42 points C and A have opposite voltages. In Fig. 43 the neutralizing condenser C_2 is connected between point C and the grid. If point B is not the midpoint of coil AC then C_2 will have a different capacity than C_1, its value being greater as the turns from B to C are fewer.

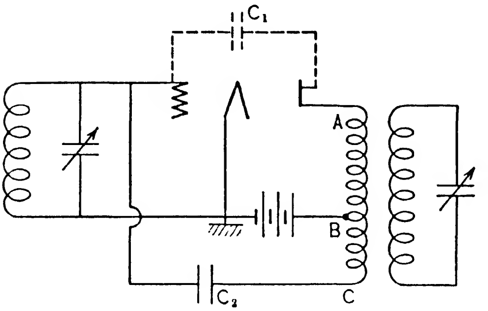


FIG. 43.—The so-called "plate form" of neutralization. Its action is fundamentally just the same as the "grid form" shown in Fig. 28.

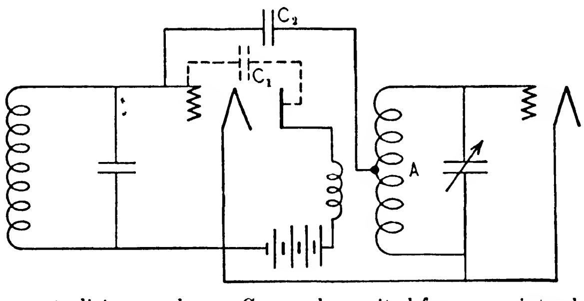


FIG. 44.—The neutralizing condenser C_2 may be excited from a point which is neither in the grid circuit nor plate circuit of the triode to be neutralized.

In Fig. 44 is shown the scheme which has been used more than any other on commercial sets. Here the point of potential opposite to that of the plate is found in the secondary of the plate-circuit transformer, that is, in the input circuit of the next tube. In Fig. 44, the point is shown as an intermediate one on the secondary coil but this point may, of course, be the end of the coil. In this scheme it is important that the primary and secondary coils of the plate transformer be connected with proper relative polarities, otherwise point A will have a voltage of the same phase as that of the plate and then of course the " neutralizing circuit " will act to give added regenerative action to the tube.

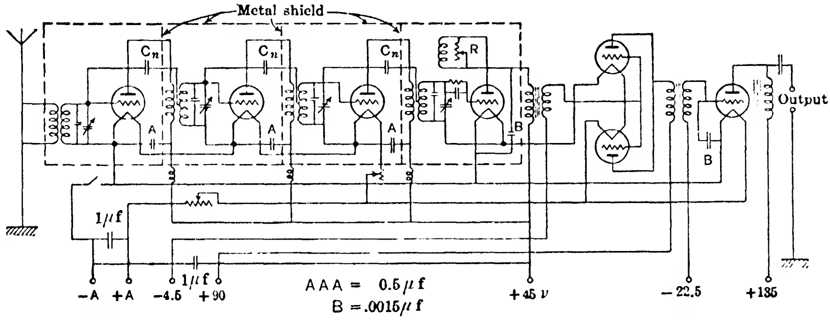


FIG. 45.—A neutralized (or balanced) radio frequency amplifier used in the radio beacon service.

For special receivers used in aeroplanes this type of circuit has found favor. Fig. 45 shows the circuit diagram of one of these as given by Pratt and Diamond.[1] There are three stages of r.f. neutralized amplification and a regenerated detector. Each of these stages is in a completely shielded compartment. As the circuits have to tune for only a narrow range (285 kc. to 350 kc.), a fixed condenser is used in each tuned circuit in addition to a small variable one of only 150 $\mu\mu f$ capacity. The regenerative action of the regenerative detector circuit is controlled by resistance R. The neutralizing condensers are shown by C_n. The set is battery operated, weighs only 18 lb., and yet delivers 6 milliwatts output when using a pole antenna only 10 feet high, having a capacity of 25 $\mu\mu f$. The field strength required, and the selectivity curve of the set, are shown in Fig. 46, and Fig. 47 shows the sensitivity throughout the frequency range for which the set is designed. The 6 milliwatts are sufficient to operate the reed indicators used in the radio beacon service. (See p. 968.)

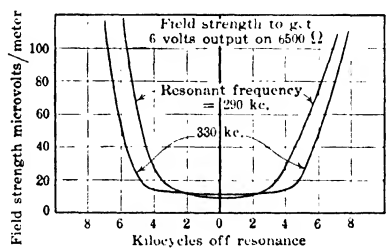


FIG. 46.—Selectivity curve of the set shown in Fig. 45.

Although the so-called "neutrodyne" receiver has already passed out of use, neutralizing condensers are still used in many modern transmitting sets, generally of the push-pull type. (See pp. 818–819.) In this type of

[1] I.R.E., Feb., 1929, p 283.

circuit, the point of "opposite potential" for furnishing the neutralizing voltage for one of the tubes is always found on the other tube of the push-pull pair.

Adjusting a Balanced Receiver.—The capacity of the balancing condenser C_2 (Fig. 44) is only a few micro-microfarads and it must be carefully adjusted to get the proper balance. It is frequently done by opening the filament circuit of the tube and sending a signal through the circuit; the filament being dead there is no plate current, so if the signal does come through it is coming through one of the condensers C_1 or C_2. The capacity of C_2 is then altered (by a screw driver having a long insulating handle) until minimum signal is heard.

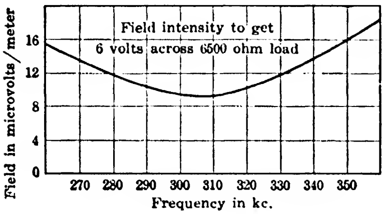


FIG. 47.—Sensitivity of the set shown in Fig. 45; a ten foot pole is the only antenna used.

Another method is to tune the amplifier circuits, input circuit of the tube being tested and the input of the succeeding tube, with the tube operating normally. With sufficient B battery voltage, the set will oscillate if the circuits are efficient. The oscillations will be heard if another oscillator is operating in the vicinity, at nearly the same frequency. When the beat note is heard the operator adjusts the value of C_2 until oscillations stop (the beat note disappears). The oscillations will be prevented by C_2 within a certain narrow range of its adjustment. These two limiting values of C_2 are noted and it is left set at a point midway between the two limits.

It will be appreciated that a set which has been accurately balanced at the factory may display the characteristics of a regenerative receiver when in use. Tubes fail and must be replaced by others and in the past the plate-grid capacity of tubes available on the market has varied within wide limits. Thus if the tube capacity with which the set was balanced was 6 $\mu\mu f$ and the new one put in has 8 $\mu\mu f$, the amplifier is almost sure to oscillate, especially at the shorter wave lengths.

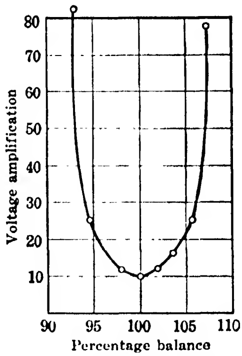


FIG. 48.—Effect of changing the neutralizing condenser of a balanced circuit.

Many of the neutralized sets on the market will not oscillate on the longer wave adjustments even if the neutralizing condenser is disconnected but will oscillate violently on the short-wave adjustments. Such a set is poorly built (in that its circuit resistances are too high) and is improperly balanced. The values of its neutralizing condensers should be changed.

Effect of Improper Setting of the Neutralizing Condenser.—Nelson [1] has analyzed the regenerative action of a poorly balanced circuit, following the method given by Beatty [2] and has experimentally tested the theoretical conclusions reach by Beatty. Fig. 48 shows the results of one of his tests. The radio-frequency stage with which he was working, when perfectly balanced, gave a voltage amplification of nearly 10. As the balancing condenser was made either too large or too small regenerative action set in, giving much greater amplification. If the condenser was changed from its proper value (a few micro-microfarads) by an amount much in excess of 5 per cent the regenerative action was so strong that the set went into oscillation. The width of this stable zone varied from 10 per cent variation for some triodes to 20 or 30 per cent for others.

Eliminating Stray Coupling.—The coupling between stages, or between parts of one stage, may tend to produce oscillations. To prevent capacitive coupling the wires of opposite electric polarities should be kept far apart, and be made small. To prevent magnetic coupling between various parts of the amplifier, wires should be short, and coils should be of astatic form or properly oriented one to the other. The axes of adjacent solenoidal coils may be made mutually perpendicular as one method but this does not lend itself so easily to manufacturing processes. If the solenoidal coils are arranged in a row, axes parallel, and making the proper angle with the line connecting the solenoids there will be zero magnetic coupling. This arrangement is indicated in Fig. 49. Hazeltine has given the value of the proper angle as 55°.

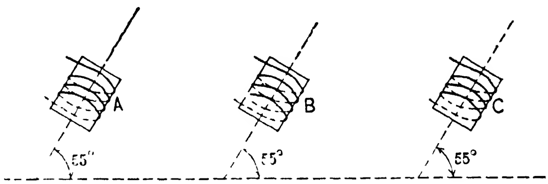


FIG. 49.—Short solenoids arranged as shown here will have practically no mutual induction.

Of course one of the expedients is to properly shield (see Fig. 45, p. 1019) successive stages from one another. Thus if a tube with its proper associated circuits is mounted on insulating material inside a copper can and the can is grounded the inside circuit is well shielded against radio-frequency disturbances. The various tuning condensers used in the r. f. stages should each be in a separate, metal enclosed chamber, the metal being grounded.

When, as is generally the case, a common B battery is used for all stages it is well to shunt it by a suitably large condenser to make its impedance low for the radio-frequency currents. As a battery is used

[1] I.R.E., Jan., 1930, p. 88.
[2] Experimental Wireless and Wireless Engineer, Jan., 1928.

up its resistance increases very much but a 1 μf condenser across its terminals will reduce the impedance for the radio currents to only a fraction of 1 ohm.

It will be noticed that all the illustrations of " power packs," that is, circuits fed by rectified alternating current, condensers for by-pass action are placed between each and every separate power lead to ground. In the radio set itself every part of the circuit which is not necessarily variable in potential is connected to ground through condensers of 0.1 μf capacity or larger.

Effect of Plate-Grid Capacity on Tuning.—It has been pointed out that the effect of the inter-electrode capacity is to cause regeneration in a circuit having tuned input and tuned output circuits and of course regeneration tends to sharpen the tuning of the circuits. However, if the regeneration is neutralized, by one of the schemes just discussed, another effect becomes noticeable.

The grid-plate capacity forms a coupling condenser for the two tuned circuits, which have a common connection at the filament. The arrangement of circuits is then exactly the same as that of Fig. 143, p. 147, and it is shown there that such a pair of coupled circuits gives two resonance peaks. The grid-plate capacity of the amplifier circuit corresponds to condenser C_3 of this figure. On p. 147 it is shown that the effect of C_3 (in giving two resonance humps) depends upon the ratio of C_3 to the capacity in the two tuned circuits.

In an actual broadcast receiver the effect of this tube-coupling capacity is not noticeable at the larger values of the tuning condensers but at the smaller values, where the capacity is about 50 $\mu\mu f$, the grid-plate capacity (which is about 10 $\mu\mu f$ on a modern amplifying tube) has a powerful enough influence to very materially broaden the resonance curve, and sometimes to actually produce two resonance peaks.

Use of Regeneration in a Balanced Amplifier.—The regenerative connection (Fig. 182, Chapter VI) of a receiving tube offers very great advantage over the non-regenerative (balanced) type of circuit in so far as sensitivity is concerned but has the very great disadvantage (in crowded communities) of acting like a miniature continuous-wave transmitter whenever it oscillates. Thus such receivers are of great harm in the broadcast field and should never be used because, even though not used in the oscillatory condition, during the adjustment of the receiver it is sure to oscillate and thus give the disagreeable whistling note in many nearby receiving sets. It is possible to get the advantage of regeneration without bothering other listeners by interposing one or two stages of balanced radio-frequency amplification between the regenerative tube and the antenna. Thus in Fig. 45, p. 1019, even if the detector should oscillate no radio-frequency could be fed to the antenna because of the three neutralized stages intervening the neutralizing condensers actually make a " one-way repeater " out of the triodes.

Amplifiers Using the Shield Grid Tube.—Practically all amplifiers are built nowadays with shield (screen) grid tubes, described on pp. 532 and 689. The effective capacity from control grid to plate in this type of tube is reduced to about 0.01 $\mu\mu f$, so small that no appreciable feed-back takes place from grid to plate.

The amplifying factor of this tetrode is very high, several hundred being the ordinary value. The plate circuit resistance (a.c.) is also very high, being generally 200,000 ohms or more. To match this high internal resistance by an outside load of equal amount is not possible, so the full amplifying power of these tubes is not obtainable. But badly as the load may match the internal resistance of the tube the available voltage amplification is much greater than can be realized with ordinary triodes, so the screen-grid tubes are almost universally used.

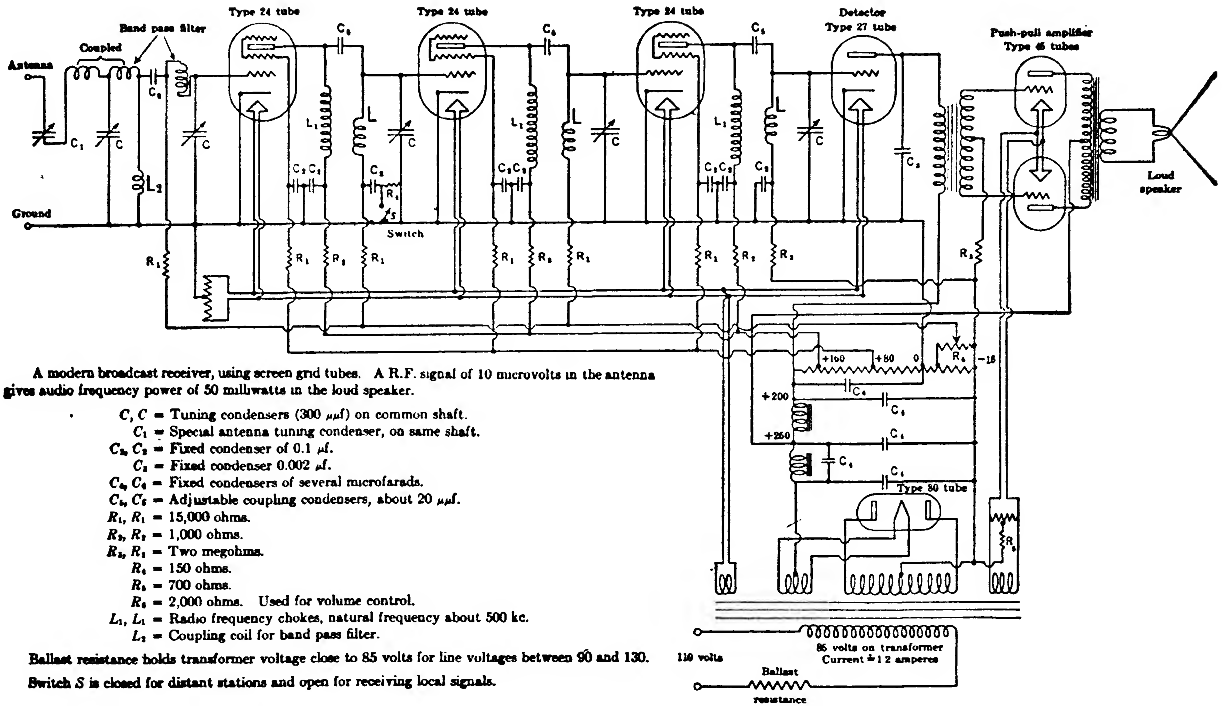
In Fig. 50 there is shown an amplifier having three of these shield grid tubes. A choke coil L_1 carries the normal plate current; it has sufficient inductance to show a very high impedance in the broadcast band. The tuned circuit connected to each input is made up of coil L and condenser C of about 300 $\mu\mu f$; at its resonant frequency this circuit shows about 50,000 ohms resistance so that about 25 per cent of the μ of the tube is utilized. The tuned circuits are coupled to the plate circuit of the preceding tubes by the small coupling condensers. Coils and condensers are in shielded compartments so that no stray fields can reach from one stage into the next.

When phonograph excitation of the audio end of the amplifier is wanted, the pick-up is inserted in the lead between condenser C_2 and ground, of the input circuit of the detector tube. This scheme then uses the detector tube as an audio amplifier stage, and gives plenty of loud speaker volume even with a low-voltage pick-up. When being so used the tuning condensers are turned to their minimum setting, and this adjustment automatically opens the lead from C_2 to ground so the phonograph attachment is ready to operate, and at the same time prevents radio-frequency signals from getting through.

By building an amplifier of screen-grid tubes, each in a separate copper-lined compartment Hull[1] reports unbelievably large amplifications. At 50 kc. he reports the voltage amplification per stage as 200, this diminishing with increasing frequency but still giving a voltage amplification of 7 per stage at 10,000 kc. He obtained an overall voltage amplification of 2 million at 50 kc. and 10,000 at 10,000 kc., the latter using five stages.

Resistance-repeating Amplifiers.—We will first discuss this type of amplifier relative to audio-frequency amplification. The diagram of Fig.

[1] See abstract in Phys. Rev., Vol. 23, p. 299, 1924.



A modern broadcast receiver, using screen grid tubes. A R.F. signal of 10 microvolts in the antenna gives audio frequency power of 50 milliwatts in the loud speaker.

C, C = Tuning condensers (300 $\mu\mu f$) on common shaft.
C_1 = Special antenna tuning condenser, on same shaft.
C_2, C_2 = Fixed condenser of 0.1 μf.
C_3 = Fixed condenser 0.002 μf.
C_4, C_4 = Fixed condensers of several microfarads.
C_5, C_5 = Adjustable coupling condensers, about 20 $\mu\mu f$.
R_1, R_1 = 15,000 ohms.
R_2, R_2 = 1,000 ohms.
R_3, R_3 = Two megohms.
R_4 = 150 ohms.
R_5 = 700 ohms.
R_6 = 2,000 ohms. Used for volume control.
L_1, L_1 = Radio frequency chokes, natural frequency about 500 kc.
L_2 = Coupling coil for band pass filter.

Ballast resistance holds transformer voltage close to 85 volts for line voltages between 90 and 130.
Switch S is closed for distant stations and open for receiving local signals.

FIG. 50.—A typical broadcast receiver using three stages of screen grid amplifiers; a push-pull stage feeds the loud speaker, through a transformer having a step down ratio of about 30, making the low resistance coil of the speaker appear as several thousand ohms to the output triodes.

51 shows such an amplifier for three stages. The incoming signal voltage is applied to the points QS and is caused to affect the grid of tube 1 through the means of the high resistance R. The grid and filament of tube 1

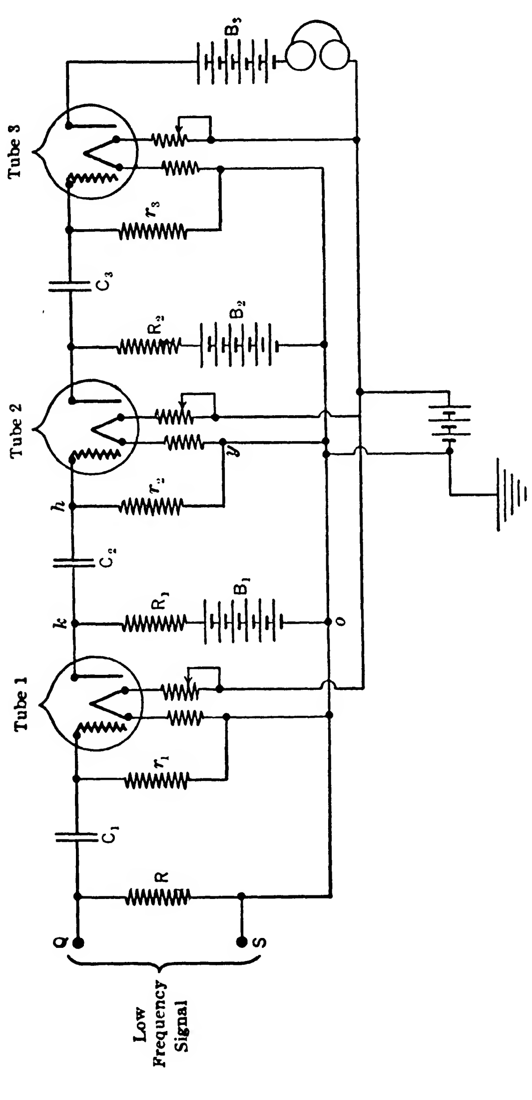


FIG. 51.—A resistance-repeating amplifier of three stages. A common power supply is now always used instead of plate batteries, and condensers of 0.1 μf or larger are connected from the lower ends of R_1 and R_2 to ground, to increase the stability of the amplifier.

are connected across the resistance R through the comparatively large condenser C_1; a leak resistance r_1 is connected from the grid to the filament. The purpose of the leak resistance and of the condenser C_1 will be explained later, but it will be presently understood that any variations of potential difference across R will be impressed upon the input circuit of tube 1 with the exception of any drop of potential which may take place in the condenser C_1.

The variations of the grid potential of tube 1 will cause a corresponding variation of the plate current in this tube, and hence a varying difference of potential will exist across the high resistance R_1. Since the point o is at constant potential it is plain that the potential difference between the points k and o will be varied and, as the battery resistance is comparatively low, the variation of this potential difference must necessarily be very nearly the same as that across R_1.

The grid and filament of tube 2 are connected across k and o through the comparatively large condenser C_2, and, therefore, any variation in the potential difference across k and o will be impressed upon the grid of tube 2, or, in other words the signal will be repeated into the second tube by means of the repeating resistance R_1.

In a similar manner the signal will be repeated from tube 2 to 3, where it will be picked up on the receivers. The purpose of the grid condensers C_2 and C_3 is to insulate the grids of tubes 2 and 3, respectively, from the batteries B_1 and B_2. Thus, if condenser C_2 were removed it is plain that the grid of tube 2 would then be connected to battery B_1 through the resistance R_1, and the battery would impress such a high positive potential upon the grid as to surely spoil the tube. A similar reasoning applies to the case of grid condenser C_1 in so far as it insulates the grid of tube 1 from any high direct electromotive force which may be to the left of the points QS; sometimes, as will be shown later, it is possible to dispense with the grid condenser C_1 and the resistances r_1 and R for the first tube.

As regards the leak resistances r_1, r_2, r_3, they are made necessary by the use of the insulating grid condensers C_1, C_2, and C_3. It has already been found in Chapter VI, p. 509, that, when a condenser is connected in series with the grid, if the grid is very highly insulated, the operation of the tube is very uncertain. The accumulation of electrons in the grid generally forces it to assume a negative potential of one or two volts, this amount depending upon filament current, etc. If a sudden pulse of e.m.f. (such as given by a "stray") is impressed on the grid it probably will accumulate sufficient electrons to force the plate current to zero and *this accumulated charge of electrons in the grid has no way of escaping.*

Of course as long as the plate current of one tube is zero the amplifier is "dead"; it is said to be "paralyzed" or "blocked." The grid of a triode should never be left "free" or "floating," as the behavior of the tube will then always be erratic. As to just how much leak resistance is required from grid to ground to make the tube stable depends upon the size of the tube and degree to which it has been pumped; it may be anything between 10^5 and 10^7 ohms for the small tubes used for amplifiers.

Suitable Values of Repeating Resistances.—Confining our attention to the repeating resistance from tube 1 to tube 2, i.e., R_1 (Fig. 51), let:

E_{g_1} = effective value of alternating voltage impressed upon the grid of first tube;

E_{g_2} = effective value of alternating voltage impressed upon the grid of second tube;

I_p = effective value of alternating component of plate current of tube 1;

μ = amplifying constant of tube 1;

R_p = a.c. resistance of plate-filament for first tube.

If we assume that the impedance of the circuit k–C_2–h–y–o is very high[1] as compared with the resistance R_1, then the impedance between the points k and o will be made up practically entirely of the resistance R_1, hence we may write

$$I_p=\frac{\mu E_{g1}}{R_p+R_1}. \qquad (20)$$

Again, assuming that the reactance C_2 is very low as compared with that of the grid-filament of tube 2, there will be a negligible drop of potential over C_2 and the voltage of the grid to filament for tube 2 will be given by:

$$E_{g_2}=I_pR_1=\frac{\mu E_{g_1}R_1}{R_p+R_1} \qquad (21)$$

and

$$\frac{E_{g_2}}{E_{g_1}}=\frac{\mu R_1}{R_p+R_1}. \qquad (22)$$

Eq. (22) shows that the ratio E_{g_2}/E_{g_1} increases continuously with increase of R_1 and it approaches a maximum which will be reached when R_1 is so large that R_p may be neglected; this maximum will be given by

$$\text{Maximum possible value of } \frac{E_{g_2}}{E_{g_1}}=\mu. \qquad (23)$$

This result is to be compared with that given by Eq. (14) applying to the case of a transformer repeating amplifier, for which

$$\text{Maximum possible value of } \frac{E_{g_2}}{E_{g_1}}=\mu\frac{n}{2}.$$

[1] At very high frequency the circuit made up of C_2 and the input circuit capacity of the triode, in series, will practically short circuit the resistance R_1 and so decrease the amplification to very low values.

In order to study more fully the relation expressed by Eq. (22) we have plotted the curve, Fig. 52, for which

$$\mu = 6$$

$$R_p = 10{,}000$$

The curve shows that it is hardly worth while to increase R_1 beyond about 30,000 ohms for this particular tube, for the gain in E_{g_2}/E_{g_1} is thereafter too small for even very large increases of R_1. Furthermore, it must not be forgotten that the insertion of a resistance in series with the plate requires a corresponding increase in the voltage of the B battery as previously pointed out. As a matter of fact such a tube would probably not be used with more than 20,000 ohms in the plate circuit. This would require a B battery of twice the voltage required if there was no IR drop in the external plate circuit and will give a voltage amplification of $\frac{2}{3}$ of μ (in the above case, 4).

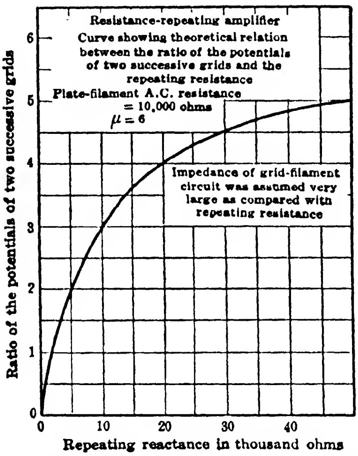


FIG. 52.—Variation in the amplifying power of a resistance-repeating tube as the value of the external resistance used in the plate circuit is varied.

As regards the first repeating resistance R it may be shown that it should be very high as compared with the resistance in series with it; the latter may be the plate-filament resistance of another tube or the resistance of a telephone line, etc.

The repeating resistances used are made up in units of small dimensions, approximately $\frac{1}{2}$ inch in diameter and 2 to 3 inches in length. There are three general types in use: Type 1 consists of a tube of insulating material wound with one layer of high-resistance wire and coated with enamel; it is made up in units up to about 5000 ohms. Type 2 consists of a tube of insulating material wound with a few turns of carbon filament containing a large percentage of clay and thus having a very high resistance; it is made up in units up to 50,000 ohms. Type 3 consists of an evacuated glass tube upon the inside walls of which there is "sputtered" a film of tungsten which is very thin and therefore of very high resistance; it is made up in units up to 2,000,000 ohms. Frequently a small strip of ink-coated paper is used, especially for the higher resistances.

In every case it must be kept in mind that no matter what type of resistance is used for repeating purposes it must have a current-carrying capacity such as will enable it to carry the average current flowing in the plate circuit of the tube wherein it is to be connected without overheating. Thus, in the case of a tube whose average plate current is 4 milliamperes a repeating resistance of 50,000 ohms should be able to dissipate 0.8 watt without overheating.

The repeating resistance should have negligible distributed capacity, for this would lower the value of its impedance and cause a reduction in the amplification.

Another important point regarding the resistances used for repeating comes up in connection with internal noises in an amplifier. It seems that some of the high-resistance units are "microphonic," that is, their resistance continually varies by a very small amount. It will be at once evident that such a resistance will give rise to noises in the amplifier, especially if the microphonic resistance is in one of the first stages of the amplifier. In general the higher the resistance the more likely is it to be microphonic.

Suitable Value of Grid Condenser.—The grid condenser must have a small reactance as compared with the circuit from grid to filament, which circuit consists of the leak resistance and the capacity and resistance of grid to filament; the point to keep in mind is that the variation of potential difference existing between the points k and o (see Fig. 51) should be made to suffer but a negligible drop over the reactance of the grid condenser, so that it may be applied very nearly in its entirety to the grid-filament circuit. For audio-frequencies the reactance of the capacity of the grid to filament is very high, i.e. (one to two million ohms), and does not appreciably affect the impedance between the grid and filament, which is almost entirely made up of the leak resistance and the internal grid to filament resistance in multiple, which make up a resistance of the order of 200,000 ohms. In this case the grid condenser may be allowed to have a reactance of 50,000 ohms without seriously affecting the grid voltage, or, in other words, for, say, 1000 cycles per second the capacity of the grid condenser may be about $1/50{,}000\times6280$ or, roughly, 3000 $\mu\mu f$.

If, however, the amplifier is used for high frequencies,[1] say $\lambda=600$ meters, then the impedance of grid to filament is made up almost wholly of the grid-filament capacity reactance, which, for the amplifying tubes generally used, is of the order of about 6000 ohms, hence the grid condenser reactance should be of the order of about 1500 ohms or less; its capacity may then be as low as 200 $\mu\mu f$ without decreasing the value of E_{g_2} more than 20 per cent. It is then apparent that smaller values of grid condenser

[1] It must be pointed out that the amplifier as arranged in Fig. 51 is not suitable for signals modulated at high frequency; the condensers in series with the grids rectify the wave trains so that in the later stages of the amplifier, only low-frequency signals occur.

capacity may be used at high than at low frequencies. In any case it is not advisable to use any larger capacity than just necessary, for in doing so, the amplifier is too likely to block for longer periods of time than necessary. If a pulse of e.m.f. is impressed on the amplifier all of these repeating condensers will become charged and so cut the various plate currents to probably zero. Before the amplifier can function the plate currents must come back to normal value and this requires that all these condensers (C_1, C_2, C_3, etc.) discharge themselves. The time required for discharge is fixed by the time constants, RC, of these condensers. Moreover if C_3 and C_2 discharge themselves before C_1 does they will charge up again when C_1 discharges, due to this discharge sending another pulse of e.m.f. through the amplifier. It is then evident that the time constant RC should be only a small fraction of the time between two " dots " of a signal, for example, if the blocking is not to interfere with reading the signal. Hence RC must be made small and this must be accomplished by making C as small as permissible, because if the leak resistance R is made small it would decrease the impedance of the grid-filament circuit so much that too large a proportion of the voltage $I_pR_{p_1}$ would be used up across the grid condenser, thus cutting down the voltage impressed on the grid. The proper relative values of R and C to keep RC small must therefore be a compromise.

Suitable Value of Leak Resistance.—The leak resistance should be as high as possible without causing any of the tubes to " block." The blocking would occur in case the grid became so negative as to make the plate current zero; the signal would, then, not go through until some of the electrons had escaped off the grid.

It is very difficult to lay down any exact rules or formulas as to the best value of the leak resistance since some of the quantities which affect it, such as the number of electrons collected on the grid, are somewhat indeterminate. It should be kept in mind, however, that a low leak resistance reduces the total impedance between points k and o on Fig. 51 and hence makes the drop over the repeating resistance very small, thus diminishing the amplification, and that a high leak resistance may cause the tube to " block." In most amplifiers the leak resistance is in the neighborhood of 1 to 5 million ohms.

Resistance-repeating Amplifier Independent of Frequency.—It has previously been shown that transformer-repeating amplifiers give a much greater increase of voltage per stage than a resistance-repeating amplifier, but it is to be pointed out that the resistance-repeating amplifier has one marked advantage over the transformer-repeating amplifier. The curves of Fig. 24 show how badly the voice may be distorted by a transformer amplifier, certain frequencies being amplified much more than others. The resistance amplifier is much superior in this respect as its amplification curve is practically flat. This feature is brought out by the curves of Fig. 53.

The amplification is essentially constant until the frequency gets so low that the reactance of the repeating condenser (C_1, C_2, and C_3 of Fig. 51) becomes large enough to compare with the grid leak. The repeating condenser should be as small as possible, if quick variations in voice amplitude are to be followed, but even so its reactance is negligible compared to the resistance of the leak except for very low frequencies. Thus if the repeating condenser has a capacity of 0.01 μf its reactance for 100 cycles is about 160,000 ohms. If the leak resistance is 1 megohm then even at 100 cycles the reactance of the condenser decreases the amplification by only 2 per cent. If, however, the condenser has a capacity of 0.01 μf and the leak resistance is only 50,000 ohms then the impedance through the condenser and leak is only about 200,000 ohms at 100 cycles and only 25 per cent of the signal voltage will appear across the grid of the next tube. The repeating condenser will absorb 75 per cent of the signal voltage.

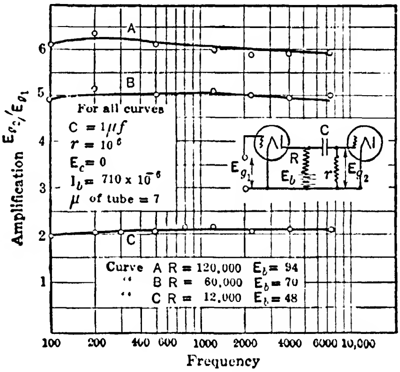


FIG. 53.—Showing the practically flat amplification characteristic of a resistance-coupled amplifier.

In Fig. 54 there is shown a resistance coupled amplifier[1] which has a voltage amplification of 3000 to 1. Up to 30 volts output its output is directly proportional to the input; its amplification is constant to within less than 5 per cent in the frequency range from 30 cycles to 100,000 cycles. Moreover there is a negligible phase shift for any frequency in this range, so that a distorted wave input will be exactly reproduced in the output. Each stage is carefully shielded from the others, and for the special circuit shown it proved advisable to have two separate power packs. All the resistors are small, of the so-called metallized type.

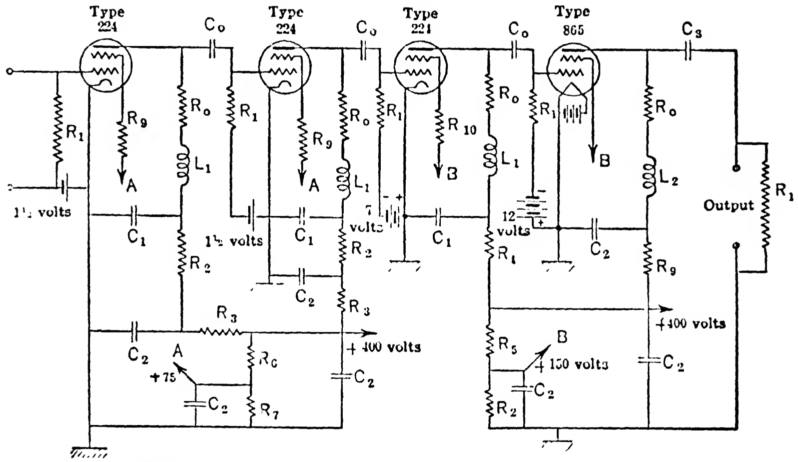
There will be increasing demand for amplifiers of this type as progress in the art of television is made. With the schemes of broadcasting television signals at present employed there must be transmitted side bands at least 100 kilocycles wide, and of course the receiving amplifier must "detect" this modulated signal and then amplify equally well all of the components of the modulation wave.

The cathode ray oscillograph (p. 44) is useful in testing such amplifiers. Its input impedance is so high that it does not appreciably disturb mag-

[1] Designed and built by Mr. H. J. Brown, E. E. Department of Columbia University.

nitude and phase relations when connected at various points of the circuit.

Inductance-repeating Amplifiers.—This type is similar to the resistance-repeating amplifier, except that instead of a resistance in the plate circuit of each amplifying tube an inductance is used whose reactance, at the frequency for which the amplifier is designed, is high. The theory upon which the repeating action from tube to tube is based is exactly



L_1 = 2.0 m.h
L_2 = 3.0 m.h
C_0 = .01 μf paper
C_1 = 1.0 " "
C_2 = 8.0 " electrolytic
C_3 = .25 " paper

R_0 = 10,000 ohms
R_1 = 250,000 "
R_2 = 50,000 "
R_3 = 10,000 "
R_4 = 25,000 "
R_5 = 40,000 "
R_6 = 100,000 "
R_7 = 15,000 "
R_9 = 2,000 "
R_{10} = 5,000 "

Heaters for three 224 tubes in series on same 7.5 volt battery as the 865

FIG. 54.—A resistance repeating amplifier which has a nearly flat characteristic between 30 cycles and 100 kilocycles. There is also a negligible change in wave form due to shift in phase of the upper harmonics. Two separate rectifying circuits supply power between ground and the points marked " + 400 volts."

the same as for the resistance-repeating amplifier and will not be gone into here again. This method of repeating has an advantage over resistance repeating in that the repeating inductance offers but little opposition to the flow of the direct current through the plate circuit and hence the B battery may be of lower voltage than if resistance repeating is used. For this reason the inductance-repeating amplifier is to be preferred to the resistance repeater for intermediate audio frequencies; but for high frequencies the distributed capacity of the inductance introduces difficulties which make it less desirable than the resistance repeater, and for very low frequencies the inductance required becomes excessively large.

Suitable Value of Repeating Inductance.—Let X_1 = reactance of the repeating inductance at the given frequency. Then, using the same symbols and making the same assumptions as in the similar discussion on the repeating resistance given on p. 1027, we have

$$\frac{E_{g_2}}{E_{g_1}}=\frac{\mu X_1}{\sqrt{R_p^2+X_1^2}}. \qquad (24)$$

The value of the ratio E_{g_2}/E_{g_1} increases continuously with increase of X_1 and has a maximum of μ which will take place when $X_1=\infty$. The relation between E_{g_2}/E_{g_1} and X_1 for a typical case is shown by the curve of Fig. 55, for which $\mu=6$ and $R_p=10{,}000$ ohms. Comparing this curve with the similar one for the resistance repeater (Fig. 52), it will be noted that the value of E_{g_2}/E_{g_1} rises much more sharply for the inductance repeater than for the other, and, as a matter of fact, for the same value of repeating impedance the resistance amplifier gives a considerably smaller ratio of E_{g_2}/E_{g_1} than does the inductance amplifier.

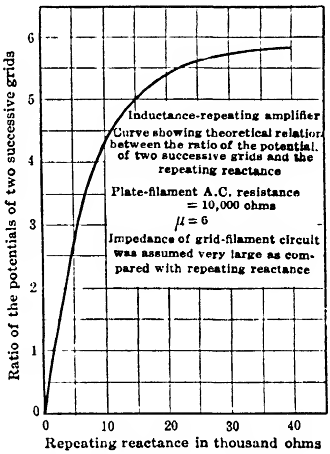


FIG. 55.—Amplifying characteristics of a tube using an inductance in the plate circuit; the amplification obtainable is much greater than with the same number of ohms of resistance.

The curve shows that there is very little to be gained by using a repeating reactance larger than about 20,000 ohms, or twice the resistance of the plate to filament. On the basis of 20,000 ohms for the repeating reactance the inductance would need to be about 60 henries for 50 cycles per second and 0.006 henry for 600 meters.

Of course the repeating inductance for audio-frequency is built on an iron core in view of its very large value. The construction of this inductance is regulated by the same principles as the construction of the repeating transformers for audio-frequency amplifiers, i.e., low iron losses and small distributed capacity together with small dimensions.

Combination of Resistance and Inductance Repeating.—In Fig. 56 is shown a circuit using a resistance in the plate circuit of one triode and an iron-core auto transformer feeding the input circuit of the next. The response of such a repeating circuit is about as shown in curve B of Fig. 57, curve A being the corresponding curve for transformer repeating. By choosing the inductance of the coil L and capacity of condenser C properly this circuit may be made to show a certain amount of resonance at a suitably low frequency; the resonance is not sharp because of the high effective resistance of coil L and the resistance R.

Amplifier for Continuous Current.—All the amplifiers so far considered fail at very low frequencies; neither inductance nor capacity can be utilized in an amplifier designed for continuous currents because one has zero impedance and the other has infinite impedance.

Lofton and White[1] have discussed various arrangements of so-called

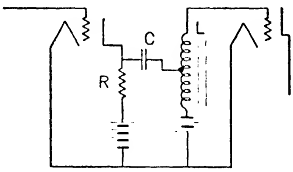


FIG. 56.—A combination resistance-inductance repeating amplifier; it gives high amplifications at the low frequencies.

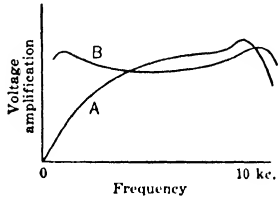


FIG. 57.—The amplifying curve of the arrangement of Fig. 56 is shown at B; curve A shows the corresponding curve for a transformer repeating amplifier.

direct-coupled amplifiers, to function at low frequencies; in Fig. 58 is shown one arrangement taken from their paper. In Fig. 59 is shown a scheme found convenient in our laboratory, utilizing the 350-volt potential battery for power supply. The circuit draws 1 ampere from the battery. The amplification curve is flat from zero to about 5000 cycles per second. By using a more suitable tube in place of the four 171 A's in parallel, the capacity to ground across the 1.2-megohm resistor could be much diminished and the frequency range could be much extended, to possibly 20 kc. Even though the input capacity of a 171A triode is only about 5 $\mu\mu f$ (including socket) the capacity reactance to ground through a condenser of 20 $\mu\mu f$ is only 800,000 ohms, at 10 kc., and as this capacity reactance is in parallel with the 1.2-megohm resistor used for repeating it is seen that the amplification must diminish before this frequency is reached.

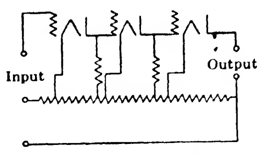


FIG. 58.—One arrangement of a continuous current amplifier.

With a 500-ohm resistive load on the output circuit the voltage amplification of the set is 31, up to about 5000 cycles; at 10,000 cycles it has fallen to 19.

The power input is 1 volt on 1 megohm, and the output is 30 volts on 500 ohms; the power amplification is then somewhat over 60 db.

This amplifier was designed for use with the Duddell oscillograph, as its output characteristic Fig. 60, shows. A device which will give 1 volt,

[1] I.R.E., April, 1930, p. 669.

on a resistance of 1 megohm, will deflect the oscillograph to full amplitude at any frequency within its range. Thus a carbon button microphone, through a suitable transformer (step-up turn ratio of 150), will make the output current of the amplifier fluctuate from 20 to 140 milliamperes, all the oscillograph vibrator is capable of handling. The microphone itself is developing only a few millivolts.

Principle of Superheterodyne, or Double Detection, Receiver.—As pointed out in the chapter on continuous wave telegraphy, if two radio-frequency voltages, f_1 and f_2, are applied to the grid of a suitably arranged detector tube there will be in the output circuit a current of frequency f_2-f_1, and this may well be of audible frequency.

This idea is used in the double detection receiver. The current generated by a local oscillator, combines with the signal current, to give a lower-frequency current, but this lower-frequency current is not of audible frequency. In various types of this kind of receiver the "difference frequency" has been from 50 kc. to 200 kc.; in the present type of broadcast receiver it is 180 kc. This beat frequency, in the range between the received radio frequency and the audio frequency, is called an *intermediate* frequency; this is abbreviated i.f., as the others are frequently abbreviated to r.f. and a.f.

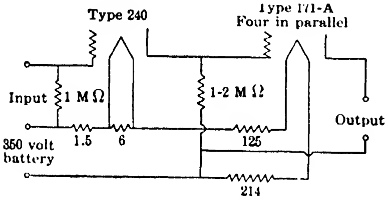


Fig. 59.—A very convenient amplifier for ordinary laboratory use; its output is just sufficient to operate a Duddell oscillograph.

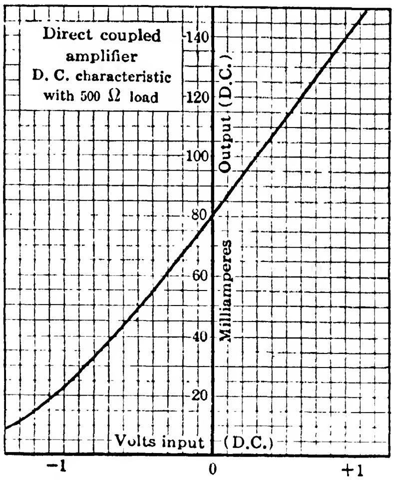


Fig. 60.—Characteristic curve of the amplifier shown in Fig. 59.

The radio-frequency current is frequently amplified in one or two tuned r.f. stages before being combined with the local oscillator to give the intermediate frequency, and sometimes the signal is combined to give i.f. in the first tube circuit of the receiver.

Modulation Carries Over to Intermediate Frequency.—If a radio frequency of f_1 cycles is modulated by the frequency p, then this current is combined with another of frequency f_2, and put through a detector, there will result the beat frequency f_1-f_2 and *this itself will be modulated with the frequency p.* This follows from Eq. (48), p. 586, which shows that the amplitude of the beat frequency current is proportional to the product of the amplitudes of the two radio-frequency currents from which it was derived. Hence if the locally generated current, of fixed amplitude, is given by $e_1 = E_1 \sin \omega_1 t$, and the modulated incoming r.f. signal is given by $e_2 = E_2 (1 - k \sin pt) \sin \omega_2 t$, and these are applied to the grid of a detector, there will be in the output of the tube a current of the form

$$e = E_1 E_2 (1 - k \sin pt) \sin (\omega_1 - \omega_2)t. \quad . \quad . \quad . \quad . \quad . \quad (25)$$

There will be other frequencies as well as this one, but as the detector output is supplied to an amplifier tuned to the frequency (ω_1-ω_2) the other frequencies are of no effect so far as the behavior of this type of receiver is concerned.

The ordinary superheterodyne receiver uses two or three stages of i.f. amplification, broadly tuned by the use of coupled tuned circuits between successive tubes. The output of this i.f. amplifier goes to a detector, generally of the plate rectification type, and from there to a push-pull output stage.

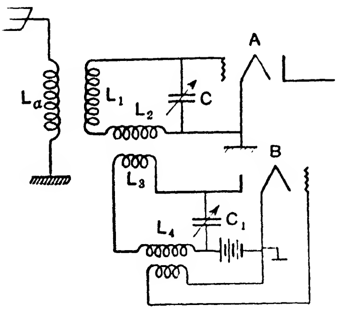


FIG. 61.—A beat frequency detector, with separate local oscillator.

This type of receiver is stable, is capable of high amplification, without objectionable frequency discrimination, and is highly selective.

Image Setting of Local Oscillator.—Evidently if a signal of 1000 kc. is to be received and the i.f. amplifier is tuned for 100 kc. the local oscillator could be set for either 1100 kc. or 900 kc.; the response would be exactly the same. However, if the tuning of the local oscillator is brought about by the same adjustment as is used for tuning the input circuit of the receiver, this double setting of the oscillator (for a given signal frequency) is done away with. The idea is illustrated in Fig. 61, condenser C tunes the input circuit of tube A for the incoming signal frequency, and condenser C_1 determines the frequency generated by the local oscillator tube B. If the condenser C_1 is controlled independently of C, then for every signal received by tube A there will be found two settings of C_1 which work equally well. However, if C_1 and C are on the same shaft this possibility is eliminated.

Requirements for Common Condenser Control.—If ordinary condensers are used for C and C_1 it will be found that the difference in frequency for which their respective circuits are tuned varies widely with different settings, if the two condensers are on the same shaft. The frequency for which a circuit is tuned varies inversely with the square root of the capacity value, and so is a peculiar type of curve, when plotted against condenser setting, if the capacity varies directly with the setting. This type of curve is shown in *A*, Fig. 62; the lowest frequency is given by the highest setting of the condenser, so the condenser scale given varies from 100 divisions down to zero.

Now if two circuits were equipped with such condensers, and these are adjusted at their full-scale settings to give to the two circuits a frequency difference of 100 kc., let us say, and then the two condensers are clamped together for common control, the difference in frequency would change from 100 kc., at maximum setting, to possibly 1000 kc. at the zero setting. This is shown by curves *A* and *B* of Fig. 62; they start 100 kc. apart at the 100-division setting of the condensers, and show an increasing frequency difference as the circuits are tuned for the higher frequencies.

However, if the straight-line frequency type of condenser is used (see p. 249), giving to the two circuits the tuning characteristics shown by curves *C* and *D*, then if they are adjusted for 100-kc. difference at any point on the scale they will maintain this difference throughout their range. The inductances used in the two circuits must be the same, and one of the *S.L.F.* condensers must have a greater capacity than the other to obtain the curves *C* and *D*.

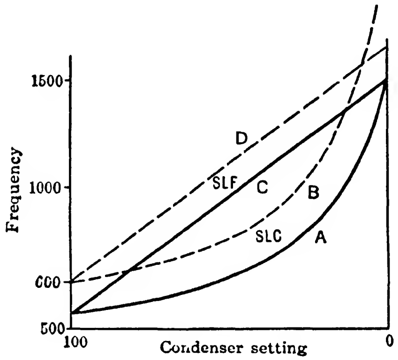


FIG. 62.—If two condensers, on the same shaft, are to maintain a constant *difference* in the tuning of their respective circuits they must be of the straight-line frequency type.

By thus having the tuning condenser C, and the oscillator condenser C_1, of the proper type, and fastened to the same shaft (or arranged for common control by some other mechanical arrangement) the "two setting" condition of the oscillator is done away with.

Elimination of Image Signal.—Let us suppose that the signal being received is 1000 kc. and the oscillator is set for 1100 kc. (100 kc. i.f. amplifier assumed). If a signal of 1200 kc. is impressed on the receiver input it too will give a beat frequency of 100 kc. and so be passed through the i.f. amplifier. Of course, the tuned circuit L_1–L_2–C will be tuned for 1000 kc., so the 1200-kc. signal will tend to be much cut down in strength; but it may be that the 1000-kc. signal is coming from a station 1000 miles away and so has a strength of a few microvolts per meter and the 1200-kc. signal (the interfering one) is from a local station and has a field strength of many millivolts per meter. In this case, the interfering signal will be much stronger than the desired one in spite of the selective action of circuit L_1–L_2–C_1.

Various schemes are possible for getting rid of such interfering signals, one of which is shown in Fig. 63. Between the grid and ground is put a circuit consisting of the *S.L.F.* condenser C_2 and coil L_3; these elements are so proportioned that their circuit is continually tuned to a frequency 200 kc. above that for which the circuit L_1–L_2–C is tuned. This will result in the grid being virtually short-circuited for the undesired signal. This extra circuit will of course decrease the strength of the desired signal somewhat, but not to a great extent. The three condensers C, C_1, and C_2, are all mounted on the same shaft for common control.

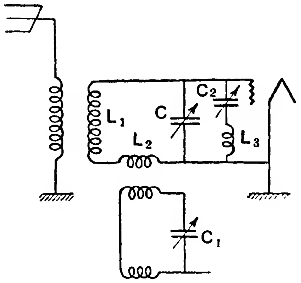


FIG. 63.—A properly designed circuit, L_3-C_2, with proper condenser will eliminate the image signal of the superheterodyne receiver.

Disadvantage of Circuit Shown in Fig. 61.—The L_1–L_2–C circuit has a high-frequency current set up in it by the local oscillator and hence will feed high-frequency power into the antenna from which it will be radiated to interfere with other listeners in the vicinity. The best scheme for eliminating this condition is to put a high-frequency stage between the antenna and circuit L_1–L_2–C, which stage uses a neutralized circuit or a screengrid tube, as shown in Fig. 64.

The Intermediate Frequency Amplifier.—If the beat frequency has been selected as 180 kc., a commonly used value, the i.f. amplifier must be tuned for this frequency. Now the amplifier should transmit equally well a band of frequencies either side of 180 kc., equal in width to the highest audio frequency which it is desired to preserve, say 8 kc. This amplifier should then transmit well all frequencies between 172 and 188 kc. This can be done to a certain degree by using the resonance phenomena of coupled tuned circuits, analyzed on pp. 133 to 150. The circuit used is shown in Fig. 65, and the curves given there show the characteristics of one circuit alone and the pair loosely coupled. Circuits L_1–C_1 and L_2–C_2 should be sharply resonant and should be coupled an amount somewhat

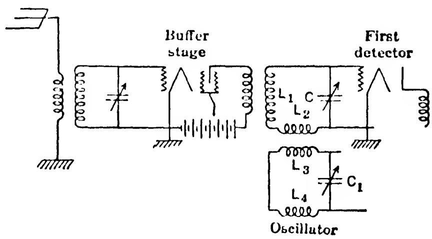


FIG. 64.—A "buffer" stage should be used to keep the locally generated high-frequency current out of the antenna.

less than that given by the fraction $(188-172)/180 = 8.9$ per cent. For the case given here, the resonance peaks of the coupled resonance curve might well come at 175 kc. and 185 kc. so that a coupling of $10/180 = 5.5$ per cent would be about right.

The two coils L_1 and L_2 are generally of the "honeycomb" type

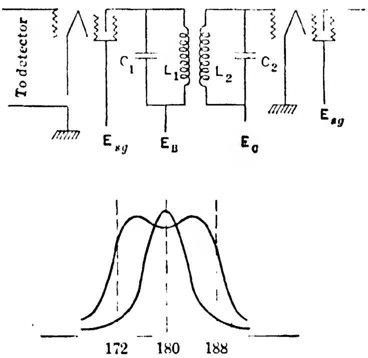


FIG. 65.—The intermediate frequency amplifier, having loosely coupled, tuned circuits, may be designed to give the desired flat-top resonance curve.

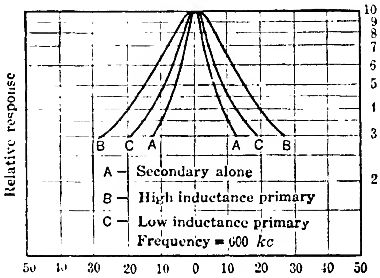


FIG. 66.—Selectivity of the r.f. stage of a superheterodyne receiver for high and low inductance primary.

shown on p. 231. They, as well as the condensers C_1 and C_2, are suitably mounted in a copper can to prevent interstage coupling. The condensers are made adjustable to a certain extent so that the respective circuits can be tuned after the set is assembled.

Performance of Superheterodyne Receiver.—Beers and Carlson [1] have analyzed in detail the performance of such a receiver and some of their results are given here. They first analyze the action of the one or two r.f. stages which precede the first detector, where the r.f. signal combines with local oscillations to give the intermediate frequency. For the primary coil of the r.f. transformers they advocate using a coil of much more inductance than had been customary, pointing out that by so doing the selectivity of the set at the two limits of the broadcast band is made more nearly the same. Ordinarily, the selectivity is very much poorer for the high frequencies. Figs. 66 and 67 bring out this comparison.

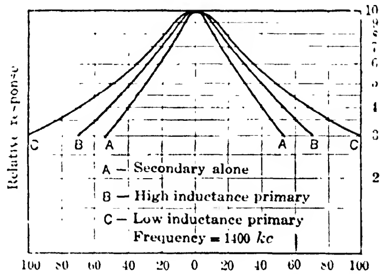


FIG. 67.—Curves similar to those of Fig. 66, for a frequency at the upper end of the broadcast spectrum.

[1] I.R.E., March, 1929, p. 501.

The tuned secondary alone gives the selectivity shown by curve A of these two diagrams, curves C show the performance of the r.f. transformer for the ordinary low inductance primary, and curves B show the action

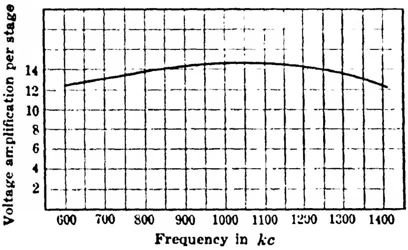


FIG. 68.—Voltage amplification (R.F.) per stage of a typical superheterodyne receiver, having some R. F. amplification ahead of the first detector.

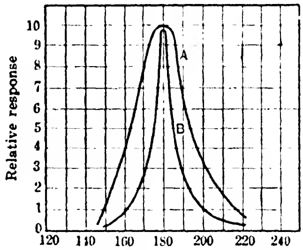


FIG. 69.—Selectivity characteristics of (B), one circuit of the I.F. amplifier of Fig. 165 and, (A), of the two circuits of Fig. 165 with the proper degree of coupling.

when a high inductance primary is used. It is seen that for the low inductance primary the resonance curve for 30 per cent of maximum response is 36 kc. wide at 600 kc. and 200 kc. wide at 1400 kc., a variation of 6 to 1; when the high inductance primary is used the corresponding variation is only 3 to 1. The reasonably uniform amplification obtainable over the broadcast range is shown by Fig. 68.

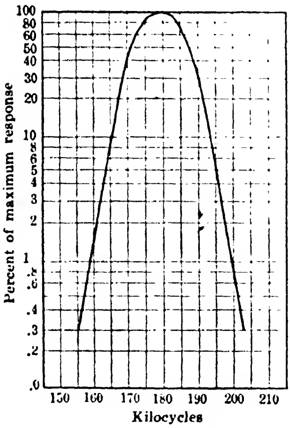


FIG. 70.—Selectivity curve of the R.F. and I.F. amplifier of a superheterodyne receiver.

One stage of the i.f. amplifier shows resonance characteristics as given in Fig. 69. One of the tuned circuits (L_1–C_1 or L_2–C_2 of Fig. 65) shows the selectivity of curve B, whereas the pair, properly coupled, show the selectivity given by curve A. The response is reasonably uniform over a frequency band about 12 kc. wide.

In Fig. 70 is shown the over all response of the combined r.f. and i.f. amplifier stages; there were two stages of r.f., the detector, and two stages of i.f. in the set giving this selectivity curve.

The second detector was supplied with the output of the i.f. amplifier; it was a type 227 tube arranged for plate-circuit rectification (grid biased properly). Its audio-frequency output was nearly proportional to the input, from the i.f. amplifier, as shown by Fig. 71. This second detector gives enough output voltage to feed directly the output stage of the set, a pair of power tubes connected in a push-pull circuit.

In Fig. 72 are shown the relative selectivities of the set for frequencies at both extremes of the broadcast band, as well as one in the middle.

Volume Control.—In the modern set the type '35 tube (variable μ) is used for both r.f. and i.f. amplifiers. The volume control is accomplished

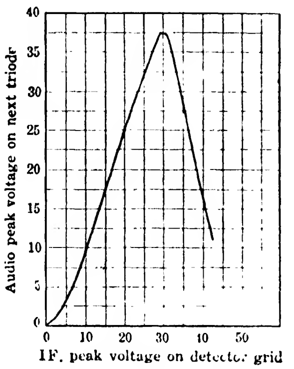


FIG. 71.—Action of the second detector, as output of the I.F. amplifier was increased.

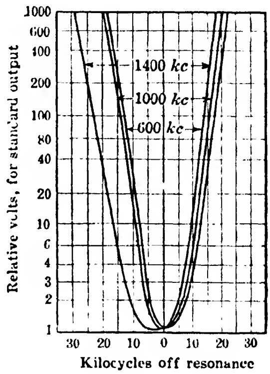


FIG. 72.—Selectivity of a superheterodyne at both extremes of the broadcast spectrum, with one intermediate frequency.

by varying the grid bias of these tubes. It is possible to do this automatically, the strength of the signal itself properly regulating the biasing voltage.

One scheme for accomplishing this is shown in Fig. 73; an auxiliary

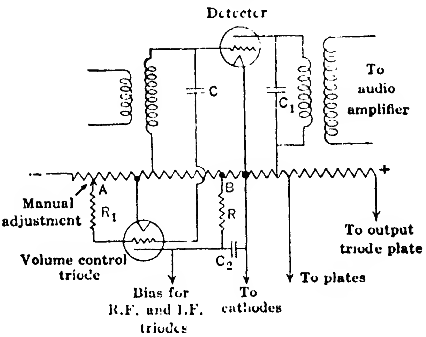


FIG. 73.—A volume control, having both automatic and manually adjusted control.

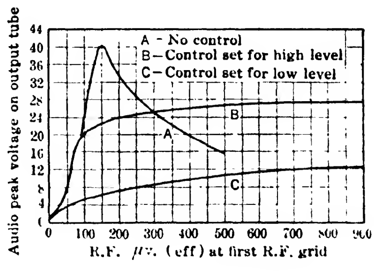


FIG. 74.—Action of the volume control of Fig. 73; curve A shows the response of the set with no control.

triode is used for controlling the bias voltage of all r.f. and i.f. tubes. The grid of this volume-control tube is controlled in two ways; with no signal coming in (detector grid not excited) the grid potential is fixed by the position of contact A, this being manually adjustable. The resistance

R_1 permits the potential of the grid to vary, however, for any one position of A. When the detector grid is excited by the signal, the grid of the control triode is also affected, through condenser C. The more this grid fluctuates (stronger signal) the greater is the average plate current of the control tube and hence the greater the drop through resistance R. But this greater drop through R results in a greater bias on the grids of the amplifying tubes.

Fig. 74 shows how this automatic control functions; as the r.f. voltage on the grid of the first r.f. tube increases from zero the output of the set increases, but when this output reaches a certain volume (determined by

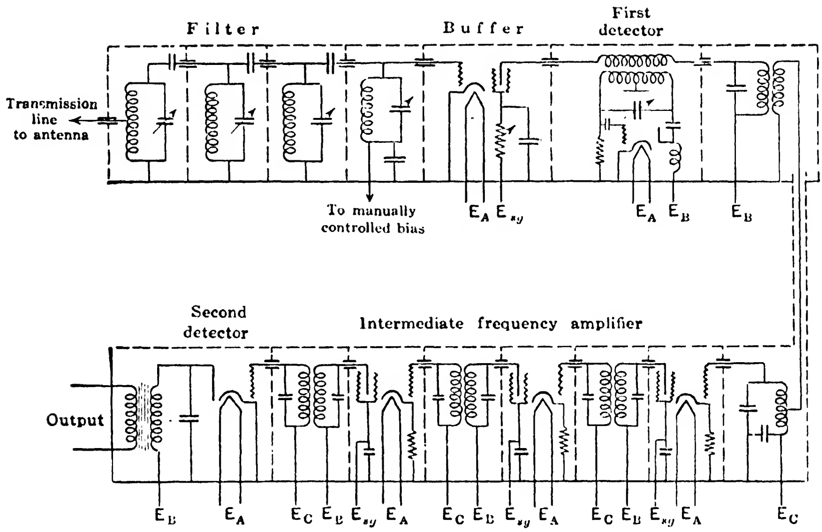


FIG. 75.—Completely shielded filtered superheterodyne receiver for short waves; the first detector is also the oscillator, as it is not worth while to use a separate oscillator when receiving short waves.

position of contact A) any increase in signal intensity has but little effect on the loud-speaker response. Without automatic control the set becomes overloaded (at the detector) with only 150 microvolts on the first r.f. tube, whereas with the control in operation, signals with ten times this intensity do not give much more than normal output, which is generally taken as 50 milliwatts.

In Fig. 75 is shown a reasonably complete diagram of a superheterodyne receiver designed to operate on a short-wave radio telephone channel from ship to shore, and vice versa; it is from a paper by Wilson and Espenschied.[1]

[1] B.S.T.J., July, 1930.

Power Levels in a Broadcast Receiver.—In Fig. 76 are shown the power levels in the various stages of a modern radio receiver using screen grid tubes for amplifiers, a '27 type tube for detector, and a power tube (or pair of them in push-pull connection) for supplying power for the loud speaker. The output tube should be capable of delivering, without appreciable distortion, as much as 5 watts of power, although on the average its output is measured in a few milliwatts. Unless the large reserve of power is available a very disagreeable type of distortion occurs during the loud parts of the program; a peculiar "rattling" noise from the loud speaker is frequently caused by tubes being overloaded.

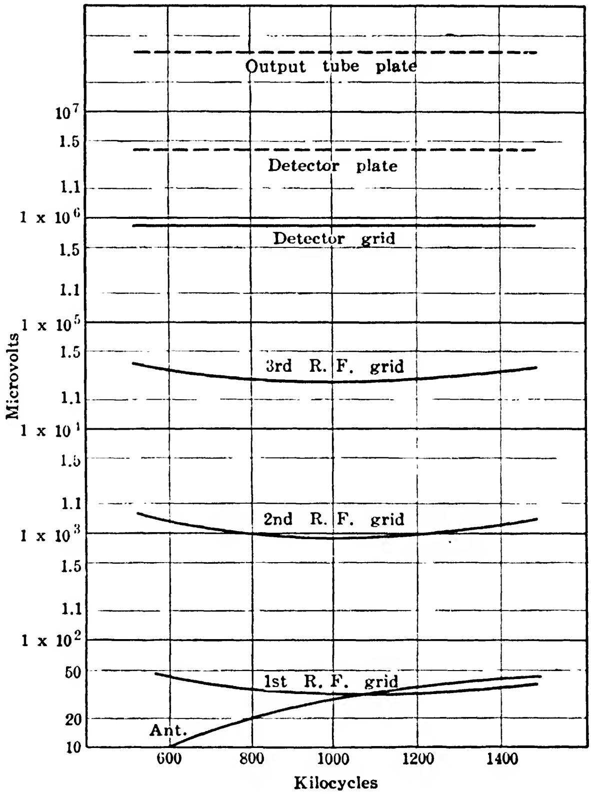


Fig. 76.—Typical voltage levels in a modern broadcast receiver.

Distortion Produced by Amplifiers.—The current form delivered by an amplifier to its loud speaker should resemble the amplitude variation of the radio-frequency current supplied by the antenna, if it is to be faithful in its reproduction. In general such is far from the actual condition; the current delivered to the loud speaker by many radio receivers is scarcely recognizable as the signal delivered to the input terminals of the amplifier. There are many causes for this distortion, only one or two of which have already been mentioned; they will all be summarized in this section.

Radio-frequency Amplifier Not Linear.—In Fig. 77 there is shown the manner in which a non-linear amplifier distorts a modulated r.f. signal; the voltage impressed on the grid is r.f. with sine wave modulation but the r.f. output of the plate circuit is by no means modulated with a simple sine wave. There are harmonics in the modulation, which will be carried through and sent out from the loud speaker. If the amplification is controlled by grid bias, the amount of r.f. signal (100 per cent modulated) which can be impressed decreases, ordinarily, with increasing bias. In Fig. 78 are shown the behaviors of types '24 and '35 tubes; if the modulation is not to be increased (by addition of second harmonic principally) by more than 20 per cent the r.f. signal voltage on the grids must be limited to the values shown in Fig. 78.

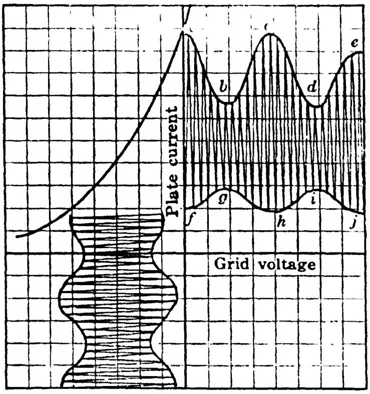


FIG. 77.—Due to the curvature of the I_p–E_g characteristic the average plate current of a R.F. amplifier increases as the grid excitation increases. This action is greater as grid bias is increased.

Tuning Too Sharp.—The correct reproduction of music as delivered over the radio broadcast channel today requires that the amplifier and its associated circuits shall transmit and amplify to an equal degree all frequencies within about 10 kc. of the carrier frequency, otherwise the high-frequency notes are lost. This requirement is satisfied by having all circuits broad enough in their resonance curves. It is practically impossible to construct a circuit with tuning so sharp that this condition is not fulfilled.

However, when the circuit is connected to a triode the regenerative action of this device may sharpen the resonance curve to an exceptional degree, thus transmitting the lower voice or music frequencies much more strongly than the upper. If there is regeneration in a receiving set it should be strictly limited in amount.

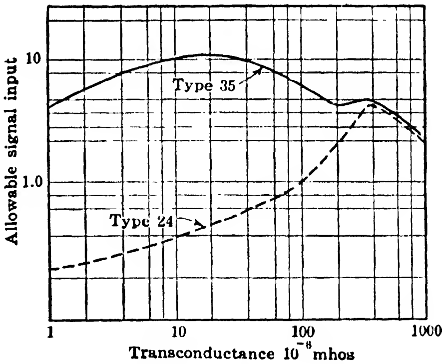


FIG. 78.—Allowable signal voltage on two types of tubes, for a given amount of modulation distortion.

Detector Overloaded.—If grid rectification is used in the detector (grid condenser and leak used) it is very easy to overload it, so that its audio-frequency output is no longer proportional to the modulated radio-frequency signal applied to its grid.

Certain tests on the type '24 tube with various conditions as to screen grid volts, and resistance in the plate circuit (across which resistance the audio output is obtained) showed results as given in Fig. 79. The response is not linear unless conditions are properly chosen, but the amount of variation from linearity shown here would probably not give enough distortion to the output for the average listener to perceive.

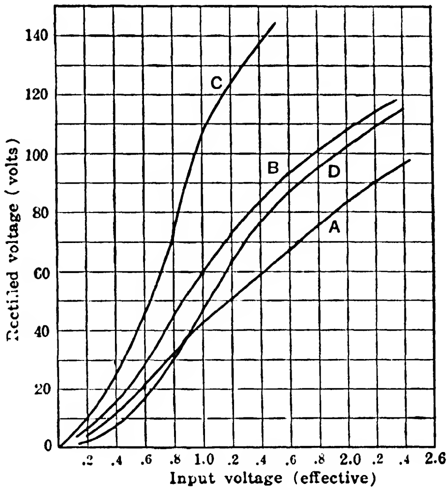
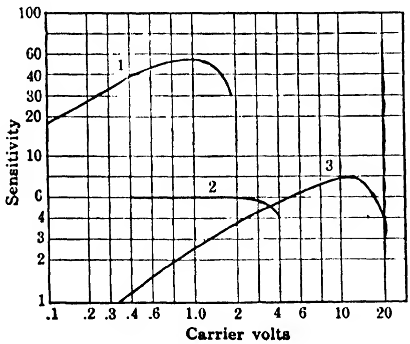


Fig. 79.—Action of '24 tetrode as plate circuit detector, with various conditions of voltage and external resistance used in the plate. Plate voltage for all curves was 300.

Curve A $E_{sg} = 67$ $R = 45{,}000$
Curve B $E_{sg} = 45$ $R = 90{,}000$
Curve C $E_{sg} = 22$ $R = 250{,}000$
Curve D $E_{sg} = 45$ $R = 110{,}000$

In Fig. 80 are shown the relative responses of the type '24 tube and type '27 for the two conditions stated. With grid rectification there is a very definite upper voltage limit, less than 4 volts for the '27 and less than 2 volts for the '24. If these limits are exceeded the response of the detector falls off very rapidly, and very bad distortion results.

If plate rectification is used (as is now customary) with high grid bias, much larger signals can be impressed, and still have the audio output proportional to the radio frequency input. Curve 3 of Fig. 80 shows this limit to be about 14 volts for a '27 tube with con-



Curve 1—Grid rectification, type '24 $R_g = 10^6$ $R_p = 90{,}000$ $E_{sg} = 40$ $E_p = 300$.

Curve 2—Grid rectification type '27. $R_g = 10_6$ $R_p = 25{,}000$ $E_p = 300$.

Curve 3 is for the type '27 triode using plate rectification. $E_c = -18$ $R_p = 200{,}000$ $E_p = 270$.

Sensitivity is A.F. volts across the resistance in the plate circuit per volt of R.F., 100% modulated, carrier impressed on grid.

Fig. 80.—Detector action of two tubes.

ditions as there given, whereas Fig. 71, p. 1041, shows it possible to get nearly linear response with as high as 30 volts on the detector grid.

Fig. 81 shows the response of a type 201A triode used as a grid bias detector, showing how the fundamental current, for various degrees of modulation of the r.f. input, increases with the amplitude of the r.f. carrier; Fig. 82 shows how the percentage distortion (mostly second harmonic) varies with input voltage, for 100 per cent modulated wave.

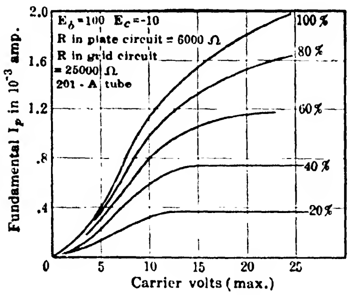


FIG. 81.—Action of 201A triode as detector, using plate rectification. Ordinates give amplitude of the alternating component of plate current, of same frequency as the modulation of carrier volts on the grid. Rectified current is practically proportional to degree of modulation.

Frequency Distortion.—When plate rectification is used nearly all the audio frequencies are acted upon alike, but when grid rectification is used the high frequencies are always rectified less efficiently than the lower ones. This action has been discussed in Chapter VI, pp. 554 and 562, and Fig. 119, p. 564, was given to show this action. Another set of results is given here, in Fig. 83. Curves *B*, *C*, *D*, and *E* show the rectifying action for different values of grid condenser and leak resistance. There is also given curve *A* which shows the rectifying action of the same tube, a type '27, when the plate-circuit external resistance is by-passed with a condenser too large.

The resistance used in the external plate circuit (for curve *A*) was 250,000 ohms, and across this there was a 150-$\mu\mu f$ condenser. Now the reactance of the 150-$\mu\mu f$ condenser at 10 kc. is only about 110,000 ohms, so it is evident that this condenser, small as it is, acted to materially cut down the plate-circuit impedance and so diminish the drop across it.

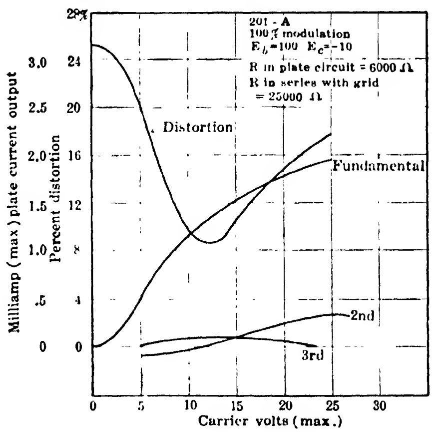


FIG. 82.—Components of alternating current in plate circuit showing highest per cent of distortion for the weak signals.

Effect of Grid Current.—In the audio-frequency portion of the amplifier the grids of the triodes should draw no current, in other words, the grids should never be allowed to swing appreciably positive. The distortion resulting therefrom is of a peculiar kind, and complex in its nature. Let us consider the ordinary case of a transformer-repeating amplifier.

As long as the grid stays negative, or swings only slightly positive it draws no current and so there is no internal impedance drop in the transformer secondary winding. The grid current, however, increases with the square of the grid potential as this goes positive and so the grid does not have to swing very far in the positive direction before the current becomes several microamperes, or even milliamperes, and this results in a distortion which adds even harmonics to the fundamental frequency, as well as to the higher frequencies during the positive part of the fundamental cycle. Furthermore, the upper frequencies are amplified less during the positive alternation of the fundamental than during the negative. These actions are well brought out in Figs. 84 to 88 of this chapter.

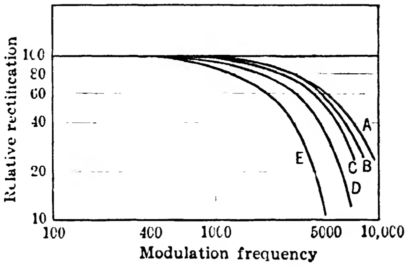


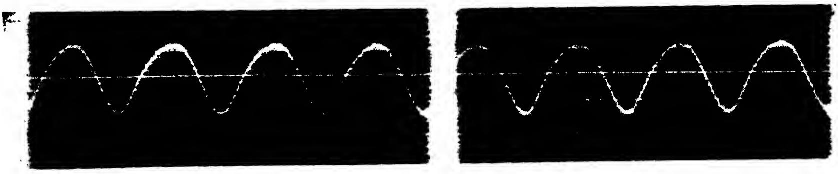
FIG. 83.—Frequency distortion of type '27 tube used as detector. Curve A—Plate rectification. Resistance in plate circuit was 250,000 ohms, shunted by a condenser of 150 $\mu\mu f$. Other curves are for grid rectification. Curve B-$R_g = 10^6$ $C_g = 50$ $\mu\mu f$. Curve C-$R_g = 2 \times 10^6$ $C_g = 50$ $\mu\mu f$. Curve D-$R_g = 10^6$ $C_g = 250$ $\mu\mu f$. Curve E-$R_g = 2 \times 10^6$ $C_g = 250$ $\mu\mu f$.

In the modern push-pull amplifiers for audio-frequency output (see p. 1006) the grids of the tubes do take large currents, but the resultant distortion, so far as even harmonics are concerned, are eliminated by the circuit action. One tube alone would generally give great distortion, whereas the pair give very little; this is shown by Figs. 29 and 30 of this chapter.

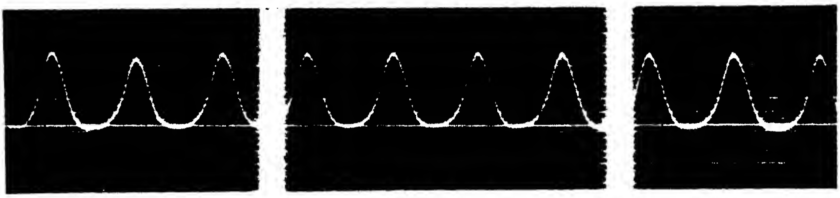
Characteristics of Transformers.—This feature of the amplifier has been thoroughly discussed in previous sections of this chapter and so needs no further analysis here.

Overloaded Output Tube.—This is probably one of the most prevalent causes today of the almost universally poor quality of loud-speaker reproduction. A triode having a full-load capacity (without distortion from overloading) of a few milliwatts is forced to deliver a hundred or more. It does this by being forced over the extremities of its plate current range but its response results in the peculiar well-known "rattle" of the average loud speaker. At the end of Chapter VI will be found a tabulation of the various output tubes showing how much power can be reasonably expected.

Loud Speaker.—There is no loud speaker available which responds equally well to all frequencies; all of them have some resonance character-



Grid voltage, second tube.
Scale, 0.294 volt per mm., a.c., max.



Grid current, second tube.
Scale, 0.771 microamp., per mm., a.c., max.

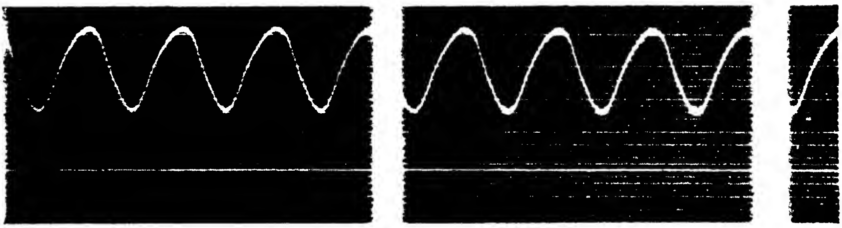
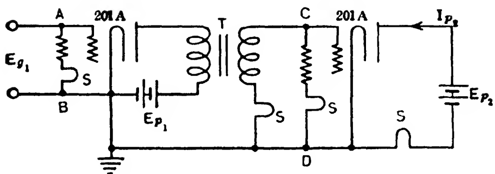


Plate current, second tube.
Scale, 0.215 milliamp., per mm., a.c., max.



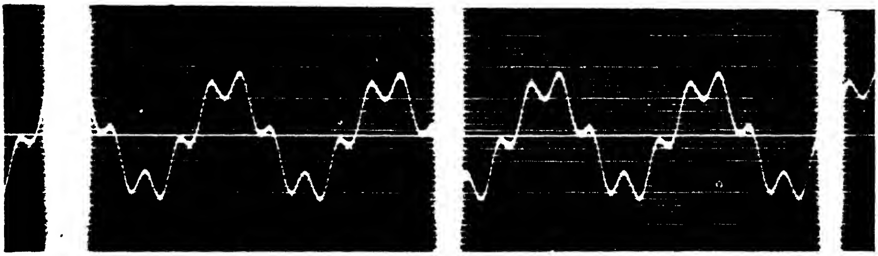
$E_{g_1} = 0.4$ volt a.c., at 100 cycles. $I_{P_2} = 2400$ microamp.
$E_{P_1} = 45$ volts. $R_{AB} = R_{CD} = 0.5$ megohm.
$E_{P_2} = 90$ volts. S = string galvanometer.
$E_{c_2} = 0.0$ volt. Transformer turn ratio = 5.

FIG. 84.—Showing distortion due to grid current flowing in the secondary of the transformer. Curves obtained by the string galvanometer. A pure sine wave was used for E_{g_1}.

istics which accentuate certain notes and repress others. In general the moving coil, or " dynamic," speaker seems to give a more faithful response than any of those using horns, or those of the double paper cone type, no matter how scientifically these may have been designed.

Oscillograms of Amplifier Performance.—The currents present in the amplifier circuits are so small that the Duddell oscillograph will not give any response; an Einthoven string galvanometer is the only convenient method of actually getting photographs of these minute currents. The author, with the help of Dr. A. Turner, succeeded in getting some interesting pictures of distortion, a few of which are given here.

The string galvanometer used had a sensitivity of about 1 microampere per millimeter deflection, and was shunted for measuring the



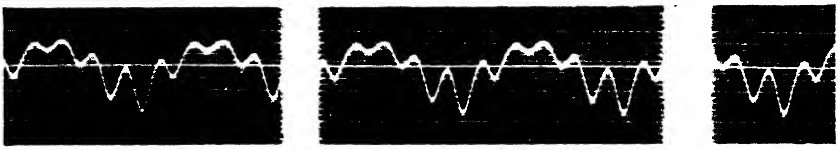
Impressed voltage.
$E_g = 0.35$ volt effective of 60 cycles.
$= 0.20$ volt effective of 300 cycles.

FIG. 85. A synthesized complex wave to be used as the "signal" voltage in the following oscillograms. The fifth harmonic was actually about 60 per cent of the fundamental but does not appear so in the wave form because the oscillograph did not respond as readily for 300 cycles as for 60 cycles.

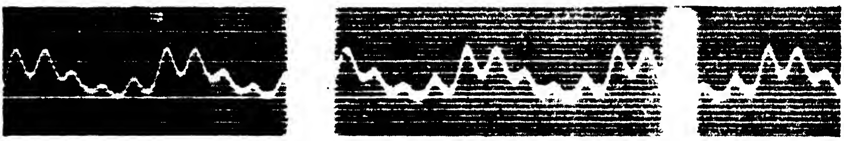
larger currents. Like all instruments of this class it responded better for the lower frequencies, as indicated in the films.

The absence of grid bias was first investigated, the signal being a pure sine wave of 60 cycles. In Fig. 84 are shown the various currents, and it is at once evident that even though a pure sine wave was impressed on the grid of the first tube, the secondary terminal voltage, impressed on the grid of the second tube, is by no means a pure sine wave. A second harmonic has been introduced by the effect of the grid current on secondary terminal voltage, and of course, this distortion carries through to the plate current.

A complex wave was built up, consisting of a 60-cycle wave with a fifth harmonic superposed and this complex wave, shown in Fig. 85, represents the signal. The actual signal had a greater percentage of fifth harmonic than shown by the oscillogram, as indicated by the calibration



Grid voltage, second tube.
Scale, 0.285 volt per mm., max., at 60 cycles.
0.895 volt per mm., max., at 300 cycles.



Grid current, second tube.
Scale, 1.282 microamps. per mm., max., at 60 cycles.
4.020 microamps. per mm., max., at 300 cycles.

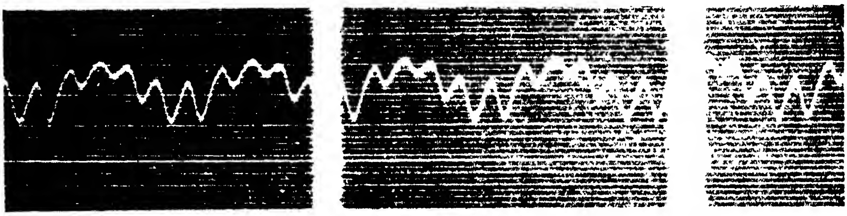
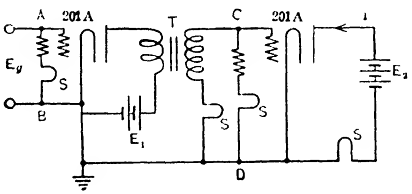


Plate current, second tube.
Scale, 0.238 milliamp. per mm., max., at 60 cycles.
0.745 milliamp. per mm., max., at 300 cycles.



$E_g = 0.35$ volt at 60 cycles $+ 0.20$ volt at 300 cycles. $I_2 = 2400$ microamps.
$E_1 = 45$ volts. $R_{AB} = R_{CD} = 0.5$ megohm.
$E_2 = 90$ volts. S = string galvanometer.
$E_{c_2} = 0.0$ volt. Transformer turn ratio = 5.

FIG. 86.—The complex wave of Fig. 85 used as signal in a transformer repeating amplifier, with no grid bias. The plate current (output of the amplifier) is far from the form of Fig. 85.

data given in Fig. 85. This calibration error of the galvanometer is, however, constant, so that it has little bearing on the interpretation of the oscillograms.

Fig. 86 shows the distortion introduced in a transformer-repeating amplifier when no grid bias is used. The grid of the second tube swings positive, draws current, and hence lowers very much the potential of the grid, compared to what it should be. The effect of this IZ drop in the transformer secondary is more marked at the more positive potentials of the grid and so we have the peculiar grid voltage shown in the upper oscillogram of Fig. 86. The second shows the grid taking a peculiar shaped current, this current being the cause of the distortion. Of course, the plate current follows the grid distortion and so we have the voltage of Fig. 85 giving the plate current of Fig. 86, by no means similar to the voltage of Fig. 85. There is another distortion due to the transformer which is better brought out in later figures.

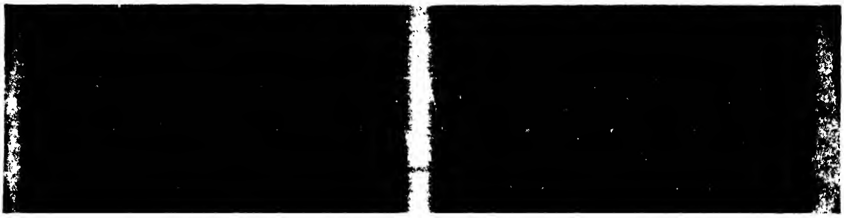
Figs. 87 and 88 show the effects of using different values of grid bias, that of Fig. 87 being not quite enough to bring the grid current to zero. In Fig. 88 there is no grid current but there become evident two other kinds of distortion. By comparing the top oscillogram with the signal voltage we see that the fifth harmonic has been amplified more than the fundamental, accentuating the ripples. This corroborates the curves of Fig. 24, p. 998.

In the plate current of Fig. 88 there appears still another type of distortion. The curvature of the E_g-I_p curve results in a greater amplification of the fifth harmonic during the positive alternation of the 60-cycle voltage than during the negative alternation. Furthermore, if the 60-cycle current be abstracted from the complex curve of I_p it will be found that a second harmonic has been introduced by the form of the E_g I_p curve.

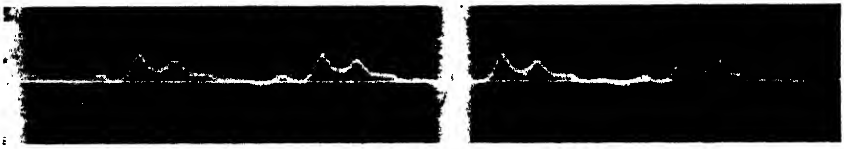
In Figs. 89 and 90 are shown the performance of a resistance-coupled amplifier. Fig. 89 with no grid bias, draws grid current and so depresses the grid potential of the second tube by an amount depending upon the signal strength. This will result in an intensity distortion of the signal, the louder parts not being amplified by the same factor as the weaker ones. Fig. 90 shows that a grid bias of 6 volts reduces the grid current to zero and so will eliminate the distortion just mentioned.

Both Figs. 89 and 90, compared with the signal voltage of Fig. 85, show how much more faithful is resistance amplification than transformer amplification.

Stability of Amplifiers—" Squealing."—A high gain amplifier, especially if of the inductance or transformer-repeating type, may produce in the telephone receivers audio-frequency sustained notes which are entirely independent of the incoming signals; this action is known as



Grid voltage, second tube.
Scale, 0.285 volt per mm., max., at 60 cycles.
0.895 volt per mm., max., at 300 cycles.



Grid current, second tube.
Scale, 1.282 microamps. per mm., max., at 60 cycles.
4.020 microamps. per mm., max., at 300 cycles.

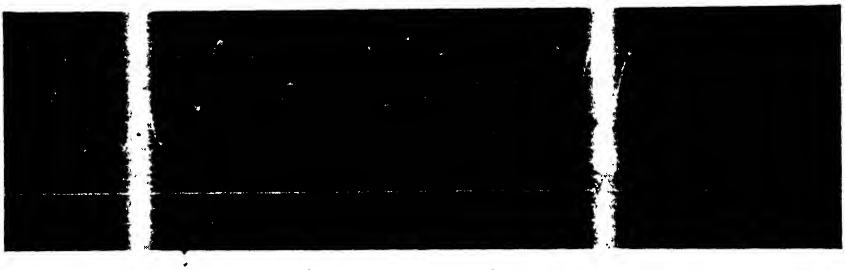
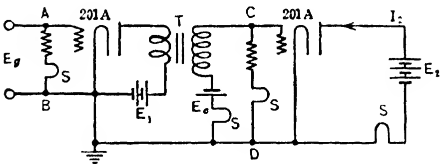
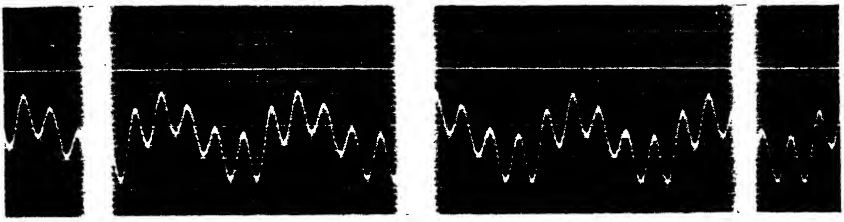


Plate current, second tube.
Scale, 0.238 milliamp., per mm., max., at 60 cycles.
0.745 milliamp., per mm., max., at 300 cycles.



$E_g = 0.35$ volt at 60 cycles and 0.20 volt at 300 cycles. $I_2 = 2400$ microamps.
$E_1 = 45$ volts. $R_{AB} = R_{CD} = 0.5$ megohm.
$E_2 = 115$ volts. S = string galvanometer
$E_{c_2} = 2.95$ volts. Transformer turn ratio = 5.

FIG. 87.—A grid bias used in the circuit of Fig. 86, decreased the grid current and somewhat improved the form of the output current.



Grid voltage, second tube.
Scale, 0.285 volt per mm., max., at 60 cycles.
0.895 volt per mm., max., at 300 cycles.

Grid current, second tube.
Scale, 1.282 microamps. per mm., max., at 60 cycles.
4.020 microamps. per mm., max., at 300 cycles.

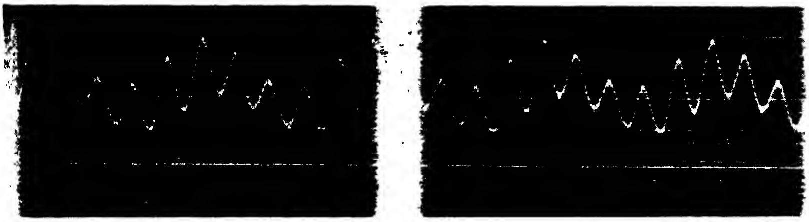
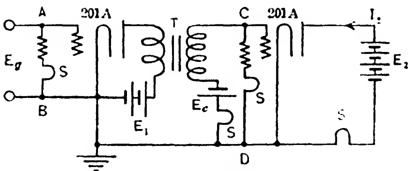
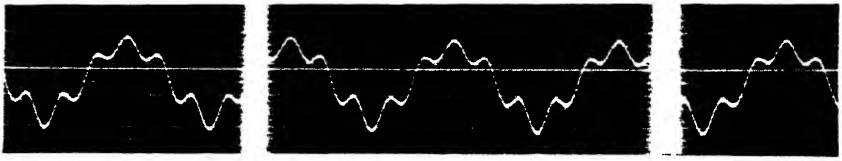


Plate current, second tube.
Scale, 0.238 milliamp. per mm., max., at 60 cycles.
0.745 milliamp. per mm., max., at 300 cycles.

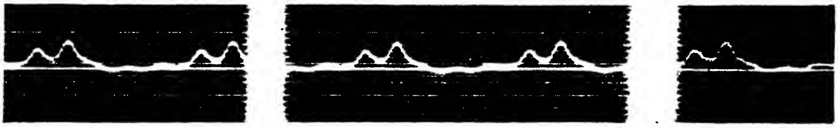


E_g = 0.35 volt at 60 cycles and 0.20 volt at 300 cycles. I_2 = 2400 microamps.
E_1 = 45 volts. $R_{AB} = R_{CD}$ = 0.5 megohm.
E_2 = 125 volts. S = string galvanometer.
E_{c_2} = 4.4 volts. Transformer turn ratio = 5.

FIG. 88.—By still further increasing the bias of the grid (over that of Fig. 87) the grid current of the amplifier tube is reduced to zero. There is still some distortion left due to the form of the I_p-E_g curve of the triode, and the characteristics of the repeating transformer. This repeated 300 cycles better than it did 60 cycles.



Grid voltage, second tube.
Scale, 0.253 volt per mm., max., at 60 cycles.
0.795 volt per mm., max., at 300 cycles.



Grid current, second tube.
Scale, 0.57 microamp. per mm., max., at 60 cycles.
1.79 microamps. per mm., max., at 300 cycles.

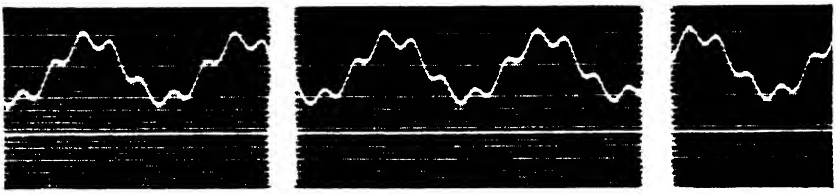
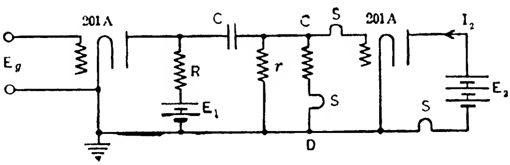


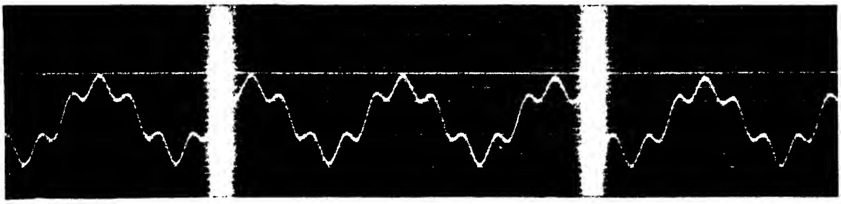
Plate current, second tube.
Scale, 0.238 milliamp. per mm., max., at 60 cycles.
0.745 milliamp. per mm., max., at 300 cycles.



$E_g = 0.35$ volt at 60 cycles, and 0.20 volt at 300 cycles.
$E_1 = 45$ volts.
$E_2 = 90$ volts.
$E_c = 0.0$ volts.
$I_2 = 1700$ microamps.

$R = 60{,}000$ ohms.
$r = 1$ megohm.
$C = 1$ microfarad.
$R_{CD} = 1$ megohm.
S = string galvanometer.

FIG. 89.—Action of a resistance-coupled amplifier with no grid bias. The arrangement draws some grid current.



Grid voltage, second tube.

Scale, 0.253 volt per mm., max., at 60 cycles.
0.795 volt per mm., max., at 300 cycles.

Grid current, second tube.

Scale, 0.57 microamp., per mm., max., at 60 cycles.
1.79 microamps., per mm., max., at 300 cycles.

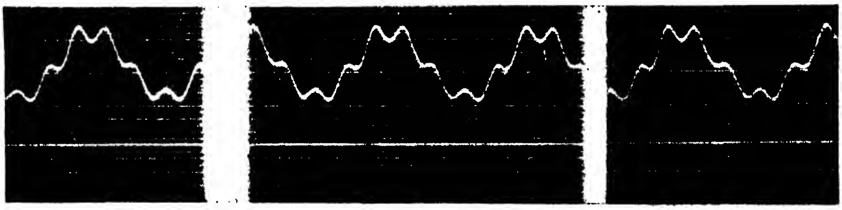
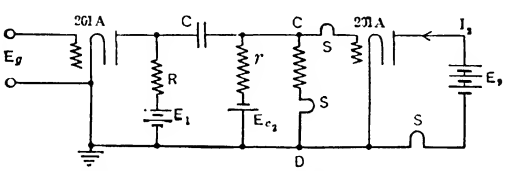


Plate current, second tube.

Scale, 0.238 milliamp. per mm., max., at 60 cycles.
0.745 milliamp. per mm., max., at 300 cycles.



$E_g = 0.35$ volt at 60 cycles, and 0.20 volt at 300 cycles.
$E_1 = 45$ volts.
$E_2 = 135$ volts.
$E_c = 6.0$ volts.
$I_2 = 1700$ microamps.

$R = 60{,}000$ ohms.
$r = 1$ megohm.
$C = 1$ microfarad.
$R_{CD} = 1$ megohm.
$S =$ string galvanometer.

FIG. 90.—The action of the circuit of Fig. 89 is improved by the addition of a grid bias. This plate current (output of amplifier) is quite similar to the "signal" voltage (Fig. 85).

"squealing," and is extremely objectionable and sometimes difficult to overcome. The squealing of an amplifier is generally due to the fact that the circuits of the various tubes are capable of oscillating, and may oscillate if the conditions are favorable; this applies to both high-frequency and low-frequency amplifiers. Thus assume, for the sake of clearness, that a single tube is connected by itself as it would be connected were it used in a low-frequency transformer-repeating amplifier and let Fig. 91 represent it. The coils of the repeating transformers T_1 and T_2 have a large amount of distributed capacity, and hence the circuit may be roughly represented by its equivalent of Fig. 92. It will be noted that L_1-C_1 is an oscillating circuit and the grid is connected across it; furthermore the circuit L_2-C_2 may have a high impedance to currents of the natural frequency of the circuit L_1-C_1. Any oscillations started in L_1-C_1 will produce a change in grid potential which will produce a change in plate current; the latter will cause a variation of plate potential to take place, in view of the impedance of L_2-C_2 being connected in series with the plate battery, and finally the variation of plate potential may, through the capacity of plate to filament and grid to filament, react back upon the grid, and may impress a higher voltage across the circuit L_1-C_1 than at first existed. Such a condition would, of course, be favorable to the maintenance of currents of the natural frequency of L_1-C_1.

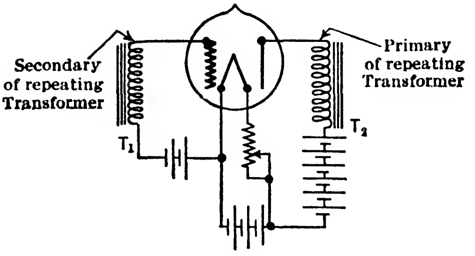


FIG. 91.—Circuit detail of a transformer-repeating amplifier; it may well be that, due to internal capacities of the coils (giving the circuit a natural period) and the coupling between plate and grid circuit inside the tube, self-sustained oscillations are set up in the circuit.

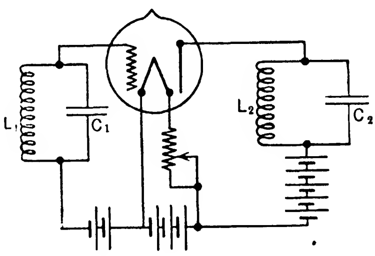


FIG. 92.—The transformer coils of Fig. 91, with their internal capacities give a circuit as shown here.

The tube may also oscillate at the natural frequency of L_2-C_2; whether it oscillates at the frequency of L_1-C_1 or L_2-C_2 depends entirely upon which of these frequencies gives correct phases of e.m.f.s and the smaller losses in the entire circuit. If the frequency at which the tube oscillates is audible, the currents produced thereby will be heard in the telephones connected in the plate circuit of the last tube of the amplifier, and will thus produce squealing. If the frequency is inaudible the telephone will give no indication of the presence of such currents, but they will nevertheless seriously interfere with the amplifying action of the tubes.

In the case of a high-frequency transformer-repeating amplifier the tube circuits may also oscillate, but they will do so at radio frequencies and will not be heard, but the efficiency of the amplifier as a whole will be reduced to practically zero by the presence of these interfering currents. On the other hand inductance-repeating high-frequency amplifiers, oscillating at radio frequency, sometimes product an audible tone in the telephone due to the fact that the grid condensers may, as a result of the oscillations, become so highly negative as to cause the plate current to become zero and thus stop the oscillations; after a time the electrons collected on the grid will leak off and the plate current will start flowing; but the oscillations will then again start in and again make the plate current zero, etc. This starting and stopping of the oscillations, with consequent pulsations in plate current, may take place at audio-frequency, in which case the amplifier will "squeal." This phenomenon is similar to the one fully discussed on p. 644 in connection with oscillating receiving tubes equipped with grid condensers.

In the discussion given above we have, for the sake of simplicity, considered the action taking place in each individual tube, which may be caused to oscillate due to the varying currents in the plate circuit of that one tube reacting back upon the oscillating circuit to which the grid and filament of the same tube are connected. Of course each tube may be caused to oscillate in the same manner at the same or a slightly different frequency from every other. What does happen, however, is that all tubes are subjected to one single frequency and the value of this frequency is the one at which it is "easiest" for the entire amplifier to oscillate, that is, the one frequency at which the losses in the whole amplifier (for a given strength of oscillation) are a minimum; of course if these losses are greater than can be supplied by the plate battery through the reactions of each plate upon each grid circuit the amplifier will fail to oscillate at that frequency or at any other frequency for which this condition prevails.

It may happen, however, that the output circuit of the amplifier is coupled, either magnetically or electrostatically, or both, to the input circuit, in which case the amplifier may oscillate, even if it would not otherwise do so. Thus, consider the three-stage transformer-repeating amplifier of Fig. 6, and assume that oscillatory currents start in the secondary of transformer T; these currents will be repeated and amplified from tube to tube; if now the plate circuit of the last tube is so related to the grid circuit of the first that the varying currents in the former can produce varying voltages in the latter, which are sufficiently large and in the right phase to increase and sustain the currents started in the secondary of transformer T, then the amplifier will oscillate. It will be easily understood that if there is coupling between the output and input circuit it is not a necessary condition, in order for the amplifier to oscillate, that the oscillations shall start in the grid circuit of the first tube, for, they may start in the grid circuit or even the plate circuit of any one of the tubes, including the last, and, in every case, the amplifying action of the apparatus may make it likely that oscillations will be sustained, even if the coupling between the output and input circuits is feeble.

Again, while in the preceding paragraphs we have assumed that the plate circuit of the last tube is coupled to the grid circuit of the first tube the amplifier may oscillate even if the plate circuit of an intermediate tube is coupled to the grid circuit of the first tube or, in general, it may oscillate if the plate circuit of any one tube is coupled to the grid circuit of any of the preceding tubes. For, as long as any currents started in the oscillatory circuit of any one tube are sustained by the reactions of the other tubes the amplifier as a whole may oscillate.

Remedies for Amplifier Squealing.—It must be stated at the outset that the more an amplifier amplifies the more likely it is to squeal; in other words, a silent amplifier is not necessarily better than one which shows tendency to self-oscillation; in fact, if a series of tubes connected in cascade show no tendency to squeal it is likely that the combination is so adjusted that the overall amplification is much lower than it should be. Even when all precautions against squealing have been taken it may be found upon testing the amplifier that it squeals most objectionably. In general the following points should be observed:

(*a*) An amplifier without any oscillatory circuits is not very likely to squeal. A resistance-repeating amplifier may be constructed practically without any oscillatory circuits, although it must be understood that even a short pair of wires form an oscillatory circuit, with a very high natural frequency to be sure, but nevertheless an oscillatory circuit. Hence even a resistance amplifier may oscillate at very high frequency and yet be " heard " if the grid condensers intermittently " block " the plate current. Resistance-repeating amplifiers (with an overall voltage amplification of about 25,000) have been constructed which do not squeal.

(*b*) Under no circumstances should the output and input circuits of an amplifier be coupled together even in the feeblest manner. It is best to use for both of these circuits short twisted leads and the output and input circuits should be kept as far apart as possible. The twisted leads should be "shielded" by enclosing them in a grounded flexible metallic casing. As a matter of fact it is advisable that all plate-circuit leads be kept from being coupled to grid-circuit leads of previous tubes; hence the leads inside of the amplifier box should be run with this very important point in view.

(*c*) Each tube and its holder may be placed in a shielded chamber, the surfaces of which are covered with copper connected to ground; this prevents any electrostatic or magnetic field from one tube from appreciably affecting the adjacent tubes or, in other words, it prevents coupling between adjacent tubes, since the energy contained in any varying fields produced by one tube is absorbed by the currents created in the surrounding copper. This precaution should always be taken in the case of high-frequency amplifiers especially.

(*d*) The balancing condensers of neutralized amplifiers should be carefully adjusted, with the tubes that are going to be used, in their proper sockets and normally operating.

(*e*) In special cases separate plate batteries and filament batteries should be used for each tube, for in this manner a means of coupling between the tubes is done away with. However, separate batteries for all tubes add so much to the weight, size, and cost of an amplifier as to make the arrangement impossible for any but special laboratory work. In any case both of these batteries should be of as low a resistance as feasible and a large condenser placed across each of them to constitute a low impedance path for high-frequency currents.

(*f*) All leads should be rigidly held in their proper places and all connections be well soldered.

(*g*) The repeating transformers, both radio and audio, should have their winding connected in proper polarity. An amplifier may generate violent oscillations with a certain connection and be quiet when the terminals of the secondary (or possibly primary) winding are interchanged. That is, with transformer secondary terminal No. 1 connected to grid, and No. 2 connected to the biasing battery, oscillations occur; by connecting terminal No. 2 to the grid and No. 1 to the biasing battery the oscillations are eliminated.

These effects are caused by capacitive couplings between the transformer windings.

(*h*) Wherever they may be put in without interfering with the amplifying action of the set, by-pass condensers of perhaps 0.1 to 1 microfarad should connect critical points of the circuit to ground.

Even when all these precautions have been taken it may be that an amplifier known to be correctly built, and of previous good behavior gives loud " sputtering " noises in the telephones, even when the input circuit is short-circuited. This may be due to a " bad " tube somewhere in the amplifier (to be discussed in the next section) or it may be due to either the A or B batteries if such are used. A storage battery is frequently used for filament heating, so it is evident that this battery has a low resistance, but it will be found that if a good amplifier (high amplification) is used with an A battery nearly discharged (say lower than 1.8

volts for a lead cell) all sorts of odd noises may be heard in the amplifier, whereas if a normally charged A battery is substituted the amplifier is quiet.

The same remark holds true regarding the B battery to an even greater degree; the small dry cells generally used for the plate battery develop a very high variable resistance towards the end of their life, and if there is one such " worn-out " cell in the battery it will result in very bad noises in the amplifier. A test of the cells with a *low-resistance* voltmeter will at once show up the defective cell.

Howling Due to Sound Wave Coupling.—There is another interesting condition which may result in an amplifier singing at audible frequency, especially likely to occur when the loud speaker is built into the same case as the audio-frequency amplifier. The sound wave from the loud speaker may be able to vibrate the elements of one of the triodes (generally the detector or first audio-amplifying tube) sufficiently to sustain the oscillations. This is quite analogous to the familiar " squealing " telephone receiver; if the receiver is taken from its hook and the diaphragm is held quite close to the transmitter it will sing with a characteristic audible note.

In general, this singing trouble can be eliminated by any expedient which makes it more difficult for sound waves to get from the loud speaker to the vibrating triode. Thus spring suspension of the triode sockets, and mounting the horn on soft felt so that the sound waves have to travel through the felt before getting to the panel on which the tubes are mounted will generally remedy the trouble. Very infrequently, if the amplification is high the sound waves traveling from the loud speaker through air will shake the tube sufficiently to maintain the singing condition. In this case it is advisable to change the detector tube, or first audio-amplifier tube, for another. The elements of the triode are probably mounted in an insecure fashion, making it easy for them to vibrate. If changing tubes does not remedy the trouble the loud speaker must be mounted farther away from the amplifier to decrease the intensity of the sound waves striking the vibrating triode.

Tube Noises.—Another feature which causes considerable difficulty in the operation of an amplifier is the " noise " produced by the tubes. The reader will realize that any slight change in the currents flowing in the plate or filament circuit of an amplifier tube, especially if it be one of the first tubes, may be so amplified as to finally produce a very large change in the plate current of the last tube, and hence a loud click in the phones. Sometimes these clicks are frequent and almost deafening as compared with the signals, hence very objectionable. As a matter of fact these noises form one of the limitations of amplifiers in so far as the number of stages is concerned, since it is almost impossible to prevent minute changes of currents in the first tubes, which, if repeated and amplified through a large number of stages, may finally "swamp" the legitimate signals.[1]

These minute changes of current in the tube circuits may take place due to several causes, the most common of which are:

(1) Sudden slight changes in the electromotive forces of the various batteries (discussed in previous section).

(2) Mechanical vibration of the elements of the tubes.

(3) A slight amount of gas causing ionization, or, what is more difficult to overcome, actual irregularities in the rate of emission of electrons from the filaments. This is probably due to surface impurities of the hot filament.

(4) Voltage due to thermal agitation of electricity.

It is evident that mechanical vibrations of the tube elements will vary the distance between the grid and filament and, of course, the plate current will change accordingly; the same is true of any changes in the distance between plate and filament and plate and grid. Hence the elements should be firmly supported. Of course, no matter how firmly supported, they may always be made to vibrate, though imperceptibly, and yet enough to be detected by the amplifier; hence the care must be exercised in supporting the elements. The importance of this point was not at first fully realized, and tubes were used for aeroplane work the elements of which were not sufficiently well supported, with the result that an amplifier consisting of these tubes became practically useless. The matter of supporting the tubes themselves is extremely important in this connection and should be given the greatest attention. Amplifier tubes are generally supported on thick pieces of soft spongy rubber or else on light springs; the point to strive after is to obtain a support such that if the tube as a whole is caused to vibrate it will do so at a very low, inaudible, frequency, and, furthermore, it will not be able to communicate the vibrations to the elements of the tubes, the natural frequency of which is very high in view of the rigid suspension of the elements.

It will be realized that the behavior of the first tube of a multi-stage amplifier must be extremely regular if it is to produce inappreciable noises in the telephones at the output end. Assuming an amplifier which multiplies the input voltage by 10^4 having a resistance (or reactance) in the first plate circuit of 50,000 ohms and assuming that a voltage of 0.02 on the grid of the last tube will give an audible signal in the telephones, it may be seen that a change in current in the plate circuit of the first tube

[1] Hull states the upper limit of amplification is about 2×10^6 in voltage; the actual irregularities in electron emission make greater amplification useless.

of only 10^{-10} ampere will produce an audible noise. This might be a variation in the plate current of the first tube of only one part in ten million! But if we try to conceive of the surface conditions of the hot metal from which the electrons are being boiled out, it seems impossible that the emission of electrons should be steady enough to eliminate such a slight irregularity. In fact it seems that with present tubes a much larger variation of the plate current is continually occurring.

Amplifier noise due to irregular effects in the tubes themselves has been well analyzed by Llewellyn.[1] After noting that rundown batteries, poor connections, microphonic action of contacts, and vibration, all give noise the remedy for which is apparent, he classifies the remaining noises as due to shot effect, thermal agitation, and ionization effects.

Shot Effect.—This effect (discovered by Schottky) is due to the irregular rate of electron emission from the filament; just as steam comes from a tea kettle in puffs, due to bursting bubbles rising from the bottom, etc., so the electrons break through the filament surface at irregular rates and so produce an irregularity in the plate current of the first tube (as well as the others) which, being amplified by all the succeeding tubes, gives a loud noise in the output. This noise sounds like that of static, coming from a disturbance a long way off ("long range static" as it is sometimes called) giving a continual hiss.

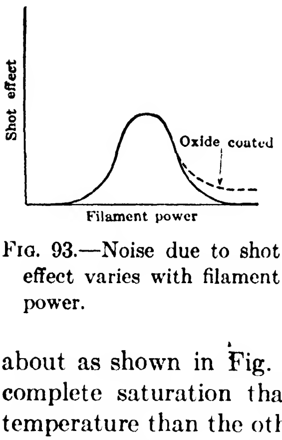


FIG. 93.—Noise due to shot effect varies with filament power.

Llewellyn shows that when the filament is heated to such an extent that space charge limits the plate current, instead of emission, the noise of the shot effect is practically zero; the noise varies with temperature of filament about as shown in Fig. 93. The pure tungsten filament, giving more complete saturation than the oxide-coated one, has less noise at high temperature than the other.

Thermal Noise.—This effect[2] is due to the heterogeneous motion of ions and electrons in a conductor due to their temperature energy. It gives a voltage of haphazard frequency distribution, indicated by the relation

$$V^2 = kT\int_0^\infty R\,d\omega \qquad (26)$$

The square of the voltage varies directly with the resistance of the conductor R, and with the absolute temperature T, and of course with the width of the frequency band through which the effect is integrated. The value

[1] I.R.E., Feb., 1930, p. 243.

[2] Johnson, Phys. Rev., Vol. 32, 1928, p. 110; and Fry, same volume, p. 97.

integrated through the audible spectrum at room temperature, is given by the relation

$$V^2 = 0.8 \times 10^{-16} R. \qquad (27)$$

Thus if a grid leak of 1 megohm is used between grid and cathode V is approximately 9 microvolts. If the amplifier is made somewhat selective, so that only part of the audible spectrum is amplified the effective thermal voltage will be correspondingly less.

The internal plate circuit of a tube is a resistance of reasonably high value, and there will be a thermal agitation voltage generated in this part of the amplifier. It appears that the proper temperature to designate for this resistance is that of the cathode; unless the grid is practically short-circuited, the plate circuit of a triode gives negligible thermal noise, because of the extra stage of amplification to which the grid noise is subjected.

Llewellyn tested for thermal noise by using a high gain amplifier employing screen-grid tubes, having a band width of 4 kc.; its over-all voltage gain was 10^6. Necessarily it was carefully shielded, and all tests using it were carried out in a shielded room.

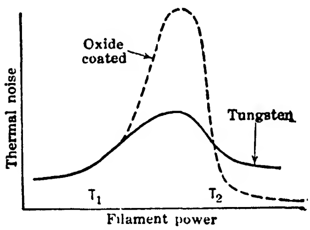


FIG. 94.—Noise due to thermal agitation varies with filament power.

Arranging the various tubes he used so that the grid circuit of each produced imperceptible noise, he found the relation between filament power and thermal noise, for tungsten filaments and oxide-coated filaments. Two of these curves are shown in Fig. 94. In the region between T_1 and T_2 the shot effect gives most of the noise and the oxide-coated filament gives the greater noise; above T_2 the space charge limitation of plate current eliminated the shot effect and the residual noise is due to thermal agitation voltage of the plate circuit. The tungsten filament, being at much higher temperature than the oxide-coated one, naturally gave the higher noise level.

Noise Due to Ionization.—Even with the best evacuation possible, there is much gas left in a vacuum tube, and the electrons on their way from the cathode to the plate will ionize some of these molecules. Every disturbance of this sort, affecting the distribution and action of the space charge between cathode and plate, will result in fluctuation of the plate current which produces noise.

Some of the positively charged ions will strike the grid, if this is held negative, and so affect the action of this electrode. Llewellyn found that those triodes which operate with a grid most highly negative, give the most noise.

In Fig. 95 is shown the plot of noise vs. change in plate current as filament current is changed. Each point of the curve represents the potential and filament conditions recommended by the manufacturer; evidently when I_p changes much with I_f, the space charge is not completely limiting the plate current, so that irregularities in filament emission may give noise. Evidently some of the tubes have too low a value of recommended filament current. However, as long as these tubes are used in an amplifier which has reasonably low gain in the a.f. end, the effect of this noise will probably be hidden under the static noise which comes in by way of the antenna.

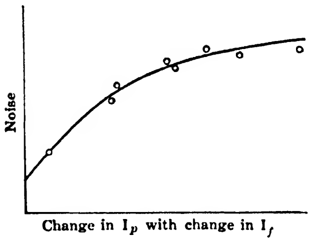


FIG. 95.—Variation of noise with filament power; it should normally be independent of this variable.

We may therefore conclude that if the input circuit of an amplifier

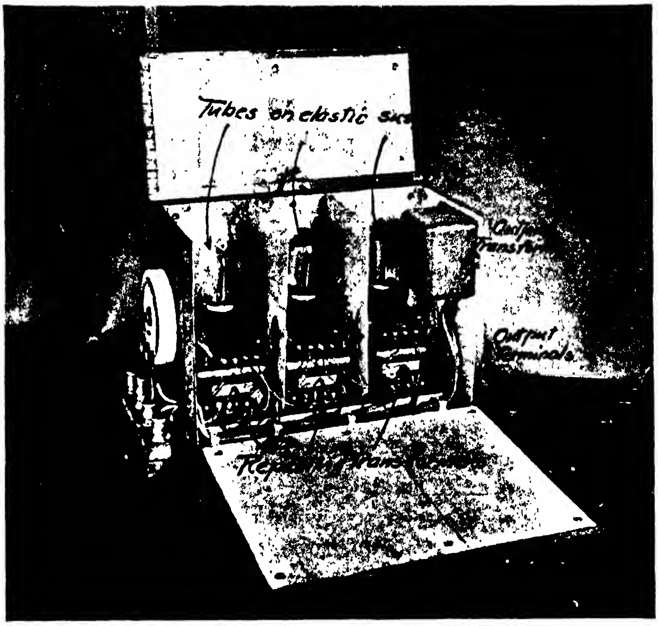


FIG. 96.—A compact transformer-repeating audio-frequency amplifier; the tubes are in spring suspensions, each in its separate metallic compartment. It has a voltage amplification of about 3000.

is not subject to interfering signals of too great an intensity, an amplifier may usefully be employed with a voltage amplification of between 10^4

and 10^5, if quiet tubes are used in the first stages; with present tubes more than this (or as much as this in most cases) is not worth while; the signal may be made louder by using perhaps one or two more tubes, but in general it is no more readable.

Arrangement of Apparatus in Amplifiers.—In Fig. 96 is shown a three-stage audio-frequency amplifier using transformer repeating. The tubes used have $\mu=7$ and the transformer ratio is about 3.5; in series with the negative lead from each filament is a small piece of resistance wire, so that the grids are held at a negative potential, with respect to each filament. The filaments are in parallel, the currents being controlled by

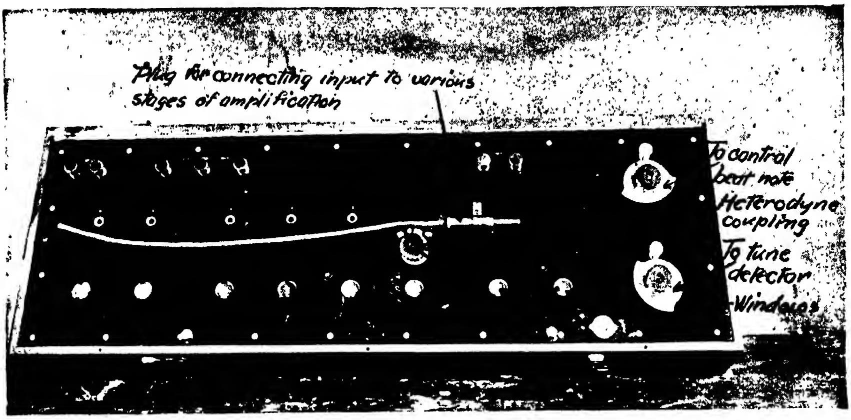


FIG. 97.—External appearance of a very well-designed amplifier for a frequency of 50,000 cycles; the input circuit can be connected to various stages by means of the plug and flexible cord. The total amplification of this instrument is about 50,000 but it can seldom be used efficiently with an amplification greater than about 5000 because of tube noises.

a common rheostat; the battery used in the plate circuit (the same battery serves all tubes) should be about 90 volts. An over-all voltage amplification of about 3000 is obtained but there is generally an audible tube noise present, making it useless for reading very weak signals.

In Figs. 97 and 98 is shown a very carefully designed amplifier having its best performance for a signal of 6000 meters wave length. Inductance repeating is used, the coils being toroids with iron-dust cores; they have a reactance of about 50,000 ohms. The tubes used have μ equal to 35 and use a plate potential of 130 volts. The total voltage amplification possible, without squealing, is about 50,000, but its useful amplification for very weak signals, is only about 5000.

This question of *useful* amplification is seldom mentioned in texts, but is really very important. It may be that two amplifiers are compared in the laboratory and it is found that one gives a voltage amplification ten times as much as the other. It may be that this comparative figure checks when different tests are made so that there is no doubt regarding its accuracy. It might be then assumed that if a signal giving a certain current in the antenna is just readable with amplifier A (the poorer one) that when amplifier B is used the signal would be readable if the antenna current were decreased to one-tenth its former value. It

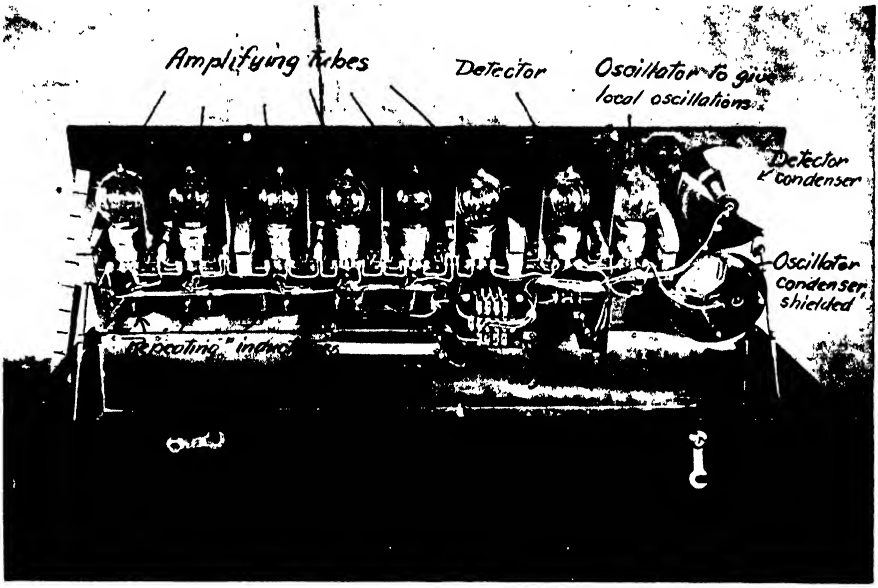


Fig. 98.—Arrangement of the apparatus of the amplifier shown in Fig. 97.

will probably be found, however, that when amplifier B is used an antenna current about one-half that used with amplifier A is the least audible signal, instead of one-tenth, as is naturally assumed. The reason for this is the " background " of noise (from tubes and other sources) present to a greater extent with B than with A. And the presence of the noisy background requires a much stronger signal in the phones when using amplifier B than is required when A is used.

INDEX

D