THEORY AND APPLICATIONS OF ELECTRON TUBES

This book is produced in full compliance with the government's regulations for conserving paper and other essential materials.

THEORY AND APPLICATIONS OF ELECTRON TUBES

BY

HERBERT J. REICH, Ph.D.
Professor of Electrical Engineering, University of Illinois; on leave to the Radio Research Laboratory, Harvard University

Second Edition

McGRAW-HILL BOOK COMPANY, Inc.
NEW YORK AND LONDON
1944

THEORY AND APPLICATIONS OF ELECTRON TUBES

PRINTED IN THE UNITED STATES OF AMERICA

THE MAPLE PRESS COMPANY, YORK, PA.

This book is dedicated
to the many scientists and engineers
upon whose work it is based

PREFACE TO THE SECOND EDITION

The preparation of the second edition of "Theory and Applications of Electron Tubes" was prompted by two principal considerations. The first, and probably more important, of these was the necessity of bringing the book up to date as regards the principal new developments of the past five years in the field which it covers. Although the press of other work connected with the defense and war efforts has made it impossible for the author to make an exhaustive search of the literature of this period, he believes that the most important subjects have been adequately treated.

The second factor that led to the preparation of the new edition was the desirability of incorporating improvements in presentation that have been suggested or that have suggested themselves during five years of use of the book in college courses. Much of the material has been rearranged, and the chapter on Modulation and Detection and that on Oscillators have been to a considerable extent rewritten. A few changes in symbols or definitions have appeared to be desirable. In response to many requests, problem answers are included in the new edition.

The author of a book in a field that develops as rapidly as that covered by this book is faced with the difficulty of keeping the book up to date without increasing the size unduly. Probably the best judges of what material should be eliminated or added are the instructors in the courses in which the book is used as a text. The author will welcome suggestions and criticisms.

HERBERT J. REICH.

CAMBRIDGE, MASS.,
August, 1944.

PREFACE TO THE FIRST EDITION

Electron tubes, which have made possible the rapid development of radio to its present state of refinement, have been assuming an increasing importance in power control and transmission, in manufacturing, in the home, and in the various branches of engineering and scientific research. The rapidly growing field of application of electronic devices has necessitated the addition of courses in theoretical and applied electronics to engineering and scientific curriculums. The need for a single book to assemble and coordinate our present knowledge of the theory and application of electron tubes led to the writing of this book.

The book is intended to give the student a sufficiently thorough grounding in the fundamental principles of electron tubes and associated circuits to enable him to apply electron tubes to the solution of new problems. The author has not attempted to discuss all applications of tubes to special problems but rather to cover basic principles and typical applications. Since it was not his purpose to write a treatise on the subject of applications of electron tubes, Class C amplification and the design of radio transmitters and receivers, which are adequately treated in books on radio engineering, have not been taken up. The basic principles that are presented, however, are applicable to radio engineering problems, as well as to industrial electronics, power control, electrical measurements, and other fields of use of tubes. Although written primarily as a text for college students, it is hoped that it will also prove to be of value to practicing engineers as a reference book.

The book is based upon mimeographed notes that have been used in the author's courses on electron tubes during the past five years. These notes have been kept up to date and have been revised as use in the classroom has indicated where improvement could be made.

A problem encountered in the preparation of the manuscript was the choice of symbols. In the main, the symbols used are those which have been standardized by the Standards Committee of the Institute of Radio Engineers. Although the use of the symbols e and E for the voltage of tube electrodes is in agreement with the practice of most writers on the subject of electron tubes in the United States during the past twenty-five years, it is not in agreement with the symbols standardized by the American Institute of Electrical Engineers nor with those used in England. Since the basic symbols that are used in this book have already been standardized by the Institute of Radio Engineers, the author feels that it would be a mistake to set up new symbols now. Because of the very large number of symbols that must be used in the analysis of tubes, difficulties are invariably encountered, regardless of the system of nomenclature that is adopted.

The series expansion for electrode currents has been made the basis of the analysis of the operation of high-vacuum tubes and associated circuits. In order to justify the use of the series expansion and several other very useful equations, the outlines of their derivations are included. The student will not be seriously handicapped by the omission of these derivations. Because the author believes that a thorough understanding of the principles of detection and modulation is of great value in the study of distortion in amplifiers, the chapter on detection and modulation precedes those on amplification. The arrangement of subject matter is such, however, that little difficulty will be experienced by the student if this chapter is studied after those on amplifiers. Equivalent-circuit and graphical methods of analysis are stressed throughout the book.

H. J. REICH.

URBANA, ILL.,
November, 1938.

CONTENTS

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

CHAPTER 12

CHAPTER 15

APPENDIX

THEORY AND APPLICATIONS OF ELECTRON TUBES

CHAPTER 1

PHYSICAL CONCEPTS

An *electron tube* is a device consisting of a number of electrodes contained within a totally or partly evacuated enclosure. The usefulness of such a device arises from its capacity to pass current, the magnitude of which may be controlled by the voltages of the electrodes. As suggested by the term *electron tube,* the operation of all types of electron tubes is dependent upon the movement of *electrons* within the tubes. The electron, which is the smallest known particle, has a mass of 9.03×10^{-28} g and a negative charge $\mathcal{E}$ of 16×10^{-20} coulomb or 4.8×10^{-10} e.s.u. The operation of many types of electron tubes also depends upon the separation and motion of other elementary particles of which matter is composed. For this reason a brief discussion of the fundamental processes governing the behavior of these elementary particles is of value in the study of electron tubes and their applications.

1-1. Excitation, Ionization, and Radiation.—The mass of the electron is greatly exceeded by that of an atom, the lightest atom, hydrogen, having a mass approximately eighteen hundred times that of the electron. The major portion of the mass of an atom is accounted for by the *nucleus,* a stable assemblage of charged and neutral particles having a net positive charge equal to the atomic number of the element. The atom also contains relatively loosely bound electrons, normally equal in number to its atomic number. The normal atom is therefore neutral.

Experiments show that, in addition to possessing kinetic energy, an atom is capable of absorbing energy internally. The internal energy appears to be associated with the configuration of the particles of which the atom is composed. The internal energy can be altered only in discrete quantities, called *quanta,* and hence the atom can exist only in definite stable *states,* which are characterized by the internal energy content. Under ordinary conditions an atom is most likely to be in that state in which the internal energy is a minimum, known as the *normal state.* If the internal energy of the atom exceeds that of its normal state, it is said to be *excited.* Excitation may be caused in a number of ways, among which is collision of the atom with rapidly moving positive or negative particles, which may give up some or all of their kinetic energy to the atom during the collision. A limiting case of excitation is *ionization,* in which the energy absorbed by the atom is sufficient to allow a loosely bound electron to leave the atom against the electrostatic force which tends to hold it within the atom. An atom that has lost one or more electrons is said to be *ionized* and is one type of *positive ion.* It is possible for excitation or ionization to take place in successive steps by the absorption of two or more quanta of energy.

The return of an excited atom to a state of lower energy content is usually accompanied by electromagnetic radiation. Since the energy of the atom can have only discrete values, the radiated electromagnetic energy corresponding to a change from one given energy state to another of lower energy is always associated with the releasing of a definite quantum of electromagnetic energy, called a *photon.* The frequency of the radiated energy is determined by the relation $W_1 - W_2 = h\nu$, in which W_1 and W_2 are the values of the internal energy of the atom in the initial and final states; h is a universal constant called *Planck's constant,* 6.55×10^{-27} erg-sec; and ν is the frequency of the radiated energy in cycles per second. From the form of this relation it is seen that the quantity $h\nu$ is the energy of the emitted photon. Each line of the emission spectrum of an element represents the transition of atoms of the element from some energy state to another of lower internal energy.

In interacting with atoms and molecules, photons exhibit some of the characteristics of material particles, a photon behaving as though it were a bundle of energy. The collision of an atom with a photon whose energy is equal to the required change of internal energy may result in excitation of the atom. If the energy of the colliding photon is equal to or greater than that necessary to remove an electron from the atom, the collision may result in ionization of the atom.

A number of attempts have been made to give a picture of the structure of atoms that will account for the phenomena of excitation, ionization, and radiation. The most successful of these, proposed by Bohr, is based upon the assumption that one or more electrons move about the central nucleus of an atom in a manner similar to the motion of the planets about the sun. To account for the observed definite values of internal energy, it was assumed that the electrons can move only in certain orbits and that the internal energy of the atom is increased when one or more electrons are abruptly displaced from given orbits to others at a greater distance from the nucleus. Radiation was assumed to result when one or more electrons jump from given orbits to others nearer to the nucleus. Any such picture is valuable principally in its ability to explain observed phenomena and to predict others. The complexity of atoms containing more than two orbital electrons limited the usefulness of the Bohr picture of the atom to the hydrogen and helium atoms.

The phenomena of excitation, ionization, and radiation are also observed in molecules. Because of the greater complexity of molecules, and the fact that their internal energy is partly associated with the vibrational and rotational motion of the atoms of which they are composed, a molecule has many more stable states than an atom of the same element.

1-2. Electron Volt.—Since the energy that a charged particle acquires in free space when accelerated by an electric field is equal to the product of the charge by the difference of potential between the initial and final positions, the difference in potential may be used as a measure of the gain in kinetic energy. Any quantity of energy may, in fact, be expressed in "electron volts." An *electron volt* is the amount of energy gained by an electron when accelerated in free space through a difference of potential of 1 volt. It is often convenient to express in electron volts the energy required to ionize or excite atoms or molecules.

1-3. Excitation and Ionization Potentials.—An *excitation potential* is the energy, expressed in electron volts, that must be given to an atom or molecule in order to cause a transition from a given state to one of higher internal energy. An *ionization potential* is the least energy, expressed in electron volts, that must be supplied to a normal or an ionized atom or molecule in order to remove an electron from the atom or molecule. Inasmuch as all atoms but hydrogen contain more than one electron, an atom or molecule may, in general, have more than one ionization potential. The first ionization potential applies to the removal of an electron from a normal atom or molecule; the second ionization potential applies to the removal of a second electron from an atom or molecule that has already lost one electron; etc. A less likely type of ionization is the simultaneous removal of two or more electrons. Table 1-I lists the first ionization potentials of some of the elements that are used in electron tubes. It should be noted that when ionization or excitation potentials are expressed in electron volts they indicate the minimum voltage that must be applied between two electrodes in order to cause ionization as the result of acceleration of electrons or other singly charged particles by the resulting field between the electrodes.

TABLE 1-I.—FIRST IONIZATION POTENTIALS IN ELECTRON VOLTS

Argon	15.69	Nitrogen	14.48	Sodium	5.12
Neon	21.47	Carbon dioxide	14.4	Rubidium	4.16
Helium	24.46	Mercury	10.38	Cesium	3.87
Hydrogen	13.53	Lithium	5.37	Magnesium	7.61
Oxygen	13.55	Potassium	4.32	Barium	5.19

1-4. Ionization.—The positively charged mass resulting from the removal of one or more electrons from an atom is only one of many types of ions. In general, an *ion* is an elementary particle of matter or a small group of such particles having a net positive or negative charge. Atoms or molecules that have lost one or more electrons, or that have picked up one or more extra electrons, and simple or complex groups of a number of atoms or molecules bearing excess positive or negative charge, are special examples of ions. This definition of an ion also includes such relatively simple particles as the electron and other elementary charged particles of which atomic nuclei are composed. The process of ionization, broadly defined, is the production of ions in gases, liquids, or solids. It may result from a number of causes, among which are

1. Collision of atoms or molecules with
   a. Electrons.
   b. Positive or negative ions of atomic or molecular mass.
   c. Excited atoms or molecules.
2. Collision of atoms or molecules with photons (photoelectric effect).
3. Cosmic radiation.
4. High temperatures in gases or vapors.
5. Chemical action.

1-5. Ionization by Moving Electrons.—One of the most important causes of ionization in electron tubes is the collision of rapidly moving electrons with atoms or molecules. In order that a single moving electron may ionize an atom or molecule it is necessary that the kinetic energy of the electron be at least equal to the first ionization potential of the atom or molecule. It is sometimes observed, however, that ions appear in a gas or vapor when the bombarding electrons have energy corresponding to the first excitation potential. The explanation of this is that the atom may be ionized in steps, each successive impact by an electron supplying sufficient energy to cause a transition to a state of higher energy. Thus, although the first ionization potential of mercury is 10.38 electron volts, ionization of mercury vapor by moving electrons begins when the colliding electrons have energy corresponding to the lowest excitation potential, 4.68 electron volts. Other more complicated processes may produce similar results. Precise measurements show that the voltages at which moving electrons begin to ionize a gas or vapor are very sharply defined.

1-6. Ionization by Positive Ions.—Ionization by positive ions is a more complicated phenomenon than ionization by electrons. One reason for this is that a positive ion which strikes a neutral atom or molecule is able to surrender not only its kinetic energy but also some or all of its own energy of ionization, thus reverting to a state of lower internal energy. A collision in which an ionized or excited atom or molecule transfers all or part of its energy of excitation to another atom or molecule is known as a *collision of the second kind.* The transferred energy may be used to excite or ionize the unexcited particle or may be converted into kinetic energy of one or both particles. In a mixture of gases a bombarding positive ion of one of the gases may ionize a neutral atom or molecule of the other, the difference in energies of ionization being supplied by or added to the kinetic energies of the two particles. Because of conservation of momentum, the large mass of the positive ion also complicates the process of ionization by positive ions. Ionization by bombardment of positive ions requires higher accelerating potentials than by electrons, and the potentials at which ionization begins are not sharply defined.

1-7. Amount of Ionization.—The ionization potential is a measure only of the minimum kinetic energy, below which a moving ion cannot ionize a normal atom or molecule by a single collision. It does not follow that every electron that acquires this amount of energy will necessarily ionize a gas through which it moves. The likelihood of ionization, which differs for different gases, is a function of the energy of the bombarding particles. The amount of ionization produced in a gas or vapor by charges that are accelerated by electric fields in the gas may be most conveniently specified by the *ionization coefficient.* The ionization coefficient is defined as the number of ionizing collisions made by an ion per centimeter of advance through the gas. It differs for different gases and for different types of ions and is a function of the gas pressure and the electric field strength. The ionization factor varies with electric field strength because increase of field strength increases the energy acquired by an ion between collisions. It varies with gas pressure because increase of gas density increases the likelihood that an ion will strike a gas particle in a given distance, but decreases the distance it moves and hence the energy it acquires between successive collisions with gas particles. Because of these conflicting effects, the ionization factor passes through a maximum value as the pressure is raised from a low value.

1-8. Photoionization.—In its narrower sense, the photo-electric effect is the release of electrons from the surface of a solid by light or other electromagnetic radiation. In its broader sense the photoelectric effect is the ionization of an atom or molecule by collision with a photon and may take place not only at the surface of a solid, but throughout a gas, liquid, or solid. The photoelectric effect and its applications will be discussed in detail in Chap. 13. The principles that govern ionization by collision also apply to photoelectric ionization. The energy of a single incident photon $h\nu$ must be at least equal to the first ionization potential of the atom. Impact of photons with atoms results not only in ionization but also in excitation. Ionization may occur in successive steps by a rapid sequence of impacts of successive photons. Excitation by photons is the inverse process to radiation, just as collision of the second kind is the inverse phenomenon to excitation or ionization by collision. Absorption spectra are an indication of the conversion of radiant energy into energy of excitation or ionization.

The frequencies of photons capable of ionizing most elements lie outside of the visible spectrum. The exceptions to this rule are the alkali metals, which are therefore used in light-sensitive electron tubes.

1-9. Ionization by Cosmic Rays.—The exact constitution of cosmic rays is still the basis of much scientific controversy. All experiments seem to indicate, however, that the primary source is outside of the earth and its atmosphere. They appear to be either electromagnetic radiation, of very short wave length, or charged particles that move with extremely high velocity. Ionization by cosmic rays at the surface of the earth results mainly from secondary rays formed in the outer atmosphere and appears to consist of both photoelectric ionization and ionization by collision. It is of importance in connection with glow and arc discharges because it is one of the sources of initial ionization that is necessary to the formation of glows in cold-cathode discharge tubes.

1-10. Space Charge and Space Current.—A group of free charges in space is called *space charge*. If the charge is of one sign only or if the density of charge of one sign in a given volume exceeds that of the other sign, the charge will give rise to an electrostatic field. The relation between the net charge within a volume and the resulting electrostatic flux is given by Gauss's law, which states that the electric flux through any surface enclosing free charges is equal to 4π times the sum of the enclosed charges.[1]

The movement of free charges in space constitutes a current, called *space current*. The current per unit area is equal to the product of the charge density and the velocity normal to the area. The conventional direction of current is in the direction of the motion of positive charges and opposite in direction to the motion of negative charges.

1-11. Deionization.—The rapid disappearance of the products of ionization is necessary for the proper functioning of certain types of electron tubes containing gas or vapor. The removal of ions from a volume of gas or vapor takes place in four ways:

1. Volume recombination.
2. Surface recombination.
3. Action of electric fields.
4. Diffusion.

[1] PAGE, L., and ADAMS, N. I., "Principles of Electricity," pp. 18, 41, D. Van Nostrand Company, Inc., New York, 1931.

There is good experimental evidence that the direct recombination of electrons with positive ions is of relatively rare occurrence in electrical discharges. Volume recombination results mainly from the attachment of electrons to neutral gas molecules to form the heavier and slower-moving negative ions, which subsequently combine with positive ions. The likelihood of attachment of electrons to neutral molecules decreases with increase in temperature, decrease in pressure, and increase in field strength. The rate of recombination of positive and negative ions is proportional to the product of the two ion densities, the constant of proportionality being called the *coefficient of recombination.* The coefficient of recombination is different for different gases.

Surface recombination occurs at conducting walls of the tube that contains the ionized gas, at the electrodes, and at surfaces of any other conducting solids that project into the ionized gas. Recombination takes place as the result of charges of opposite sign that are induced on conducting surfaces or pulled out by strong fields at the surfaces. On insulating surfaces or on conductors that are electrically isolated, charges may also accumulate and build up an electric field that repels charges of the same sign and attracts those of opposite sign. Surface recombination is one of the main factors affecting deionization in glow- and arc-discharge tubes and in circuit breakers of the "deion" type.

The electric field in the vicinity of the electrodes carries ions to the electrode surfaces, where surface recombination may take place. Within the main part of the gas, the electric field can change the ion density in a given small volume only if the ion density is not uniform throughout the tube, or at least in the vicinity of the volume under consideration. If the ion density is uniform, the field will sweep as many new ions into one side of the volume as it removes from the other. If an ion density gradient does exist, the action of the field may either increase or decrease the density in the given volume. An example of deionization by electric fields is the removal of ions in glow- and arc-discharge tubes when applied potentials are reduced below the value necessary to maintain a glow or arc.

Diffusion of ions results from the fact that ions, like gas molecules, have a random motion which carries them from point to point. If the ion density in an element of volume is greater than in adjacent elements, then on the average more ions leave the given element than enter it; if, on the other hand, the ion density is uniform throughout the region under consideration, then on the average just as many ions enter a given element of volume in unit time as leave it. Thus, deionization by diffusion, like deionization by electric fields, is dependent upon nonuniform density. Because of surface recombination and nonuniform production of ions throughout the tube, ion density differences are set up in glow- and arc-discharge tubes which enable diffusion and electric fields to be effective deionizing agents.

1-12. Free Electrons in Metals. Electron Affinity.—Some of the electrons of metallic atoms are very loosely bound to the atoms. When groups of such atoms are massed together into solids or liquids, the loosely bound electrons can readily pass from atom to atom. Because of their random temperature velocity these *free electrons* are constantly moving about within the mass, and, although at some particular instant they may be loosely bound to particular atoms, on the average they experience no force in any particular direction. It is these electrons that make metallic conduction possible and that play indispensable roles in thermionic emission.

An electron that happens to pass through the surface of the metal in the course of its random motion will, while it is still close to the surface, induce a positive charge on the surface of the metal. This induced charge results in a force that tends to return the electron into the metal, the so-called *image force*. In order to escape from the metal the electron must give up a certain amount of kinetic energy. The kinetic energy that an electron loses in passing through the surface of a metal and far enough away to be beyond the range of the image forces is called the *electron affinity* of that metal and is represented by the symbol w.

The electron affinity is different for different materials and varies greatly with the condition of the surface and with impurities contained in the substance, particularly at the surface. Schottky has shown theoretically that the electron affinity should be reduced by strong electric fields at the surface and that it should vary from point to point on a given surface because of microscopic irregularities, being smaller at projections and larger in hollows.[1] Table 1-II lists representative values of electron affinity for a number of metals commonly used in electron tubes.

TABLE 1-II.—ELECTRON AFFINITY w IN ELECTRON VOLTS

Tungsten	4.52	Magnesium	2.7
Platinum	5.0	Nickel	4.0
Tantalum	4.1	Sodium	1.9
Molybdenum	4.3	Mercury	4.4
Carbon	4.5	Calcium	2.5
Copper	4.0	Barium	2.0
Thorium	3.0	Thorium on tungsten	2.63

Oxide-coated nickel....0.5 to 1.5

1-13. Contact Difference of Potential.[2]—A potential difference, called *contact difference of potential*, exists between the surfaces of two

[1] SCHOTTKY, W., *Physik. Z.*, **12**, 872 (1914), **20**, 220 (1919); *Ann. Physik*, **44**, 1011 (1914); *Z. Physik*, **14**, 63 (1923).

[2] An interesting discussion of contact potential and other small effective electrode voltages is given by R. M. Bowie, *Proc. I.R.E.*, **24**, 1501 (1936).

different metals that are in contact or are connected through an external circuit, as shown in Fig. 1-1. Contact difference of potential results from the fact that the electron affinities of the two metals differ. It is very nearly equal to the difference between the electron affinities divided by the charge of an electron. To prove this, suppose that an electron starts from any point a close to the surface of one of the metals, passes into this surface, thence around the circuit, and out through the surface of the second metal to any point b close to its surface. In passing into the first metal the electron does work equal to $-w_1$, the electron affinity of that metal. In crossing the junction between the metals it does a small amount of work because of the Peltier potential difference, which always exists at the junction of two different metals and which is also a function of the electron affinities of the metals. In passing out of the surface of the second metal it does work equal to w_2, the electron affinity of that metal. If the small amount of energy that the electron loses as the result of impacts with atoms and molecules in the metals (I^2R loss) is neglected and the Peltier potential difference is called V_P, the total amount of work done is $\mathcal{E}V_P + w_2 - w_1$, where $\mathcal{E}$ is the charge of an electron. $\mathcal{E}V_P$ may be shown to be small in comparison with $w_2 - w_1$, and so the work done is approximately $w_2 - w_1$. Since the potential difference between two points may be defined as the work done in moving a unit charge between the points, it follows that a difference of potential approximately equal to $(w_2 - w_1)/\mathcal{E}$ exists between any two points a and b close to the surfaces of the two metals. If the connection between the metals is made through an external circuit containing one or more additional metals, the proof is altered only in respect to the additional Peltier potential differences, which are negligible.

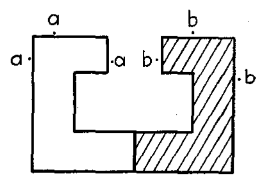


FIG. 1-1.—Production of contact difference of potential by metals of unequal electron affinities.

Inasmuch as the various electrodes in electron tubes may be made of different metals, contact potentials may exist. The contact potential differences may be of the order of 1 to 4 volts, as seen from Table 1-II, and must sometimes be taken into account when applied voltages are small or when high accuracy is necessary in the analysis of electron tube behavior.

1-14. Emission of Electrons and Other Ions from Solids.—The presence of ions in a given volume may result not only from ionization processes in that volume but also from the introduction of ions produced or existing elsewhere. The mechanism by which ions are introduced may be diffusion, action of electric fields, or the emission of electrons or positive ions from the surfaces of solids or liquids. There are five ways in which ions may be emitted from the surfaces of solids or liquids:

1. Thermionic emission.
2. Photoelectric emission.
3. Secondary emission.
4. Field emission.
5. Radioactive disintegration.

Thermionic and photoelectric emission will be discussed in detail in later chapters and therefore need no further consideration at this point.

Secondary Emission.—When ions or excited atoms impinge upon the surface of a solid, electrons may be ejected from the surface. These are called *secondary electrons*, and the phenomenon is termed *secondary emission.* Some secondary emission as the result of electron bombardment appears to take place when the energy of the impinging electrons is less than the electron affinity of the emitter. For this reason it seems likely that the energy that makes possible the escape of a secondary electron is obtained not only from the impinging electron, but also from the thermal energy of the emitter. The number of secondary electrons ejected per primary electron increases with the velocity of the primary electrons and may become great enough to have an appreciable effect upon the behavior of electron tubes at accelerating voltages as low as 5 to 10 volts. At several hundred volts the emission passes through a maximum and then continues to fall with further increase of accelerating voltage. This may be because the primary electrons penetrate farther into the solid and transfer most of their energy to electrons that are so far from the surface that they collide with atoms or molecules before reaching it. The number of secondary electrons emitted per primary electron is in general higher for surfaces having a low electron affinity. It also depends upon the condition of the emitting surface, being reduced by degassing the emitter and by carbonizing the surface. A single primary electron may eject as many as 8 or 10 secondary electrons. The primary electrons may be absorbed, reflected, or scattered by the surface. The number of secondary electrons released by a single impinging particle is less for positive-ion bombardment than for electron bombardment, and emission does not appear to take place unless the energy of the impinging ions exceeds the electron affinity of the emitter. The phenomenon of secondary emission is an important one in all types of electron tubes.

Field Emission.—Field emission is the emission of electrons as the result of intense electric fields at a surface. It is probably an important factor in the operation of certain types of arc discharges. The field strengths required to pull electrons through surfaces at ordinary temperatures are so great that the phenomenon is difficult to produce by direct means.

Emission Resulting from Radioactive Disintegration.—The emission of positive or negative ions in the disintegration of radioactive substances is of importance in glow- and arc-discharge tubes because the presence of minute traces of radioactive materials in the tube walls and electrodes may thus produce a small amount of residual ionization, which makes possible the initial flow of current.

1-15. Electron Dynamics.—The motion of ions, including electrons, is governed by the same laws as the motion of larger masses that can be observed directly. Analyses of the dynamics of masses are based primarily upon Newton's second law, which may be stated symbolically by the equation

$$f = ma \tag{1-1}$$

in which f is the force in dynes acting upon a mass of m grams and a is the resulting acceleration in centimeters per second per second. Aside from forces resulting from the collision of charged particles with other charged or uncharged masses, the forces acting upon charged particles are electrostatic and electromagnetic. An *electric* or a *magnetic* field is said to exist in a region in which electric or magnetic forces, respectively, act. The *electric intensity* (*electric force*) at any point is a vector quantity which is given, both in magnitude and in direction, by the force (mechanical) per unit positive charge which would act on a charged particle placed at this point. The *magnetic intensity* (*magnetic force*) at a point may be defined as the vector quantity which is measured in magnitude and direction by the force (mechanical) that would be exerted on a unit magnetic pole placed at the point.

1-16. Motion of an Electron with Zero Initial Velocity in a Uniform Electric Field.—In the simplest case commonly encountered in electron tubes, the electrostatic force acting upon an ion results from the application of a potential difference to two parallel plane electrodes whose area is large in comparison with their separation. Except near the edges, the electric intensity between such plates is constant throughout the space between them. Since difference of potential between two points may be defined as the work done in moving a unit charge between the points, it follows that the potential difference is equal to

$$E = \int_0^d F\,dx \qquad \text{e.s.u. (statvolts)} = 300 \int_0^d F\,dx \qquad \text{volts} \tag{1-2}$$

where F is the electric intensity in e.s.u. and d is the electrode spacing in centimeters. Since F is constant,

$$E = F \int_0^d dx = Fd \qquad \text{e.s.u.} \tag{1-3}$$

and

$$F = \frac{E}{d} \qquad \text{e.s.u.} \tag{1-4}$$

The force exerted upon an electron between two such plates is equal to the force exerted upon a unit charge times the charge of the electron:

$$f_e = \frac{E\mathcal{E}}{d} \qquad \text{dynes} \tag{1-5}$$

where E is expressed in e.s.u. (statvolts), $\mathcal{E}$ in e.s.u. (statcoulombs), and d in centimeters. It follows from Eqs. (1-1) and (1-5) that the acceleration of the electron is

$$a_e = \frac{f_e}{m_e} = \frac{E\mathcal{E}}{m_e d} \qquad \text{cm/sec}^2 \tag{1-6}$$

in which m_e is the mass of the electron in grams. Since, by definition, acceleration is the rate of change of velocity, the velocity at any instant after the electron starts moving under the influence of the field is

$$v_e = \int_0^t a_e\,dt = \frac{E\mathcal{E}}{m_e d}\int_0^t dt = \frac{E\mathcal{E}t}{m_e d} \qquad \text{cm/sec} \tag{1-7}$$

The distance moved by the electron in the time t is

$$s = \int_0^t v_e\,dt = \frac{E\mathcal{E}}{m_e d}\int_0^t t\,dt = \frac{E\mathcal{E}}{2m_e d}t^2 \qquad \text{cm} \tag{1-8}$$

From Eq. (1-8) it follows that the time taken for the electron to move from one electrode to the other under the sole influence of the electric field is

$$t_d = d\sqrt{\frac{2m_e}{E\mathcal{E}}} \qquad \text{sec} \tag{1-9}$$

The velocity with which it strikes the positive plate is

$$v_d = \frac{E\mathcal{E}t_d}{m_e d} = \sqrt{\frac{2E\mathcal{E}}{m_e}} \qquad \text{cm/sec} \tag{1-10}$$

The energy with which it strikes and which is converted into heat in the positive electrode is

$$K.E. = \tfrac{1}{2}m_e v_d^2 = E\mathcal{E} \qquad \text{ergs} \tag{1-11}$$

Equation (1-11) might have been obtained directly, since the difference of potential between the electrodes is the energy acquired by a unit charge moved under the sole influence of the field and, if the charge does not collide with other particles on the way, this energy can appear only in kinetic form.

1-17. Motion of an Electron in a Uniform Electric Field. Initial Velocity Parallel to the Field.—If an electron leaves one of the plates with initial velocity v_o parallel to the field, the velocity at any instant thereafter is

$$v_e = v_o \pm a_e t \qquad \text{cm/sec} \tag{1-12}$$

in which the minus sign is used if the direction of the electric force is such as to reduce the initial velocity. The distance moved in the time t is

$$s = \int_0^t (v_o \pm a_e t)dt = v_o t \pm \tfrac{1}{2}a_e t^2 \qquad \text{cm} \tag{1-13}$$

If the field reduces the initial velocity, the maximum distance moved can be found by differentiating Eq. (1-13). The time taken for the electron to move through this distance is found by equating ds/dt to zero.

$$0 = \frac{ds}{dt} = v_o - a_e t_{\max} \tag{1-14}$$

$$t_{\max} = \frac{v_o}{a_e} \qquad \text{sec} \tag{1-15}$$

$$s_{\max} = \frac{v_o^2}{a_e} - \frac{v_o^2}{2a_e} = \frac{v_o^2}{2a_e} \qquad \text{cm} \tag{1-16}$$

The electron will reach the second electrode if

$$s_{\max} \equiv \frac{v_o^2}{2a_e} \equiv \frac{m_e d}{2E\mathcal{E}} v_o^2 \geqq d \tag{1-17}$$

that is, if

$$\tfrac{1}{2}m_e v_o^2 \geqq E\mathcal{E} \tag{1-18}$$

Equation (1-18) follows directly from the law of conservation of energy, since the electron can reach the second electrode only if its initial kinetic energy $\frac{1}{2}m_e v_o^2$ equals or exceeds the energy it would lose in moving between the electrodes, *i.e.*, $E\mathcal{E}$.

It also follows from the law of energy conservation that energy gained by a charged particle while moving in vacuum under the action of an electric field must be supplied by the source of potential applied to the electrodes. Conversely, energy given up by an ion in moving in an electric field is returned to the source, or converted into I^2R loss in conductors joining the electrodes to the source, or into heat of impact if the electron strikes an electrode or other surface.

1-18. Motion of Electrons in Nonuniform Electric Fields.—In general, the electric intensity in electron tubes is not uniform, and so the integration of Eqs. (1-2), (1-7), (1-8), and (1-13) is less simple. When space charge becomes appreciable, the motion of individual electrons is also affected by the fields set up by other electrons. This important phenomenon will be discussed in detail in Chap. 2.

Inspection of the methods used in deriving Eqs. (1-5) to (1-18) shows that the equations may be applied to other ions than electrons by making suitable changes in the values of charge and mass.

1-19. Importance of Transit Time.—The time taken for electrons and other ions to move between electrodes is so small that it may be neglected in many analyses of the operation of electron tubes (see Probs. 1-1 and 1-2). It cannot be neglected, however, when the electrode voltages alternate at frequencies so high that the time of transit is of the same order of magnitude as the periods of the voltages. It is also of importance in the study of the deionization of gas- or vapor-filled tubes.

1-20. Electric Field Normal to Initial Direction of Motion.—Up to this point it has been assumed that the initial velocity of the electron is

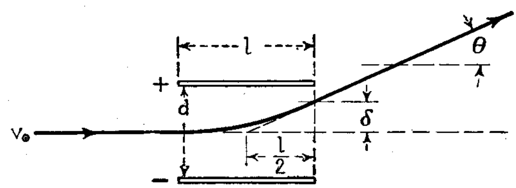


FIG. 1-2.—Deflection of electron beam by electric field.

parallel to the electric field. Under this assumption there is no change in the direction of motion in vacuum. In cathode-ray oscillograph and television tubes, electron beams are deflected by electric fields perpendicular to the initial direction of motion. The arrangement is essentially that shown in Fig. 1-2. Electrons enter the space between the deflecting electrodes with a velocity v_o parallel to the electrode surfaces, and are deflected by the electric field between the electrodes, which have a length l, a separation d, and a potential difference E.

The acceleration produced by the field is normal to the initial velocity and its magnitude is given by Eq. (1-6) as $E\mathcal{E}/m_ed$ cm/sec². Since the velocity normal to the electric field remains constant, the time taken for an electron to move through the deflecting field, which is assumed to be uniform between the plates and zero on either side, is l/v_o. The vertical displacement at the point where the electron leaves the plates is given by Eq. (1-8). It is

$$\delta = \frac{E\mathcal{E}}{2m_ed}t^2 = \frac{E\mathcal{E}l^2}{2v_o^2m_ed} \qquad \text{cm} \tag{1-19}$$

The vertical velocity when the electron leaves the deflecting plates is given by Eq. (1-7)

$$v_y = \frac{E\mathcal{E}t}{m_e d} = \frac{E\mathcal{E}l}{v_o m_e d} \qquad \text{cm/sec} \tag{1-20}$$

The horizontal velocity, which is unaffected by the field, is still v_o. Therefore, the final direction of motion makes an angle with the initial direction given by the relation

$$\tan\theta = \frac{v_y}{v_o} = \frac{E\mathcal{E}l}{v_o^2 m_e d} \tag{1-21}$$

But by Eq. (1-19),

$$\frac{\delta}{\frac{1}{2}l} = \frac{E\mathcal{E}l}{v_o^2 m_e d} = \tan\theta \tag{1-22}$$

Examination of Fig. 1-2 shows, therefore, that after deflection the electrons move as though they had passed through a point midway between and midway along the deflecting plates.

1-21. Motion of an Electron in a Magnetic Field Normal to the Initial Velocity.—Like a current-carrying conductor, an electron moving normal to a magnetic field experiences a force perpendicular to the field and to the direction of motion of the charge. The magnitude of the orce is given by the relation

$$f_h = B\mathcal{E}'v \qquad \text{dynes} \tag{1-23}$$

in which B is the flux density in gauss, $\mathcal{E}'$ is the charge of an electron in *electromagnetic* units, and v is the velocity in centimeters per second.

If the electronic charge is expressed in e.s.u., Eq. (1-23) may also be written

$$f_h = \frac{B\mathcal{E}v}{3 \times 10^{10}} \qquad \text{dynes} \tag{1-24}$$

The action of a magnetic field differs from that of an electric field in that the force on a moving charge in an electric field is always parallel to the field, whereas the force on a moving charge in a magnetic field is normal to the field and to the instantaneous velocity. If an electron enters a uniform magnetic field with an initial velocity v_o normal to the field, it will be deflected at right angles to the field by the force $B\mathcal{E}'v_o$ dynes. Since the acceleration is normal to the velocity, the speed remains constant. If ρ is the instantaneous radius of curvature, the radial acceleration is v_o^2/ρ. Therefore, by Eq. (1-1),

$$B\mathcal{E}'v_o = \frac{m_e v_o^2}{\rho} \tag{1-25}$$

and

$$\rho = \frac{v_0 m_e}{B\mathcal{E}'} = 3 \times 10^{10}\,\frac{v_0 m_e}{B\mathcal{E}} \qquad \text{cm} \tag{1-26}$$

Since B is assumed to be constant, the electron moves with constant speed along a path of constant radius of curvature, *i.e.*, along a circular path, in a plane perpendicular to the field. The dependence of the radius of curvature upon the velocity and upon the mass makes possible the separation of charged particles of different velocities or masses.

If the initial velocity also has a component parallel to the field, the electron will describe a spiral path around an axis parallel to the field. A similar spiral motion results if the initial velocity is parallel to the field but some other force, such as repulsion between two or more electrons, produces an acceleration normal to the field. This principle may be used in preventing the spreading of and in focusing a beam of electrons. Electron beams may also be focused by the use of nonuniform electric fields such as exist between adjacent cylinders of unequal diameter when a difference of potential exists between them.[1] These methods are used in focusing electron beams in cathode-ray oscillograph and television tubes and in electron microscopes (see Secs. 2-4 and 15-17).

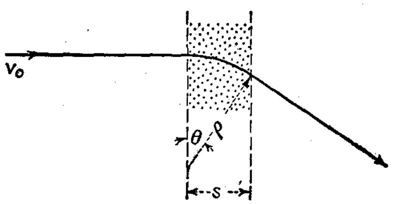


FIG. 1-3.—Deflection of electron beam by magnetic field.

In cathode-ray oscilloscope and television tubes, electron beams may be deflected by magnetic fields normal to the initial velocity of the electrons that comprise the beams. Although a general analysis is complicated, a useful expression for the beam deflection angle may be readily derived under the assumption that the field is uniform throughout a distance s and zero elsewhere, as shown in Fig. 1-3. Inspection of Fig. 1-3 shows that

$$\sin \theta = \frac{s}{\rho} \tag{1-27}$$

in which ρ, the radius of curvature of the circular path within the field, is given by Eq. (1-26). Substitution of Eq. (1-26) in Eq. (1-27) gives the relation:

$$\sin \theta = \frac{sB\mathcal{E}}{3 \times 10^{10} v_o m_e} \tag{1-28}$$

For small deflection angles, Eq. (1-28) simplifies to

$$\theta = \frac{sB\mathcal{E}}{3 \times 10^{10} v_o m_e} \tag{1-29}$$

In the range of θ in which Eq. (1-29) is valid, the path of the electrons

[1] See, for instance, I. G. MALOFF, and D. W. EPSTEIN, "Electron Optics in Television," McGraw-Hill Book Company, Inc., New York, 1939.

after deflection by the magnetic field is the same as though they had originated at the center of the region in which the field acts.

1-22. Crossed Electric and Magnetic Fields.—If an electron is sent through electric and magnetic fields that are perpendicular to each other and to the initial velocity of the electron, as shown in Fig. 1-4, the electron is acted upon by a force $\mathcal{E}F$ dynes caused by the electric field and a force $B\mathcal{E}v_o/(3 \times 10^{10})$ dynes caused by the magnetic field. These forces are both normal to the surfaces of the deflecting plates. If they are equal in magnitude and opposite in direction, the electron is undeflected. This is true if

$$\mathcal{E}F = \frac{B\mathcal{E}v_o}{3 \times 10^{10}} \qquad (1\text{-}30)$$

or

$$v_o = 3 \times 10^{10} \frac{F}{B} \quad \text{cm/sec} \qquad (1\text{-}31)$$

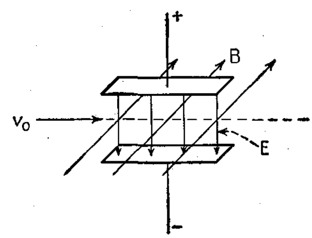


FIG. 1-4.—Path of an electron through balanced electric and magnetic fields.

in which F is measured in e.s.u. and B is measured in gauss. This phenomenon may obviously be used to measure the speed of electrons or other charged particles.

Problems

1-1. *a.* Find the time of transit of an electron between parallel plane electrodes having a separation of 0.2 cm and a potential difference of 250 volts.

$$\mathcal{E} = 4.8 \times 10^{-10} \text{ e.s.u.}$$

b. Find the energy delivered to the positive electrode by the electron.

1-2. *a.* An electron at the surface of a plane electrode is accelerated toward a second parallel plane electrode by a 200-volt battery, the polarity of which is reversed without loss of time 10^{-9} sec after the circuit is closed. If the electrode separation is 1.5 cm, on which electrode will the electron terminate its flight?

b. What will become of the kinetic energy that it acquires during its acceleration?

1-3. *a.* An electron having initial kinetic energy of 10^{-9} erg at the surface of one of two parallel plane electrodes and moving normal to the surface is slowed down by the retarding field caused by a 400-volt potential applied between the electrodes. Will the electron reach the second electrode?

b. What will become of its initial energy?

1-4. The electrons that comprise a beam of cathode rays are given their velocity v_o by means of a potential difference of 500 volts impressed between the electron source and an accelerating anode. Determine the difference of potential that must be impressed between two deflecting electrodes 3 cm long and 1 cm apart in order to deflect the beam through an angle of 20 degrees.

1-5. The electrons that comprise a beam of cathode rays are given their velocity v_o by means of a potential difference of 1500 volts impressed between the electron source and an accelerating anode. Determine the flux density that must exist throughout a distance of 2 cm in order to deflect the beam through an angle of 10 degrees.

Bibliography

WATERMAN, A. T.: Fundamental Properties of the Electron, *Elec. Eng.*, **53,** 3 (1934).

HUGGINS, M. L.: Structure of Atoms and Molecules, Electronic Theory of Valence, *Elec. Eng.*, **53,** 851 (1934).

BRODE, R. B.: Quantitative Study of the Collisions of Electrons with Atoms, *Rev. Modern Phys.*, **5,** 257 (1933).

COMPTON, K. T., and LANGMUIR, I.: Electrical Discharges in Gases; Part I, *Rev. Modern Phys.*, **2,** 123 (1930).

LANGMUIR, I., and COMPTON, K. T.: Electrical Discharges in Gases; Part II, *Rev. Modern Phys.*, **3,** 191 (1931).

TONKS, L.: Electrical Discharges in Gases, *Elec. Eng.*, **53,** 239 (1934).

DARROW, K. K.: Electrical Discharges in Gases—II, Ions in Dense Gases, *Elec. Eng.*, **53,** 388 (1934).

SLEPIAN, J., and MASON, R. C.: Electrical Discharges in Gases—III, Self-maintained Discharges, *Elec. Eng.*, **53,** 511 (1934).

HULL, A. W.: Fundamental Electrical Properties of Mercury Vapor and Monatomic Gases, *Elec. Eng.*, **53,** 1435 (1934).

DARROW, K. K., *Bell System Tech. J.*, **11,** 576 (1932); **12,** 91 (1933); *Rev. Sci. Instruments*, **4,** 6 (1933); *Elec. Eng.*, **54,** 808 (1935).

Physics Staff, University of Pittsburgh: "An Outline of Atomic Physics," John Wiley & Sons, Inc., New York, 1933.

DARROW, K. K.: "Electrical Phenomena in Gases," Williams & Wilkins Company, Baltimore, 1932.

SLEPIAN, J.: "Conduction of Electricity in Gases," Educational Department, Westinghouse Electric and Manufacturing Company, East Pittsburgh, 1933.

THOMSON, J. J., and THOMSON, G. P.: "Conduction of Electricity Through Gases," Cambridge University Press, London, 1928–1932.

EMELÉUS, K. G.: "The Conduction of Electricity through Gases," E. P. Dutton & Company, Inc., New York, 1929.

TOWNSEND, J. S.: "Electricity in Gases," Oxford University Press, New York, 1915; "Motion of Electrons in Gases," 1925.

LOEB, L. B.: "Fundamental Processes of Electrical Discharge in Gases," John Wiley & Sons, Inc., New York, 1939.

COBINE, J. D.: "Gaseous Conductors," McGraw-Hill Book Company, Inc., New York, 1941.

CHAPTER 2

THERMIONIC EMISSION; THE HIGH-VACUUM THERMIONIC DIODE

The operation of most electron tubes is dependent upon thermionic emission. The theory of thermionic emission is, therefore, of great importance in the study of electron tubes. It is the purpose of this chapter to discuss the basic principles of thermionic emission, the construction of practical emitters, and the flow of electron space current in high-vacuum tubes containing two electrodes.

2-1. Theory of Thermionic Emission.—Richardson's theory of the emission of electrons from hot bodies is in many respects analogous to the kinetic theory of vaporization.[1] Heat possessed by a metal is believed to be stored not only in the kinetic energy of random motion of atoms and molecules, but also in the kinetic energy of free electrons. As a result of collisions between electrons or between electrons and atoms or molecules, the speed and direction of motion of a given electron are constantly changing. The random motion of the electrons causes some of them to strike the inner surface of the metal. Electrons that arrive at the surface will escape from the metal if they have a component of velocity toward the surface equal to or greater than u_w, corresponding to a kinetic energy $\frac{1}{2}m_e u_w^2$ equal to the electron affinity w. The number of electrons that reach the surface in unit time with a normal component of velocity equal to or greater than u_w is proportional to the fraction of all the free electrons throughout the metal that have such velocities. At room temperatures this fraction is extremely small, and so no thermionic emission is detectable. As the temperature of the emitter is increased, however, the average velocity of the free electrons increases, and so the number having velocities equal to or greater than u_w is increased. This can be seen from the velocity-distribution curves of Fig. 2-1. These are theoretical curves so constructed that the area under a given curve between any two velocities u and $u + \Delta u$ is proportional to the fraction of the free electrons having velocities lying within this range at the temperature for which the curve is constructed. The area lying to the right of a given positive value of u, such as u_w, is proportional to the fraction of the electrons having velocities in excess of this value. Figure 2-1

[1] Richardson, O. W., *Proc. Cambridge Phil. Soc.*, **11**, 286 (1901); "Emission of Electricity from Hot Bodies," rev. ed., Longmans, Green & Company, New York, 1921.

shows that the fraction of the electrons having velocities toward the surface equal to or greater than u_w increases with temperature. Hence the rate at which electrons escape from the metal, *i.e.*, the emission current, increases with temperature. Observable emission currents are obtained at temperatures in excess of 1000°K. Emission current is also increased by reduction of electron affinity, since reduction of electron affinity lowers the velocity u_w and therefore increases the number of electrons whose velocity toward the surface exceeds u_w. This is clearly shown by Fig. 2-1, since reduction of work function w means that u_w is moved to the left and the area under the curve to the right of u_w is increased.

Electrons that escape will have resultant velocities made up of the excess perpendicular to the surface, plus the original components parallel to the surface, which are not altered by the surface forces. If the emitted electrons are not drawn away by an external field, they will form a

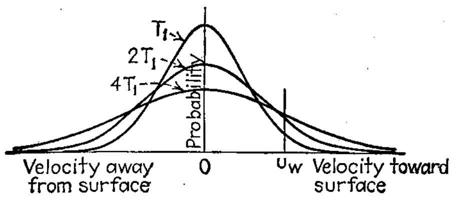


FIG. 2-1.—Maxwellian distribution curves of velocities normal to emitter surface at three temperatures.

space charge, the individual particles of which are moving about with random velocities. Because the initial average normal velocity of the electrons after emission is away from the surface and because of the mutual repulsion of like charges, electrons drift away from the surface. Collisions between electrons cause some of them to acquire velocity components toward the emitter, where they may reenter the surface with a gain of kinetic energy equal to the electron affinity. Another factor responsible for the return of electrons to the emitter is the electrostatic field set up by the negative space charge and, if the emitter is insulated, by the positive charge that it acquires as the result of loss of electrons. This field increases with the density of space charge, and equilibrium is established when only enough electrons can move away from the surface to supply the loss by diffusion of the space charge. If diffusion can then be prevented, just as many electrons return to the metal in unit time as leave it. Figure 2-2 gives a rough picture of the electron distribution under equilibrium conditions.

If a second, cold electrode is placed near the emitting surface in vacuum and connected to the emitter through a galvanometer, as shown in Fig. 2-3, the meter will indicate the small current resulting from the drift of electrons from the emitter to the second electrode. These electrons return to the emitter through the galvanometer and prevent the emitter from becoming positively charged. This phenomenon, first observed by Edison, is called the *Edison effect.* When the second electrode is made positive with respect to the emitter by the addition of a battery, as shown in Fig. 2-4, the current is increased. As the voltage is gradually raised, it is found that at any emitter temperature there is a more or less definite voltage beyond which the current is nearly constant, all emitted electrons being drawn to the collector. This current is called the *saturation current,* and the corresponding voltage, the *saturation voltage.* The lack of increase of current beyond saturation voltage is

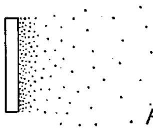


FIG. 2-2.—Distribution of electrons near an emitting surface.

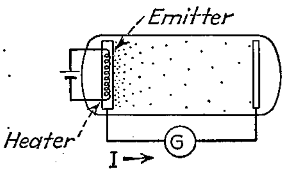


FIG. 2-3.—Flow of anode current as the result of diffusion of electrons from cathode to anode without the application of anode voltage (Edison effect).

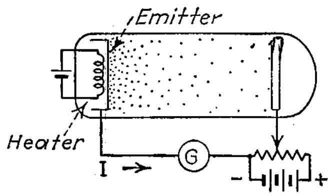


FIG. 2-4.—Use of anode voltage to increase anode current.

spoken of as *voltage saturation.* Saturation current varies with the temperature and electron affinity of the emitter. The negative emitter in Fig. 2-4 is called the *cathode;* and the positive collector, the *anode* or *plate.* An electron tube containing only a cathode and an anode is called a *diode.* Although the electrons move from cathode to anode, the current, according to convention, is said to flow from anode to cathode within the tube.

2-2. Richardson's Equation.—By means of classical kinetic theory and by thermodynamic theory, Richardson derived two slightly different equations for saturation current as a function of temperature.[1] It is not possible experimentally to determine which of Richardson's equations is correct, but this was later done theoretically by M. v. Laue, S. Dushman, and A. Sommerfeld. The equation that is now believed to be correct is

$$I_s = AT^2\epsilon^{-w/kT} \tag{2-1}$$

in which I_s is the saturation current per unit area of emitter, T is the absolute temperature, w is the electron affinity of the emitter, k is Boltzmann's universal gas constant, and A is a constant, probably universal for pure metals. The value of k is 8.63×10^{-5} electron volt/degree and the

[1] RICHARDSON, *loc. cit.*

theoretical value of A for pure metals is 60.2. The form of the curve that represents Richardson's equation, shown in Fig. 2-5*a*, is determined practically entirely by the exponential factor.

It is important to note that Richardson's equation holds only for the saturation current and that the anode voltage must, therefore, be high enough at all times so that all emitted electrons are drawn to the anode. If the anode voltage is fixed at some value E_1, while the temperature is raised, then at some temperature the current will begin to be limited by space charge in a manner similar to that when there is no accelerating voltage. Further increase of temperature will then have no effect upon the current. This temperature is called the *saturation temperature*, and the failure of the current to increase at higher temperature is spoken of as *temperature saturation*. If the emitter were homogeneous and the electrostatic field constant over the surface of the emitter, the advent

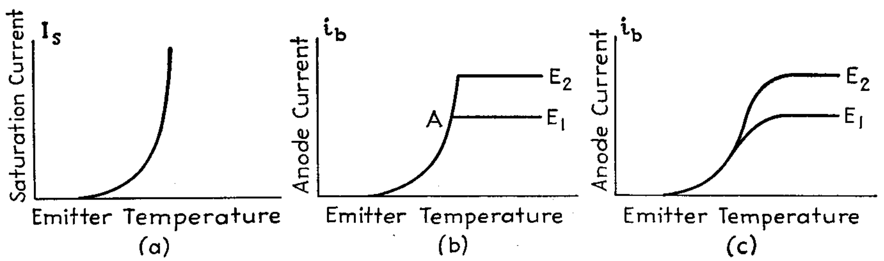


FIG. 2-5.—Curves of anode current *vs.* emitter temperature: (*a*) saturation emission current; (*b*) theoretical curves at two values of anode voltage; (*c*) experimental curves at two values of anode voltage.

of saturation would be abrupt, as indicated at point A of Fig. 2-5*b*. Actually, because of variations of temperature and electron affinity and of electrostatic field, saturation does not take place over the whole cathode surface at the same temperature, and so experimentally determined curves bend over gradually, as shown by the curves of Fig. 2-5*c*. At higher anode voltage E_2, saturation occurs at a higher temperature. If the voltage is increased with temperature, the current will continue to rise with temperature, as in Fig. 2-5*a*, until the temperature becomes sufficiently high to vaporize the emitter.

Since only those electrons which have relatively high energies can escape from the metal, thermionic emission necessarily results in the reduction of the average energy of the remaining electrons and molecules, and hence of the temperature of the emitter. Heat must be supplied continuously to the emitter in order to prevent its temperature from falling as the result of emission. The cooling effect of emission current is plainly visible in filaments in which the emission current is comparable with the heating current, as in the type 30 tube.

Richardson's equation shows that the emission current which can be obtained at any temperature varies inversely with the electron affinity. The exponential form of the equation shows that small changes in temperature or electron affinity may result in large changes of emission current. A 15 per cent reduction of electron affinity produces an eight- or tenfold increase of emission in the working range of temperature. The ratio of the emission current in milliamperes per square centimeter to the heating power in watts per square centimeter is called the *emission efficiency*. Emission efficiency increases with decrease of electron affinity.

A satisfactory practical source of thermionic emission must satisfy two requirements: it must have a high emission efficiency and it must have a long life. Thermal losses can be reduced by proper design of the emitter and by reduction of emitter temperature. (Cathodes of special design, which give very low thermal losses, can be used in gaseous discharge tubes. These will be discussed in Sec. 12-16.) The life of an emitter increases with the difference between the normal operating temperature and the vaporization or melting temperatures of the metal or metals of which it is constructed. Since low operating temperature is made possible by low electron affinity, it is evident that the choice of emitters of low electron affinity is favorable to long life, as well as to high efficiency.

2-3. Pure Metallic Emitters.—Pure metals having low electron affinities, such as the alkali metals or calcium, cannot be used as emitters in electron tubes because they vaporize excessively at temperatures at which appreciable emission is obtained. Only two pure metals, tantalum and tungsten, are suitable for use as practical emitters. Although the electron affinity of tantalum is lower than that of tungsten, tantalum is more sensitive to the action of residual gases and has lower vaporization temperature. Tantalum is therefore seldom used. Pure metallic emitters are now used only in large high-voltage (above 3500 volts) power tubes, in which they are found to have longer life than the special emitters that are used successfully in small tubes.

2-4. Thoriated Tungsten Emitters.—The presence of impurities in a metal may produce a marked change in the value of its electron affinity. This is usually attributed to the formation of thin layers of these impurities at the surface. Such a layer may produce very high fields at the surface by virtue of the fact that it may be electropositive or electronegative relative to the main metal. Thus the presence of an absorbed layer of oxygen, which is electronegative with respect to tungsten, results in a field that opposes the emission of electrons and therefore increases the electron affinity of tungsten. The presence, on the other hand, of a monatomic layer of thorium atoms or ions on the surface of tungsten reduces its electron affinity remarkably. It is of interest to note that the electron affinity of thoriated tungsten may be even lower than that of pure thorium (see Table 1-II).

The reduction of the electron affinity of tungsten as the result of introduction of small amounts of thorium was first observed by Langmuir in 1914 in the course of a study of the properties of tungsten filaments.[1] Thorium oxide is introduced into the tungsten in the course of its manufacture. Subsequent high temperature converts a portion of the thorium oxide into metallic thorium, which diffuses to the surface. Investigations by Dushman, Becker, and others[2] indicated that the lowest value of w and the highest value of the constant A in Richardson's equation are obtained when the tungsten is completely, or perhaps very nearly, covered with a single layer of thorium atoms.

Thoriated tungsten shows no increase of emission over that of pure tungsten until it is *activated.* The activation process is performed after evacuation of the tube. It consists first in "flashing" the emitter for a few moments at a temperature of 2500 to 2800°K. This high temperature reduces some of the thorium oxide to thorium. The temperature is then kept for some minutes at about 2200°K, which allows the metallic thorium to diffuse to the surface. The best value of diffusing temperature is determined by the rates of diffusion of thorium to the surface, and of evaporation from the surface. If the temperature is too high, the evaporation exceeds the diffusion, resulting in deactivation. The emitter is normally operated at temperatures that do not exceed 2000°K, which is sufficiently low so that evaporation of thorium from the surface is negligible. If the emitter is accidentally operated at such a high temperature that the whole supply of thorium diffuses to the surface and evaporates, it can be reactivated by repeating the original activation process. This may be done several times before all the thorium oxide is used up.

A useful tool in the study of the phenomenon of activation is the "electron microscope."[3] This consists of the emitter and means for accelerating the electrons and for focusing them upon a screen[4] which fluoresces under the impact of electrons. It has been shown that the

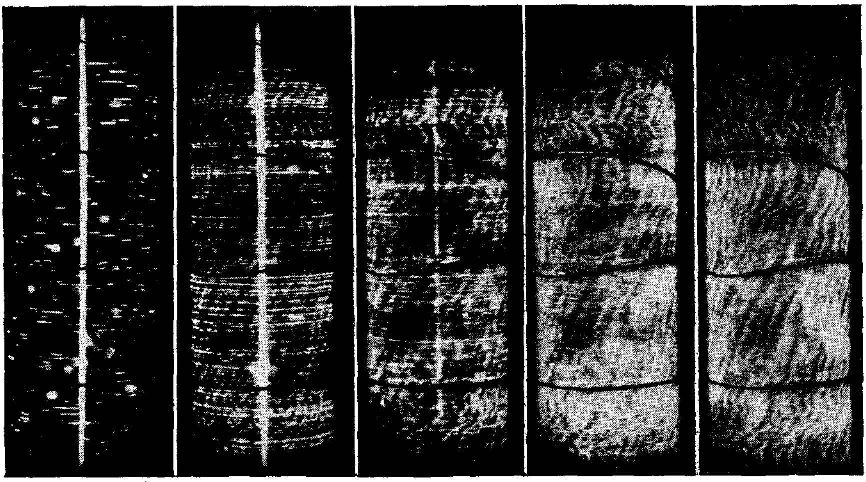
[1] LANGMUIR, I., *Phys. Rev.*, **4**, 544 (1914).

[2] DUSHMAN, S., and EWALD, J., *Phys. Rev.*, **29**, 857 (1927); BECKER, J. A., *Trans. Am. Electrochem. Soc.*, **55**, 153 (1929).

[3] KNOLL, M., and RUSKA, E., *Ann. Physik*, **12**, 607 (1932). The electron microscope has many applications besides that mentioned here. Recent instruments may be used in place of ordinary microscopes in the study of matter and give higher magnification than can be attained with light. For a survey and a bibliography of this subject, see R. P. Johnson, *J. Applied Physics*, **9**, 508 (1938).

[4] See, for instance, I. G. MALOFF and D. W. EPSTEIN, "Electron Optics in Television," McGraw-Hill Book Company, Inc., New York, 1938.

action of nonuniform electromagnetic and electrostatic fields upon electron beams is similar to the action of lenses upon light.[1] Thus it is possible to obtain on the screen a sharp enlarged image which shows clearly the individual points of emission of the cathode.[2] Similar results are achieved by use of a straight filament at the axis of a cylindrical glass tube, the inner surface of which is covered with a fluorescent material.[3] The coated surface, which acts as the anode, is maintained at a positive



(a) (b) (c) (d) (e)

FIG. 2-6.—Typical activation behavior of thoriated tungsten. Image (a) immediately after 10 sec. at 2800°K, (b) after additional 4 min. at 1850°K, (c) after 20 min. at 1850°K, (d) after 30 min. at 1850°K, (e) after 70 min. at 1850°K. All pictures were made at 1200°K. The decreasing exposure time is evidenced by reduction of apparent brilliance of the filament. (*Courtesy of R. P. Johnson.*)

potential of several thousand volts by means of a wire helix coiled inside the tube in contact with the coating. Electrons emitted by the filament are attracted radially toward the anode coating, where they produce a magnified image of the electron emission at the surface of the filament. Figure 2-6 gives a series of photographs of the screen of such a tube, showing the activation of thoriated tungsten. (The bright vertical line is caused by light from the filament, and the dark lines by the shadow of the helix.)

[1] BUSCH, H., *Ann. Physik*, **81**, 974 (1926); MALOFF, I. G., and EPSTEIN, D. W., *Proc. I.R.E.*, **22**, 1386 (1934); EPSTEIN, D. W., *Proc. I.R.E.*, **24**, 1095 (1936); MALOFF and EPSTEIN, "Electron Optics in Television," *op. cit.*

[2] See, for instance, E. BRÜCHE and H. JOHANNSON, *Ann. Physik*, **15**, 145 (1932); M. KNOLL, *Electronics*, September, 1933, p. 243.

[3] JOHNSON, R. P., and SHOCKLEY, W., *Phys. Rev.*, **49**, 436 (1936). See also *Electronics*, January, 1936, p. 10, March, 1937, p. 23.

The presence of even small amounts of gas has a very destructive effect upon a thoriated tungsten emitter. This may result either from direct chemical action, such as oxidation, or from the removal of thorium from the surface by the bombardment of positive ions. The sensitiveness of thoriated emitters to the action of gases and the rate of evaporation of thorium may be greatly reduced by heating the emitter in an atmosphere of hydrocarbon vapor, which causes the formation of a shell of tungsten carbide. Because of the reduction of the rate of evaporation of the thorium, a carbonized emitter can be operated at a much higher temperature, with consequent increase of emission current and efficiency. At this higher temperature the increase of diffusion makes possible the con-

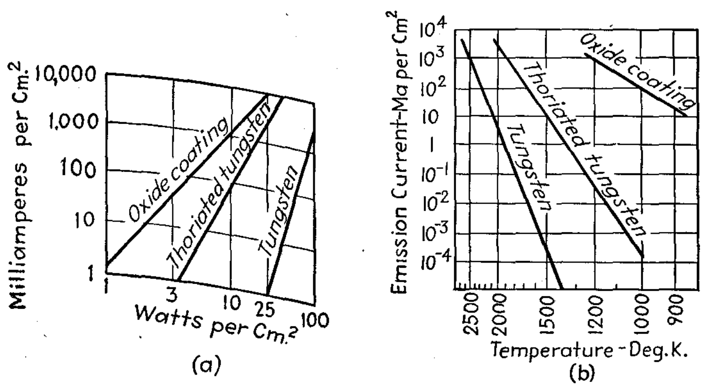


FIG. 2-7.—(a) Typical curves of emission current *vs.* heating power; (b) typical curves of emission current *vs.* emitter temperature.

tinuous replacement of thorium removed from the surface by the action of gas molecules or ions. Figure 2-7 shows that the emission efficiency of thoriated tungsten is much higher than that of pure tungsten and that a given emission may be obtained at a much lower temperature. Because of the lower electron affinity, higher emission efficiency, and longer life of oxide-coated emitters, thoriated tungsten emitters are now used very little in receiving tubes.

2-5. Oxide-coated Emitters.—By far the most widely used emitters in small high-vacuum tubes are oxide-coated emitters, first used by Wehnelt.[1] Although the process of manufacture of oxide-coated cathodes varies considerably, it consists, in general, in coating a core metal, usually nickel or alloys of nickel and other metals, with one or more layers of a mixture of barium and strontium carbonates (approximately 60 per cent barium carbonate and 40 per cent strontium carbonate). The carbonates

[1] WEHNELT, A., *Ann. Physik*, **14**, 425 (1904). For an excellent review of the subject of oxide-coated emitters see J. P. Blewett, *J. Applied Physics*, **10**, 668 and 831 (1939).

may be suspended in water, although a binder such as collodion or a mixture of one part of Zapon varnish in 20 parts of amyl acetate is usually used. The mixture may be applied to the core by spraying or by dipping or dragging the core through the mixture. When a thick coating is desired, the mixture is preferably applied in the form of several thin coatings heated sufficiently between applications to burn out the binder. After application of the carbonate coating, the emitter is mounted in the tube, which is then evacuated, and the emitter is heated electrically to a temperature of about 1400°K. The high temperature reduces the carbonates to oxides, the liberated carbon dioxide being removed by the pumps. The temperature is then lowered somewhat and voltage is applied to the anode for some time, during which the emission builds up to its proper value. The normal operating temperature ranges from 1000 to 1300°K.

Many experiments have been performed to determine what takes place during the activation process and from which part of the emitter electrons are emitted. Reduction of the oxides to pure metal may result from chemical reaction, from electrolysis of the oxides, or from the bombardment of positive ions formed in the gas between the anode and the emitter by electrons accelerated by the applied field. Perhaps all three of these processes occur. Free metal formed throughout the oxide diffuses toward the surface. Although particles of free metal are distributed throughout the coating in a completed emitter, most recent evidence appears to indicate that the emission takes place at the outer surface.

The electron affinity, saturation emission current, and emission efficiency of oxide-coated emitters are greatly dependent upon the manufacturing processes. The electron affinity ranges from 0.5 to 1.5 electron volts. Typical curves of saturation current *vs.* temperature and emission efficiency *vs.* heating power are shown in Fig. 2-7. Examination of these curves shows that the saturation current and emission efficiency are considerably higher than those of thoriated tungsten. The relatively low temperature at which oxide-coated cathodes can be operated is a distinct advantage in most applications.

The emission from oxide-coated cathodes is reduced or destroyed by the presence of oxygen, due to oxidation of the active metal or to removal of the active metal or even the complete coating by positive-ion bombardment. Another cause of damage to oxide-coated cathodes is the development of hot spots. Because of nonuniform activation of the emitter, emission is not uniform over the surface. The flow of emission current through the oxide coating, which has high resistance, raises its temperature. Since the temperature rise is greatest at points of the cathode at which the emission is high, the emission increases still more at these points. If the current is not limited by space charge, the action may become cumulative and the current and temperature increase to such an extent that the coating is removed. In filamentary cathodes the local rise in temperature may be so great as to melt the filament. Hot spots are most likely to occur at high anode voltages. The tendency toward the formation of hot spots is also increased by deactivation caused by the flow of excessive space current.

When full emission current is drawn from an oxide-coated emitter, the current first falls rapidly and then slowly approaches a steady value. This decay of current is thought to be caused by electrolytic removal of barium from the surface or by electrolytic deposition of oxygen on the surface. The initial emission is recovered if the emitter is heated without the flow of space current. The useful life of oxide-coated emitters, which is several thousand hours, is terminated by a rather sudden decay in emission to a very low value. This may be caused by evaporation of free barium and of the supply of barium oxide that furnishes free barium during the active life of the emitter. The useful life of a vacuum tube containing an oxide-coated cathode may also be terminated by the liberation of gas from the emitter. Schade has pointed out that the peak current that can be obtained from an oxide-coated cathodes greatly exceeds the steady current.[1] Transient peak currents of 25 amp/cm^2 have been observed from well-activated cathodes. The stable peak emission over an extended period is usually less than one-third of this value.

Formation of hot spots and the likelihood of ionization of gas emitted during the life of oxide-coated emitters make them unsuitable for use in high-voltage transmitting tubes.

2-6. Cesiated Tungsten Emitters.—A fourth type of emitter, not used commercially in thermionic tubes, is produced by depositing a monatomic layer of cesium on tungsten. Because the ionizing potential of cesium vapor is less than the electron affinity of tungsten, the tungsten removes an electron from a cesium atom which strikes it, leaving a positive ion which is held to the tungsten surface by the resulting electrostatic field. The force of adhesion is even greater if the tungsten is first covered with a monatomic layer of oxygen, which is electronegative with regard to tungsten. The strong electrostatic field between the cesium ions and the tungsten or oxygen reduces the electron affinity to the comparatively low value of 0.7 electron volt or less. Because cesium melts at a temperature only slightly above room temperature, the cesium vapor is obtained by merely introducing a small amount of cesium into the evacuated tube, the subsequent vaporization being sufficient to coat the filament.

[1] SCHADE, O. H., *Proc. I.R.E.*, **31**, 341 (1943).

The low electron affinity of the tungsten-oxygen-cesium emitter makes possible high emission currents at a temperature of only 1000°K. This type of emitter has several disadvantages, however, which make it impractical for use in commercial tubes. As the result of the high vapor pressure of cesium at operating temperatures of the tube, the characteristics of the tube are influenced by tube temperature. Too high temperature vaporizes the cesium, causing temporary reduction in emission, or even removal of the oxygen layer with permanent reduction of emission. Except at very low anode voltages, ionization of the cesium vapor occurs, resulting in fluctuations of anode current. The presence of positive ions is also detrimental to the action of amplifier tubes for other reasons, which will be discussed (Secs. 2-8, 6-7, 13-16).

2-7. Mechanical Construction of Cathodes.—Cathodes used in high-vacuum thermionic tubes are divided into two general classes: filamentary and indirectly heated. Figure 2-8 shows the form of typical filamentary cathodes. Early vacuum tubes used only filamentary cathodes. When filamentary cathodes are operated on alternating current, the stray alternating electrostatic field and the alternating voltage across the filament cause an alternating component of plate current that may be objectionable. This difficulty led to the development of the indirectly heated, or *heater-type*, cathode.

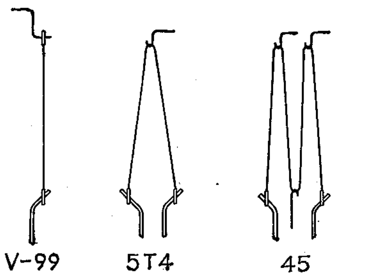


FIG. 2-8.—Typical filamentary cathodes. (*Courtesy of Radio Corporation of America.*)

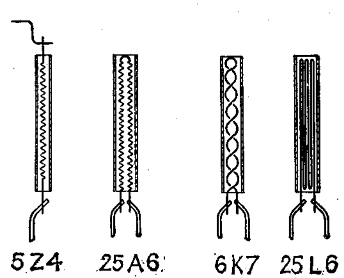


FIG. 2-9.—Structure of typical heater-type cathodes. (*Courtesy of Radio Corporation of America.*)

Indirectly heated cathodes used in receiving tubes consist of an oxide-coated cylindrical sleeve, usually of nickel, within which is some form of heater. The most common types of heaters are illustrated in Fig. 2-9. The 5Z4 and 25A6 heater coils are helically wound. That of the 6K7 type of cathode is wound in a reverse helix. After being wound and formed, the coils are coated with a refractory insulating material and inserted into the sleeve. The heater of the 25L6 type of cathode is covered with a refractory insulating coating of sufficient adherence to permit the wire to be bent into the desired shape after it has been coated. Because of the small magnetic field produced by the

6K7 type of heater, there is very little 60-cycle plate-current variation, or "hum," when the heater is operated on alternating current. The loop type of helical heater exemplified by the 25A6 heater and the folded type of heater used in the 25L6 cathode make possible the use of enough wire for 25-volt operation. The advantage of the 25L6 construction is its low cost. More complicated cathode structures used in arc-discharge tubes will be discussed in Chap. 12.

The advantages of heater-type cathodes over filamentary cathodes are the much lower alternating component of plate current resulting from a-c operation of the heaters, and the possibility of using a single source of power to heat a number of cathodes between which a difference of potential exists. The small effect of the alternating heating current upon the plate current results in part from the fact that the indirectly heated cathode is a unipotential surface, and in part from the small magnitude of the stray alternating fields caused by the flow of heater current. A disadvantage of indirectly heated cathodes is their much longer heating time. Because oxide-coated emitters do not stand up in high-voltage tubes and because indirect heating cannot be used with pure tungsten or thoriated tungsten emitters, heater-type cathodes are not used in tubes that require high plate voltage.

2-8. Effects of Gas upon Emission and Space Currents.—The deleterious effects of gas upon emitters of various types as the result of chemical action, adsorption of thin layers of gas upon the surface, and positive-ion bombardment have already been mentioned. When the anode voltage is sufficiently high to produce ionization of the gas, other effects become apparent. If anode current is at first limited by space charge, the appearance of positive ions tends to neutralize the negative space charge surrounding the filament, thus increasing the anode current. If the anode voltage is high enough to give saturation current initially, increase of current occurs because the electrons and ions produced by bombardment of neutral gas molecules by the emitted electrons add to the current. Unfortunately this increase of current is likely to be accompanied by a number of undesirable effects. Currents through ionized gases usually fluctuate, resulting in "noise" in tubes used for amplification. The relatively low velocity of positive and negative ions produces a lag in the response of current to changes of voltage. Variations of gas pressure resulting from changes of temperature or from the absorption or emission of gas from the walls and electrodes may cause the characteristics of the tube to vary. Finally, in a gassy tube, positive-ion current flows to an electrode to which a negative voltage is applied. When the anode current is controlled by means of a negative voltage applied to an electrode through a high resistance, the flow of current through the resistance may cause an objectionable voltage drop and, under certain circumstances, may even result in damage to the tube (see Secs. 6-7, 12-20, and 13-16).

In the manufacture of high-vacuum tubes, many precautions are taken to ensure the removal of gas from walls and electrodes. The electrodes are thoroughly cleaned and are then heated for several minutes in an atmosphere of hydrogen, which removes oxygen and water vapor. After the tube is assembled and connected to the pumps, the electrodes are heated to about 800 or 1000°C by high-frequency induction in order to remove other occluded gases. Residual gas is removed by the use of *getters*, which are active chemical substances such as barium, magnesium, aluminum, and tantalum, having the property of combining with gases when they are vaporized. In glass tubes a small amount of the getter is mounted in such a position that it will be heated and vaporized or "flashed" during the inductive heating of the elements. By proper location of the getter, the vapor can be prevented from condensing in places where it might cause electrical leakage or undesirable primary or secondary emission. The effectiveness of the getter results not only from its chemical combination with gases during flashing, but also from subsequent absorption of gases by getter that has condensed on the walls of the tube. The action of the getter during flashing is increased by ionization of the gas by means of voltages applied between the electrodes or by the radio-frequency field of the induction heater. To ensure the removal of gas from the walls, tubes are baked during the process of manufacture. On machines that exhaust and seal the tubes separately, the bulbs are heated in ovens during exhaustion. On "Sealex" machines, the tubes are sealed and exhausted on the same machine, the heat from sealing being used to drive gases from the bulbs during exhaustion.

In tubes with metal envelopes, the shielding action of the shell makes it impossible to heat the electrodes and the getter by induction. The electrodes may be heated by radiation from the shell, which is heated by gas flames. Although the getter may be fastened to the inside of the shell and vaporized by heating the shell locally, another method has been developed that requires less critical control.[1] A short length of tantalum ribbon, which connects the shell to its terminal pin in the base, is coated with a mixture of barium and strontium carbonates. While the tube is on the pumps, the temperature of the tantalum wire is raised electrically to about 1100°C. This converts the carbonates into oxides. After the tube has been sealed off, the tantalum wire is heated to a temperature in excess of 1200°C. This causes the tantalum to reduce the oxides to pure metallic barium and strontium, which vaporize. Since the vapor moves in straight lines, it can be directed as desired by means of shields

[1] LEDERER, E. A., and WAMSLEY, D. H., *RCA Rev.*, **2**, 117 (1937). This article also discusses gettering methods used in glass tubes.

and by proper location of the tantalum wire. This type of getter is called *batalum* [see (11) in Fig. 3-11*b*].

2-9. Limitation of Anode Current by Space Charge.—The effect of space charge in limiting space current and the increase of anode current resulting from an accelerating anode potential have already been mentioned in connection with the theory of thermionic emission. Before proceeding to a discussion of the quantitative relation between anode current and anode potential in a two-element tube, it is of interest to discuss further the physical picture underlying the phenomenon. The behavior of the emitted thermionic electrons is complicated by their initial velocities. For this reason it is best first to formulate a theory on the assumption that the initial velocities are zero and then, when they are taken into consideration, to see in what manner the results should be altered. For the present, therefore, initial velocities will be assumed to be zero. It will be further assumed that both cathode and anode are homogeneous, constant-potential, parallel planes of large area, and hence that the electric field over the surface of the cathode may be assumed to be uniform.

Electrons that leave the cathode constitute a space charge that exerts a retarding field at the cathode. The net field at the surface of the cathode is the difference between this retarding field and the accelerating field produced by the positive voltage of the anode. The number of electrons in the space, and hence the retarding component of field at the cathode, increase with the anode current. When the positive anode voltage is applied, the anode current builds up with great rapidity to such a value that the average retarding field at the cathode caused by the space charge is equal to the accelerating field caused by the anode voltage, making the average field zero at the cathode. Increase of emission then does not raise the anode current, as the additional emitted electrons merely reenter the cathode. If it were possible in some manner to increase the density of space charge by increasing the current beyond this equilibrium value, or if the anode voltage were reduced slightly, then the net field at the cathode would be a retarding one. For an instant, all emitted electrons would be prevented from moving away from the cathode, and the current and space-charge density would be automatically reduced to a value that would again make the average field at the cathode zero. An increase of anode voltage causes the accelerating field to exceed the retarding field. The number of electrons moving to the anode then increases until the retarding field again equals the accelerating field.

At first thought it may not seem plausible that there can be a steady flow of electrons to the anode when both the velocities of emitted electrons and the average electrostatic field are zero at the cathode. It is only the *time average* field, however, that is zero at any point on the cathode. The *instantaneous* field at any point may vary in a random manner between positive and negative values. Immediately after one or more electrons have entered some point of the anode, the net field at a corresponding point of the cathode may be positive, causing one or more electrons to move away from the cathode. These electrons produce a retarding field behind them, which prevents the departure of more electrons from that point until the entrance of other electrons into the anode again results in an accelerating field. Many electrons are entering and leaving the space at any instant, so that the field fluctuations are rapid and haphazard.

2-10. Child's Law.—The foregoing descriptive explanation shows that, if an ample supply of electrons is available at the cathode, the anode current in a diode varies with the voltage applied between the anode and cathode. A mathematical analysis of this phenomenon was first made by Child.[1] The equation relating the anode current and voltage is called *Child's law.* The general derivation for electrodes of any size and shape is too difficult to yield a useful equation, and so only the relatively simple cases such as those applying to plane parallel electrodes of large area and to long concentric cylinders are ordinarily considered. In deriving Child's law for plane parallel electrodes the following assumptions are made:

1. The cathode temperature is high enough at all points so that more electrons are emitted than are drawn to the anode; *i.e.*, the current is limited by space charge.
2. The cathode and anode are parallel plates whose area is large as compared to their spacing; *i.e.*, the electrostatic field is normal to the electrode surface and uniform over the surface of any plane parallel to the electrodes.
3. The surfaces of the anode and cathode are equipotential surfaces.
4. The space between the cathode and the anode is sufficiently free of gas so that electrons do not lose energy by collision with gas molecules in moving from the cathode to the anode.
5. Emitted electrons have zero initial velocity after emission.

Under these assumptions the following three equations may be written:

$$\frac{d^2V}{dx^2} = -4\pi\rho \tag{2-2}$$

$$\mathcal{E}V = \tfrac{1}{2}m_e v^2 \tag{2-3}$$

$$i_b = -\rho v A \tag{2-4}$$

in which V is the potential, relative to the cathode, at a distance x from the cathode; ρ is the (negative) density of electron space charge at a

[1] CHILD, C. D., *Phys. Rev.*, **32**, 498 (1911).

distance x from the cathode; $\mathcal{E}$ and m_e are the charge and the mass, respectively, of an electron; v is the velocity dx/dt of an electron at a distance x from the cathode; i_b is the anode current; and A is the area of the electrodes.

Equation (2-2) combines in symbolic form the definitions of potential difference and electric field. It is a special form of Poisson's equation, one of the most important fundamental laws of electrostatics, and may be derived directly from Gauss's law[1] (see Sec. 1-10). Equation (2-3) states that the energy gained by an electron in moving from the cathode to a distance x from the cathode under the influence of the electric field appears entirely in the form of kinetic energy. Equation (2-4) is a symbolic formulation of the definition of the magnitude of an electric current as the rate of flow of charge.

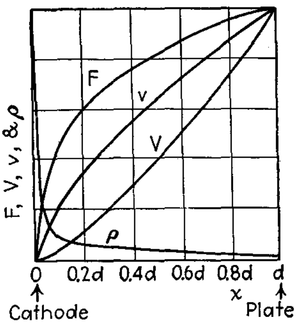


FIG. 2-10.—Variation of field strength F, potential V, electron velocity v, and space-charge density ρ with distance x from plane cathode. Zero initial velocity. Arbitrary units.

In the solution of the simultaneous differential Eqs. (2-2), (2-3), and (2-4), the following boundary conditions must be applied: At the cathode, the potential V, the average electric field $\partial V/\partial x$, and the velocity v are zero. At the anode, where x is equal to the cathode-to-anode spacing d, the potential is equal to e_b, the applied anode voltage. Solution of the equations and substitution of numerical values of $\mathcal{E}$ and m_e give the following equation for the anode current of a diode:[2]

$$i_b = 2.34 \times 10^{-6} \frac{Ae_b^{3/2}}{d^2} \qquad \text{amp} \qquad (2\text{-}5)$$

By combining Eq. (2-5) with Eqs. (2-2), (2-3), and (2-4), theoretical expressions may be derived for density of space charge, electron velocity, electric field strength, and potential as functions of distance from the cathode. Curves derived from these are shown in Fig. 2-10.

Child's law for concentric cylinders whose length is large as compared to their spacing is

$$i_b = 14.68 \times 10^{-6} \frac{he_b^{3/2}}{br} \qquad \text{amp} \qquad (2\text{-}6)$$

in which r is the radius of the anode, h is the length of the electrodes, and b is a factor whose value depends upon the ratio of the radius of the anode to that of the cathode. b has the approximate value $\frac{1}{4}$ for a ratio 2, $\frac{1}{2}$ for a ratio 3, and 0.9 for a ratio 8. If the plate diameter is large as compared

[1] PAGE, L., and ADAMS, N. I., "Principles of Electricity," p. 83, D. Van Nostrand Company, Inc., New York, 1931.

[2] PAGE and ADAMS, *op. cit.*, p. 297.

to that of the cathode, Eq. (2-6) reduces to the approximate form

$$i_b = 14.7 \times 10^{-6} \frac{he_b^{3/2}}{r} \qquad \text{amp} \tag{2-7}$$

Equations (2-5) to (2-7) show the importance of close spacing between cathode and anode if large currents are desired at small anode voltages.

2-11. Deviations from Child's Law Observed in Practical Diodes.—Deviations from Child's law result from the failure of practical diodes to satisfy the assumptions made in its derivation. Since the temperature of the cathode is fixed by considerations of emission efficiency and life, there is always a saturation voltage above which the current is not limited by space charge but by filament emission. If other assumptions were satisfied, the saturation voltage would be quite definite and the current-voltage curve would be as shown by the dotted lines of Fig. 2-11. Because of variations in temperature, electron affinity, and field strength over the cathode surface, the anode voltage at which voltage saturation takes place is not the same for all points of the cathode. The curve of anode current *vs.* anode voltage therefore bends over gradually, as shown by the full line of Fig. 2-11. Above saturation the current is not entirely constant but continues to rise somewhat with anode voltage. This is explained by reduction of electron affinity with increase of external field (see Sec. 1-12), and lack of homogeneity of the surface of the emitter. The effect is particularly noticeable with oxide-coated cathodes.

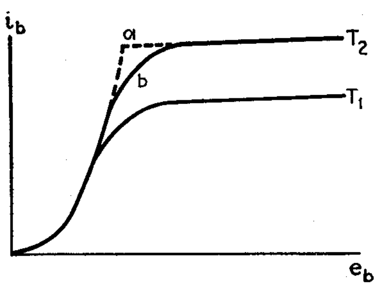


FIG. 2-11.—Curves of anode current i_b *vs.* anode voltage e_b at two values of emitter temperature.

The assumptions of uniform field and equipotential cathode are satisfied fairly closely in heater-type diodes with cylindrical plates. The voltage drop in filamentary cathodes may be shown to change Child's $\frac{3}{2}$-power law into a $\frac{5}{2}$-power law at anode voltages relative to the negative end of the filament that are less than the voltage of the positive end of the filament. This tends to make the lower part of the i_b-e_b curve steeper. The exact effect upon the curve of failure to satisfy the assumption of uniform field is complicated and impossible to predict completely.

2-12. Effect of Initial Velocities of Emitted Electrons.—Modified forms of Child's law which take into consideration the initial velocities of emitted electrons have been derived by Schottky, Langmuir, and others.[1] For the purpose of this book, a qualitative explanation of the effect of initial velocities is sufficient. Let it first be supposed that

[1] SCHOTTKY, W., *Physik. Z.*, **15**, 526 (1914); *Ann. Physik*, **44**, 1011 (1914); LANGMUIR, I., *Phys. Rev.*, **21**, 419 (1923); DAVISSON, C., *Phys. Rev.*, **25**, 808 (1925).

the electrons emerge with zero velocity. Under equilibrium conditions the space current and the space-charge density assume such values that the *average* field at the cathode is zero (see Sec. 2-9). The field and potential distributions in the interelectrode space are as shown in Fig. 2-10. Now let the emitted electrons suddenly have initial velocities which, for the sake of simplicity, are assumed to be the same for all electrons. Electrons that, without initial velocity, would have reentered the cathode now move toward the anode in spite of the fact that the average field is zero. As a result, the current and the space-charge density increase. The retarding field of the space charge now exceeds the accelerating field of the anode, giving a net retarding field at the cathode surface which slows up the electrons in the vicinity of the cathode. Equilibrium results when the retarding field in the vicinity of the cathode is sufficiently high so that the electrons are brought to rest in a plane a short distance s from the cathode. The average field in this plane is zero, but instantaneous fluctuations allow just enough electrons to pass to give the required anode current. The behavior is similar to that which would obtain if the initial velocity were zero and the cathode were moved toward the anode by the distance s. Because of the random distribution of electron velocities the phenomenon is actually more complicated than this simplified picture indicates. The simple theory shows, however, that the effect of initial velocities is to increase the anode current corresponding to any anode voltage and is, therefore, equivalent to that of a small increase of anode voltage. Because the electrons emerge with velocities corresponding to kinetic energy of the order of a volt or less, the effect is appreciable only for low anode voltages. The field and potential distributions throughout the interelectrode space are plotted in Fig. 2-12. That the potential must pass through a minimum where the field is zero follows from the fact that the field at any point may be expressed as the space derivative of the potential at that point.

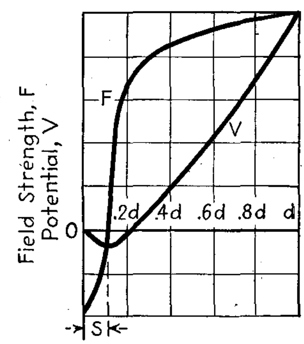


FIG. 2-12.—Variation of field strength and potential with distance from cathode for plane parallel electrodes of separation d. Initial velocities of emitted electrons considered.

The lower part of the i_b-e_b curve of Fig. 2-11 is raised slightly as the result of initial velocities of emitted electrons, and a small negative voltage must be applied in order to reduce the anode current to zero. Theoretical equations relating anode current and voltage at negative anode voltages have been derived by Schottky[1] and Davisson.[2] Because

[1] SCHOTTKY, *loc. cit.;* LANGMUIR, *loc. cit.*

[2] DAVISSON, *loc. cit.*

of their complicated form and because of failure to satisfy in practice the assumptions made in their derivation, these are seldom of great practical value. At negative anode voltages that are high enough to reduce the anode current to the order of 50 μa or less, the anode current of diodes with unipotential cathodes follows closely the empirical relation

$$i_b = k_1 \epsilon^{k_2 e_b} \tag{2-8}$$

in which k_1 and k_2 are constants for a given tube. The current departs materially from this exponential law as the negative anode voltage is reduced in magnitude in the vicinity of zero voltage, particularly at high cathode temperatures. Experimental curves corresponding to Eq. (2-8) were first obtained by Germer.[1]

Unless the plate of a highly evacuated diode becomes hot enough to emit electrons, increase of negative anode voltage beyond the value that reduces the current to zero has no further effect upon the anode current. The fact that anode current flows in one direction only is employed in the application of diodes to detection and to power rectification, which will be discussed in Chaps. 9 and 14.

2-13. Relation of Richardson's and Child's Laws.—It should be noted that Richardson's equation and Child's law apply to two different conditions of operation of two-element tubes. Richardson's equation holds only under voltage saturation, whereas Child's law applies only under temperature saturation. In most applications, vacuum tubes are used in such a manner that temperature saturation prevails.

2-14. Shot Effect.—The random motion of electrons causes rapid variations of the number of electrons that pass from the cathode to the anode in unit time, and thus produces fluctuations of anode current. This phenomenon, which may be readily detected by the use of sufficient amplification, is called the *shot effect.* It is one of the factors that limit amplification by vacuum tubes (see Sec. 6-28).

Heating of the Plate.—The kinetic energy acquired by electrons in moving from the cathode to the plate is converted into heat when the electrons strike the plate. The average current that a vacuum tube can pass is limited by the temperature of the plate at which absorbed gas is driven out of the plate or electron emission takes place from the plate. The energy that is converted into heat at the plate is equal to the time integral of the product of the plate current and plate voltage. Radiation of heat from the plate is increased by blackening its outer surface.

2-15. Classification of Tubes.—Electron tubes may be classified in a number of ways. These classifications include those based upon the process involved in the emission of electrons from the cathode, the degree of evacuation, the number of electrodes, and the type of application.

[1] GERMER, L. H., *Phys. Rev.*, **25**, 795 (1925).

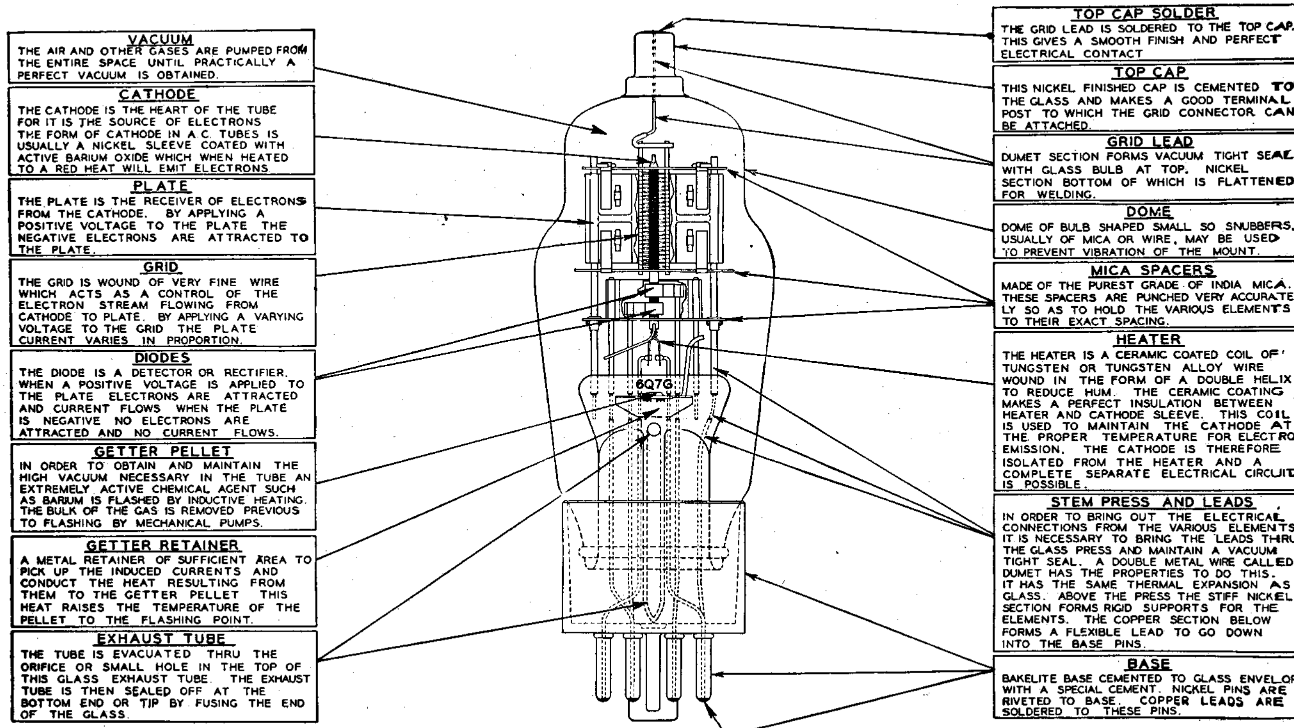


FIG. 2-13.—Structure of typical glass receiving tube. *(Courtesy of Ken-rad Tube and Lamp Corp.)*

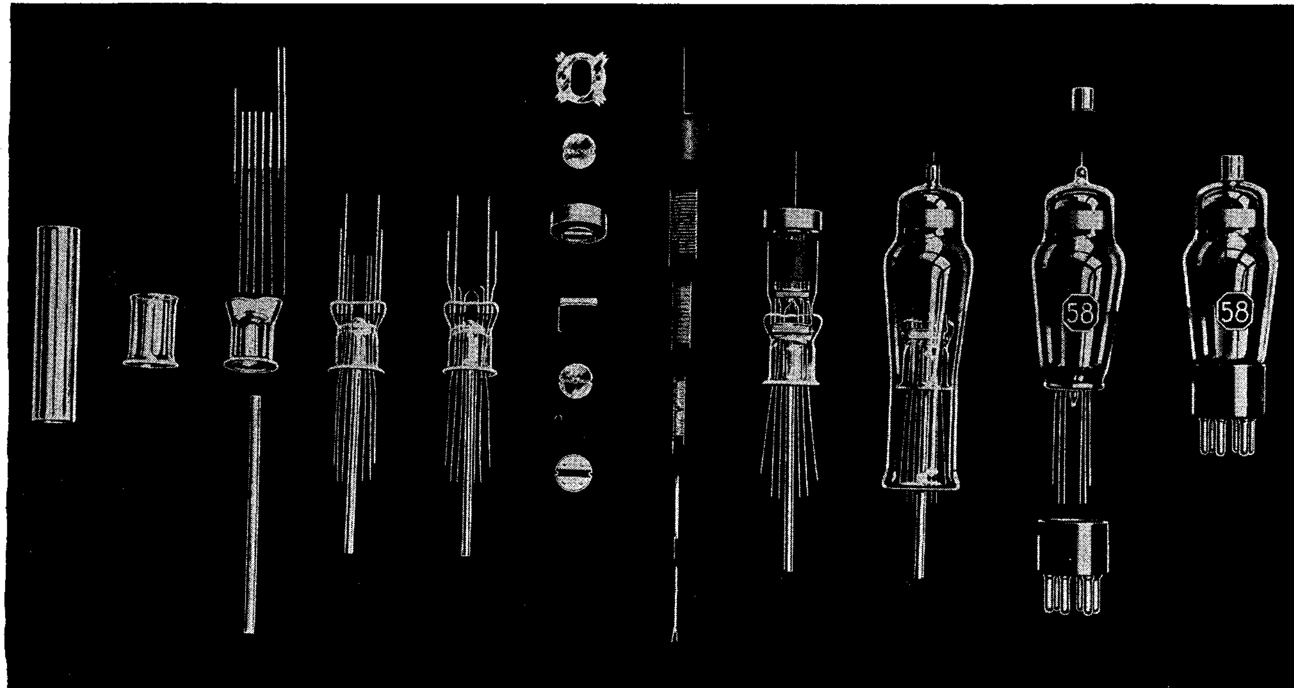


Fig. 2-14.—Assembly of typical glass receiving tube. (*Courtesy of Radio Corporation of America.*)

Included in the first classification are the thermionic tube and the phototube. A *thermionic tube* is an electron tube in which the electron or ion

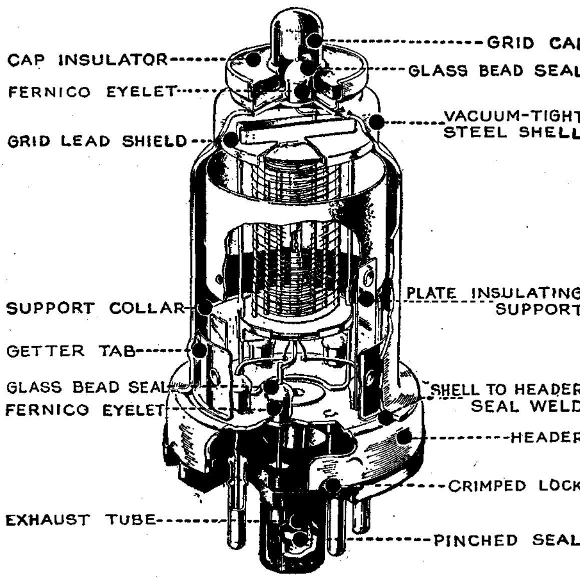


FIG. 2-15.—Structure of typical metal receiving tube.

emission is produced by the heating of an electrode. A *phototube* is an electron tube in which electron emission is produced directly by radiation falling upon an electrode. According to degree of evacuation, electron

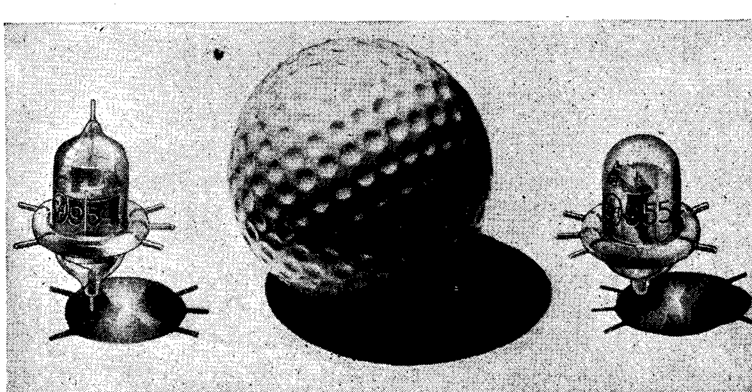


FIG. 2-16.—Size of typical "acorn" tubes shown in comparison with a golf ball. (*Courtesy of Radio Corporation of America.*)

tubes are classified as high-vacuum tubes and gas- or vapor-filled tubes. A *high-vacuum tube* (*vacuum tube* or *pliotron*) is an electron tube evacuated to such a degree that its electrical characteristics are essentially unaffected by gaseous ionization. A *gas-filled* or *vapor-filled tube* (*gas tube*) is an electron tube in which the pressure of the gas or vapor is such as to affect appreciably the electrical characteristics of the tube. According to the number of electrodes, tubes are classified as *diodes*, *triodes*, *pentodes*, etc. For convenience or economy, or for reduction of space or weight, two or more sets of elements may be enclosed in a single envelope. Thus, there are duplex (double) triodes, duplex-diode pentodes, triode pentodes, etc. The diverse classification of electron tubes according to application will be made in later chapters.

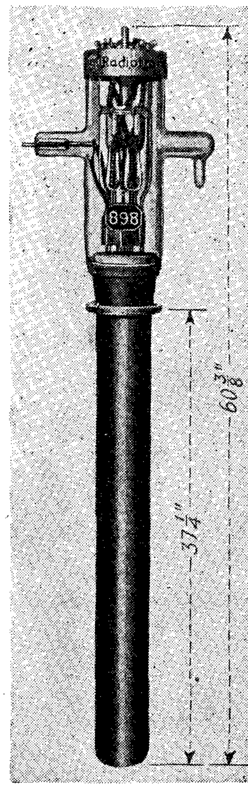


FIG. 2-17.—One-hundred-kilowatt water-cooled transmitting tube. (*Courtesy of Radio Corporation of America.*)

2-16. Structure of Tubes.—Tubes are made with both glass and metal envelopes.[1] The principal advantages of metal tubes lie in their greater mechanical strength and in the fact that the electrodes are permanently and completely shielded without the use of an external shield. Furthermore, they do not require on the inside of the envelope the conducting coating that must be used in glass tubes to prevent the wall from acquiring a positive charge as the result of secondary emission caused by the impact of electrons that pass around the plate. A disadvantage of metal tubes is that the shells become so hot in operation that they cannot be conveniently handled. This is of importance in the routine factory testing of radio receivers. Another minor disadvantage is the impossibility of determining visually whether the heater is in operation. Glass tubes appear to be somewhat more reliable. In rectifiers, particularly, metal tubes are likely to give difficulty as the result of short circuits. Figures 2-13 and 2-14 show the construction of a glass receiving tube; Figs. 2-15 and 3-11*b* show typical metal receiving tubes.

The great range in size of vacuum tubes is illustrated by Figs. 2-16 and 2-17. Figure 2-16 shows a typical *acorn* tube, developed for use at very high frequencies, at which it is essential to keep lead capacitance and inductance as small as possible.[2] No base is used on the acorn type of tube, connections being made directly to the electrode leads. Figure

[1] PIKE, O. W., and METCALF, G. F., *Electronics*, October, 1934, p. 312. See also *Electronics*, September, 1935, p. 31.

[2] SALZBERG, B., and BURNSIDE, D. G., *Proc. I.R.E.*, **23**, 1142 (1935).

2-17 shows a 100-kw water-cooled transmitting tube of a type that is used in large broadcasting stations.

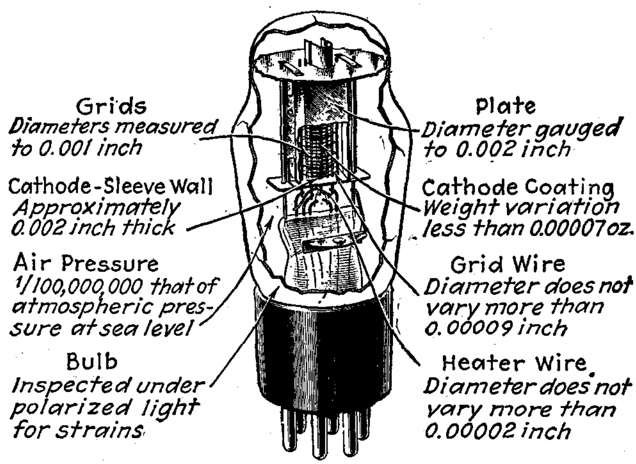


FIG. 2-18.—Materials used in typical radio receiving tubes. The complex nature of the structure of the modern vacuum tube and of the manufacturing processes are well illustrated by a consideration of the materials that are used.

Gases.—Argon, carbon dioxide, chlorine, helium, hydrogen, illuminating gas, neon, nitrogen, and oxygen.

Metals and Compounds.—Alumina, aluminum, ammonium chloride, arsenic trioxide, barium, barium carbonate, barium nitrate, borax, boron, cesium, calcium, calcium aluminum fluoride, calcium carbonate, calcium oxide, carbon, chromium, cobalt, cobalt oxide, copper, iridium, iron, lead, lead acetate, lead oxide, magnesia, magnesium, mercury, misch metal, molybdenum, monel, nickel, phosphorus, platinum, potassium, potassium carbonate, silica, silicon, silver, silver oxide, sodium, sodium carbonate, sodium nitrate, tantalum, thorium, thorium nitrate, tin, titanium, tungsten, zinc, zinc chloride, and zinc oxide.

Accessories.—Bakelite, ethyl alcohol, glass, glycerine, isolantite, lava, malachite green, marble dust, mica, nigrosine, petroleum jelly, porcelain, rosin, shellac, synthetic resin, and wood fiber.

Bibliography

RICHARDSON, O. W.: "Emission of Electricity from Hot Bodies," Longmans, Green & Company, New York, 1916.

DUSHMAN, S.: Thermionic Emission, *Rev. Modern Phys.*, **2,** 381 (1930).

COMPTON, K. T., and LANGMUIR, I.: *Rev. Modern Phys.*, **3,** 191 (1931).

RIEMANN, A. L.: "Thermionic Emission," John Wiley & Sons, Inc., New York, 1934.

STILES, W. S.: *Dept. Sci. Ind. Research (Brit.), Special Rept.* 11, London, 1932.

DUSHMAN, S.: Electron Emission, *Elec. Eng.*, **53,** 1054 (1934).

CHAFFEE, E. L.: "Theory of Thermionic Vacuum Tubes," Chaps. IV, V, McGraw-Hill Book Company, Inc., New York, 1933.

KOLLER, L. R.: "The Physics of Electron Tubes," 2d ed., Chaps. I-IV, McGraw-Hill Book Company, Inc., New York, 1937.

MCARTHUR, E. D.: "Electronics and Electron Tubes," Chap. III, John Wiley & Sons, Inc., New York, 1936.

CHAPTER 3

GRID-CONTROLLED HIGH-VACUUM TUBES

3-1. The greatest single advance in the development of vacuum tubes undoubtedly came with the introduction by De Forest of a control electrode between the cathode and the plate of the diode.[1] The principal value of such a control electrode arises from the fact that relatively large plate current and power may be controlled by small variations of voltage of the control electrode relative to the cathode without the expenditure of appreciable power in the control circuit. A three-electrode vacuum tube containing an anode, a cathode, and a control electrode is called a *triode*. Figure 3-1 shows the cross sections of typical high-vacuum triodes.

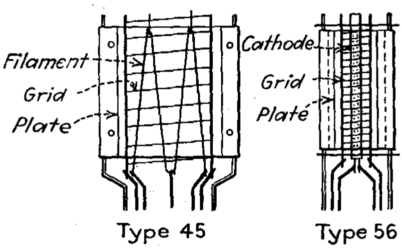


FIG. 3-1.—Electrode structure of typical filamentary and heater-type triode receiving tubes.

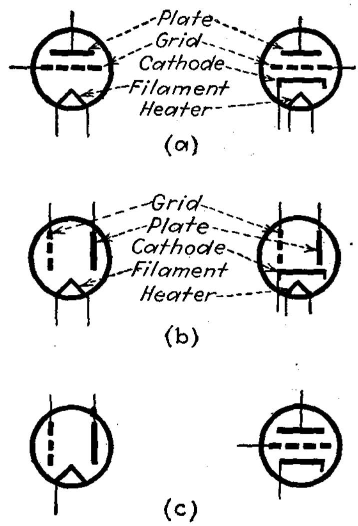


FIG. 3-2.—Symbols for filamentary and heater-type high-vacuum triodes.

The form of the control electrode in early tubes led to the use of the term *grid* for this electrode. A grid is now defined more broadly as an electrode that contains openings through which electrons or ions may pass. Many vacuum tubes contain two or more grids. In numerous applications of multigrid tubes the voltages of all but one grid are kept constant. This grid, called the *control grid,* serves to vary the plate (or other electrode) current by means of changes of voltage applied to it. The behavior of multigrid tubes with all grid voltages but that of the control grid fixed is in many respects similar to that of a triode.

Standard symbols for filamentary and heater-type triodes are shown

[1] DE FOREST, LEE, U. S. Patent 841387 (1907); U. S. Patent 879532 (1908).

in Fig. 3-2*a*.[1] Often circuit diagrams can be simplified by the use of the modified forms of Fig. 3-2*b*. In most circuit diagrams the filament and heater connections are of secondary importance and may be omitted. Whenever possible, therefore, simplified symbols such as those of Fig. 3-2*c* will be used. Except for the omission of the grid, diode symbols are the same as those used for triodes. Figure 14-20 contains the symbol used for a diode with two anodes.

3-2. Theory of Grid Action in Triodes. Equations for Plate and Grid Currents.—Electrodynamic analyses of the action of the grid in

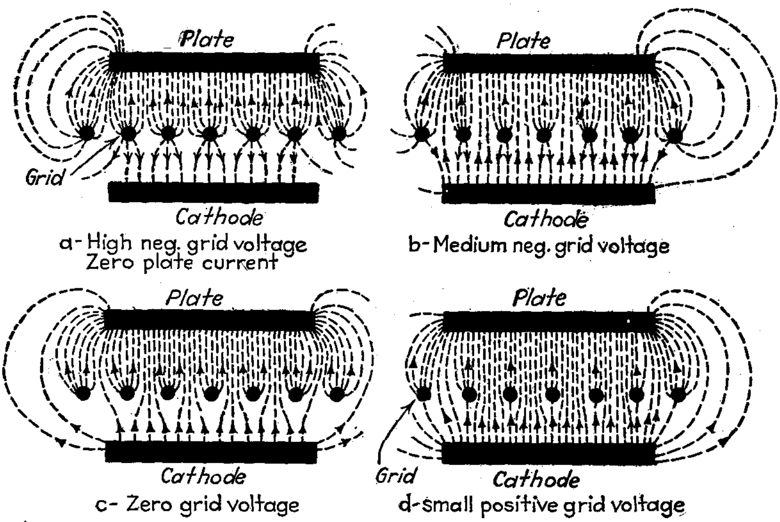


FIG. 3-3.—Approximate field distribution resulting from applied voltages in a triode with plane parallel electrodes for fixed positive plate voltage and four values of grid voltage. Arrows indicate the direction of the force on an electron.

controlling the plate current of a triode have been developed.[2] For the purposes of this book it will be better to present a brief qualitative discussion of the phenomenon of grid control and to base subsequent derivations upon empirically determined facts.

In Fig. 3-3 is shown the approximate field distribution resulting from applied electrode voltages in a triode with plane cathode and anode. Lines are used in the customary manner to indicate the electrostatic field, but, contrary to convention, the arrows indicate the direction in which electrons are urged by the field.[3] The plate voltage is assumed to be posi-

[1] "Standards on Electronics," p. 15, Institute of Radio Engineers, New York, 1938.

[2] See, for instance, E. L. CHAFFEE, "Theory of Thermionic Vacuum Tubes," Chap. VII, McGraw-Hill Book Company, Inc., New York, 1933.

[3] No special effort has been made in Fig. 3-3 to depict the field distribution accurately. Figure 3-3 is derived from more complete and carefully constructed diagrams shown on pp. 175 and 176 of "Theory of Thermionic Vacuum Tubes," by E. L. Chaffee.

tive and constant in value. If the grid is made sufficiently negative with respect to the cathode, all lines of force terminate on the grid, and no field exists directly between the plate and the cathode. This is illustrated in Fig. 3-3*a*. The field at all points of the cathode is in the direction to return emitted electrons to the cathode and so the plate current is zero. If the grid is slightly less negative, some field will extend directly from the anode to the cathode, as shown in Fig. 3-3*b*, and there will be a force tending to carry electrons from certain points of the cathode to the anode. Further decrease of negative grid potential increases the areas of the cathode over which the field tends to remove electrons and strengthens the average field over these areas. As long as the grid is negative with respect to the cathode, electrons can reach the grid only if they have sufficient kinetic energy to overcome the retarding field terminating on the grid. Because the initial velocity of emission of the electrons is small, electron current to the grid is zero unless the negative grid voltage is appreciably less than a volt. If the tube contains traces of gas, there may be a small positive-ion current to the grid when the grid is negative. At zero grid potential, illustrated by Fig. 3-3*c*, no point of the cathode experiences a retarding field. Initial velocities cause some electrons to strike the grid, giving a small grid current. When the grid is positive, as in Fig. 3-3*d*, there is an accelerating field over the whole surface of the cathode. A portion of this field terminates on the grid, causing appreciable grid current to flow.

As in the diode, the net field at the cathode is actually the resultant of the retarding field produced by the space charge and the field caused by the electrode voltages. Equilibrium is established when the average net field is zero at a short distance from the cathode. Increase of applied accelerating field causes the net field to become positive and thus allows more electrons to go to the plate. The space current and space charge increase until the average net field is again zero a short distance from the cathode.

Since the field at the cathode depends upon the potential both of the grid and of the plate, the plate current is a function of both grid and plate voltages. This may be stated symbolically by the functional equation

$$i_b = f(e_b, e_c) \tag{3-1}$$

in which i_b is the plate current, e_b is the plate voltage, and e_c is the grid voltage. Because of the screening action of the grid, only a portion of the field from the plate extends directly to the cathode, whereas there is nothing to intercept the field between the grid and the cathode. One is led to guess, therefore, that the plate current is affected more by changes of grid voltage than of plate voltage, *i.e.*, that the grid voltage is μ times as effective as the plate voltage in controlling the plate current, μ

being a factor greater than unity. μ is not necessarily constant. The effect of initial velocity of emitted electrons is the same as though a small increase were made in either grid or plate potential. The contact potentials may either increase or decrease the effective field at the cathode. It is convenient to combine the effects of initial velocity and contact potentials into a single quantity ϵ, an equivalent voltage that would produce the same effect upon plate current as the initial velocity plus the contact potentials. ϵ is ordinarily so small in comparison with externally applied potentials that it may be neglected. This analysis leads to the assumption that the plate current is a function of $(e_b + \mu e_c + \epsilon)$. This fact may be expressed by the functional equation

$$i_b = F(e_b + \mu e_c + \epsilon) \qquad (T_f = \text{const.}) \tag{3-2}$$

where T_f is the temperature of the cathode.

The use of Eq. (3-2) is justified by the fact that it is verified experimentally. Sometimes it is possible to make use of the more explicit approximate law

$$i_b = A(e_b + \mu e_c)^n \tag{3-3}$$

in which A is a constant. The exponent n varies considerably with grid and plate voltages, the values ranging roughly between 1.2 and 2.5 for negative values of e_c. When either grid or plate voltage is maintained constant, however, the variation of n is often so small that n may be assumed to be constant over certain ranges. In some analyses, n is assumed to be equal to 1.5 when $e_c = 0$, although actual values may depart appreciably from this value.

μ is called the *plate amplification factor*, or simply the *amplification factor* of the tube. It is a measure of the relative effectiveness of the grid and plate voltages in controlling the plate current. The amplification factor will be shown to be related to certain of the characteristic curves of a triode, and it will be defined mathematically on the basis of this relationship (see Sec. 3-5). The value of μ depends upon the shape and spacing of the electrodes,[1] and to some extent upon the plate current; it may also be made to vary with electrode voltages (see Sec. 3-7). Electrodynamic analysis, based upon the assumption that the electrodes are parallel and of infinite extent, that the grid wire spacing is large compared to the diameter of the grid wire, and that the space charge between electrodes is zero, yields the following approximate formula for μ:[2]

$$\mu = \frac{2\pi p}{a \log_\epsilon \left(2 \sin \dfrac{\pi r}{a}\right)} \tag{3-4}$$

[1] KUSUNOSE, Y., *Proc. I.R.E.*, **17**, 1706 (1929).

[2] SCHOTTKY, W., *Arch. Elektrotech.*, **8**, 21 (1919); SALZBERG, B., *Proc. I.R.E.*, **30**, 134 (1942).

in which p is the distance between the planes of the grid and the plate, a is the distance between adjacent grid wires, and r is the radius of the grid wire. The following approximate expression for μ applies to an electrode structure in which the anode and the cathode are long cylinders and the grid is helical.[1]

$$\mu = \frac{2\pi \, \rho_g/\rho_p \, (\rho_p - \rho_g)}{a \log_\epsilon a/2\pi r} \tag{3-5}$$

in which ρ_p and ρ_g are the radii of the anode and grid, respectively, and a is the spacing of the grid wires. Because of factors neglected in the derivation of Eqs. (3-4) and (3-5) and also because modern tube design is based largely upon empirical methods, the value of these equations lies mainly in the indication that they give of the general effects of triode tube dimensions upon the amplification factor.

The effect of space charge is to cause a variation of amplification factor with plate current or electrode voltages. Except in the case of "variable-mu" tubes (see Sec. 3-7), in which the grid wires are not equally spaced, the variation of μ with electrode voltages is sufficiently small so that μ may often be assumed to be constant over the working range of current and voltages. At positive grid voltages, however, diversion of current from the plate to the grid lowers μ.

The grid current of a triode is also a function of the grid and plate voltages. This fact may be expressed by the functional equation

$$i_c = G(e_c + \mu_g e_b + \epsilon) \qquad (T_f = \text{const.}) \tag{3-6}$$

in which μ_g, the *grid amplification factor*, is less than unity.

3-3. Time of Transit of Electrons.—Because of its small mass, the acceleration of an electron is so rapid that the time taken for electrons to pass from the cathode to the plate may usually be neglected and the response of electrode currents to changes of electrode voltages considered to be instantaneous. At the very high frequencies corresponding to wave lengths of the order of a few meters or less, however, time of transit must be taken into consideration.[2] Since the operation of vacuum tubes at ultrahigh frequencies will not be discussed in detail in this book, the time of transit of electrons will usually be neglected.

3-4. Static and Dynamic Characteristics.—Theoretical and practical studies of the performance of vacuum tubes and vacuum-tube circuits are

[1] ABRAHAM, H., *Arch. Elektrotech.*, **8**, 42 (1919); KING, R. W., *Phys. Rev.*, **15**, 256 (1920).

[2] BENHAM, W. E., *Phil. Mag.*, **5**, 641 (1928); LLEWELLYN, F. B., *Proc. I.R.E.;* **21**, 1532 (1933); **22**, 947 (1934); **23**, 112 (1935); CHAFFEE, J. G., *Proc. I.R.E.*, **22**, 1009 (1934); FERRIS, W. R., *Proc. I.R.E.*, **24**, 82, 105 (1936); NORTH, D. O., *Proc. I.R.E.*, **24**, 108 (1936); BRAINERD, J. G., KOEHLER, G., REICH, H. J., and WOODRUFF, L. F., "Ultra-high-frequency Techniques," D. Van Nostrand Company, Inc., New York, 1942.

greatly facilitated by the use of curves relating the electrode currents and voltages, called *characteristics.* A *static electrode characteristic* is a relation, usually shown by a graph, between the voltage and the current of that electrode, other electrode voltages being maintained constant. A *static transfer characteristic* is a relation, usually shown by a graph, between the voltage of one electrode and the current of another electrode, all other voltages being maintained constant. Unless otherwise specified, the term *transfer characteristic* is understood to apply to characteristics relating control-grid voltage and plate current, which are the most frequently used transfer characteristics.

Strictly, a *static* characteristic is one obtained with steady voltages, whereas a *dynamic* characteristic is one obtained with alternating voltages. Inasmuch as all voltages but one are specified to be constant in the above definitions, the characteristics obtained with alternating voltages differ from those obtained with direct voltages only when the frequency is so high that tube capacitances and electron transit time cause appreciable out-of-phase components of current. The term *dynamic transfer characteristic* has come to be applied to a transfer characteristic obtained with alternating control-grid voltage when the electrode current under consideration passes through an external impedance, called the *load impedance,* the supply voltage for that current being maintained constant. Voltage drops in the load impedance cause the electrode voltage to differ from the supply voltage, and the electrode voltage to vary with current. In general, *IR* drop in the load causes the transfer characteristic to be affected by load even when the characteristic is derived by using steady voltages. Extension of the term *dynamic transfer characteristic* to include such a characteristic may be justified by considering it to be a limiting curve obtained as the frequency of alternating voltage is made to approach zero.

There are four sets of static characteristics of triodes. They are the plate characteristics i_b *vs.* e_b at constant values of e_c; the grid characteristics i_c *vs.* e_c at constant values of e_b; the grid-plate transfer characteristics i_b *vs.* e_c at constant values of e_b; and the plate-grid transfer characteristics i_c *vs.* e_b at constant values of e_c.[1] The behavior of a triode is completely specified by either the plate and grid families of characteristics or the two families of transfer characteristics.

[1] Since the static characteristics are constructed by plotting corresponding values of direct voltages and currents, the letters used in representing these voltages and currents should be capitals to be entirely in accord with the system of nomenclature used in this book (Sec. 3-17). In most applications of the characteristic curves, however, the currents and voltages are assumed to vary, and so must be represented by lower-case symbols. For this reason, lower-case symbols have been used in Eqs. (3-1) to (3-6) and will be used for all characteristic curves.

In Figs. 3-4*a* and 3-4*b* are shown typical characteristics for a triode. The bending of the i_b-e_c curves of Fig. 3-4*b* at positive values of e_c, particularly noticeable for low values of plate voltage, is caused by diver-

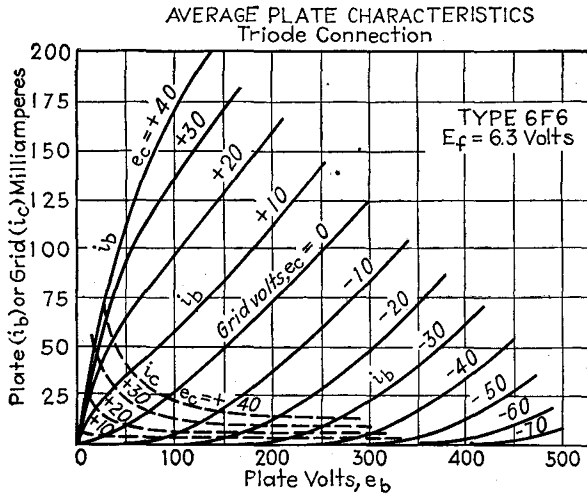


FIG. 3-4*a*.—Typical triode plate characteristics and plate-grid transfer characteristics.

sion to the grid of electrons emitted by the cathode, and by secondary emission. Secondary electrons emitted by the plate when the grid voltage exceeds the plate voltage are drawn to the grid and constitute a current opposite in direction to the normal plate current. Typical curves of i_b and i_c *vs.* e_b at large positive grid voltages are shown in Fig. 3-4*c*. The reversal of curvature of the plate characteristics of Fig. 3-4*c* between the plate voltages of 50 and 100 for grid voltages of 45 and above is also the result of secondary emission.

The points at which the characteristic curves intercept the voltage axes are called *cutoff points*, and the corresponding voltages, the *cutoff voltages*. An approximate relation between grid and plate voltages at plate current cutoff can be derived from Eq. (3-3). When $i_b = 0$, $(e_b + \mu e_c)^n = 0$. This can be true only if $e_b = -\mu e_c$. Because μ is never strictly constant, the accuracy of this relation depends upon the voltages at which μ is evaluated. Considerable error may result when the usual published value of μ is used, but the relation is often useful in making a

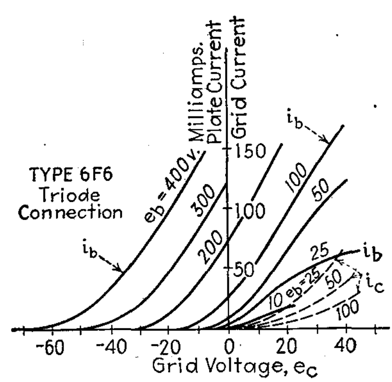


FIG. 3-4*b*.—Typical triode grid-plate transfer characteristics and grid characteristics (same tube as for Fig. 3-4*a*).

rapid approximate determination of cutoff voltages. Because of lack of homogeneity of the emitter, voltage drops in the cathode, variation of the electric field at various points of the cathode, and distribution of initial velocities of emission, the intersections of the characteristic curves with the voltage axis at cutoff are not sharp. Although some grid current usually flows at negative grid voltages lower than half a volt to a volt, the simplifying assumption is often made that grid current cutoff occurs at zero grid voltage.

The transfer characteristics of Fig. 3-4*b* may be derived from the plate characteristics of Fig. 3-4*a*, and vice versa. In Chap. 4 it will be shown that much essential information concerning the performance of

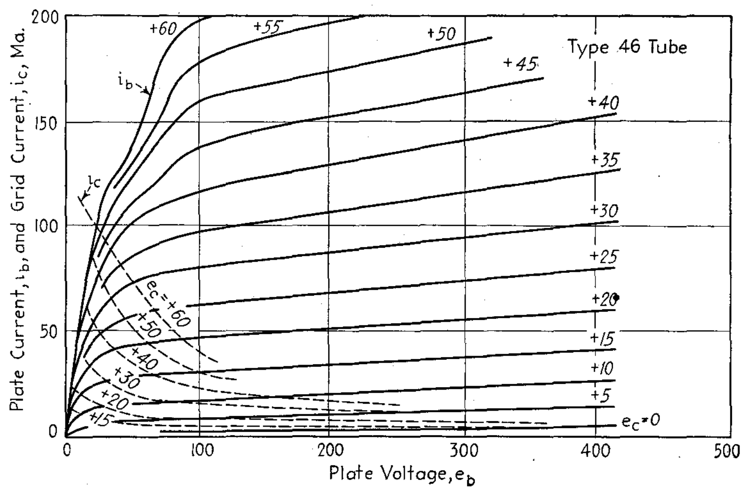


FIG. 3-4*c*.—Typical triode plate characteristics and plate-grid transfer characteristics at high positive grid voltages.

tubes in circuits may be obtained from the plate characteristics. It is important to note that in tubes with filamentary cathodes the grid and plate voltages are measured with respect to the negative end of the filament.

The curves of plate and grid current against cathode temperature of a triode are similar to those of plate current against cathode temperature of a diode, but voltage at which saturation becomes apparent depends upon both plate voltage and grid voltage. Since triodes are almost always operated above temperature saturation, these curves are of comparatively little practical value in connection with ordinary applications.

3-5. Tube Factors.—The mathematical and graphical analyses of the operation of vacuum tubes and vacuum-tube circuits require the use of certain tube *factors* whose numerical values are dependent upon the construction of the tube and upon the electrode voltages and currents, and which serve as indices of the ability of given tubes to perform specific functions. Although only triodes have been discussed so far; it will be convenient also to present at this point general definitions that apply to tubes with more than three electrodes.

Mu-factor is the ratio of the change in one electrode voltage to the change in another electrode voltage, under the conditions that a specified current remains unchanged and that all other electrode voltages are maintained constant. It is a measure of the relative effect of the voltages of two electrodes upon the current in the circuit of any specified electrode. As most precisely used, the term refers to infinitesimal changes. Symbolically, mu-factor is defined by the equation

$$\mu_{jkl} = -\frac{\partial e_j}{\partial e_k} \qquad (i_l,\ e_l,\ e_m,\ \text{etc.},\ = \text{const.}) \tag{3-7}$$ [1]

The most important mu-factor is the control-grid–plate mu-factor, or *amplification factor.* Amplification factor is the ratio of the change in plate voltage to a change in control-grid voltage under the conditions that the plate current remains unchanged and that all other electrode voltages are maintained constant. It is a measure of the effectiveness of the control-grid voltage, relative to that of the plate voltage, upon the plate current. The sign is taken as positive when the voltage changes of the two electrodes must be of opposite sign. As most precisely used, the term refers to infinitesimal changes. Symbolically, amplification factor is defined by the equation

$$\mu = -\frac{\partial e_b}{\partial e_c} \qquad (i_b = \text{const.}) \tag{3-8}$$ [2]

It must be proved that the factor μ which appears in Eqs. (3-2) and (3-3) is the same as that defined by Eq. (3-8). Examination of Eq. (3-2) shows that, unless the plate current is independent of electrode voltages, which is true in practice only when the current is zero or saturated, i_b is constant only when $e_b + \mu e_c + \epsilon$ is constant. Since tubes are seldom used continuously either with zero plate current or with saturation plate

[1] Although the partial derivative implies that other variables are held constant, for the sake of emphasis it seems advisable in this and following equations to indicate the constant parameters in parentheses.

[2] The symbolic definitions of tube factors are usually written in terms of the alternating components of electrode voltages and currents, rather than in terms of the total values. Since the difference between the instantaneous value of the alternating component of a varying quantity and the instantaneous total value of the quantity is equal to the average value, the derivative of which is zero, derivatives of the alternating component and of the total quantity are equivalent. In order to show more closely the relation of the tube factors to the characteristic curves, and to simplify derivations based upon Eq. (3-2), the tube factors will be defined in terms of the total values of currents and voltages (see Secs. 3-16 and 3-17 for symbols).

current, i_b is constant when

$$e_b + \mu e_c + \epsilon = \text{const.} \tag{3-9}$$

Differentiation of Eq. (3-9) shows that

$$-\left.\frac{de_b}{de_c}\right]_{i_b = \text{const.}} = \mu \tag{3-10}$$

or

$$\mu = -\frac{\partial e_b}{\partial e_c} \tag{3-11}$$

Since Eq. (3-11) is identical with Eq. (3-8), the factor μ of Eq. (3-2) is the same as that of Eq. (3-8).

(*A-c*) *electrode conductance* is the ratio of the change in the current in the circuit of an electrode to a change in the voltage of the same electrode, all other electrode voltages being maintained constant. As most precisely used, the term refers to infinitesimal changes as indicated by the defining equation[1]

$$g_j = \frac{\partial i_j}{\partial e_j} \qquad (e_k,\ e_l, \text{ etc.} = \text{const.}) \tag{3-12}$$

(See also Sec. 3-26.)

(*A-c*) *electrode resistance* r_j is the reciprocal of electrode conductance.

The electrode conductance that is used most frequently in the analysis of vacuum tubes and vacuum-tube circuits is the *plate conductance*[1]

$$g_p = \frac{\partial i_b}{\partial e_b} \qquad (e_c = \text{const.}) \tag{3-13}$$

(*A-c*) *plate resistance* is the reciprocal of plate conductance.[1]

$$r_p \equiv \frac{1}{g_p} = \frac{\partial e_b}{\partial i_b} \qquad (e_c = \text{const.}) \tag{3-14}$$

(See also Sec. 3-26.)

Transconductance is the ratio of the change in the current in the circuit of an electrode to the change in the voltage of another electrode, under the condition that all other voltages remain unchanged. As most

[1] The a-c electrode conductance defined by Eq. (3-12) must not be confused with the d-c electrode conductance, which is defined as the ratio of the total or direct electrode current to the total or direct electrode voltage. Similarly, the a-c plate resistance defined by Eq. (3-14) must not be confused with the d-c plate resistance, e_b/i_b. D-c electrode conductances and resistances are rarely of value and never essential in the analysis of vacuum tubes and associated circuits. Hence the a-c conductances and resistances are usually referred to simply as conductances and resistances, the adjective "a-c" being omitted.

precisely used, the term refers to infinitesimal changes as indicated by the defining equation

$$g_{jk} = \frac{\partial i_j}{\partial e_k} \qquad (e_j,\ e_l, \text{ etc.} = \text{const.}) \tag{3-15}$$

The transconductance most frequently used in the analysis of vacuum tubes and vacuum-tube circuits is the *grid-plate transconductance* (*mutual conductance*), which is defined symbolically as

$$g_m = \frac{\partial i_b}{\partial e_c} \qquad (e_b = \text{const.}) \tag{3-16}$$

Unless otherwise specified, the term *transconductance* usually refers to control-grid–plate transconductance and will be so used in the remainder of this book.

Grid Factors.—In many applications of vacuum tubes the operating voltages are such that no conduction current flows to the control grid and all electrode voltages except those of the control grid and the plate are kept constant. Under these conditions, vacuum-tube problems and derivations can be treated by the use of only three of the factors that have been defined: μ, r_p, and g_m. If conduction current flows to the control grid, it may be necessary to make use of the corresponding control-grid factors, which are defined symbolically as follows:

$$\textit{Grid amplification factor } \mu_g = -\frac{\partial e_c}{\partial e_b} \qquad (i_c = \text{const.}) \tag{3-17}$$

$$\textit{Grid conductance } g_g = \frac{\partial i_c}{\partial e_c} \qquad (e_b = \text{const.}) \tag{3-18}$$

$$\textit{Plate-grid transconductance } g_n = \frac{\partial i_c}{\partial e_b} \qquad (e_c = \text{const.}) \tag{3-19}$$

Because of the effect of space charge and because of division of the total cathode current between the grid and the plate, μ_g is not in general the reciprocal of μ.

Proof That $g_m = \mu/r_p$.—Only two of the plate factors are independent. This may be shown by taking the partial derivatives of the plate current, as expressed by Eq. (3-2).

$$g_m = \frac{\partial i_b}{\partial e_c} = \mu F'(e_b + \mu e_c + \epsilon) \tag{3-20}$$

$$g_p = \frac{\partial i_b}{\partial e_b} = F'(e_b + \mu e_c + \epsilon) \tag{3-21}$$

Dividing Eq. (3-20) by Eq. (3-21) gives

$$\frac{g_m}{g_p} = \mu \qquad \text{or} \qquad g_m = \frac{\mu}{r_p} \tag{3-22}$$

A similar derivation, based on Eq. (3-6), shows that

$$g_n = \frac{\mu_g}{r_g} \tag{3-23}$$

3-6. Relation of Tube Factors to Characteristic Curves.—The definitions state that g_p, g_m, and μ are the slopes of the i_b-e_b, i_b-e_c, and e_b-e_c curves, respectively, at points corresponding to the given voltages. Values of these factors may, therefore, be determined accurately by measuring the slopes of the static characteristics at points corresponding to the electrode voltages, and approximately by taking the ratios of small increments of current and voltage corresponding to points on the characteristics. All three factors may be determined from a single family of characteristics. The most accurate method of obtaining the three plate

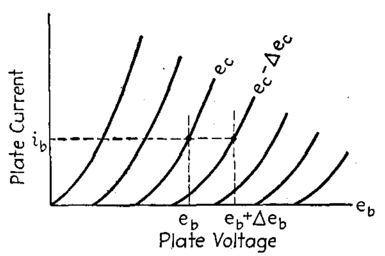


FIG. 3-5.—Method of determining triode amplification factor from the family of plate characteristics.

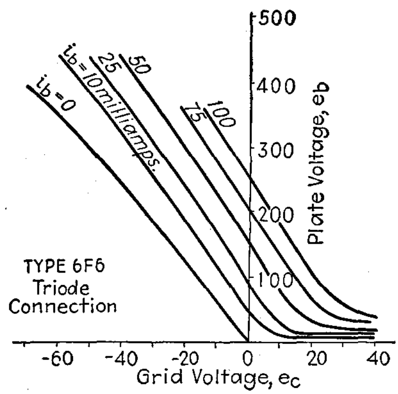


FIG. 3-6.—Typical triode e_b-e_c characteristics. (Derived from Fig. 3-4a.)

factors of a triode directly from the plate family of characteristics is to find r_p from the reciprocal of the slope of the tangent to the i_b-e_b curve at the point corresponding to the given electrode voltages, μ from the ratio of Δe_b to Δe_c at the constant current of the point, as shown in Fig. 3-5, and g_m from the ratio of μ to r_p. Curves of e_b *vs.* e_c at constant i_b for a typical triode are shown in Fig. 3-6.[1] The practically constant slope of the curves except at low plate voltages and high positive grid voltages indicates that over the normal operating range the amplification factor is nearly constant. Except in tubes designed to have variable amplification factor (Sec. 3-7), the variation of μ over the normal range of voltages does not exceed 10 to 15 per cent in triodes. Because of the small variation of μ at negative grid voltages, fairly large voltage incre-

[1] e_b-e_c curves are not usually used in the solution of vacuum-tube problems and are shown here only to point out the nearly constant value of the amplification factor in the normal range of voltages.

ments may be used without great error in determining its value from the plate characteristics. Some increase in accuracy is gained by using a grid-voltage increment such that the point at which μ is desired is at the center of the increment, rather than at one side as in Fig. 3-5.

3-7. Sharp-cutoff and Remote-cutoff Grids. Variable-mu Tubes.—Thus far it has been assumed that the grid-wire spacing and diameter and the spacing of the grid from the cathode are uniform throughout the length of the grid. When this is true the field strength does not vary greatly over the cathode surface and so the negative grid voltage necessary to prevent electrons from going to the plate at any value of plate voltage is very nearly the same for all points of the cathode. The static transfer characteristic therefore approaches the grid-voltage axis relatively sharply. For this reason a grid of uniform structure is called a *sharp-cutoff grid.* If some dimension of the grid, such as the spacing between the wires, varies along the grid, on the other hand, the field at the cathode varies correspondingly at the cathode surface. A greater negative grid voltage is required to prevent electrons from leaving the cathode at points corresponding to portions of the grid where the spacing is large than to portions where the spacing is small. Cutoff consequently takes place at different values of grid voltage at different parts of the cathode and so the static transfer characteristic approaches the axis gradually. Such a grid is known as a *gradual-cutoff* or *remote-cutoff* (*supercontrol*) *grid.* Because the mu-factor corresponding to an elementary length of the grid varies along the grid and because the mu-factor of the entire grid varies greatly with electrode voltages, such a grid is also termed a *variable-mu* grid. A tube that has a variable-mu control grid is called a variable-mu tube.

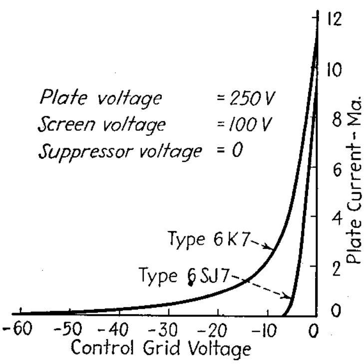


FIG. 3-7.—Comparison of transfer characteristics of similar tubes having remote-cutoff (6K7) and sharp-cutoff (6SJ7) grids.

Figure 3-7 shows the static transfer characteristics of two comparable tubes, one of which has a sharp-cutoff control grid, and the other a remote-cutoff control grid. Multigrid tubes may contain both one or more sharp-cutoff grids and one or more remote-cutoff grids. The advantages of remote-cutoff grids and tubes are discussed in Secs. 6-19 and 6-25 in connection with their use in voltage amplifiers.

3-8. Multigrid Tubes.—Many desirable characteristics can be attained in vacuum tubes by the use of more than one grid. The most common types of multigrid tubes are the tetrode and the pentode. A *tetrode* is a four-electrode type of thermionic tube containing an anode, a cathode, a control electrode, and an additional electrode, which is ordinarily a grid. A *pentode* is a five-electrode type of thermionic tube containing an anode, a cathode, a control electrode, and two additional electrodes, which are ordinarily grids.

The symbols for tetrodes and pentodes are similar to those for triodes, the various grids being shown in the relative positions that they occupy in the tubes. A special symbol, shown in Fig. 3-8*b*, is often used for screen-grid tetrodes.

The Screen-grid Tetrode.—One stimulus to the development of multigrid tubes was the necessity of reducing the capacitance between the grid and plate of the triode. If a vacuum tube used in a voltage amplifier has high grid-plate capacitance, the relatively large variations of plate voltage may induce appreciable variations of grid voltage. If the phase relations are correct, this induced grid voltage may add to the impressed alternating voltage in such a manner as to cause the amplifier to oscillate (see Sec. 10-28). This difficulty imposes a limit upon the amplification that can be attained in radio-frequency amplifiers. For some years the problem was solved by "neutralizing." Neutralization consists in connecting the grid through a small variable condenser to a point in the output circuit whose voltage is opposite in phase to that of the plate. The condenser may be adjusted so that it balances out, or neutralizes, the effect of the grid plate capacitance. Difficulties of adjustment, circuit complications, and the cost of patent royalties made it advantageous to solve the problem by removing the cause, rather than by counteracting it. This was accomplished by introducing between the control grid and the plate another grid, the *screen grid*, the purpose of which is to shield the grid from the plate, and thus reduce the grid-to-plate capacitance.[1] Further reduction in capacitance between the grid and the plate was attained by placing the control-grid terminal at the top of the tube, instead of on the base. The screen-grid tetrode proved to have other characteristics which are fully as important as its low grid-to-plate capacitance.

The general construction of the elements of a type 24A screen-grid tetrode is shown in Fig. 3-8*a*. The screen grid consists of two cylinders of fine-mesh screening, one of which is between the plate and the control grid and the other outside of the plate. These two cylinders are joined at the top by an annular disk, which completes the shielding. The

[1] SCHOTTKY, W., *Arch. Elektrotech.*, **8**, 299 (1919); U. S. Patent 1537708; BARKHAUSEN, H., *Jahrb. drahtl. Tel. u. Tel.*, **14**, 43 (1919); HOWE, G. W. O., *Radio Rev.*, **2**, 337 (1921); HULL, A. W., and WILLIAMS, N. H., *Phys. Rev.*, **27**, 432 (1926); HULL, A. W., *Phys. Rev.*, **27**, 439 (1926); WARNER, J. C., *Proc. I.R.E.*, **16**, 424 (1928) (with 22 references); PRINCE, D. C., *Proc. I.R.E.*, **16**, 805 (1928); WILLIAMS, N. H., *Proc. I.R.E.*, **16**, 840 (1928); PIDGEON, H. A., and MCNALLY, J. O., *Proc. I.R.E.*, **18**, 266 (1930).

potential of the screen is normally intermediate between the quiescent potentials of the cathode and the plate. The positive voltage of the screen draws the electrons away from the cathode. Some of these electrons strike the screen and result in a screen current which usually performs no useful function; the rest pass through the screen grid and into the field of the plate, which causes them to be drawn to the plate. Since the electrostatic field of the plate terminates almost completely on the screen, the capacitance between the plate and the grid is very small. Furthermore, variations of plate voltage have little effect on the plate current. The control-grid voltage, on the other hand, is just as effective as in the triode. The change in plate current resulting from a change in plate voltage at constant grid voltage is small, and the ratio of the change in plate voltage to the change in grid voltage, necessary to produce a given change in plate current, is very high. It follows from the definitions of plate resistance and amplification factor that the screen-grid tetrode has high plate resistance and high amplification factor. By proper choice of control-grid structure and spacing of electrodes the transconductance can also be made high. A screen-grid tetrode can therefore be designed to have the same transconductance as that of an equivalent triode and very much higher amplification factor and plate resistance.

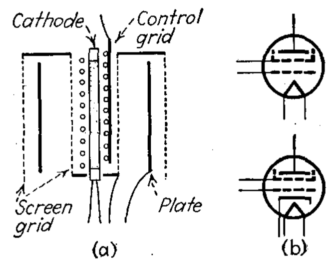


FIG. 3-8.—(a) Electrode structure of heater-type screen-grid tetrode. (b) Tube symbols for filamentary and heater-type screen-grid tetrodes.

In Fig. 3-9 is shown a family of plate characteristics for a typical screen-grid tetrode, the type 24A. The negative slope of the characteristics at plate voltages lower than the screen voltage is the result of secondary emission from the plate. At zero plate voltage there is a small plate current which results from those electrons which pass through the screen with sufficient velocity to reach the plate. As the plate voltage is raised, more and more electrons are drawn to the plate after passing through the screen. The velocity with which they strike the plate increases with the plate voltage and, when e_b is about 10 volts, becomes sufficiently high to produce appreciable secondary emission from the plate. Because the screen is at a higher voltage than the plate, these secondary electrons are drawn to the screen. Since the secondary electrons move in the direction opposite to that of the primary electrons, they reduce the net plate current. If the plate is not treated to reduce secondary emission, the number of secondary electrons leaving the plate may exceed the number of primary electrons that strike the plate, and so the plate current may reverse in direction. This is shown by the dashed curve of Fig. 3-9, which is for the old type 24A tube, with untreated plate.

It is to be expected that all secondary electrons emitted by the plate will return to the plate when the plate voltage is higher than the screen voltage. The rise in plate current starting at voltages considerably lower than the screen voltage shows, however, that many secondary electrons return to the plate while the screen is still positive relative to the plate. The reason for this is the retarding field at the plate produced by electron space charge between the screen and the plate. (It will be explained in Sec. 3-11 how tubes may be designed so that this space charge prevents secondary electrons emitted by the plate from going to the screen even at plate voltages much lower than the screen voltage.)

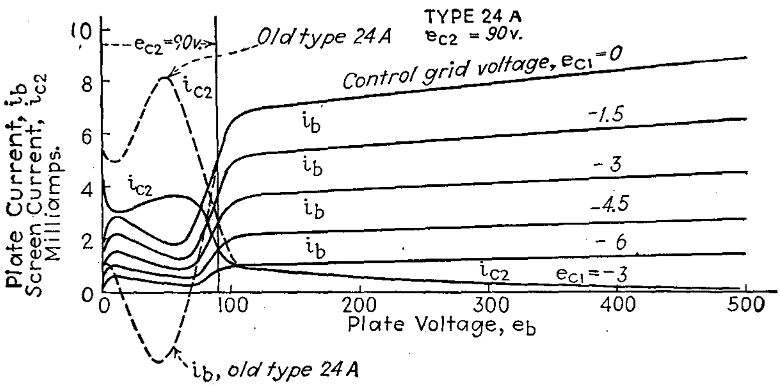


FIG. 3-9.—Typical screen-grid tetrode plate characteristics at 90-volt screen voltage, e_{c2}.

As the plate voltage approaches the screen voltage, the field at the plate produced by the screen voltage becomes less than the retarding field of the space charge and so the slower-moving secondary electrons are returned to the plate. At voltages higher than that at which all secondary electrons are returned to the plate, secondary emission from the plate has no effect upon the plate current, which is then determined almost entirely by the screen- and control-grid voltages. Since very little of the plate field penetrates to the cathode, further increase of plate voltage has only a small effect upon the plate current. The increase of plate current at plate voltages higher than the screen voltage is accounted for partly by increase in the number of secondary electrons from the screen that are drawn to the plate.

3-9. The Space-charge Tetrode.—Instead of using the inner grid of a tetrode as the control electrode and applying a positive voltage to the second grid, it is possible to operate the tube by applying a small positive voltage to the inner grid and using the second grid as the control electrode.[1] The positive voltage on the first grid overcomes the effect of the space charge in the vicinity of the cathode, and thus increases the plate current and the transconductance. Some of the electrons are drawn to the positive inner space-charge grid, but the remainder pass through this grid and into the region controlled by the second grid and the plate. The effect is in some respects the same as though the cathode were placed much closer to the control grid in a triode. A high negative voltage on the second grid prevents the electrons from passing to the plate and returns them to the positive space-charge grid. As the negative control-grid voltage is reduced, more electrons pass to the plate and fewer to the space-charge grid. Thus, the plate current increases, and the space-charge–grid current decreases with decrease of negative voltage on the control grid. Figure 3-10 shows typical curves of plate current and of first-grid current as a function of second-grid voltage.

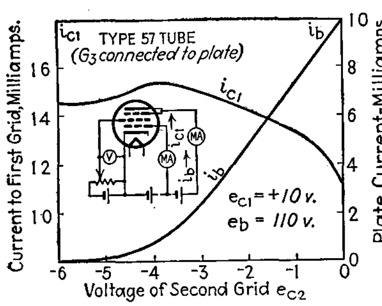


FIG. 3-10.—Characteristics showing first-grid current i_{c1} and plate current i_b of a space-charge tetrode as a function of second-grid voltage e_{c2}.

Although the transconductance of a space-charge tetrode is greater than that of a triode with a similar cathode, the relatively high current to the space-charge grid results in a less efficient use of the cathode current. Because more recently developed pentodes have much better characteristics than space-charge tetrodes, space-charge tetrodes are now used only in special applications, some of which will be discussed in later chapters.

3-10. The Pentode.—For most applications the curved portions of the characteristic curves of screen-grid tetrodes at plate voltages lower than the screen voltage are undesirable. In amplifiers, excessive distortion results if the tube is operated in this region and, if the circuit contains inductance and capacitance, oscillation may occur (see Sec. 10-17). Restriction of operation to the region to the right of the plate-current dip reduces the output voltage or power that can be obtained at a given value of operating plate voltage.

By the use of a ribbed plate and special treatment to reduce secondary emission, it is possible to design tetrodes whose characteristic curves do not have portions with negative slope. The type 48 tetrode is an example of such a tube. The effects of secondary emission can also be reduced or eliminated by preventing the secondary electrons emitted by

[1] ARDENNE, M. VON, *Hochfrequenztech. u. Elektroakustik*, **42**, 149 (1933). See also *Wireless Eng.*, **11**, 93 (1934) (abstr.); I. LANGMUIR, U. S. Patent 1558437, filed Oct. 29, 1913; WARNER, *loc. cit.*

the plate from going to the screen. This can be done by placing between the screen and the plate a third grid, called the *suppressor grid.* Figure 3-11*a* shows the arrangement of the electrodes of a typical suppressor pentode of the voltage amplifier type, the 57. The purpose of the shield in the dome of the tube is to shield the control-grid lead and terminal from the plate. This shield is connected internally to the cathode and shaped so as to act as the continuation of an external shield which may be placed around the tube. In voltage pentodes of more recent design, such as that shown in Fig. 3-11*b*, special precautions in placing and shielding the leads have made it possible to connect the control grid to a base pin instead of to a terminal at the top of the tube.[1] In power pentodes the control-grid terminal is in the base of the tube.

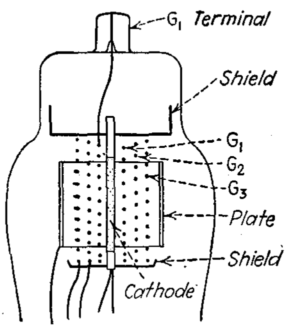


FIG. 3-11*a*.—Electrode structure of the type 57 pentode. G_1G_2, and G_3 are normally the control, screen, and suppressor grids, respectively.

Figure 3-12 shows a series of plate characteristics of the 57 tube at constant control-grid and screen-grid voltages for a number of suppressor-grid voltages. It can be seen that the secondary-emission dips move to the left and become less pronounced as the suppressor-grid voltage is reduced below the screen voltage. When the suppressor voltage is zero, *i.e.*, when the suppressor is connected to the cathode, the secondary-emission effects are almost entirely absent.

The explanation of the action of the suppressor is simple. When the suppressor is connected to the cathode, the field between the plate and the suppressor is always such as to move electrons toward the plate. The secondary electrons removed from the plate have sufficiently low velocity of emission so that even at low plate voltages few can permanently leave the plate against this retarding field. Velocity acquired by the primary electrons in the space between the cathode and the screen carries most of them through the screen and suppressor and thence into the field beyond the suppressor, which draws them to the plate.

The additional shielding effect of the suppressor grid results in grid-plate capacitance that is even lower than that of tetrodes and in plate resistance and amplification factor that are even greater than those of tetrodes. It can be seen from Fig. 3-12 that the plate resistance of the suppressor pentode can be varied by means of negative suppressor voltage.

In some pentodes the suppressor is permanently connected to the cathode internally; in others, on the other hand, all three grids may have

[1] KELLY, R. L., and MILLER, J. F., *Electronics*, September, 1938, p. 26.

external connections in order that the grids may be used in various ways. When the second and third grids are connected to the plate, the tube has ordinary triode characteristics. When the first and second grids are used together as the control grid and the third is connected to the plate, the tube acts as a triode with very high amplification factor and low plate current. Other special applications of pentodes will be dis-

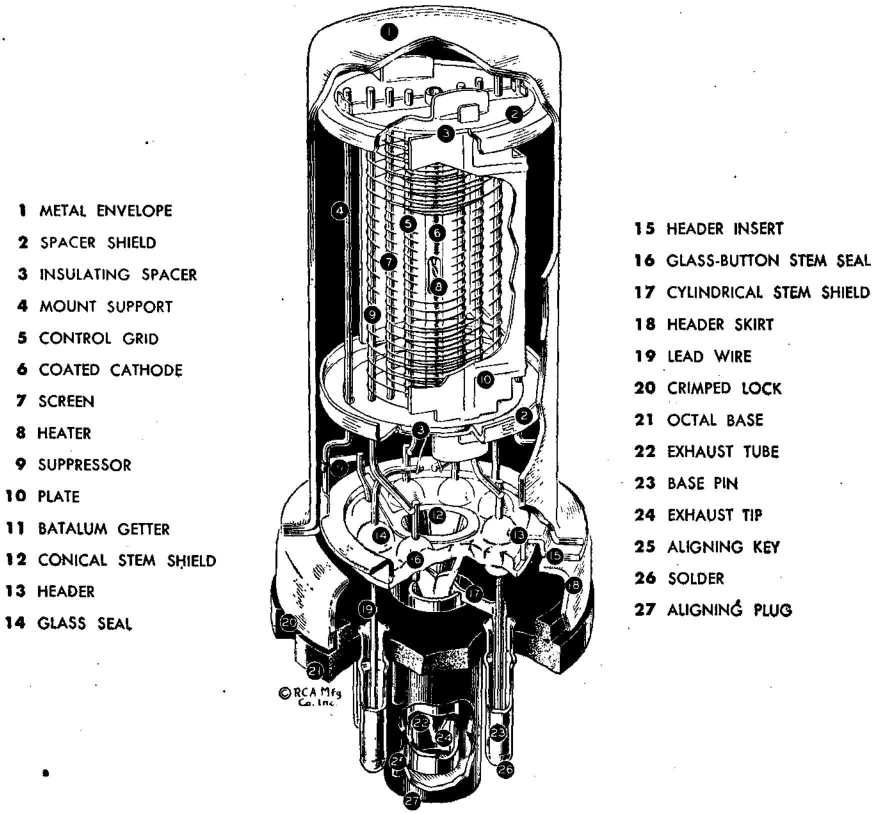


FIG. 3-11*b*.—Typical metal voltage-amplifier pentode in which the control-grid connection is brought out through the base. Note the conical stem shield 12. (*Courtesy of Radio Corporation of America.*)

cussed in later chapters (see end of Sec. 6-1 and Secs. 10-3, 10-4, 10-12, 10-23, 10-39, 12-6, and 12-33).

Figures 3-13 and 3-14 show the plate characteristics for two types of suppressor pentodes, the 57 and the 2A5 (see also Figs. A-13 to 15 and A-19, pages 677–678). It will become apparent from material to be presented in this and later chapters that ideal characteristics for amplifier tubes would be straight, parallel, and equidistant for all values of plate voltage. The gradual bending of the characteristics at low plate voltages, particularly

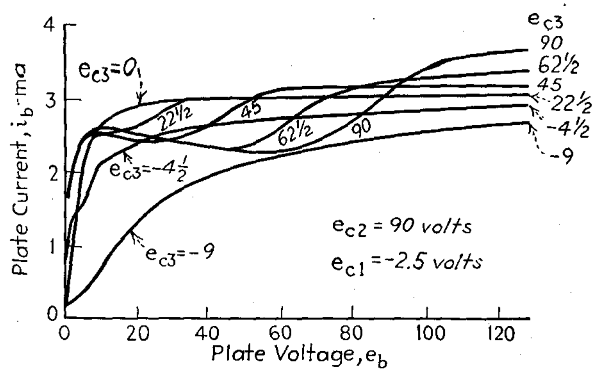


FIG. 3-12.—Plate characteristics of the type 57 pentode, showing the effect of variation of suppressor voltage e_{c3} at 90-volt screen voltage e_{c2}.

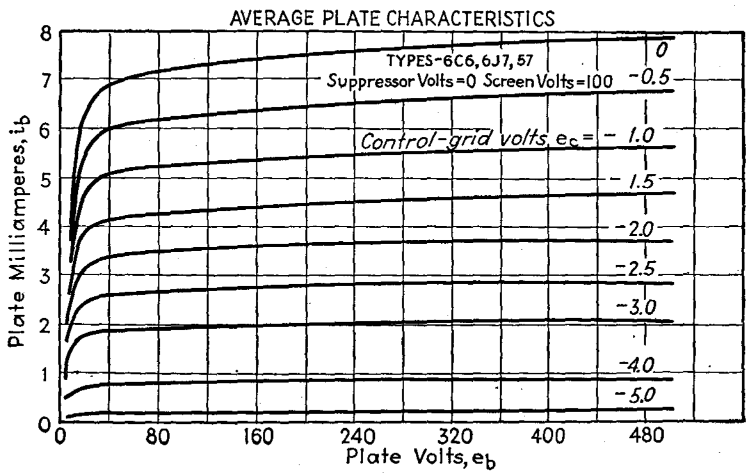


FIG. 3-13.—Plate characteristics of a typical voltage pentode.

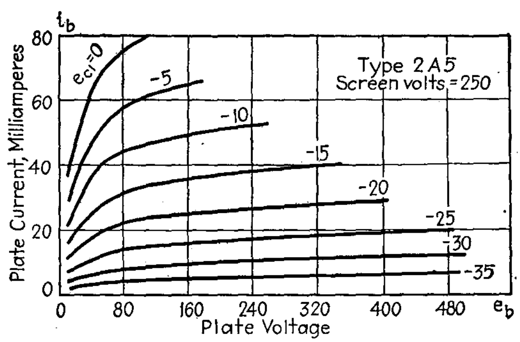


FIG. 3-14.—Plate characteristics of typical power pentode.

noticeable in the characteristics of pentodes designed to give large power output, such as the type 2A5, causes distortion in amplifiers and is, therefore, objectionable. It is the result of nonuniformity of the field in the plane of the suppressor grid. Because of the shielding action of the grids, the plate voltage has relatively little effect upon the number of electrons arriving in the plane of the suppressor grid. As the plate voltage is raised, however, more and more electrons arriving at the suppressor are drawn to the plate. The comparatively constant "saturation" plate current is obtained when all electrons arriving at the plane of the suppressor are drawn to the plate. Because of nonuniformity of field in the plane of the suppressor, saturation is attained at different plate voltages for various points in this plane, and so the characteristics bend over gradually and have broad knees. This difficulty is avoided in the *beam power pentode.*[1]

3-11. Beam Pentodes.—In the beam tube the secondary electrons are returned to the plate by the repulsion of negative space charge between the screen and the plate. This space charge is accentuated by the retarding field of the plate when the plate potential is lower than the screen potential. By proper design the space charge may be made so dense as to cause the formation of a virtual cathode [*i.e.*, a plane of zero average field and zero electron velocity (see Sec. 2-12)] near the plate at low plate voltages. For all values of plate voltage less than the screen voltage a potential minimum is formed which may be kept sufficiently lower than the plate potential so that secondary electrons from the plate are returned to the plate. The action is similar to that of a suppressor grid; but, if the density of space charge and the electron velocity at the virtual cathode are uniform and the distance of the virtual cathode from the plate is everywhere the same, saturation takes place simultaneously at all points in the plane of the virtual cathode, and the knee of the plate characteristic is sharp. The virtual cathode and the plate act in a manner similar to a diode, the plate current being limited by space charge at low plate voltages. As in a diode, saturation at low plate voltage, *i.e.*, a low-voltage knee, requires that the virtual cathode shall be close to the plate.

In the beam power tube, of which the 6L6 is a typical example, the required electron density is achieved by confining the electrons to beams. The homogeneity of space charge and electron velocity is attained by proper design of the contours of the cathode, grids, and plate and by correct choice of the ratio of screen-plate to screen-cathode spacing (2.9)

[1] SCHADE, O. H., *Proc. I.R.E.*, **26**, 137 (1938). See also F. BELOW, *Z. Fernmeldetech.* **9**, 113 (1928); R. S. BURNAP, *RCA Rev.*, **1**, 101 (1936); J. F. DREYER, JR., *Electronics*, April, 1936, p. 18; J. H. O. HARRIES, *Electronics*, May, 1936, p. 33; B. SALZBERG and A. V. HAEFF, *RCA Rev.*, **2**, 336 (1938).

and beam angle (approximately 60 degrees). The electrons are confined to beams by means of beam-forming plates, as shown in Fig. 3-15, which are at cathode potential. The flattened cathode gives a larger effective area than a round cathode and so results in a higher transconductance.

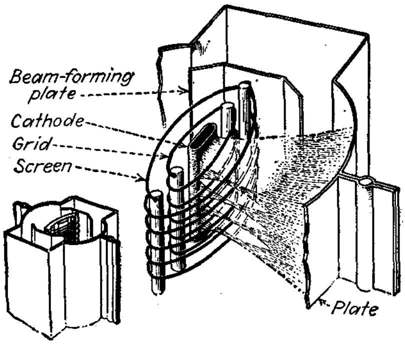


FIG. 3-15.—Electrode structure of the type 6L6 beam power pentode. (*Courtesy of Radio Corporation of America.*)

The screen current of beam power tubes is much lower than that of suppressor pentodes. The screen and control grid have equal pitch and are proportioned and assembled so that the screen grid is hidden from the cathode by the control grid, and the individual beam sheets formed by the control grid are focused in the plane of the screen. Very few, therefore, of the electrons moving toward the plate can strike the screen directly. Furthermore, because of the sharpness of the beams and the uniform fields, few electrons acquire tangential velocity at the expense of velocity normal to the electrode planes. The number of electrons that miss the plate at low plate voltage and return to the screen is therefore also small. Low screen current results in a number of

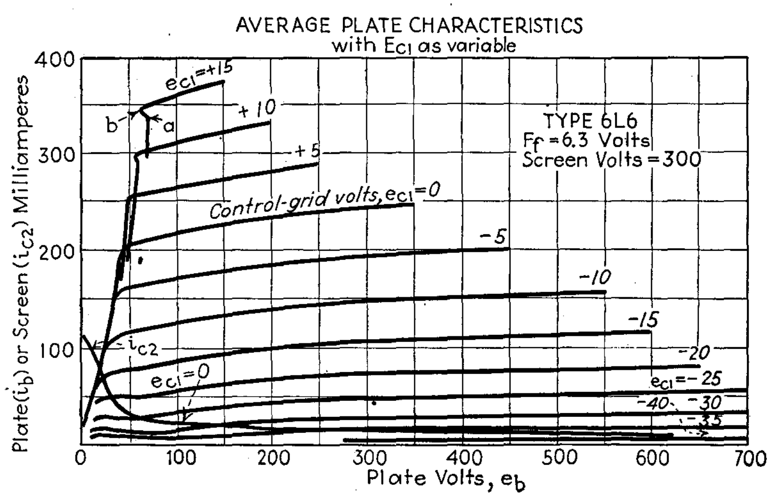


FIG. 3-16.—Plate characteristics of the type 6L6 beam power pentode.

advantages, among which are low screen dissipation and consequent larger power rating without danger of electron emission from the screen because of high screen temperature, and the possibility of using low-power resistors in the voltage divider that supplies the screen voltage.

Figure 3-16 shows the plate characteristics of the 6L6. These characteristics are straight and parallel over a much greater range of plate voltage than those of suppressor pentodes, and the knees are sharper. The peculiar shape of the knees of the curves for positive control-grid voltages is the result of two possible distributions of space charge at the same value of plate voltage.[1] As the plate voltage is gradually increased from a low value, the plate current increases continually until the point *a* is reached, at which it jumps abruptly to the higher value corresponding to this voltage. When the plate voltage is then gradually reduced, the current decreases continuously until the point *b* is reached, at which it falls abruptly to the lower value corresponding to this plate voltage. The abrupt changes in current are accompanied by changes in space-charge distribution.

3-12. Equations for Plate Current of Multigrid Tubes.—Plate-current equations similar to Eqs. (3-1) and (3-2) may be written for tetrodes and pentodes. In most circuits, however, only the control-grid and plate voltages are varied, and Eqs. (3-1) and (3-2) may be applied.

Determination of Tetrode and Pentode Factors from Characteristics. The plate characteristics of tetrodes and pentodes are not ordinarily shown for sufficiently small increments of control-grid voltage to make possible the direct determination of μ from the plate characteristics. Furthermore, the plate resistance of voltage-amplifier tetrodes and pentodes, such as the 24A (Fig. 3-9) and the 6J7 (Fig. 3-13), is so high that the plate characteristics are nearly horizontal throughout the working range of voltages and the plate resistance can be determined only approximately from the tangent to the plate characteristic at the point corresponding to the given voltages. The plate resistance of power suppressor pentodes such as the 2A5 (Fig. 3-14), however, and of beam pentodes such as the 6L6 (Fig. 3-16) is sufficiently low so that it can be determined with fair accuracy from the tangent to the plate characteristic.

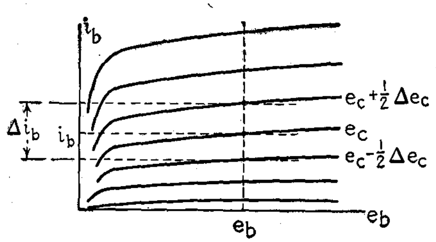


FIG. 3-17.—Graphical determination of transconductance of pentodes.

An approximate value of g_m can be found from increments of plate current and grid voltage at constant plate voltage, as determined from the plate characteristics. Since g_m is not constant, accuracy in the values of g_m found by this method is dependent upon the use of very small increments. Accuracy also makes it desirable to use an increment of

[1] SALZBERG and HAEFF, *loc. cit.*

grid voltage such that the point at which g_m is desired lies at the center of the increment, as shown in Fig. 3-17. If the grid voltage intervals corresponding to the available plate characteristics are large, it may be advisable to construct the static transfer characteristic corresponding to the given plate voltage and to find g_m accurately from the slope of this characteristic at the given plate voltage. Equation (3-22) may be used to find an approximate value of μ from g_m in conjunction with the approximate value of r_p determined from the plate characteristic.

3-13. Duplex Tubes and Tubes with More than Three Grids.—The behavior of the individual units of duplex tubes is no different than when these units are enclosed in separate envelopes. For this reason duplex tubes require no further discussion. Special tubes having more than three grids will be treated in later chapters in connection with their applications (see Secs. 6-18, 6-19, 9-12*A*, and 9-22).

3-14. Applied Voltages in Grid and Plate Circuits.—In most applications of vacuum tubes, one or more alternating voltages are applied to the control-grid and plate circuits, in addition to the steady voltages, as shown in Fig. 3-18. The steady applied voltages E_{cc}, E_{bb}, and E_{ff} are called the *grid supply voltage* or *C-supply voltage*, the *plate supply voltage* or *B-supply voltage*, and the *filament* or *heater supply voltage*, respectively. v_g and v_p, the applied alternating voltages in the grid and plate circuits, are called the *grid excitation voltage* and *plate excitation voltage*, respectively. Because of impedance drops resulting from the flow of electrode currents or other circuit currents, the electrode voltages in general differ from the voltages impressed upon the electrode circuits. This will be discussed in detail in Secs. 3-17 and 4-1.

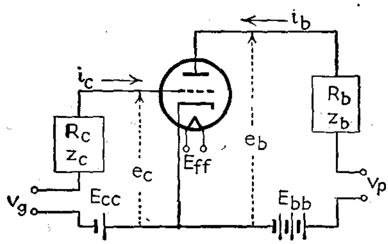


FIG. 3-18.—Triode grid and plate circuits, showing supply voltages, impressed alternating voltages, and electrode voltages.

All electrode voltages are measured with respect to the cathode, and an electrode voltage is said to be positive if the electrode is positive relative to the cathode. Similarly, supply voltages, instantaneous values of exciting voltages, and instantaneous values of voltages across load impedances are ordinarily measured relative to the cathode side of these voltages and are called positive if they tend to make the electrode positive.

3-15. Form of Alternating Plate-current Wave.—Figure 3-19 shows the manner in which a sinusoidal grid voltage causes the plate current to vary when the load is nonreactive. The wave of plate current is constructed by projecting from the wave of grid voltage to the transfer characteristic at various instants throughout the cycle. The significance of the *time axis* of the wave of varying plate current and of the current I_{bt} corresponding to this axis should be noted. I_{bt} is the current assumed at the instants in the cycle when the alternating grid voltage is zero. The reason why I_{bt} differs from the steady current I_{bo}, assumed when the excitation is zero, will be explained in detail in Chap. 4 (Secs. 4-8 to 4-10).

Curvature of the transfer characteristic in general causes the wave of varying plate current to be asymmetrical even though the exciting voltage is sinusoidal. It will be proved in later sections that the wave of alternating plate current, measured relative to the time axis, in general contains not only a fundamental component of the same frequency as the grid excitation voltage, but also harmonics of that frequency, and a steady component. When the load is nonreactive, the wave of plate voltage is of the same form as the wave of plate current.

Usually it is most convenient to measure the instantaneous value of the alternating component of plate or grid current or voltage relative to the time axis of the wave. Occasionally, however, it is necessary to measure the instantaneous value with respect to the average value. An instantaneous value measured relative to the axis differs from the value measured relative to the average in that the former contains the steady component of the alternating current or voltage, whereas the latter does not. Unless otherwise specifically stated, the terms *alternating plate current* and *alternating grid current*, and the corresponding terms for voltages, will be understood to refer to values measured relative to the time axes.

In some applications of vacuum tubes the curvature of the characteristic or the amplitude of the alternating plate current, or both, are small enough so that the alternating plate current is essentially sinusoidal, as shown in Fig. 3-20. The difficulty in making a rigorous analysis sometimes necessitates the assumption that the alternating plate current is sinusoidal, even though it is known to be distorted.

Most circuit elements dealt with in the power field of electrical engineering are linear elements, *i.e.*, elements in which the currents flowing through the elements are proportional to the voltages across the elements. The currents that flow as the result of simultaneous application of direct and alternating voltages or of two or more alternating voltages of different frequencies are entirely independent. Hence, when several voltages are applied simultaneously, the net result may be determined by making relatively simple analyses involving only direct currents or alternating currents of only one frequency at a time and superimposing the individual results. Comparatively few symbols suffice in dealing with such circuits. Vacuum tubes, on the other hand, are nonlinear circuit elements. The direct currents that flow in a circuit containing a nonlinear element depend not only upon the impressed direct voltages but also upon the impressed alternating voltages. Similarly, the alternating currents depend upon both the alternating and the direct impressed voltages. Furthermore, in addition to currents having the same frequencies as the impressed voltages, alternating currents will in general flow whose frequencies differ from those of the impressed alternating voltages. It is apparent that the analysis of vacuum-tube circuits may be considerably more complicated than that of power circuits containing only linear circuit elements.

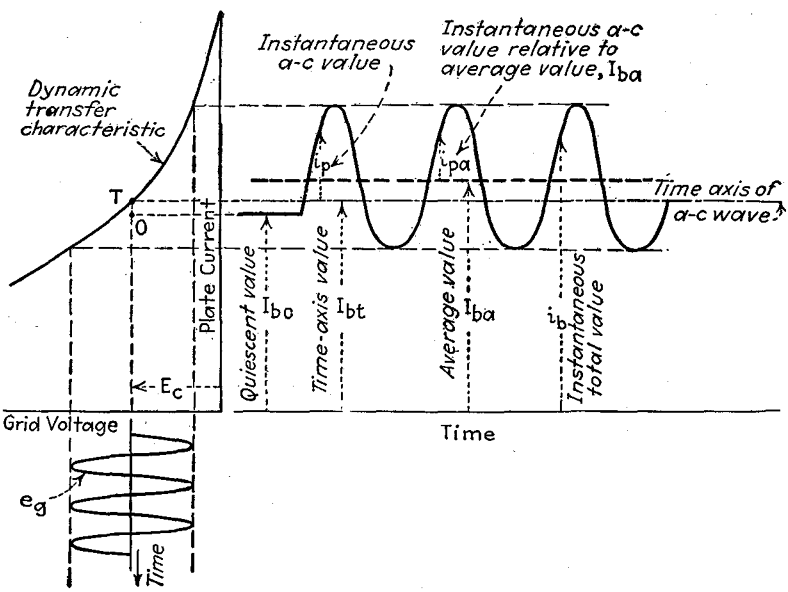


FIG. 3-19.—Plate-current relations for asymmetrical wave of plate current.

3-16. Symbols.—Because it is often necessary to consider simultaneously the various direct and alternating voltages and currents in vacuum-tube circuits, a large number of symbols must be used. This is unfortunate but usually unavoidable. The symbols used in this book are based upon those standardized by the Institute of Radio Engineers.

The student will find it helpful to learn the following rules, according to which these symbols are formed:

1. The subscripts $_c$ and $_g$ refer to the grids or to the grid circuits, and the subscripts $_b$ and $_p$ to the plate or to the plate circuit.
2. Lower-case letters indicate instantaneous values of varying quantities, and capital letters indicate steady (direct) values and average, r-m-s, and crest values of varying quantities (see Figs. 3-19 and 3-20).

3. Lower-case letters with subscripts $_c$ and $_b$ indicate *total* instantaneous values of varying quantities (see Figs. 3-19 and 3-20).

4. Lower-case letters with subscripts $_g$ and $_p$ indicate instantaneous values *of the alternating components* of varying quantities. If the wave form of the varying quantity is asymmetrical, second subscripts $_a$ and $_t$ are added to differentiate between an instantaneous value measured relative to the average value and an instantaneous value measured relative to the value corresponding to the axis of the wave (see Fig. 3-19).

5. Capital letters with subscripts $_c$ and $_b$ indicate direct or average values. Second subscripts $_o$, $_a$, and $_t$ are added to differentiate between quiescent, average, and time-axis values of an asymmetrical wave of total varying plate current or voltage (see Fig. 3-19).

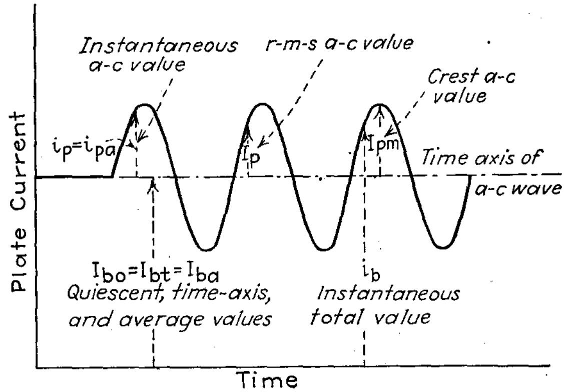


FIG. 3-20.—Plate-current relations for sinusoidal wave of plate current.

6. Capital letters with subscripts $_g$ and $_p$ indicate r-m-s or crest values of alternating quantities. Crest values of sinusoidal alternating quantities are distinguished from r-m-s values by the addition of the second subscript $_m$ (see Fig. 3-20).

In tetrodes and pentodes and in tubes having more than three grids, it is necessary to distinguish between the voltages and between the currents of the various grids. This is done by means of the addition of a number in the subscript to indicate the number of the grid. Thus e_{c1}, e_{c2}, e_{c3}, etc., indicate the instantaneous total voltages of the first, second, and third, etc., grids, grid 1 being nearest the cathode. Usually the first grid serves as the control electrode, and e_{c1} indicates the control-grid voltage. In many analyses it is necessary to speak only of the control-grid voltage. It is then convenient to omit the number in the subscript, even though the tube contains more than one grid. In the work that follows, therefore, it will be understood that, when no number appears in the subscript, reference is to the control grid and that the term *grid* refers to the control grid.

The symbols are as follows:

Control-grid supply voltage	E_{cc1} or E_{cc}
Screen-grid supply voltage	E_{cc2}
Plate supply voltage	E_{bb}
Filament or heater supply voltage	E_{ff}
Instantaneous value of grid excitation voltage	v_g
Instantaneous value of plate excitation voltage	v_p
Instantaneous total grid voltage	e_c
Instantaneous total plate voltage	e_b
Instantaneous total grid current	i_c
Instantaneous total plate current	i_b
Quiescent (zero excitation) value of grid voltage	E_{co}
Quiescent (zero excitation) value of plate voltage	E_{bo}
Quiescent (zero excitation) value of grid current	I_{co}
Quiescent (zero excitation) value of plate current	I_{bo}
Voltage corresponding to time axis of wave of varying plate voltage	E_{bt}
Current corresponding to time axis of wave of varying plate current	I_{bt}
Average value of grid voltage (grid bias)	E_c (or E_{ca})
Average value of plate voltage	E_{ba}
Average value of grid current	I_{ca}
Average value of plate current	I_{ba}
Instantaneous value of alternating component of grid voltage	e_g
Instantaneous value of alternating component of plate voltage, measured relative to the time axis of the wave of alternating plate voltage	e_p
Instantaneous value of alternating component of plate voltage, measured relative to average plate voltage	e_{pa}
Instantaneous value of alternating component of grid current	i_g
Instantaneous value of alternating component of plate current, measured relative to the time axis of the wave of alternating plate current (see Fig. 3-19)	i_p
Instantaneous value of alternating component of plate current, measured relative to average plate current (see Fig. 3-19)	i_{pa}
Effective value of alternating component of grid voltage	E_g
Effective value of alternating component of plate voltage	E_p
Effective value of alternating component of grid current	I_g
Effective value of alternating component of plate current	I_p
Crest value of sinusoidal alternating component of grid voltage (grid swing)	E_{gm}
Crest value of sinusoidal alternating component of plate voltage	E_{pm}
Crest value of sinusoidal alternating component of grid current	I_{gm}
Crest value of sinusoidal alternating component of plate current	I_{pm}
Impedance of the grid circuit at angular frequency ω	$\begin{cases} (z_c)_\omega = (r_c)_\omega + j(x_c)_\omega \\ \text{or} \quad z_c = r_c + jx_c \end{cases}$
Impedance of the plate circuit (load impedance) at angular frequency ω	$\begin{cases} (z_b)_\omega = (r_b)_\omega + j(x_b)_\omega \\ \text{or} \quad z_b = r_b + jx_b \end{cases}$
Admittance of the grid load at angular frequency ω	$\begin{cases} (y_c)_\omega \equiv \dfrac{1}{(z_c)_\omega} = (g_c)_\omega - j(b_c)_\omega \\ \text{or } y_c \equiv \dfrac{1}{z_c} = g_c - jb_c \end{cases}$

Admittance of the plate load at angular frequency ω. $\begin{cases}(y_b)_\omega \equiv \dfrac{1}{(z_b)_\omega} = (g_b)_\omega - j(b_b)_\omega \\ \text{or } y_b \equiv \dfrac{1}{z_b} = g_b - jb_b\end{cases}$

D-c resistance of the grid load . . . . R_c
D-c resistance of the plate load . . . . R_b
Effective value of the alternating voltage drop in the plate circuit impedance . . . . E_{zb}

3-17. Current and Voltage Relations in the Grid and Plate Circuits.—The following relations are apparent from Fig. 3-18 and from Fig. 3-19 and the similar wave of plate voltage:

$$i_b = I_{bt} + i_p = I_{ba} + i_{pa} \tag{3-24}$$
$$e_b = E_{bt} + e_p = E_{ba} + e_{pa} \tag{3-25}$$
$$E_{ba} = E_{bb} - I_{ba}R_b \tag{3-26}$$
$$e_p = v_p - \left(L_b \frac{di_p}{dt} + r_b i_p + \frac{1}{C_b}\int i_p\,dt\right) \tag{3-27}$$

in which L_b, r_b, and C_b are the equivalent series inductance, resistance, and capacitance of the plate load. For the special case in which the wave of alternating plate current is symmetrical, as illustrated by the sinusoidal current wave of Fig. 3-20, I_{ba} and I_{bt} are identical with I_{bo}, and E_{ba} and E_{bt} are identical with E_{bo}. i_p and i_{pa} are then also identical, as are e_p and e_{pa}. For a symmetrical wave, therefore, Eqs. (3-24), (3-25), and (3-26) reduce to the following:

$$i_b = I_{bo} + i_p \tag{3-28}$$
$$e_b = E_{bo} + e_p \tag{3-29}$$
$$E_{bo} = E_{bb} - I_{bo}R_b \tag{3-30}$$

Equation (3-30) also holds when the excitation voltage is zero. When the plate current is sinusoidal, Eq. (3-27) may be written more conveniently in terms of effective values, as follows:

$$E_p = V_p - I_p z_b \tag{3-31}$$

In many circuits the plate excitation voltage V_p is zero, and Eq. (3-31) reduces to

$$E_p = -I_p z_b \tag{3-32}$$

The following equations are apparent from Fig. 3-18 and from a diagram analogous to that of Fig. 3-19:

$$i_c = I_{ct} + i_g = I_{ca} + i_{ga} \tag{3-33}$$
$$e_c = E_{ct} + e_g = E_c + e_{ga} \tag{3-34}$$
$$E_c \text{ (or } E_{ca}) = E_{cc} - I_{ca}R_c \tag{3-35}$$
$$e_g = v_g - \left(L_c \frac{di_g}{dt} + r_c i_g + \frac{1}{C_c}\int i_g\,dt\right) \tag{3-36}$$

in which L_c, r_c, and C_c are the equivalent series inductance, resistance, and capacitance of the grid load. In many vacuum-tube circuits, currents other than grid current flow through impedances contained in the grid circuit (see Sec. 4-1). The expression for e_g given by Eq. (3-36) must then be replaced by one including the alternating voltages resulting from the flow of these additional currents.

In many vacuum-tube circuits the control grid is maintained sufficiently negative so that grid current does not flow. Unless some other direct current flows through a resistance contained in the grid circuit, Eq. (3-35) then simplifies to

$$E_c(\text{or } E_{ca}) = E_{cc} \tag{3-37}$$

and, unless some other alternating current flows through an impedance contained in the grid circuit, Eq. (3-36) becomes

$$e_g = v_g \tag{3-38}$$

which may also be written in terms of effective values as follows:

$$E_g = V_g \tag{3-39}$$

The method of evaluating E_g, when E_g results wholly or in part from the flow of alternating currents through impedances contained in the grid circuit, will be explained in detail in Sec. 4-1.

3-18. Static and Dynamic Operating Points.—The steady values of electrode voltages and currents assumed when the excitation voltages are zero are called the *quiescent* or *static operating* voltages and currents. The point on the static characteristics corresponding to given static operating voltages and currents is termed the *static operating point* (or *quiescent point*). The static operating point will be indicated on the characteristic curves by the letter O and will sometimes be referred to as the O-point. The average values of electrode voltages and currents assumed with excitation are called the *dynamic operating* voltages and currents. The corresponding point on the plate characteristics is termed the *dynamic operating point* and will be indicated by the letter A. The dynamic and static operating points coincide when the wave of plate current is symmetrical. Static and dynamic operating points will be discussed in detail in Secs. 4-7 to 4-13.

The amplitude of the alternating control-grid voltage E_{gm} is called the *grid swing*. The direct component of control-grid voltage E_c is called the *grid bias* or *C bias*. Because the symbol E_c has long been used to represent the grid bias, this symbol, rather than the alternative E_{ca}, will be used throughout the remainder of this book. Grid bias may be obtained by the use of a battery or other source of fixed voltage, or it may be produced by the flow of cathode current through a resistance, as in Figs. 5-7 and 5-8. The former is called *fixed bias*, the latter *self-bias*. The term *bias* is also sometimes applied to the direct component of voltage of electrodes other than the control grid.

3-19. Generation of Harmonics.—If a sine wave of excitation voltage is applied to an electrode of a vacuum tube, it is found that the plate current in general contains not only an alternating component of the same frequency as that of the applied voltage but also components whose frequencies are equal to harmonics of the impressed frequency. Usually there is also a change in average plate current. If two or more sinusoidal voltages are impressed simultaneously the wave form of the alternating plate current is even more complicated, containing the applied frequencies

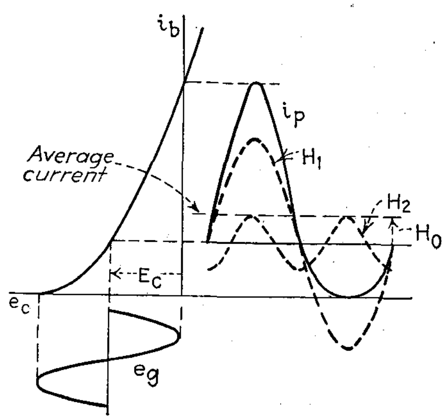


FIG. 3-21.—Generation of second-harmonic and steady components of plate current by a tube with a parabolic transfer characteristic.

and their harmonics and also frequencies equal to the sums and differences of the applied frequencies and their integral multiples.[1]

The generation of new frequencies by a vacuum tube is associated with the fact that it is not a linear circuit element, *i.e.*, that the characteristic curves are not linear. The appearance of plate-current components of frequencies different from the impressed frequency as the result of curvature of the transfer characteristic can be readily demonstrated graphically. The transfer characteristic of Fig. 3-21 is parabolic, and the exciting voltage sinusoidal, as indicated by the curve of e_g. The wave of i_p is constructed by finding from the transfer characteristic the values of instantaneous plate current corresponding to instantaneous grid voltages at various instants of the cycle. The dotted curves show

[1] It is evident, therefore, that it is not in general possible to apply to the vacuum tube the superposition theorem, which states that the current that flows through a linear circuit element as the result of a number of simultaneously impressed voltages of different frequencies is equal to the sum of the currents that would flow if the various voltage components were applied individually.

the fundamental, steady, and second-harmonic components into which the alternating plate current i_p may be resolved.

3-20. Series Expansion for Alternating Plate Current.—Theoretically, it should be possible to predict from Eq. (3-2) and corresponding tetrode and pentode equations the form of the alternating plate current corresponding to an exciting voltage of known wave form. Practically, however, $F(e_b + \mu e_c + \epsilon)$ is so complicated in form that Eq. (3-2) is of little or no value for this purpose. The behavior of a nonlinear circuit element can in general be analyzed mathematically most readily by expressing the alternating current in the form of an infinite power series. For a two-terminal element, in which the current depends upon only one voltage, the series involves only the impressed voltage, the element impedance and its derivatives at the operating point, and the circuit impedance. It may be derived by the application of Taylor's expansion for a function of one variable to the functional equation of current. Because the plate current of a vacuum tube depends upon all the electrode voltages, a complete series expansion applicable to a tube with three or more electrodes must involve all electrode voltages, as well as various tube factors and their derivatives at the operating point.[1] The general series expansion for tetrodes and pentodes is complicated, therefore, particularly if control-grid current is assumed to flow and the control-grid circuit contains an impedance. When all electrode voltages except those of the control grid and the plate are constant, however, as is usually true, the action of multigrid tubes is similar to that of triodes, and the general form of the series reduces to that for a triode. The triode expansion may consequently be applied to any tube in which only the plate and control-grid voltages vary. Furthermore, most problems in which the series expansion is of value can be adequately treated by the use of the form derived under the assumption that control-grid current does not flow. A further simplification results from the assumption that μ is constant, which is approximately true in many tubes. For the general forms of the series expansion for plate current of multigrid tubes, and for the similar expansions for currents to other electrodes, the student should refer to the work of Llewellyn[2] and others.[3]

3-21. Series Expansion. *Resistance Load.*—The series expansion for the plate current of a triode with negative grid voltage, constant amplification factor, and nonreactive load has the following form:

$$i_p = a_1 e + a_2 e^2 + a_3 e^3 + \cdots \tag{3-40}$$

[1] CARSON, J. R., *Proc. I.R.E.*, **7**, 187 (1919).

[2] LLEWELLYN, F. B., *Bell System Tech. J.*, **5**, 433 (1926).

[3] BRAINERD, J. G., *Proc. I.R.E.*, **17**, 1006 (1929); CAPORALE, P., *Proc. I.R.E.*, **18**, 1593 (1930); BONER, M. O., *Phys. Rev.*, **39**, 863 (1932); BENNETT, W. R., *Bell System Tech. J.*, **12**, 228 (1933); ESPLEY, D. C., *Proc. I.R.E.*, **22**, 781 (1934); BARROW, W. L., *Proc. I.R.E.*, **22**, 964 (1934).

where

$$\left.\begin{aligned} a_1 &= \frac{\mu}{r_p + r_b} \\ a_2 &= -\frac{\mu^2 r_p}{2(r_p + r_b)^3}\frac{\partial r_p}{\partial e_b} \\ a_3 &= \frac{\mu^3 r_p}{6(r_p + r_b)^5}\left[(2r_p - r_b)\left(\frac{\partial r_p}{\partial e_b}\right)^2 - r_p(r_p + r_b)\frac{\partial^2 r_p}{\partial e_b^2}\right] \end{aligned}\right\} \tag{3-41}$$

and

$$e = e_g + \frac{v_p}{\mu} \tag{3-42}$$

For common high-vacuum triodes, Eq. (3-40) converges rapidly enough so that the required accuracy is usually obtained with only a few terms of the series if the current amplitude is not too great. $\partial r_p/\partial e_b$ may be evaluated by plotting a curve of r_p *vs.* e_b determined from the static plate characteristic corresponding to the given operating grid bias. $\partial r_p/\partial e_b$ is the slope of this r_p-e_b curve at the point corresponding to the operating plate voltage. Other higher-order derivatives may be evaluated in a similar manner from curves of derivatives of the next lower order, but the accuracy rapidly decreases with increase of order of the derivative. The value of the series expansion lies not so much in the direct solution of numerical problems, as in the general analysis of the operation of vacuum tubes. In later chapters the series expansion will be applied to the analysis of the operation of vacuum tubes as amplifiers, detectors, modulators, and oscillators.

3-22. Derivation of Series Expansion.[1]—Equations (3-40) and (3-41) are derived by applying Taylor's theorem for two variables to Eq. (3-2). This gives the following equation for alternating plate current:

$$i_p = \frac{1}{1!}\left(\frac{\partial i_b}{\partial e_c}e_g + \frac{\partial i_b}{\partial e_b}e_p\right) + \frac{1}{2!}\left(\frac{\partial i_b}{\partial e_c}e_g + \frac{\partial i_b}{\partial e_b}e_p\right)^2 + \frac{1}{3!}\left(\frac{\partial i_b}{\partial e_c}e_g + \frac{\partial i_b}{\partial e_b}e_p\right)^3 + \cdots \tag{3-43}$$

in which a term of the form $\left(\frac{\partial i_b}{\partial e_c}e_g + \frac{\partial i_b}{\partial e_b}e_p\right)^n$ is a convenient symbolical representation of the terms[2]

[1] The following outline of the derivation of Eqs. (3-40) and (3-41) is included in order to explain the general method and justify the use of these equations throughout the remainder of the book. The student will not be seriously handicapped by omitting Sec. 3-22.

[2] OSGOOD, W. F., "Advanced Calculus," p. 172, The Macmillan Company, New York, 1925.

$$\frac{\partial^n i_b}{\partial e_c{}^n} e_g{}^n + \frac{n}{1!} \frac{\partial^n i_b}{\partial e_c{}^{n-1} \partial e_b} e_g{}^{n-1} e_p + \frac{n(n-1)}{2!} \frac{\partial^n i_b}{\partial e_c{}^{n-2} \partial e_b{}^2} e_g{}^{n-2} e_p{}^2 + \frac{n(n-1)(n-2)}{3!} \frac{\partial^n i_b}{\partial e_c{}^{n-3} \partial e_b{}^3} e_g{}^{n-3} e_p{}^3 + \cdots \tag{3-44}$$

and all derivatives are evaluated at the static operating point.

Although the amplification factor varies with operating voltages, for many purposes sufficient accuracy is obtained by assuming it to be constant. Under this assumption a term of the form $\partial^n i_b / \partial e_c{}^{n-m} \partial e_b{}^m$ becomes $\mu^{n-m} \dfrac{\partial^n i_b}{\partial e_b{}^n}$,* and Eq. (3-43) reduces to

$$i_p = (e_p + \mu e_g) \frac{\partial i_b}{\partial e_b} + \frac{(e_p + \mu e_g)^2}{2!} \frac{\partial^2 i_b}{\partial e_b{}^2} + \frac{(e_p + \mu e_g)^3}{3!} \frac{\partial^3 i_b}{\partial e_b{}^3} + \cdots \tag{3-45}$$

But, at the operating point,

$$\left.\begin{aligned} \frac{\partial i_b}{\partial e_b} &= \frac{1}{r_p} \\ \frac{\partial^2 i_b}{\partial e_b{}^2} &= \frac{\partial}{\partial e_b}\left(\frac{1}{r_p}\right) = -\frac{1}{r_p{}^2} \frac{\partial r_p}{\partial e_b} \\ \frac{\partial^3 i_b}{\partial e_b{}^3} &= \frac{1}{r_p{}^3}\left[2\left(\frac{\partial r_p}{\partial e_b}\right)^2 - r_p \frac{\partial^2 r_p}{\partial e_b{}^2}\right] \end{aligned}\right\} \tag{3-46}$$

Therefore,

$$i_p = \frac{e_p + \mu e_g}{1! r_p} - \frac{(e_p + \mu e_g)^2}{2! r_p{}^2} \frac{\partial r_p}{\partial e_b} + \frac{(e_p + \mu e_g)^3}{3! r_p{}^3}\left[2\left(\frac{\partial r_p}{\partial e_b}\right)^2 - r_p \frac{\partial^2 r_p}{\partial e_b{}^2}\right] + \cdots \tag{3-47}$$

The plate resistance r_p and its derivatives in Eq. (3-47) are evaluated at the operating point.

Vacuum tubes are usually used with some form of load impedance in the plate circuit. It is convenient to consider first the relatively simple case in which the load is a resistance. For resistance load, Eq. (3-27) is

$$e_p = v_p - i_p r_b \tag{3-48}$$

and

$$e_p + \mu e_g = v_p + \mu e_g - i_p r_b = \mu\left(e_g + \frac{v_p}{\mu}\right) - i_p r_b \tag{3-49}$$

or

$$e_p + \mu e_g = \mu e - i_p r_b \tag{3-50}$$

* This may be shown by applying $\partial^n i_b / \partial e_c{}^{n-m} \partial e_b{}^m$ and $\partial^n i_b / \partial e_b{}^n$ to Eq. (3-2).

Substituting Eq. (3-50) in Eq. (3-47) gives

$$i_p = \frac{\mu e - i_p r_b}{1!\,r_p} - \frac{(\mu e - i_p r_b)^2}{2!\,r_p^2}\frac{\partial r_p}{\partial e_b} + \frac{(\mu e - i_p r_b)^3}{3!\,r_p^3}\left[2\left(\frac{\partial r_p}{\partial e_b}\right)^2 - r_p\frac{\partial^2 r_p}{\partial e_b^2}\right] + \cdots \tag{3-51}$$

Equation (3-51) cannot be conveniently used, because i_p appears in an involved manner on the right side of the equation. An explicit expression for i_p may be obtained by assuming that i_p is of the form of Eq. (3-40)

$$i_p = a_1 e + a_2 e^2 + a_3 e^3 + \cdots \tag{3-40}$$

The coefficients of Eq. (3-40) may be evaluated by substituting (3-40) in (3-51), giving

$$\sum_{n=1}^{\infty} a_n e^n = \frac{\mu e - r_b \sum\limits_{n=1}^{\infty} a_n e^n}{1!\,r_p} - \frac{\left(\mu e - r_b \sum\limits_{n=1}^{\infty} a_n e^n\right)^2}{2!\,r_p^2}\frac{\partial r_p}{\partial e_b} + \frac{\left(\mu e - r_b \sum\limits_{n=1}^{\infty} a_n e^n\right)^3}{3!\,r_p^3}\left[2\left(\frac{\partial r_p}{\partial e_b}\right)^2 - r_p\frac{\partial^2 r_p}{\partial e_b^2}\right] + \cdots \tag{3-52}$$

Equation (3-52) is an identity, which must hold for all values of e. This can be true only if for every value of n the summation of terms in e^n on the left side of the identity is equal to the summation of terms of e^n on the right side. Terms in e on the right side arise only from the first term of Eq. (3-52). Therefore,

$$a_1 e = \left(\frac{\mu}{r_p} - \frac{a_1 r_b}{r_p}\right) e \tag{3-53}$$

Terms in e^2 arise from the first two terms of the right side of Eq. (3-50). Therefore,

$$a_2 e^2 = -\left[\frac{a_2 r_b}{r_p} + \frac{(\mu - a_1 r_b)^2}{2 r_p^2}\frac{\partial r_p}{\partial e_b}\right] e^2 \tag{3-54}$$

By the solution of Eqs. (3-53) and (3-54) and a similar equation in e^3, the first three coefficients of Eq. (3-40) may be shown to have the values given by Eqs. (3-41).

3-23. Harmonic Generation and Intermodulation.—Before proceeding to a discussion of the more general problem of a tube with impedance load it is necessary to show that the presence of the second- and higher-order terms of the series is associated with the production of components of alternating plate current of frequencies other than those which are contained in the applied signal. Consider, for instance, the simple case in which the excitation voltage has only a single frequency. Then

$$\begin{aligned} e &= E_m \sin \omega t \\ e^2 &= E_m^2 \sin^2 \omega t = \tfrac{1}{2}E_m^2 - \tfrac{1}{2}E_m^2 \cos 2\omega t \\ e^3 &= \tfrac{3}{4}E_m^3 \sin \omega t - \tfrac{1}{4}E_m^3 \sin 3\omega t \end{aligned} \tag{3-55}$$

Thus the second-order term of the series gives rise to a steady component and to a second-harmonic component in the alternating plate current. The third-order term gives rise to a fundamental and a third-harmonic component of alternating plate current. The production of harmonics in the plate current has been shown graphically in Fig. 3-21 for the simple case in which the static characteristics are assumed to obey a parabolic law, $i_b = A(e_b + \mu e_c)^2$, and z_b is assumed to be zero. If r_b is zero, a_1, a_2, and a_3, as given by Eqs. (3-41), reduce to $\mu \dfrac{\partial i_b}{\partial e_b}$, $\dfrac{\mu^2}{2!}\dfrac{\partial^2 i_b}{\partial e_b^2}$, and $\dfrac{\mu^3}{3!}\dfrac{\partial^3 i_b}{\partial e_b^3}$, respectively. Similarly, any coefficient a_n reduces to $\dfrac{\mu^n}{n!}\dfrac{\partial^n i_b}{\partial e_b^n}$. Since $\dfrac{\partial^3 i_b}{\partial e_b^3}$ and higher-order derivatives of plate current with respect to plate voltage are zero when the static characteristics obey a parabolic law, the plate-current series contains only the first- and second-order terms. Consequently the alternating plate current should consist of fundamental, steady, and second-harmonic components when the excitation is sinusoidal. This is in agreement with Fig. 3-21.

If the excitation voltage contains more than one frequency, the plate current contains not only the impressed frequencies and their harmonics, but also frequencies equal to the sums and differences of the impressed frequencies and their integral multiples, as may be shown by expanding

$$(E_1 \sin \omega_1 t + E_2 \sin \omega_2 t + E_3 \sin \omega_3 t + \cdots)^n.$$

These are called *intermodulation* frequencies. *Intermodulation* is defined as the production, in a nonlinear circuit element, of frequencies equal to the sums and differences of integral (1, 2, 3, etc.) multiples of two or more frequencies which are transmitted to that element. It should be noted that the harmonic and intermodulation frequencies contained in the output are not present in the impressed excitation but are generated by the nonlinear circuit element. It is also important to note that the fact that the intermodulation frequencies are equal to the sums and differences of multiples of the impressed frequencies does not imply that these sum and difference frequencies are the result of interaction of the generated harmonics. The sum and difference frequencies and the harmonics are generated simultaneously by the nonlinear element. It will be shown in later chapters that intermodulation in vacuum-tube circuits is sometimes desirable and sometimes objectionable.

The production of intermodulation frequencies can be demonstrated in a striking manner by a simple laboratory experiment. The voltages from two audio-frequency oscillators are filtered to remove harmonics and are applied in series to the grid circuit of a vacuum tube. The voltage developed across a plate load resistance (preferably considerably smaller than the plate resistance) is applied to the input of a low-pass filter (0 to 3000 cps, for instance), the output of which goes to headphones or, through an amplifier, to a loud-speaker. The oscillator frequencies are made high enough so that the fundamental components of the output voltage cannot pass through the filter. Various combinations of oscillator frequencies can be found at which one or more frequencies are heard in the phones, the pitch of which varies with the tuning of either oscillator. The frequencies are always found to be equal to the difference between one oscillator frequency or one of its multiples and the other oscillator frequency or one of its multiples. Since the application of oscillator harmonics to the grid is prevented by filtering the oscillator voltages, the frequencies that are heard in the output are generated by the tube. A similar experiment, performed with a high-pass filter in the output, demonstrates the production of intermodulation frequencies equal to the sums of the oscillator frequencies and their multiples.

3-24. Series Expansion. *Impedance Load and Variable Amplification Factor.*—The derivation of the series expansion for the more general case of impedance load, although similar in form to that for resistance load, is considerably more complicated. Since the coefficients of the series involve the load impedance, which depends upon the frequency, there must be a coefficient a_n for each frequency arising from the expansion of $e^n = (E_h \sin \omega_h t + E_k \sin \omega_k t + \cdots)^n$, instead of a single coefficient a_n, as with resistance load. This fact may be indicated by writing Eq. (3-40) in the form

$$i_p = \Sigma a_1 e + \Sigma a_2 e^2 + \Sigma a_3 e^3 + \cdots \tag{3-56}$$

An excellent development of the series expansion for the general case of impedance load and variable amplification factor has been given by Llewellyn, who has derived the first two coefficients.[1] These are

$$\left.\begin{aligned} (a_1)_h &= \frac{\mu}{r_p + (z_b)_h} \\ (a_2)_{h-k} &= \\ &\frac{\mu \dfrac{\partial \mu}{\partial e_b}[r_p{}^2 - (z_b)_h(z_b)_k] + \dfrac{\partial \mu}{\partial e_c}[r_p + (z_b)_h][r_p + (\bar{z}_b)_k] - \mu^2 r_p \dfrac{\partial r_p}{\partial e_b}}{2[r_p + (z_b)_h][r_p + (\bar{z}_b)_k][r_p + (z_b)_{h-k}]} \end{aligned}\right\} \tag{3-57}$$

[1] LLEWELLYN, *loc. cit.*

in which

$$(z_b)_h = r_b + j(x_b)_h \qquad (\bar{z}_b)_k = r_b - j(x_b)_k$$
$$(z_b)_{h-k} = r_b + j(x_b)_{h-k}$$

and r_p and μ and their derivatives are evaluated at the operating point. The subscript $_h$ in the symbol $(x_b)_h$ indicates that the reactance of the load is to be evaluated at the frequency $2\pi h$. The coefficient $(a_2)_{h+k}$ is similar to $(a_2)_{h-k}$ except that $(\bar{z}_b)_k$ and $(z_b)_{h-k}$ are replaced by $(z_b)_k$ and $(z_b)_{h+k}$, respectively. The coefficients for the second-harmonic and steady components may be written in the forms $(a_2)_{h+h}$, $(a_2)_{k+k}$, $(a_2)_{h-h}$, $(a_2)_{k-k}$, etc., and found from the formulas for $(a_2)_{h+k}$ and $(a_2)_{h-k}$ by replacing k by h wherever it appears in these formulas.

Because the plate current contains components having different frequencies, the series expansion cannot be stated in terms of effective values of currents and voltages. Since an imaginary component of instantaneous current has no apparent meaning, some question arises as to the significance of the use of complex coefficients in conjunction with instantaneous values of voltage. No difficulty arises, however, if the coefficients are considered as a form of shorthand notation. The magnitude of each coefficient can be used to determine the amplitude of the particular component of plate current, and the angle can be used to determine the phase angle of that component of current. Consider, for instance, the simple case in which e is of the form $E \sin \omega t$. The first coefficient a_1 is $\mu/[r_p + r_b + j(x_b)_\omega]$, and the first term in Eq. (3-56) is $\mu e/[r_p + r_b + j(x_b)_\omega]$, which should be interpreted as a convenient shorthand method of expressing the instantaneous value of the fundamental component of plate current,

$$\frac{\mu E}{\sqrt{(r_p + r_b)^2 + (x_b)_\omega{}^2}} \sin\left[\omega t - \tan^{-1}\frac{(x_b)_\omega}{r_p + r_b}\right].$$

The second term of the series consists of a steady component and a second-harmonic component. The latter is of the complex form

$$\frac{b}{c + jd}\, e^2,$$

which should be interpreted as $\dfrac{bE^2}{\sqrt{c^2 + d^2}} \sin\left(2\omega t - \tan^{-1}\dfrac{d}{c}\right)\cdot$ If the exciting voltage contains two frequencies f_h and f_k whose values are, by way of example, 60 and 100 cps, respectively, the exciting voltage is of the form $E_h \sin 120\pi t + E_k \sin (2000\pi t + \theta)$. The first term of Eq. (3-56) then becomes[1] $\mu e_h/[r_p + (z_b)_{120\pi}] + \mu e_k/[r_p + (z_b)_{2000\pi}]$, which

[1] The symbols $(z_b)_{120\pi}$ and $(x_b)_{120\pi}$ indicate that the impedance and the reactance of the load are to be evaluated at the frequency 120 cps.

should be interpreted as

$$\frac{\mu E_k}{\sqrt{(r_p + r_b)^2 + (x_b)_{120\pi}{}^2}} \sin\left[120\pi t - \tan^{-1}\frac{(x_b)_{120\pi}}{r_p + r_b}\right] + \frac{\mu E_k}{\sqrt{(r_p + r_b)^2 + (x_b)_{2000\pi}{}^2}} \sin\left[2000\pi t + \theta - \tan^{-1}\frac{(x_b)_{2000\pi}}{r_p + r_b}\right].$$

The second term of the series gives rise to a steady component, a 120-cycle component, a 2000-cycle component, a 1060-cycle component, and a 940-cycle component. The amplitude of the latter component can be determined, for example, by multiplying the product E_hE_k by the magnitude of the coefficient $(a_2)_{940}$, which is found by substituting $(z_b)_{2000\pi}$, $(\bar{z}_b)_{120\pi}$, and $(z_b)_{1880\pi}$ in Eq. (3-59).[1] The amplitudes of the other components can be found in a similar manner.

When only the first term of the series expansion is taken into consideration, *i.e.*, when the tube is assumed to be a linear circuit element, the plate current contains only the impressed frequencies and Eq. (3-56) reduces to

$$i_p = \frac{\mu e}{r_p + z_b} \tag{3-58}$$

for each frequency contained in the excitation voltage e. Effective values of current and voltage may then be used and Eq. (3-58) written in the form

$$I_p = \frac{\mu E}{r_p + z_b} = \frac{\mu E_g + V_p}{r_p + z_b} \tag{3-59}$$

No step in the derivation of Eqs. (3-40) and (3-41) restricts these equations to alternating exciting voltage. They apply equally well when e is a small change of applied grid or plate voltage. Equations (3-56) and (3-57), on the other hand, involve impedances that are evaluated at specific frequencies. They cannot, therefore, be used when e is not a periodic function of time. The sudden application of grid or plate excitation or of direct voltages will in general result in the production of transient components of plate current when the grid or plate circuits contain reactance. The theory of the transient behavior of tube circuits is beyond the scope of this book.[2]

3-25. Relation of Series Coefficients to Dynamic Tube Characteristics.—A study of Eqs. (3-41) and (3-57) shows that the introduction of load impedance into the plate circuit reduces the ratios of the coefficients

[1] The amplitude of a plate-current component whose frequency is $nf_h \pm mf_k$ is equal to $E_h{}^nE_k{}^m$ multiplied by the magnitude of the coefficient $(a_{n-m})_{h\pm k}$.

[2] JACKSON, W., *Phil. Mag.*, **13**, 143, 735 (1932); SCHLESINGER, K., *E. N. T.*, **38**, 144 (1931).

of the higher-order terms of the series to the first-order coefficient. This can be shown most readily for the special case of resistance load. For resistance load, Eqs. (3-57) give the following expression for the ratio of the coefficient of the second term to the coefficient of the first term of the series:

$$\frac{a_2}{a_1} = \frac{\frac{\partial \mu}{\partial e_b}(r_p{}^2 - r_b{}^2) + \frac{\partial \mu}{\partial e_c}\frac{(r_p + r_b)^2}{\mu} - \mu r_p \frac{\partial r_p}{\partial e_b}}{2(r_p + r_b)^2} \tag{3-60}$$

As r_b is increased, the ratio is reduced and approaches the limiting value, $\frac{1}{2}\left(\frac{1}{\mu}\frac{\partial \mu}{\partial e_c} - \frac{\partial \mu}{\partial e_b}\right)$. If μ is practically constant, as is true over certain ranges in many tubes, the ratio of a_2 to a_1 decreases rapidly and becomes negligibly small as r_p becomes large. If μ is constant, the introduction into the plate circuit of a load resistance equal to twice the plate resistance reduces a_1 to $\frac{1}{3}$ its value for zero load resistance and a_2 to $\frac{1}{27}$ its value for zero load resistance. The ratios of a_3 and higher-order coefficients to a_1 decrease even more rapidly with increase of load resistance than that of a_2 to a_1. A similar analysis for the more general case of impedance load, although not quite so simple, also shows that in general the amplitudes of the second- and higher-order terms of the series are decreased relative to the amplitude of the first term by an increase of load impedance and, if μ were constant, could be reduced to any desired degree by making z_b/r_p sufficiently high.

Equations (3-40) and (3-41) show that for resistance load the dynamic transfer characteristic would be a straight line if the series contained only the first term and that the presence of the higher-order terms is associated with curvature of the dynamic transfer characteristic. If only the first term were present, a change of grid voltage Δe_c would result in a proportional change of plate current Δi_b, indicating a linear relation between i_b and e_c. The higher-order terms of the series destroy this linearity, since contributions to Δi_b from the higher-order terms are not proportional to Δe_c. Reduction of the higher-order terms of the series by increase of load resistance is, therefore, accompanied by reduction of the curvature of the dynamic transfer characteristic.

That the transfer characteristic of a triode is straightened in the negative range of grid voltage by the introduction of plate load resistance and, hence, that the coefficients of the second and higher terms of the series are reduced with respect to the first term can also be shown by constructing dynamic transfer characteristics for different values of load resistance. Although the dynamic transfer characteristics can be most readily obtained from the plate family of characteristics in a manner that will be explained in Chap. 4, it is instructive at this point to derive one from the static transfer characteristics. O represents the static operating point in Fig. 3-22. With no external resistance in the plate circuit a change of grid voltage from E_c to $E_c + \Delta e_c$ would cause the current to change to i_b', corresponding to a new point a on the same static characteristic. With resistance in the plate circuit, however, the increase of plate current accompanying the change in grid voltage increases the IR drop in the plate circuit and thus reduces the voltage of the plate to a new value $E_{bo} - R_b\Delta i_b$, corresponding to point a'. Similarly, if the grid voltage is changed to $E_c - \Delta e_c$, the point shifts to b'. By plotting corresponding values of i_b and e_c it is possible to obtain the complete dynamic transfer characteristic, shown by the dashed line. Figure 3-23 shows the dynamic transfer characteristics of a type 56 triode for various values of resistance load, the plate supply voltage being increased with R_b in such a manner as to maintain the same static operating point. Figure 3-24 shows similar curves for constant plate supply voltage.

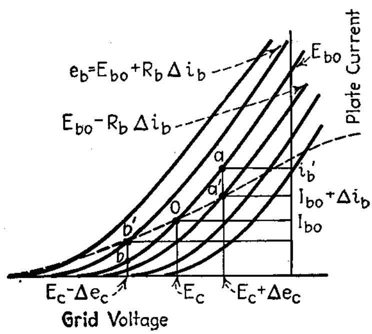


FIG. 3-22.—Construction of a dynamic transfer characteristic from the family of static transfer characteristics.

In the positive range of grid voltage, the introduction of load resistance into the plate circuit may increase the curvature of the transfer

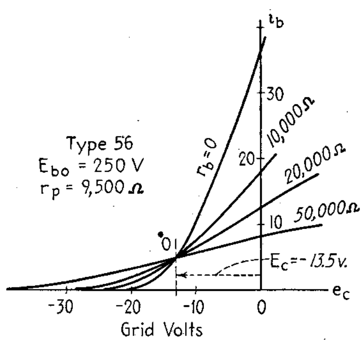


FIG. 3-23.—Dynamic transfer characteristics of the type 56 triode for four values of load resistance at fixed operating voltages. $r_b = R_b$.

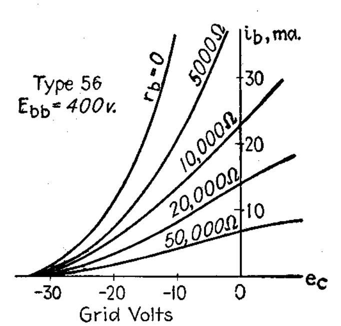


FIG. 3-24.—Dynamic transfer characteristics of the type 56 triode for five values of load resistance at fixed plate supply voltage.

characteristic. The reason for this is made apparent by reference to Fig. 3-6. Since the amplification factor is equal to the slope of the e_b-e_c characteristics, it is evident from Fig. 3-6 that the amplification factor varies greatly in the positive region of grid voltage. Hence the ratio

a_2/a_1 given by Eq. (3-60) does not approach zero as r_b is increased, but increases with r_b at large values of r_b. The increase of curvature with increase of load resistance can also be explained readily physically. As pointed out on page 49, the bending over of the transfer characteristics at positive grid voltages results in part from diversion of electrons to the grid. At a given positive grid voltage, the number of electrons diverted to the grid increases as the plate voltage is decreased. The increase of *IR* drop resulting from an increase of load resistance decreases the plate voltage. More electrons are diverted to the grid, and so the plate current is decreased. Increase of load resistance therefore increases the bending over of the transfer characteristics at positive grid voltages and may even cause them to have maxima.

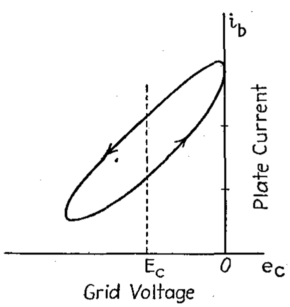


FIG. 3-25.—Typical dynamic transfer characteristic for inductive load.

At large values of grid excitation voltage, the amplification factor of pentodes is not constant throughout the negative-grid range of operation and, although the curvature of the transfer characteristic first decreases with increase of load resistance, a value is reached beyond which curvature again increases (see Sec. 6-25).

When the load contains reactance, as well as resistance, the behavior is complicated by the phase difference between the plate voltage and plate current. If the impedance is sufficiently high, the dynamic transfer characteristic obtained with sinusoidal excitation is nearly elliptical. For lower values of impedance the dynamic transfer characteristic resembles an ellipse but has a curved axis, as shown for inductive reactance in Fig. 3-25. As the reactance is reduced, the path of operation gradually changes into the curve obtained for pure resistance (see Sec. 4-6).

3-26. Dynamic Plate Resistance.—The *dynamic plate resistance* is the quotient of the alternating plate voltage by the inphase component of the alternating plate current, all other electrode voltages being maintained constant. The quadrature component of current, which results from electrode capacitances and, at ultrahigh frequencies, from electron transit time, is negligible at low audio frequency. The inphase component of plate current corresponding to a given alternating plate voltage when other electrode voltages are constant may be found from Eqs. (3-40) to (3-42) by making e_g and r_b zero. When r_b is zero $v_p = e_p$, and Eq. (3-40) becomes

$$i_p = \frac{e_p}{r_p} - \frac{e_p^2}{2r_p^2}\frac{\partial r_p}{\partial e_b} + \frac{e_p^3}{6r_p^3}\left(2\frac{\partial r_p}{\partial e_b} - r_p\frac{\partial^2 r_p}{\partial e_b^2}\right) + \cdots \tag{3-61}$$

Because of curvature of the static plate characteristics, the derivatives of the plate resistance are not zero. When the excitation voltage is high, therefore, the second- and higher-order terms of Eq. (3-61) are not negligible. Contributions of the third- and other odd-order terms to the fundamental component of plate current cause the ratio of the alternating plate voltage to the fundamental component of alternating plate current to differ from r_p, the static plate resistance. As e_p is reduced, however, the ratio of the higher-order terms to the first term decreases, and e_p/i_p approaches r_p. If the amplitude of the plate excitation voltage is small enough, the dynamic plate resistance approximates the value determined from the slope of the static characteristic at the operating point (at frequencies low enough so that the electron transit time is negligible). The static plate resistance may, therefore, be found dynamically by using a sufficiently small plate excitation voltage. In a similar manner it may be shown that the dynamic measurement of transconductance and of variable amplification factor requires the use of small excitation voltages.

Problems

3-1. From the static plate characteristics of the type 6J5 tube find the values of μ, r_p, and g_m at the point $E_b = 200$ volts, $E_c = -8$ volts.

3-2. From the static plate characteristics of the type 6SJ7 tube, find the values of μ, r_p, and g_m at the point $E_b = 200$ volts, $E_c = -2$ volts.

3-3. *a.* From the following data find approximate values of μ, r_p, and g_m at the point $E_b = 180$ volts, $E_c = -12.5$ volts. (Note that only two of these factors can be found directly from the data given.)

E_b, volts	E_c, volts	I_b, ma
180	−12.5	7.5
160	−10.0	7.5
180	−12.3	7.84

b. From the following data find approximate values of μ, r_p, and g_m at the point $E_b = 250$ volts, $E_c = -16$ volts.

E_b, volts	E_c, volts	I_b, ma
250	−16	2.0
220	−14	2.0
260	−16	3.0

3-4. Determine the frequencies of all components of alternating plate current associated with the first three terms of the plate-current series when the frequencies 60, 100, and 900 cps are simultaneously impressed upon the grid of a vacuum tube.

3-5. Equation (3-2) suggests that it should be possible to express the alternating plate current by the following series, in which μ is assumed to be constant:

$$i_p = k_1(e_p + \mu e_g) + k_2(e_p + \mu e_g)^2 + k_3(e_p + \mu e_g)^3 + \cdots \tag{3-62}$$

By differentiating both sides of this equation successively and evaluating each derivative at the operating point, show that $k_n = \left.\frac{\partial^n i_b}{\partial e_b{}^n}\right]_0 \times \frac{1}{n!}$, and thus derive Eq. (3-45).

3-6. *a.* Show that, if the amplification factor is constant and the load impedance zero, the coefficients a_n of the series expansion for i_p reduce to $\frac{\mu^n}{n!}\frac{\partial^n i_b}{\partial e_b{}^n}$ [see Eq. (3-45)].

b. Show that, if the static characteristics are assumed to be parabolic, all coefficients of terms of the series expansion for i_p of higher order than the second are zero.

c. By means of the series expansion show that for parabolic characteristics, constant amplification factor, and zero load impedance the amplitudes of the steady and second-harmonic components of plate current are one-fourth the amplitude of the fundamental component when the grid bias E_c is one-half the cutoff value and $e = e_g = E_c \sin \omega t$.

3-7. Show that Eqs. (3-57) reduce to Eqs. (3-41) if μ is constant and the load is nonreactive.

3-8. Evaluate a_1 and a_2 graphically for a type 6J5 triode with resistance load at rated operating voltages when (*a*) $r_b = 0$; (*b*) $r_b = r_p$; (*c*) $r_b = 2r_p$.

CHAPTER 4

METHODS OF ANALYSIS OF VACUUM TUBES AND VACUUM-TUBE CIRCUITS

Since the application of vacuum tubes is governed in part by the extent to which it is possible to analyze the operation of tubes and associated circuits, a study of the methods of analysis is of considerable importance. This chapter will, therefore, present analytical and graphical methods of analysis of high-vacuum tubes and their circuits.

4-1. The Equivalent Plate Circuit.[1]—In the solution of many vacuum-tube problems, particularly when the excitation voltage is small and the load impedance high, sufficient accuracy is obtained by making use of only the first term of the series expansion for the plate current [Eqs. (3-58) and (3-59)]. To a first approximation the alternating plate current is then given by the equations

$$I_p = \frac{\mu E}{r_p + z_b} = \frac{\mu E}{r_p + r_b + jx_b} \tag{4-1}$$

$$= \frac{\mu E}{\sqrt{(r_p + r_b)^2 + x_b^2}} \underline{/\theta} \tag{4-2}$$

where

$$E = \frac{V_p}{\mu} + E_g \qquad \text{and} \qquad \theta = -\tan^{-1} \frac{x_b}{r_p + r_b}$$

In these equations r_p is the plate resistance at the operating point, and z_b is the impedance of the plate load at the frequency of the applied voltage. The value of I_p given by Eqs. (4-1) and (4-2) is the same as that which would flow in the simple series circuit of Fig. 4-1*a*. This *equivalent plate circuit* may, therefore, be used to find the approximate value of the fundamental component of plate current and of the currents in the various branches of the load impedance. This fact may be expressed in the form of the *equivalent-plate-circuit theorem for amplification,* which states that the fundamental component of alternating plate current is approximately equal to the current that flows as the result of the application of

[1] Miller, J. M., *Bur. Standards Bull.* **15,** 367 (1911); *Proc. I.R.E.*, **6,** 141 (1918); Van der Bijl, H. J., *Phys. Rev.*, **12,** 171 (1918); *Proc. I.R.E.*, **7,** 97 (1919); Chaffee, E. L., *Proc. I.R.E.*, **17,** 1633 (1929); Goodhue, W. M., *Electronics*, December, 1933, p. 341; Richter, W., *Electronics*, March, 1936, p. 19; Landon, V. D., *Proc. I.R.E.*, **18,** 294 (1930).

the voltage $V_p + \mu E_g$ to an equivalent circuit consisting of the external load impedance in series with a constant resistance equal to the a-c plate resistance of the tube at the operating point.

Equation (4-1) may also be written in the form

$$E_{zb}(g_p + y_b) = V_p g_p + E_g g_m \tag{4-3}$$

Equation (4-3) suggests the parallel equivalent circuit of Fig. 4-1*b*. The ease of combining admittances sometimes makes it advantageous to use Eq. (4-3) and the parallel equivalent circuit, in place of Eq. (4-1) and the series equivalent circuit, particularly when the load impedance consists of a number of parallel branches (see Appendix, Sec. A-1).

By making use of the first term of the series expansion for the grid current, equations analogous to Eqs. (4-1) and (4-3) may be derived for the grid circuit of a triode.

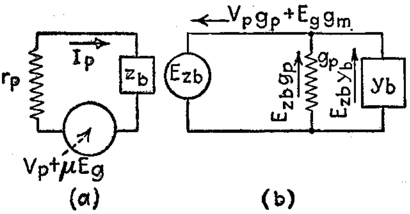


FIG. 4-1.—Series and parallel equivalent plate circuits.

$$I_g = \frac{V_g + \mu_g E_p}{r_g + z_c} \tag{4-4}$$

and

$$E_{zc}(g_g + y_c) = V_g g_g + E_p g_n \tag{4-5}$$

The equivalent grid circuits are of the same form as Figs. 4-1*a* and 4-1*b*, the subscripts $_p$, $_g$, and $_b$ being replaced by $_g$, $_p$, and $_c$, respectively and μ by μ_g. Equations (4-1) to (4-5) may also be written in terms of crest values of currents and voltages. Similar equations may be written for any electrode of a multigrid tube.

Equations (4-1) to (4-5) are based upon Fig. 3-18, in which the load impedance is assumed to be free of e.m.fs. and the plate excitation voltage V_p is assumed to be in series with z_b. The load impedance is in general made up of a number of branches that may contain other e.m.fs., in addition to or in place of the e.m.f. in series with the entire impedance. Application of the principle of superposition, which is valid under the assumption that the second- and higher-order terms of the plate current series may be neglected, shows that the presence of additional e.m.fs. within the load does not affect the currents caused by E_g or by a voltage V_p in series with z_b. The equivalent circuit will, therefore, give the correct values of currents if the additional voltages are included in the equivalent circuit in the positions that they occupy in the actual circuit. The equivalent circuit must include all portions of the actual circuit that are conductively, inductively, or capacitively coupled to the plate.

If the grid circuit contains impedances through which alternating current flows, as in Figs. 4-2 and 4-30, the value of the alternating grid voltage E_g may differ from the excitation voltage V_g applied from an external source. V_g may, in fact, be derived entirely from the flow of alternating currents through impedances contained in the grid circuit. In general, therefore, it is necessary to evaluate E_g before the network equations for the equivalent circuit can be solved. E_g is usually found most readily from the actual circuit and is equal to the vector sum of all alternating voltages between the cathode and the grid along any continuous path.

The use of equivalent circuits greatly simplifies the solution of many problems involving tubes and associated circuits. Once the equivalent circuit has been formed, it is not even necessary to know that a vacuum tube is involved, as fundamental currents and voltages at various points may be computed by the ordinary methods of a-c circuit analysis. The validity of the equivalent circuit in a given problem depends upon the desired accuracy. If only qualitative results are desired, it is often possible to apply the equivalent circuit even when harmonic production is high. Where a high degree of accuracy is essential, the use of the equivalent circuit is valid only at small amplitudes. It cannot be applied when the conditions of operation are such that current ceases to flow during a portion of the cycle of exciting voltage (see Sec. 5-15). The equivalent circuit obviously gives no indication of the production of harmonics and intermodulation frequencies. The acceptance of the equivalent circuit is, therefore, tantamount to the assumption that the tube is a linear circuit element.

Since the grid conducts only when it is positive and since the grid characteristics are not linear, the flow of grid current through external impedance distorts the alternating grid voltage. Furthermore, the grid current varies with plate voltage. For these reasons the simple equivalent plate circuit is not applicable with accuracy when grid current flows through a grid-circuit impedance. Ordinarily, however, in circuits in which appreciable grid current is allowed to flow, the conditions of operation are such that the equivalent circuit would not be applicable even if no grid current flowed. In such circuits it is necessary to resort to graphical or experimental methods of analysis.

4-2. Construction of Equivalent Plate Circuits.—Use of the following procedure ensures the correct formation and solution of series equivalent plate circuits. (See Appendix, Sec. A-1, for parallel-circuit procedure.) With suitable changes of symbols it may be applied to the circuit of any electrode.

1. Insert the equivalent voltage μE_g and the a-c plate resistance r_p in series between the plate and the cathode, indicating the polarity of the equivalent voltage to be such that the cathode side of the voltage is positive, as in Fig. 4-2*b*.[1]

[1] The polarities indicate the vector relations between the voltages, a change in sign being equivalent to a reversal in the direction of the vector representing the voltage. The polarities may also be construed to apply to the instantaneous voltages at some

If alternating voltage is impressed upon more than one grid, there is an equivalent voltage in series with r_p for each grid that is excited.

2. Assume polarities for the grid and plate excitation voltages V_g and V_p. It is convenient to indicate the polarities to be such as to make the grid and plate positive relative to the cathode. (When excitation voltages that are in phase opposition are applied simultaneously to two electrode circuits, however, only one of the excitation voltages can be indicated as having such polarity as to make its electrode positive relative to the cathode.)

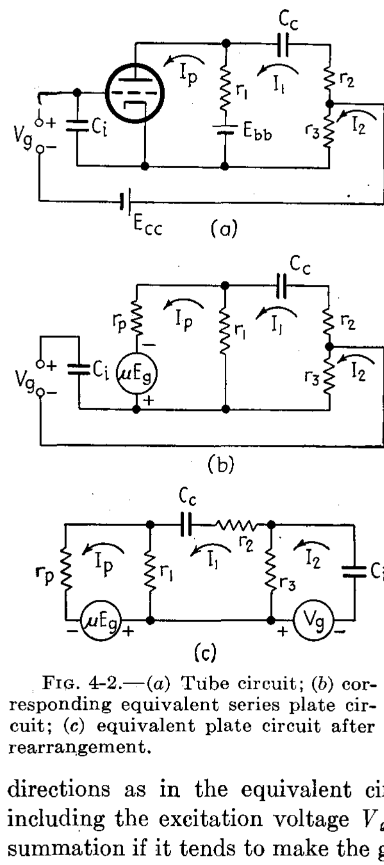


FIG. 4-2.—(a) Tube circuit; (b) corresponding equivalent series plate circuit; (c) equivalent plate circuit after rearrangement.

3. Assume positive directions for the various alternating circuit currents. It is sometimes convenient to choose the positive direction of plate current as that in which the equivalent voltage μE_g tends to cause it to flow, *i.e.*, through r_p from plate to cathode, as shown in Fig. 4-2, but this is not essential.

4. Delete the tube symbol (or show dotted), all electrode supply voltages and other direct voltages, and all circuit branches not coupled to the plate (such as the screen circuit of a screen-grid tetrode, as it is usually used).

5. Redraw the resulting equivalent circuit in the form in which it may be most readily analyzed.

6. Express the alternating grid voltage E_g as the vector sum of all alternating voltages between the cathode and the grid along any continuous path. This can usually be done most conveniently by reference to the original circuit, all currents and excitation voltages being indicated in the same directions as in the equivalent circuit. Any voltage that contributes to E_g, including the excitation voltage V_g, must be preceded by a positive sign in this summation if it tends to make the grid positive relative to the cathode, and by a negative sign if it tends to make the grid negative relative to the cathode. If alternating voltage is impressed upon more than one grid, each equivalent grid voltage in the equivalent circuit must be evaluated in this manner. It should

instant in the cycle. When g_m is positive, a positive increment of grid voltage produces a positive increment of plate current. Likewise, when r_p is positive, a positive increment of plate voltage produces a positive increment of plate current. Hence, when E_g makes the grid positive relative to the cathode, the polarity of the equivalent voltage μE_g must be such as to cause plate current of the same direction as is produced by a positive plate voltage. This is ensured by making the polarity of the equivalent voltage such that it tends to cause current to flow into the plate and calling E_g positive if it makes the grid positive relative to the cathode.

be noted that V_g may be zero in some circuits, E_g being derived entirely from the flow of alternating currents through impedances contained in the grid circuit, as in Fig. 4-30*m*.

7. Write network equations for each loop of the equivalent circuit. These equations, in conjunction with the expression for E_g found in step (6), may be solved simultaneously to find all circuit currents and voltages. A negative value of any current found in this manner merely indicates that the current is opposite in direction to the assumed positive value, *i.e.*, that its phase is opposite to that assumed in the equivalent circuit.

The method of constructing equivalent plate circuits and of evaluating E_g can be shown best with the aid of specific examples. Application of steps (1) to (4) of the above procedure to Fig. 4-2*a* gives the equivalent circuit of Fig. 4-2*b*, which may be rearranged in the more convenient form of Fig. 4-2*c*. The most direct path between the cathode and the grid is through C_i. The only alternating voltage between the cathode and the grid along this path is that resulting from the flow of I_2 through C_i. Hence $E_g = -I_2/j\omega C_i$. The minus sign must be used because the flow of I_2 in the indicated positive direction causes C_i to charge in such polarity as to make the grid negative relative to the cathode. An alternative path from cathode to grid is through r_3 and V_g. Summation of voltages along this path shows that E_g is also equal to

$$+V_g + (I_2 - I_1)r_3.$$

I_2r_3 is positive because the flow of I_2 in the indicated positive direction produces a voltage across r_3 that tends to make the grid positive relative to the cathode. Either of these two expressions for E_g, together with the three equations obtained by summing voltages in the three loops of the equivalent circuit, may be solved simultaneously to find the values of the currents in terms of the circuit constants and V_g. If C_i were omitted, I_2 would be zero and E_g would be equal to $V_g - I_1r_3$. The right-hand loop of the equivalent circuit of Fig. 4-2*c* would then also be absent.

Alternating voltage may be applied simultaneously to two or more electrodes in order to produce an alternating plate current, as in the circuit of Fig. 4-3*a*. In this circuit the excitation voltage V_{g1} is impressed upon the first grid, but the flow of I_1 through the resistance r produces an alternating component of suppressor voltage. The equivalent circuit, shown in Fig. 4-3*b*, therefore contains the two equivalent voltages μE_{g1} and $\mu_{p3}E_{g3}$.[1] E_{g1} is equal to V_{g1}, and E_{g3} is equal to $-I_1r$. Since the screen voltage is constant, the equivalent plate circuit does not contain an equivalent screen voltage. If the screen circuit contained an impedance, however, the flow of alternating screen current through this

[1] $\mu_{p3} = -\partial e_p/\partial e_{g3}$ at constant i_p.

impedance would produce an alternating component of screen voltage, which would affect the alternating plate current. The equivalent plate circuit would then contain a third equivalent voltage, $\mu_{p2}E_{g2}$. Since E_{g2} would have to be found from the equivalent screen circuit, which would in itself contain three equivalent voltages, two of which would depend upon the plate current, it is evident that the analysis of the circuit would be very involved. Fortunately, however, this difficulty is seldom

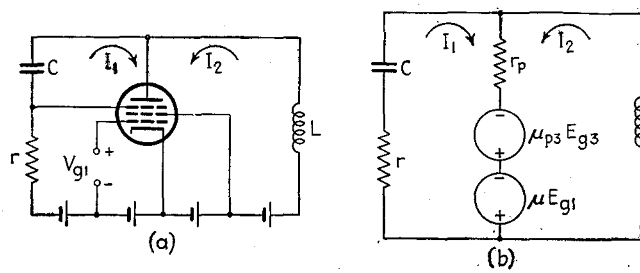


FIG. 4-3.—(*a*) Pentode circuit with excitation applied to both first and third grid circuits; (*b*) equivalent plate circuit.

encountered in practice, because the impedance of the screen circuit is ordinarily made negligible at the frequency of the excitation voltage.

Figure 4-4*a* shows a circuit in which the excitation is applied to the plate circuit. Through voltage-divider action of the circuit impedances, a portion of the impressed voltage V_p is also applied to the grid. The seven-step procedure outlined at the beginning of this section may be applied, V_p being for convenience indicated in such polarity as to make the plate positive. The equivalent circuit is shown in Fig. 4-4*b*. Flow

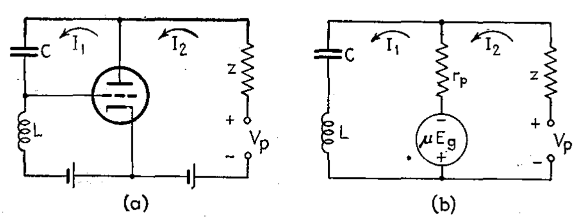


FIG. 4-4.—(*a*) Triode with excitation applied to both plate and grid circuits; (*b*) equivalent plate circuit.

of I_1 in the indicated positive direction tends to make the grid positive. Hence $E_g = +j\omega LI_1$.

4-2A. Equivalent Circuits for Electrodes Other Than the Plate.—The procedure of Sec. 4-2 is not restricted to the plate circuit, but may, with suitable changes in symbols, be used in forming the equivalent circuit of any electrode of an amplifier tube excited by alternating voltages impressed upon the circuit of that or of any other electrode. In the circuit of Fig. 4-5*a*, for instance, the excitation is impressed upon the circuit of the third grid in order to vary the second-grid current. The equivalent second-grid circuit, formed by a five-step procedure analogous to that outlined in Sec. 4-2 for plate circuits, is shown in Fig. 4-5*b*. Because no alternating voltages are impressed upon the first grid or upon the plate, the equivalent circuit contains only one equivalent voltage, $\mu_{23}E_{g3}$.[1] The alternating third-grid voltage E_{g3} is equal to $+V_{g3} - I_1r_1$.

Any two of the three tube factors relating two electrodes of a multigrid tube may be negative (see Secs. 3-8, 3-9, 6-39, 10-3, and 10-23). Negative tube factors do not change the method of forming the equivalent circuit or of evaluating the alternating grid voltages but may affect the numerical

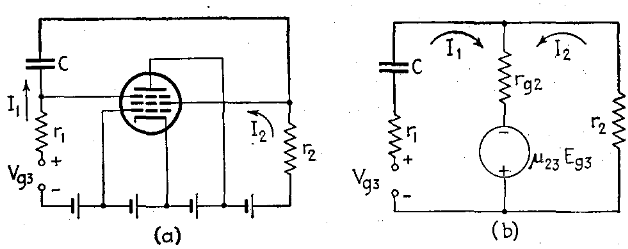


FIG. 4-5.—(*a*) Pentode with excitation applied to the suppressor circuit in order to produce an alternating screen current; (*b*) equivalent screen circuit.

solution of the equivalent-circuit equations or the interpretation of an equation obtained by the solution of the circuit equations. The fact that tube factors are negative need not be taken into account until numerical values are introduced into the equations or their solution.

4-3. Tube Capacitances and Admittances.—The complete equivalent circuit is not so simple as that which has been presented up to this point. The small interelectrode capacitances and the conductances caused by surface leakage or by electron or ion currents between grid and cathode or grid and plate cannot always be neglected. Fortunately this fact does not invalidate the equivalent-circuit theorems, inasmuch as the interelectrode capacitances and conductances act the same as equal capacitances and conductances connected externally between the electrodes. It is permissible, therefore, to add them to the simple equivalent grid and plate circuits as parts of the external impedances. In modern tubes of good manufacture the leakage conductances are so small that they can usually be neglected in comparison with other tube and circuit impedances. Figure 4-6 shows the complete equivalent circuit of a triode (or any tube in which only the plate and control-grid voltages are allowed to vary) with negligible leakage conductances.

[1] $\mu_{23} = -\delta e_{g2}/\delta e_{g3}$ at constant i_{g2}.

The rather involved formulas for the total grid and plate admittances which may be derived by the solution of the network of Fig. 4-6[1] are seldom useful. This is partly because of the complexity of the formulas and partly because the flow of grid current through z_c during only a portion of the cycle causes the alternating grid voltage to depart markedly from sinusoidal form and thus destroys the accuracy of the formulas. More frequently used are the somewhat less complicated formulas for the case

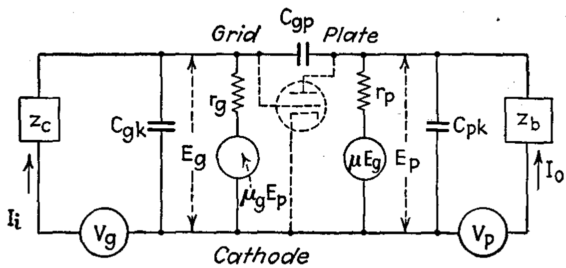


FIG. 4-6.—Complete equivalent circuit of triode with negligible electrode leakage conductances. C_{gk}, C_{gp}, and C_{pk} are the interelectrode capacitances.

in which the grid bias is sufficiently great to prevent the flow of thermionic grid current. These are[2]

$$Y_g = \frac{\omega^2 C_{gp}{}^2(g_p + g_b) + \omega C_{gp} g_m[\omega(C_{gp} + C_{pk}) - b_b]}{[\omega(C_{gp} + C_{pk}) - b_b]^2 + [g_p + g_b]^2} + j\left\{\omega(C_{gk} + C_{gp}) + \frac{\omega C_{gp} g_m (g_p + g_b) - \omega^2 C_{gp}{}^2[\omega(C_{gp} + C_{pk}) - b_b]}{[\omega(C_{gp} + C_{pk}) - b_b]^2 + [g_p + g_b]^2}\right\} \tag{4-6}$$

$$Y_p = g_p + \frac{\omega^2 C_{gp}{}^2 g_c + \omega C_{gp} g_m[\omega(C_{gp} + C_{gk}) - b_c]}{[\omega(C_{gp} + C_{gk}) - b_c]^2 + g_c{}^2} + j\left\{\omega(C_{pk} + C_{gp}) + \frac{\omega C_{gp} g_m g_c - \omega^2 C_{gp}{}^2[\omega(C_{gp} + C_{gk}) - b_c]}{[\omega(C_{gp} + C_{gk}) - b_c]^2 + g_c{}^2}\right\} \tag{4-7}$$

where

$$Y_g = \frac{I_i}{E_g} \quad \text{and} \quad Y_p = \frac{I_o}{E_p}$$

Simplifications that facilitate the use of Eqs. (4-6) and (4-7) can usually be made in particular applications.

[1] CHAFFEE, E. L., *Proc. I.R.E.*, **17**, 1633 (1929); COLEBROOK, F. M., *Wireless Eng.*, **10**, 657 (1933). See also J. M. MILLER, *Bur. Standards Sci. Paper* 351 (1919); S. BALLANTINE, *Phys. Rev.*, **15**, 409 (1920).

[2] These formulas neglect electron transit time, which must be taken into consideration at ultrahigh frequencies (see footnote 1, p. 97).

Inspection of Eq. (4-6) shows that the effective input capacitance of a vacuum tube is increased by the flow of plate current. An insight into the physical reason for this change, as well as useful approximate expressions for the effective input capacitance and the input conductance, can be obtained by an approximate analysis based upon Fig. 4-7. The input current I_i is the vector sum of two components: I_1, which flows through C_{gk}, and I_2, which flows through C_{gp}. The voltage which causes I_1 is E_g, whereas that which causes I_2 is the vector difference[1] between E_g and E_{zb}. E_{zb} in general differs from E_g in both phase and magnitude. This fact may be expressed symbolically by the relation $E_{zb} = AE_g$, where A is a vector quantity that is ordinarily greater than unity in magnitude.[2] Then the vector difference between E_g and E_{zb} is $E_g(1 - A)$ and the input current is

$$I_i = I_1 + I_2 = E_g j\omega C_{gk} + E_g j\omega C_{gp}(1 - A) \qquad (4\text{-}8)$$

and the input admittance is

$$Y_g = \frac{I_i}{E_g} = j\omega[C_{gk} + C_{gp}(1 - A)] \qquad (4\text{-}9)$$

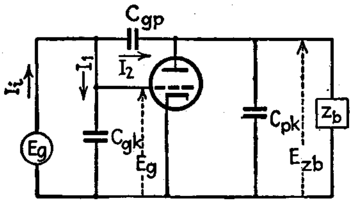


FIG. 4-7.—Circuit of a triode with impedance load, showing interelectrode capacitances.

Reactance in the load z_b in general causes A to have an imaginary component and hence Y_g to have a conductive component. When z_b is a pure resistance r_b, which is small in comparison with the reactance of C_{pk} and C_{gp} in parallel, however, the imaginary component of A is negligible and the equivalent plate circuit shows A to have the value $-\mu r_b/(r_p + r_b)$. Y_g is then purely capacitive and the equivalent input capacitance is

$$C_i = \frac{Y_g}{j\omega} = C_{gk} + C_{gp}(1 - A) = C_{gk} + C_{gp}\left(1 + \frac{\mu r_b}{r_b + r_p}\right) \qquad (4\text{-}10)$$

For pure reactance load x_b, small in comparison with the reactance of $C_{pk} + C_{gp}$, the equivalent plate circuit shows A to be $-j\mu x_b/(r_p + jx_b)$. The input admittance is

$$Y_g = j\omega\left[C_{gk} + C_{gp}\left(1 + \mu\frac{x_b^2}{r_p^2 + x_b^2}\right)\right] - \mu\frac{r_p x_b \omega C_{gp}}{r_p^2 + x_b^2} \qquad (4\text{-}11)$$

The effective input capacitance is

$$C_i = C_{gk} + C_{gp}\left(1 + \mu\frac{x_b^2}{r_p^2 + x_b^2}\right) \qquad (4\text{-}12)$$

[1] It is the vector difference, rather than the vector sum, because both E_g and E_{zb} are measured relative to the cathode.

[2] It will be seen in Chap. 5 that A is called the voltage amplification (see Sec. 5-4).

Equations (4-10) and (4-12) may be written in the form

$$C_i = C_{gk} + C_{gp}(1 + |A|) \tag{4-13}$$

If the reactance of $C_{pk} + C_{gp}$ is large in comparison with r_p, C_i increases with the ratio of r_b or x_b to r_p, approaching the limiting value

$$\text{Max. } C_i = C_{gk} + C_{gp}(1 + \mu) \tag{4-13a}$$

The input conductance for reactive load, given by the last term of Eq. (4-11), is zero when x_b is zero, rises to a maximum value at $x_b = r_p$, and then falls with further increase of x_b. When the load is an inductance L_b, the input conductance is negative and has an approximate maximum value

$$\text{Max. } G_g = -\tfrac{1}{2}\mu\omega C_{gp} \tag{4-13b}$$

It should be remembered that this analysis is an approximation and that Eqs. (4-9) to (4-13b) yield sufficiently accurate results only when the reactance of $C_{pk} + C_{gp}$ is considerably larger than r_p and z_b. The effece of the capacitance $C_{pk} + C_{gp}$ is usually to decrease A and thus to reduce the effective input capacitance below the value given by Eq. (4-13a). This equation is, therefore, useful in indicating approximately the largest value of C_i that is likely to be obtained. At 10,000 cps the reactance of $C_{pk} + C_{gp}$ is of the order 1 to 3 megohms in receiving tubes and so does affect the value of Y_g of pentodes, tetrodes, and high-mu triodes even at audio frequencies. For accurate values of input admittance it is necessary to use Eq. (4-6). Equations (4-9) to (4-13b) may be obtained from Eq. (4-6) by making b_b or g_b zero and assuming that $\omega(C_{gp} + C_{pk})$ is much larger than g_p and g_b. The approximate analysis shows that the increase of effective input capacitance over that obtained when the cathode is cold results from the action of E_{zb} in sending current through C_{gp}.

The fact that the input conductance may be negative when the plate load contains inductance is of importance because oscillation may take place when an oscillatory circuit shunts a negative resistance (see Secs. 10-17 and 10-21). It will be shown in Chap. 6 that the input capacitance must be taken into consideration in the analysis of vacuum-tube amplifiers at frequencies above about 2000 cps. Special tubes having very low interelectrode capacitance have been designed for use at high radio frequencies. The 954 and 955 tubes shown in Fig. 2-16 are examples of such tubes.

It was pointed out in the foregoing analysis that the current that flows into C_{gp} is caused to flow by the vector difference of E_{zb} and E_g. If the plate load is nonreactive, E_{zb} is 180 degrees out of phase with E_g, the input current leads E_g by 90 degrees, and the input admittance is purely susceptive. If the load contains reactance, on the other hand, E_{zb} is not exactly in phase opposition to E_g, and so the input current has a component that is in phase with E_g, and the admittance has a conductive component. At frequencies so low that electron transit time has a negligible effect, the input admittance can have an appreciable conductive component only when the load (including C_{pk}) has an appreciable reactive component. If the frequency of the input voltage is so high that the transit time is of the same order of magnitude as the period, the electrons arrive at the plate an appreciable fraction of the cycle after they leave the cathode. Since the number that leave the cathode is determined by the grid voltage at the instant that they leave, it follows that the plate current lags the grid voltage. E_{zb} is, therefore, not 180 degrees out of phase with E_g, even when the load is nonreactive. Hence the input admittance has a conductive component. The effective input conductance resulting from transit time when the load is nonreactive is

$$G_g = Kg_m f^2 T_t^2 \tag{4-13c}$$

in which K is a parameter that depends upon the electrode spacing and upon the operating voltages, g_m is the transconductance, f is the frequency, and T_t is the electron transit time.[1] Because of the grid circuit power loss that is associated with it, input conductance resulting from electron transit time is one of the factors that limit the performance of amplifiers and oscillators at ultrahigh frequencies.

4-4. Graphical Methods.—Many vacuum-tube problems can be solved most readily by graphical methods based upon the plate or transfer families of characteristics.[2] In certain cases there is some advantage in the use of the transfer characteristics, but the usefulness of the plate characteristics is much broader. This results partly from the fact that with nonreactive load the locus of corresponding values of plate current and plate voltage assumed during the cycle with a given load is a straight line, and partly because certain areas of the i_b-e_b diagram are proportional to power supplied or expended in various parts of the plate circuit, as will be shown in Sec. 4-21 and Chaps. 7 and 8.

4-5. Static Load Line.—When the excitation voltages are zero, the plate current and voltage assume their quiescent values, which are related by Eq. (3-30). This equation may be written in the form

$$I_{bo} = \frac{E_{bb} - E_{bo}}{R_b} \tag{4-14}$$

[1] FERRIS, W. R., *Proc. I.R.E.*, **24,** 82 (1936); NORTH, D. O., *Proc. I.R.E.*, **24,** 108 (1936).

[2] WARNER, J. C., and LOUGHREN, A. V., *Proc., I.R.E.*, **14,** 735 (1926); GREEN, E., *Wireless Eng.*, **3,** 402, 469 (1926); GRÜNWALD, E., *T.F.T.*, **22,** 306 (1933); COCKING, W. T., *Wireless Eng.*, **11,** 655 (1934).

If R_b is constant, Eq. (4-14) is that of a straight line intersecting the voltage axis at the point $e_b = E_{bb}$ and having a negative slope in amperes per volt equal to the reciprocal of the d-c load resistance. It is possible that the d-c load resistance may vary with I_{bo}. Equation (4-14) is then that of a curve passing through the point on the voltage axis where $e_b = E_{bb}$. Static load lines for fixed and variable load resistance are illustrated in Figs. 4-8*a* and 4-8*b*.

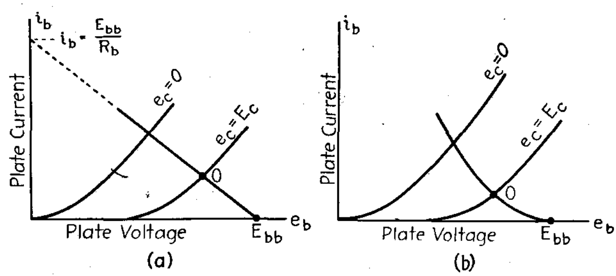


FIG. 4-8.—(*a*) Static load line for a fixed d-c load resistance; (*b*) d-c load resistance that varies with current.

The line represented by Eq. (4-14) is the locus of all static operating points that can be assumed with the given d-c load resistance and plate-supply voltage, and is called the *static load line*. From a comparison of Eqs. (3-26) and (3-30) it follows that all corresponding values of average plate current and average plate voltage assumed with excitation must also lie on the static load line.

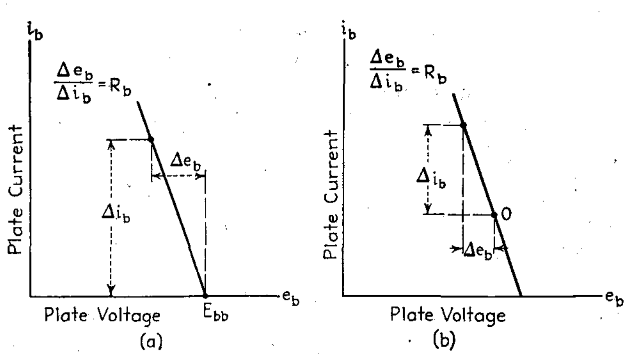


FIG. 4-9.—Method of constructing the static load line. (*a*) Plate supply voltage given; (*b*) static operating point given.

If the plate supply voltage is known, the static load line for constant d-c load resistance may be constructed by joining the point $e_b = E_{bb}$ on the voltage axis with the point $i_b = E_{bb}/R_b$ on the current axis. Usually, however, the slope of the static load line is so great that its intercept on the current axis is beyond the current range covered by the plate diagram.

Some other convenient point must then be used. Since the slope of the static load line is equal to $1/R_b$, it follows that $\Delta e_b/\Delta i_b = R_b$, where Δe_b and Δi_b are the difference in voltage and current of any two points on the line. Hence the line can be constructed by drawing it through the point E_{bb} on the voltage axis and through any other point whose coordinates differ from those of the point E_{bb} by amounts Δe_b and Δi_b such that $\Delta e_b/\Delta i_b = R_b$, as shown in Fig. 4-9*a*. Any convenient value of Δe_b, such as 10 or 100 volts, may be used in this relation to find Δi_b. After the static load line has been drawn, the static operating point for a particular grid bias may be determined from the intersection of the static load line with the static characteristic corresponding to the given bias.

If the static operating current and voltages, rather than the plate supply voltage, are given, the static load line is drawn through the static operating point and any other point whose coordinates differ from those of the operating point by amounts Δe_b and Δi_b such that $\Delta e_b/\Delta i_b = R_b$, as shown in Fig. 4-9*b*. Although the required plate supply voltage may then be found from the intersection of the static load line with the voltage axis, it may be determined with greater accuracy from Eq. (4-14).

4-6. Dynamic Path of Operation. Dynamic Load Line.—The locus of all corresponding values of instantaneous plate current and plate voltage assumed during the cycle with a given value of load impedance is called the *dynamic path of operation.* Under the assumption that the alternating plate current is sinusoidal, the instantaneous alternating plate current and plate voltage are

$$i_p = I_{pm} \sin \omega t \tag{4-15}$$

$$e_p = -I_{pm}|z_b| \sin (\omega t + \theta) \tag{4-16}$$

(see Eq. 3-32) where

$$\theta = \tan^{-1} \frac{x_b}{r_b} \tag{4-17}$$

Expanding $\sin (\omega t + \theta)$ and substituting Eq. (4-15) in Eq. (4-16) gives

$$-e_p = i_p|z_b| \cos \theta + |z_b| \sqrt{I_{pm}^2 - i_p^2} \sin \theta \tag{4-18}$$

Transposing, squaring, and substituting $r_b/|z_b|$ and $x_b/|z_b|$ for $\cos \theta$ and $\sin \theta$, simplify Eq. (4-18) to

$$e_p^2 + 2e_p i_p r_b + i_p^2|z_b^2| = I_{pm}^2 x_b^2 \tag{4-19}$$

This is the equation of an ellipse whose center is at the operating point. It follows that the dynamic path of operation is an ellipse whose center is the operating point, as shown in Fig. 4-10. When $i_p = I_{pm}$, $e_p = E_p'$ and Eq. (4-19) reduces to $I_{pm}/E_p' = -1/r_b$, indicating that the slope of the line joining the operating point with the points of tangency of the ellipse to the upper and lower horizontal tangents is the reciprocal of the a-c load resistance.

The presence of harmonics in the plate current and voltage causes the path of operation to depart from the elliptical form indicated by an analysis based only upon the fundamental component. Graphical methods of analysis are most useful when harmonic production is not negligible. Since the true path of operation cannot be determined until the harmonic content is known, however, and the path of operation is needed for the determination of harmonic content, graphical analysis is rather complicated when the load reactance is taken into consideration.[1] For this reason it is customary to make the assumption that the load is nonreactive. Although the results attained under this assumption are not rigorous, they are of considerable value in predicting the performance of vacuum-tube circuits. If $x_b = 0$, Eq. (4-19) reduces to

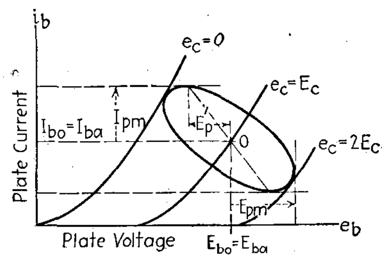


FIG. 4-10.—Path of operation for a load containing reactance. Distortion is assumed to be negligible.

$$e_p + i_p r_b = 0 \tag{4-20}$$

which is the equation of a straight line through the static operating point having a negative slope equal to the reciprocal of the a-c load resistance. This line, which represents the locus of all corresponding values of plate current and voltage assumed during the cycle with the given resistance load, is called the *dynamic load line*. If the a-c load resistance varies with plate current, as is usually true if the plate is coupled to the grid of a tube that passes grid current, then the dynamic load line is a curve whose slope at every point is the reciprocal of the a-c resistance corresponding to the current at that point, as in Fig. 8-24. The dynamic load line will be indicated by the letters *M-N*, as in Figs. 4-14 to 4-21.

4-7. Construction of the Dynamic Load Line. The Plate Diagram.—In most vacuum-tube circuits the a-c resistance of the load differs from the d-c resistance. Figure 4-11 shows a circuit in which the a-c resistance exceeds the d-c resistance. The d-c resistance R_b is merely the resistance R_1 of the primary of the transformer. The a-c resistance r_b, however, is equal to $R_1 + \dfrac{M^2\omega^2(R_2 + R)}{(R_2 + R)^2 + \omega^2 L_2^2}$, which, for an iron-core transformer, can usually be simplified to $R_1 + (R_2 + R)/n^2$, where n is

[1] GREEN, *loc. cit.;* ARDENNE, M. VON, *Proc. I.R.E.*, **16**, 193 (1928); BARCLAY, W. A., *Wireless Eng.*, **5**, 660 (1928); PREISMAN, A., *RCA Rev.*, **2**, 124, 240 (1937) (with bibliography); PREISMAN, A., "Graphical Constructions for Vacuum Tube Circuits," McGraw-Hill Book Company, Inc., New York, 1943.

the secondary-to-primary turn ratio (see Sec. 7-21). Figure 4-12 shows a circuit in which the a-c resistance is less than the d-c resistance. In applications of this circuit to vacuum tubes, the condenser C is usually so large that its reactance is negligible in comparison with R_2, and so the a-c resistance is that of R_1 and R_2 in parallel. The d-c resistance, on the other hand, is R_1.

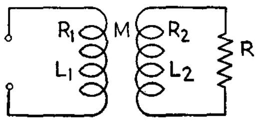


FIG. 4-11.—Circuit in which the a-c resistance between the terminals exceeds the d-c resistance.

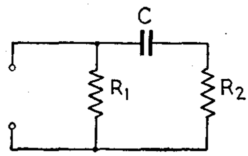


FIG. 4-12.—Circuit in which the a-c resistance between the terminals is less than the d-c resistance.

Since the slope of the a-c load line is equal to $1/r_b$, the coordinates of any two points on this line must differ by amounts Δe_b and Δi_b such that $\Delta e_b/\Delta i_b = r_b$. If the dynamic operating point is known, the line may be constructed by drawing it through the operating point and any other point whose coordinates differ from those of the operating points by amounts such that $\Delta e_b/\Delta i_b = r_b$.

The family of plate characteristics, together with the static and dynamic load line, is called the *plate-circuit diagram*, or simply the *plate diagram*. The use of plate diagrams in the prediction of vacuum-tube amplifier performance will be treated in later chapters. For reasons already given, the reactive component of load impedance will be neglected in the construction of plate-circuit diagrams in the derivations and analyses presented in the remainder of this book.

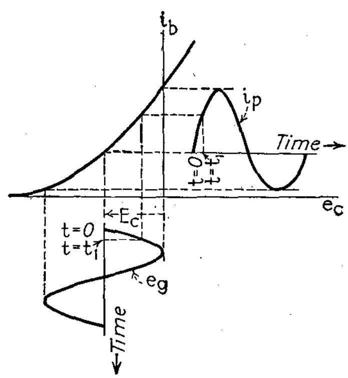


FIG. 4-13.—Use of the dynamic transfer characteristic in the graphical construction of the plate-current wave.

The dynamic transfer characteristic may be readily derived from the plate diagram by plotting corresponding values of i_b and e_c for the points at which the dynamic load line intersects the static characteristics. The dynamic transfer characteristic can then be used to determine the wave form of the plate current for a given wave of alternating grid voltage. From the values of instantaneous grid voltage at various instants in the cycle the corresponding values of current are determined by reference to the curve, or the current wave may be constructed by projection, as shown in Fig 4-13.

It is sometimes convenient to indicate the range over which the grid voltage varies by showing the wave of alternating grid voltage on an extension of the plate characteristic through the static operating point, as in Fig. 4-14. It should be noted that because the static characteristics for equal intervals of grid voltage are not usually equally spaced, the wave of plate current cannot be accurately constructed by projecting from the grid voltage wave to the load line and thence to the plate current wave.

4-8. Plate Diagram for a Load Having Equal A-c and D-c Resistance.—Figure 4-14 shows the plate diagram for a circuit in which the load is a pure resistance that does not vary with frequency. Since the a-c and d-c load resistances are equal, the static and dynamic load lines

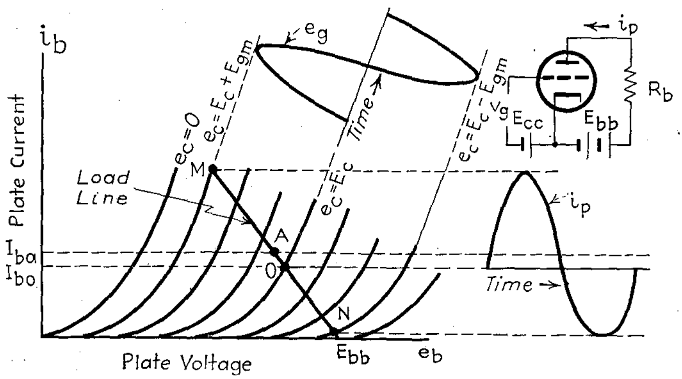


FIG. 4-14.—Triode plate diagram. Pure resistance load. $r_b = R_b$.

coincide. The path of operation and shape of the static characteristics are such that the current amplitude of the positive half cycle exceeds that of the negative. The average plate current therefore increases with excitation and the dynamic operating point A lies above the static operating point O. In general, the position of A relative to O depends upon the region of operation and the shape of the static characteristics. Shift of the operating point with excitation is predicted by the series expansion for alternating plate current, since the even-order terms of the series give rise to direct components of current, the magnitudes of which increase with the excitation amplitude.

4-9. Simplified Plate Diagram for a Load Having Unequal A-c and D-c Resistances.—Figure 4-15 shows an approximate plate diagram for a circuit in which the a-c resistance of the load exceeds the d-c resistance.[1] In the construction of this diagram it is assumed that the dynamic load line passes through the static operating point. The wave of alter-

[1] In Fig. 4-15 values of instantaneous plate voltage in excess of the supply voltage result from voltage induced in the inductance of the load coupling transformer.

nating plate current i_p, derived from the dynamic transfer characteristic in the manner explained in Sec. 4-7, shows, however, that the average plate current exceeds the static plate current and hence that the dynamic operating point lies above the static operating point O. It will be shown in Sec. 4-10 that the dynamic operating point must lie on the dynamic load line. Since the static load line is the locus of all corresponding values of average plate current and average plate voltage, the dynamic operating point must also lie on the static load line. The static and dynamic load line must, therefore, intersect at the dynamic operating point, rather than at the static operating point, and the diagram of Fig. 4-15 is not correct. In many analyses of vacuum circuits, particularly when triodes are used, it is nevertheless permissible to assume that the static and dynamic load lines intersect at the static operating

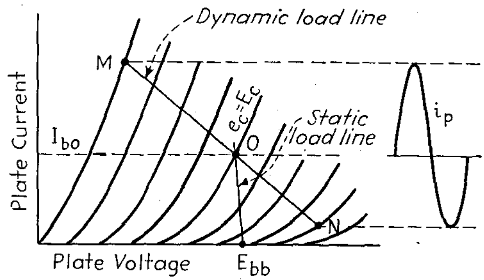


FIG. 4-15.—Simplified plate diagram for a circuit in which the a-c resistance of the load exceeds the d-c resistance.

point, *i.e.*, that the static and dynamic operating points coincide. Errors resulting from neglecting the shift of the operating point with excitation in the graphical analysis of vacuum-tube circuits are appreciable only when there is a relatively large change of average plate current with excitation, as there may be, for instance, in pentode circuits. Construction of the plate diagram is considerably more complicated when shift of the operating point is taken into consideration, as will become apparent in the following sections.

4-10. Exact Plate Diagram for Loads Having Unequal A-c and D-c Resistances.—The instantaneous value of plate current must become equal to the average value I_{ba} twice in each cycle. Similarly, the instantaneous value of plate voltage must become equal to the average value E_{ba} twice in each cycle. When the load is nonreactive, the load voltage and plate current are in phase and have identical wave form. The instantaneous values of plate current and plate voltage must, therefore, become equal to the average values of current and voltage at the same instants, and so the dynamic load line, which is the locus of all possible simultaneous values of i_b and e_b assumed during the cycle, must pass through the dynamic operating point A, at which the current and voltage have the average values I_{ba} and E_{ba}. As already pointed out, the dynamic operating point A must also lie on the static load line, and so the dynamic and static load lines intersect at the dynamic operating point. Unless the

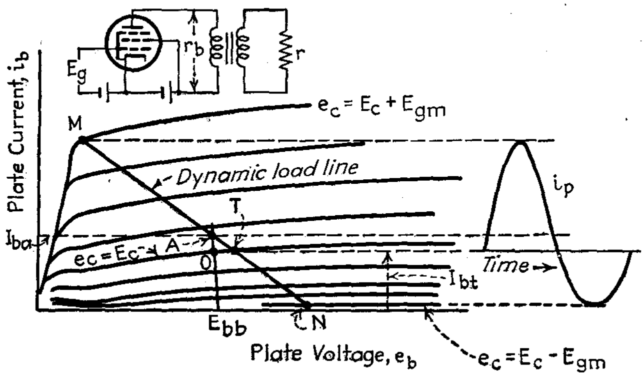


FIG. 4-16.—Exact plate diagram for a pentode with nonreactive load. $r_b > R_b$. Conditions of operation such that the average plate current increases with excitation.

wave of plate current is symmetrical, the dynamic operating point lies either above or below the static operating point.

Equations (3-58) and (3-59) show that, with nonreactive load, the instantaneous values of alternating grid voltage and alternating plate current are zero simultaneously. The time axis of the alternating-plate-current wave must therefore intersect the dynamic load line at a point

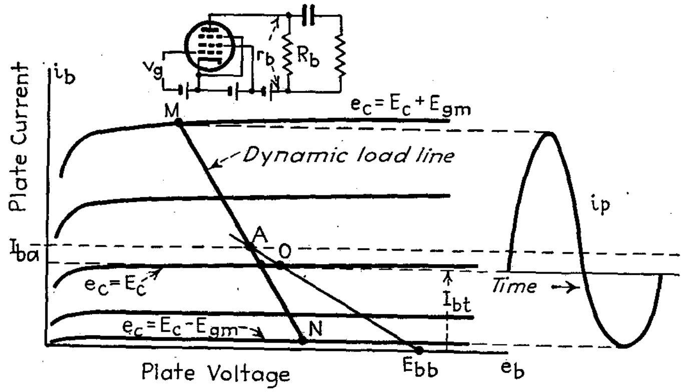


FIG. 4-17.—Plate diagram for a pentode with nonreactive load. $r_b < R_b$. Conditions of operation such that the average plate current increases with excitation.

on the diagram at which the instantaneous alternating grid voltage is zero, *i.e.*, at which $e_c = E_c$. In Figs. 4-16 to 4-18, this is the point T determined by the intersection of the dynamic load line with the static characteristic corresponding to $e_c = E_c$.

Typical plate diagrams are illustrated by Figs. 4-16 to 4-18. The diagram of Fig. 4-16 is for a circuit in which the a-c load resistance exceeds the d-c load resistance, such as that of Fig. 4-11 or one in which the load resistance is coupled to the plate by means of a choke and a condenser, as in Fig. 5-17. The diagram of Fig. 4-17 is for the circuit of

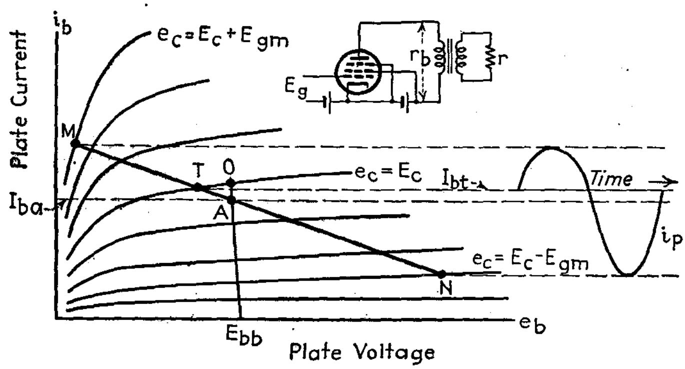


FIG. 4-18.—Plate diagram for a pentode with nonreactive load. $r_b > R_b$. Conditions of operation such that average plate current decreases with excitation.

Fig. 4-12, in which the a-c load resistance is less than the d-c load resistance. Figure 4-18 is an example of a pentode plate diagram in which the operating point is chosen so that the average plate current decreases when the tube is excited.

When the bias is obtained by the use of a cathode resistor,[1] the change in bias with average cathode current results in appreciable reduction in

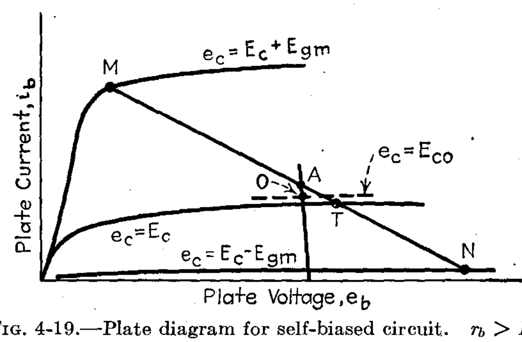


FIG. 4-19.—Plate diagram for self-biased circuit. $r_b > R_b$.

the shift of the operating point. A plate diagram for self-bias operation is illustrated in Fig. 4-19, in which the separation of the points is exaggerated for the sake of clearness. The point O lies at the intersection of the static load line with the static characteristic corresponding to the static bias E_{co}. The point T lies at the intersection of the dynamic load line

[1] See Sec. 5-8.

with the static characteristic corresponding to the dynamic bias E_c. With triodes, $E_{co} = I_{bo}R_{cc}$ and $E_c = I_{ba}R_{cc}$, where R_{cc} is the biasing resistance. With tetrodes and pentodes the action is more complicated than with triodes, since the bias is produced not only by the average plate current but also by the screen current, which varies with bias and may vary appreciably with excitation.

It should be noted that dependence of a-c load resistance upon frequency causes the slope of the dynamic load line to change with frequency. Furthermore, departure of the plate-current wave from sinusoidal form increases with amplitude of excitation voltage, so that the difference between static and dynamic average plate current also increases with excitation amplitude. The form of the plate diagram therefore changes with frequency and with amplitude of excitation.

In graphical analyses of vacuum-tube circuits the average plate current I_{ba} is determined from the dynamic load line, from the dynamic transfer characteristic, or from the wave of plate current derived from the dynamic load line. Since the value of I_{ba} must be known in order to locate the dynamic load line, the dynamic operating point and dynamic load line must be located by trial. If the dynamic operating point is correctly chosen, the average plate current, as determined from the dynamic load line or the wave of plate current derived from it, must be equal to the current I_{ba} at the assumed dynamic operating point. Location of the dynamic operating point is greatly facilitated by the use of graphical methods of harmonic analysis. For this reason it is desirable to study graphical methods of harmonic analysis before discussing practical methods of locating the dynamic load line and analyzing the plate diagram.

4-11. Graphical Analysis of Plate Current.—Graphical methods of analyzing the plate current of a vacuum tube are based upon the assumption that for sinusoidal excitation voltage the plate current contains a finite number of frequency components, or upon the equivalent assumption that the plate current may be expressed as a finite series. If as many instantaneous values of current can be determined graphically as there are components, or terms in the series, simultaneous equations can be set up from which the amplitudes may be computed. A number of interesting and useful methods have been developed, which differ from one another mainly as to choice of the points of the cycle at which the currents are evaluated and as to whether the instantaneous currents are measured with respect to alternating current zero or with respect to the instantaneous values that would obtain at those instants in the cycle if there were no distortion.[1] The accuracy and convenience of these methods depend upon the manner in which the selected currents are

[1] See bibliography at end of chapter.

chosen. The currents may be evaluated at equal time intervals of the excitation cycle, as exemplified by the method of Lucas; they may be evaluated at equal grid-voltage intervals, as in Espley's method; or they may be evaluated at such instants as to give the highest accuracy. The advantage of the second method lies in the fact that the grid voltages at which the current is evaluated may often be made to coincide with those of the static plate characteristics, so that the currents may be read directly from the intersections of the dynamic load line with these static characteristics. Chaffee has developed a more general method which includes all three of these special methods.[1]

In the treatment that follows, it will be assumed that the load is nonreactive. The path of operation is then a straight line, and evaluating the current at a given number of points on the path of operation is equivalent to evaluating at twice that number of instants in the fundamental cycle. It will sometimes be convenient to differentiate between methods of analysis according to the number of instantaneous values of current that are used to determine the amplitudes of the fundamental and harmonic frequency components. Thus a seven-point analysis is one in which the current is evaluated at seven points of the path of operation or of the dynamic transfer characteristic, or at 14 instants in the fundamental period.

4-12. Derivation of Formulas for Graphical Analysis.—In the following analysis the excitation voltage is assumed to be impressed upon the control grid in order to vary the plate current. The form of the analysis is similar when the excitation is applied to any electrode in order to vary the current of that or any other electrode. With suitable changes of symbols, therefore, the equations derived in this section may be applied to the analysis of the alternating currents flowing in any electrode when the excitation is sinusoidal and the load nonreactive. Since the instantaneous alternating voltage across a nonreactive load is proportional to the instantaneous alternating plate current, the equations may also be applied to the analysis of load voltage. The use of the equations may, in fact, be extended to the harmonic analysis of any quantity that varies periodically as the result of the sinusoidal variation of another quantity that is related to the first through a single-valued curve of known form.

Substitution of the sinusoidal voltage $E_{gm} \sin \omega t$ for the excitation voltage e in the series expansion for i_p [Eq. (3-40)], gives

$$i_p = a_1 E_{gm} \sin \omega t + a_2 E_{gm}^2 \sin^2 \omega t + a_3 E_{gm}^3 \sin^3 \omega t + a_4 E_{gm}^4 \sin^4 \omega t + \cdots \qquad (4\text{-}21)$$

which may be written in the form

[1] CHAFFEE, E. L., *Rev. Sci. Instruments*, **7**, 384 (1936).

$$i_p = H_0 + H_1 \sin \omega t - H_2 \cos 2\omega t - H_3 \sin 3\omega t + H_4 \cos 4\omega t + H_5 \sin 5\omega t - \cdots \quad (4\text{-}22)$$

in which H_n is the amplitude of the nth harmonic component of the alternating plate current and H_0 is the steady component of alternating plate current. As previously explained, the time axis of the alternating-plate-current wave passes through the point T, and i_p is measured with respect to the current I_{bt}, corresponding to the time axis. The total instantaneous plate current is

$$i_b = I_{bt} + i_p = I_{bt} + H_0 + H_1 \sin \omega t - H_2 \cos 2\omega t - H_3 \sin 3\omega t + H_4 \cos 4\omega t + H_5 \sin 5\omega t - \cdots \quad (4\text{-}23)$$

Formulas must be derived for the coefficients of Eq. (4-23) in terms of selected values of instantaneous plate current. The accuracy of these

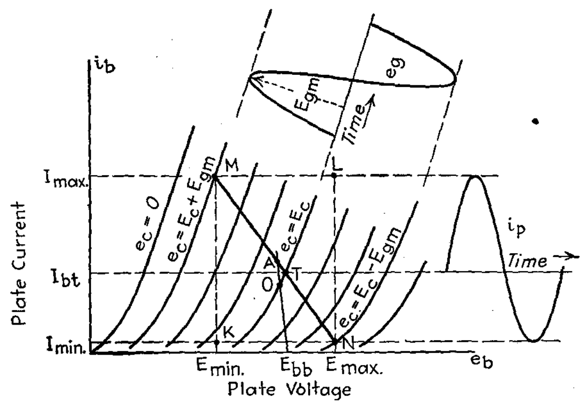


FIG. 4-20.—Diagram used in the derivation and application of Eqs. (4-29).

formulas for harmonic amplitudes increases with the number of points of the fundamental cycle at which the current is evaluated and also depends upon the location of these points in the cycle. The variation of the instantaneous total plate current with amplitude of a given harmonic is greatest at the instants at which the harmonic has its crest value. Highest accuracy is therefore obtained if the currents used to determine the amplitude of a given harmonic correspond to the instants at which the harmonic has its crest value.

In order to explain the method and thus justify the use of the formulas, Espley's formulas will be derived for the simple case in which the third and higher harmonics are assumed to be negligible. Under this assumption the alternating plate current may be expressed by two terms of the plate-current series. Plate excitation voltage is assumed to be zero. The total instantaneous plate current is

$$i_b = I_{bt} + a_1 e_g + a_2 e_g^2 \quad (4\text{-}24)$$

The following relations are apparent from Fig. 4-20:

$$i_b = I_{\max} \quad \text{when} \quad e_g = E_{gm} \tag{4-25}$$
$$i_b = I_{\min} \quad \text{when} \quad e_g = -E_{gm} \tag{4-26}$$

Substituting Eqs. (4-25) and (4-26) in Eq. (4-24) and solving the resulting simultaneous equations gives

$$a_1 = \frac{I_{\max} - I_{\min}}{2E_{gm}} \qquad a_2 = \frac{I_{\max} + I_{\min} - 2I_{bt}}{2E_{gm}^2} \tag{4-27}$$

Substituting Eq. (4-27) in the first two terms of Eq. (4-21) and expanding $\sin^2 \omega t$ gives

$$i_b = I_{bt} + \tfrac{1}{4}(I_{\max} + I_{\min} - 2I_{bt}) + \tfrac{1}{2}(I_{\max} - I_{\min}) \sin \omega t \\ - \tfrac{1}{4}(I_{\max} + I_{\min} - 2I_{bt}) \cos 2\omega t \tag{4-28}$$

The average plate current and the amplitudes of the fundamental and second-harmonic components of plate current are

$$\left.\begin{aligned} H_0 &= \tfrac{1}{4}(I_{\max} + I_{\min} - 2I_{bt}) \\ I_{ba} &= I_{bt} + H_0 = \tfrac{1}{4}(I_{\max} + I_{\min} + 2I_{bt}) \\ H_1 &= \tfrac{1}{2}(I_{\max} - I_{\min}) \\ H_2 &= \tfrac{1}{4}(I_{\max} + I_{\min} - 2I_{bt}) \end{aligned}\right\} \tag{4-29}$$

It is important to note that Eqs. (4-29) will give sufficiently accurate values of steady, fundamental, and second-harmonic components of plate current only when higher harmonics are negligible. If the higher harmonics cannot be neglected, it is necessary to use formulas based upon a greater number of terms of the series expansion for i_p. To derive equations that include harmonics up to the nth, n terms of the expansion are used, and the series is evaluated at $n + 1$ values of instantaneous grid voltage. Espley's method for four harmonics requires the determination of the instantaneous plate currents corresponding to $e_g = +E_{gm}$, $e_g = +\frac{1}{2}E_{gm}$, $e_g = 0$, $e_g = -\frac{1}{2}E_{gm}$, and $e_g = -E_{gm}$. These five values of current will be represented by the symbols I_1, $I_{\frac{1}{2}}$, I_{bt}, $I_{-\frac{1}{2}}$, and I_{-1}, respectively. The following expressions give the amplitudes of the components of plate current:

$$\left.\begin{aligned} I_{ba} &= \tfrac{1}{6}(I_1 + 2I_{\frac{1}{2}} + 2I_{-\frac{1}{2}} + I_{-1}) \\ H_1 &= \tfrac{1}{3}(I_1 + I_{\frac{1}{2}} - I_{-\frac{1}{2}} - I_{-1}) \\ H_2 &= \tfrac{1}{4}(I_1 - 2I_{bt} + I_{-1}) \\ H_3 &= \tfrac{1}{6}(I_1 - 2I_{\frac{1}{2}} + 2I_{-\frac{1}{2}} - I_{-1}) \\ H_4 &= \tfrac{1}{12}(I_1 - 4I_{\frac{1}{2}} + 6I_{bt} - 4I_{-\frac{1}{2}} + I_{-1}) \end{aligned}\right\} \tag{4-30}$$

I_1, $I_{\frac{1}{2}}$, $I_{-\frac{1}{2}}$, and I_{-1} are determined from the intersections of the dynamic load line with the static characteristics corresponding to $e_c = E_c + E_{gm}$, $e_c = E_c + \frac{1}{2}E_{gm}$, $e_c = E_c - \frac{1}{2}E_{gm}$, and $e_c = E_c - E_{gm}$, respectively, as shown

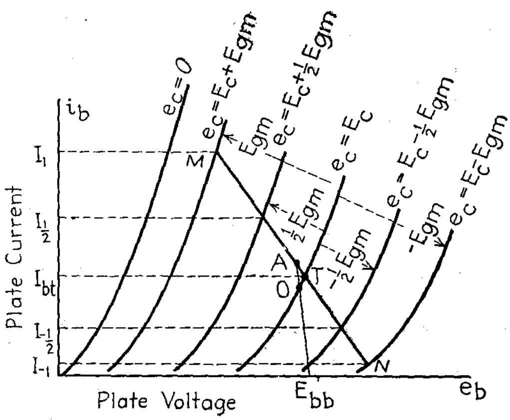


FIG. 4-21.—Use of the plate diagram in the application of Eqs. (4-30).

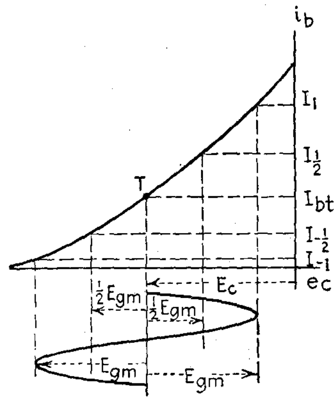


FIG. 4-22.—Use of the dynamic transfer characteristic in the application of Eqs. (4-30).

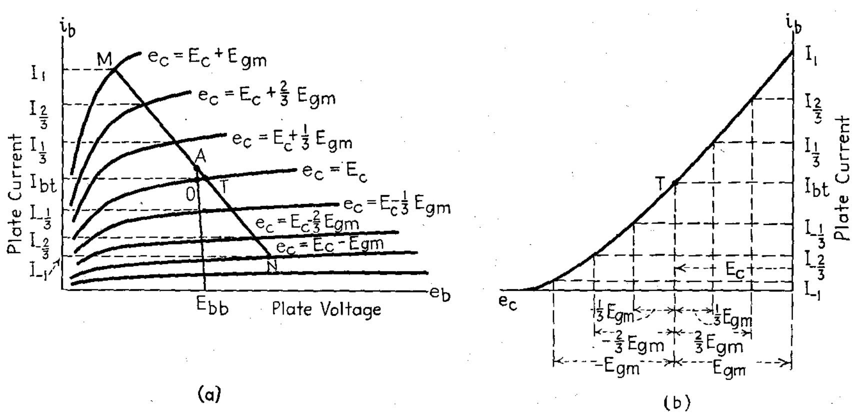


FIG. 4-23.—Use of (a) the plate diagram and (b) the dynamic transfer characteristic in the application of Table 4-I.

in Fig. 4-21. If the static characteristics corresponding to these values of grid voltage are not available, the currents may be read from the dynamic transfer characteristic, as shown in Fig. 4-22.

For six harmonics, Espley's method requires that the plate current be evaluated at the seven points of the load line or dynamic transfer

TABLE 4-I

	I_1	$I_{2/3}$	$I_{1/3}$	I_{bt}	$I_{-1/3}$	$I_{-2/3}$	I_{-1}
I_{ba}	0.131	0.295	−0.105	0.359	−0.105	0.295	0.131
H_1	0.261	0.394	−0.07	0	0.07	−0.394	−0.261
H_2	0.218	0.19	−0.475	0.133	−0.475	0.19	0.218
H_3	0.176	−0.141	−0.246	0	0.246	0.141	−0.176
H_4	0.12	−0.295	0.105	0.141	0.105	−0.295	0.12
H_5	0.063	−0.253	0.316	0	−0.316	0.253	−0.063
H_6	0.032	−0.19	0.475	−0.633	0.475	−0.19	0.032

characteristic corresponding to $e_g = 0$, $\pm \frac{1}{3}E_{gm}$, $\pm \frac{2}{3}E_{gm}$, and $\pm E_{gm}$, as indicated in Figs. 4-23*a* and 4-23*b*. The coefficients by which these currents must be multiplied are given in Table 4-I. To find the amplitude of the harmonic listed in any row of this table, the coefficients appearing in that row are multiplied by the graphically determined numerical values of the currents listed at the heads of the columns, and added.

Espley's equations are convenient to use if the instantaneous grid voltages at which the currents are found correspond to the given static characteristics; but the accuracy, particularly of the even harmonics, is not the highest that may be obtained, since the instantaneous currents do not in general correspond to the instants at which the harmonics have crest values. Maximum seven-point accuracy in the evaluation of each harmonic would require a different set of seven instantaneous currents for each harmonic and would therefore greatly increase the work of analysis. The following equations, in which the currents are evaluated at or near the instants at which the harmonics have crest values, have five- and seven-point

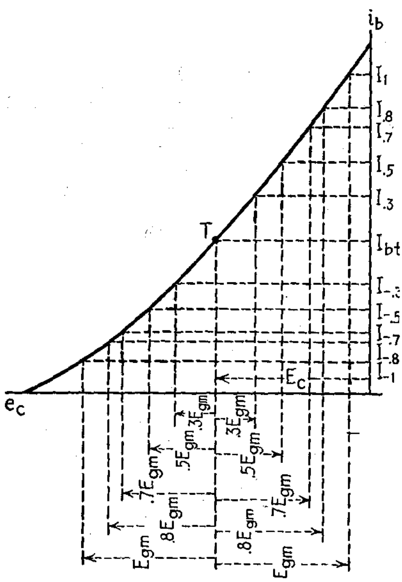


FIG. 4-24.—Use of the dynamic transfer characteristic in the application of Eqs. (4-31).

accuracy. The currents must be found corresponding to plus and minus 0.3, 0.5, 0.7, 0.8, and full grid swing, as shown in Fig. 4-24. If the currents are indicated by the symbols I_1, I_{-1}, $I_{.8}$, $I_{-.8}$, etc., the harmonic amplitudes are

$$\left.\begin{aligned}
I_{ba} &= \tfrac{1}{6}(I_1 + 2I_{.5} + 2I_{-.5} + I_{-1}) \\
H_2 &= \tfrac{1}{4}(I_1 - 2I_{bt} + I_{-1}) \\
H_3 &= \tfrac{1}{6}(I_1 - 2I_{.5} + 2I_{-.5} - I_{-1}) \\
H_4 &= \tfrac{1}{8}(I_1 - 2I_{.7} - 2I_{bt} - 2I_{-.7} - I_{-1}) \\
H_5 &= 0.095I_1 - 0.197I_{.8} + 0.207I_{.3} - 0.207I_{-.3} \\
&\qquad + 0.197I_{-.8} - 0.095I_{-1} \\
&\cong 0.1(I_1 - 2I_{.8} + 2I_{.3} - 2I_{-.3} + 2I_{-.8} - I_{-1}) \\
H_1 &= \tfrac{1}{2}(I_1 - I_{-1} - H_3 - H_5) \\
&\cong \tfrac{1}{30}(7I_1 + 6I_{.8} + 10I_{.5} - 6I_{.3} + 6I_{-.3} \\
&\qquad - 10I_{-.5} - 6I_{-.8} - 7I_{-1})
\end{aligned}\right\} \quad (4\text{-}31)$$

These formulas for I_{ba}, H_2, and H_3 are the same as those given under Eqs. (4-30). That for H_4 may be derived by a modification of Espley's method or by that of Chaffee, under the assumption that $\sin 45° = 0.700$. Those for H_5 and H_1 may be derived by Espley's or Chaffee's method for seven points.

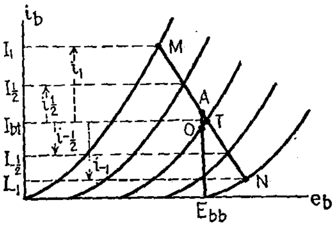


FIG. 4-25.—Diagram showing the relation between instantaneous currents measured relative to zero and relative to I_{bt}.

Sometimes it is convenient or necessary to measure the instantaneous alternating currents (with respect to the current corresponding to the time axis I_{bt}), rather than the instantaneous total currents, in making graphical analyses. The following relations are apparent from Fig. 4-25:

$$\left.\begin{aligned}
I_1 &= I_{bt} + i_1 & \qquad I_{-\frac{1}{2}} &= I_{bt} - i_{-\frac{1}{2}} \\
I_{\frac{1}{2}} &= I_{bt} + i_{\frac{1}{2}} & \qquad I_{-1} &= I_{bt} - i_{-1}
\end{aligned}\right\} \quad (4\text{-}32)$$

in which the lower-case symbols indicate currents measured relative to I_{bt}. By means of Eqs. (4-32), Eqs. (4-30) may be transformed into

$$\left.\begin{aligned}
I_{ba} &= \tfrac{1}{6}(i_1 + 2i_{\frac{1}{2}} + 6I_{bt} - 2i_{-\frac{1}{2}} - i_{-1}) \\
H_1 &= \tfrac{1}{3}(i_1 + i_{\frac{1}{2}} + i_{-\frac{1}{2}} + i_{-1}) \\
H_2 &= \tfrac{1}{4}(i_1 - i_{-1}) \\
H_3 &= \tfrac{1}{6}(i_1 - 2i_{\frac{1}{2}} - 2i_{-\frac{1}{2}} + i_{-1}) \\
H_4 &= \tfrac{1}{12}(i_1 - 4i_{\frac{1}{2}} + 4i_{-\frac{1}{2}} - i_{-1})
\end{aligned}\right\} \quad (4\text{-}33)$$

If the dynamic transfer characteristic is symmetrical about E_c, then

$i_1 = i_{-1}$ and $i_{1/2} = i_{-1/2}$, and Eqs. (4-33) reduce to

$$\left.\begin{array}{ll} I_{ba} = I_{bt} = I_{bo} & H_1 = \frac{2}{3}(i_1 + i_{1/2}) \\ H_2 = H_4 = 0 & H_3 = \frac{1}{3}(i_1 - 2i_{1/2}) \end{array}\right\} \qquad (4\text{-}34)$$

In a similar manner, for a symmetrical dynamic transfer characteristic, the seven-point equations corresponding to Table 4-I become

$$\left.\begin{array}{l} H_0 = H_2 = H_4 = H_6 = 0 \\ H_1 = 0.522i_1 + 0.787i_{2/3} - 0.141i_{1/3} \\ H_3 = 0.351i_1 - 0.281i_{2/3} - 0.492i_{1/3} \\ H_5 = 0.127i_1 - 0.506i_{2/3} + 0.633i_{1/3} \end{array}\right\} \qquad (4\text{-}35)$$

For a symmetrical characteristic, Eqs. (4-31) are

$$\left.\begin{array}{l} H_1 = \frac{1}{15}(7i_1 + 6i_{.8} + 10i_{.5} - 6i_{.3}) \\ H_3 = \frac{1}{3}(i_1 - 2i_{.5}) \\ H_5 = \frac{1}{5}(i_1 - 2i_{.8} + 2i_{.3}) \end{array}\right\} \qquad (4\text{-}36)$$

Equations (4-34), (4-35), and (4-36) show that with symmetrical dynamic transfer characteristics the dynamic and static operating points coincide and the even harmonics are zero. These equations will be found useful in the analysis of *push-pull* amplifiers, which will be defined in Chap. 5. The dynamic transfer characteristics of push-pull amplifiers are of the form shown in Fig. 4-26. I_{bt} is seen to be zero in this case.

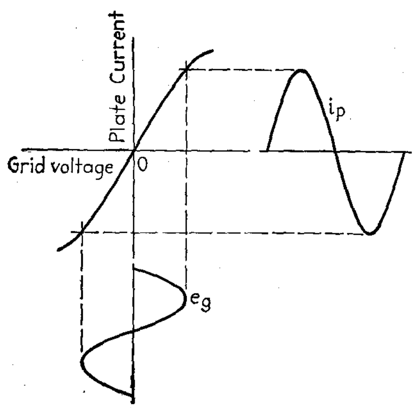


FIG. 4-26.—Symmetrical wave of plate current produced when the tranfer characteristic is symmetrical.

4-13. Choice of Equations.—The choice of the equations to be used in a particular analysis depends largely upon the accuracy required and upon the convenience of application. The number of static plate characteristics intersected by the load line may be such as to make it convenient to apply a seven-point analysis when five-point accuracy is sufficient. For very high accuracy, or in the determination of a large number of harmonics, some form of "schedule" method of analysis may have to be used.[1]

4-14. Field of Application of Equations for Harmonic Analysis.—As pointed out at the beginning of Sec. 4-11, the development of the equations of Secs. 4-11 and 4-12 was made for the special case of excita-

[1] MALTI, M. G., "Electric Circuit Analysis," p. 188, John Wiley & Sons, Inc., New York, 1930; KNIGHT, A. R., and FETT, G. H., "Introduction to Circuit Analysis," Chap. 14, Harper & Brothers, New York, 1943.

tion applied to the control-grid circuit in order to vary the plate current. By suitable changes of symbols the equations may be applied to the analysis of the current or voltage of any other electrode or of any quantity that varies periodically as the result of the sinusoidal variation of another quantity related to the first through a single-valued curve of known form.

4-15. Significance of Algebraic Signs of Numerical Values of Current Components.—It should be noted that the algebraic sign of the amplitude of a particular harmonic may turn out to be either positive or negative. The significance of the algebraic signs is indicated by Eq. (4-22). A reversal of sign merely indicates a 180-degree shift in the phase of a given harmonic. If the amplitude is positive, the harmonic adds to the fundamental at the instant when the fundamental has its positive crest value; if the amplitude is negative, the harmonic subtracts from the fundamental at that instant.

4-16. Nonsinusoidal Excitation Voltage.—The formulas for harmonic content developed in this chapter are all based upon the assumption that the alternating grid voltage is sinusoidal. They are not, therefore, of value in analyzing the alternating plate current when the grid voltage is not sinusoidal. It is possible, however, to construct a wave of plate current by means of the dynamic transfer characteristic and to analyze it by well-known methods of wave analysis.[1] This neglects the fact that the a-c load resistance and slope of the load line may be different for each input frequency component. An approximate analysis may be made under the assumption that the components of the wave of plate current are the same as would be obtained if the various components of the grid voltage were applied separately and the corresponding output components added. This would be valid if there were no intermodulation. It gives no indication of the intermodulation frequencies and neglects the fact that the dynamic operating point is different for each component than it would be for the resultant grid voltage. Ordinarily, adequate indication of the performance of the tube and circuit may be obtained by graphical methods based upon sinusoidal excitation, with full excitation voltage.

4-17. Mechanical Aids to Harmonic Measurement.—Scales that make possible the direct reading of percentage second and third harmonic have been described by D. C. Espley and L. I. Farren.[2] These scales, which are based upon a five-point analysis, are particularly useful when a large number of graphical determinations of harmonic content must be made.

[1] MOULLIN, E. B., *Wireless Eng.*, **8**, 118 (1931).

[2] ESPLEY, D. C., and FARREN, L. I., *Wireless Eng.*, **11**, 183 (1934); SARBACHER, R. L., *Electronics*, December, 1942, p. 52.

4-18. Percentage Harmonic and Distortion Factor.[1]—The percentage of a given harmonic is defined as the ratio of the amplitude of the harmonic to the amplitude of the fundamental, multiplied by 100. Thus, if harmonics above the second are negligible, Eqs. (4-29) indicate that the second harmonic percentage is

$$\text{Percentage } H_2 = \frac{I_{\max} + I_{\min} - 2I_{bt}}{2(I_{\max} - I_{\min})} \times 100 \tag{4-37}$$

The relative importance of different harmonics in producing audible distortion of speech and music depends to some extent upon the fundamental frequency. A given percentage of harmonic is in general more objectionable the higher the order of the harmonic. High-order harmonics result in unpleasant sharpness of tone or a hissing sound. Actually, however, the intermodulation frequencies, rather than the harmonics, are largely responsible for the disagreeable effects. The principal reason for this is that the intermodulation frequencies are in general inharmonically related to the impressed frequencies. A second reason is that some of the intermodulation frequencies corresponding to impressed frequencies near the upper end of the audio-frequency band may fall near the center of the band, whereas the harmonics fall outside of the band. The amplification of succeeding amplifiers and the sensitivity of the ear may be greater to these mid-band intermodulation frequencies than to the impressed frequencies.[2]

For a given value of maximum grid swing, the amplitude of a sum- or difference-frequency component associated with any term of the series does not exceed that of the corresponding harmonic produced when a single frequency is applied. The harmonic amplitudes at a given grid swing may, therefore, be used as a measure of distortion. A fairly satisfactory index of audible distortion is the *distortion factor*, which is defined by the equation

$$\delta = \frac{\sqrt{H_2^2 + H_3^2 + H_4^2 + \cdots}}{H_1} \tag{4-38}$$

Distortion factors up to about 0.05 are ordinarily not objectionable, and considerably larger values are frequently tolerated. Specific values will be discussed in Chaps. 6 to 8.

4-19. Practical Procedure. Location of Load Line. Harmonic Determination.—If the position of the static operating point is not known, the first step in the construction of a plate diagram is to locate

[1] MASSA, FRANK, *Proc. I.R.E.*, **21**, 682 (1933); Federal Radio Commission Rules and Regulations, Secs. 103, 139; NASON, C. H. W., *Radio Eng.*, **13**, 20 (1933); GRAFFUNDER, W., KLEEN, W., and WEHNERT, W., *Electronics* (abst.), November, 1935, p. 48; BARTLETT, A. C., *Wireless Eng.*, **12**, 70 (1935).

[2] MASSA, F., *Electronics*, September, 1938, p. 20.

the static operating point by the method explained in Sec. 4-5. The dynamic operating point must then be determined by trial and the dynamic load line drawn. There are several ways of making use of graphical methods of harmonic analysis in locating the dynamic operating point. The most obvious consists in guessing the location of the dynamic operating point and drawing a tentative dynamic load line through this point. By means of this tentative load line or the dynamic transfer characteristic obtained from it, I_{ba} is determined by means of Eqs. (4-29), (4-30), or (4-31), or Table 4-I. If the computed value of I_{ba} is equal to the value at the assumed dynamic operating point, the assumed point is correct. If the computed value exceeds the value at the assumed point, the point should be raised; if the computed value is less than the value at the tentative dynamic operating point, on the other hand, the point should be lowered. The procedure is repeated until the assumed and computed values of I_{ba} agree. It is sometimes of help to plot a curve of computed against assumed I_{ba}. The two values will be equal at the point at which the curve intersects a 45-degree line from the origin (equal scales being used).

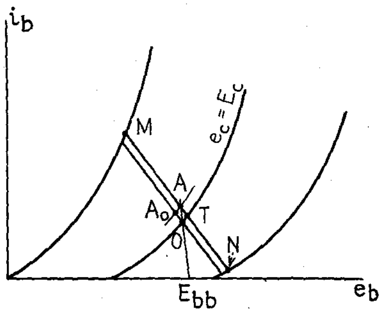


FIG. 4-27.—Diagram showing a method of locating the dynamic operating point.

A very simple construction, illustrated in Fig. 4-27, is helpful in locating the dynamic operating point.[1] A tentative load line is drawn through the static operating point or through a tentative dynamic operating point. With this load line, I_{ba} is computed by means of one of the graphical formulas. A point A_o is then located on the tentative dynamic load line, such that the current corresponding to this point is equal to the computed value of I_{ba}. A close approximation to the correct dynamic operating point is given by the intersection of the static load line and the static plate characteristic through A_o. The correctness of this new operating point should be checked by comparing its current with the value of I_{ba} computed graphically from the new dynamic load line. If necessary, the procedure should be repeated.

When self-bias is used,[2] the process of locating the dynamic operating point is complicated by the necessity of using the dynamic value of bias in locating point T. The construction shown in Fig. 4-27 is then of little value, but the use of a curve of assumed against computed values of I_{ba}, as explained above, is helpful. With self-biased tetrodes and

[1] BEDFORD, L. H., *Wireless Eng.*, **8**, 599 (1931).

[2] See Sec. 5-8.

pentodes, the graphical analysis involves so many factors, including dependence of screen current upon bias and excitation, that it is usually simpler to make laboratory measurements. The graphical location of the operating point may be simplified somewhat by assuming that the screen current is not affected by the change in bias, or that the variation of screen current with bias is proportional to the variation of plate current.

When the dynamic operating line has been located, the fundamental and harmonic components of alternating plate current or voltage may be determined by application of graphical formulas.

Voltage amplifiers and triode power amplifiers are ordinarily used in such a manner that the steady component of alternating plate current does not exceed 5 per cent of the amplitude of the fundamental component. When this is so, it is usually not necessary to go to the trouble of locating the dynamic operating point, sufficiently accurate results being obtained from a load line drawn through the static operating point as in Fig. 4-15. Kilgour has shown, however, that considerable error may result when this approximation is used in the analysis of power pentode amplifiers.[1]

4-20. Operating Point for Tubes with Filamentary Cathodes.—In obtaining the data for the static characteristics of a tube with a filamentary cathode, the filaments are always operated on direct voltage, and the grid and plate voltages are measured with respect to the negative end of the filament. When such a tube is operated with alternating filament voltage, the grid and plate circuits are connected to the filament through a center tap on the filament transformer or through a center-tapped resistor shunting the filament (see Sec. 5-13), and the operating grid bias is specified with respect to the center tap. Under d-c operation the grid is negative with respect to the center of the filament by half the filament voltage, and so the same plate current is obtained when the grid bias is lower than under a-c operation. Therefore the numerical value of bias used in locating the operating point graphically under a-c operation should be less than the voltage specified or applied between the grid and center of the filament by one-half the filament voltage. In families of plate characteristics furnished by tube manufacturers the static characteristic corresponding to specified operating voltages with respect to the center of the filament is often shown as a dotted curve. The filament voltage is ordinarily so much smaller than the plate voltage that no correction need be made in plate voltage in locating the operating point.

4-21. Graphical Determination of Power Output.[2]—Ordinarily only the fundamental power output is of importance. This is

[1] KILGOUR, C. E., *Proc. I.R.E.*, **19**, 42 (1931).

[2] For a discussion of the determination of power output in circuits using alternating plate supply voltage, see W. A. Schwaizmann, *Electronics*, August, 1943, p. 94.

$$P_o = (0.707H_1)^2 r_b = \tfrac{1}{2}H_1^2 r_b \tag{4-39}$$

where H_1 is the graphically determined value of fundamental plate-current amplitude. If harmonics higher than the second can be neglected, the fundamental amplitude given by Eqs. (4-29) is a close approximation. The corresponding equation for fundamental power output is

$$P_o = \tfrac{1}{8}(I_{\max} - I_{\min})^2 r_b \tag{4-40}$$

But since the slope of the dynamic load line is $1/r_b$ it can be seen from Fig. 4-20 that $(I_{\max} - I_{\min})r_b = E_{\max} - E_{\min}$.

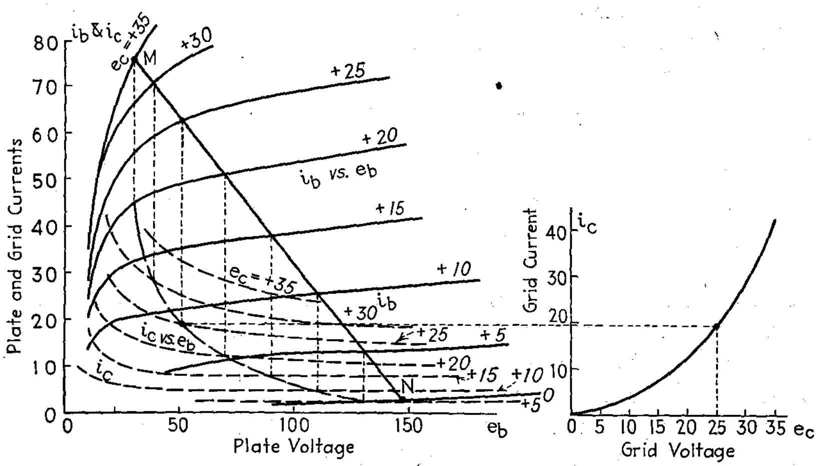


FIG. 4-28.—Use of the plate diagram and the static i_c-e_b characteristics in the graphical construction of the dynamic grid characteristic.

Therefore Eq. (4-40) may be transformed into

$$P_o = \tfrac{1}{8}(I_{\max} - I_{\min})(E_{\max} - E_{\min}) \tag{4-41}$$

This is equal to one-eighth of the area of the rectangle *KMLN* in Fig. 4-20.

4-22. Dynamic Grid Characteristics.—The grid current corresponding to a given value of grid voltage depends upon the plate voltage. When load is used in the plate circuit, the plate voltage varies with grid voltage in a manner that is determined by the load impedance. Therefore the form of the curve of grid current *vs.* grid voltage varies with the load impedance. Figure 4-28 shows the method of deriving a dynamic grid characteristic for a given plate load from the curves of i_b and i_c *vs.* e_b. For a given value of grid voltage the intersection of the load line with the plate characteristic for that grid voltage determines the corresponding value of plate voltage. The grid current at this plate voltage is then read from the static i_c-e_b characteristic for the given grid voltage. The dynamic grid-current curve is drawn by plotting values of i_c, determined in this manner, against e_c. Thus in Fig. 4-28 the plate voltage corresponding to a grid voltage of 25 for the given plate load resistance is 50 volts. At $e_b = 50$ volts, the grid-current curve for $e_c = 25$ volts indicates a grid current of 19.5 ma. These values determine one point on the dynamic grid-current characteristic.

When the grid is allowed to assume positive voltages and the grid circuit contains appreciable impedance, the instantaneous voltage of the grid will be less than the applied excitation voltage by the impedance drop. This will cause the wave form and phase angle of the grid voltage to differ from those of the grid excitation voltage. The effects of the resulting grid-circuit distortion and methods of determining it from the characteristic curves will be discussed in Chap. 8.

4-23. Starting and Stopping Transients.—In a circuit of the type shown in Fig. 4-11, the average plate current cannot increase instantaneously from I_{bo} to I_{ba}. Because of inductance in the primary of the transformer there will always be transient conditions just after the voltage e_g is applied or removed from the grid circuit. When the voltage is first applied, any tendency for the steady plate current to increase results in the production of an induced voltage $-L(dI_b/dt)$, in the primary of the transformer. This voltage opposes the plate supply voltage, producing an effect equivalent to a reduction of the supply voltage to a new value $E_{bb} - L(dI_b/dt)$. The static load line is shifted to the left sufficiently so that the current I_{ba}' of the transient dynamic operating point A' is equal to the initial steady current I_{bo}. This is illustrated in Fig. 4-29. The average current immediately starts to increase exponentially, and points A' and T' approach their steady positions A and T. When e_g is removed from the grid circuit, an induced voltage is added to E_{bb} of sufficient magnitude to give a new static operating point O'', for which I_{bo}'' is equal to I_{ba}. The plate current immediately begins to decrease, and O'' moves down the static characteristic for $e_c = E_c$, eventually reaching O. The capacitance in the circuit of Fig. 4-12 also results in transients. Any increase in the steady component of plate current increases the steady IR drop through R_1 and thus unbalances the voltages in the branch containing C, which in turn causes charging current to flow into C until the increase in condenser voltage is equal to the product of R_1 by the increase in steady plate current. This

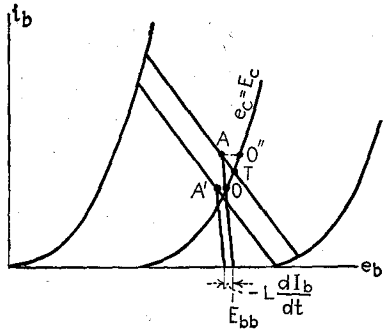


FIG. 4-29.—Diagram showing the effect of starting and stopping transients upon the shift of the operating point.

charging current in effect diverts a portion of the plate current from R_1 so that, while C is charging, the drop through R_1 is less than its final value. The result is the same as though R_1 were shunted during the transient time by a resistance that increased exponentially from an

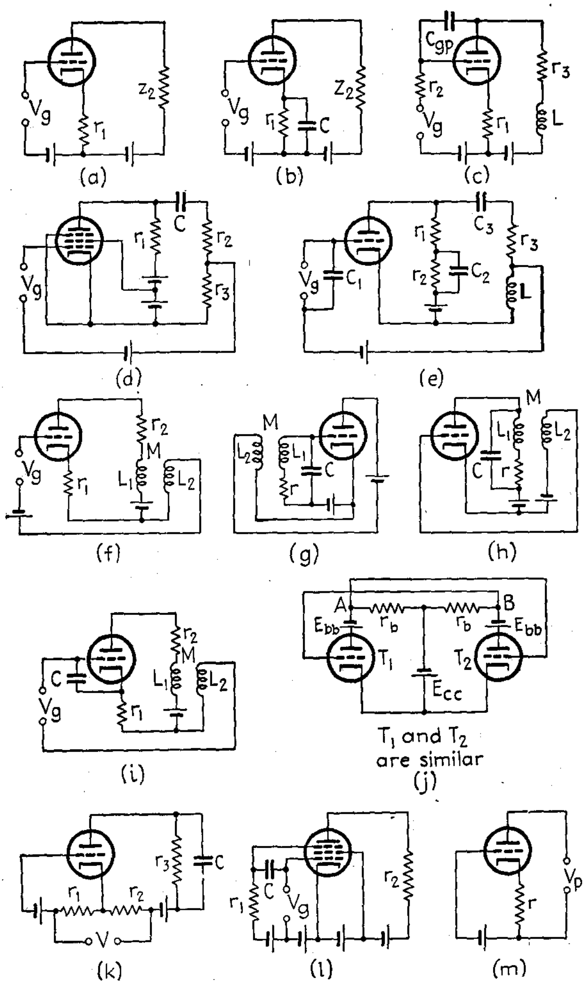


FIG. 4-30.—Diagrams for Prob. 4-1.

initial value R_2 to an infinite value. On the plate-circuit diagram this is shown by a change in the slope of the static load line from an initial slope of $1/R_1 + 1/R_2$ to its final slope $1/R_1$. Not only is a transient associated with the change in steady plate current when the load contains reactance, but there is also in general a transient for each frequency contained in the plate current.[1] These transients may under certain conditions prove to be very undesirable in oscillographic work.

Problems

4-1. *a.* Construct the equivalent plate circuits for the circuits of Fig. 4-30, indicating the assumed polarities of all impressed and equivalent voltages.

b. Write the network equations for the equivalent circuits.

c. Express E_g as a summation of voltages between the cathode and the grid.

d. Form the parallel equivalent plate circuits (see Appendix, Sec. A-1).

4-2. *a.* Construct the equivalent second-grid circuit for the circuit of Fig. 4-31.

b. Write the expression for E_{g3} in terms of V_{g2}, r_c, and C_c.

4-3. *a.* Construct the equivalent plate circuit for the circuit of Fig. 4-32.

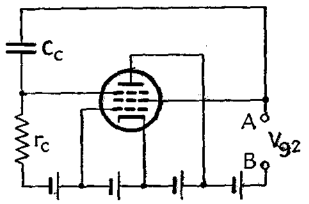


FIG. 4-31.—Diagram for Probs. 4-2 and 4-4.

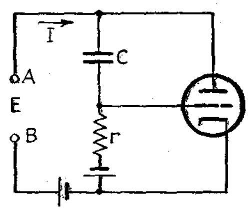


FIG. 4-32.—Diagram for Prob. 4-3.

b. By means of the equivalent circuit derive an expression for I in terms of E and thus show that the circuit acts like an admittance of value

$$y_e = \frac{r(1 + \mu) + r_p - j/\omega C}{r_p(r - j/\omega C)} \tag{4-42}$$

c. Show that when μr is large in comparison with r_p the circuit acts like a resistance r_e in parallel with a reactance x_e whose values are

$$r_e = \frac{r_p(r^2C^2\omega^2 + 1)}{r^2C^2\omega^2(1 + \mu) + 1} \tag{4-43}$$

$$x_e = -\frac{rC\omega + 1/rC\omega}{g_m} \tag{4-44}$$

d. Show that when $rC\omega$ is large, Eq. (4-43) reduces to

$$r_e = \frac{r_p}{1 + \mu} \tag{4-45}$$

4-4. *a.* Making use of the similarity between the equivalent circuits of Probs. 4-2 and 4-3, write equations equivalent to Eqs. (4-43) and (4-44) for the circuit of Fig. 4-31.

b. The current to the second grid in Fig. 4-31 increases with a positive increment of second grid voltage, which is positive, but decreases with a positive increment of third grid voltage, which is negative. Making use of this fact, show that the effective parallel resistance between A and B may be negative, and that the effective parallel reactance may be inductive.

c. Show that, when r is large in comparison with r_{g2}, the minimum value of negative resistance is equal to $r_{g2}/(1 + \mu_{23})$ and that the negative resistance approaches this value when the resistance r is large in comparison with the reactance of C.

[1] JACKSON, W., *Phil. Mag.*, **13**, 143, 735 (1932).

4-5. *a.* By means of the equivalent plate circuit, show that the effective resistance between points A and B of Fig. 4-30j is equal to[1]

$$r_e = \frac{2r_p r_b}{r_p + r_l(1 - \mu)} \tag{4-46}$$

b. Determine the condition that makes r_e negative.

4-6. *a.* Under the assumption that z_b is a pure resistance and that the reactances of C_{gp} and C_{pk} are very high in comparison with the load resistance r_b, construct the vector diagram for the circuit of Fig. 4-7 at a frequency sufficiently low so that electron transit time need not be taken into consideration. Show vectors for E_g, I_p, E_{zb}, the resultant voltage $E_g - E_{zb}$ impressed upon C_{gp}, and the current I_2 into C_{gp}.[2] Note that the condenser current is purely capacitive relative to E_g.

b. Repeat (*a*) for a load having zero resistance and a capacitive reactance. Note that the condenser current has a component that is conductive and a component that is capacitive relative to E_g.

c. Repeat (*a*) for a load having zero resistance and an inductive reactance. Note that the condenser current has a negative conductive component and a capacitive component relative to E_g.

d. Repeat (*a*) for a nonreactive load at a frequency so high that electron transit time causes the plate current to lag the grid voltage. Note that the condenser current has a conductive component.

4-7. *a.* Construct the approximate plate diagram for a type 45 tube when $E_{bb} = 280$ volts, $E_c = -50$ volts, $R_b = 500$ ohms, and $r_b = 3333$ ohms.

b. Derive a dynamic transfer characteristic from the plate diagram.

4-8. If the static operating voltages of a type 45 tube are $E_{bo} = 240$ volts, $E_c = -50$ volts; the plate supply voltage is 250 volts; and the a-c load resistance is 5000 ohms,

a. Find μ, r_p, and g_m at the static operating point.

b. Find the d-c resistance of the load.

c. Construct the approximate plate diagram.

d. Find the amplitudes of the fundamental and second-harmonic components of plate current at grid swings of 50 volts and 40 volts.

e. Find the fundamental power output at grid swings of 50 volts and 40 volts.

4-9. *a.* Locate the dynamic operating point and dynamic load line for a type 2A5 tube connected as a pentode and operated with a load that has a d-c resistance of 300 ohms and an a-c resistance of 5000 ohms. $E_{bo} = 250$ volts, $E_c = -15.0$ volts, and $E_{gm} = 15.0$ volts. (Note that the third, fourth, and fifth harmonics cannot be neglected.) (See p. 62 for 2A5 characteristics or, preferably, refer to a tube manual.)

b. Find H_1, H_2, H_3, H_4, and H_5.

c. Find the required B-supply voltage.

d. Find the fundamental power output.

e. Find the harmonic amplitudes and the power output under the assumption that the operating point does not shift when the tube is excited.

4-10. Repeat Prob. 4-9 for self-biased operation, assuming that the screen current remains constant at 6.5 ma. $E_{co} = -15.0$ volts.

[1] Stewart, J. A., "The Roberts Neutralizer Circuit" (Bachelor's thesis), Purdue University, June, 1935.

[2] Since an increase of plate current causes the plate end of the load impedance to become more negative relative to the cathode end, and E_{zb} is assumed to be measured relative to the cathode end, $E_{zb} = -I_p r_b$. Hence the direction of E_{zb} is opposite to that of E_g, and the direction of $E_g - E_{zb}$ is the same as that of E_g.

4-11. A type 53 tube is used with a pure resistance load of 1000 ohms and a plate supply voltage of 120 volts. Construct the corresponding dynamic grid characteristic.

Supplementary Bibliography

Theory of Harmonic Production in Tubes

Caporale, P.: *Proc. I.R.E.*, **18**, 1593 (1930).
Boner, M. O.: *Phys. Rev.*, **39**, 863 (1932).
Bennett, W. R.: *Bell System Tech. J.*, **12**, 228 (1933).
Espley, D. C.: *Proc. I.R.E.*, **22**, 781 (1934).
Bartlett, A. C.: *Wireless Eng.*, **12**, 70 (1935).

Harmonic Analysis

Fischer-Hinnen, J.: *Elektrotech. Z.*, **22**, 396 (1901).
Lincoln, P. M.: *Elec. J.*, **5**, 286 (1908).
Turner, P. K.: *Wireless Eng.*, **7**, 371 (1930).
Scroggie, M. G.: *Wireless Eng.*, **8**, 372 (1931).
Roder, H.: *Proc. I.R.E.*, **19**, 1481 (1931).
Greenwood, W.: *Wireless Eng.*, **8**, 428 (1931).
Lucas, G. S. C.: *Wireless Eng.*, **8**, 595 (1931); **9**, 75 (1932).
Ford, J. R.: *Proc. I.R.E.*, **20**, 359 (1932).
Gehrts, A.: *E.N.T.*, **10**, 436 (1933); see also abstr., *Wireless Eng.*, **11**, 93 (1934).
Van der Ven, A. J. H.: *Physik. Ber.*, **15**, 48 (1934).
Espley, D. C., and Oliver, D. A.: *Radio Eng.*, February, 1934, p. 11.
Terebesi, P.: *Arch. Elektrotech.*, **28**, 195 (1934).
Espley, D. C., and Farren, L. I.: *Wireless Eng.*, **11**, 183 (1934).
Kleen, W.: *E.N.T.*, **11**, 293 (1934).
Mouromtseff, I. E., and Kozanowski, H. N.: *Proc. I.R.E.*, **22**, 684 (1934).
Ferris, W. R.: *Proc. I.R.E.*, **23**, 510 (1935).
Hutcheson, J. A.: *Electronics*, January, 1936, p. 16.
Chaffee, E. L.: *Rev. Sci. Instruments*, **7**, 384 (1936).
Macfadyen, K. A.: *Wireless Eng.*, **15**, 310 (1938).
Scott, H. J., and Black, L. J.: *Proc. I.R.E.*, **26**, 449 (1938).
Fairweather, A., and Williams, F. C.: *Wireless Eng.*, **16**, 57 (1939).
Wheeler, H. A.: *Proc. I.R.E.*, **27**, 359, 384 (1939).
Thomson, W. T.: *Electrical Eng.*, **58**, 488 (1939).
Jonker, J. L. H.: *Wireless Eng.*, **16**, 274, 344 (1939).
van der Ven, A. J. H.: *Wireless Eng.*, **16**, 383, 444 (1939).
Bloch, A.: *Wireless Eng.*, **16**, 592 (1939).
Nims, A. A.: *Electronics*, May, 1939, p. 23.
Frommer, J., and Rédl, A.: *Wireless Eng.*, **17**, 4 (1940).

CHAPTER 5

AMPLIFIER DEFINITIONS, CLASSIFICATIONS, AND CIRCUITS

The many applications of thermionic electron tubes may be divided into amplification, modulation and detection, generation of alternating voltage, power rectification, current and power control, and measurement. The subject of amplification will be treated in this and the following three chapters. The purpose of the present chapter is to define terms that must be used in the discussion of amplifiers, to classify different types of amplifiers, and to show basic circuits of amplifiers. Chapters 6 to 8 will deal with the characteristics and the design of amplifiers.

5-1. Signal.—The term *signal* is applied to any alternating voltage or frequency impressed upon the input of an amplifier or other four-terminal network. It is also applied to the resulting fundamental components of output voltage or current, as distinguished from harmonic or intermodulation components.

5-2. Amplifiers.—An *amplifier* may be defined as a device for increasing the amplitude of electric voltage, current, or power, through the control, by the input, of power supplied to the output circuit by a local source. A vacuum-tube amplifier is one that employs vacuum tubes to effect the control of power from the local source.

Amplifiers are but one type of the general four-terminal network. Other types include the transformer and the filter. Wherever possible, the definitions and discussions given in this chapter are worded so as to apply also to other four-terminal networks.

5-3. Amplifier Distortion.—If an amplifier, or other four-terminal network, is distortionless, the application of a periodic wave of any form to the input terminals will result in the production of an output wave that is a replica of the input wave. In general, such a periodic wave will consist of a fundamental and one or more harmonics. In order that the output wave form shall be identical with that of the input, three conditions must be satisfied: (1) The output must contain only the frequencies contained in the input. (2) The output must contain all frequencies contained in the input, and the relative amplitudes of the various components must be the same as in the input. (3) If any component of the output is shifted in phase relative to the corresponding component of the input, all components must be shifted by the same number of electrical degrees of the fundamental cycle of that component. Examination of a fundamental wave and its nth harmonic shows that (3) is equivalent to saying that the phase shift of the nth harmonic, measured in electrical degrees of its own cycle, must be either n times the phase shift of the fundamental or an integral multiple of 180 degrees. This can be true only if the phase shift for sinusoidal input is either proportional to the frequency, or zero or 180 degrees at all frequencies.[1] Failure to satisfy these three conditions results in three corresponding types of distortion: nonlinear distortion (amplitude distortion), frequency distortion (frequency discrimination), and phase distortion, any one of which alters the wave form of the output relative to that of the input.

Nonlinear distortion (*amplitude distortion*) is the generation in an amplifier or other four-terminal network of frequencies not present in the impressed signal. It is usually associated with a nonlinear relation between output and input amplitudes. In vacuum-tube amplifiers it is the result of curvature of the dynamic tube characteristics. The generation of harmonics and intermodulation frequencies was discussed in Chap. 4. A nonlinear relation between the output and input amplitudes when the dynamic transfer characteristic is curved is predicted by the fact that the third and higher odd-order terms of the series expansion for plate current [Eq. (3-56)] give rise to fundamental components of plate current, the amplitudes of which vary as the cube or higher odd power of the excitation voltage. For this reason, unless the coefficients of all odd-order terms of the series are negligibly small, the voltage, current, and power output are not proportional to the exciting voltage. Nonlinear distortion is objectionable in the amplification of speech and music mainly because intermodulation frequencies are in general inharmonically related to the impressed frequencies and, therefore, produce unpleasant discords.[2] Although it is the inharmonic intermodulation frequencies, rather than the harmonics, that are objectionable, nonlinear distortion is most readily measured and specified in terms of the distortion factor, defined in Sec. 4-18. Nonlinear distortion can be minimized by proper choice of tubes, load impedances, and operating voltages, and by avoiding too high excitation voltage.

Excessive nonlinear distortion may occur when the exciting voltage and grid bias are such that the normal range of operation on the dynamic transfer characteristic is exceeded. Figure 5-1 shows how improper choice of grid bias or the use of excessive excitation may result in the flattening of one or both peaks of the wave of alternating plate current. A similar flattening of the positive peak may result from the flow of grid

[1] FRY, T. C., *Physik. Z.*, **23**, 273 (1922).

[2] BARTLETT, A. C., *Wireless Eng.*, **12**, 70 (1935); BARROW, W. L., *Phys. Rev.*, **39**, 863 (1932); ESPLEY, D. C., *Proc. I.R.E.*, **22**, 78 (1934).

current. In the circuit of Fig. 3-18, the flow of grid current through the grid-circuit impedance z_c causes the instantaneous alternating grid voltage e_g to be less than the instantaneous grid excitation voltage v_g during the time in which the grid is positive. In the circuit of Fig. 5-2, the excitation voltage is produced by the flow of alternating current i_z through the grid-circuit impedance z_c. During the part of the cycle in which the grid is positive, a portion i_g of the current i flows into the grid of the tube and so reduces i_z below the value that it would have if grid current did not flow. The alternating grid voltage is therefore also reduced below the value it would have if the grid did not conduct. Flattening of the peaks of alternating plate current as the result of excessive grid excitation voltage is called *overloading*. Flattening of the plate current peaks is an indication of the generation of harmonics and intermodulation frequencies of large amplitude.

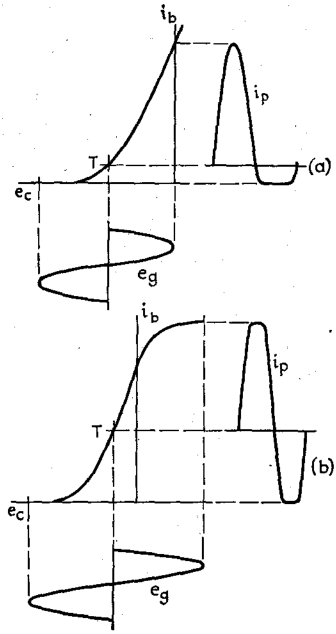


FIG. 5-1.—Distortion of waves of plate current as the result of overloading.

Frequency distortion in an amplifier or other four-terminal network is the variation of amplification or sensitivity with frequency of the impressed signal. In a vacuum-tube amplifier it results from dependence of circuit and interelectrode impedances upon frequency.[1] It can be minimized by proper design of input, output, and interstage coupling circuits, being least in amplifiers in which the circuits do not contain reactance. The difficulty of preventing frequency distortion increases with the width of the frequency band for which the amplifier is designed and with amplification per stage. Although frequency distortion may not produce disagreeable effects in the amplification of music, it impairs fidelity of tone and may prevent the reproduction of the sounds of some instruments. By eliminating the high frequencies essential to the reproduction of consonants, it may make reproduced speech difficult to understand.

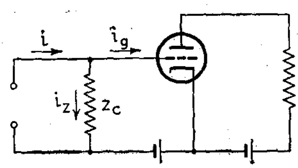


FIG. 5-2.—Diversion of current from the grid-circuit impedance as the result of flow of grid current when the grid is positive.

[1] At ultrahigh-frequency electron transit time also causes frequency and phase distortion.

Phase distortion is the shifting of the phase of the output voltage or current of an amplifier or other four-terminal network relative to the input voltage or current by an amount that is not proportional to frequency. Like frequency distortion, it results from the reactance of electrodes and circuit elements and can be made negligible by designing the coupling and other circuit elements so that the reactances have negligible effect throughout the desired frequency range.[1] This causes the phase shift to approximate zero or 180 degrees throughout the frequency range. Because phase distortion usually cannot be detected by ear it is ordinarily the least objectionable type of distortion in the amplification of sound. It is objectionable, however, in the amplification of television signals and in the use of amplifiers in the oscillographic study of voltage and current wave form.

Since the variations of amplification and of phase of the output voltage of an amplifier with frequency are both caused by tube and circuit reactances, the phase of the output voltage is ordinarily constant throughout any range of frequency in which the amplification is constant.

It is of interest to note that the human ear has both noticeable nonlinear and frequency distortion.

5-4. Amplifier Classification.—Vacuum-tube amplifiers are commonly classified in four ways: (1) according to use, (2) according to circuits, (3) according to frequency range, and (4) according to the portion of the cycle during which plate current flows.

Voltage, Current, and Power Amplifiers.—When classified as to type of service, amplifiers are termed *voltage amplifiers*, *current amplifiers*, or *power amplifiers*, depending upon whether they are designed to furnish voltage, current, or power output. The effectiveness with which a voltage amplifier accomplishes its function is indicated by its *voltage amplification*. *Voltage amplification* is the ratio of the signal voltage available at the output terminals of an amplifier, transformer, or other four-terminal network, to the signal voltage impressed at the input terminals. It will be represented by the symbol A. Methods of measuring voltage amplification are discussed in Sec. 15-41.

Current amplification is the ratio of the signal current produced in the output circuit of an amplifier, transformer, or other four-terminal network to the signal current supplied to its input circuit. The effectiveness with which a current amplifier accomplishes its function may sometimes be specified by its current amplification, but usually the input impedance is so high that this term has no useful significance unless the input is shunted by a specified impedance. The performance can be indicated better by the *current sensitivity*, which is defined as the ratio

[1] See footnote, p. 126.

of the signal current produced in the output circuit of an amplifier or other four-terminal network to the signal voltage impressed at the input terminals. Current sensitivity is measured in mhos.

Power amplification is the ratio of the power delivered to the output circuit of an amplifier, or other four-terminal network containing a source of local power, to the power supplied to its input circuit. Because of the high input impedance of many types of power amplifiers, this term may have no useful significance unless the input is shunted by a specified impedance. It is usually better, therefore, to specify the performance of a power amplifier by its *power sensitivity*, which is defined as the ratio of the signal-frequency power delivered by the output circuit of an amplifier or other four-terminal network containing a local source of power, to the square of the effective value of the signal voltage impressed at the input terminals.[1] Power sensitivity is measured in mhos.

A *frequency-response characteristic* is a graph that relates the amplification or sensitivity of an amplifier or other four-terminal network with the frequency of the impressed signal. The term is usually applied to a graph of voltage amplification as a function of frequency of the impressed signal.

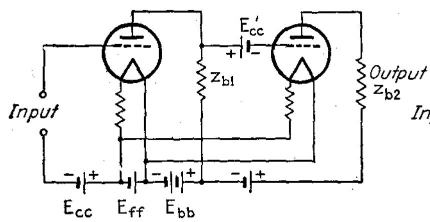


FIG. 5-3.—Direct-coupled amplifier.

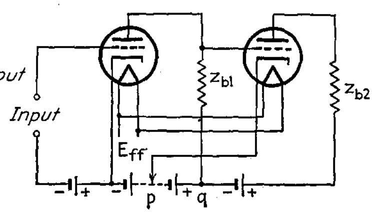


FIG. 5-4.—Direct-coupled amplifier.

In amplifiers designed to furnish voltage output, all stages, including the final one, should be voltage amplifiers. In amplifiers designed to furnish power output, on the other hand, the last stage must be a power amplifier. All other stages are voltage amplifiers unless the final power tube operates in such a manner that grid current flows during part of the cycle, in which case the next to the last stage must also be a power amplifier. The theory and characteristics of voltage, current, and power amplifiers will be considered in Chaps. 6 to 8.

5-5. Amplifier Circuits.—The classification of amplifiers according to circuits is based upon the number of stages, upon the type of circuit used to couple successive stages, and upon whether each stage uses a single tube or a symmetrical arrangement of two tubes. The coupling circuit serves a dual function. It converts the alternating plate current of one tube into alternating voltage to excite the grid of the following

[1] BALLANTINE, STUART, *Proc. I.R.E.*, **18**, 452 (1930).

tube; and, by preventing application of direct plate voltage of one tube to the grid of the following tube, it allows the proper operating voltages of all electrodes to be maintained. There are three fundamental types of coupling: *direct coupling, impedance-capacitance coupling,* and *transformer coupling.*

5-6. Direct Coupling.—A direct-coupled amplifier is one in which the plate of a given stage is connected to the grid of the next stage either directly or through a biasing battery. Basic circuits of two-stage amplifiers with direct coupling are shown in Figs. 5-3 and 5-4. The biasing battery adjacent to the grid of the second tube in the circuit of Fig. 5-3 is necessitated by the fact that the plate of the first tube is positive relative to its cathode, whereas the grid of the second tube must be negative relative to its cathode. The need for this battery is avoided in the circuit of Fig. 5-4 by making the cathode of the second tube positive with respect to the cathode of the first tube. The voltage of the cathode of the second tube in Fig. 5-4 is adjusted so that the voltage between p and q is less than the drop through z_{b1} by the required grid bias of the second tube.

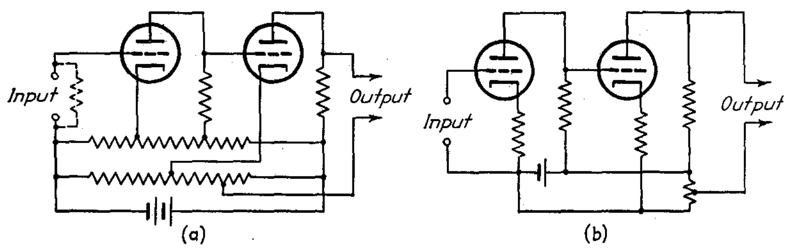


FIG. 5-5.—Direct-coupled amplifiers with a single source of B and C voltages.

The several voltage sources of Fig. 5-4 may be replaced by a single voltage source and a voltage divider.[1] In order to prevent the application to the grid of the first tube of voltage changes or alternating voltages resulting from the flow of plate current of the second tube through the voltage divider, it is desirable to use two voltage dividers, as shown in Fig. 5-5*a*. Direct voltage is prevented from appearing between the output terminals by connecting the lower output terminal to a point on the voltage divider that is at the same potential as the plate of the output tube. The flow of alternating plate current of each tube through its voltage divider applies to the grid of that tube an alternating voltage in phase opposition to the impressed signal voltage and thus causes some loss of amplification. This difficulty may be avoided by the use of push-pull circuits, which will be discussed in Sec. 5-10.[2] If the

[1] LOFTIN, E. H., and WHITE, S. Y., *Proc. I.R.E.*, **16**, 281 (1928); **18**, 669 (1930).
[2] GOLDBERG, H., *Trans. Am. Inst. Elec. Eng.*, **59**, 60 (1940).

amplifier is to be used only for the amplification of alternating voltages, the loss of amplification can be prevented in the circuit of Fig. 5-5*a* by connecting a condenser between each cathode and the negative side of the voltage supply. In order to avoid frequency distortion these condensers must be large enough so that their reactances are small at the lowest frequency to be amplified. Another circuit that makes possible the use of a single voltage source is shown in Fig. 5-5*b*. In this circuit the grid biasing voltages are obtained by the flow of plate current through resistors adjacent to the cathodes. The voltage drop across the cathode resistor of the second stage must be equal to the plate voltage of the first tube minus the magnitude of the grid bias of the second tube. This circuit is particularly applicable when the second tube is a power tube requiring a relatively high grid bias. When the amplifier is used only in the amplification of alternating voltages, loss in amplification resulting from alternating voltage produced across the cathode resistors may be prevented by shunting them with condensers of low reactance. In the amplification of direct voltages this type of circuit is of principal value in push-pull form, in which the cathode resistors do not cause loss of amplification. Such a circuit is shown in Fig. 6-44. The use of cathode resistors to produce grid bias will be discussed in further detail in Sec. 5-8.

Another form of direct-coupled amplifier is similar to the circuit of Fig. 5-6, except that the coupling condenser C_c is replaced by a glow-discharge tube. Such a tube is characterized by a practically constant voltage drop over a wide range of current (see Sec. 11-4). Variations of voltage of the plate of one tube are therefore passed on to the grid of the following tube and the difference in direct voltage between these electrodes appears as voltage drop in the glow tube. The objection to such a circuit arises from the fact that the current that flows through the glow tube has small random fluctuations which cause relatively large disturbances in the amplifier output. Although these disturbances may be very serious in amplifiers used for the production of sound or in other applications, glow-tube coupling is sometimes used in vacuum-tube instruments and in control apparatus. An example of the use of glow-tube coupling is afforded by the circuit of Fig. 6-37.

Direct-coupled amplifiers respond down to zero frequency, *i.e.*, they amplify changes of direct voltage. Although the small frequency distortion and the response at zero frequency are advantages of the direct-coupled amplifier, the response to changes of steady voltage makes it difficult to use more than two stages in the circuits of Figs. 5-3 to 5-5. Small changes in the operating voltages of the first tube are amplified to such an extent that it is hard to maintain correct grid bias in the final stage of an amplifier having three or more stages. This difficulty may be reduced, however, by the use of inverse feedback (see Sec. 6-36) and push-pull circuits.[1] A practical three-stage circuit is shown in Fig. 6-44.

5-7. Impedance-capacitance Coupling.—In the impedance-capacitance–coupled amplifier of Fig. 5-6 the plate voltage is kept from the grid of the succeeding tube by the use of a coupling condenser C_c. If the reactance of this condenser is small in comparison to the grid coupling impedance z_2, practically the full voltage developed across z_1 is applied to the grid of the second tube. z_2 is almost always a resistance. The plate coupling impedance z_1 may be a resistance, an inductive reactance, a resonant circuit, or a more complicated type of impedance. When resistance is used, the amplifier is termed a *resistance-capacitance–coupled* amplifier, or simply a *resistance-coupled* amplifier.

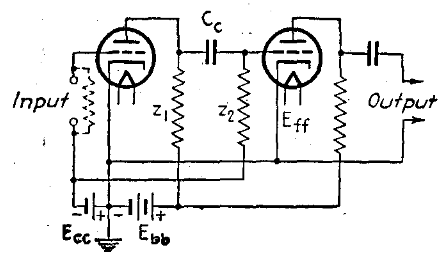


FIG. 5-6.—Impedance-capacitance–coupled amplifier.

5-8. Use of Cathode Resistors to Provide Bias.—It is unnecessary to use a separate voltage source to supply grid-bias voltages. The plate and grid supply voltages of Fig. 5-6 may be replaced by a single voltage source and voltage divider, or the bias voltages may be obtained

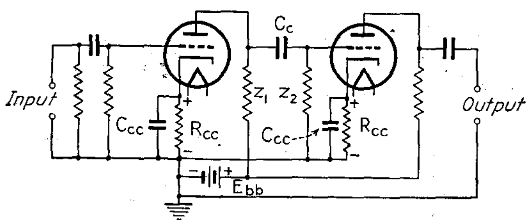


FIG. 5-7.—Impedance-capacitance–coupled amplifier with self-biasing resistors R_{cc}.

by the use of cathode resistors, as in Fig. 5-7. The steady components of plate and screen currents through a biasing resistor R_{cc} cause a steady voltage drop that is of such polarity as to make the grid negative with respect to the cathode. The tube is said to be *self-biased.* The correct value of biasing resistance for any stage is equal to the required bias divided by the sum of the static operating plate and screen currents. If the resistance alone is used, the signal voltage produced across this resistance is also applied to the grid and, being opposite in phase to the input voltage, reduces the amplification. Although the effects of the

[1] CLAPP, J. K., *Gen. Radio Expt.*, **9**, February, 1939, p. 1; GOLDBERG, *loc. cit.;* GOODWIN, C. W., *Yale J. Biol. and Med.*, **14**, 101 (1941).

out-of-phase voltage applied to the grid are not necessarily without benefit (see Sec. 6-29), the loss in amplification may be prevented by shunting the resistance with a *by-pass* condenser C_{cc}, whose reactance is small at signal frequency. To ensure that the amplification will not fall off at low frequencies, this condenser must have sufficiently high capacitance so that the alternating voltage across R_{cc} is negligible at the lowest frequency to be amplified. Because the required biasing voltage is usually relatively small, a low-voltage, high-capacitance electrolytic condenser (25-volt, 25-μf, for instance) may ordinarily be used.

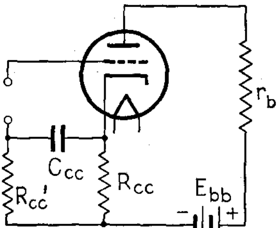


FIG. 5-8.—Single-stage amplifier with self-biasing resistor and decoupling filter.

Another method of preventing the application of out-of-phase signal voltage to the grid is by the *decoupling* circuit of Fig. 5-8. The condenser C_{cc} and resistor R_{cc}' serve as a simple filter by acting as a voltage divider for any alternating voltage appearing across R_{cc}.[1] If the reactance of C_{cc} is small in comparison with the resistance R_{cc}' at the lowest frequency to be amplified, a negligible portion of the alternating voltage will appear across C_{cc} at this and higher frequencies. The condenser charges up to a voltage equal to the direct voltage across R_{cc} and thus applies this amount of biasing voltage to the grid circuit. The advantage of the circuit of Fig. 5-8 over that of Fig. 5-7 is that it requires a smaller value of C_{cc}. Its disadvantage is the need for an additional resistor.

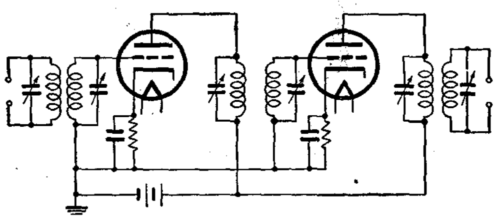


FIG. 5-9.—Doubly tuned transformer-coupled amplifier.

The by-pass condenser in the circuit of Fig. 5-7 and the condenser-resistance filter in the circuit of Fig. 5-8 also serve to reduce hum by preventing ripple voltage, which appears across R_{cc} as the result of a poorly filtered B supply (see Chap. 14), from being applied to the grid. In some circuits grid bias is derived from the nearly constant voltage

[1] POUND, F. J. A., *Wireless Eng.*, **9,** 445 (1932); KINROSS, R. I., *Wireless Eng.*, **10,** 612 (1932); STEVENS, B. J., *Wireless Eng.*, **11,** 129 (1934); CLARKE, G. F., *Wireless Eng.*, **11,** 370 (1934); WILLIAMS, EMRYS, *Wireless Eng.*, **11,** 600 (1934). See also Sec. 6-38.

drop across a glow-discharge tube (see Sec. 12-4 and Figs. 14-21*c*, 14-22, and 14-23).

5-9. Transformer Coupling.—Basic circuits of transformer-coupled amplifiers are shown in Figs. 5-9 to 5-11. The amplifiers of Figs. 5-9

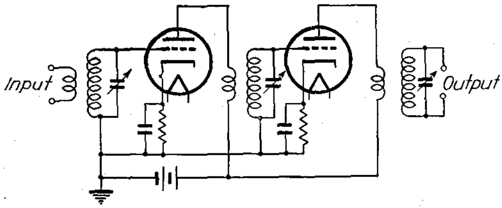


FIG. 5-10.—Singly tuned transformer-coupled amplifier.

and 5-10, which incorporate resonant circuits, are used principally at radio frequencies. The transformers may be either air-core or iron-core. The amplifier of Fig. 5-11, which employs untuned iron-core transformers, is used principally at audio frequencies.

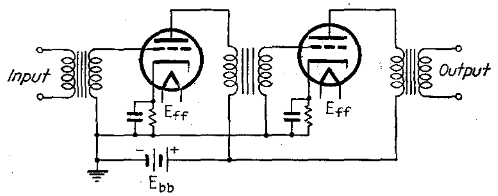


FIG. 5-11.—Untuned transformer-coupled amplifier.

5-10. Push-pull Amplifiers.—The circuits illustrated in Figs. 5-3 to 5-11 use a single tube in each stage of amplification. Such amplifiers are said to be *single sided*. It is also possible to use two tubes in each stage, connected so that the alternating grid voltages of the two tubes are opposite in phase, but the output voltages of the two tubes add, as in the single-stage circuit of Fig. 5-12. An amplifier that uses such a symmetrical arrangement of two tubes is called a *push-pull* amplifier.

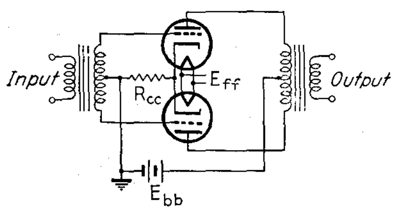


FIG. 5-12.—Single-stage push-pull amplifier.

Push-pull amplifiers have a number of advantages over single-sided amplifiers. (1) The principal advantage of push-pull amplifiers is that they do not introduce even harmonics of the impressed signal or the accompanying intermodulation frequencies associated with the even-order terms of the series expansion. (2) The reduction of even-order nonlinear distortion makes possible the use of higher bias and grid swing and hence the development of greater power output than that of the same two tubes used in parallel in a single-sided amplifier. (3) Hum voltage resulting from insufficient filtering of the B supply, and fluctuations of supply voltage, are applied in the same phase to the two tubes of a push-pull amplifier and hence the effects balance out of the output. (4) In push-pull transformer-coupled amplifiers the steady plate currents pass through the coupling-transformer primary in opposite directions and so do not tend to saturate the core if the two tubes have identical characteristics. For this reason a smaller core can be used. (5) The fundamental components of alternating plate currents of the two tubes, which are opposite in phase, cancel in the biasing resistance R_{cc} and therefore do not produce alternating voltage across this resistance. Even if the alternating currents of the two tubes are not exactly equal in amplitude, any alternating voltage produced across R_{cc} is applied in the same phase to the two grids and so this voltage does not produce any effect upon the output voltage. Hence the by-pass condenser may be omitted without loss of amplification, and danger of frequency distortion resulting from the variation of reactance of this condenser with frequency may be avoided. (6) Similarly, the fundamental components of plate current of the two tubes cancel in the impedance of the source of plate voltage and any small alternating voltage produced in this impedance if the currents are not exactly equal is applied to the two tubes in the same phase. Loss of amplification in individual stages or danger of oscillation as the result of feedback of signal voltage from any stage to a preceding stage (see Sec. 6-38) is thus avoided.

The symmetry of push-pull amplifiers can be made sufficiently close by choice of similar tubes and by minor circuit adjustments, so that the negative half of the output wave may be considered to have the same form as the positive half. A periodic wave that has this type of symmetry contains no even harmonics.[1] It follows, therefore, that a properly balanced push-pull amplifier does not generate even harmonics of the applied frequencies. This is also shown by Eqs. (4-34) to (4-36).

That a symmetrical push-pull amplifier does not generate even harmonics or even-order intermodulation frequencies may be proved rigorously by means of the series expansion for alternating plate current. Since the circuit is assumed to be symmetrical, the coefficients for the series are the same for the currents of both tubes. Let e_g and i_p be the alternating grid voltage and alternating plate current of one tube and e_g' and i_p' those of the other tube. Then

$$i_p = \Sigma a_1 e_g + \Sigma a_2 e_g^2 + \Sigma a_3 e_g^3 + \Sigma a_4 e_g^4 - \cdots \tag{5-1}$$

$$i_p' = \Sigma a_1 e_g' + \Sigma a_2 e_g'^2 + \Sigma a_3 e_g'^3 + \Sigma a_4 e_g'^4 - \cdots \tag{5-2}$$

[1] MALTI, M. G., "Electric Circuit Analysis," p. 176, John Wiley & Sons, Inc., New York, (1930).

But $e_g' = -e_g$. Therefore

$$i_p' = -\Sigma a_1 e_g + \Sigma a_2 e_g^2 - a_3 e_g^3 + \Sigma a_4 e_g^4 - \cdots \tag{5-3}$$

Because the plate currents flow through the primary of the coupling transformer in opposite directions, the voltage across the secondary is proportional to the difference between i_p and i_p'.

$$e_o = A(i_p - i_p') = 2A(\Sigma a_1 e_g + \Sigma a_3 e_g^3 + \Sigma a_5 e_g^5 + \cdots) \tag{5-4}$$

in which A, the constant of proportionality, is different for each frequency component of the expanded series. If $e_g = E_m \sin \omega t$ is substituted in Eq. (5-4) and the powers of $\sin \omega t$ are expanded, the resulting series contains only the fundamental and odd-harmonic frequencies. If e_g is the sum of several voltages of different frequencies, the series contains also odd-order intermodulation frequencies, but no even-order intermodulation frequencies. With triodes, the coefficients of the third- and higher-order terms of the series can be made small enough so that the output of a push-pull amplifier has negligible harmonic and intermodulation content. With pentodes a great reduction of amplitude distortion can be effected by the use of push-pull circuits operated under conditions that give small third-harmonic content but that would give high second-harmonic content with single-sided amplifiers.

If the output transformer of Fig. 5-12 is center tapped, it can be used to excite a following push-pull amplifier stage. Impedance-capacitance coupling may also be used, as shown in Fig. 5-13. In this circuit there can be no external connection between either side of the input and the cathode circuit, as this would in effect short-circuit the a-c input to one tube. For this reason the input can be grounded only at the midpoint of the input resistor. Although push-pull stages may be used throughout a multistage amplifier when it is desired to reduce distortion to a minimum, it is more common to use the push-pull connection only in the final stage, in which the current amplitude, and hence the distortion, are large.

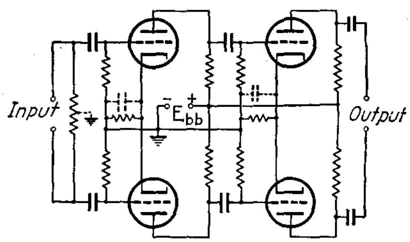


FIG. 5-13.—Resistance-capacitance-coupled push-pull amplifier.

5-11. Phase Inverters.—The requirement that the exciting voltages of the two tubes of a push-pull amplifier shall be opposite in phase necessitates the use of special circuits in coupling a single-sided stage to a push-pull stage by means of resistance-capacitance coupling.[1] One

[1] AUGHTIE, F., *Wireless Eng.*, **6**, 307 (1929); DAVIDSON, P. G., *Wireless Eng.*, **6**, 437 (1929); SHORTT, H. L., *Radio Eng.*, January, 1935, p. 14.

circuit that may be used for this purpose is shown in Fig. 5-14. In this circuit, the 180-degree difference in phase of the exciting voltages of the push-pull tubes is obtained by splitting the plate resistor of the first stage into two parts and inserting the B-supply voltage between them.[1] The signs in Fig. 5-14 show the polarities of the instantaneous voltages across the various resistors when the instantaneous input voltage is of such polarity as to swing the grid voltage of the first tube in the positive direction. This circuit has the disadvantage that the input cannot be grounded if the voltage supply is grounded. This disadvantage may be avoided by connecting the lower input terminal and the lower end of the input resistor to the negative side of the supply voltage E_{bb}, instead of as shown in Fig. 5-14.[2] The voltage developed across the lower plate resistor of the first tube is then applied to the grid in series with the input

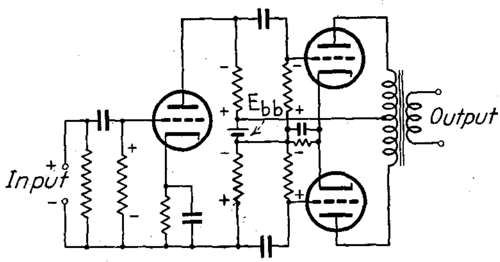


FIG. 5-14.—Amplifier circuit in which resistance-capacitance coupling is used to couple a single-sided stage to a push-pull stage.

voltage. Since these voltages are opposite in phase, the resultant grid voltage is small and the amplification of the first stage is approximately unity. The reduction of amplification, which may be undesirable, is accompanied by a decrease of nonlinear distortion in the first stage (see Sec. 6-30).

A second type of phase-inverter circuit, in which the phase reversal is accomplished by means of an extra tube, is shown in Fig. 5-15. The signs indicate the relative polarities of the instantaneous alternating voltages throughout the circuit. The alternating grid voltage of tube T_2, which is obtained from the plate load resistance of T_1, is opposite in phase to the voltage impressed upon the grid of T_1. Hence the output voltages of the two tubes are opposite in phase. The slider of the voltage divider is set so that $(r_1 + r_2)/r_1$ is equal to the voltage amplification of T_1. The alternating grid voltages and, therefore, the output voltages, of the two tubes are then equal in amplitude. A twin type of tube is usually used for T_1 and T_2. A simple method of balancing the circuit is to connect a pair of headphones across the biasing resistor and to

[1] TULAUSKAS, L., *Electronics*, May, 1933, p. 134.
[2] MCPROUD, C. G., and WILDERMUTH, R. T., *Electronics*, October, 1940, p. 50.

adjust the voltage divider so that only harmonic and other distortion frequencies are heard.

A third type of phase-inverter circuit, called the *cathode phase inverter*, is shown in basic form in Fig. 5-16*a*.[1] The alternating grid voltage for T_2, which is obtained from the resistance in the cathode circuit of T_1, is

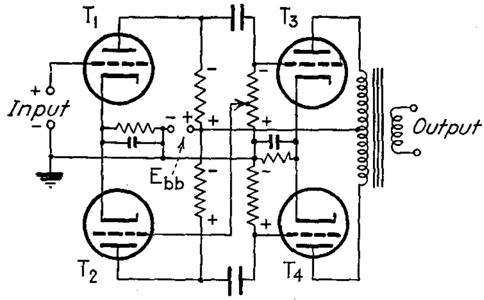


FIG. 5-15.—Use of a phase-inverting tube in coupling a single-sided stage to a push-pull stage.

opposite in phase to the alternating grid voltage of T_1, and so the two output voltages (measured relative to the common output terminal) are opposite in phase. A simple analysis, based upon the equivalent plate circuits (see Prob. 5-1), shows that the ratio of E_{o1} to E_{o2} is equal to $r_{b1}[1 + (r_p + r_{b2})/(\mu + 1)r_k]/r_{b2}$. This ratio can be made equal to unity by making r_{b2} larger than r_{b1}. When it is desired, for the sake of sim-

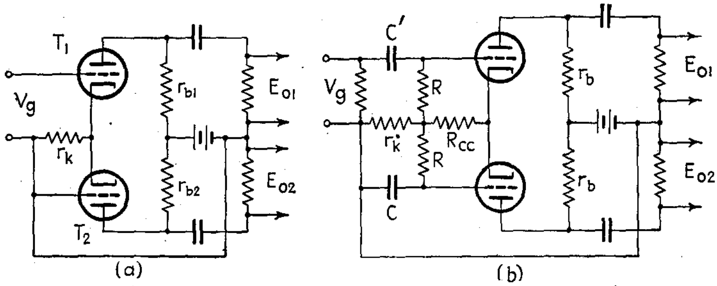


FIG. 5-16.—Cathode phase inverter (*a*) with a common cathode coupling and biasing resistor and (*b*) with coupling resistance greater than the biasing resistance.

plicity, to make r_{b1} and r_{b2} of equal value r_b, the amplitudes of the two output voltages can be made to approach equality by making $(\mu + 1)r_k$ large in comparison with $r_p + r_b$. The voltage amplification is then approximately $\mu r_b/2(r_p + r_b)$. In order to make the two output voltages essentially equal in this manner, it is necessary to use values of r_k considerably larger than those required to provide the correct grid bias. In

[1] SCHMITT, O. H., *J. Sci. Instruments*, **15**, 136 and 234 (1938); *Rev. Sci. Instruments*, **12**, 548 (1941).

the modified circuit of Fig. 5-16*b* the bias is provided by the resistance R_{cc}, whereas the alternating grid voltage of T_2 is taken from across both r_k and R_{cc} through the condenser C, the reactance of which is negligible in comparison with the resistance R.

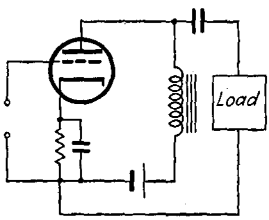


FIG. 5-17.—Choke-condenser–coupled load.

5-12. Output Circuits.—It is often advantageous or necessary to prevent the steady component of plate current of the output tube from passing through the loud-speaker or other load. This may be done by using an output transformer as in Figs. 5-9 to 5-12, 5-14, and 5-15, a resistance-capacitance network as in Figs. 5-7 and 5-13, or a condenser-choke combination as in Fig. 5-17.

5-13. Use of Center-tapped Filament Transformer with Filamentary Cathodes.—When tubes with filamentary cathodes are operated from an a-c filament supply, the grid and plate circuits are connected to the cathode through a center tap on the filament transformer or through a center-tapped resistor shunted across the filament. The reason for the center connection is to prevent a-c output at supply frequency as the result of a variation of voltage of one end of the filament relative to the grid. The center connection is equivalent to a connection to the mid-point of the filament. When the alternating voltage drop in the filament causes the voltage of the grid relative to one end of the filament to become more negative, it causes the grid voltage relative to the other end to become less negative. If the filament is symmetrical, the decrease of space current from one end of the filament is offset by an equal increase from the other end, and the plate current remains constant. Slight deviation of the filament from symmetry with respect to the mid-point can be compensated by the use of a resistor with an adjustable tap.

5-14. Frequency Range.—According to frequency range, amplifiers are classified as *wide-band* and *narrow-band*. The presence of circuit inductances and capacitances restricts the frequency range over which any amplifier can amplify uniformly. The width of the response band may be reduced to any desired amount by the use of tuned circuits or band-pass filters in the amplifier. For most applications of audio-frequency amplifiers it is desirable to amplify uniformly over as great a frequency range as possible, and so untuned amplifiers are the rule. Tuned amplifiers, which are indispensable in radio-frequency amplification, are seldom of value in audio-frequency work unless it is desired to emphasize or repress certain frequencies. Amplifiers incorporating band-pass filters are of value at both radio and audio frequencies.[1]

[1] BUTTERWORTH, S., *Wireless Eng.*, **7**, 536 (1930).

They may be used in separating and isolating various portions or components of a wide band of frequencies which are simultaneously produced or transmitted.

5-15. Class A, Class AB, Class B, and Class C Amplifiers.—According to the portion of the cycle during which plate current flows, amplifiers are classified as Class A, Class AB, Class B, and Class C. A *Class A amplifier* is one in which the grid bias and alternating grid voltage are such that plate current in the tube, or in each tube of a push-pull stage, flows at all times. A *Class AB amplifier* is one in which the grid bias and alternating grid voltages are such that plate current in the tube, or in each tube of a push-pull stage, flows for appreciably more than half but less than the entire electrical cycle. A *Class B amplifier* is an amplifier in which the grid bias is approximately equal to the cutoff value, so that the plate current is approximately zero when no exciting grid voltage is applied and so that plate current in the tube, or in each tube of a push-pull stage, flows for approximately one-half of each cycle when an alternating grid voltage is applied. A *Class C amplifier* is an amplifier in which the grid bias is appreciably greater than the cutoff value, so that the plate current in each tube is zero when no alternating grid voltage is applied and so that plate current flows in each tube for appreciably less than one-half of each cycle when an alternating grid voltage is applied. The suffix 1 may be added to the letter or letters of the class identification to denote that grid current does not flow during any part of the input cycle, and the suffix 2 to denote that grid current flows during some part of the cycle.

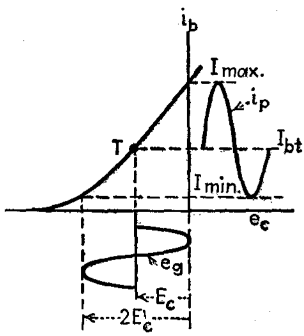


FIG. 5-18.—Single-sided Class A1 operation.

Class A1 operation is illustrated in Fig. 5-18. According to the definition of Class A amplification, the lower limit of instantaneous total plate current is zero. Practically, however, distortion caused by curvature of the dynamic transfer characteristic at low values of plate current limits the minimum plate current in a single-sided amplifier to values that ordinarily are not less than one-fifteenth of the maximum plate current and that may be considerably larger. Since grid current starts flowing when the grid voltage is zero or even slightly negative, the upper limit of plate current is approximately that corresponding to zero grid voltage. The largest amplitude of alternating plate current is evidently obtained when the grid bias and signal voltage have such values that the grid voltage is equal to or slightly less than zero at the positive crest of exciting voltage and the plate current has the minimum value of $I_{\min}$, consistent with allowable amplitude distortion, at the negative crest of exciting voltage.

Although larger amplitude of alternating plate current can be obtained with Class A2 operation than with Class A1, this advantage is offset by the higher value of static operating current, which increases the power consumption and the loss within the tube (see Sec. 7-2), and by the complications arising from the flow of grid current (see Sec. 8-9). Class A2 operation is, therefore, seldom used. Class A1 operation is used in voltage amplifiers and in many audio-frequency power amplifiers. It results in lower distortion than do other types of operation but is less efficient than Class AB, B, or C operation.

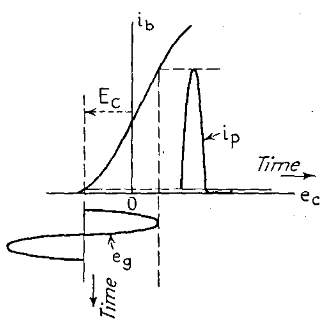


FIG. 5-19.—Single-sided Class B2 operation.

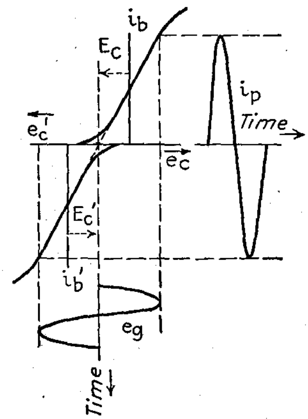


FIG. 5-20.—Push-pull Class B2 operation.

Class B2 operation is illustrated in Figs. 5-19 and 5-20. If a single tube were used as a Class B audio-frequency amplifier, all, or nearly all, of the negative half of the cycle would be cut off, as shown in Fig. 5-19, and amplitude distortion would be excessive. When two tubes are used in push-pull, however, the grid bias can be adjusted so that each tube functions in alternate half cycles during a little more than half the cycle, and the resultant current through the load is very nearly a replica of the exciting grid voltage, as shown in Fig. 5-20. In radio-frequency amplifiers, tuned circuits may be used in the plate circuit to suppress harmonics in the output and thus make possible the use of a single-sided amplifier. Class B operation is used in a-f amplification only in power amplifiers. Class B amplifiers are capable of delivering large amounts of power at higher efficiency than Class A amplifiers, but the harmonic content is greater. Difficulties of design result from the flow of grid current at the high excitation voltage required in order to take full advantage of the capabilities of Class B amplifiers. The characteristics of Class B audio amplifiers and the special problems encountered in their design will be treated in detail in Chap. 8.

Class AB operation is intermediate between that of Class A and Class B. Single-sided Class AB1 and AB2 operation are illustrated in Figs. 5-21*a* and 5-21*b*. Because plate current in a single tube does not flow during the entire cycle, it is necessary to use push-pull circuits to avoid excessive distortion in Class AB a-f amplifiers. Often the grid is not allowed to swing positive, since the flow of grid current results in complications which may offset the advantages (see Sec. 8-9). Considerably greater power output and higher efficiency can be obtained with Class AB1 amplifiers than with Class A1 amplifiers. Class AB amplifiers will be treated in Chaps. 7 and 8.

Because of the fact that plate current flows during less than 180 electrical degrees it is impossible to use Class C operation in a-f amplifiers

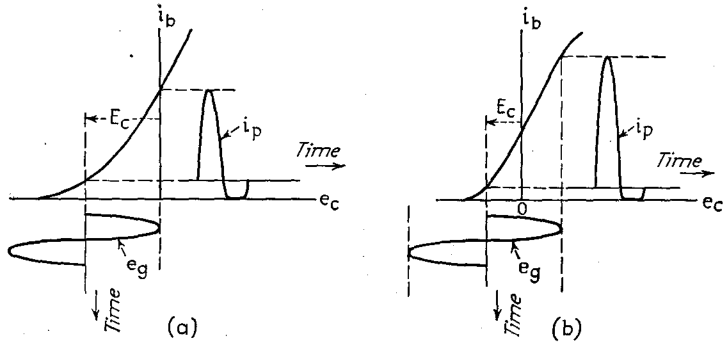


FIG. 5-21.—(*a*) Single-sided Class AB1 operation. (*b*) Single-sided Class AB2 operation.

designed for wide frequency bands. In r-f amplifiers, harmonics are suppressed by the use of resonant circuits. The efficiency of Class C amplifiers is greater than that of the other classes.

5-16. The Decibel.—Although current, voltage, and power amplification, as well as the magnitude of a given voltage, current, or power, relative to reference values, can be expressed as an ordinary ratio, it has been found to be far more convenient to make use of logarithmic ratios. The unit that has been adopted in this country is the decibel, which is indicated by the symbol db and is defined as follows:

1. The ratio of two amounts of power P_2 and P_1 is said to be n db if

$$n = 10 \log_{10} \frac{P_2}{P_1} \tag{5-5}$$

2. The ratio of two voltages E_2 and E_1 or two currents I_2 and I_1 is said to be n db if

$$n = 20 \log_{10} \frac{E_2}{E_1} \qquad \text{or} \qquad n = 20 \log_{10} \frac{I_2}{I_1} \tag{5-6}$$

If P_2 exceeds P_1, n is positive, P_2 is said to be "up" n db with respect to P_1, and n is said to indicate a *gain;* if P_1 exceeds P_2, n is negative, P_2 is said to be "down" n db with respect to P_1, and n is said to indicate a *loss.*

By substituting $P = I^2z \cos\theta$ and $P = \dfrac{E^2}{z}\cos\theta$ in Eq. (5-5) the student may show that Eqs. (5-5) and (5-6) give the same number of decibels if the impedances in which P_1 and P_2 are developed have the same magnitude and phase angle.

It follows from the definition of a logarithm that if a system is made up of a number of units whose decibel gain or loss is n_1, n_2, n_3, etc., the over-all gain or loss of the system is

$$n = n_1 + n_2 + n_3 + \cdots \tag{5-7}$$

due account being taken of the signs of n_1, n_2, n_3, etc.

It should be noted that the decibel always refers to the ratio of two amounts of power, voltage, or current. A single quantity of power, voltage, or current can be specified in decibels by agreeing to express the ratio always with respect to a fixed reference value, called *zero level.* In telephone engineering practice, 6 mw has long been accepted as zero power level. The somewhat more convenient value of 1 mw, as well as other values, is often used. The term *volume unit* is used in place of "decibel" in specifying power when 1 mw is used as zero power level.[1]

The small numbers in which it is possible to express the decibel gain corresponding to large amplification ratios and the ease of adding and subtracting, as compared with multiplying and dividing, constitute two advantages of the use of the decibel. Furthermore, the change in gain of an amplifier in decibels is a better index of the effect of the sound output upon the ear, than the corresponding change in amplification.[2]

A chart for determining decibel gain corresponding to power, voltage, and current ratios is given on page 676.

Problems

5-1. *a.* By the use of equivalent plate circuits, show that the voltage amplification of tube T_1 in the circuit of Fig. 5-16*a* is

$$A = -\frac{\mu r_{b1}[r_p + r_{b2} + (\mu + 1)r_k]}{(r_p + r_{b1})(r_p + r_{b2}) + (2r_p + r_{b1} + r_{b2})(\mu + 1)r_k} \tag{5-8}$$

and that the ratio of the output voltages of the two tubes is

$$\frac{E_{o1}}{E_{o2}} = -\left[1 + \frac{(r_p + r_{b2})}{(\mu + 1)r_k}\right]\cdot\frac{r_{b1}}{r_{b2}} \tag{5-9}$$

[1] AFFEL, H. A., CHINN, H. A., and MORRIS, R. M., *Electronics*, February, 1939, p. 28; CHINN, H. A., GANNETT, D. K., and MORRIS, R. M., *Proc. I.R.E.*, **28**, 71 (1940).

[2] PERRY, S. V., *RMA Tech. Bull.* 1, Nov. 1, 1940.

b. Show that when $r_{b1} = r_{b2} = r_b$ and $(\mu + 1)r_k >> r_p + r_b$, Eqs. (5-8) and (5-9) reduce to

$$A = -\frac{\mu r_b}{2(r_p + r_b)} \tag{5-10}$$

$$\frac{E_{o1}}{E_{o2}} = -1 \tag{5-11}$$

5-2. The voltage produced by the flow of 1 ma through a resistance of 1000 ohms is amplified to give a voltage of 100 volts across a 10,000-ohm resistance. Find the voltage gain, current gain, and power gain in decibels.

CHAPTER 6

ANALYSIS AND DESIGN OF VOLTAGE AND CURRENT AMPLIFIERS

Voltage amplifiers are usually operated in such a manner that amplitude distortion is small. Much can be learned regarding their performance, therefore, by taking into account only the fundamental components of plate current and making use of the equivalent-plate-circuit theorem. The results of such an approximate analysis are closely verified by laboratory measurements. When it is necessary to determine the harmonic content or to make more accurate predictions regarding the amplification, the graphical methods explained in Chap. 4 may be employed.

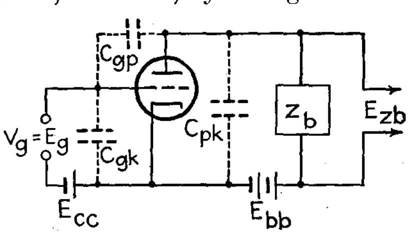


FIG. 6-1.—Single-sided amplifier with impedance load.

6-1. Voltage Amplification of Tube with Impedance Load.—The simplest form of vacuum-tube voltage amplifier consists of a single tube with an impedance in the plate circuit, as shown in Fig. 6-1. Figure 6-2 shows the equivalent plate circuit. As far as its shunting effect upon z_b is concerned, C_{gp} may be replaced by an equivalent capacitance

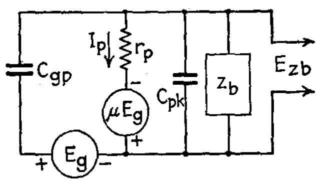


FIG. 6-2.—Equivalent circuit for the amplifier of Fig. 6-1.

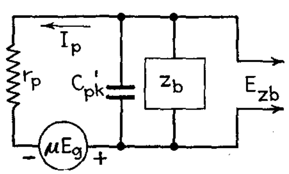


FIG. 6-3.—Simplified equivalent circuit for the amplifier of Fig. 6-1.

$C_{gp}' = C_{gp}(E_g + E_{zb})/E_{zb}$, in parallel with C_{pk}. Since E_{zb} is usually large in comparison with E_g, C_{gp}' is approximately equal to C_{gp}. The equivalent circuit may then be simplified to that of Fig. 6-3, in which $C_{pk}' = C_{pk} + C_{gp}$. At low frequency the reactance of C_{pk}' is so high that its effect may be neglected. Under this assumption the voltage amplification is

$$A = \frac{E_{zb}}{V_g} = \frac{E_{zb}}{E_g} = -\frac{I_p z_b}{E_g} = -\frac{\mu z_b}{r_p + z_b} \tag{6-1}$$

The manner in which $|A/\mu|$ varies with the ratio $|z_b/r_p|$ with pure resistance load and with pure reactance load is shown in Fig. 6-4. The voltage amplification approaches the amplification factor when the load impedance becomes large in comparison with the plate resistance. Inductance load has the advantage that loss of direct voltage resulting from IR drop in the load may be kept to a minimum by making the d-c resistance small. This advantage is more than offset, however, by dependence of load impedance upon frequency, which tends to make the amplification fall and to advance the phase of the output voltage with respect to the exciting voltage at low frequency.

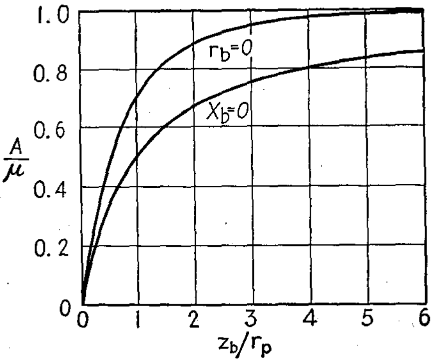


FIG. 6-4.—Curves showing the manner in which the amplification of a single tube with pure resistance load and with pure reactance load varies with load impedance.

At frequencies which are so high that the reactance of C_{pk}' is comparable with the plate resistance, the effective load impedance z_b consists of the parallel combination of z_b and the reactance of C_{pk}'. The voltage amplification therefore falls off at high frequency, and the phase of the output voltage is retarded. The upper limit of the range of uniform amplification can be raised at the expense of amplification, by reducing z_b. With resistance load it is not difficult to obtain uniform amplification from zero frequency well into the radio-frequency range. Special methods of improving the high-frequency response will be discussed.

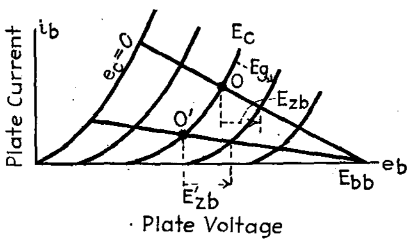


FIG. 6-5.—Plate diagram of a triode with pure resistance load, showing the variation of voltage output with load resistance at constant excitation.

For triodes with resistance load, the manner in which A varies with r_b is complicated by the fact that, as this resistance is increased, the path of operation is lowered and the plate resistance is increased. Thus, the plate resistance at the operating point O' of Fig. 6-5 is higher than that at the point O. The amplification does not increase so rapidly, therefore, as would be indicated by Eq. (6-1) if the plate resistance were assumed to be constant. The manner in which the voltage amplification varies with load resistance can be determined most readily graphically. The voltage amplification is roughly equal to the difference in plate voltage of points of intersection of the load line with two adjacent static plate characteristics, divided by the difference of grid voltage of the two characteristics. This is the ratio E_{zb}/E_g or E_{zb}'/E_g in Fig. 6-5. The accuracy of this method increases, of course, as the interval between the characteristics is decreased. It can be seen from Fig. 6-5 that the rate at which the amplification is increased with load resistance is small at high values of load resistance. For this reason and because of the falling off of amplification at high frequency when r_b is large, the load resistance is usually limited to 500,000 ohms or less in practice, even with triodes having high plate resistance. The voltage amplification that can be realized in practice with triodes approximates 80 per cent of the amplification factor.

For pentodes with resistance load, the action is complicated by the fact that at low values of plate current the spacing of the static characteristics rapidly becomes smaller as the negative grid voltage is increased. The resulting curvature of the dynamic transfer characteristic causes amplitude distortion. This difficulty can be reduced by increasing the plate supply voltage, but the required voltage rapidly becomes prohibitive as the load resistance is increased. Practical values of pentode load resistance are also limited to about 500,000 ohms. Since pentode plate resistances are well in excess of a megohm, the realizable voltage amplification is much smaller than the amplification factor, 350 representing the approximate limit. Because the plate resistance of pentodes is considerably higher than the usable load resistance, Eq. (6-1) reduces to $A = g_m z_b$ for pentodes.

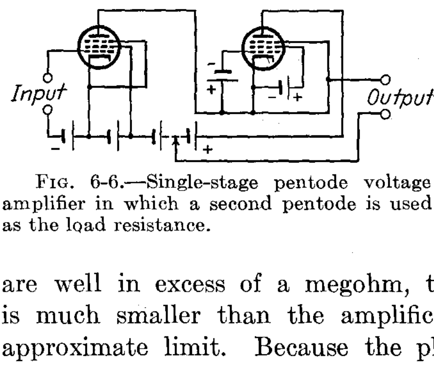


FIG. 6-6.—Single-stage pentode voltage amplifier in which a second pentode is used as the load resistance.

Values of voltage amplification approximating the amplification factor of pentodes can be obtained by using the plate resistance of a second pentode as the load resistance,[1] as shown in Fig. 6-6. The a-c resistance of the pentode that is used as load is high, but its electrode voltages may be adjusted so that the d-c drop through it is small, and so that its plate current is equal to the desired operating plate current of the amplifier pentode. In this manner it is possible to obtain high a-c load resistance without lowering the operating plate voltage and current of the amplifier pentode to values that give excessive distortion. The interelectrode capacitances of the two tubes cause the amplification to fall at the upper end of the audio-frequency range.

[1] HERD, J. F., *Wireless Eng.*, **8**, 192 (1931); MEISSNER, E. R., *Electronics*, July, 1933, p. 195; HORTON, J. W., *J. Franklin Inst.*, **216**, 749 (1933); SCHMITT, O., *Rev. Sci. Instruments*, **4**, 661 (1933).

6-2. Voltage Amplification of Multistage Amplifiers.—The over-all amplification of a multistage amplifier is equal to the product of the amplifications of the individual stages. (The over-all gain is equal to the sum of the gains of the individual stages.) In forming the equivalent circuit of each stage, the stage must be considered to consist of the tube and all circuit elements coupled to the plate, including the tube capacitances and the input admittance of the following tube.[1] Different types of coupling may, of course, be used in the different stages. The form of the plate load of the final stage depends upon the purpose for which the amplifier is designed.

6-3. Voltage Amplification of Direct-coupled Amplifier.—The direct-coupled amplifier illustrated in Figs. 5-3 to 5-5 consists of two stages of the simple amplifier of Fig. 6-1. Because of the effect of the input capacitance of the second tube upon the amplification of the first stage, the amplification begins to fall off at a lower frequency than with a single stage; otherwise the behavior is similar. The over-all amplification is the product of the amplifications of the two stages (the sum of the decibel gains), which may be determined from the equivalent circuit of Fig. 6-3. The direct-coupled amplifier is a special form of an impedance-capacitance–coupled amplifier in which the coupling capacitance and grid-leak impedance are infinite. Since the impedance-capacitance–coupled amplifier is analyzed in the following sections, the direct-coupled amplifier need not be treated in further detail. If the coupling and output impedances of this type of amplifier are nonreactive, the amplification is independent of frequency from zero frequency up to frequencies at which the effect of interelectrode capacitances becomes apparent. This type of amplifier gives the least frequency and phase distortion and is the only type of multistage amplifier that will amplify at zero frequency, *i.e.*, that will amplify changes in direct voltage or current.

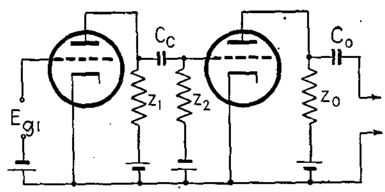


FIG. 6-7.—Two-stage impedance-capacitance–coupled amplifier.

6-4. Impedance-capacitance–coupled Voltage Amplifier.—The basic circuit of a two-stage impedance-capacitance–coupled amplifier is shown in Fig. 6-7. The final stage may be considered as a special form of impedance-capacitance–coupled stage in which one or two of the coupling elements are absent. If the amplifier is used to furnish voltage output only, *i.e.*, if the external impedance across the output terminals

[1] The student is cautioned against attempting to combine the equivalent circuits of two or more stages into a single equivalent circuit.

is infinite, the load of the final stage may be considered to be merely z_o and the amplification may be found by the use of Eq. (6-1).

The equivalent circuit of an impedance-capacitance-coupled stage is shown in Fig. 6-8*a*. Because the effective plate-to-cathode capacitance C_{pk1} is very much smaller than the capacitance of the coupling condenser C_c, the effect of C_{pk1} upon the behavior of the amplifier is the same as though C_{pk1} were connected in parallel with the input impedance z_{g2} of the following tube.[1] Further simplification results from the assumption that the input conductance of the following tube is negligible. Then the admittance is [see Eq. (4-13)]

$$Y_{g2} = B_{g2} = j\omega C_{i2} = j\omega[C_{gk2} + (1 + |A_2|)C_{gp2}] \tag{6-2}$$

in which $|A_2|$ is the magnitude of the voltage amplification of the following stage, and C_{gk2} and C_{gp2} are the interelectrode capacitances of the following

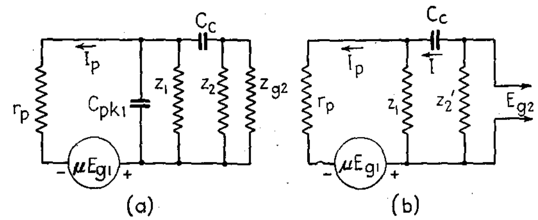


Fig. 6-8.—Equivalent plate circuit for the first stage of the amplifier of Fig. 6-7.

tube. The effective input capacitance C_{i2} of the following tube, in parallel with the effective plate-to-cathode capacitance C_{pk1}, may be replaced by an equivalent capacitance

$$C_2 = C_{pk1} + C_{gk2} + (1 + |A_2|)C_{gp2} \tag{6-3}$$

For convenience in analysis the parallel combination of z_2 and C_2 may be replaced by the equivalent impedance

$$z_2' = \frac{z_2}{1 + j\omega C_2 z_2} \qquad \text{or} \qquad y_2' = y_2 + j\omega C_2 \tag{6-4}$$

Figure 6-8*b* shows the simplified equivalent circuit. Summation of voltages in the circuit of Fig. 6-9 yields the following equations

$$I_p(r_p + z_1) - Iz_1 = \mu E_{g1} \tag{6-5}$$

$$I\left(z_1 + z_2' - \frac{j}{\omega C_c}\right) = I_p z_1 \tag{6-6}$$

Examination of Fig. 6-7 shows that the second grid is made more negative when I_p and I are increased by making the grid of the first tube less

[1] Strictly, C_{pk1}', as defined on p. 144, rather than C_{pk1}, should be used in this analysis. In tubes suitable for use in voltage amplifiers, however, C_{gp1} is so much smaller than C_{pk1} that little error results from using C_{pk1} instead of C_{pk1}'.

negative. Therefore, $E_{g2} = -Iz_2'$, and the voltage amplification is

$$A = \frac{E_{g2}}{E_{g1}} = -\frac{Iz_2'}{E_{g1}} \tag{6-7}$$

Solution of the simultaneous equations (6-5) to (6-7) gives the following general expression for the voltage amplification of a stage of an impedance-capacitance–coupled amplifier:

$$A = -\frac{\mu z_1 z_2'}{(z_1 + z_2')r_p + z_1 z_2' - \dfrac{j(r_p + z_1)}{\omega C_c}} \tag{6-8}$$

Equation (6-8) may be readily transformed into the equivalent equation

$$A = -\frac{\mu}{(y_1 + y_2')r_p + 1 - j\dfrac{y_2'}{\omega C_c}(y_1 r_p + 1)} \tag{6-9}$$

Equations (6-8) and (6-9) and their rationalized equivalents are too complicated to show clearly the influence of the various circuit constants upon the voltage amplification. The formulas for special forms of the impedance-capacitance–coupled amplifier are considerably simpler.

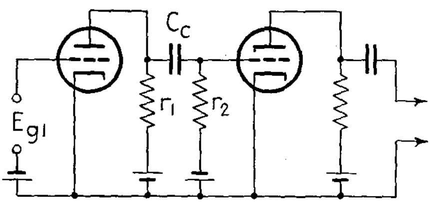


FIG. 6-9.—Two-stage resistance-capacitance–coupled amplifier.

6-5. Resistance-capacitance–coupled Amplifier.—The most commonly used type of impedance-coupled amplifier is the resistance-capacitance–coupled amplifier, in which the coupling impedance z_1 and grid coupling (grid-leak) impedance z_2 are resistances r_1 and r_2, as shown in Fig. 6-9. Figure 6-10 shows a typical frequency-response curve of such an amplifier. It can be seen from this curve that the amplification remains constant over a considerable range of frequency but falls on either side of this range. The range of constant amplification is called the *mid-band range* of frequency. The amplification in this range is called the *mid-band amplification* and is represented by the symbol A_m. The mid-band range usually extends from below 100 cps to above five thousand cps, and may be greatly extended in both directions.

Examination of Fig. 6-9 shows that the reduction of voltage amplification at low frequencies must result from the voltage-dividing action of

C_c and r_2. Since the reactance of C_c increases as the frequency is reduced, more of the alternating voltage across r_1 appears across the condenser and less is applied to the grid of the following tube. The reduction of amplification at high frequencies, on the other hand, must result from the shunting effect of C_2, the effective input capacitance of the following stage. As the frequency goes up, the resultant impedance of the parallel combination of r_1, r_2, and C_2 goes down, and so the voltage amplification falls, as predicted by Eq. (6-1) and shown by Fig. 6-4. The constant amplification in the mid-band frequency range indicates that within this range the reactance of the series capacitance C_c is so small that C_c may be considered to be short-circuited and that the impedance of the parallel capacitance C_2 is so high that C_2 may be considered to be absent. Furthermore, the effect of C_2 is negligible in the frequency range in which C_c reduces the amplification, and the effect of C_c is negligible in the range in which C_2 reduces the amplification. Hence the mid-band, high, and low ranges of frequency may be considered separately in analyzing the behavior of the amplifier.

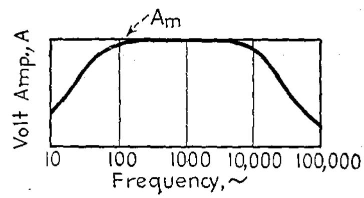


FIG. 6-10.—Typical shape of the frequency-response characteristic of a resistance-capacitance-coupled amplifier.

Although expressions for the voltage amplification of a resistance-capacitance-coupled amplifier can be readily obtained by making appropriate simplifications in Eq. (6-9), it is more instructive to derive them directly from equivalent circuits. Since both C_c and C_2 have negligible effect in the mid-band range, the mid-band load of the first stage of the circuit of Fig. 6-9 consists of r_1 and r_2 in parallel. The voltage amplification can be found from Eq. (6-1). If the numerator and denominator of the right side of Eq. (6-1) are divided by $z_b r_p$, the equation becomes

$$A = -\frac{\mu/r_p}{\dfrac{1}{r_p} + \dfrac{1}{z_b}} = -\frac{g_m}{\dfrac{1}{r_p} + \dfrac{1}{z_b}} \tag{6-10}$$

But since the effective load impedance is that of r_1 and r_2 in parallel,

$$\frac{1}{z_b} = \frac{1}{r_1} + \frac{1}{r_2} \tag{6-11}$$

Substitution of Eq. (6-11) in Eq. (6-10) gives for the mid-band amplification

$$A_m = -\frac{g_m}{\dfrac{1}{r_p} + \dfrac{1}{r_1} + \dfrac{1}{r_2}} = -g_m r_h \tag{6-12}$$

in which

$$r_h = \frac{1}{\dfrac{1}{r_p} + \dfrac{1}{r_1} + \dfrac{1}{r_2}} = \frac{r_1 r_2 r_p}{r_1 r_2 + r_1 r_p + r_2 r_p} \tag{6-13}$$

The approximate equivalent circuit of the first stage of the amplifier of Fig. 6-9 at frequencies above the mid-band range is that of Fig. 6-11. The voltage amplification can be found from Eq. (6-10) if it is noted that the load is the parallel combination of r_1, r_2, and C_2.

$$\frac{1}{z_b} = \frac{1}{r_1} + \frac{1}{r_2} + j\omega C_2 \tag{6-14}$$

Substitution of Eq. (6-14) in Eq. (6-10) gives for the high-frequency amplification (see also Sec. A-1, page 671)

$$A_h = -\frac{g_m}{\dfrac{1}{r_p} + \dfrac{1}{r_1} + \dfrac{1}{r_2} + j\omega C_2} = -\frac{g_m}{\dfrac{1}{r_h} + j\omega C_2} = \frac{A_m}{1 + j\omega r_h C_2} \tag{6-15}$$

The subscript in A_h indicates that Eq. (6-15) holds at frequencies above the mid-band range. It can be seen from Eq. (6-15) that as the frequency is reduced at frequencies above the mid-band range, the voltage amplification approaches the mid-band value $g_m r_h$.

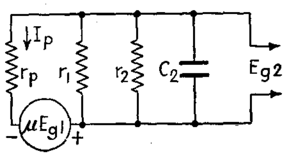


FIG. 6-11.—Simplified equivalent plate circuit for one stage of a resistance-coupled amplifier at high frequencies.

In the theoretical determination of the frequency response of amplifiers and in the design of amplifiers to meet specified frequency-response requirements, it is convenient to know the complex ratio of the amplification at any frequency to the mid-band amplification. Equation (6-15) may be written in the form

$$\frac{A_h}{A_m} = \frac{1}{1 + j\omega r_h C_2} = \frac{1}{\sqrt{1 + (\omega r_h C_2)^2}} (\cos \gamma - j \sin \gamma)$$
$$= \frac{1}{\sqrt{1 + (\omega r_h C_2)^2}} \underline{/\gamma} \tag{6-16}$$

where

$$\gamma = -\tan^{-1} \omega r_h C_2 \tag{6-17}$$

γ is the angle of the phase shift of the output voltage relative to the mid-band phase of the output voltage. (In the mid-band frequency range, the output voltage is 180 degrees out of phase with the input voltage.) The magnitude of the ratio of the high-frequency amplification to the mid-band amplification is seen from Eq. (6-16) to be

$$\left|\frac{A_h}{A_m}\right| = \frac{1}{\sqrt{1 + (\omega r_h C_2)^2}} = \cos \gamma \tag{6-18}$$

$|A_h/A_m|$ can be evaluated most readily at a given value of $\omega r_h C_2$ by finding the cosine of the angle whose tangent is $\omega r_h C_2$. Curves derived from Eqs. (6-17) and (6-18) are shown in Fig. 6-12. These curves give the relative amplification and the relative phase shift at high frequencies

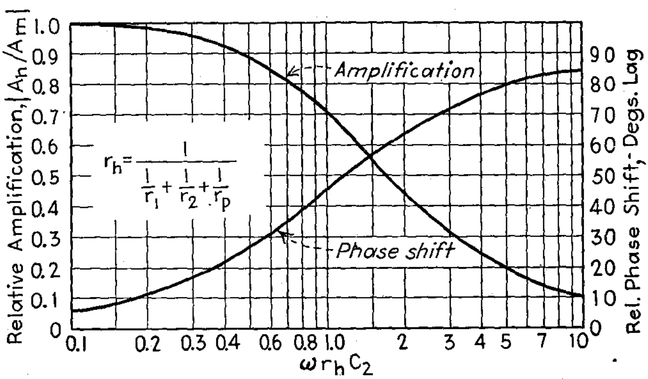


FIG. 6-12.—Generic curves of relative amplification and relative phase shift of a resistance-capacitance-coupled stage at high frequencies.

with respect to the mid-band values, as a function of frequency and circuit constants.

The approximate equivalent circuit of the first stage of the amplifier of Fig. 6-9 at frequencies below the mid-band range is shown in Fig. 6-13.

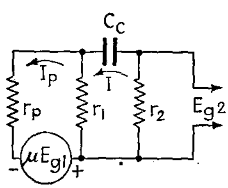


FIG. 6-13.—Simplified equivalent plate circuit for one stage of a resistance-coupled amplifier at low frequencies.

Summation of voltages in this circuit yields the following equations:

$$I_p(r_p + r_1) - Ir_1 = \mu E_{g1} \tag{6-19}$$

$$I\left(r_1 + r_2 - \frac{j}{\omega C_c}\right) = I_p r_1 \tag{6-20}$$

The output voltage is $E_{g2} = -Ir_2$ and the voltage amplification is

$$A_l = \frac{E_{g2}}{E_{g1}} = \frac{Ir_2}{E_{g1}} \tag{6-21}$$

Solution of the simultaneous equations (6-19), (6-20), and (6-21) gives the voltage amplification at low frequencies:

$$A_l = -\frac{\mu r_1 r_2}{r_1 r_2 + r_1 r_p + r_2 r_p - \frac{j(r_p + r_1)}{\omega C_c}}$$

$$= -\frac{g_m r_h}{1 - \frac{j(r_p + r_1)}{(r_1 r_2 + r_1 r_p + r_2 r_p)\omega C_c}} = -\frac{A_m}{1 - \frac{j}{\omega r_l C_c}} \tag{6-22}$$

where

$$r_l = \frac{r_1 r_2 + r_1 r_p + r_2 r_p}{r_p + r_1} = r_2 + \frac{1}{1/r_1 + 1/r_p} \tag{6-23}$$

The subscript in A_l indicates that Eq. (6-22) holds at frequencies below the mid-band range. Equation (6-22) shows that the low-frequency amplification approaches the mid-band value $g_m r_h$ as the frequency is increased at frequencies below the mid-band range.

Equation (6-22) may be written in the form

$$\frac{A_l}{A_m} = \frac{1}{\sqrt{1 + 1/(\omega r_l C_c)^2}} \underline{/\alpha} \tag{6-24}$$

where

$$\alpha = \tan^{-1} \frac{1}{\omega r_l C_c} \tag{6-25}$$

α is the phase angle of the output voltage relative to the mid-band phase of the output voltage. The magnitude of the ratio of the low-frequency amplification to the mid-band amplification is

$$\left|\frac{A_l}{A_m}\right| = \frac{1}{\sqrt{1 + 1/(\omega r_l C_c)^2}} = \cos \alpha \tag{6-26}$$

Curves derived from Eqs. (6-25) and (6-26) are shown in Fig. 6-14. These curves give the relative amplification and the relative phase shift

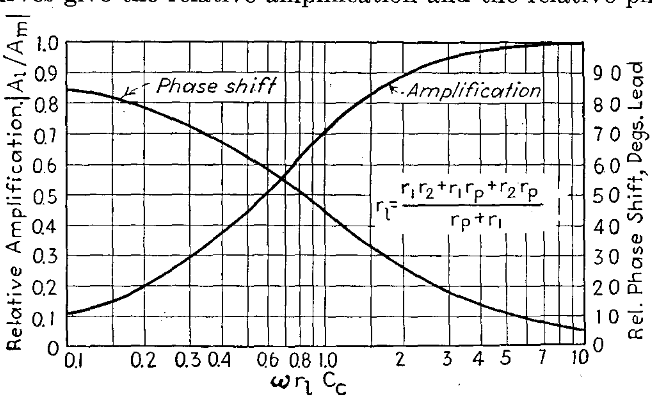


FIG. 6-14.—Generic curve of relative amplification and relative phase shift of a resistance-capacitance–coupled stage at low frequencies.

at low frequencies with respect to the mid-band values, as a function of frequency and circuit constants.

The manner in which the voltage amplification and phase shift of a resistance-capacitance–coupled amplifier vary with the circuit constants and with frequency is completely specified by Eqs. (6-12), (6-17), (6-18), (6-25), and (6-26) (subject to the limitation that the effective input conductance of the following tube is negligible) and by the curves of Figs.

6-12 and 6-14, which are derived from these equations. The method of using these equations and curves can be best shown by computing the voltage amplification of one stage of a resistance-capacitance-coupled amplifier in the mid-band range and at a frequency above and one below the mid-band range. The amplifier will be assumed to have the following circuit and tube constants:

$$\begin{aligned} \mu &= 100 & r_1 &= 200{,}000 \text{ ohms} \\ r_p &= 100{,}000 \text{ ohms} & r_2 &= 500{,}000 \text{ ohms} \\ C_2 &= 200\ \mu\mu\text{f} & C_c &= 0.01\ \mu\text{f} \end{aligned}$$

r_h, as determined from Eq. (6-13), is 58,800 ohms, and g_m is

$$\frac{100}{100{,}000} = 10^{-3} \text{ mho.}$$

The mid-band voltage amplification is found by Eq. (6-12) to be −58.8. At a frequency of 10,000 cps, $\omega r_h C_2$ is

$$2\pi \times 10^4 \times 58{,}800 \times 2 \times 10^{-10} = 0.738.$$

Figure 6-12 shows that, at this value of $\omega r_h C_2$, $|A_h/A_m|$ is approximately 0.8. Hence $|A_h|$ is $0.8 \times 58.8 = 47$ at 10,000 cps. By means of Eq.

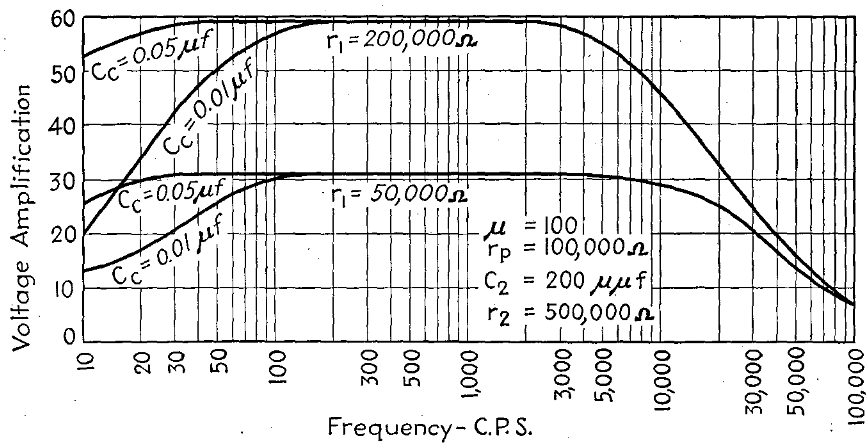


FIG. 6-15.—Theoretical frequency-response characteristics for a typical resistance-coupled amplifier.

(6-23), r_l is found to be 566,000 ohms. At a frequency of 50 cps, $\omega r_l C_c$ is $2\pi \times 50 \times 5.66 \times 10^5 \times 10^{-8} = 1.78$. Figure 6-14 shows that, at this value of $\omega r_l C_c$, $|A_l/A_m|$ is approximately 0.86. Therefore $|A_l|$ is $0.86 \times 58.8 = 50.6$. Theoretical frequency-response curves for the tube factors used in this illustrative problem and for four combinations of circuit constants, including those used in the problem, are shown in Fig. 6-15.

6-6. Effect of Circuit Constants upon Resistance-capacitance-coupled Amplifier Response.—Equation (6-12) may be put into the form

$$A_m = -\frac{\mu}{1 + r_p(1/r_1 + 1/r_2)} \tag{6-27}$$

which shows that the maximum amplification, the mid-band value, increases with increase of amplification factor μ, coupling resistance r_1, and grid-leak resistance r_2, and decreases with increase of plate resistance r_p. Examination of Fig. 6-14 and Eq. (6-23) shows that the mid-band range, throughout which the amplification is independent of frequency, can be extended at the low-frequency end by increasing r_2 and C_c and, less effectively, by increasing r_1 or r_p. Examination of Fig. 6-12 and Eq. (6-13), on the other hand, shows that the mid-band range is extended in the high-frequency direction by *decreasing* r_1, r_2, r_p, or C_2. For convenient reference, these facts are shown in tabular form in Table 6-I, which lists effects produced by increasing the various tube factors and circuit constants. Improvement of characteristics, consisting of widening of the range of uniform amplification or of increase of maximum amplification, is indicated by a plus sign in this table; a narrowing of the range of constant amplification or a reduction of maximum amplification is indicated by a minus sign; 0 signifies that variation of a given quantity has no effect upon the amplification in a particular portion of the frequency band. The lowering of the upper limit of uniform amplification listed in the bottom row of the table results from the increase of C_2, shown by Eq. (6-3), and so is actually included in the effect tabulated in the preceding row.

TABLE 6-I.—EFFECT OF TUBE FACTORS AND CIRCUIT CONSTANTS UPON RESISTANCE-CAPACITANCE–COUPLED AMPLIFIER RESPONSE

Increase of	Lower frequency limit of mid-band amplification	Mid-band amplification, A_m	Upper frequency limit of mid-band amplification
μ	0	+	0
r_p	+	−	−
C_c	+	0	0
r_1	+	+	−
r_2	+	+	−
C_2	0	0	−
A of following stage	0	0	−

The improvement of low-frequency response resulting from increase of C_c, and the effect of decrease of r_1 upon the upper limit of uniform amplification and upon amplification throughout the whole frequency range, are clearly shown by Fig. 6-15.

The value of C_2 for any stage of an amplifier can be readily determined experimentally by the aid of Fig. 6-12 [or Eq. (6-18)], the tube of the following stage being used as a vacuum-tube voltmeter to indicate the output voltage[1] (see page 597). The voltmeter tube is first biased to cutoff. The input to the stage under test is then applied and adjusted to give a convenient small value of plate current of the voltmeter tube at some frequency within the mid-band range. The frequency is then raised until the input voltage must be increased in some convenient ratio at constant output voltage, as indicated by the current of the voltmeter tube. The value of $\omega r_h C_2$ corresponding to the voltage ratio may be read from Fig. 6-12; and, since the frequency is known and r_h may be computed, C_2 may be found. A convenient input voltage ratio is 1.41, for which $\omega r_h C_2$ is unity. A separate vacuum-tube voltmeter may be used in place of the following tube if correction is made for the voltmeter capacitance and conductance.

6-7. Practical Considerations in the Design of Resistance-capacitance–coupled Audio Amplifiers.[2]—A study of Table 6-I reveals that the design of a practical multistage resistance-capacitance–coupled amplifier must involve the compromise between high amplification and uniform amplification. Increasing the amplification by increasing the resistance of r_1 or r_2 results in a lowering of the high-frequency limit of uniform amplification. Furthermore, in a multistage amplifier, increase of the amplification of each stage affects the high-frequency response of the preceding stage adversely because of increase of effective input capacitance. For these reasons it is usually better to increase the amplification by increasing the number of stages rather than by raising the amplification of each stage. The adverse effect of reduction of r_2 upon the response at low frequency can be compensated by an increase of C_c. A practical limit to the maximum size of C_c may exist because of increase of leakage with condenser size, which tends to raise the direct grid voltage of the following tube; and because of the internal inductance of certain types of condensers, which may resonate with the capacitance at a frequency within the mid-band or high-frequency range. The effect of condenser inductance upon the frequency response may be prevented by shunting the large condenser with a small one.

Another point that must be taken into consideration in the choice of the coupling resistor of a resistance-coupled amplifier is the effect of the large direct voltage drop in the coupling resistor. This is even more objectionable than it is in a single-stage amplifier or in a direct-coupled amplifier because the a-c resistance of the coupling circuit is less than the d-c resistance, as explained on page 101. It can be seen from Fig. 6-16

[1] SEELEY, S. W., and KIMBALL, C. N., *RCA Rev.*, **2**, 171 (1937).

[2] See also Sec. 6-38.

that too large a value of plate coupling resistance puts the operating point O in such a location that comparatively small grid excitation extends the path of operation into a region of high amplitude distortion and may even cause the plate current to be cut off during the negative peaks of grid excitation voltage. The operating point O', obtained with a smaller coupling resistance r_1 or higher B-supply voltage, allows the application of higher exciting voltage without overloading.

It is ordinarily not practicable to use plate coupling resistance in excess of either 500,000 ohms or five times the plate resistance. Necessity for uniform response at high frequency may require the use of considerably lower values. The allowable plate supply voltage is limited because of danger of damage to the tube, particularly during the warming-up period of the cathode, before the flow of plate current reduces the plate voltage below the supply voltage. The maximum operating plate voltage specified by tube manufacturers is usually approximately half the value that the tube can safely stand. With pentodes, the distortion resulting from the low operating plate voltage in resistance-coupled amplifiers can be reduced by lowering the screen voltage. This lowers the plate current and thus raises the plate voltage. As a rule, the screen voltage should be approximately one-fifth of the plate supply voltage.

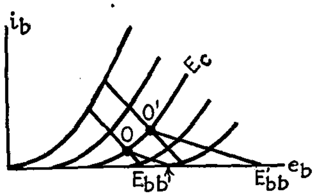


FIG. 6-16.—Plate diagram for resistance-capacitance-coupled stage showing how low plate supply voltage results in high nonlinear distortion.

There is also usually a limit to the size of the grid-leak resistance that can be used. With oxide-coated cathodes it is practically impossible to prevent small amounts of barium or strontium from being deposited on the control grid during manufacture or use of the tube. The temperature of the control grid may become sufficiently high to allow some emission to take place. The emitted electrons are drawn to the cathode and plate, both of which are at higher potential than the grid. The resulting current through the grid resistor is in the direction to make the grid voltage less negative. This causes the plate current to increase and thus raises the plate temperature and, as the result of heat radiated from the plate, that of the grid. Furthermore, if the grid voltage rises sufficiently so that the grid is positive during a portion of the cycle, electrons will be drawn from the cathode to the grid. The impact of these electrons upon the grid also contributes to the rise of grid temperature and may at the same time cause secondary emission from the grid. If the number of electrons leaving the grid as the result of primary and secondary emission exceeds the number entering the grid, the grid potential rises farther. The action tends to become cumulative, particularly if secondary emission from the grid is large, and the plate and grid temperatures may become so high as to result in damage to the tube. The action is complicated by ionization of residual gas in the tube, which allows positive-ion current to flow when the grid is negative and increases electron current to the grid when the grid is positive. The direction of the positive-ion current is such as to cause the grid potential to rise. Positive-ion current may, therefore, contribute to the initial rise of grid voltage. Cumulative rise of grid voltage and plate current may be prevented by making the resistance of the grid resistor low enough so that the initial rise in grid potential caused by grid emission and positive-ion current is small. The maximum grid-circuit resistance that can be safely used with most tubes having oxide-coated cathodes is $\frac{1}{2}$ to 1 megohm. For tubes in which the operating plate current is large, such as those designed to deliver large power output, the allowable grid-circuit resistance may be considerably lower, especially if fixed bias is used. Self-bias tends to prevent the cumulative increase of plate current and thus allows the use of higher resistance than is safe with fixed bias. The maximum allowable resistance is usually specified in the tube-operating data provided by the tube manufacturer.

The transient response is often an important consideration in the design of a resistance-capacitance–coupled amplifier, particularly when it is necessary to amplify voltage pulses without distortion. If a very large coupling condenser is used in order to make amplification possible at very low frequencies, the time taken for the condenser to charge and discharge may be so great that the duration of the transient set up by the sudden application of steady excitation or of a voltage pulse may be excessive. Too large a value of the product C_cr_2 also results in another undesirable effect. Any temporary disturbance of high amplitude may cause one or more of the grids to swing positive, causing a flow of electrons to these grids that leaves the sides of the condensers adjacent to the grids negatively charged. This makes the negative grid voltage greater than the operating value and may even bias one or more grids beyond cutoff and thus reduce the amplification to zero. If the product C_cr_2 is large, the time taken for the charge to leak off through the grid resistors may be so long that the amplifier is inoperative for an appreciable time.

6-8. Design Procedure for Resistance-capacitance–coupled Voltage Amplifiers.—Since circuit constants suitable for use with different types of tubes in resistance-capacitance–coupled amplifiers are specified by tube manufacturers,[1] it is seldom necessary to design special circuits. An understanding of the correct design procedure is, however, valuable.

The high-frequency amplification of a given stage depends upon the effective input capacitance C_2 of the following stage, which depends in turn upon the voltage amplification of the latter. For this reason it is

[1] See, for instance, the RCA Tube Manual.

most convenient to design the last stage of the amplifier first and to proceed stage by stage toward the input stage. Although the reduction of amplification that can be tolerated at the limits of the desired frequency band depends upon the purpose for which the amplifier is to be used, the design can usually be considered to be good if the ratios $|A_h/A_m|$ and $|A_l/A_m|$ for the entire amplifier do not fall below 0.9. Since the values of these ratios for the entire amplifier are equal to the products of the corresponding ratios for the individual stages, the minimum allowable ratios for the individual stages increase with the number of stages.

The following is a satisfactory procedure in the design of a multistage voltage amplifier:

1. Choose a tube for each stage whose amplification factor exceeds the desired voltage amplification of the stage. Because the grid swing of the last stage of a multistage voltage amplifier may be large, it is often good practice to use in the last stage a tube requiring a higher grid bias than that of the preceding stages, *i.e.*, one having a relatively low amplification factor. (The allowable grid swing is usually about $\frac{1}{2}$ to $\frac{3}{4}$ volt less than the grid bias.)
2. Choose the grid coupling resistance r_2 of each stage so that it does not exceed the maximum value specified by the manufacturer for the following tube.
3. Choose the plate coupling resistances r_1. These should not as a rule exceed in value either the grid coupling resistances, five times the plate resistances, or 500,000 ohms.
4. By means of Eq. (6-12), determine the voltage amplification of each stage in the mid-band range. For tubes with high plate resistance, this approximates the product of g_m and the resistance of r_1 and r_2 in parallel. The over-all amplification in the mid-band range should then be found and compared with the desired over-all amplification.
5. By means of Fig. 6-12 or Eq. (6-18), find the ratio $|A_h/A_m|$ for the final stage at the upper limit of the desired frequency range. The effective capacitance across the output of the last stage (including the plate-cathode capacitance) should be used for C_2. If the $|A_h/A_m|$ is too low at this frequency, the output load resistance should be reduced and the amplification recomputed in the mid-band range and at the upper limit of the desired frequency range.
6. If capacitance coupling is used in the output, the coupling condenser C_c should be chosen to give the required amplification at the low-frequency limit of the desired frequency range, as determined from Fig. 6-14 or Eq. (6-26).
7. Using the amplification of the final stage in Eq. (6-3), find the value of C_2 for the next-to-the-last stage. Although C_2 may be found at the upper limit of the frequency band, the design will be conservative if the mid-band amplification of the last stage is used in evaluating C_2. The slightly larger value of C_2 found in this manner will correct in part for the failure to take into account distributed circuit capacitances.
8. Apply steps (5), (6), and (7) to the next-to-the-last stage and, in succession, to the preceding stages, using in each stage the value of C_2 found from the mid-band amplification and interelectrode capacitances of the stage that follows it.

It should be noted that the values of g_m, r_p, and μ that are used in computing the amplification are those at the operating point. They can be determined most readily graphically. The operating point is found from the intersection of the static operating line with the static plate characteristic corresponding to the grid bias. The static load line must, of course, correspond to a resistance equal to the sum of all d-c resistances between the plate supply and the plate. When the amplifier is being designed on the basis of fixed supply voltage, the voltage needed for bias must be taken into account. As the required value of bias is not known until the operating point has been chosen, a tentative estimate must be made of the manner in which the total voltage is divided between the plate supply voltage and the grid bias. If the required bias turns out to differ materially from the estimated value, a second, more accurate, estimate may be made.

As tetrode and pentode plate characteristics are usually published for only one screen voltage, it may be necessary to make experimental determinations of characteristics at other screen voltages. With pentodes, the procedure can be simplified considerably by making use of the fact that, since secondary emission effects are negligible, the potential distribution between the cathode and the plate and the distribution of space current between the various electrodes are not altered when all electrode voltages are changed in the same proportion. Plate characteristics obtained at zero suppressor voltage and at one value of screen voltage can be used at another value of screen voltage if the plate and control-grid voltage scales are changed in the same ratio as the two values of screen voltage, and the proper change is made in the current scale.[1] The factor by which the current scale must be multiplied can be determined from a curve of plate current *vs.* plate voltage corresponding to zero control-grid and suppressor voltage and to a fixed ratio of plate voltage to screen voltage, say 2 to 1. The conversion factor for the current scale is equal to the ratio of the plate currents at points on this curve corresponding to the original and the new screen voltages.

When grid current is allowed to flow in a resistance-coupled amplifier, a part of the current that would otherwise flow through the grid coupling resistor flows through the tube during the portion of the cycle in which grid current flows. The resulting reduction in IR drop across the grid resistor tends to flatten the positive crests of alternating grid voltage and thus causes nonlinear distortion (see Fig. 5-2).

Ordinarily the grid swing of a voltage amplifier is so small that the nonlinear distortion is well within the allowable value. When the grid swing is large, however, it may be advisable to check the final stage for harmonic content. By application of the graphical methods outlined

[1] *RCA Laboratory Series*, No. UL-7.

in Chap. 4 it is possible to determine the grid bias that will give the greatest allowable grid swing without exceeding the allowable distortion. In making such a determination and in finding the maximum crest output voltage graphically, it must be borne in mind that grid current may start flowing when the grid is $\frac{1}{2}$ to $\frac{3}{4}$ volt negative. Grid current therefore limits the allowable grid swing to values that are smaller than the grid bias. This is of particular importance in amplifiers using high-mu tubes, for which the grid bias is small. The maximum crest output voltage may be appreciably less than AE_c (see Probs. 6-5 and 6-6).

6-9. Effect of Plate Supply Filter and of Self-biasing Impedance.—In order to reduce the operating plate voltage of one or more stages of an amplifier or in order to prevent undesirable effects resulting from the interaction of two or more stages, it is sometimes necessary to use a filter of the type illustrated by the heavy lines of Fig. 6-47 in series with the plate coupling resistor. This will be discussed in Sec. 6-38. Increase of reactance of the filter condenser causes the effective coupling impedance of the stage to rise with decrease of frequency and thus tends to result in an increase of amplification at low frequency. The low-frequency response may also be affected by variation of self-biasing impedance with frequency. The increase of reactance of the condenser shunting the cathode resistor tends to cause the voltage amplification to fall at low frequency as the result of degeneration (see Secs. 6-29 and 6-36). The effect is complicated by the application of the alternating voltage across the biasing impedance not only to the control grid but also to the screen grid. The low ratio of screen-plate transconductance to grid-plate transconductance ordinarily justifies neglecting the effect of the alternating voltage applied to the screen. In an accurate analysis the effects of the filter and biasing impedances must be taken into consideration by making use of a complete equivalent circuit that includes these impedances.

Usually sufficient accuracy is obtained under the assumption that the effects of the filter and the biasing impedance upon the relative amplification and phase shift at low frequency are independent of each other and of the effect of the coupling condenser C_c. The resultant relative amplification is then the product of the relative amplifications indicated by the three equivalent circuits in which the biasing impedance and the filter impedance, the biasing impedance and the coupling condenser reactance, and the filter impedance and the coupling condenser reactance, respectively, are assumed to be zero. The resultant phase shift is the sum of the phase shifts given by these three simplified circuits. By using the parallel combination of r_1 and r_2 for z_2 in Fig. 4-30*b*, the student may show that the relative amplification and phase shift resulting from only the biasing impedance is given by the equation

$$\frac{A_l}{A_m} = \frac{\sqrt{m^4 + (k^2 + 2k + 2)m^2 + (k + 1)^2}}{m^2 + (k + 1)^2} \Big/ \tan^{-1} \frac{km}{m^2 + k + 1} \qquad (6\text{-}28)$$

where

$$k = \frac{R_{cc}(\mu + 1)}{r_p + r_1 r_2/(r_1 + r_2)}$$
$$m = R_{cc}\omega C_{cc}$$

and R_{cc} and C_{cc} are the biasing resistance and capacitance. The student may also find it of interest to derive an expression for the relative amplification at low frequency when only the plate-supply filter is considered, the biasing impedance and coupling condenser reactance being assumed to be zero.

Equation (6-28) makes possible the determination of the size of the bias by-pass condenser that will prevent the amplification from falling below a given value at a given frequency. If the relative amplification at 50 cps, for instance, is to be not less than 0.8, and Fig. 6-14 shows the coupling condenser to reduce the relative amplification to 0.95 at this frequency, then the value of A_l/A_m given by Eq. (6-28) must exceed $0.8/0.95 = 0.843$ at this frequency. The corresponding value of by-pass capacitance may be found from the equation or from curves derived from it at various values of k.

A third factor that may influence the amplification is impedance in the screen circuit, which causes the screen voltage to vary with screen current. This is particularly important in amplifiers using tetrodes, the screen current of which varies greatly with plate voltage. For this reason it is advisable to use with tetrodes a voltage divider of low resistance, rather than a voltage-dropping resistor, in reducing the screen voltage below the plate supply voltage. A voltage-dropping resistor may be used with pentodes. In either case, ample by-pass capacitance should be provided in order to reduce the effective impedance to a low value at the lowest frequency to be amplified.

6-10. Compensated Amplifiers.[1]—Drooping of the frequency-response curve of a resistance-capacitance–coupled amplifier at the lower end of the frequency range can be prevented by shunting a portion of the coupling resistor r_1 with a condenser, as shown in Fig. 6-17*a*. This causes the coupling impedance to rise at low frequency and thus compensate for the loss of gain resulting from the voltage-dividing action of the

[1] LANE, H. M., *Proc. I.R.E.*, **26**, 722 (1932); ROBINSON, G. D., *Proc. I.R.E.*, **21**, 833 (1933); BEARDSALL, J., *Television and Short-wave World* (London), **5**, 95 (1936); WALKER, L. E. Q., *Television and Short-wave World*, **9**, 305 (1936); SEELEY, S. W., and KIMBALL, C. N., *RCA Rev.*, **2**, 171 (1937); NAGY, P., *Television and Short-wave World*, **10**, 160, 220 (1937); FREEMAN, R. L., and SCHANTZ, J. D., *Electronics*, August, 1937, p. 22; EVEREST, F. A., *Electronics*, January, 1938, p. 16, May, 1938, p. 24.

coupling condenser and grid leak. In this manner low-frequency amplification can even be made to exceed the mid-band value. With the circuit of Fig. 6-17*b* the amplification can be made to increase at high frequency relative to the mid-band amplification at the expense of mid-band and low-frequency amplification. A similar effect is obtained by the use of low by-pass capacitance across the cathode self-biasing resistor.

The high-frequency response of a resistance-coupled amplifier can be greatly improved by the use of inductance in series with the plate

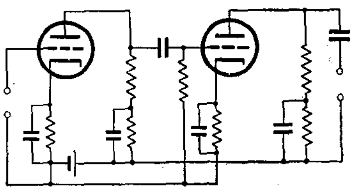


FIG. 6-17*a*.

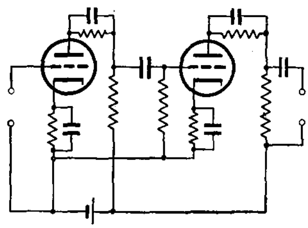


FIG. 6-17*b*.

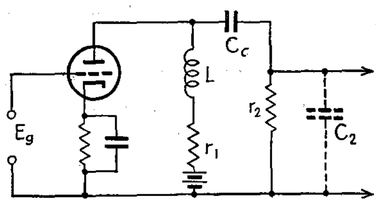


FIG. 6-17*c*.

Fig. 6-17.—*a*. Use of capacitance across a portion of the coupling resistance to improve low-frequency response. *b*. Use of resistance shunted by capacitance in series with the plates to improve high-frequency response. *c*. Use of inductance in series with the plate coupling resistance to improve high-frequency response.

coupling resistance, as shown in Fig. 6-17*c*. The inductance L is chosen so that it resonates with the effective input capacitance of the following tube in the vicinity of the frequency at which the amplification would otherwise start to fall appreciably. In this manner the high-frequency limit of the range of uniform response can be raised considerably and, if desired, the high load impedance resulting from parallel resonance can be made to produce a high-frequency peak in the response curve. It follows from the discussion in Sec. 6-6 that, in order to ensure uniform response up to very high frequency, the ratio of r_2 and r_p to z_1 must be large in amplifiers using high-mu tubes.[1] Under this assumption, the student may derive the following relations for the mid-band and high-

[1] $r_p >> z_1$ because r_p is large in high-mu tubes and z_1 must be relatively small in order to ensure adequate amplification at high frequencies. r_2 must be large in order to ensure adequate amplification at low frequencies without excessively large C_c.

frequency amplification and phase shift of the compensated amplifier of Fig. 6-17c (see Prob. 6-8).

$$A_m = -g_m r_1 \tag{6-29}$$

$$A_h = A_m \frac{\sqrt{1 + \left[(1 - K)\dfrac{\omega}{\omega_u} + K^2\left(\dfrac{\omega}{\omega_u}\right)^3\right]^2}}{\left(\dfrac{\omega}{\omega_u}\right)^2 + \left[K\left(\dfrac{\omega}{\omega_u}\right)^2 - 1\right]^2} \underline{/\gamma} \tag{6-30}$$

where

$$\gamma = \tan^{-1}\left[(1 - K)\frac{\omega}{\omega_u} + K^2\left(\frac{\omega}{\omega_u}\right)^3\right] \tag{6-31}$$

$$\omega_u = \frac{1}{r_1 C_2} \tag{6-32}$$

$$K = \frac{L}{r_1^2 C_2} \tag{6-33}$$

ω_u is defined by Eq. (6-32) as the frequency at which the reactance of the shunting capacitance C_2 is equal to the coupling resistance r_1. Curves of $|A_h/A_m|$ and of γ for several values of K are shown in Fig. 6-18 as a function of ω/ω_u. The great improvement resulting from the use of compensating inductance is evident from a comparison of the curves for $K = 0.41$ or 0.5 with those for $K = 0$, corresponding to an uncompensated amplifier. It can be seen from Eq. (6-32) that the frequency ω_u, beyond which the amplification falls rapidly, can be increased at a given value of C_2 only by decreasing r_1, which also reduces the mid-band amplification. There is a limit, therefore, to the frequency range over which the amplifier can be made to have uniform response.

The compensating circuits of Figs. 6-17a and 6-17c are used in amplifiers for cathode-ray oscilloscopes and in "video" amplifiers for television, in which uniform amplification is essential over a frequency band extending from low audio frequencies well into the radio-frequency range.

6-11. The Cathode-follower Amplifier.—The amplification of a stage of a resistance-coupled amplifier begins to fall off appreciably in the high-frequency range at a frequency at which the reactance of the input capacitance of the following stage becomes less than about three times r_h, the effective output impedance of the given stage, defined by Eq. (6-13). Increasing the amplification of a given stage by increasing r_1 and r_2 therefore has two objectionable effects in an ordinary multistage resistance-coupled amplifier. The high amplification increases the effective input capacitance of the stage in question and thus affects the high-frequency response of the preceding stage adversely; and the high value of the coupling resistances allows the input impedance of the following stage to fall below $3r_h$ at a comparatively low frequency and thus affects the high-frequency response of the given stage adversely. These adverse effects can be prevented by making the effective output

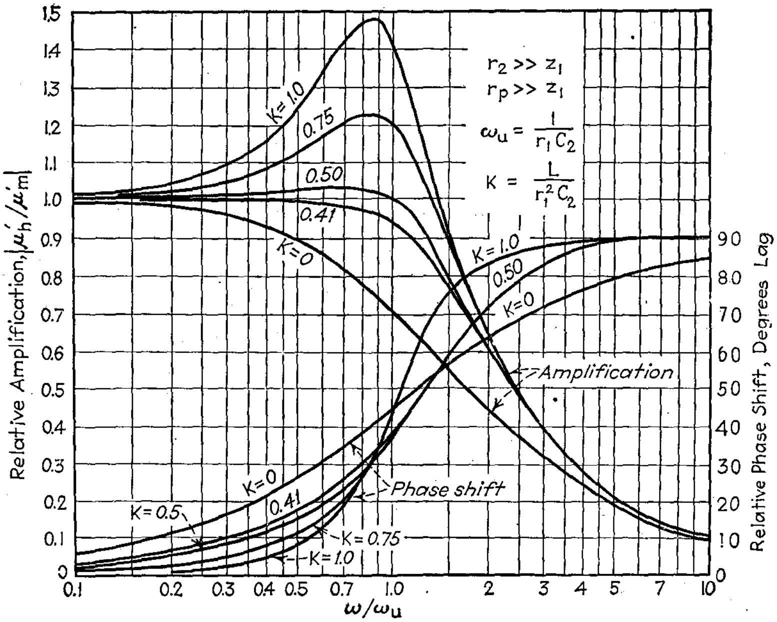


FIG. 6-18.—Generic curves of high-frequency relative amplification and relative phase shift at five values of $K = \frac{L}{r_1{}^2C_2}$ for an inductance-compensated resistance-capacitance-coupled amplifier.

impedance of the preceding stage low by the use of low coupling resistances, and by making the effective input capacitance of the following stage low by making its amplification small [see Eq. (6-3)] by the use of

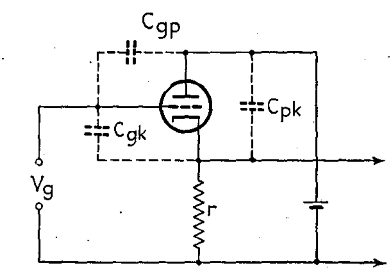


FIG. 6-19a.—Basic circuit of the cathode-follower amplifier.

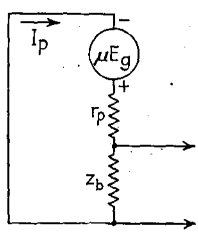


FIG. 6-19b.—Equivalent plate circuit of the cathode-follower amplifier.

low load impedance. Although the low-gain stages that precede and follow the high-gain stage may contribute little or nothing to the over-all amplification, they make possible higher uniform amplification over a given wide frequency band than can be attained in the conventional multistage amplifier. Unfortunately, however, the use of low resistances in the plate circuits of the stages that precede and follow the high-gain stage causes the curvature of the transfer characteristics of these stages to be high and thus results in high nonlinear distortion. This difficulty can be avoided by using in these stages the cathode-follower amplifier shown in basic form in Fig. 6-19*a*.[1] In this amplifier the plate load resistor is adjacent to the cathode and also forms a portion of the grid circuit, so that the output voltage, in addition to the input voltage, is impressed upon the grid circuit. Although the cathode-follower amplifier has a voltage amplification that is less than unity, it has low effective input capacitance, high input impedance, low output terminal impedance, and low nonlinear distortion.

6-11A. Voltage Amplification of the Cathode-follower Amplifier.—It can be seen from Fig. 6-19*a* that the grid-plate capacitance of the tube shunts the input and therefore does not affect the voltage amplification. The plate-cathode capacitance C_{pk} shunts the load (cathode) resistor r. The grid-cathode capacitance C_{gk} shunts the load resistor through the input voltage. Since the voltage across r opposes V_g in the grid circuit, C_{gk} can be replaced by a smaller equivalent capacitance in parallel with r in determining the voltage amplification. In many practical circuits the output voltage is of the order of $0.9V_g$, and this equivalent capacitance is therefore of the order of only $0.1C_{gk}$. For values of r that are usually used, the effect of the interelectrode and circuit capacitances upon the voltage amplification is ordinarily negligible at frequencies below about 1 megacycle. At higher frequencies the effect of the capacitances can be taken into account by shunting r in the equivalent circuit by C_{pk} and by the approximate equivalent value of C_{gk} or by using the exact equivalent circuit, in which C_{gk} is shown in its actual position. In the equivalent plate circuit of Fig. 6-19*b* z_b is assumed to include C_{pk} and the equivalent of C_{gk}, as well as any additional external load, coupling, or wiring impedances that may shunt r.

Solution of the equivalent circuit shows that the voltage amplification is

$$A = -\frac{\mu z_b}{r_p + z_b(1 + \mu)} \tag{6-34}$$

Curves of A *vs.* r_b/r_p derived from Eq. (6-34) for nonreactive load at

[1] PREISMAN, A., *RCA Rev.*, **2,** 430 (1938); LOCKHART, C., *Television and Short-wave World,* **13,** 492 (1940); WILLIAMS, E., *Wireless World,* **47,** 176 (1941); HANNEY, E. A., *Wireless World,* **48,** 164 (1942); NORDICA, C. F., *Radio,* August, 1942, p. 28; LOCKHART, C. E., *Electronic Engineering* (London), **15,** 287 (December, 1942), **16,** 375 (February, 1943), **16,** 21 (June, 1943); RICHTER, W., *Electronics,* November, 1943, p. 112; SHAPIRO, D. L., *Proc. I.R.E.*, **32,** 263 (1944).

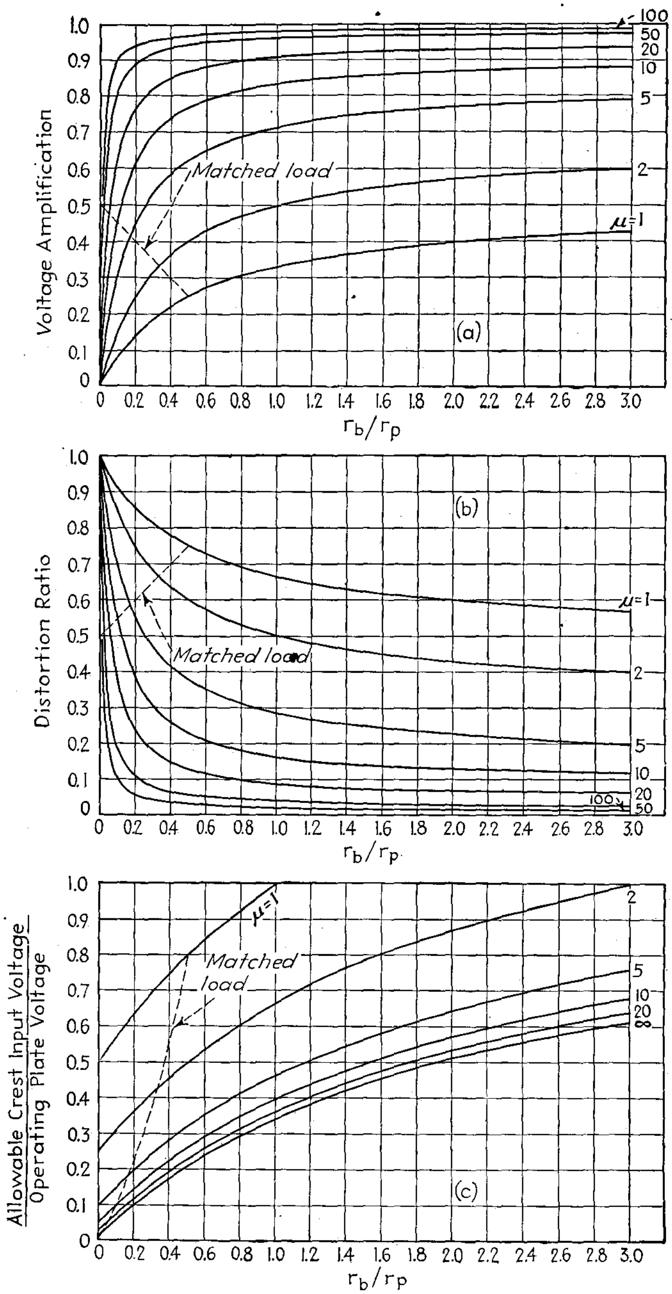


FIG. 6-20.—(*a*) Voltage amplification of a cathode-follower amplifier as a function of r_b/r_p. (*b*) Ratio of nonlinear distortion in a cathode-follower amplifier to that in a conventional amplifier using the same tube and load. (*c*) Ratio of allowable crest input voltage to operating plate voltage of a cathode-follower amplifier.

several values of μ are shown in Fig. 6-20a. For values of μ and r_p found in ordinary tubes, A is of the order of 0.9 or higher when r_b is equal to 10,000 ohms.

Figure 6-21 shows a resistance-capacitance–coupled cathode-follower amplifier stage. C_2 is the sum of the effective input capacitance of the following stage and the effective output capacitance of the given stage (approximately $C_{pk} + 0.1C_{gk}$). The voltage amplification at frequencies in and above the mid-band range can readily be found from Eq. (6-34) if it is noted that the load is the parallel combination of r_1, r_2, and C_2. In the mid-band range the effect of C_2 is negligible and z_b is the resistance of r_1 and r_2 in parallel. The mid-band amplification can therefore be determined from Fig. 6-20a by using $r_1r_2/(r_1 + r_2)$ for r_b. The amplification at frequencies above the mid-band range is

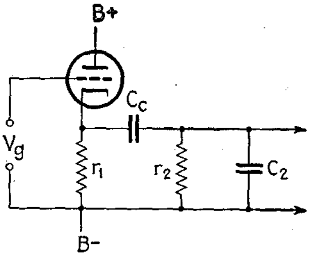


FIG. 6-21.—Resistance-capacitance coupled cathode-follower amplifier stage.

$$A_h = -\frac{g_m}{g_m + 1/r_p + 1/r_1 + 1/r_2 + j\omega C_2} = -\frac{g_m}{1/r_h' + j\omega C_2} \tag{6-34A}$$

where

$$r_h' = r_h/(g_m r_h + 1) \tag{6-34B}$$

and r_h is defined by Eq. (6-13).

Equation (6-34A) differs from Eq. (6-15) only in that r_h' replaces r_h. Hence the relative amplification and the relative phase shift at frequencies above the mid-band range can be determined from Eqs. (6-16) and (6-17) or the curves of Fig. 6-12 if r_h' is used in place of r_h. It follows that, at any value of A_h/A_m, $\omega r_h' C_2$ for a cathode-follower stage is equal to $\omega r_h C_2$ for a conventional stage. Since r_h' is normally much smaller than r_h, the product of the effective capacitance C_2 and the upper limiting frequency of uniform response is much greater for a cathode-follower stage than for a conventional stage. A considerably higher frequency limit may therefore be attained with a cathode-follower stage, even when the stage is followed by a high-gain stage having a larger value of effective input capacitance. If, for example, $C_2 = 250$ $\mu\mu$f, $r_1 = 10{,}000$ ohms, $r_2 = 100{,}000$ ohms, and the tube is a 6J5 operated at normal voltages, the response of a cathode-follower stage is flat to approximately 500 kc. For the same value of C_2, 50 kc represents a good value for the approximate upper limit of mid-band amplification of a conventional stage.

The relative amplification and relative phase shift of the resistance-capacitance–coupled cathode-follower amplifier at frequencies below the mid-band range can be found from Eqs. (6-24) and (6-25) or from Fig. 6-14.

6-11B. Effective Input Capacitance of the Cathode-follower Amplifier. Examination of Fig. 6-19*a* shows that the interelectrode-capacitance current flowing through the source of V_g is made up of a component that flows through C_{gp} and a component that flows through C_{gk}. The former component is equal to $V_g j\omega C_{gp}$. The latter component is caused to flow by the vector sum of V_g and the voltage across r, which is equal in magnitude to AV_g and is opposite in phase to V_g. Hence this component of current is equal to $V_g(1 - A)j\omega C_{gk}$, and the total current through the source is $j\omega V_g[C_{gp} + (1 - A)C_{gk}]$. The effective input capacitance is

$$C_i \doteq C_{gp} + (1 - A)C_{gk} \tag{6-35}$$

When z_b is zero, A is also zero and C_i has its maximum value, $C_{gp} + C_{gk}$. As z_b is increased, A approaches 1 in magnitude. For nonreactive load, A is real, and C_i approaches its minimum value, C_{gp}. In many practical cathode-follower circuits A approximates 0.9 and C_i has the approximate value $C_{gp} + 0.1C_{gk}$. Comparison of Eqs. (6-35) and (4-13) shows that the effective input capacitance of a cathode-follower stage is very much lower than that of an ordinary amplifier stage. The grid-plate capacitance of voltage-amplifier pentodes is of the order of 0.005 $\mu\mu$f and the grid-cathode capacitance of the order of 6 $\mu\mu$f. Hence very low effective input capacitances may be obtained with such tubes.

6-11C. Effective Output Terminal Impedance of the Cathode-follower Amplifier.—The effective output terminal impedance or admittance can be determined by finding the current that flows as the result of application of an alternating voltage V to the output terminals, as shown in Fig. 6-22*a*. Under the assumption that the circuit is linear, *i.e.*, that the equivalent plate circuit is valid, any current flowing in the plate circuit as the result of application of additional exciting voltage to the grid circuit is independent of the current that flows as the result of V. The grid-excitation voltage may therefore be assumed to be zero and the grid circuit to contain only the internal impedance z_c of the source of grid-excitation voltage.

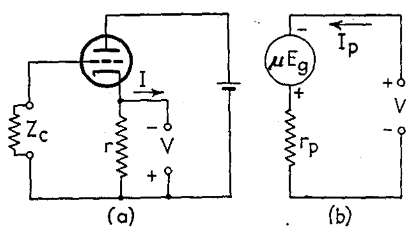


FIG. 6-22.—Circuit for the determination of output terminal impedance of the cathode-follower amplifier. (*a*) Actual circuit; (*b*) equivalent plate circuit.

It can be seen from Fig. 6-22*a* that the output terminals are shunted not only by the plate-cathode path of the tube, but also by r, by C_{pk}, by the series combination of C_{gp} and z_c, and by any external load or coupling circuits connected between the terminals. If the output admittance of the tube itself is known, it may be added to the admittances of any or all of the other paths to find the resultant admittance. The effective admittance of the tube alone is seen from the equivalent circuit of Fig. 6-22*b* to be

$$Y_o = \frac{I_p}{V} = \frac{(V + \mu E_g)}{r_p V} \tag{6-36}$$

The alternating grid voltage E_g is equal to the impressed voltage V less the voltage drop through z_c resulting from the flow of current through the grid-cathode capacitance C_{gk}. Since the capacitance of C_{gk} is of the order of 4 $\mu\mu$f, its reactance is about 40,000 ohms at 1 megacycle. If z_c is nonreactive and less than about 10,000 ohms in magnitude, the alternating grid voltage E_g exceeds $0.97V$ at frequencies below this value. Little error then results from assuming that E_g is equal to V, and Eq. (6-36) becomes

$$Y_o = \frac{(1 + \mu)}{r_p} \tag{6-37}$$

The reduction of E_g resulting from impedance drop in z_c at high frequencies when z_c is large reduces Y_o below the value given by Eq. (6-37). This effect is, however, usually small or entirely absent.

Equation (6-37) shows that the output terminal admittance of the tube is somewhat greater than the transconductance. Actual values of g_m are large enough so that the output terminal impedance Z_o, *i.e.*, the reciprocal of Y_o, may be smaller than 170 ohms with single tubes and does not exceed 1000 ohms under the assumptions made in the derivation of Eq. (6-37). In many applications the admittance of the load and of the capacitances that shunt the output terminals are so small in comparison with the tube admittance Y_o that they have little effect upon the resultant admittance. Equation (6-37) then gives a close approximation to the over-all admittance. The interelectrode capacitances and cathode-to-ground wiring capacitance may begin to have an appreciable effect upon the resultant output admittance at frequencies above a few megacycles.

6-11D. Cathode-follower Amplifier Circuits and Applications.—Figure 6-23 shows two practical cathode-follower circuits in which the grid bias can be adjusted to the correct value when r exceeds the value required to give the correct grid bias. The resistance r' should be large in comparison with the reactance of C at the lowest frequency to be amplified. Circuit (*a*) is superior to circuit (*b*) when high input impedance is of primary importance. The proper bias can also be obtained by connecting the lower input terminal to a tap on a voltage divider across the voltage supply.

An amplifier incorporating cathode-follower stages is shown in Fig. 6-24. The low effective input capacitance of the cathode-follower stage

T_3 prevents falling off of amplification and large shift in phase of the output of the preceding high-gain stage T_2 at high frequencies. The low effective output impedance of the cathode-follower stage T_3 prevents the input capacitance of the following high-gain stage T_4 from causing the amplification to fall off at high frequencies. The combined use of the cathode-follower stages and inductance in the plate circuits of the high-gain stages makes possible uniform amplification throughout a frequency

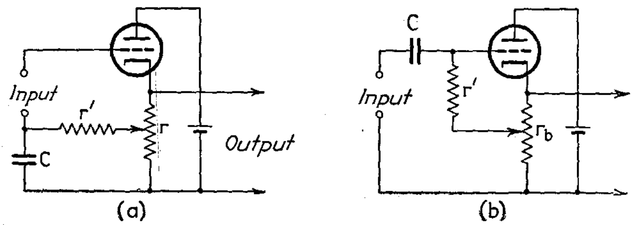


FIG. 6-23.—Cathode-follower amplifiers with bias adjustment.

range extending from low audio frequency well into the radio-frequency range.

The use of the cathode-follower stage T_1 in the amplifier of Fig. 6-24 illustrates another important application of the cathode-follower amplifier. If the input were connected directly to the grid of the high-gain stage T_2, the high effective input capacitance of this stage might react unfavorably upon the source of the input voltage at high frequencies. Although the input stage T_1 actually reduces the over-all amplification slightly, its high input impedance is a distinct advantage in many appli-

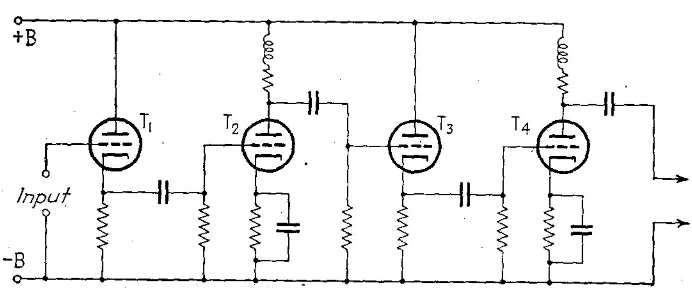


FIG. 6-24.—Four-stage amplifier incorporating cathode-follower input and coupling stages.

cations, as in cathode-ray oscilloscopes (see Sec. 15-21), which must meet the requirement that negligible current be drawn from the source of input voltage. A second advantage of the cathode-follower input stage may result from its ability to handle large input voltages without overloading. This is so because the alternating grid voltage is equal to the difference between the impressed voltage and the voltage across r, which is almost as great as the impressed voltage. A cathode-follower stage is also used to advantage as the last stage of an amplifier when the output voltage must be made independent of external impedance connected across the output terminals of the amplifier. Since Z_0 may readily be made as low as a few hundred ohms, the output voltage will be practically unaffected by impedances of the order of several thousand ohms or more shunted across the output.

The high input impedance and low output impedance of a cathode-follower stage also make possible its use in place of an impedance-matching transformer in matching a high-impedance line or network to a low-impedance line or network. The input impedance can be made to have any desired value within a wide range by means of impedance shunting the input terminals. The output terminal impedance can be made to have the desired value by proper choice of the tube and by adjusting the transconductance by means of the grid bias. The advantage of the cathode-follower amplifier over an impedance-matching transformer is that it is unaffected by changes of frequency over a very wide range of frequency. It will be shown in Sec. 6-11E that a disadvantage of the cathode-follower amplifier in impedance matching is that relatively little power or current can be developed in the load without excessive amplitude distortion.

The fact that the output voltage of a cathode-follower stage is in phase with the input voltage, rather than in phase opposition[1] as in the ordinary type of resistance-capacitance–coupled stage, is occasionally useful.

6-11E. Design of Cathode-follower Amplifiers.—The cathode-follower amplifier is a form of inverse-feedback amplifier, which will be discussed in detail in Secs. 6-29 to 6-37. It will be shown in Secs. 6-30 to 6-32 that distortion may be low in inverse-feedback amplifiers. The curves of Fig. 6-20*b* show the ratio of the distortion factor of a cathode-follower amplifier to that of an ordinary single-stage amplifier having the same tube and load. The curves show that this ratio is low for high values of amplification factor. It should be borne in mind, however, that high-mu tubes have high plate resistance. Since low load impedance is desirable in order to minimize the effect of interelectrode capacitances at high frequencies, the use of a high-mu tube may result in a low ratio of load impedance to plate resistance, which is favorable to high nonlinear distortion. For this reason it may be desirable in high-frequency circuits to use tubes having only medium values of amplification factor and relatively low plate resistance, such as the 6J5.

In order to avoid excessive distortion as the result of grid-circuit overloading (see Sec. 5-3), the grid swing must be kept less than the bias. The ratio of the crest input voltage to the grid swing is given by the equation

[1] Both voltages are measured relative to the common terminal (lower terminals in Fig. 6-19*a*).

$$\frac{V_{gm}}{E_{gm}} = \frac{(1 + k + \mu k)}{k + 1} \tag{6-38}$$

where $k = r_b/r_p$. The maximum allowable value of E_{gm} is E_c, which is roughly equal to $E_{bo}(k + 1)/\mu(k + 2)$ (see Eq. 7-19). Hence

$$\text{Max. } V_{gm} = \frac{1 + k(\mu + 1)}{k + 1} E_c = \frac{1 + k(\mu + 1)}{\mu(k + 2)} E_{bo} \tag{6-39}$$

Curves of Max. V_{gm}/E_{bo} *vs.* r_b/r_p are shown in Fig. (6-20*c*) for various values of μ. Figure 6-20*c* shows that large values of input voltage can be used if μ is low and r_b/r_p is high. Since, as already pointed out, low r_b is desirable, high r_b/r_p requires low r_p, which is also obtained in low-mu tubes. The 6J5 tube, used with a 10,000-ohm load resistance, can handle a crest input voltage equal to approximately $0.4E_{bo}$. It can be seen from Fig. 6-20*c* that use of a tube with higher μ, such as the 6SF5, gives some decrease in allowable input voltage. The higher mu tube has the disadvantage, moreover, of requiring a much larger load resistance.

It was pointed out in Sec. 6-11B that very low values of effective input capacitance can be obtained in a cathode-follower amplifier by using a voltage-amplifier pentode. The low allowable input voltage of such a tube may, however, offset the advantage of low input capacitance. Equation (6-37) shows that low output impedance is obtained with tubes having high transconductance.

In choosing r, it should be borne in mind that a low value is desirable in order to minimize the effect of tube capacitances and in order to prevent large direct voltage drop in the resistance. A high value is desirable, on the other hand, in order to make possible high excitation without excessive nonlinear distortion and in order to make the amplification nearly unity. The best value in a given amplifier will depend upon which of these factors is the most important in the application for which the amplifier is designed. A value of approximately 10,000 ohms is satisfactory in many applications, and there is usually little or no advantage in the use of larger values.

When a cathode-follower amplifier is used in impedance matching, the load resistor r is necessary only if the external load does not provide a path for the direct plate current. The external load is then matched when its admittance is equal to the resultant admittance of the tube and r. Often, however, r is unnecessary, or the effective output impedance of the tube is so small in comparison with r that the external load is nearly matched when its admittance is equal to the tube admittance, as given by Eq. (6-37). Then

$$\frac{z_L}{r_p} = \frac{1}{(1 + \mu)} \tag{6-39A}$$

where z_L is the impedance of the external load. The dashed lines in Figs. 6-20*a*, *b*, and *c* show values that satisfy this relation when the load is nonreactive. It can be seen from Fig. 6-20*c* that, except at very low μ, the ratio of load resistance to plate resistance and the allowable input voltage are both very low. Figure 6-20*b* shows that the reduction of nonlinear distortion resulting from inverse feedback is also small. Nonlinear distortion may, therefore, be excessive unless the input amplitude is much smaller than the value found from Eq. (6-39) or Fig. 6-20*c*.

Although the crest input voltage that may be impressed without grid overloading may be made of the order of one-fourth the plate supply voltage by the use of a high-transconductance, low-mu tube, such as the 2A3, the high nonlinear distortion resulting from low ratio of load impedance to plate resistance may still be large unless the input amplitude is small in comparison with this value. This may present a problem if the source cannot be shunted with a voltage divider. The problem may sometimes be solved by the use of two cathode-follower stages in succession. The first stage uses a tube that can handle a large input voltage, such as the 6J5. The second stage, which uses a high-transconductance tube, is matched to the load, the input voltage to this stage being adjusted to a sufficiently low value by means of an adjustable tap on the load resistor of the first stage. The low allowable grid excitation of a cathode-follower matching stage imposes a limitation upon the power and current that can be delivered to the load. This disadvantage is not avoided by the use of the two-stage circuit, since the first stage merely makes possible the reduction of excitation of the matching stage without the use of a voltage divider across the input source.

6-11F. Filter-type Wide-band Amplifiers.—Another type of wide-band amplifier that has proved very useful in television is the filter-type amplifier, in which complex band-pass filters are used to couple stages. The analysis and design of such amplifiers is beyond the scope of this book.[1]

6-12. Transformer-coupled Audio Amplifier. Preliminary Analysis. The analysis and design of the transformer-coupled audio-frequency amplifier are complicated by the action of distributed winding capacitance, which reduces the amplification at high frequency because it shunts the output, and may increase it over some portion of the frequency range as the result of resonance with leakage inductance. Before analyzing this type of amplifier rigorously, it is instructive to make a qualitative study of its behavior at low, intermediate, and high audio frequencies.

[1] See, for instance, H. A. WHEELER, *Proc. I.R.E.*, **27**, 429 (1939). This reference contains a bibliography of 21 items.

1. *Amplification at Low and Intermediate Frequencies.*—Below a few hundred cycles the reactance of the distributed capacitance of the windings is so high (approximately 10^6 ohms) that its effect may be neglected. In Class A1 amplification the admittance of the following tube is also negligible. The secondary of the transformer may therefore be considered to be unloaded. The equivalent circuit, neglecting the small core losses, is that of Fig. 6-25. z_1, the primary impedance, is equal to $r_1 + j\omega L_1$, where L_1 is the total primary inductance. The magnitude of the voltage induced across the primary is

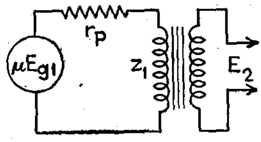


FIG. 6-25.—Simplified equivalent plate circuit of a transformer-coupled audio-frequency stage at low and mid-band frequencies.

$$|E_1| = \frac{\mu E_{g1}\omega L_1}{r_p + z_1} = \mu E_{g1}\frac{\omega L_1}{\sqrt{(r_1 + r_p)^2 + \omega^2 L_1^2}} \tag{6-40}$$

The magnitude of the secondary voltage is

$$|E_2| = n|E_1| = \frac{n\mu E_{g1}}{\sqrt{1 + \left(\frac{r_1 + r_p}{\omega L_1}\right)^2}} \tag{6-41}$$

where n is the ratio of secondary to primary turns. The magnitude of the voltage amplification of the stage is

$$|A| = \frac{|E_2|}{|E_{g1}|} = \frac{n\mu}{\sqrt{1 + \left(\frac{r_1 + r_p}{\omega L_1}\right)^2}} \tag{6-42}$$

As ω is increased $|A|$ increases, approaching the limiting value $n\mu$. The curve relating low-frequency amplification with frequency is of the same form as the upper curve of Fig. 6-4 (the scale of ordinates being multiplied by n). The frequency at which $|A|$ approaches the value $n\mu$ goes down as the ratio $L_1/(r_p + r_1)$ is increased and may be below 100 cps in a well-designed amplifier. Equation (6-42) shows that the loss of amplification at low frequency is the result of falling of the primary reactance with frequency.

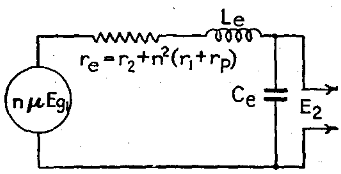


FIG. 6-26.—Simplified equivalent plate circuit of a transformer-coupled audio-frequency stage at high frequency.

2. *Amplification at High Frequency.*—When the frequency becomes so high that the distributed capacitance cannot be neglected, the circuit approximates that of Fig. 6-26, in which C_e represents an equivalent capacitance replacing the primary, secondary, and interwinding capacitances of the transformer and the input capacitance of the following tube;

L_e is the total equivalent leakage inductance; and r_e is the total equivalent winding and plate resistance, all quantities being referred to the secondary. The voltage applied to the grid of the following tube is that across the condenser. The magnitude of the voltage amplification is

$$|A| = \frac{|E_2|}{|E_{g1}|} = \frac{n\mu}{\omega C_e \sqrt{r_e^2 + (\omega L_e - 1/\omega C_e)^2}} \tag{6-43}$$

This has a maximum value at approximately the frequency that makes the reactive term under the radical zero. The voltage amplification at resonance is therefore approximately

$$A_r = \sqrt{\frac{L_e}{C_e}} \cdot \frac{\mu n}{r_e} \tag{6-44}$$

By proper design it is possible to make the factor $\sqrt{L_e/C_e}/r_e$ approximately equal to unity, so that there is little rise in amplification because of resonance, and the amplification is substantially constant up to 10,000 cps or higher.

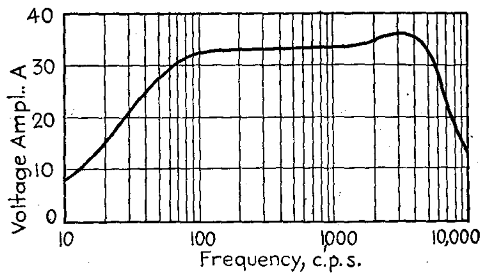


FIG. 6-27.—Frequency-response curve of a transformer-coupled audio-frequency amplifier.

Equation (6-43) shows that A falls above resonance frequency. Both the factor ωC_e, outside of the radical, and the factor ωL_e, under the radical, increase with frequency, so that the amplification eventually falls to zero. Physically, the increase of ωC_e is associated with the shunting effect of the distributed capacitance, and the increase of ωL_e with the loss in voltage in leakage reactance. Figure 6-27 shows the frequency-response curve of a typical transformer-coupled amplifier.

6-13. Transformer-coupled Audio Amplifier. Rigorous Analysis.—Figure 6-28 is an equivalent circuit that closely represents a stage of transformer-coupled amplification. The assumptions are made that the distributed capacitances of the primary and secondary and between primary and secondary may be replaced by lumped capacitances C_1, C_2, and C_{12}. r_{he} is an equivalent resistance to take account of the hysteresis and eddy-current core losses. r_1 and r_2 are the primary and secondary resistances. L_1 and L_2 are the total primary and secondary inductances, and M is the mutual inductance between windings. Since the purpose of the following discussion is primarily to make a qualitative analysis that will point out the principal characteristics of a typical amplifier and how they are influenced by circuit and tube constants, it will be permissible to make certain simplifying assumptions that would not be valid if exact results were required. An excellent complete analysis based upon an equivalent circuit that shows the primary and secondary leakage inductances has been made by Koehler.[1] For a complete anal-

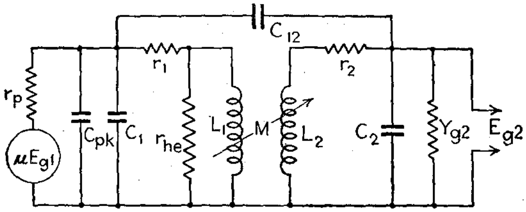


FIG. 6-28.—Complete equivalent plate circuit for a stage of a transformer-coupled audio-frequency amplifier.

ysis and a discussion of design problems, the reader should refer to Koehler's paper.

The primary distributed capacitance C_1 and the plate-to-cathode capacitance C_{pk} may be neglected in an approximate analysis, since the impedance corresponding to $C_1 + C_{pk}$ is much higher than other impedances in parallel with it. C_{12} may be replaced by an equivalent capacitance C_2' in parallel with C_2. (Below resonance primary and secondary terminal voltages are approximately in phase or 180 degrees out of phase, and their amplitude ratio is n. Under this condition the value of C_2' that will cause the same secondary current as C_{12} is $C_{12}(n \pm 1)/n$. In the vicinity of resonance the phase of the secondary terminal voltage shifts, and the magnitude becomes greater, so that C_2' changes.) Y_{g2} may be neglected in comparison with the very much higher admittance of the distributed capacitance, and C_2 and C_2' replaced by a single equivalent capacitance C_e. r_1 and r_p may be replaced by a single resistance r_1'. r_{he} is high enough so that it may be neglected in an approximate analysis. With these simplifications the circuit of Fig. 6-28 may be replaced by the less complicated equivalent circuit of Fig. 6-29.

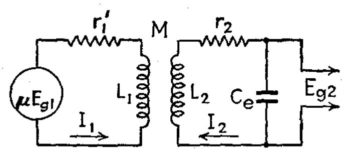


FIG. 6-29.—Simplified equivalent plate circuit for a stage of transformer-coupled audio-frequency amplification.

The following equations are obtained from the network of Fig. 6-29:

$$I_1 z_1 + I_2 j\omega M = \mu E_{g1} \tag{6-45}$$
$$I_2 z_2 + I_1 j\omega M = 0 \tag{6-46}$$

[1] KOEHLER, GLENN, *Proc. I.R.E.*, **16**, 1742 (1928).

$$A = \frac{E_{g2}}{E_{g1}} = \frac{I_2 \dfrac{1}{j\omega C_e}}{E_{g1}} \tag{6-47}$$

where

$$z_1 = r_1' + j\omega L_1 \qquad \text{and} \qquad z_2 = r_2 + j\left(\omega L_2 - \frac{1}{\omega C_e}\right) \tag{6-48}$$

The solution of Eqs. (6-45) to (6-47) yields the relation

$$A = -\frac{M\mu}{C_e z_1(z_2 + M^2\omega^2/z_1)} \tag{6-49}$$

or

$$A = -\frac{M\mu}{C_e z_2(z_1 + M^2\omega^2/z_2)} \tag{6-50}$$

Since the polarity of the secondary voltage depends upon the relative directions of winding, the minus sign in Eqs. (6-49) and (6-50) may be dropped.

The coefficient of coupling k is defined by the equation

$$k = \frac{M}{\sqrt{L_1 L_2}} \tag{6-51}$$

If the leakage inductance is small or proportional to the number of turns,

$$L_1 = \frac{L_2}{n^2} \tag{6-52}$$

Substituting Eqs. (6-48), (6-51), and (6-52) in Eq. (6-49) gives

$$A = \frac{L_1 n \mu k}{C_e(r_1' + j\omega L_1)\left[r_2 + j\left(\omega L_2 - \dfrac{1}{\omega C_e}\right) + \dfrac{k^2 L_1 L_2 \omega^2}{r_1' + j\omega L_1}\right]} \tag{6-53}$$

1. *Response at Low Frequency.*—At frequencies below 200 or 300 cps the term $1/\omega C_e$ is so much larger than the other terms of the second factor of the denominator of Eq. (6-53) that the other terms may be neglected and Eq. (6-53) simplified to

$$A_l = \frac{j\omega L_1}{r_1' + j\omega L_1} nk\mu \tag{6-54}$$

For a typical well-designed audio transformer, k is about 0.998, so that Eq. (6-54) may be further simplified to

$$A_l = \frac{j\omega L_1}{r_1' + j\omega L_1} n\mu \tag{6-55}$$

As the frequency is increased from a low value, A_l increases, approaching the limiting value $n\mu$, and the phase of the output voltage approaches that of the excitation voltage. This limiting value may be attained at a frequency of 100 cps or less and maintained up to several thousand cycles or more. As in the analysis of resistance-capacitance–coupled amplifiers, it will be called the *mid-band amplification* and will be represented by the symbol A_m.

$$A_m = n\mu \tag{6-56}$$

The substitution of Eq. (6-56) in Eq. (6-55) gives

$$\frac{A_l}{A_m} = \frac{j\omega L_1}{r_1' + j\omega L_1} = \frac{\omega L_1}{\sqrt{r_1'^2 + (\omega L_1)^2}} \underline{/\phi} \tag{6-57}$$

where

$$\phi = \tan^{-1} \frac{r_1'}{\omega L_1} \tag{6-58}$$

ϕ is the angle of the phase shift of the output voltage relative to the mid-band phase of the output voltage. (In the mid-band frequency range,

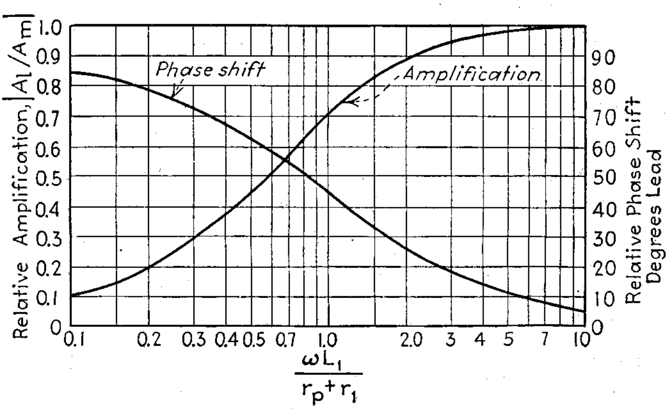


FIG. 6-30.—Generic curves of relative amplification and relative phase shift for transformer-coupled audio-frequency amplifiers at low frequency.

the output voltage is either in phase with or 180 degrees out of phase with the input voltage.) The magnitude of the ratio of the low-frequency amplification to the mid-band amplification is

$$\left|\frac{A_l}{A_m}\right| = \frac{\omega L_1}{\sqrt{r_1'^2 + (\omega L_1)^2}} = \cos \phi \tag{6-59}$$

Figure 6-30 shows curves of $|A_l/A_m|$ and ϕ *vs.* $\omega L_1/r_1'$ derived from Eqs. (6-59) and (6-58).

At first thought, a maximum value of A might be expected at or near the natural resonant frequency of the secondary. That this is not true results from the close coupling between primary and secondary. The value of A at the natural frequency of the secondary can be found by substituting $\omega^2 = 1/L_2C_e$ in Eq. (6-50) and combining it with Eqs. (6-51) and (6-52). This gives

$$A = \frac{n\mu}{\dfrac{n^2C_e r_1' r_2}{kL_2} + k + j\sqrt{\dfrac{C_e}{L_2}}\cdot\dfrac{r_2}{k}} \tag{6-60}$$

The values of resistance, inductance, and distributed capacitance of audio transformers are such that the first and third terms of Eq. (6-60) are negligible in comparison with k, so that $A = n\mu/k$. Since k is practically unity, A is approximately equal to $n\mu$, the mid-band amplification.

2. *Response of Transformer-coupled Amplifier at High Frequency.*—At frequencies above 1000 or 2000 cps, ωL_1 is so much larger than r_1' that $(r_1')^2$ may be neglected in comparison with $\omega^2L_1^2$. With this approximation, Eq. (6-53) may be put into the form

$$A_h = \frac{n\mu k}{j\omega C_e\{r_2 + k^2n^2r_1' + j[\omega L_2(1 - k^2) - 1/\omega C_e]\}} \tag{6-61}$$

The factor $(1 - k^2)$ is called the *leakage factor.* When the leakage inductance is small or proportional to the number of turns,

$$(1 - k^2)L_2 = L_s + n^2L_p = L_e \qquad (6\text{-}62)^1$$

L_s and L_p are the secondary and primary leakage inductances, and L_e is the total equivalent leakage inductance, referred to the secondary. Since k is very nearly unity, the total equivalent resistance of the windings and the tube, referred to the secondary, is

$$r_2 + k^2n^2r_1' \cong r_2 + n^2r_1' = r_2 + n^2(r_1 + r_p) = r_e \tag{6-63}$$

[1] This may be shown by assuming the total primary and secondary inductances to be made up of inductances L_1' and L_2', having perfect coupling (unity coupling coefficient), and leakage inductances L_p and L_s, having zero coupling.

$$\begin{aligned}(1 - k^2)L_2 &= \left(1 - \frac{M^2}{L_1L_2}\right)L_2 = L_2 - \frac{M^2}{L_1}\\ &= L_2 - \frac{L_1'L_2'}{L_1} \qquad \left(\text{since } \frac{M^2}{L_1'L_2'} = 1\right)\\ &= L_2' + L_s - \frac{L_1'L_2'}{L_1' + L_p}\\ &= L_2' + L_s - \left(L_2' - \frac{L_pL_2'}{L_1' + L_p}\right) = L_s + L_p\frac{L_2'}{L_1' + L_p}\end{aligned}$$

If the leakage inductance is small, L_p may be neglected in comparison with L_1'. Furthermore, if the leakage inductance is small or is proportional to the number of turns, $L_2'/L_1' = n^2$, and so

$$(1 - k^2)L_2 = L_s + n^2L_p$$

Hence Eq. (6-61) may be written

$$A_h = \frac{n\mu}{j\omega C_e[r_e + j(\omega L_e - 1/\omega C_e)]} \tag{6-64}$$

Equation (6-64), which is equivalent to Eq. (6-43), has a maximum value near the frequency $\omega_0 \doteq 1/\sqrt{L_e C_e}$, at which the equivalent leakage inductance L_e resonates with the equivalent distributed capacitance C_e. Equation (6-64) may be put into the form

$$\frac{A_h}{A_m} = \frac{A_h}{n\mu} = \frac{1}{j\frac{\omega}{\omega_0}\frac{1}{Q_0} - \left(\frac{\omega}{\omega_0}\right)^2 + 1} \tag{6-65}$$

where

$$Q_0 = \frac{\omega_0 L_e}{r_e} = \frac{1}{r_e}\sqrt{\frac{L_e}{C_e}} \tag{6-66}$$

Equation (6-65) gives the vector ratio of high-frequency amplification to mid-band amplification $n\mu$, as a function of the ratio of the angular frequency to the resonant angular frequency ω_0. Curves of magnitude of amplification and of phase shift relative to the mid-band values as determined from Eq. (6-65) for three values of Q_0 are shown in Fig. 6-31.[1]

Subject to the approximations that have been made in the analysis, the amplification and phase shift of a transformer-coupled audio amplifier are completely specified by Eqs. (6-56), (6-58), (6-59), and (6-65), or by the curves of Figs. 6-30 and 6-31. The numerical values of the quantities that appear in these equations can be readily determined by laboratory measurement.

The response of a self-biased transformer-coupled amplifier is influenced at low frequency by the self-biasing impedance. The effect is complicated by the possibility of resonance between the primary inductance and the capacitance of the by-pass condenser. The student will find it instructive to derive curves of relative amplification and relative phase shift at low frequency by analyzing the equivalent circuit obtained by inserting the parallel combination of R_{cc} and C_{cc} in the equivalent circuit of Fig. 6-25. It should be noted that the alternating voltage appearing across R_{cc} is impressed upon the grid.

Sometimes it is advantageous to apply the plate supply voltage through a choke and to couple the transformer primary through a condenser. This prevents direct current from passing through the primary and thus makes possible the use of smaller transformers. Because the transformer may then be connected directly to the cathode, instead of to the positive side of the B supply, the alternating plate current need not

[1] TERMAN, F. E., *Electronics*, June, 1937, p. 34.

pass through the biasing resistor, and loss of amplification may be avoided. Resonance between the transformer primary inductance and the coupling condenser capacitance may increase the amplification at low frequency. The low effective load impedance at resonance, however, results in relatively high nonlinear distortion (see Sec. 3-25). In order to minimize capacitance from plate to ground, the coupling condenser should be placed in the cathode lead of the transformer primary.

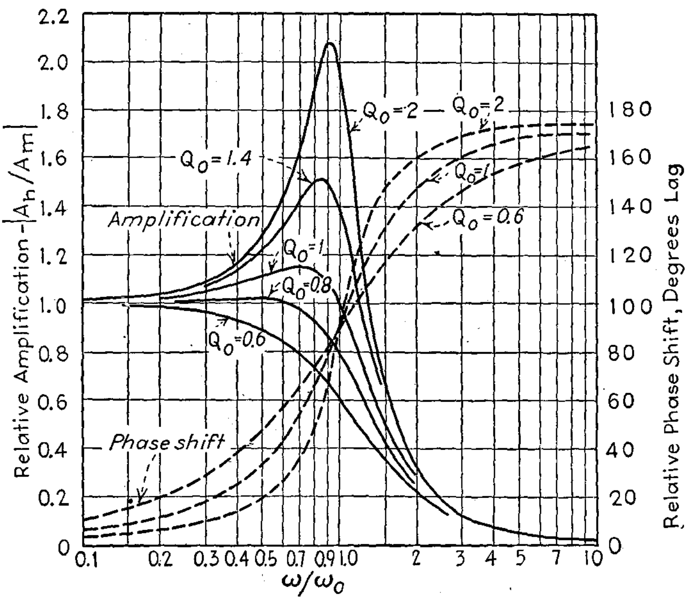


FIG. 6-31.—Generic curves of relative amplification and relative phase shift for transformer-coupled audio-frequency amplifiers at high frequency. $\omega_0 = \dfrac{1}{\sqrt{L_eC_e}}$; $Q_0 = \dfrac{\omega_0 L_e}{r_e}$.

The response of transformer-coupled amplifiers can be altered by shunting the secondary of one or more transformers with resistance or capacitance.

6-14. Transformer Design.—The correct designing of an audio-frequency coupling transformer is not a simple matter, as a compromise must be made between conflicting requirements. Equation (6-59) and the curve of Fig. 6-30 show that for satisfactory response at low audio frequencies the primary inductance must be high. Even with high-permeability cores this requires the use of a large number of turns. A high resonance frequency ω_0, beyond which amplification falls rapidly, calls for low leakage inductance and small distributed capacitance. The leakage inductance and distributed capacitance can be minimized by proper methods of winding and careful choice of shape and arrangement of coils,[1] but with any type of winding the increase of distributed capacitance with number of turns affects the high-frequency response adversely. Since a high transformation ratio requires either a small number of primary turns or a large number of secondary turns, or both, it follows that the increase in voltage amplification which results from the use of a high transformation ratio can be obtained only at the expense of poor response at low or high frequency, or both. In early receivers for broadcast reception, high amplification was obtained by the use of audio transformers with small primaries. As a result, these receivers were noted for their lack of bass tones. Well-designed modern transformers for Class A voltage amplifiers do not ordinarily use transformation ratios higher than 3 to 1. Equation (6-65) and the curves of Fig. 6-31 show that the shape of the resonance peak can be controlled by choice of the ratio

$$\frac{\omega_0 L_e}{n^2(r_p + r_1) + r_2} = \frac{1}{n^2(r_p + r_1) + r_2} \cdot \sqrt{\frac{L_e}{C_e}}.$$

Since this ratio contains r_p, the height and sharpness of the resonance peak are affected by the plate resistance. For this reason and because high plate resistance is unfavorable to good response at both low and high frequencies, it is important to use a-f transformers with the types of tubes for which they are designed, or with tubes having similar characteristics.

6-15. Comparison of Resistance-capacitance and Transformer-coupled Audio Amplifiers.—The fact that uniform amplification in excess of the amplification factor of the tube can be obtained with transformer coupling if the plate resistance is not too high made the use of transformer coupling decidedly advantageous before the development of tubes with high amplification factors. Unfortunately, however, high amplification factor is accompanied by high plate resistance, and so high-mu tubes like the type 6SJ7 pentode cannot be satisfactorily used with transformer coupling. The response is poor at low frequency because of the low ratio of primary reactance to plate resistance; and at high frequency because of the low value of $\frac{1}{r_e}\sqrt{\frac{L_e}{C_e}}$. A voltage amplification of 50 represents about the highest that can be obtained with transformer coupling without sacrifice of response at low or high frequency. With resistance-capacitance coupling, even though the stage amplification is much lower than the amplification factor of the tube, the high amplification factors of tubes like the 6SJ7 make it possible to obtain an amplification of 100 or more per stage. Because likelihood of oscillation makes it difficult to use more than two transformer-coupled stages on a common

[1] KOEHLER, *loc. cit.*

voltage supply, higher amplification at audio frequencies can as a rule be obtained with resistance-capacitance coupling than with transformer coupling. Resistance-capacitance coupling is considerably cheaper than transformer coupling, weighs less, and requires less space. Because of the lower direct voltage drop in the plate circuit, and because voltage induced in the transformer inductance allows the instantaneous plate voltage to exceed the B-supply voltage (see Figs. 4-15 to 4-18), higher voltage output may be obtained with a given plate supply voltage with transformer output than with resistance output. Transformer coupling to the load may be necessary in the output stage when the load impedance is so low as to cause excessive nonlinear distortion without impedance matching (see Sec. 7-21).

6-16. Inductance-capacitance Coupling.—Inductance-capacitance coupling has the advantage over resistance-capacitance coupling of lower direct voltage drop in the plate load. Because there is only one winding, it is possible to obtain somewhat higher inductance than in the primary of a transformer. Although this higher inductance improves the low-frequency response over that of a transformer-coupled amplifier, it is still too low to prevent the amplification from falling off at low frequency with high–amplification-factor tubes. The high-frequency response is also somewhat better than in transformer-coupled amplifiers because of the lower distributed capacitance of the single winding. These advantages are offset by the fact that the voltage is not stepped up as it may be by a transformer. Inductance-capacitance coupling is now seldom used. It may be analyzed by the methods that have been used in this chapter.

6-17. Choice of Tubes for Audio-frequency Voltage Amplifiers. Equations (6-1), (6-12), and (6-56) show that the maximum amplification obtainable with a voltage amplifier is proportional to the amplification factor of the tube. For this reason the amplification factor may in general be considered to be the "figure of merit" of a triode used in an audio-frequency voltage amplifier. Other factors such as plate resistance and plate current are also of importance in choosing the tubes for a voltage amplifier. It was shown in preceding sections that plate resistance reduces the attainable gain of direct- or resistance-capacitance–coupled amplifiers and causes frequency distortion in transformer-coupled amplifiers. High plate current causes high direct voltage drop in the coupling resistor of direct- or resistance-capacitance–coupled amplifiers and thus necessitates the use of higher plate supply voltage. It tends to cause nonlinear and frequency distortion in transformer-coupled amplifiers because of core saturation. It also increases the power consumption in the plate circuit.

Higher voltage amplification is in general attainable with pentodes than with triodes in resistance-coupled amplifiers, but nonlinear distortion may be greater and, because of higher plate resistance, it is more difficult to design for uniform response over a wide band. Since the amplification attainable with pentodes approximates $\frac{r_1 r_2}{r_1 + r_2} g_m$, the transconductance may be considered to be the figure of merit of pentodes in resistance-coupled voltage amplifiers. As pointed out in Sec. 6-15, the high plate resistance of pentodes makes them unsuitable for use in transformer-coupled a-f amplifiers.

6-18. Control of Amplification of Audio-frequency Amplifiers.—The most common method of controlling the amplification of audio-frequency voltage amplifiers is by means of a voltage divider that varies the excitation voltage of one of the tubes. In transformer-coupled amplifiers the voltage divider is shunted across the secondary of the input transformer, as in Fig. 6-32*a*; in resistance-coupled amplifiers it may serve as the grid coupling resistance, as in Fig. 6-32*b*. In the transformer-coupled circuit,

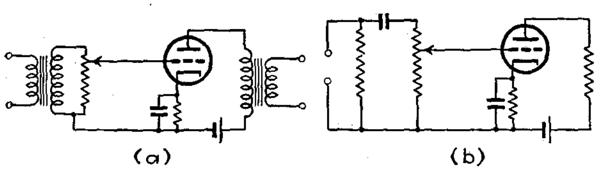


FIG. 6-32.—Methods of varying the amplification of transformer- and resistance-coupled audio-frequency amplifiers.

the resistance of the voltage divider should be high in order to prevent appreciable loading of the transformer. Potentiometers of 500,000 ohms resistance or higher are commonly used. In order to keep the grid swing of all stages as low as possible and thus to minimize nonlinear distortion, the amplification control is usually used in the grid circuit of the first stage. Occasionally, however, the amplitude control is used in a later stage so as to reduce noise caused by the slider of the voltage divider.

Because the voltage amplification of resistance-coupled amplifiers using pentodes or other high-plate-resistance tubes is nearly proportional to g_m, the amplification of such amplifiers can be readily controlled by varying the transconductance by means of the grid bias. In order to make it possible to reduce the transconductance to a low value without danger of having the grid swing beyond cutoff, the signal is applied to a remote-cutoff grid (see Fig. 3-7). Although the gain-control voltage may be applied to the same grid as the signal, less control voltage is required if it is applied to an additional sharp cutoff grid. It may be applied simultaneously to both grids. The 6L7 type of tube is suitable for this purpose. This type of control has the advantage that it may be accomplished by means of a rheostat or potentiometer that is located at a distance from the amplifier and is by-passed by a condenser at the amplifier.

If two tubes controlled in this manner and excited by separate signal voltages use a common plate load resistance, a gradual transition may be made from one signal to the other by raising one control voltage and lowering the other.

6-19. Volume Compression and Expansion.[1]—The wide range of power level of speech creates a problem in transoceanic radio telephony. If the amplification is adjusted so that the loudest sound will fully modulate the transmitter, the fainter sounds may not be distinguishable above static and other interfering noises. This difficulty can be overcome by automatically increasing the amplification of the weaker signals in order to reduce the signal amplitude range and raise the average level. At the receiving end the process must be reversed by automatically raising the amplification with signal amplitude, so that the range of power level of the output is proportional to the power-level range of the original sounds.

A similar problem is encountered in the recording of phonograph records. The range of sound level that can be recorded is much smaller than that which is necessary for the faithful reproduction of orchestral music (for which the power ratio corresponding to the weakest and strongest sounds is very great). If the recorded sound is too faint, needle scratch is objectionable. If it is too loud, on the other hand, there is danger that the needle may cut from one groove to the next. The amplitude range must, therefore, be reduced in the recording amplifiers and should, preferably, be increased by the reproducing amplifier.

These results can be accomplished by means of automatic gain control of audio-frequency amplifiers. The a-f output is applied to a detector, the output of which controls the bias of the amplifier grids. Whether the amplitude range is increased or decreased depends upon whether the direct voltage output of the detector is applied to the grids in such a manner as to increase or to decrease the amplification with increase of amplifier output. A typical "volume expander" circuit is shown in Fig. 6-33. Since the voltage amplification of a resistance-coupled amplifier with small ratio of load to plate resistance is proportional to transconductance [see Eqs. (6-1) and (6-12)], increase of grid bias reduces the amplification. To prevent excessive amplitude distortion when the transconductance is reduced to a low value by high negative bias, the signal is applied to a remote-cutoff grid. Need for large gain-control voltage is avoided by applying the gain-control voltage to an additional

[1] MATHES, R. C., and WRIGHT, S. B., *Elec. Eng.*, **53,** 860 (1934); BALLANTINE, S., *Proc. I.R.E.*, **22,** 612 (1934); SOWERBY, A. L. M., *Wireless World*, **34,** 150 (1934); *Radio Eng.*, November, 1934, p. 7, THOMAS, T. S. E., *Wireless Eng.*, **12,** 493 (1935); SINNETT, C. M., *Electronics*, November, 1935, p. 14; HALLMAN, L. B., JR., *Electronics*, June, 1936, p. 15; RCA Application Note 53.

sharp-cutoff grid. This is made possible by the use of the 6L7 type of tube.

6-20. Tone Control.[1]—Differences in the acoustical properties of different rooms, nonuniform response of loud-speakers, individual preferences, and the fact that the sensitivity of the ear falls off at low and high frequencies often make it desirable or necessary to change the frequency characteristics of audio-frequency amplifiers. Minor changes in the characteristic of an amplifier can be obtained by the use of simple filters. One commonly used filter consists of a condenser in series with a variable resistance, shunted across some point of the amplifier, usually between grid and cathode of one of the tubes. Decrease of resistance cuts down the response at the higher frequencies. Sometimes tone

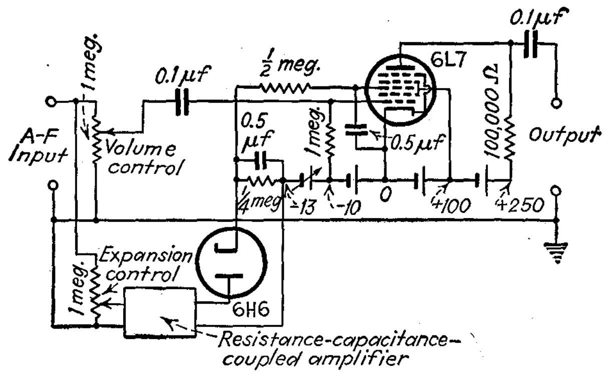


FIG. 6-33.—Circuit diagram of a volume expander.

control is combined with the manual gain control in such a manner that the amplification is reduced more at the middle of the audio range than at the upper and lower ends. This tends to correct for the apparent loss of low and high tones at low sound level caused by the nonuniform response of the ear. Some amplifiers are provided with separate tone controls for the low and high frequencies and, in more elaborate systems, separate amplifiers and loud-speakers may be used to cover two or more portions of the entire a-f range. Considerable flexibility of tone control may be attained in this manner.

6-21. Resistance-coupled Radio-frequency Amplifiers.—The analysis that has been made of resistance-capacitance–coupled amplifiers is valid at radio frequencies if the frequencies are not so high that the time of transit of electrons from cathode to plate must be taken into consideration.[2]

[1] SCROGGIE, M. G., *Wireless Eng.*, **9**, 3 (1932); Hazeltine Service Corp., *Radio Eng.*, June, 1935, p. 7; COLEBROOK, F. M., *Wireless Eng.*, **10**, 4 (1933) (with bibliography of seven items); CALLENDER, M. V., *Proc. I.R.E.*, **20**, 1427 (1932).

[2] LLEWELLYN, F. B., *Proc. I.R.E.*, **21**, 1532 (1933); **23**, 112 (1935).

Tuned Radio-frequency Amplifiers.—Transformer-coupled r-f amplifiers are commonly of two types: doubly tuned and singly tuned. In doubly tuned amplifiers, tuning capacitance is used in parallel with both the primary and the secondary of the transformer, as shown in Fig. 5-9. In singly tuned amplifiers, only the secondary is shunted by tuning capacitance, as shown in Fig. 5-10. Radio-frequency transformers are now made either with air cores or with special low-loss iron cores. The primary and secondary are coupled loosely enough so that the capacitance between primary and secondary is negligible.

6-22. Doubly Tuned Radio-frequency Amplifier.—The equivalent circuit of the doubly tuned amplifier, at frequencies below those at which time of transit of electrons must be considered, is shown in Fig. 6-34*a*.[1]

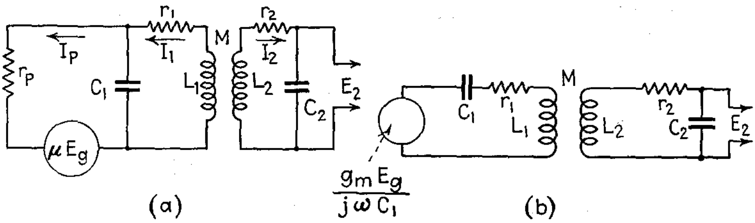


FIG. 6-34.—(*a*) Equivalent plate circuit for doubly tuned r-f amplifier; (*b*) Approximate equivalent circuit applicable to tetrodes and pentodes.

Solution of the network equations for this equivalent circuit gives the following expression for voltage amplification:

$$A = -\frac{\mu M/C_2}{(1 + j\omega C_1 r_p)(z_1 z_2 + \omega^2 M^2) + jz_2/\omega C_1} \tag{6-67}$$

in which $z_1 = r_1 + j\omega L_1 + 1/j\omega C_1$ and $z_2 = r_2 + j\omega L_2 + 1/j\omega C_2$. For reasons that will be explained, screen-grid tubes or pentodes are far superior to triodes in tuned radio-frequency amplifiers. In practice the reactance of the primary condenser in the vicinity of the frequency for which the amplifier is designed is so much smaller than the plate resistance of a tetrode or pentode that little error results if $1/r_p$ is neglected in comparison with $j\omega C_1$ in Eq. (6-67). Equation (6-67) then simplifies to

$$A = \frac{j(g_m M/\omega C_1 C_2)}{z_1 z_2 + M^2\omega^2 + z_2/r_p\omega^2 C_1{}^2} \tag{6-68}$$

The form of Eq. (6-68) is too complicated to yield readily to an analysis to determine the manner in which A varies with frequency. With typical circuit constants and plate resistance, however, the third term in the denominator is sufficiently small in comparison with the sum

[1] r_2 includes an equivalent resistance to take into account the effective input conductance of the following tube, and C_2 includes the effective input capacitance of the following tube.

of the other two terms so that it may be neglected. It is most convenient, therefore, first to neglect this term and later to determine under what conditions and in what manner the result should be modified in order to take its effect into account. Equation (6-68) may then be simplied to

$$A = \frac{j(g_m M/\omega C_1 C_2)}{z_1 z_2 + M^2\omega^2} \tag{6-69}$$

A general indication of the form of the curve of A *vs.* ω can be obtained by considering the special case, often met in practice, in which the primary and secondary inductances, capacitances, and resistances are equal. Then $z_1 = z_2 = r + j(\omega L - 1/\omega C)$, and Eq. (6-69) becomes

$$A = \frac{j\dfrac{g_m M}{\omega C^2}}{\left[j\omega(L + M) + \dfrac{1}{j\omega C} + r\right]\left[j\omega(L - M) + \dfrac{1}{j\omega C} + r\right]} \tag{6-70}$$

The effect of ω in the numerator being neglected, A will have a maximum at approximately the two frequencies that make the reactance zero in the two factors of the denominator of Eq. (6-70). The curve of A *vs.* frequency will, therefore, in general have two peaks, as shown by curve a in Fig. 6-35, whose separation increases with M.

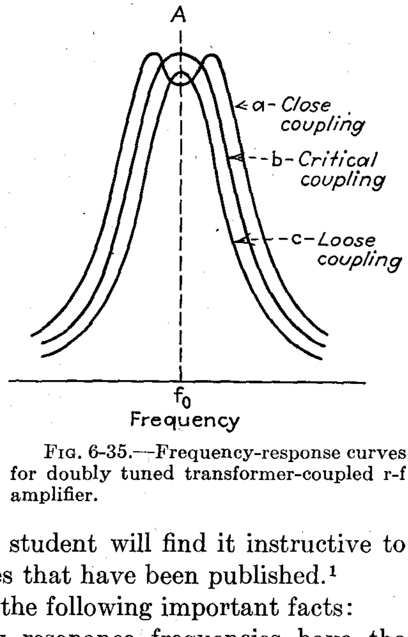


FIG. 6-35.—Frequency-response curves for doubly tuned transformer-coupled r-f amplifier.

The amplification given by the approximate equation (6-69) is the same as that of the equivalent circuit of Fig. 6-34*b*, which is, therefore, an approximate equivalent circuit of the doubly tuned amplifier used with tetrodes and pentodes. Since the complete analysis of this circuit and of Eq. (6-68) is fairly long and involved, it will not be given. The student will find it instructive to refer to some of the excellent analyses that have been published.[1]

The results of such analyses show the following important facts:

1. If the primary and secondary resonance frequencies have the same value f_0, the two frequencies corresponding to the peaks of the response curve have the approximate values

[1] See, for instance, W. L. EVERITT, "Communication Engineering," p. 496, McGraw-Hill Book Company, Inc., New York, 2d ed., 1937; E. A. GUILLEMIN, "Communication Networks," Vol. I, p. 323, John Wiley & Sons, Inc., New York, 1931.

$$f = f_0\left(1 \pm \tfrac{1}{2}\sqrt{k^2 - \frac{r_1 r_2}{\omega_0^2 L_1 L_2}}\right) \tag{6-71}$$

2. When $\omega_0^2 M^2 > r_1 r_2$, the height of the peaks does not change as the value of M is changed by varying the coupling, other parameters being held constant.

3. The ratio h of the height of the peaks to that of the intervening trough is related to the circuit parameters by the equation

$$h \pm \sqrt{h^2 \pm 1} = k\sqrt{\frac{\omega_0^2 L_1 L_2}{r_1 r_2}} = \sqrt{\frac{\omega_0^2 M^2}{r_1 r_2}} \tag{6-72}$$

in which k is the coefficient of coupling, $\sqrt{M^2/L_1L_2}$.

4. h is unity when $\omega_0^2 M^2 = r_1 r_2$, indicating that the height of the peaks is equal to that of the trough. At this *critical coupling*, therefore, the two peaks merge into one.

The peak voltage amplification at critical coupling, which is the same as that corresponding to the two peaks of the response curve above critical coupling, is found from Eq. (6-69) to be

$$A \cong \tfrac{1}{2} g_m \sqrt{\frac{L_1}{r_1 C_1}\frac{L_2}{r_2 C_2}} = \tfrac{1}{2} g_m \sqrt{Q_1 Q_2 \sqrt{\frac{L_1}{C_1}\frac{L_2}{C_2}}} \tag{6-73}$$

where

$$Q_1 = \frac{\omega_0 L_1}{r_1} \qquad \text{and} \qquad Q_2 = \frac{\omega_0 L_2}{r_2}$$

Q_1 and Q_2, which may be shown to be the ratio of the energy stored per cycle to the energy dissipated per cycle in the primary and secondary circuits, respectively, are a measure of the efficiency of the resonant circuits.

5. The height of the single peak falls, and the peak becomes sharper, as the coupling is decreased below the critical value.

Typical curves of A *vs.* ω for coupling greater than, equal to, and less than the critical value are shown in Fig. 6-35.

The importance of the third term in the denominator of Eq. (6-68) can be determined by analyzing a practical case. Laboratory measurements of a typical air-core 450-kc transformer gave the following approximate values

$$L_1 = L_2 = 2 \times 10^{-3} \text{ henry}$$
$$r_1 = r_2 = 150 \text{ ohms at } 450 \text{ kc}$$
$$C_1 = C_2 = 60 \times 10^{-12} \text{ farad}$$
$$w_0 = \frac{1}{\sqrt{L_1 C_1}} = 289 \times 10^4 \text{ rad/sec}$$
$$Q_1 = \frac{\omega_0 L_1}{r_1} = 38$$

For a 6K7 r-f pentode, r_p is approximately 1 megohm. With these values the sum of the first and second terms of the denominator at critical coupling is 45×10^3 (ohms)^2 at the secondary resonance frequency, and the third term approximately 5×10^3 (ohms)^2. In this example comparatively little error results from dropping the third term. Coils of better design, however, may have values of Q of 100 or more, and with different operating voltages, or a different type of tube, the plate resistance may be reduced appreciably. The third term may then be of the same order of magnitude as the other two terms of the denominator, so that it cannot be neglected.

The third term in the denominator of Eq. (6-68) is a minimum at the secondary resonance frequency, which, in transformers with similar primary and secondary constants, corresponds to the frequency of the trough of the resonance curve. This term, therefore, tends to raise the trough of the curve. It also increases the critical coupling above the value $M = \sqrt{r_1 r_2}/w_0$ indicated by the analysis in which this term is neglected.

The doubly tuned amplifier with coupling slightly in excess of the critical value amplifies very nearly uniformly over the range of frequencies lying between the two peaks, and the amplification falls rapidly outside of this range. For this reason this type of amplifier is often called a band-pass amplifier. The response curve can be made to approximate square-top form more nearly by using a closely coupled doubly tuned stage in conjunction with a singly tuned stage or a loosely coupled doubly tuned stage.[1] A square-topped response curve is desirable in amplifiers used for radio reception, which should amplify uniformly at all frequencies contained in the modulated wave and cut off sharply at lower and higher frequencies.

Advantage may also be taken of the change in width of the frequency band of a doubly tuned amplifier with change in coupling. It is evident that for a value of coupling corresponding to curve a of Fig. 6-35 the frequency band over which amplification is essentially constant is considerably wider than for coupling corresponding to curve b. Adjustable coupling thus makes possible the choice of a narrow band for high selectivity in the reception of a weak modulated carrier wave in the presence of a strong carrier wave in an adjacent frequency channel or of a wide band for high quality in the reception of a strong modulated wave. This change in selectivity may be accomplished automatically by the use of vacuum-tube coupling in addition to inductive coupling.[2]

Subject to the assumptions made in its derivation, Eq. (6-68) shows that the amplification of a doubly tuned amplifier is proportional to the transconductance of the tube.

[1] LOH, HO-SHON, *Proc. I.R.E.* **26,** 469 (1938).
[2] MAYER, H. F., *Electronics*, December, 1936, p. 32.

6-23. Singly Tuned Amplifier.—For the sake of analysis, the transformer-coupled circuit with tuned secondary may be considered as a special form of the doubly tuned circuit. It is convenient for this purpose to write Eq. (6-67) in the form

$$A = -\frac{\mu M/C_2}{(1 + j\omega C_1 r_p)[(j\omega L_1 + r_1)z_2 + \omega^2 M^2] + r_p z_2} \tag{6-74}$$

In practice the primary of the transformer is sufficiently small so that both its reactance and resistance are negligible in comparison with r_p. If this is assumed to be true and C_1 is made zero, Eq. (6-74) reduces to

$$A = -\frac{\mu M/C_2}{z_2 r_p + \omega^2 M^2} \tag{6-75}$$

Differentiation of Eq. (6-75) with respect to M at constant ω shows that A has a maximum value at a given frequency when $\omega^2 M^2 = z_2 r_p$. The value of coupling at which this relation is satisfied is termed *optimum coupling*. Differentiation of Eq. (6-75) with respect to ω at constant M shows that A is a maximum for a given value of M at approximately the secondary resonance frequency. At this frequency, which is determined by the relation $\omega_0^2 = 1/L_2C_2$, the secondary reactance is zero, and the criterion for optimum coupling reduces to $\omega_0^2 M^2 = r_2 r_p$. Substitution of $\omega_0 = \sqrt{1/L_2C_2}$, $z_2 = r_2$, and $\omega_0^2 M^2 = r_2 r_p$ in Eq. (6-75) gives for the optimum amplification at resonance

$$\text{Optimum } A = -\tfrac{1}{2}\sqrt{\frac{\mu g_m L_2}{r_2 C_2}} = -\tfrac{1}{2}\sqrt{\mu g_m Q_2 \sqrt{\frac{L_2}{C_2}}} \tag{6-76}$$

It is evident from Eq. (6-76) that for high voltage amplification a radio-frequency amplifier with tuned secondary should use a tube with both high amplification factor and high transconductance, that the Q of the secondary should be high, and that the ratio of secondary inductance to capacitance should be high.

The curve of amplification as a function of frequency is of the same form as the curve of admittance *vs.* frequency of a simple series circuit containing inductance, capacitance, and resistance. The sharpness of the resonance peak increases with decrease of M and ω_0 and with increase of L_2/r_2 and r_p.

Dividing Eq. (6-73) by Eq. (6-76) shows (subject to the approximations made in the development of these equations) that the ratio of the optimum amplification of the doubly tuned amplifier to that of the singly tuned amplifier is equal to $\sqrt{L_1/r_1 r_p C_1}$. For typical transformers used with r-f pentodes this ratio is somewhat less than unity. In fixed-frequency amplifiers, however, the advantage of the sharpness of the response curve of the doubly tuned amplifier, which may be made to approximate rectangular form, more than offsets the disadvantage of the slightly lower amplification. In applications in which the frequency of the input changes, on the other hand, the greater simplicity of tuning of the singly tuned amplifier is an advantage.

It is evident from Eqs. (6-73) and (6-76) that high voltage amplification necessitates the use of low effective transformer resistances (high-Q coils) and high L/C ratio. The skin effect, radiation loss, dielectric loss, and tube input conductance resulting from electron transit time cause the effective a-c resistances r_1 and r_2 to increase with frequency. Distributed winding capacitances, circuit capacitances, and tube capacitances limit the extent to which C_1 and C_2 can be reduced and thus necessitate decreasing the L/C ratios as the resonant frequency of the amplifier is increased. The attainable voltage amplification of tuned r-f amplifiers therefore decreases as the frequency is increased. The decrease of amplification is accompanied by decrease of selectivity. These facts are very important in the design of r-f amplifiers and are one of the reasons why amplification in modern radio receivers is accomplished mainly at frequencies lower than those of the received signals.

6-24. Choice of Tubes for Tuned Radio-frequency Amplifiers.—Tetrodes and pentodes are used practically universally in tuned r-f amplifiers, in preference to triodes. The advantages of tetrodes and pentodes are threefold. The high values of amplification factor and transconductance result in high voltage amplification; the low grid-to-plate capacitance reduces the tendency of the amplifiers to oscillate as the result of the feeding back of amplified voltage to the input; and the lower effective input capacitance lowers the frequency limit of amplification and reduces the attainable amplification and sharpness of tuning at very high frequencies.

By the use of coils having a high ratio of inductance to resistance, the amplification of each stage of a tuned r-f amplifier using high-transconductance tubes can be made much greater than the amplification factor of the tube.

Control of Amplification.—The dependence of voltage amplification of tuned r-f amplifiers upon transconductance of the tube makes possible the variation of amplification by means of operating voltages. Usually the most effective method, particularly when the gain is controlled automatically by means of tubes, is to vary the grid bias of the amplifier tubes. Automatic control of amplification is discussed in Secs. 6-19 and 9-26.

6-25. Distortion in Radio-frequency Amplifiers.—Curvature of the transfer characteristic of tubes used in r-f amplifiers results in several types of distortion in the amplification of modulated voltages, most of which are caused by intermodulation of the various frequency components in the impressed modulated voltage or in two modulated voltages simultaneously impressed upon the amplifier.[1] In the reception of radio signals this produces such undesirable effects as nonlinear distortion of the audio-frequency output, superposition of the a-f modulation of an undesired station in an adjacent channel upon the desired a-f output of the station to which the receiver is tuned, and the reception of stations lying outside of the tuning range of the receiver when the receiver is tuned to the sum or difference of their frequencies. Another objectionable effect is increased hum in the output of a-c–operated receivers as the result of intermodulation between the desired input signal and 60- or 120-cycle voltages not completely removed from the output of the power supply and therefore impressed upon the amplifier grids. These various types of intermodulation may be shown to be associated with the second- and higher-order terms of the series expansion for alternating plate current.

In tuned r-f amplifiers using tubes with high plate resistance, such as tetrodes and pentodes, the load impedance at resonance may be small in comparison with the plate resistance of the tube. Under this condition the series expansion for alternating plate current reduces to

$$i_p = \frac{\partial i_b}{\partial e_c} e_g + \frac{1}{2!}\frac{\partial^2 i_b}{\partial e_c^2} e_g^2 + \frac{1}{3!}\frac{\partial^3 i_b}{\partial e_c^3} e_g^3 + \cdots \tag{6-77}$$

This shows clearly that the various types of distortion which occur in tuned r-f amplifiers increase with input amplitude and with curvature of the static transfer characteristics. In automatic control of amplification, the output is maintained constant by increasing the bias with input (Sec. 9-26). At high input the operating point is then close to cutoff. Curvature is high in this region, and there is even danger of reducing the plate current to zero during part of the cycle. Hence distortion may be excessive at high inputs. It may be reduced by the use of variable-mu tubes. The transfer characteristics of variable-mu tubes (see Sec. 3-7) approach the voltage axis so gradually that, even though the operating bias is made high in order to reduce the transconductance and hence the voltage amplification, high excitation amplitude may be used without danger that the current will be reduced to zero. At low bias, on the other hand, the transconductance has the high value necessary for the amplification of weak signals. Some tubes, of which the 6L7 is an example, have both a remote-cutoff grid, to which the signal is applied, and a sharp-cutoff grid, which may be used to control the amplification.

6-26. Current Amplification.—Figure 6-36 shows a single-stage current amplifier. The current amplification is

[1] HARRIS, SYLVAN, *Proc. I.R.E.*, **18**, 350 (1930); BALLANTINE, STUART, and SNOW, H. A., *Proc. I.R.E.*, **18**, 2102 (1930).

$$\frac{I_p}{I_i} = \frac{\mu E_g}{r_p + z_b} \cdot \frac{1}{I_i} = \frac{\mu z_c}{r_p + z_b} \tag{6-78}$$

The current amplification increases as z_b is decreased, approaching the limiting value $g_m z_c$. The current sensitivity is $\mu/(r_p + z_b)$, which approaches the value g_m as z_b is decreased. This indicates the importance of transconductance in tubes used for current amplification. In a multistage current amplifier, all but the last stage are voltage amplifiers and require tubes with high amplification factor to give high over-all current sensitivity. Special tubes and circuits for the amplification and measurement of small currents are described in Sec. 15-11.

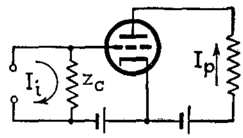


FIG. 6-36.—Single-stage current amplifier.

Current amplifiers are useful in the operation of relays[1] and electromagnetic oscillographs.[2] One common application is in the amplification of the current of phototubes used to control relays. Circuits for this purpose are analyzed in Chap. 13. Another application of current amplifiers is in the regulation of generator voltage.[3] A typical circuit is shown in basic form in Fig. 6-37, in which T_1 is a voltage amplifier, T_2 is a current amplifier, and T_3 is a glow tube that reduces the voltage applied to the grid of T_1 from the positive terminal of the generator and thus lowers the required biasing battery voltage (see Sec. 12-4). The

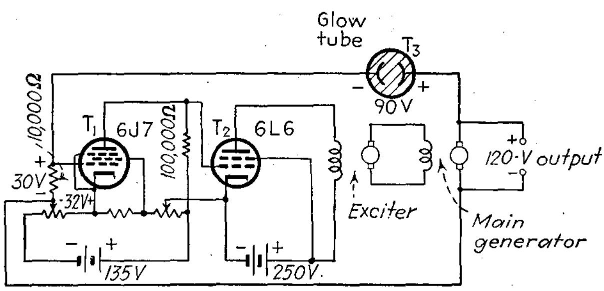


FIG. 6-37.—Vacuum-tube voltage regulator for d-c generator.

action is as follows: Reduction of terminal voltage makes the grid of T_1 more negative and so makes that of T_2 less negative. The resulting increase of plate current of T_2 raises the field current of the exciter, which

[1] GEORGE, E. E., *Electronics*, August, 1937, p. 19; DUDLEY, B., *Electronics*, May, 1938, p. 18; MUEHTER, M. W., *Electronics*, December, 1933, p. 336.

[2] JACKSON, W., *Wireless Eng.*, **11**, 64 (1934); REICH, H. J., and MARVIN, G. S., *Rev. Sci. Instruments*, **2**, 814 (1931); WALDORF, S. K., *J. Franklin Inst.*, **213**, 605 (1932).

[3] VAN DER BIJL, H. J., "The Thermionic Vacuum Tube and Its Applications," p. 371, McGraw-Hill Book Company, Inc., New York, 1920; VERMAN, L. C., and REICH, H. J., *Proc. I.R.E.*, **17**, 2075 (1929); VERMAN, L. C., and RICHARDS, L. A., *Rev. Sci. Instruments*, **1**, 581 (1930).

in turn raises the excitation of the main generator and restores the terminal voltage to nearly its original value. Obviously the exciter may be eliminated if the current of T_2 is sufficient to excite the generator directly. T_1 may also be eliminated if very high constancy of voltage is not required. By the addition of a rectifier, the circuit may be adapted to alternator regulation.

6-27. Shielding.—The high amplification obtained in radio-frequency amplifiers and in audio-frequency amplifiers using tetrodes or pentodes usually necessitates the complete shielding of the individual stages in multistage amplifiers to prevent oscillation. It may also be necessary to shield input leads to the amplifier to prevent the picking up of stray fields and to prevent oscillation as the result of feedback from the output to the input.

6-28. Limit of Amplification.—Amplification is usually limited by a tendency of the amplifier to oscillate as the result of the feeding back of some of the output voltage to the input. This can be minimized by shielding, by careful placing of apparatus, and by the use of chokes and by-pass condensers and of more than one power supply. By means of successive amplifiers operated from separate power supplies, very high gain or power output can be obtained. Because of inherent fluctuations of currents in amplifiers, however, there is a lower limit to the input voltage or current that can be amplified to a given output level. Current and voltage fluctuations in an amplifier are called *noise*. Noise results from a number of causes, among which are the shot effect, ionization, secondary emission, vibration of tube elements, imperfect contacts, fluctuations of resistance of old or damaged batteries, incomplete filtering of power supplies, variation of resistance of circuit elements, and random motion of electrons through high resistances (thermal agitation). The shot effect (see Sec. 2-14) can be kept to a minimum by operating tubes well above temperature saturation. Ionization and secondary emission, which have been greatly reduced in modern tubes, can be eliminated when necessary by the use of very low voltages. "Microphonic" effects resulting from the vibration of tube elements have been largely eliminated in modern tubes by improving the methods of mounting and supporting the elements. Current fluctuations produced by the random motion of electrons in resistances are in some respects similar to those produced by the shot effect. The r-m-s output voltage of an amplifier resulting from the random motion of electrons in a resistance is given by the formula

$$E = 2\sqrt{kT\int_{f_1}^{f_2} R|A|df} \qquad (6\text{-}79)$$

in which k is Boltzmann's constant (1.38×10^{-23} erg/degree); T is the absolute temperature in degrees K; R is the resistive component of the impedance in which the disturbance arises, measured at the frequency f; $|A|$ is the magnitude of the voltage amplification between the resistance and the output at the frequency f; and f_1 and f_2 are the lower and upper frequencies at which the amplification may be assumed to have fallen to zero. It is evident from Eq. (6-79) that, in order to minimize noise, it is advisable to make the frequency range of the amplifier only wide enough to cover the desired band.

6-29. Inverse Feedback Amplifiers.[1]—Much improvement in the operating characteristics of an amplifier may be obtained by feeding a portion of the output voltage back to the input. This process is called *feedback.* If it is accomplished in such a manner that the voltage fed back subtracts from the impressed voltage and thus reduces the alternating grid voltage, the output voltage is reduced and feedback is said to be *inverse.* Since the voltage fed back is greatest at frequencies at which the amplification is highest, the reduction in output voltage for a given input voltage is greatest at these frequencies. Inverse feedback therefore tends to reduce frequency distortion. The voltage fed back also contains nonlinear distortion and noise components which are amplified by the amplifier and cancel a portion of the nonlinear distortion and noise components in the output. Hence, inverse feedback also reduces nonlinear distortion and noise generated in the amplifier. Other equally important improvements result from the use of inverse feedback. The benefits may be summarized as follows:

1. Reduction of nonlinear distortion.
2. Reduction of frequency distortion.
3. Reduction of phase distortion.
4. Reduction of noise.
5. Increase of stability (reduction of variation of amplification or of current or power sensitivity with operating voltages and tube age).

In Fig. 6-38, which represents a feedback amplifier in general form, A is the amplification of the amplifier without feedback and β is the feedback-network attenuation, or ratio of the feedback voltage to the total output voltage. Both A and β are in general complex quantities. The product $A\beta$, which is called the *feedback factor,* is assumed to be positive if the vector sum of the impressed voltage and the feedback voltage is larger than the impressed voltage.

The total output voltage e_o consists of two parts: that resulting from the applied input e_i, and that resulting from the feedback input βe_o. The

[1] NYQUIST, H., *Bell System Tech. J.*, **11**, 126 (1932); BAGGALLY, W., *Wireless Eng.*, **10**, 413 (1933); BLACK, H. S., *Elec. Eng.*, **53**, 114 (1934); *Bell Laboratories Record,* **12**, 290 (1934). For additional references see the supplementary bibliography at the end of this chapter.

first is made up of the fundamental output $e_f = Ae_i$, harmonic and intermodulation components e_h, and noise e_n. The second contains three corresponding components, plus new harmonic and intermodulation components that are generated from the original distortion which is fed back. These secondary distortion components are ordinarily so small that they may be neglected. The amplitude distortion e_h is usually generated mainly in the final stage, and its amplitude may be assumed to be a function of the amplitude of the fundamental output voltage e_f.

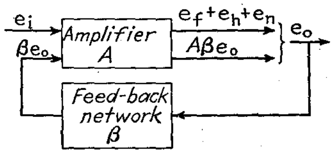


FIG. 6-38.—Schematic diagram of feedback amplifier.

The manner in which the noise output is affected by feedback depends upon the point in the amplifier at which it is introduced. Hence the analysis of the effect of feedback upon noise can best be considered independently of nonlinear distortion. This is made possible by the fact that noise is in general independent of signal amplitude. If the noise output is disregarded, the output voltage is

$$e_o = e_f + e_h + A\beta e_o \tag{6-80}$$

$$e_o = \frac{e_f + e_h}{1 - A\beta} = \frac{Ae_i}{1 - A\beta} + \frac{e_h}{1 - A\beta} \tag{6-81}$$

Without feedback the output voltage would be $e_f + e_h$.

Equation (6-81) shows that, for a given applied input voltage e_g, feedback changes both the fundamental and the harmonic components of the output voltage by the factor $1/(1 - A\beta)$.

The gain of an amplifier is increased, and feedback is said to be *positive*, or *regenerative*, when $|1 - A\beta| < 1$. The gain is decreased, and feedback is said to be *negative, inverse*, or *degenerative* when $|1 - A\beta| > 1$. When $|1 - A\beta| = 1$, the gain is unaffected by feedback.[1]

6-30. Reduction of Nonlinear Distortion by Inverse Feedback.—Suppose that a voltage e_i' is applied to the amplifier without feedback. This results in an output (noise neglected)

$$e_o' = e_f' + e_h' = Ae_i' + e_h' \tag{6-82}$$

To produce the same fundamental output with feedback requires an input voltage $e_i = (1 - A\beta)e_i'$. For this input, the output with feedback, as given by Eq. (6-81), is

[1] The student should note specially that the criterion for determining whether feedback is positive or negative is the magnitude of $1 - A\beta$ relative to unity, and not the magnitude or phase angle of the feedback factor $A\beta$. The feedback is, in fact, negative when $A\beta$ is positive and real, if its magnitude exceeds 2. This will be discussed in connection with polar diagrams (Sec. 6-34).

$$e_o = \frac{A(1 - A\beta)e_i'}{1 - A\beta} + \frac{e_h}{1 - A\beta} = e_f' + \frac{e_h}{1 - A\beta} \tag{6-83}$$

But since harmonic generation is a function of the amplitude of the fundamental output, which is the same for both amplifiers, $e_h = e_h'$. Therefore,

$$e_o = e_f' + \frac{e_h'}{1 - A\beta} \tag{6-84}$$

This shows that, for the same fundamental output, the harmonic and intermodulation output are changed in the ratio $1/(1 - A\beta)$ by the use of feedback. If $|1 - A\beta| > 1$, the nonlinear distortion is reduced. Thus nonlinear distortion is reduced by negative feedback.

Feedback does not reduce distortion resulting from the flow of grid current in the input stage of an amplifier when the grid is driven positive.[1] Since the grid-current distortion is a function of the alternating grid voltage e_g, and the distortion and the signal are both amplified in the ratio A between the input grid and the output, the amplified distortion is a function of the signal output and is unaffected by feedback if the signal output remains constant. Feedback does, however, reduce grid-current distortion set up in stages other than the first, the reduction being greatest for distortion set up in the output stage.

6-31. Reduction of Frequency Distortion by Inverse Feedback.—Equation (6-81) gives the voltage amplification or ratio of the fundamental output to the input of a feedback amplifier as

$$A' = \frac{e_o}{e_i} = \frac{A}{1 - A\beta} \tag{6-85}$$

It can be seen from Eq. (6-85) that for very small values of the denominator, such as may be obtained with positive feedback, a small change in the denominator, resulting from variation of A, may produce a very large change in e_o/e_i. As $A\beta$ is increased beyond the value 2 or 3, however, the importance of 1 diminishes in comparison with $A\beta$, particularly if the imaginary component of $A\beta$ is large in comparison with the real component. The ratio e_o/e_i approaches the limiting value $-1/\beta$, which is independent of A. Frequency distortion may, therefore, be reduced by the use of negative feedback.

If

$$|A\beta| >> 1, \tag{6-86}$$

Eq. (6-85) simplifies to

$$A' = \frac{e_o}{e_i} = -\frac{1}{\beta} \tag{6-87}$$

Equation (6-87) has three important implications:

[1] See pp. 125–126.

1. If $|A\beta| >> 1$, the gain is inversely proportional to β, which means that the frequency-response characteristic of the feedback amplifier with high feedback factor is the inverse of the response characteristic of the feedback network. By proper design of the feedback network the amplifier may be made to have any desired frequency response. In particular, it is possible to correct for nonuniform response of any network preceding or following a feedback amplifier by using an *identical* network as the feedback network of the amplifier. The design of this identical network is usually much simpler than the design of a network having inverse characteristics. Furthermore, if feedback is obtained through a resistive network, β is independent of frequency, and the relative phase shift and dependence of gain upon frequency may be made negligible.

2. If $|A\beta| >> 1$, amplification is independent of the load impedance in the output of the amplifier, provided that the load does not form a part of the feedback network. When β is made dependent upon load impedance, the amplification can be caused to vary in a desired manner with load impedance.

3. If $|A\beta| >> 1$, the amplification is independent of A and, therefore, of tube factors. The amplification is unaffected by variations of battery voltage or aging of tubes. Hence the gain stability of an amplifier is improved by feedback if the feedback factor is large.

Equation (6-87) also holds under the more general condition that

$$|1 - A\beta| = |A\beta| \tag{6-88}$$

of which Eq. (6-86) is a special form. If $A\beta$ is written in the form $a + jb$, Eq. (6-88) is satisfied if

$$\sqrt{(1 - a)^2 + b^2} = \sqrt{a^2 + b^2} \tag{6-89}$$
$$a = \tfrac{1}{2} \tag{6-90}$$

Thus the amplification with feedback may also be made to depend only on β by making the real component of the complex feedback factor equal to $\frac{1}{2}$. Since $A\beta$ changes with frequency, this condition can in general be satisfied at only one frequency and from a practical point of view is, therefore, of only incidental interest.

6-32. Reduction of Phase Shift by Inverse Feedback.—Reduction of amplifier phase shift by inverse feedback may be readily shown for the case in which β is a real quantity, *i.e.*, when there is no phase shift in the feedback network. Let

$$A = a + jb = |A|\underline{/\theta} \tag{6-91}$$

in which $|A|$ is the magnitude of the voltage amplification and θ is the phase angle of the output voltage of the amplifier, relative to the input voltage, without feedback.

$$\theta = \tan^{-1}\frac{b}{a} \tag{6-92}$$

and

$$|A| = \sqrt{a^2 + b^2} \tag{6-93}$$

The amplification with feedback is

$$A' = \frac{E_o}{E_i} = \frac{A}{1 - A\beta} = \frac{a + jb}{1 - (a + jb)\beta} = \frac{a - (a^2 + b^2)\beta + jb}{(1 - a\beta)^2 + b^2\beta^2} \tag{6-94}$$

in which E_0 and E_i are the output and input voltages, respectively, of the feedback amplifier. The phase angle of the output voltage relative to the input voltage, with feedback, is

$$\alpha = \tan^{-1}\frac{b}{a - (a^2 + b^2)\beta} = \tan^{-1}\frac{b/a}{1 - A\beta\sqrt{1 + b^2/a^2}}$$
$$= \tan^{-1}\frac{\tan\theta}{1 - A\beta\sqrt{1 + \tan^2\theta}} \tag{6-95}$$

When the feedback is negative, *i.e.*, when $|1 - A\beta| > 1$, $|A\beta|$ must be either negative or greater than 2. Examination of Eq. (6-95) shows that α is then less than θ. Hence the phase shift of the output voltage relative to the input voltage is reduced by inverse feedback. A similar analysis may be made for the more general case in which β is complex.

6-33. Reduction of Noise by Feedback.—An analysis similar to that which was made for the effect of feedback upon nonlinear distortion shows that if the output is kept constant by increasing the input, A being kept constant, the effect of inverse feedback is to reduce the noise-to-signal ratio of the output by the factor $1/(1 - A\beta)$, regardless of the point at which the noise is introduced. If the output is maintained constant by increasing the voltage amplification A of the amplifier, on the other hand, rather than by increasing the input voltage, the noise-to-signal ratio is reduced in the ratio $1/(1 - A\beta)$ only when the change of A is made in the portion of the amplifier preceding the point at which the noise is introduced. It is unaffected by feedback if the change in A is made in the portion of the amplifier that follows the point at which the noise is introduced. At constant excitation and constant output, therefore, negative feedback is effective in reducing hum resulting from poor filtering of the power supply of the final stage, but has little or no effect upon noise-to-signal ratio of hum or other noises originating in early stages, such as those caused by thermal agitation and microphonic effects.

6-33A. Effect of Inverse Feedback upon Allowable Input Voltage.—Examination of Fig. 6-38 shows that the ratio of the impressed voltage to the alternating grid voltage of a feedback amplifier is $(1 - A\beta)$. The voltage that can be impressed without overloading the amplifier is therefore increased in the ratio $(1 - A\beta)$ by inverse feedback. It has already been pointed out that the impressed voltage must be increased in this ratio in order to produce the same output as without feedback.

6-34. Polar Diagram.—Black has shown that much important information regarding the performance of a feedback amplifier can be determined from a polar diagram of $A\beta$ drawn for all frequencies from 0 to ∞.[1] $A\beta$ may be expressed in the form $a + jb$, or in the equivalent polar form $|A\beta|\underline{/\theta}$, where

$$|A\beta| = \sqrt{a^2 + b^2} \qquad \text{and} \qquad \theta = \tan^{-1} \frac{b}{a}.$$

In a multistage amplifier

$$\left.\begin{aligned} A\beta &= |\beta||A_1||A_2||A_3| \cdots |A_n| \cdots \underline{/\theta_\beta + \theta_1 + \theta_2 + \theta_3 + \cdots \theta_n + \cdots} \\ &\text{where} \\ \beta &= a_\beta + jb_\beta = |\beta|\underline{/\theta_\beta} \\ A_n &= a_n + jb_n = |A_n|\underline{/\theta_n} \end{aligned}\right\} \tag{6-96}$$

The polar diagram may be constructed by plotting the imaginary part b, of $A\beta$, against the real part a, or by plotting $|A\beta|$, the length of the radius vector against the polar angle θ. $|A\beta|$ is the product of the magnitude of β and the magnitudes of the gains of individual stages, and θ is the sum of the phase angle of β and those of the individual stages of the amplifier.

For ordinary forms of resistance- and transformer-coupled amplifiers the magnitudes of the stage amplifications and the corresponding phase angles can be determined from Figs. 6-12, 6-14, 6-30, and 6-31 or the equations from which they are derived. Consider, for instance, the transformer-coupled amplifier of Fig. 6-40*b*, in which the feedback network is a simple resistance voltage divider. Assume that the mid-band amplification is 30 and that $r_f = r_f'$. Then $\beta = r_f/(r_f + r_f') = 0.5$ and in the mid-band range $|A\beta| = 30 \times 0.5 = 15$. Since the voltage fed back is opposite in phase to the impressed voltage, $\theta = 180$ degrees. Below the mid-band range, corresponding values of $|A\beta|$ and θ are found from Fig. 6-30. Thus, when $L_1/(r_p + r_1)$ is 0.3, $|A_l/A_m| = 0.3$ and the phase shift relative to the mid-band phase is 73 degrees leading. Hence $|A\beta| = 30 \times 0.3 \times 0.5 = 4.5$ and $\theta = 180 - 73 = 253$ degrees. Above the mid-band range Fig. 6-31 is used in a similar manner. Thus, if Q_0 is 1.0, $|A_h/A_m|$ is 1.1 when ω/ω_0 is 0.5, and the corresponding phase angle relative to the mid-band value is 33 degrees lagging. Therefore

[1] BLACK, *loc. cit.*

$|A\beta| = 30 \times 1.1 \times 0.5 = 16.5$ and $\theta = 180 - 33 = 147$ degrees. Other corresponding values of $|A\beta|$ and θ are found in a similar manner.

Typical polar diagrams are shown in Fig. 6-39. Diagram a is that for the resistance-capacitance–coupled stage obtained when z_b in Fig. 6-40a is a resistance. Diagram b is for the transformer-coupled stage of Fig. 6-40b, and diagrams c and d are for the two-stage feedback amplifier of Fig. 6-41b. In these diagrams the phase angle is indicated relative to the position of the input voltage vector. The circuits of Figs. 6-40 and 6-41 are such that the feedback voltage at mid-band

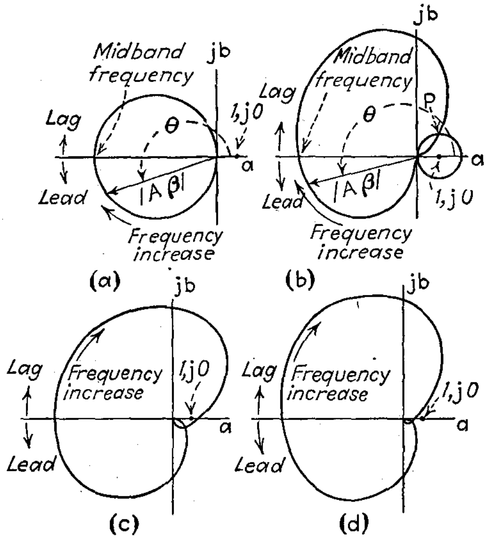


FIG. 6-39.—Polar diagrams for (a) the circuit of Fig. 6-40a with nonreactive load, (b) the circuit of Fig. 6-40b, and (c) and (d) the circuit of Fig. 6-41b.

frequencies is opposite in phase to the applied voltage. The feedback vector is, therefore, negative at these frequencies and is drawn to the left of the origin. It should be noted that equal arc lengths or angles of the polar curve do not necessarily represent equal frequency ranges. The vector will, for instance, remain fixed in magnitude and direction throughout any frequency range for which the amplification of the amplifier and the attenuation of the feedback network are constant.

By means of a simple vector diagram, or from the relation

$$\sqrt{(1 - a)^2 + b^2} = 1,$$

the student may show that the locus of all values of $A\beta$ for which

$$|1 - A\beta| = 1$$

is a circle of unit radius on the polar diagram, having its center at the point 1, $j0$. From the definitions of degenerative and regenerative feedback it follows that for all frequencies at which $A\beta$ terminates within this circle the amplifier is regenerative, and for all values of frequency for which $A\beta$ terminates outside of this circle the amplifier is degenerative. Figure 6-39*b*, for example, shows that the amplifier of Fig. 6-40*b* is regenerative at frequencies higher than that corresponding to point P; it is degenerative at lower frequencies, and is unaffected by feedback at the frequency corresponding to P. The beneficial effects of the feedback are obtained only in the frequency range in which the feedback is degenerative.

6-35. Oscillation.—Equation (6-85) shows that the amplification of a feedback amplifier is infinite if $A\beta = 1 + j0$. Under this condition, oscillation takes place. Nyquist has shown[1] that this is not the only condition under which the amplifier will oscillate but that oscillation will occur if the polar diagram formed by plotting $A\beta$ and its conjugate encloses the point 1, $j0$. An inspection of Fig. 6-39 shows that such a polar diagram cannot enclose the point 1, $j0$, if the absolute value of $A\beta$ is less than unity when the feedback voltage is in phase with the input voltage ($\theta = 0$). This requirement can be met with ease in several common types of amplifiers. Practical negative-feedback amplifiers are designed so that the feedback voltage is opposite in phase to the input voltage at the middle of the frequency band, where the response curve is flat. Oscillation will not occur if the phase shift of the feedback voltage, relative to its phase at the middle of the band, at no time reaches 180 degrees, or if the absolute value of $A\beta$ falls below unity before the relative phase shift reaches 180 degrees. Figures 6-12, 6-14, 6-30, and 6-31 show that the high values of relative phase shift occur at frequencies for which A is small, so that it is possible to prevent oscillation in one-, two- or three-stage resistance-coupled inverse-feedback amplifiers and in single-stage transformer-coupled inverse-feedback amplifiers.

6-36. Practical Inverse Feedback Circuits.—Since the over-all amplification of a feedback amplifier is $A/(1 - A\beta)$, reduction of distortion and noise and increase of supply-voltage stability are obtained with less sacrifice of over-all amplification if the feedback factor $A\beta$ is increased by increasing A, rather than by increasing β. For this reason the advantages of inverse feedback increase with amplification per stage and with the number of stages. Because very high stage gain can be realized with tetrodes and pentodes, they are almost always used in at least the voltage amplifying stages of a multistage feedback amplifier. In the diagrams of Figs. 6-40, 6-41, and 6-42, all tubes are shown as triodes in order to eliminate those portions of the circuits not actually involved in feedback. Tetrodes and pentodes merely require additional

[1] NYQUIST, *loc. cit.*

fixed grid voltages. The condensers shunting the self-biasing resistors are assumed to be sufficiently large so that negligible feedback is produced by the resistors that they shunt.

Figure 6-40a shows one of the two basic types of single-stage negative feedback circuits. For the special case in which z_b is a resistance, the polar diagram, shown in Fig. 6-39a, is readily drawn by the use of Figs. 6-12 and 6-14. This diagram does not intersect a circle of unit radius drawn about the point 1, $j0$, and does not enclose this point. This type of amplifier is, therefore, always degenerative and cannot oscillate. If z_b contains inductive reactance, the relative phase shift approaches 180 degrees lead at low frequency, and the diagram is in the fourth

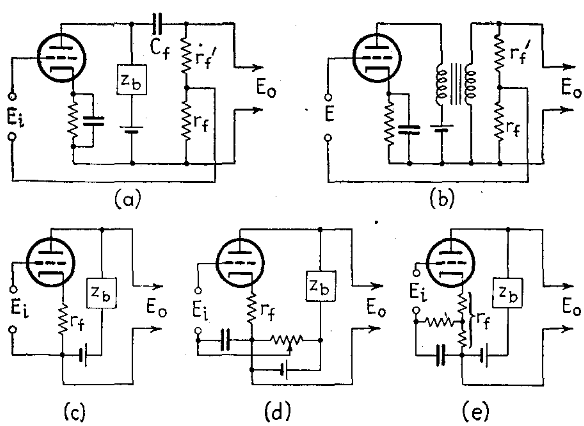


FIG. 6-40.—Single-stage single-sided inverse-feedback amplifiers.

quadrant at low frequency. The diagram does not enclose the point 1, $j0$, but at low frequency is within the unit circle about this point. The amplifier is therefore regenerative at low frequency but does not oscillate. The impedance z_b in Fig. 6-40a may be the primary of an output transformer. The feedback voltage may also be taken from the secondary of the transformer, as in Fig. 6-40b. No isolating condenser need then be used. The polar diagram for this circuit, derived by the use of Figs. 6-30 and 6-31, is that of Fig. 6-39b. This diagram shows that the circuit is regenerative at high and low frequencies, where the gain has fallen to low values, but that oscillation does not take place.

The second type of single-stage circuit, shown in Fig. 6-40c, makes use of the bias resistor to provide feedback voltage. $\beta = r_f/z_b$ and $A = -\mu z_b/(r_p + z_b + r_f)$. The polar diagram may be derived from these expressions for β and A. The feedback factor can be increased above that obtainable with the biasing resistor alone by use of the modified circuits of Figs. 6-40d and 6-40e.

Figure 6-41*a* shows the basic inverse feedback circuit that is ordinarily used in a two-stage resistance-capacitance–coupled amplifier. The feedback voltage must be applied to the first tube in a different manner from that of Fig. 6-40*a* because of the added phase shift which takes place in the second stage. At frequencies for which the reactance of the condenser C_f is negligible, $\beta = -r_f/(r_f + r_f')$. Figure 6-41*b* shows a modification of Fig. 6-41*a* in which the use of an output transformer makes the coupling condenser unnecessary. Polar diagrams for this amplifier are shown in Figs. 6-39*c* and 6-39*d*. Diagram *c* is for the case in which gain of the resistance-coupled stage begins to fall off at about the resonance frequency of the transformer stage. Diagram *d* is for the case in which the gain of the resistance-coupled stage is uniform to frequencies at which the amplification of the transformer stage has

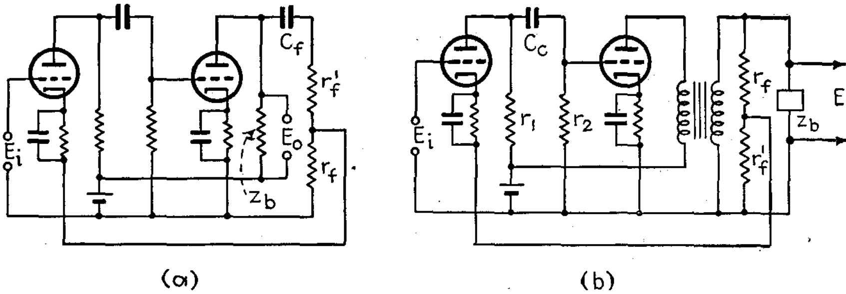


FIG. 6-41.—Two-stage single-sided inverse-feedback amplifiers.

fallen to a low value. Both amplifiers are regenerative at low and high frequencies, but only the former will oscillate.

In a three-stage resistance-capacitance–coupled amplifier the relative phase shift approaches 270 degrees at very low and very high frequency. To prevent oscillation when inverse feedback is used, it is necessary to ensure that at the extremes of the frequency band, where the phase shift is high, the amplification has fallen sufficiently so that $|A\beta| < 1$. This may be accomplished by designing one stage to have nearly uniform amplification, and hence small phase shift, up to and beyond the frequency limits at which the amplification of the other two stages falls to such a low value that $|A\beta| < 1$.[1] Because the phase shift in a transformer-coupled stage may approach 180 degrees at high frequency unless the transformer secondary is loaded, it is not feasible to use inverse feedback in multistage transformer-coupled amplifiers unless low shunting resistances are used across the secondaries.[1]

Figure 6-42 shows two methods by which negative feedback may be applied to a push-pull stage. Circuit *a* corrects for loss of gain at low frequency caused by falling primary reactance[1] but does not correct for a

[1] TERMAN, F. E., *Electronics*, January, 1937, p. 12.

falling high-frequency characteristic caused by distributed capacitance of the secondary of the output transformer. Circuit *b* corrects for both, but some difficulty may be experienced in obtaining sufficient feedback voltage with a low-impedance secondary load such as a loud-speaker voice coil. The use of feedback on only one side of a push-pull amplifier should be avoided, and less distortion is obtained in push-pull amplifiers with transformer output when the feedback is between points at which the amplifier is single sided.[1]

The circuits of Figs. 6-40*a*, 6-40*b*, 6-41, 6-42*a*, and 6-42*b* give all the benefits that have been discussed and in addition result in increased damping of transient oscillations of a loud-speaker,[2] used as the load. Solution of equivalent plate circuits for these circuits, when the load impedance is replaced by a generator of zero internal resistance and voltage E, shows that the effective internal impedance between the output

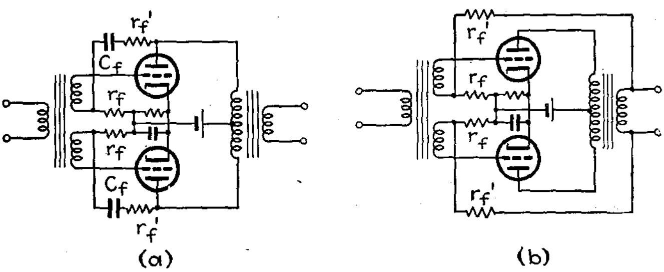


FIG. 6-42.—Single-stage push-pull inverse-feedback amplifiers.

terminals of the amplifier is changed by the factor $1/(1 - A\beta)$ by feedback. Negative feedback therefore reduces the effective resistance shunting the loud-speaker and thereby increases the speaker damping.

In the circuits of Figs. 6-40*c*, 6-40*d*, and 6-40*e*, β varies with load impedance, and therefore feedback does not correct for frequency discrimination caused by variation of load impedance. These circuits do, however, reduce harmonic distortion and hum, and increase the gain stability. Solution of the equivalent plate circuits for the circuits of Figs. 6-40*c*, 6-40*d*, and 6-40*e* with a generator in place of the load shows that the effective terminal impedance is $r_p + r_f(1 + A)$. The increase of effective resistance shunting the loud-speaker reduces the damping with these circuits.

Ample improvement in amplifier characteristics and performance is usually attained with a negative-feedback factor equal to 3 or 4 (approximately 10 to 12 db) in the mid-band frequency range. For an amplifier designed for carrier telephone service, Black reported distortion energy

[1] DAY, J. R., and RUSSELL, J. B., *Electronics*, April, 1937, p. 16.

[2] DREYER, J. F., JR., *Electronics*, April, 1936, p. 18; July, 1936, p. 30; MARTIN, L., *Radio Eng.*, May, 1937, p. 13.

75 db below signal energy, as compared with 35 db without feedback, and gain stability better than 0.0007 db per volt change of plate voltage. Figure 6-43 shows frequency-response curves of an amplifier with and without feedback. A Class B amplifier with inverse feedback, and distortion curves for this amplifier, are shown in Figs. 8-28 and 8-29.

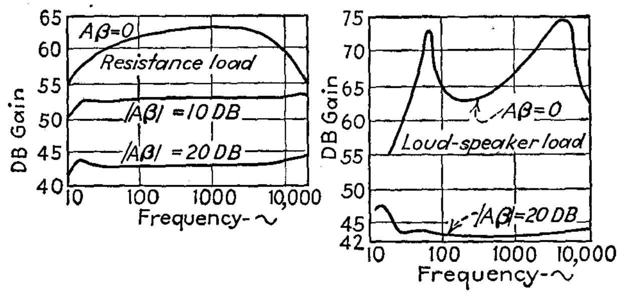


FIG. 6-43.—Frequency-response curves showing the reduction of frequency distortion that results from inverse feedback.

Figure 6-44 shows a push-pull direct-coupled amplifier in which inverse feedback is used to stabilize against changes of supply voltage and tube characteristics, to reduce distortion, and to afford a simple means of controlling amplification.[1] The feedback is controlled by means of the variable resistance R_1. The potentiometer P and variable resistance R_2

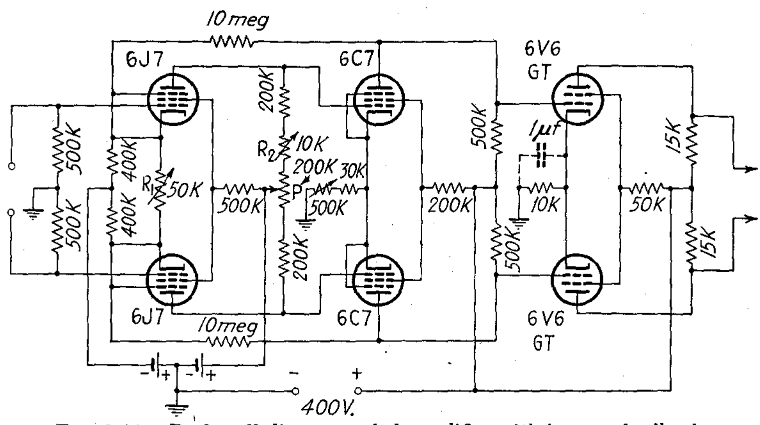


FIG. 6-44.—Push-pull direct-coupled amplifier with inverse feedback.

are coarse and fine balancing controls. The amplification of this amplifier ranges from 10^3 to 10^6 and the response curve is flat up to 10,000 cps.

6-37. Use of Inverse Feedback to Obtain High Selectivity.—Amplifiers having very high selectivity may be obtained by the use of inverse feedback by taking the feedback voltage from the output of the amplifier

[1] GOODWIN, C. W., *Yale J. Biol. and Med.*, **14**, 101 (1941).

through some form of resonance bridge, such as those discussed in Sec. 15-16.[1] The output of the bridge is zero at resonance but increases rapidly as the frequency departs from the resonance value. The amplification is therefore high at resonance and falls rapidly on either side of resonance. A typical circuit is shown in Fig. 6-45.[2] In this circuit, the effective Q, which is a measure of the selectivity of the circuit, is $Q_o[1 + AR_3/(R_2 + R_3)]$, in which Q_o is the Q of the tuned arm of the bridge and A is the voltage amplification of the amplifier without feedback. Values of effective Q ranging from 2000 to 5000 may be readily realized with this circuit at audio and low radio frequencies. By means of the potentiometer P the selectivity may be varied without changing the amplification at resonance. The circuit of Fig. 6-45 and similar

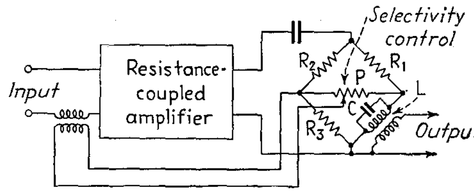


FIG. 6-45.—The use of inverse feedback to obtain high selectivity.

circuits are used in oscillators (see Sec. 10-39), wave analyzers (see Sec. 15-38), and other instruments.

6-38. Use of By-pass Condensers.—The presence of impedance in the voltage supply of a multistage amplifier or in a common lead between the voltage supply and the plates or other electrodes may result in objectionable feedback or even oscillation.[3] An example of a circuit in which a resistance common to several stages may lead to oscillation is given by Fig. 6-46. The resistance r may be that of a worn battery, a poorly designed power supply, or of a voltage-dropping resistor. The arrows indicate the directions of the instantaneous alternating plate currents of the three tubes resulting from an instantaneous applied voltage that swings the grid voltage of the first tube in the positive direction. Because of amplification, $i_{p1} + i_{p3}$ exceeds i_{p2}, and so the drop through r is in the direction shown. This voltage drop is added to that caused by the flow of i_{p1} through the plate resistor of the first stage, and the alternating voltage applied to the grid of the second tube is greater than it would be if r were zero. A number of other similar feedback effects take place between stages because of r, but these are of smaller magnitude. If r is large and the gain of the amplifier high, the

[1] AUGUSTADT, U. S. Patent 2106785; SCOTT, H. H., *Proc. I.R.E.*, **26**, 226 (1938); TERMAN, F. E., BUSS, R. R., HEWLETT, W. R., and CAHILL, F. C., *Proc. I.R.E.*, **27**, 649 (1939).

[2] TERMAN, BUSS, HEWLETT, and CAHILL, *ibid.*

[3] ANDERSON, J. E., *Proc. I.R.E.*, **15**, 195 (1927); TAMM, R., *E.T.Z.*, **57**, 631 (1930).

feedback may be sufficiently great to cause oscillation. This may be prevented by shunting r with a condenser C, whose reactance is small throughout the amplification range. Because wound paper condensers have appreciable inductance, large condensers of this type should be shunted by small condensers that are effective at radio frequencies.

A common source of feedback in audio-frequency amplifiers is B-supply regulation. Because of variation of voltage drop through the

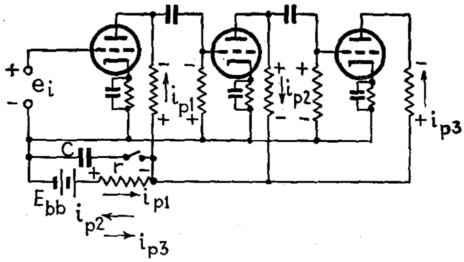


FIG. 6-46.—Circuit diagram showing the manner in which common plate supply impedance causes regeneration in a multistage resistance-coupled amplifier with an odd number of stages.

rectifier tubes and filter, the terminal voltage of a B-supply varies with the current drain. This type of feedback causes low-frequency oscillation, aptly termed *motorboating*. Feedback resulting from B-supply regulation can be reduced by using in each stage a *decoupling* filter consisting of a resistance and a condenser, as shown by the heavy lines of Fig. 6-47.[1] Better filtering may be obtained, particularly in radio-frequency amplifiers, by the use of chokes instead of resistors. Chokes have the additional advantage of not reducing the operating plate voltage greatly. Better methods of preventing motorboating are to make the gain of the amplifier small enough at very low frequency so that this low-frequency oscillation cannot occur,[2] or to use inverse feedback. When voltage-dropping resistors must be used to reduce the electrode voltages of the tubes of a number of stages, it is in general advisable to use individual resistors and by-pass condensers for the different tubes.

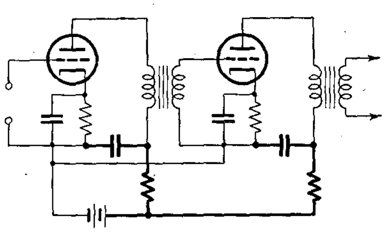


FIG. 6-47.—Plate decoupling filters (heavy lines).

The tendency toward motorboating is much less in push-pull amplifiers than in single-sided amplifiers, since voltage drops in voltage-supply impedances are normally in phase opposition for currents of the two sides

[1] BAGGALLY, W., *Wireless Eng.*, **11**, 179 (1934); COCKING, W. T., *Wireless World*, **29**, 322, (1931). See also p. 132.

[2] RCA Application Note 67.

of the amplifier. There is, however, the possibility that voltage from the output may be fed back to the input in such a manner that the grid voltages of the two input tubes are in phase. The two tubes in each stage respond to this type of excitation as though they were in parallel, instead of in push-pull, and so motorboating may result.[1] This difficulty is prevented by the use of inverse feedback. It is of interest to note that the feedback control in the circuit of Fig. 6-44 is such that it does not reduce the inverse feedback insofar as action of the tubes in parallel is concerned.

6-39. Use of Vacuum Tubes as Variable-impedance Elements.[2]— Vacuum tubes may be used in obtaining an impedance the magnitude and phase angle of which may be readily varied. Inductive reactance may be produced by circuits containing only capacitances and resistances, and capacitive reactance by circuits containing only inductances and resistances. Negative effective resistance may also be readily produced.

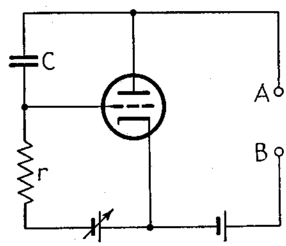


FIG. 6-48.—Reactance-tube circuit.

Figure 6-48 shows one type of impedance-conversion circuit. In this circuit the alternating plate voltage is the voltage across C and r, whereas the alternating grid voltage is that across r. Since the voltage across r leads that across the series combination of C and r, and the grid voltage has more effect upon the plate current than does the plate voltage, the plate current leads the plate voltage. The tube therefore acts as an impedance having a capacitive component. Solution of the equivalent plate circuit shows that, between the input terminals A and B, the circuit acts like a parallel combination of capacitive reactance x_e and resistance r_e of values (see Prob. 4-3).[3]

$$x_e = -\frac{r_p(r^2\omega^2C^2 + 1)}{C(\mu r + r_p)} \tag{6-97}$$

$$r_e = \frac{r_p(r^2\omega^2C^2 + 1)}{r^2\omega^2C^2(1 + \mu) + r_pr\omega^2C^2 + 1} \tag{6-98}$$

The effective capacitance between A and B is

$$C_e = \frac{C}{\omega^2r^2C^2 + 1}(g_mr + 1) \tag{6-99}$$

[1] GOODWIN, *loc. cit.*

[2] TRAVIS, C., *Proc. I.R.E.*, **23**, 1125 (1935); SHEAFFER, C. F., *Proc. I.R.E.*, **28**, 66 (1940); REICH, H. J., *Proc. I.R.E.*, **30**, 288 (1942); BRUNETTI, C., and GREENOUGH, L., *Proc. I.R.E.*, **30**, 542 (1942). See also Secs. 10-23 and 10-24 and Probs. 4-3 to 4-5.

[3] Expressions may, of course, also be derived for equivalent series resistance and reactance.

Ordinarily g_m is so much greater than unity that the second term in the parentheses may be neglected. C_e then has its maximum value when $r = 1/\omega C$ and the maximum value of effective capacitance is

$$\text{Max. } C_e = \frac{g_m}{2\omega} \tag{6-100}$$

The corresponding value of effective shunting resistance is

$$r_e \text{ at max. } C_e = \frac{2}{g_m} \tag{6-101}$$

Since the transconductance depends upon the grid bias, the effective capacitance can be changed by means of the grid bias.

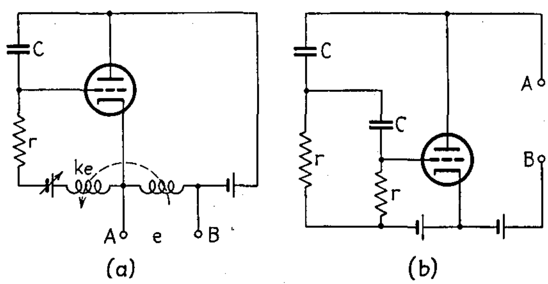


FIG. 6-49.—Variants of the reactance-tube circuit of Fig. 6-48.

Equation (6-100) indicates that large effective capacitance necessitates the use of tubes with high transconductance. Equation (6-101) shows, however, that the effective shunting resistance varies inversely with transconductance. It follows that the use of a high-transconductance tube in order to obtain high effective capacitance results in low effective shunting resistance, which may be undesirable in some applications of the circuit.

If rg_m and μ are large in comparison with unity, the effective conductance is zero (the effective shunting resistance is infinite) when the grid voltage is approximately 90 degrees out of phase with the plate voltage. This phase relationship can be attained in the circuit of Fig. 6-48 only by making $r\omega C$ zero. The grid voltage and, therefore, C_e are then also zero. Figure 6-49 shows two circuits in which the required phase relation can be attained without reducing the alternating grid voltage to such an extent that the effective capacitance is too small to be useful. Circuit *a* incorporates a transformer of such turn ratio and coupling that the application of the voltage E to the plate circuit results in the application of a voltage kE to the grid circuit, in phase opposition to the voltage applied to the plate circuit.[1] Solution of the equivalent plate circuit

[1] SHEAFFER, *loc. cit.*

shows that, under the assumptions that $rg_m >> k + 1$ and that the primary impedance of the transformer is infinite, the effective input conductance is zero when

$$(rC\omega)^2 = \frac{k\mu - 1}{\mu + 1} \tag{6-102}$$

If $k\mu$ is considerably larger than unity, the effective input resistance is infinite when $(rC\omega)^2$ is equal to k, *i.e.*, when the resistance r is equal to $\sqrt{k}$ times the reactance of the condenser C. This is also the approximate criterion for maximum effective capacitance, which is given by the relation

$$\text{Max. } C_e \cong \frac{\sqrt{k}g_m}{\omega} \tag{6-103}$$

Comparison of Eqs. (6-100) and (6-103) shows that, when k exceeds $\frac{1}{4}$, the circuit of Fig. 6-49*a* gives greater maximum effective capacitance than the circuit of Fig. 6-48. The fact that the effective resistance shunting this maximum effective capacitance is theoretically infinite may be a great advantage in some applications of the circuit of Fig. 6-49. Because the primary impedance of the transformer is not infinite, r_e cannot be made infinite in practice. It may, however, be made large. The type of transformer used in this circuit depends upon the frequency at which the circuit is to be used. An audio interstage coupling transformer may be used at audio frequencies, and singly or doubly tuned transformers at radio frequencies.

In the circuit of Fig. 6-49*b* the proper phase relation between the grid and plate voltages is achieved by the use of a two-stage resistance-capacitance coupling network. Analysis of this circuit shows that, when rg_m is large in comparison with unity, the effective admittance between A and B is given by Eq. (6-117). If μ equals or exceeds 15, the effective conductance can be made zero by proper choice of $r\omega C$. For values of μ equal to 20 and 100, respectively, the conductance is zero when $r\omega C$ is equal to 0.53 and 0.96, and the corresponding values of effective capacitance C_e are $0.147g_m/\omega$ and $0.32g_m/\omega$. At large values of μ the effective capacitance approaches the limiting value $g_m/3\omega$. It should be noted that increasing the value of μ decreases the amplitude of the voltage that can be impressed upon the circuit without overloading. A power pentode such as the 6F6, which combines relatively high values of μ, g_m, and allowable grid swing, is a good tube for this circuit.

Although the tubes in Figs. 6-47 and 6-48 are shown as triodes, the plate and grid may be replaced by any two electrodes of a multielectrode tube if the control electrode is maintained sufficiently negative to prevent the flow of electrons to it. In particular, the plate and grid of Fig. 6-48 may be replaced by the screen and suppressor, respectively, of a pentode, the plate and first grid voltages being maintained constant, as shown in Fig. 6-50. Appropriate changes of symbols then transform Eqs. (6-97) and (6-98) into (see Prob. 4-4)

$$x_e = -\frac{r_{g2}(r^2C^2\omega^2 + 1)}{C\omega(\mu_{23}r + r_{g2})} \tag{6-104}$$

$$r_e = \frac{r_{g2}(r^2C^2\omega^2 + 1)}{r^2C^2\omega^2(1 + \mu_{23}) + r_{g2}r\omega^2C^2 + 1} \tag{6-105}$$

in which r_{g2} is the screen resistance, and μ_{23} is the screen-suppressor mu-factor. Because an increase of negative suppressor voltage reduces the number of electrons that go to the plate, the screen current increases with a negative increment of suppressor voltage. μ_{23} is therefore negative. Since the magnitude of μ_{23} is greater than unity, r_e may be negative; and since $|\mu_{23}|r$ may be greater than r_{g2}, the numerical value of x_e may be positive, *i.e.*, the effective reactance between A and B may be inductive. The circuit thus acts like a negative resistance in parallel with (or in series with) an inductance.[1]

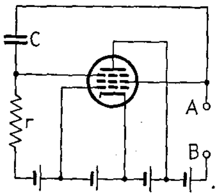


FIG. 6-50.—Pentode reactance-tube circuit.

Expressions for x_e and r_e may be readily derived for circuits in which the capacitance is replaced by an inductance and for those in which the position of the condenser or inductance is interchanged with that of the resistance.[2]

A second type of circuit that may be used in changing the phase angle or magnitude of an impedance is shown in Fig. 6-51. By the use of inverse feedback and of low load impedance in the final stage of the amplifier, the voltage amplification A is made independent of r and C throughout the frequency range in which the circuit is to be used. The output voltage AE is then in phase with or in phase opposition to the impressed voltage E. A is assumed to be positive when the relative polarities of the instantaneous input and output voltages are as indicated by the signs in Fig. 6-51. Under the assumption that the input impedance of the amplifier is infinite, the current resulting from the application of the voltage E is

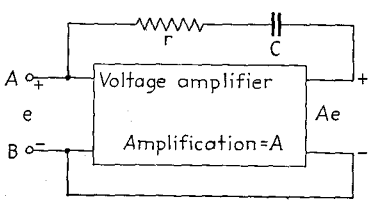


FIG. 6-51.—Amplifier-type reactance-tube circuit.

[1] Actually the circuit acts like an inductance only in that the current lags the voltage. Since the instantaneous value of the reactive voltage between A and B is proportional to the integral of the current, rather than to the rate of change of current, it is more correct to say that the circuit acts like a negative capacitance.

[2] REICH, *loc. cit.*

$$I = \frac{E - AE}{r + 1/j\omega C} \tag{6-106}$$

and between the points A and B the circuit acts like a parallel combination of reactance x_e and resistance r_e of values

$$x_e = -\frac{r^2\omega^2C^2 + 1}{C(1 - A)} \tag{6-107}$$

$$r_e = \frac{r^2\omega^2C^2 + 1}{r\omega^2C^2(1 - A)} \tag{6-108}$$

or like a series combination of reactance x_e' and resistance r_e' of values

$$x_e' = \frac{1}{(1 - A)\omega C} \tag{6-109}$$

$$r_e' = \frac{r}{1 - A} \tag{6-110}$$

If the amplifier contains an odd number of stages, the output voltage is in phase opposition to the input voltage (both being measured relative to the common terminals) and the numerical value of A is negative in Eqs. (6-107) and (6-108). The effective resistance is then positive and the reactance capacitive. The equivalent capacitance is

$$C_e = \frac{(1 - A)C}{r^2\omega^2C^2 + 1} \tag{6-111}$$

At a given value of C, C_e has its maximum value when r is zero. The effective resistance shunting C_e is then infinite and the effective capacitance is

$$\text{Max. } C_e = (1 - A)C \tag{6-112}$$

It is evident that very large values of effective capacitance can be obtained with this circuit if A is large. Since A may be varied by means of the bias of one or more grids of a stage to which inverse feedback is not applied, the value of the effective capacitance may be readily varied. The input voltage E must, of course, be maintained low enough to prevent overloading of the amplifier.

If the amplifier of Fig. 6-51 contains an even number of stages, the numerical value of A is positive in Eqs. (6-107) to (6-110) and the circuit acts like an inductive reactance[1] in parallel with (or in series with) a negative resistance, the value of which may be made low by the use of a high-gain amplifier. If the condenser C in the circuit of Fig. 6-51 is replaced by an inductance, the circuit acts like an effective capacitance in parallel with (or in series with) a resistance.

[1] See footnote 1, p. 214.

Figure 6-52 shows another circuit by means of which the magnitude or phase angle of an impedance may be varied with the aid of a vacuum tube. Applications of this circuit are limited for the following reasons: It does not give an inductive reactance when x is capacitive, it cannot be used to produce a negative resistance, the effective capacitance is always less than the capacitance of x, and the effective shunting resistance cannot be made infinite.

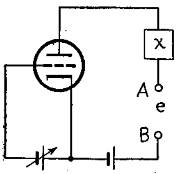


FIG. 6-52.—Series reactance-tube circuit.

Problems

6-1. Show that the voltage amplification of a single-stage amplifier with nonreactive load r_b and self-biasing resistor R_{cc} without a by-pass condenser (see Fig. 4-30*a*) is

$$A = -\frac{\mu r_b}{r_p + r_b + (\mu + 1)R_{cc}} \tag{6-113}$$

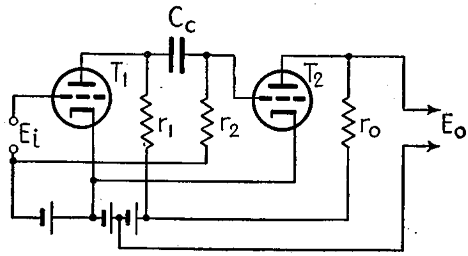


FIG. 6-53.—Diagram for Prob. 6-2.

6-2. *a.* Draw the equivalent circuits for the amplifier of Fig. 6-53.

b. Neglecting interelectrode capacitances, derive a general expression for the voltage amplification of each stage.

c. T_1 and T_2 are type 6J5 tubes, operated at a plate supply voltage of 250 volts and a bias of -6 volts, and the circuit elements have the following values:

$$r_1 = 50{,}000 \text{ ohms} \qquad C_c = 0.005\ \mu\text{f}$$
$$r_2 = 250{,}000 \text{ ohms} \qquad r_o = 50{,}000 \text{ ohms}$$

Find the gain in decibels of each stage and of the amplifier at 60, 100, and 1000 cps. Check by means of Fig. 6-14.

d. Repeat (*c*), substituting type 6SJ7 tubes for type 6J5. Use the values $\mu = 1500$ and $r_p = 1.5$ megohms.

e. Discuss the suitability of these amplifiers for use at frequencies between 50 and 3000 cps.

f. Determine the size of the cathode resistors necessary to provide the required bias when type 6J5 tubes are used.

6-3. Design a two-stage resistance-capacitance–coupled voltage amplifier using 6SF5 tubes, which will give the highest voltage amplification consistent with uniform response in the range from 100 to 10,000 cps. The output of the amplifier is to be applied to the deflecting plates of a cathode-ray oscillograph. The capacitance of these plates and their leads is 5 $\mu\mu$f.

6-4. *a.* Plot a frequency-response curve for the amplifier of Prob. 6-2, using type 6J5 tubes and the circuit constants listed in part (*c*) and taking interelectrode capacitances into account. The output of the second stage is shunted by a 4-$\mu\mu$f capacitance. $C_{gk} = C_{gp} = 3.4\ \mu\mu$f; $C_{pk} = 3.6\ \mu\mu$f.

b. What should be done in order to reduce frequency distortion in this amplifier?

6-5. By graphical methods determine the maximum crest fundamental voltage output and the voltage amplification of a type 6J5 triode in class A1 operation under the following operating conditions:

$$E_{bb} = 250 \text{ volts}$$
$$E_c = -6 \text{ volts}$$
$$r_b = R_b = 100{,}000 \text{ ohms}$$

The harmonic content must not exceed 5 per cent. Grid current starts flowing at $e_c = -0.5$ volt.

6-6. *a.* By graphical methods determine the maximum crest fundamental voltage output and the voltage amplification of a type 6SF5 triode in class A1 operation under the following operating conditions:

$$E_{bb} = 300 \text{ volts}$$
$$E_c = -1.5 \text{ volts}$$
$$r_b = R_b = 500{,}000 \text{ ohms}$$

The harmonic content must not exceed 5 per cent. Grid current starts flowing at $e_c = -0.75$ volt.

b. Repeat (*a*) for a 6SF5 tube in the first stage of the circuit of Fig. 6-52.

$$E_{bb} = 300 \text{ volts} \qquad r_1 = 500{,}000 \text{ ohms}$$
$$E_c = -1.5 \text{ volts} \qquad r_2 = 1{,}000{,}000 \text{ ohms}$$
$$f = 1000 \text{ cps} \qquad C_c = 0.01\ \mu\text{f}$$

6-7. *a.* By graphical methods find the voltage amplification of a type 6SF5 triode for the following operating conditions:

$$E_{bb} = 250 \text{ volts} \qquad r_b = R_b = 1 \text{ megohm}$$
$$E_c = -1.5 \text{ volts} \qquad E_{gm} = 1 \text{ volt}$$

b. Find r_p and μ at the operating point and find A by the use of Eq. (6-1).

c. Find the required resistance of the self-biasing resistor.

d. By means of Eq. (6-113) of Prob. 6-1 find the voltage amplification when the self-biasing resistor is used without a by-pass condenser.

6-8. *a.* Assuming that $r_2 >> r_1 + j\omega L$, show that the high-frequency amplification of the circuit of Fig. 6-17*c* is

$$A_h = \frac{\mu(r_1 + j\omega L)}{j\omega C_2 r_p(r_1 + j\omega L) + r_p + r_1 + j\omega L} \tag{6-114}$$

b. Show that if $r_p >> r_1 + j\omega L$, Eq. (6-114) reduces to

$$A_h = g_m \frac{r_1 + j\omega L}{j\omega C_2 r_1 - \omega^2 L C_2 + 1} \tag{6-115}$$

c. Transform Eq. (6-115) into Eq. (6-30).

6-9. An audio-frequency interstage coupling transformer has the following constants:

Primary inductance L_1	25 henrys
Equivalent leakage inductance, referred to the secondary, L_e	2 henrys
Equivalent distributed capacitance, referred to the secondary, C_e	80 $\mu\mu$f
Primary resistance r_1	850 ohms
Secondary resistance r_2	10,000 ohms
Turn ratio n	3

a. This transformer is used to couple a type 6J5 tube to a type 2A3 tube. The operating voltages of the 6J5 are $E_{bo} = 250$ volts and $E_c = -8$ volts. Construct a frequency-response curve for the first tube and transformer.

b. Show that excessive frequency distortion results when the type 6J5 tube is replaced by a type 6SJ7 tube used as a pentode.

6-10. *a.* For the transformer whose constants are given on page 190 determine the coefficient of coupling that will give a 10-kc separation of resonance peaks with a type 6SK7 tube.

b. Find the voltage amplification corresponding to the peaks and to the trough of the response curve when the control-grid voltage is -3 volts.

6-11. Design a two-stage amplifier for use with an electromagnetic oscillograph that requires a crest alternating current of 50 ma for full deflection. Resistance of the element is 5 ohms. In order to make possible the study of transients at very low frequency, the response should be uniform down to zero frequency. Full deflection must be obtained with an input voltage that does not exceed 1 volt, and amplitude distortion must not exceed 5 per cent. Specify all circuit constants and voltages.

6-12. The terminal voltage of a 120-volt d-c generator is regulated by applying the voltage to the input of a two-stage amplifier in such a manner that decrease of terminal voltage increases the field excitation and thus raises the voltage. The action is as follows: If the terminal voltage tends to fall by 1.05 volts with increase of load, the amplifier increases the field current sufficiently to raise the induced voltage by 1 volt, so that the net change in terminal voltage is only 0.05 volt. The average field current is 50 ma. In the normal range of field current a 1-ma change results in a 0.1-volt change of induced voltage.

a. Choose a suitable tube to control the 50-ma field current.

Field resistance = 2000 ohms.

b. Determine the change of grid voltage of this tube that is necessary in order to raise the induced voltage by 1 volt.

c. Choose a first-stage amplifier circuit and tube such that the 0.05-volt drop in terminal voltage will give the necessary change in the grid voltage of the second tube.

d. Draw a circuit diagram, specifying all circuit constants, tube operating voltages, and supply voltages.

6-13. Construct a polar diagram for the circuit of Fig. 6-41*a* used with the following tubes and circuit constants:

T_1 6J7	$r_1 = 500{,}000$ ohms	$C_f = 0.1$ μf
T_2 6F6 (pentode)	$r_2 = 1$ megohm	$r_f = 100{,}000$ ohms
$z_b = r_b = 7000$ ohms	$C_c = 0.01$ μf	$r_f' = 1000$ ohms

6-14. The amplifier shown in Fig. 6-54 "motorboats" (oscillates at a frequency of 2 or 3 cps) when used with a B-supply of poor regulation. The motorboating stops when the switch S is opened but not when the first tube is removed. Explain.

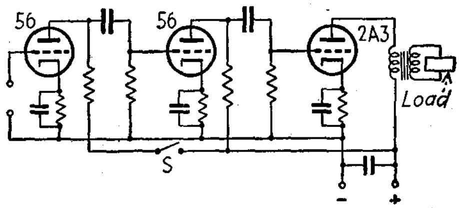


FIG. 6-54.—Diagram for problem 6-14.

6-15. Design a voltage amplifier incorporating cathode-follower stages and having a voltage gain of 85 db from 100 cps to 500 kc and an effective input capacitance not to exceed 3 $\mu\mu$f.

6-16. *a.* Show that the alternating grid voltage in the circuit of Fig. 6-49*b* is

$$E_g = \frac{a^4 - a^2 + 3ja^3}{a^4 + 7a^2 + 1} V \tag{6-116}$$

where $a = r\omega C$.

b. Show that, if the product rg_m is sufficiently large in comparison with unity so that the effect of the coupling circuit upon the effective admittance is negligible, the effective admittance between A and B is

$$Y_e = \frac{(\mu + 1)a^4 + (7 - \mu)a^2 + 1 + 3j\mu a^3}{(a^4 + 7a^2 + 1)r_p} \tag{6-117}$$

c. Plot a curve of a *vs.* μ for values of a that make the effective conductance zero.

d. Plot a curve of effective capacitance C_e *vs.* μ when a is chosen so that the effective conductance is zero.

Supplementary Bibliography

The bibliography on amplifiers, particularly resistance-coupled voltage amplifiers, is too large to be listed in entirety. The following partial list may prove helpful:

Direct-coupled Amplifiers

Engineering Staff, Aerovox Corp.: *Electronics*, March, 1938, p. 42.
TATEL, H., MONCTON, H. S., and LUHR, O.: *Rev. Sci. Instruments*, **9**, 229 (1938).
KORFF, S. A.: *Rev. Sci. Instruments*, **9**, 256 (1938).
TREVIÑO, S. N., and OFFNER, F.: *Rev. Sci. Instruments*, **11**, 412 (1940).
Electronics, March, 1938, p. 42.
MÜLLER, F., and DÜRICHEN, W.: *Z. Elektrochem.*, **42**, 31 (1936); *Electronics* (abstr.), November, 1936, p. 62.

Resistance-capacitance–coupled Amplifiers

MOULLIN, E. B.: *Wireless Eng.*, **8**, 118 (1931).
THOMAS, H. A.: *Wireless Eng.*, **8**, 11 (1931).
BARCLAY, W. A.: *Wireless Eng.*, **8**, 362 (1931).
LUCK, D. G. C.: *Proc. I.R.E.*, **20**, 1401 (1932).
POUND, F. J. A.: *Wireless Eng.*, **9**, 445 (1933).
COLEBROOK, F. M.: *Wireless Eng.*, **10**, 69 (1933).
CURTIS, W. F.: *Proc. I.R.E.*, **24**, 1230 (1936).
Electronics, February, 1937, p. 29.

KELLOGG, E. W., and PHELPS, W. D.: *Electronics*, February, 1937, p. 22; February, 1938, p. 26.
EVEREST, F. A.: *Electronics*, January, 1938, p. 16; May, 1938, p. 24.
RHOAD, E. J.: *Communications*, January, 1938, p. 11.
JOHNSTONE, D. M.: *Wireless Eng.*, **15**, 208 (1938).
PREISMAN, A.: *RCA Rev.* **2**, 421 (1938).
HEROLD, E. W.: *Communications*, August, 1938, p. 11.
SHEAFFER, C. F.: *Electronics*, September, 1938, p. 14.
SEELEY, S. W., and KIMBALL, C. N.: *RCA Rev.*, **3**, 290 (1939).
BEDFORD, A. V., and FREDENDALL, G. L.: *Proc. I.R.E.*, **27**, 277 (1939).
POLLACK, D.: *Electronics*, April, 1939, p. 38.
PREISMAN, A.: *RCA Rev.*, **3**, 473 (1939).
NIMS, A. A.: *Electronics*, May, 1939, p. 23.
RAMO, S.: *Communications*, May, 1939, p. 16.
EPHRAIM, B.: *Communications*, June, 1939, p. 12.
WHEELER, H. A.: *Proc. I.R.E.*, **27**, 429 (1939).
LIU, Y. J., and TRIMMER, J. D.: *Electronics*, September, 1939, p. 35.
FARRINGTON, J. F.: *R.M.A., Eng.*, **4**, 13, (1939).
FREEMAN, W. H.: *Electronics*, January, 1940, p. 35.
EVEREST, F. A., and JOHNSTON, H. R.: *Proc. I.R.E.*, **28**, 71 (1940).
HAEFF, A. V.: *Proc. I.R.E.*, **28**, 126 (1940).
BARCUS, L. M.: *Electronics*, June, 1940, p. 44.
STRONG, C. E.: *Electronics*, July, 1940, p. 14.
TRIMMER, J. D., and LIU, Y. J.: *Electronics*, July, 1940, p. 22.
FARRINGTON, J. F.: *Communications*, September, 1940, p. 3.

Transformer-coupled Amplifiers

DIAMOND, H., and WEBB, J. S.: *Proc. I.R.E.*, **15**, 767 (1927).
AUGHTIE, F., and COPE, W. F.: *Wireless Eng.*, **8**, 177 (1931).
KLIPSCH, P. W.: *Proc. I.R.E.*, **24**, 219 (1936).

Transient Response of Amplifiers

OATLEY, C. W.: *Wireless Eng.*, **8**, 307 (1931).
SCHLESINGER, K.: *Hochfrequenztech. u. Elektroakustik.*, **38**, 144 (1931).
SMITH, C. H.: *Wireless Eng.*, **10**, 296 (1933).
KROKER, G., and FELGENTREU, S.: *T.F.T.*, **22**, 277 (1933).
NOWOTNY, W.: *Arch. Elektrotech.*, **27**, 832 (1933); *Wireless Eng.* (abstr.), **10**, 340 (1933), **11**, 147 (1934).
BUILDER, G.: *Wireless Eng.*, **12**, 246 (1935).
PUCKLE, O. S.: *Wireless Eng.*, **12**, 251 (1935).

Low-frequency Amplifier Using High-frequency Carrier

BLACK, L. J., and SCOTT, H. J.: *Proc. I.R.E.*, **28**, 269 (1940).

Radio-frequency Amplifiers

LOFTIN, E. H., and WHITE, S. Y.: *Proc. I.R.E.*, **14**, 605 (1926).
SMITH, V. G.: *Proc. I.R.E.*, **15**, 525 (1927).
WHEELER, H. A., and MACDONALD, W. A.: *Proc. I.R.E.*, **19**, 738, 804 (1931).
BAYLY, B. DE F.: *Proc. I.R.E.*, **19**, 873 (1931).
NELSON, J. R.: *Proc. I.R.E.*, **19**, 1233 (1931).
POLYDOROFF, W. J.: *Proc. I.R.E.*, **12**, 690 (1933).
LANGLEY, R. H.: *Electronics*, July, 1931, p. 8.
CROSSLEY, A.: *Electronics*, November, 1933, p. 298.

Feedback Amplifiers

PETERSON, E., KREER, J. G., and WARE, L. A.: *Proc. I.R.E.*, **22,** 1191 (1934).
MARINESCO, M.: *Wireless Eng.*, **12,** 375 (1935).
DOME, R. B.: *R.M.A. Engineer*, November, 1936, p. 2.
FRITZINGER, G. H.: *Proc. I.R.E.*, **26,** 207 (1938).
REID, D. G.: *Wireless Eng.*, **14,** 588 (1937).
BRAYSHAW, G. S.: *Wireless Eng.*, **14,** 597 (1937).
FISHER, C. B.: *R.M.A. Eng.*, November, 1937, p. 19 (with extensive bibliography).
SCHADE, O. H.: *Proc. I.R.E.*, **26,** 137 (1938).
FARREN, L. I.: *Wireless Eng.*, **15,** 23 (1938).
FROMMER, J.: *Wireless Eng.*, **15,** 20, 90 (1938).
MARINESCO, M.: *Wireless Eng.*, **15,** 21 (1938).
BARTLETT, A. C.: *Wireless Eng.*, **15,** 90 (1938).
STEVENS, B. J.: *Wireless Eng.*, **15,** 143 (1938).
FRITZINGER, G. H.: *Proc. I.R.E.*, **26,** 207 (1938).
GINZTON, E. L.: *Proc. I.R.E.*, **26,** 1367 (1938).
MAYER, H. F.: *Proc. I.R.E.*, **27,** 213 (1939).
YOUNG, L. G.: *Electronics*, August, 1939, p. 20.
BARTELS, H.: *Electronics*, January, 1940, p. 74 (abstr.).
PEDERSON, P. O.: *Proc. I.R.E.*, **28,** 59 (1940).
STEWART, H. H., and POLLOCK, H. S.: *Electronics*, February, 1940, p. 19.
RIVLIN, R. S.: *Wireless Eng.*, **17,** 298 (1940).
BODE, H. W.: *Bell System Tech. J.*, **19,** 421 (1940).
SANDEMAN, E. K.: *Wireless Eng.*, **17,** 342 (1940).

Current Fluctuations in Tubes and Resistances

FRY, T. C.: *J. Franklin Inst.*, **199,** 203 (1925).
HULL, A. W., and WILLIAMS, N. H.: *Phys. Rev.*, **25,** 147 (1925); **27,** 432 (1926).
JOHNSON, J. B.: *Phys. Rev.*, **32,** 97 (1928).
NYQUIST, H.: *Phys. Rev.*, **32,** 110 (1928).
LLEWELLYN, F. B.: *Proc. I.R.E.*, **18,** 243 (1930).
CASE, N. P.: *Proc. I.R.E.*, **19,** 963 (1931) (with bibliography of 20 items).
THATCHER, E. W.: *Phys. Rev.*, **40,** 114 (1932).
WILLIAMS, N. H., and THATCHER, E. W.: *Phys. Rev.*, **40,** 121 (1932).
MCNALLY, J. O.: *Proc. I.R.E.*, **20,** 1263 (1932).
MOULLIN, E. B., and ELLIS, H. W. M.: *Wireless Eng.*, **11,** 21 (1934).
JOHNSON, J. B., and LLEWELLYN, F. B.: *Elec. Eng.* (survey with bibliography), **53,** 1449 (1934).
PEARSON, G. L.: *Bell Lab. Record*, **14,** 56 (1935).
STINCHFIELD, J. M.: RCA Application Note 25.
HAMILTON, H. G.: *Electronics*, August, 1937, p. 26.
ROWLAND, E. N.: *Proc Cambridge Phil. Soc*, **32,** 580 (1936); **33,** 344 (1937).
SCHOTTKY, W.: *Z. Physik*, **104,** 248 (1937).
SCHOTTKY, W., SPENKE, E., JACOBY, H., and KIRCHGESSNER, L.; *Wiss. Veröff. a. d. Siemens-Werken*, **16,** 1 (1937).
PERCIVAL, W. S., and HORWOOD, W. L.: *Wireless Eng.*, **15,** 128, 202, 213, 268, 322, 440 (1938).
RACK, A. J.: *Bell System Tech. J.*, **17,** 592 (1938).
SILVERMAN, D.: *Electronics*, February, 1939, p. 34 (an extensive bibliography covering noise and noise measurement).
JANSKY, K. G.: *Proc. I.R.E.*, **27,** 763 (1939).
THOMPSON, B. J.: *RCA Rev.*, **4,** 269 (1940).
NORTH, D. O.: *RCA Rev.*, **4,** 441 (1940); **5,** 106 (1940); **5,** 244 (1940).

CHAPTER 7

CLASS A AND CLASS AB1 POWER AMPLIFIERS

7-1. Methods of Analysis of Power Amplifiers.—The amplitude of the alternating plate current of power-amplifier tubes under rated conditions of operation is so great that nonlinear distortion is not negligible. A rigorous analysis based upon the series expansion for plate current should, therefore, take into account the second- and higher-order terms. By considering only the first term of the series, however, it is possible to derive a number of approximate relations, which, although they do not yield accurate numerical results, are useful in making a qualitative study of the effects of various circuit and tube factors upon the performance of Class A power amplifiers. An examination of Eqs. (3-41) and (3-57) shows that the second- and higher-order terms of the series expansion are absent only when r_p and μ are constant. Neglecting the second- and higher-order terms of the series expansion is therefore equivalent to assuming the static plate characteristics to be linear, parallel, and equidistant. Much of this chapter will be devoted to the derivation of useful approximate relations based upon the first term of the series expansion and upon ideal static plate characteristics of constant slope and spacing. Considerable error may result from the application of these formulas to numerical problems. For this reason graphical analyses or, preferably, laboratory measurements must be made when accurate results are essential or when it is necessary to determine nonlinear distortion.

As explained in Chap. 4, graphical analyses become very complicated when load reactance is taken into consideration. Although loud-speakers and other ordinary loads are not nonreactive, the assumption is usually made that the reactive component of load impedance may be neglected and the path of operation drawn as a straight line of negative slope equal to the reciprocal of the effective a-c load resistance. This assumption is justified because it leads to results which, although not accurate, are of great value in the analysis and design of amplifiers and which could otherwise be obtained only with great difficulty. Beyond Eq. (7-9), the treatment presented in this chapter is based upon nonreactive load impedance. For methods that take reactance into consideration, the student will find it profitable to refer to the literature on this subject.[1]

[1] Green, E., *Wireless Eng.*, **3**, 402, 469 (1926); Ardenne, von, M., *Proc. I.R.E.*, **16**, 193 (1928); Barclay, W. A., *Wireless Eng.*, **5**, 660 (1928); Whitehead, C. C.,

7-2. Power Relations in Vacuum-tube Plate Circuits.—Although the vacuum tube is replaced by an equivalent generator in series with the plate resistance in analyses based upon the equivalent plate circuit, power in an actual plate circuit can come from only one source, the B supply. The power supplied to the plate circuit by the source of direct plate voltage is

$$P_i = \frac{1}{T}\int_0^T E_{bb}i_b\,dt = E_{bb}\frac{1}{T}\int_0^T i_b\,dt = E_{bb}I_{ba} \tag{7-1}$$

where T is the period of the fundamental component of plate current. This input power is converted into a-c and d-c power developed in the load and into plate dissipation (heat) in the tube. This fact may be stated in the form of the equation

$$P_i = P_p + P_o + I_{ba}{}^2R_b \tag{7-2}$$

in which P_p is the plate dissipation, P_o is the a-c output power developed in the load, and $I_{ba}{}^2R_b$ is the d-c power developed in the load.

Equation (7-2) may be rewritten in the form

$$P_p = P_i - P_o - I_{ba}{}^2R_b \tag{7-3}$$

which states that the plate is not called upon to dissipate the portion of the input power that appears as output. If the input power remains essentially constant, the plate dissipation must rise as the output power is decreased by decreasing the excitation voltage. This is true in Class A amplifiers, in which the average plate current I_{ba} under excitation does not differ greatly from the quiescent plate current I_{bo}. Plate dissipation in Class A operation is thus a maximum when the excitation voltage is zero. At zero excitation, I_{ba} becomes I_{bo} and P_i, as given by Eq. (7-1), becomes $E_{bb}I_{bo}$. Equation (7-3) then becomes

$$\text{Max. Class A} \quad P_p = I_{bo}E_{bb} - I_{bo}{}^2R_b = I_{bo}(E_{bb} - I_{bo}R_b) \tag{7-4}$$

By means of Eq. (3-30), Eq. (7-4) may be transformed into

$$\text{Max. Class A} \quad P_p = I_{bo}E_{bo} \tag{7-5}$$

In Class A amplification, therefore, the plate dissipation will always be within the allowable value if the *zero-excitation* plate dissipation $I_{bo}E_{bo}$ does not exceed the rated maximum plate dissipation. In Class B and Class C operation, on the other hand, the power input falls more rapidly

---

Wireless Eng., **10,** 78 (1933); SÖCHTING, F., *Wireless Eng.* (abstr.), **10,** 165 (1933); PREISMAN, A., *RCA Rev.*, **2,** 124, 240 (1937) (with bibliography); JONKER, J. L. H., *Wireless Eng.*, **16,** 274, 344, (1939); FAIRWEATHER, A., and WILLIAMS, F. C., *Wireless Eng.*, **16,** 57 (1939). PREISMAN, A., "Graphical Construction for Vacuum Tube Circuits," McGraw-Hill Book Company, Inc., New York, 1943.

than the power output as the excitation is reduced, becoming zero at zero excitation, and so the *full-excitation* plate dissipation must not exceed the allowable plate dissipation. As explained in Chap. 2, the energy lost in plate dissipation first appears as kinetic energy of the electrons, which is in turn converted into heat when the electrons strike the plate. The allowable dissipation depends upon the maximum temperature that the plate can acquire without giving off absorbed gas or emitting electrons. When not definitely specified, the maximum allowable plate dissipation of small receiving tubes may be assumed to equal the product of the maximum values of rated operating plate current and plate voltage.

7-3. Plate-circuit Efficiency.—*Plate-circuit efficiency* is defined as the ratio of the fundamental a-c power output to the power input of the plate circuit

$$\eta_p = \frac{P_o}{P_i} \tag{7-6}$$

Plate-circuit efficiency is important for two reasons. First, increase of plate-circuit efficiency may result in reduction of first cost of an amplifier and associated power supply. Since the allowable plate dissipation is related to tube size, the required size of tube tends to decrease with increased plate efficiency. Increased efficiency reduces the required capacity of the B supply. Secondly, economy of operation increases with plate-circuit efficiency, both because of reduction of power input to the plate circuit, and because of possible reduction of cathode heating power as the result of reduced tube size. Economy of operation is of particular importance in battery-operated amplifiers and in very large a-c–operated amplifiers.

In a power amplifier, the load is usually coupled to the plate by means of an output transformer or a choke-condenser filter. Since the primary resistance of a well-designed output transformer or the resistance of an output choke is only a few hundred ohms, the d-c power $I_{ba}{}^2R_b$ developed in the load, which serves no useful purpose, is ordinarily negligible. The output transformer or filter, therefore, not only prevents any objectionable effects that might result from the flow of direct current through the load, but also increases the plate-circuit efficiency.

It is readily proved that the maximum theoretical efficiency of Class A amplifiers with transformer-coupled nonreactive load is 50 per cent when the distortion is negligible. In Fig. 7-1 is shown the path of operation of a Class A amplifier with transformer-coupled nonreactive load. Since the discussion is not limited to any special type of tube, no static characteristics are shown. If the distortion is negligible, the waves of plate current and voltage are sinusoidal, and the power output is $\frac{1}{2}I_{pm}E_{pm}$. Since distortion is assumed negligible, I_{ba} is equal to I_{bo}. If the d-c resistance of the plate circuit is negligible, E_{bb} is equal to E_{bo}. The power input, as given by Eq. (7-1), is $I_{bo}E_{bo}$. Equation (7-6) gives the plate-circuit efficiency as $I_{pm}E_{pm}/2I_{bo}E_{bo}$. As the path of operation is extended, I_{pm} and E_{pm} increase. Their maximum values are obtained when the path extends to the current axis in one direction and the voltage axis in the other. Then the plate current is just reduced to zero at the negative peak of excitation voltage, and the plate voltage is reduced to zero at the positive peak of excitation voltage. Under this condition of operation $I_{pm} = I_{bo}$ and $E_{pm} = E_{bo}$. The plate-circuit efficiency is then $\frac{1}{2}$, or 50 per cent. Actual Class A efficiencies are less than 50 per cent (see Secs. 7-10 and 7-15). Class B and C efficiencies may exceed 50 per cent.

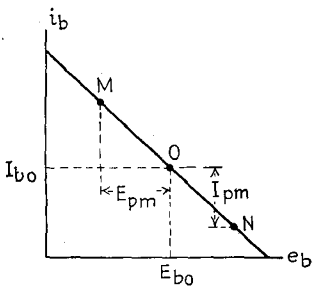


FIG. 7-1.—Generalized Class A path of operation.

7-4. Power Output.—If only the first term of the series expansion for plate current is considered and V_p is zero the r-m-s plate current is

$$I_p = \frac{\mu E_g}{r_p + z_b} \tag{7-7}$$

The approximate power output is

$$P_o = I_p{}^2 r_b = \frac{\mu^2 E_g{}^2 r_b}{(r_p + r_b)^2 + x_b{}^2} \tag{7-8}$$

Equation (7-8) may be written in the form

$$P_o = g_m \mu E_g{}^2 \frac{m \cos \theta}{m^2 + 2m \cos \theta + 1} \tag{7-9}$$

where $\cos \theta = r_b/|z_b|$ and $m = r_p/|z_b|$. Equation (7-9) indicates that tubes used to deliver power should have a high product of amplification factor by transconductance. It also shows that the power output depends upon the load power factor and upon the ratio of the load impedance to the plate resistance and is proportional to the square of the excitation voltage. As would be expected, the power output at a load impedance of given magnitude is greatest for nonreactive load, for which $\cos \theta = 1$.

Differentiation of Eq. (7-9) with respect to m, all other quantities being held constant, shows that the power output is a maximum at a fixed value of grid swing when $z_b = r_p$. From the definition of power sensitivity it follows that this is the criterion for maximum power sensitivity. It should be noted that this criterion for maximum power output holds only when the grid swing is constant and sufficiently small so that nonlinear distortion is negligible. In most applications of power amplifiers the grid swing is large enough so that nonlinear distortion is not negligible. It will be shown that the grid swing must then be varied with load resistance in order to obtain the greatest output consistent with allowable distortion, and that the best value of load impedance is not equal to the plate resistance.

7-5. Optimum Power Output.—*Optimum power output* is defined broadly as the greatest power output that can be obtained at a given operating plate voltage, subject to limiting restrictions on operation. The load resistance that gives optimum power is called *optimum load resistance.* Restrictions on operation that must be taken into consideration in the determination of optimum power and optimum load resistance include allowable nonlinear distortion, allowable grid current, available grid-circuit input power, and grid and plate dissipation.

7-6. Class A Power Amplifiers.—The flow of grid current involves a number of difficulties, which will be discussed in detail in Chap. 8. In Class A amplifiers the objectionable effects of the flow of grid current more than offset the advantages resulting from the increase in excitation voltage that may be obtained by allowing the grid to swing positive. For this reason Class A2 operation is seldom used. It will not be discussed further.

In Class A1 amplifiers the operating restrictions are that nonlinear distortion and plate dissipation must not exceed the allowable values and that grid current must not flow. Since the nonlinear distortion at a given grid swing depends upon the load resistance and the location of the operating point, the maximum excitation voltage that can be used at a given operating plate voltage, without exceeding the allowable nonlinear distortion, varies with load resistance and with grid bias. Under the assumption that grid current starts flowing at zero grid voltage,[1] the grid swing must not exceed the grid bias. The limiting grid swing at a given operating plate voltage and load resistance is greatest when the grid bias is adjusted so that the nonlinear distortion is equal to the allowable value when the grid swing is equal to the grid bias. This value of grid swing will be called *full grid swing,* and the corresponding power output *full power output.*

Determination of optimum Class A power output and load resistance is complicated by the dependence of nonlinear distortion upon load resistance. As the load resistance is varied, the grid bias and excitation voltage must be adjusted so as to maintain full grid swing. Optimum

[1] As pointed out in Secs. 3-4 and 6-8, grid current may start flowing when the grid is slightly negative. Because the voltage at which grid current starts flowing is small in comparison with the bias of power tubes, the error involved in this assumption is ordinarily small.

Class A1 power output is defined as the greatest power output that can be obtained when the grid bias and excitation voltage are adjusted so that the harmonic content has a specified value, grid current does not flow, and allowable plate dissipation is not exceeded.[1] Usually 5 per cent is accepted as maximum allowable harmonic content.[2]

7-7. Triode Class A1 Power Amplifiers.—The harmonic content of the alternating plate current of single-sided triode Class A amplifiers is almost entirely second harmonic. That the grid bias and excitation voltage of a Class A1 triode power amplifier must be varied with load resistance in order to maintain full power output can therefore be shown by means of Eq. (4-37), which gives the percentage second harmonic when the third and higher harmonics are negligible.

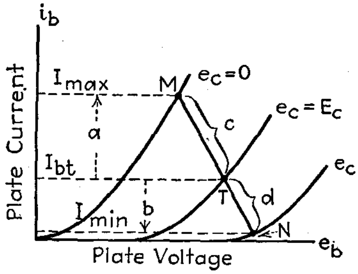


FIG. 7-2.—Triode plate diagram for Class A1 operation.

$$\text{Percentage } H_2 = \frac{I_{\max} + I_{\min} - 2I_{bt}}{2(I_{\max} - I_{\min})} \times 100 \qquad (4\text{-}37)$$

In Fig. 7-2 the ratio of the positive to the negative swing of plate current is

$$\frac{a}{b} = \frac{I_{\max} - I_{bt}}{I_{bt} - I_{\min}} \qquad (7\text{-}10)$$

Combining Eqs. (4-37) and (7-10) gives

$$\text{Percentage } H_2 = 100\,\frac{\frac{a}{b} - 1}{2\left(\frac{a}{b} + 1\right)} \qquad (7\text{-}11)$$

Equation (7-11) shows that the second-harmonic distortion increases with the ratio a/b. The shape of triode characteristics is such that the ratio a/b, and hence the second-harmonic distortion, increases as the operating point is lowered at constant load resistance and decreases as the load line is made less steep by increasing the load resistance with a fixed operating point.

As the load resistance is increased at a given value of plate supply voltage, the operating point can be lowered and the grid swing increased without exceeding the allowable distortion. This can be explained by

[1] Thus optimum P_o is the maximum full P_o obtained as r_b is varied.

[2] MASSA, FRANK, *Proc. I.R.E.*, **21**, 682 (1933).

reference to Fig. 7-3, in which, for simplicity, shift of the operating point with change of excitation is neglected. Let MN be the original load line, for which the distortion has the limiting value. Then let the load resistance be increased so that the new path of operation is $M''N''$. Because the static characteristics of a triode are nearly parallel, the ratio a/b is more nearly equal to unity for the new load line, and so the amplitude distortion is less than the allowable value. The operating point may, therefore, be lowered to O' by increasing the negative grid bias to the value E_c', giving a lower path of operation $M'N'$, for which nonlinear distortion is equal to the limiting value. This has two beneficial effects. (1) The increase of bias increases the allowable grid swing from the value E_c to the new value E_c' and thus raises the power output obtained with the new resistance over the value obtained with the new resistance and original bias. (2) The reduction of operating plate current decreases the plate dissipation and increases the plate-circuit efficiency.

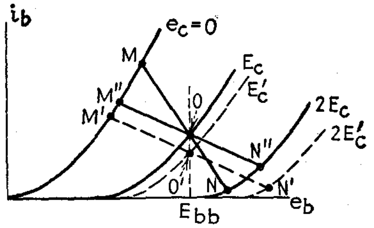


FIG. 7-3.—Triode plate diagram showing the necessity of varying bias and grid swing with load resistance.

7-8. Theoretical Value of Optimum Load and Optimum Power Output of Class A1 Triode with Transformer-coupled Resistance Load and Specified Plate Supply Voltage.[1]—Several investigators have derived expressions for optimum load resistance and optimum power output based upon empirical equations for plate current. Because no single simple equation accurately expresses the plate current throughout the range of operation, most of these derivations have proved to be of little practical value. More useful approximate expressions have been derived under the assumption that the static characteristics are linear throughout the operating range. Although accurate determinations can be made only by experimental or graphical methods, these expressions are often of considerable value in making rough computations and in predicting the general effect of various parameters upon the performance of power amplifiers.

In the determination of optimum operating conditions it is possible either that the operating plate voltage is specified at a fixed value, or that it may be adjusted to the optimum value consistent with allowable plate dissipation and distortion. In the design of amplifiers using relatively small tubes, the operating plate voltage is usually fixed. This case will be analyzed first.

[1] Since condenser-choke coupling, illustrated in Fig. 5-17, is equivalent to coupling through a transformer having unity turn ratio, Sec. 7-8 and subsequent sections also apply to condenser-choke–coupled load.

The allowable plate dissipation determines the maximum operating plate current that can be used at a given operating plate voltage and must, therefore, be taken into consideration in determining the optimum load resistance, bias, and power output. Often, however, the allowable dissipation is high enough so that the operating point and load resistance are determined entirely by the considerations of amplitude distortion, flow of grid current, and power output. It is convenient, therefore, first to assume that the allowable plate dissipation is so high that it does not affect the choice of operating point, and then to see in what manner the procedure must be modified if the allowable dissipation limits the plate current to a value less than that determined on the basis of harmonic content, flow of grid current, and power output.

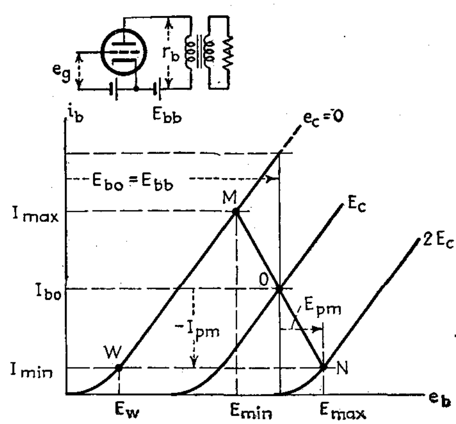


FIG. 7-4.—Idealized triode plate diagram used in determining theoretical optimum load resistance and grid bias at fixed E_{bo}.

One of the most frequently used rules for the choice of load resistance in triode power amplifiers with transformer-coupled load was first given by W. J. Brown.[1] Brown's derivation is based upon the assumptions that the operating plate voltage is fixed and equal to the plate supply voltage, that the minimum plate current I_{min} is not varied with load resistance, that the plate characteristics are straight, equidistant, parallel lines, and that the allowable plate dissipation is large enough so that it does not affect the choice of operating point.[2] The following derivation is a modification of that of Brown.

Assuming that the plate characteristics are straight, equidistant, parallel lines is equivalent to assuming that the plate resistance and

[1] BROWN, W. J., *Proc. Phys. Soc. London*, **36**, 218 (1924). See also NOTTINGHAM, W. B., *Proc. I.R.E.*, **29**, 620 (1941).

[2] Since the tube characteristics are assumed to be linear, the static and dynamic operating points coincide.

amplification factor are constant. The plate current is therefore given by the first term of the series expansion.

$$I_{pm} = \frac{\mu E_g}{r_p + r_b} \tag{7-12}$$

At full grid swing, the grid swing is equal to the bias. Therefore, at full grid swing,

$$I_{pm} = \frac{\mu E_c}{r_p + r_b} \tag{7-13}$$

Inspection of Fig. 7-4 shows that

$$E_{\max} - E_{bo} = E_{pm} = -I_{pm}r_b = -\frac{\mu E_c r_b}{r_p + r_b} \tag{7-14}$$

Points W and N have the same value of i_b but differ in grid voltage by $2E_c$. For the assumed ideal characteristics, μ is constant, and

$$\frac{\partial e_b}{\partial e_c} = \frac{\Delta e_b}{\Delta e_c}$$

at constant i_b. Hence,

$$\mu = -\frac{\Delta e_b}{\Delta e_c} \text{ at const. } i_b \tag{7-15}$$

$$\mu = -\frac{E_{\max} - E_w}{2E_c - 0} \tag{7-16}$$

$$E_{\max} = E_w - 2\mu E_c \tag{7-17}$$

Substituting Eq. (7-17) in Eq. (7-14) gives

$$E_{bo} - E_w = -\mu E_c \left(2 - \frac{r_b}{r_p + r_b}\right) = -\mu E_c \frac{r_b + 2r_p}{r_b + r_p} \tag{7-18}$$

$$E_c = -\frac{E_{bo} - E_w}{\mu} \cdot \frac{r_b + r_p}{r_b + 2r_p} \tag{7-19}$$

Substitution of Eq. (7-19) in Eq. (7-13) gives

$$I_{pm} = -\frac{E_{bo} - E_w}{r_b + 2r_p} \tag{7-20}$$

The power output at full grid swing is

$$\text{Full } P_o = \tfrac{1}{2}I_{pm}^2 r_b = \tfrac{1}{2}(E_{bo} - E_w)^2 \frac{r_b}{(r_b + 2r_p)^2} \tag{7-21}$$

The power output is a maximum with respect to r_b when $dP_o/dr_b = 0$. Differentiation of Eq. (7-21) with respect to r_b shows that this is true when the load resistance is equal to twice the plate resistance. Therefore

$$\text{Optimum } r_b = 2r_p \tag{7-22}$$

Substitution of Eq. (7-22) in Eq. (7-19) gives the optimum bias,

$$\text{Optimum } E_c = -\tfrac{3}{4}\,\frac{E_{bo} - E_w}{\mu} \tag{7-23}$$

Since $I_{\min}$ does not actually remain constant as r_b is varied but may be decreased with increase of r_b, it is of interest also to derive an expression for E_c under the assumption that the ratio of $I_{\max}$ to $I_{\min}$ remains constant as r_b is varied, which is roughly true. Such a derivation yields the following equations:

$$\text{Optimum } r_b = \frac{k}{k-1}\,2r_p \tag{7-24}$$

$$\text{Optimum } E_c = -\frac{3k-1}{4k}\cdot\frac{E_{bo}}{\mu} \tag{7-25}$$

where $k = I_{\max}/I_{\min}$.

Examination of plate diagrams of triode power tubes shows that under standard operating conditions E_w, the plate voltage corresponding to $I_{\min}$ and zero grid voltage, never differs greatly from $E_{bo}/10$ for 5 per cent H_2, and that $I_{\max}/I_{\min}$ ranges from approximately 5 to 15. For these values, Eq. (7-23) calls for a bias of $-0.675(E_{bo}/\mu)$, and Eq. (7-25) a bias of from $-0.7(E_{bo}/\mu)$ to $-0.73(E_{bo}/\mu)$. Values that are checked experimentally within 5 per cent are given by the following equation, which is intermediate between Eqs. (7-23) and (7-25):

$$E_c = -0.7\,\frac{E_{bo}}{\mu} \tag{7-26}$$

The value of optimum load resistance given by Eq. (7-24) does not ordinarily differ greatly from that given by Eq. (7-22). Since both equations are approximations, therefore, it is usually preferable to use the simpler equation (7-22), even though Eq. (7-24) may approximate more closely the experimentally determined value.

Substituting Eq. (7-22) in Eq. (7-9) and letting $x_b = 0$ gives

$$P_o = \tfrac{2}{9}\mu g_m E_g^2 \tag{7-27}$$

But for optimum Class A1 amplification the grid swing is approximately equal to E_c so that $E_g \cong 0.707E_c$ and

$$\text{Optimum } P_o \cong \tfrac{1}{9}\mu g_m E_c^2 \tag{7-28}$$

From the definition of power sensitivity it follows that the theoretical power sensitivity corresponding to optimum load resistance is

$$\text{P. S. at optimum load} = \tfrac{2}{9}\mu g_m \tag{7-29}$$

In spite of the fact that the characteristics of triodes do not satisfy the hypotheses made in the foregoing analysis, Eqs. (7-21) and (7-22) are roughly verified by laboratory measurements. A curve of power output as a function of r_b/r_p as determined from Eq. (7-21) is shown in Fig. 7-5. Experimentally determined curves of power output as a function of r_b/r_p for a type 45 triode at three values of second-harmonic distortion are shown in Fig. 7-6. The similarity between the theoretical and the experimental curves is apparent. Furthermore, Fig. 7-6 shows that optimum load for the 45 tube at 5 per cent second harmonic is very nearly twice the a-c plate resistance at the operating point. It may be shown by graphical studies based on ideal static characteristic curves which follow simple power laws that curvature of the characteristics tends to reduce the optimum value of r_b and that increase of $I_{\min}$ tends to

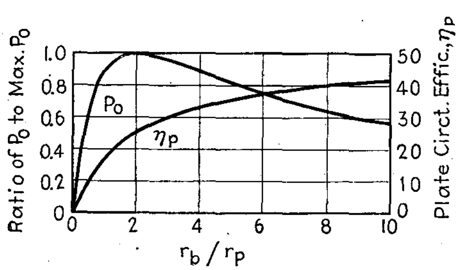


FIG. 7-5.—Theoretical curves of full power output and of plate-circuit efficiency for an ideal triode in Class A1 operation.

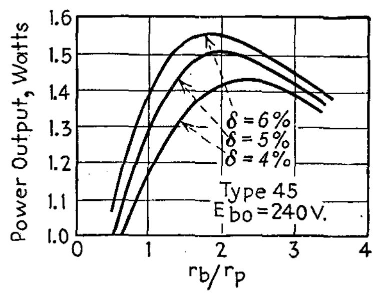


FIG. 7-6.—Experimentally determined curves of full power output of type 45 triode in Class A1 operation.

increase optimum r_b. The close agreement between Fig. 7-5 and Eq. (7-22) is probably explained by compensation of these two effects. The agreement is not so close with other types of triodes, practical values of r_b/r_p ranging from 2 to 3 for optimum output at 5 per cent second harmonic.

At first thought, Eq. (7-22) may be surprising in view of Maxwell's rule that maximum transfer of power from a generator to a load takes place when the load impedance is equal to the generator impedance. This rule assumes, however, that the voltage of the generator is not varied as the load impedance is varied, whereas in the derivation of Eq. (7-22) the exciting voltage was assumed to vary with load impedance in such a manner that the grid swing had the largest value consistent with allowable distortion and with the restriction on grid current. In Sec. 7-4 it was shown that the maximum output at fixed grid swing, or maximum power sensitivity, is obtained when the load impedance is a resistance equal to the plate resistance.

Optimum Pure Resistance Load.—An analysis similar to that used in the derivation of Eq. (7-22) shows that, if the load is a resistance inserted directly in the plate circuit, the optimum load resistance for a

Class A1 triode amplifier is equal to the plate resistance. Because of loss of direct voltage and power in the load, this type of load is seldom used.

7-9. Optimum Power Output in Terms of Operating Plate Voltage.—At optimum load resistance, Eq. (7-21) reduces to

$$\text{Optimum } P_o = \frac{(E_{bo} - E_w)^2}{16r_p} \tag{7-30}$$

$$\cong 0.05\,\frac{E_{bo}^2}{r_p} \tag{7-31}$$

Although the numerical values of optimum power indicated by Eq. (7-31) are usually not greatly in error, its principal value lies in the fact that it indicates the manner in which optimum power output of triodes depends upon plate supply voltage and plate resistance. It is evident from Eq. (7-31) that, in order to give high power output at a given plate supply voltage, a triode should have low plate resistance. Variation of plate resistance by changing the grid structure and electrode spacing of a triode is accompanied by a roughly proportional change in amplification factor. Triodes with low plate resistance therefore have low amplification factor and a low product of transconductance by amplification factor. For this reason it is difficult to obtain both high power output and high power sensitivity [see Eq. (7-29)] with triodes.

7-10. Theoretical Plate-circuit Efficiency.—Accurate determinations of plate-circuit efficiency, like those of power output, must be made by graphical methods or laboratory measurements. It is profitable, however, to determine the largest value of plate-circuit efficiency that could be obtained with idealized linear triode characteristics. To simplify the derivation, $I_{\min}$ will be assumed to be zero.

With linear characteristics, $I_{ba} = I_{bo}$, and Eq. (7-1) becomes

$$P_i = E_{bb}I_{bo} \tag{7-32}$$

Because the primary resistance of well-designed output transformers is small, E_{bb} may be replaced by E_{bo}.

$$P_i \cong E_{bo}I_{bo} \tag{7-33}$$

Figure 7-4 shows that, if $I_{\min}$ is zero, I_{bo} is equal in magnitude to I_{pm} and E_w is zero. Then from Eq. (7-20),

$$I_{bo} = \frac{E_{bo}}{r_b + 2r_p} \tag{7-34}$$

Substituting Eq. (7-34) in Eq. (7-33) gives

$$P_i = \frac{E_{bo}^2}{r_b + 2r_p} \tag{7-35}$$

But if E_w is zero, Eq. (7-21) becomes

$$P_o = \frac{\frac{1}{2}E_{bo}{}^2 r_b}{(r_b + 2r_p)^2} \tag{7-36}$$

The plate-circuit efficiency for ideal linear triode characteristics is

$$\eta_p = \frac{P_o}{P_i} = \frac{r_b}{2(r_b + 2r_p)} \tag{7-37}$$

The ideal theoretical plate-circuit efficiency, as determined from Eq. (7-37), is plotted in Fig. 7-5 as a function of r_b/r_p. It can be seen from Eq. (7-37) or Fig. 7-5 that the plate-circuit efficiency of triodes in Class A1 amplification increases with load resistance[1] and that the maximum possible ideal value of 50 per cent is obtained with infinite load, at which the power output is zero. At optimum load the plate-circuit efficiency for ideal linear characteristics is 25 per cent. The facts that the static plate characteristics are not linear and that $I_{\min}$ and E_w are not zero reduce actual values of plate-circuit efficiency of Class A1 triode power amplifiers at optimum load to values that range from about 20 to 23 per cent.

7-11. Limitation of Power Output by Plate Dissipation.—At a given value of operating plate voltage the use of optimum load determined by Eq. (7-22) and the corresponding optimum grid bias specified by Eq. (7-26) may give such a high value of static operating plate current that the zero-excitation plate dissipation exceeds the allowable dissipation. The zero-excitation plate dissipation may be reduced to the allowable value by reducing the operating plate voltage, the load resistance being maintained constant at approximately twice the plate resistance. This results in little change in plate-circuit efficiency. The zero-excitation plate dissipation may also be reduced to the allowable value by increasing the load resistance and grid bias, as explained in Sec. 7-7, and thus reducing the plate current. Equation (7-37) shows that the plate-circuit efficiency increases with increase of load resistance. This means that reduction of plate dissipation by increase of load resistance and grid bias is accompanied by less reduction in power output than when it is effected by decrease of operating plate voltage. For this reason, power triodes in Class A1 amplification, when operated at high plate voltage, are often used with an a-c load resistance that is considerably higher than twice the plate resistance. When the power output is limited by plate dissipation, the operating plate current for a given operating voltage is determined from the allowable plate dissipation, and

[1] Since the negative grid bias is increased with r_b, I_{ba} and hence P_i decrease with increase of r_b.

the load resistance is chosen to have the lowest value consistent with allowable distortion.

7-12. Graphical Determination of Power-triode Operating Conditions and Performance at a Specified Operating Plate Voltage.—The work of determining a triode load line corresponding to a given harmonic content can be greatly simplified by the use of the so-called *power-output rule*, first described by K. S. Weaver.[1] This device is based upon the fact that for a given second-harmonic content and negligible third and higher harmonics, the lengths of the portions of the load line lying above and below the operating point have a fixed ratio. This is readily proved by reference to Fig. 7-2. By similar triangles,

$$\frac{c}{d} = \frac{a}{b} \tag{7-38}$$

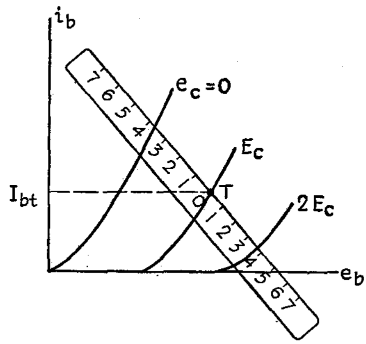


FIG. 7-7.—Method of using power-output rule in drawing the load line corresponding to optimum load.

Combining Eq. (7-38) with Eq. (7-11) gives

$$\frac{c}{d} \doteq \frac{2\delta_2 + 1}{1 - 2\delta_2} \tag{7-39}$$

where δ_2 is the ratio of the second-harmonic amplitude to the fundamental amplitude. For the commonly accepted value of 5 per cent second harmonic, Eq. (7-39) becomes

$$\frac{c}{d} = 1.22 \tag{7-40}$$

The power-output rule has two uniform scales which are marked off in opposite directions from a common origin. The length of the divisions of one of these scales is arbitrary, and the length of those of the other is 1.22 times as great. The load line corresponding to 5 per cent second harmonic at full grid swing is found by placing the origin of the rule on the operating point and turning the rule so that the intersections of the edge of the rule with the static characteristics for $e_c = 0$ and $e_c = 2E_c$ give equal readings on the two scales, as shown in Fig. 7-7. The power output is determined from the load line in the usual manner. With the aid of Eq. (7-39) it is a simple matter to construct rules for any percentage second harmonic.

Since the operating conditions specified in manufacturers' tube manuals are satisfactory for most ordinary applications of power tubes,

[1] WEAVER, K. S., *QST*, November, 1929, p. 14. See also GREENWOOD, W., *Wireless Eng.*, **8,** 429 (1931); ESPLEY, D. C. and FARREN, L. I., *Wireless Eng.*, **11,** 183 (1934), *Radio Eng.*, September, 1934, p. 20.

it is not often necessary to make determinations of triode operating points and performance. When it is desired to operate a power triode at a specified operating plate voltage that differs from values listed in tube manuals, the following procedure may be used:

1. Estimate the approximate location of the operating point and determine graphically the value of μ at this point.
2. By means of Eq. (7-26), determine the optimum grid bias.
3. Using the plate current corresponding to the given operating plate voltage and the optimum grid bias, determine the plate dissipation. If the plate dissipation is within the allowable value, the operating point is the point corresponding to these voltages. If the plate dissipation exceeds the allowable value, however, the operating point is that determined by the operating plate voltage and the maximum allowable operating plate current, which is equal to the allowable plate dissipation divided by the operating plate voltage. The magnitude of the grid bias corresponding to the operating point then exceeds the value determined by means of Eq. (7-26). It is ordinarily permissible to make an adjustment up to 5 per cent or so in the value of the grid bias in order to use a value for which a plate characteristic curve is available.
4. Using a power-output rule, draw a dynamic load line corresponding to 5 per cent second harmonic content (or other allowable value) through the operating point. If a power-output rule is not at hand, the load line may be drawn corresponding to a load resistance equal to twice the plate resistance, and the harmonic content determined graphically by means of Eq. (4-37). If the harmonic content differs considerably from the allowable value, the load line should be redrawn and again checked.
5. The dynamic operating point may be located and the true dynamic load line drawn. Ordinarily this added work is not justified by the small increase in accuracy of the results.
6. Determine the power output graphically by the method explained in Sec. 4-21. The power output may be checked roughly by the use of Eqs. (7-28) and (7-31).
7. Find the power input to the plate circuit by means of Eq. (7-1) or (7-2). Find the plate circuit efficiency by dividing the graphically determined power output by the power input.

7-13. Determination of Optimum Operating Conditions for a Power Triode with Unspecified Operating Plate Voltage.—It is sometimes necessary to determine the optimum load, bias, and performance of a power triode when the allowable plate dissipation and harmonic content are specified, but the operating plate voltage is not specified. The theoretical expression for optimum load resistance may be derived by expressing E_{bo} in terms of r_b in Eq. (7-21) and setting the derivative of P_o with respect to r_b equal to zero. The resulting expression is, however, very complicated. The following analysis is much simpler:[1]

[1] NOTTINGHAM, *loc. cit.*

Equation (7-21) may be written in the form

$$P_o = (I_{bo} - I_{\min})^2 r_b \tag{7-41}$$

For optimum P_o, $dP_o/dr_b = 0$. Therefore, since $I_{\min}$ is constant,

$$\frac{dI_{bo}}{dr_b} = -\frac{I_{bo} - I_{\min}}{2r_b} \tag{7-42}$$

$I_{pm} = I_{bo} - I_{\min}$ and, when the plate dissipation is equal to the maximum allowable value P_{pm}, $P_{pm} = E_{bo}I_{bo}$. Hence, at maximum plate dissipation, Eq. (7-20) may be written in the form

$$(I_{bo} - I_{\min})(r_b + 2r_p) = E_w - \frac{P_{pm}}{I_{bo}} \tag{7-43}$$

Differentiation of Eq. (7-43) with respect to r_b gives

$$I_{bo} - I_{\min} + (r_b + 2r_p)\frac{dI_{bo}}{dr_b} = \frac{P_{pm}}{I_{bo}^2}\cdot\frac{dI_{bo}}{dr_b} \tag{7-44}$$

Substitution of Eq. (7-42) in Eq. (7-44) gives

$$r_b = 2r_p + \frac{P_{pm}}{I_{bo}^2} \tag{7-45}$$

Substitution of Eq. (7-45) in Eq. (7-43) gives

$$\frac{P_{pm}I_{\min}}{I_{bo}^2} = 4r_p(I_{bo} - I_{\min}) - E_w \tag{7-46}$$

Equation (7-46) relates the operating plate current I_{bo} with the known values of P_{pm}, $I_{\min}$, r_p, and E_w at optimum load. In using Eq. (7-46) it is necessary to assume a reasonable value of $I_{\min}$ and to determine from the plate characteristics a value of r_p that holds approximately within the operating range. The left-hand and right-hand sides of Eq. (7-46) are then plotted against I_{bo}. The intersection of the two resulting curves gives the optimum value of I_{bo} corresponding to the given value of P_{pm} and the assumed values of $I_{\min}$ and r_p. r_b may then be found from Eq. (7-45). $E_{bo} = P_{pm}/I_{bo}$. E_c may be found from Eq. (7-19). It is advisable to check r_p at the resulting operating point against the assumed value, and to check the graphically determined harmonic content against the allowable value. If either the plate resistance or the harmonic content differs appreciably from the assumed value, the procedure may be repeated, using new values of r_p and $I_{\min}$. After the dynamic load line has been constructed, the power output and plate-circuit efficiency are found in the usual manner.[1]

[1] An alternative, often easier, method of determining optimum operating conditions at unspecified plate voltage is to apply the procedure of Sec. 7-12 at various plate voltages and plot a curve of P_o *vs.* E_{bo}.

7-14. Conversion Equations for Power Triodes.—Under the assumption that the plate current of triodes varies as the $\frac{3}{2}$ power of the plate voltage, conversion equations can be derived by means of which it is possible to make rapid approximate determinations of operating current, grid bias voltage, plate factors, and power output at any operating plate voltage when the values are known at one value of operating plate voltage. The plate current is approximated by the equation

$$i_b = Ae_b^{3/2} \tag{7-47}$$

The plate resistance is found by differentiating Eq. (7-47) with respect to e_b and evaluating at the operating point.

$$r_p = \frac{2}{3A} E_{bo}^{-1/2} \tag{7-48}$$

The transconductance is

$$g_m = \frac{\mu}{r_p} = \tfrac{3}{2}\mu A E_{bo}^{1/2} \tag{7-49}$$

The power output is

$$P_o \cong \frac{E_{bo}^2}{16r_p} = \frac{3A}{32} E_{bo}^{5/2} \tag{7-50}$$

The grid bias is given by Eq. (7-26).

From Eqs. (7-26) and (7-47) to (7-50) the following relations are readily found:

$$\left.\begin{array}{ccc} \dfrac{I_{bo}'}{I_{bo}} = \left(\dfrac{E_{bo}'}{E_{bo}}\right)^{3/2} & \dfrac{E_c'}{E_c} = \dfrac{E_{bo}'}{E_{bo}} & \dfrac{P_o'}{P_o} = \left(\dfrac{E_{bo}'}{E_{bo}}\right)^{5/2} \\ \dfrac{r_p'}{r_p} = \left(\dfrac{E_{bo}}{E_{bo}'}\right)^{1/2} & & \dfrac{g_m'}{g_m} = \left(\dfrac{E_{bo}'}{E_{bo}}\right)^{1/2} \end{array}\right\} \tag{7-51}$$

E_{bo}' and E_{bo} are any two values of operating plate voltage. In spite of the fact that triode characteristics do not obey a $\frac{3}{2}$-power law and that Eqs. (7-26) and (7-30), which are used in deriving Eq. (7-50), are based upon linear characteristics, Eqs. (7-51) are sufficiently accurate to be useful when E_{bo}' and E_{bo} do not differ too greatly. Curves based upon Eqs. (7-51) are shown on page 677.

7-15. Class A Tetrode and Pentode Power Amplifiers.—In Sec. 7-9 it was pointed out that in triodes the low plate resistance required for high power output cannot be attained without correspondingly low amplification factor, which results in low power sensitivity. This difficulty is avoided by the use of tetrodes or pentodes. Because of the greater ease of preventing with pentodes the undesirable shape of plate characteristics resulting from secondary emission, tetrodes are used comparatively little in power amplifiers. Much of the following discussion of power pentodes also applies, however, to power tetrodes.

Pentodes have two advantages over triodes in Class A1 power amplification: They have higher plate-circuit efficiency and power output and they have higher power sensitivity. The higher plate-circuit efficiency of pentodes is associated with the shape of the characteristics, which is such as to allow the path of operation to extend more nearly to zero plate voltage, as required for the ideal maximum efficiency of 50 per cent. Physically, the higher efficiency of pentodes is explained by the effect of the positive screen voltage, which reduces the plate voltage necessary to produce a given plate current and thus allows a greater percentage of the plate supply voltage to appear across the load.

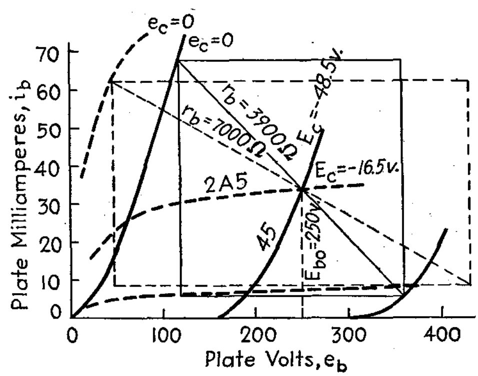


FIG. 7-8.—Comparison of plate diagrams of the type 45 power triode and the type 2A5 power pentode.

It is instructive to compare in Fig. 7-8 the plate diagram of the type 45 triode with that of the type 2A5 power pentode, which has the same values of static operating plate voltage and current. The plate voltage of the pentode is reduced to a much lower value at the peak of plate current, and the area of the rectangle enclosing the path of operation of the pentode is considerably larger than that enclosing the path of operation of the triode. Inasmuch as these areas are to a first approximation proportional to the power output, and the input plate power is the same for both, the 2A5 power output and plate-circuit efficiency are greater than those of the 45. (This analysis neglects shift of the dynamic operating points.) Efficiencies of about 35 per cent are readily obtainable with pentodes.

The higher power sensitivity of pentodes is possible because, with a given transconductance, pentodes can be designed with much higher amplification factors than triodes (see Secs. 3-10 and 3-11). In pentodes the high amplification factor results mainly from the shielding action of the screen and suppressor grids. Although these grids shield the cathode effectively from the plate, the positive screen-grid voltage produces a high positive field at the cathode, so that a comparatively high negative control-grid voltage is necessary in order to reduce the plate current to zero. The introduction of the screen and suppressor grids thus increases the amplification factor without reducing the full grid swing. Equation (7-9) shows that the combination of high amplification factor and high transconductance is necessary for high power sensitivity and that, if high amplification factor and high transconductance can be obtained without reduction of full grid swing, high power sensitivity will be obtained without sacrifice of power output.

7-16. Power Output, Distortion, and Optimum Load of Class A Pentode Power Amplifiers.—The shape of the characteristics of pentodes is such that it is not possible to derive simple theoretical equations for optimum load resistance, optimum power, and optimum bias, as was done for triodes. For this reason the correct operating conditions must be determined entirely by graphical methods or by laboratory measurement.

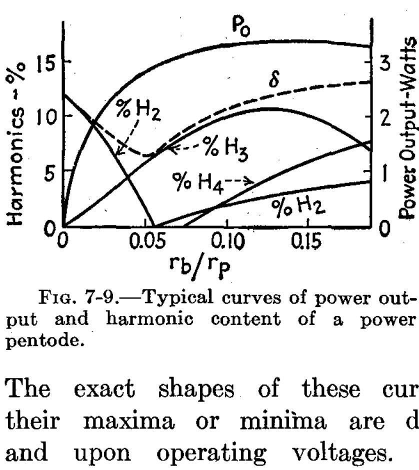


FIG. 7-9.—Typical curves of power output and harmonic content of a power pentode.

Figure 7-9 illustrates typical curves of power output and harmonic content as a function of load resistance for a suppressor pentode in single-sided Class A1 operation. The exact shapes of these curves and the relative positions of their maxima or minima are dependent upon the type of tube and upon operating voltages. The large percentage of second harmonic at low load impedance is associated with the crowding together of the plate characteristics at low plate current, which flattens the negative peak of alternating plate current, as shown by Fig. 4-16. At high load impedance, on the other hand, the large percentage of second harmonic results from the crowding together of the characteristics at low plate voltage, which flattens the positive peak of alternating plate current, as shown by Fig. 4-18. The second harmonic is zero at approximately the load resistance for which the positive and negative swings of plate current are equal. It is possible to design and operate a pentode so that the load for maximum power also gives minimum second harmonic, but this is in general not true. The form of the static plate characteristics is such that the minimum value of total harmonic content usually ranges from 6 to 10 per cent at full grid swing.

Since the load for maximum power output is, in general, different from that for minimum distortion, choice of the load must involve a compromise between large power output and small distortion. Because of the relatively high minimum distortion, distortion is ordinarily the determining factor, and the most suitable load does not differ greatly from that corresponding to zero second harmonic. It should be noted that the triode–optimum-load relation $r_b = 2r_p$ *does not hold for pentodes,* optimum power output being obtained for load resistances ranging from about $r_p/4$ to $r_p/8$. Considerable reduction in harmonic content can be achieved by using two tubes in push-pull and choosing a load resistance lower than that corresponding to zero second harmonic. The even harmonics are eliminated by the push-pull connection. By reduction of load impedance, the third harmonic may then be cut down to 5 per cent or less without excessive loss of power output (see Sec. 7-31). Further reduction of distortion may be achieved by the use of inverse feedback.

7-17. Disadvantages and Problems of Pentode Power Amplification. Nonlinear distortion is considerably more objectionable in suppressor pentodes than in triodes. This results not only from the greater distortion factor, but also from the presence of large percentages of third and higher harmonics, as well as second. Associated with the large percentage of odd harmonics is variation of power sensitivity with signal amplitude because of contributions to the fundamental output by the third and other odd terms of the plate-current series. The shape of pentode static plate characteristics is such that the coefficient of the third term of the series expansion for alternating plate current is negative. The fundamental component of plate current associated with the third term of the series expansion is therefore opposite in phase to that associated with the first term and reduces the total fundamental plate current. Since the contribution of the third term increases with the cube of the excitation voltage, whereas that of the first term increases only with the first power of the excitation voltage, this type of distortion tends to build up the amplitude of the weaker components of the output, relative to the stronger, and so affects the fidelity of reproduction. Figure 7-10 compares the variation of power sensitivity of triodes and pentodes.[1] The effective input capacitance of power pentodes is in general greater than that of comparable power triodes and thus causes greater frequency distortion in the preceding stage.

The variation of loud-speaker impedance with frequency is more objectionable with pentodes than with triodes. Because of the large variation of harmonic content with load impedance it is difficult to keep the harmonic content low over a wide range of frequency. Furthermore, the load resistance required to keep the nonlinear distortion within

[1] GLESSNER, J. M., *Proc. I.R.E.*, **19**, 1391 (1931).

allowable limits is usually considerably less than that which gives maximum power output, and so the power output and power sensitivity vary appreciably with load, as may be seen from Fig. 7-9. The effective tube resistance shunting the loud-speaker is also considerably higher with pentodes than with triodes. For this reason the loud-speaker is damped less with pentodes, and "booming" caused by loud-speaker resonance may be objectionable.

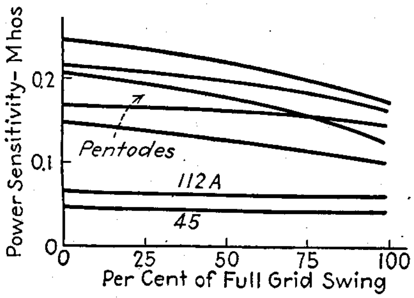


FIG. 7-10.—Variation of power sensitivity of triodes and pentodes with grid swing.

As pointed out in Chap. 6, these different types of distortion can be greatly reduced by the use of negative feedback. The distortion resulting from the variation of speaker impedance can also be reduced by shunting the primary of the output transformer with a series combination of capacitance and resistance.[1] Curve *a* of Fig. 7-11 shows the variation of impedance of a loud-speaker with frequency. Curve *b* shows the improvement obtained by shunting the primary of the output transformer with a 5000-ohm resistor in series with a 0.1-μf condenser. The reduction of pentode distortion by the use of push-pull Class A and Class AB circuits will be discussed later in this chapter and in Chap. 8.

As the result of higher amplification, the degenerative effect of a self-biasing resistor is greater in pentodes than in triodes. For this reason a

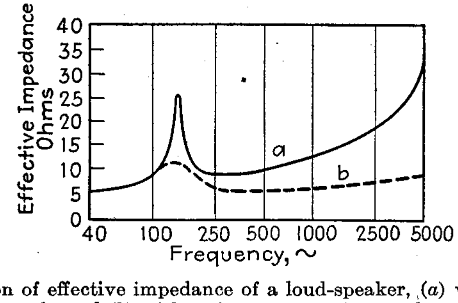


FIG. 7-11.—Variation of effective impedance of a loud-speaker, (*a*) without shunting network, and (*b*) with resistance-capacitance shunt.

larger by-pass condenser is required to prevent degeneration. Inasmuch as degeneration reduces nonlinear and frequency distortion, there is an advantage in omitting the by-pass condenser, but it is more advantageous to use a form of inverse feedback which also reduces the effective resistance shunting the loud-speaker (see page 207).

[1] TULAUSKAS, L., *Electronics*, **3**, 142 (1931).

7-18. Effect of Operating Plate Voltage upon Pentode Power Output. Power pentodes are designed to operate with a screen voltage equal to the plate voltage. The derivation of Eqs. (7-51) is based upon assumptions that do not hold in connection with pentodes. Nevertheless, when the screen voltage is made equal to the operating plate voltage, these equations and the curves based upon them (see page 677) appear to hold with sufficient accuracy for values of E_{bo}' in the approximate range from $\frac{1}{2}E_{bo}$ to $2E_{bo}$ to make them of value in approximate determinations. This is probably explained, in part at least, by the facts that the plate current is determined largely by the screen voltage, that the plate current *vs.* screen voltage characteristics are similar in form to triode plate characteristics, and that the potential distribution between cathode and anode is unaltered when all electrode voltages are changed in proportion.

7-19. Beam Power Tubes as Class A1 Amplifiers.—The characteristic curves of a typical beam power tube, the 6L6, were shown in Fig. 3-16.

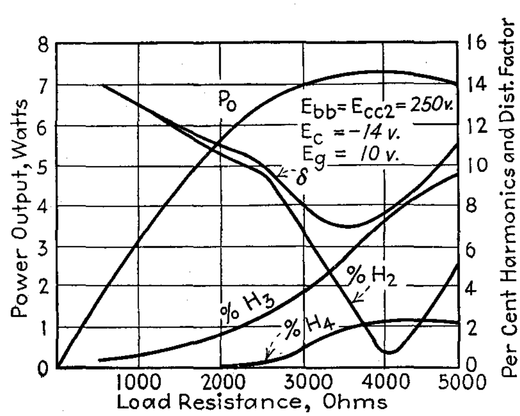


FIG. 7-12.—Power output and distortion curves for the type 6L6 beam power pentode in Class A1 operation.

The shape of these characteristics is such that at a given operating plate voltage even greater power output without excessive nonlinear distortion can be obtained than with suppressor pentodes. The static characteristic curves are more nearly straight and parallel at low plate voltages than those of suppressor pentodes. Figure 7-12 shows curves of power output and distortion against load resistance for a single 6L6 under typical operating voltages with grid swing equal to the grid bias. In order to reduce nonlinear distortion, beam power tubes are usually used in push-pull circuits (see Sec. 7-31). Inverse feedback may also be used to advantage.

7-20. Screen Dissipation of Tetrodes and Beam Pentodes.—The large increase of screen current of power tetrodes and beam pentodes as the plate voltage swings to low values necessitates precautions to ensure that the allowable screen dissipation is not exceeded. The dynamic screen-current curve may be constructed in a manner analogous to that illustrated in Fig. 4-28 for control-grid current. The screen dissipation under excitation may then be found by taking the product of the direct screen voltage and the average screen current, determined graphically from the dynamic screen-current characteristic. For the degree of accuracy ordinarily required in the determination of screen dissipation the average screen current may be found by the use of Eq. (4-29), I_{max} being the screen current corresponding to the control-grid and plate voltages at the positive crest of exciting voltage, I_{bt} being replaced by the quiescent value of screen current, and I_{min} being zero.

Because of the reduction of minimum plate voltage with increase of load, the maximum and average screen currents increase with plate load resistance. Screen dissipation thus limits the maximum load resistance that can be used. It is important to note that the load should never be removed from the secondary of the output transformer of a tetrode or beam pentode, as the resulting increase of effective load impedance may result in damage to the tube because of excessive screen dissipation. For the same reason the plate voltage of tetrodes and pentodes should not be removed without also removing the screen voltage.

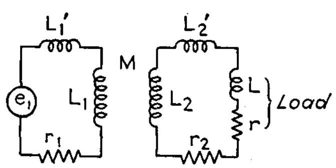


FIG. 7-13.—Equivalent circuit of a loaded transformer.

7-21. Turn Ratio of Output Transformer.—The loud-speaker or other device that is to be operated by the power output of a vacuum tube may have an impedance that differs materially from the desired optimum value. In order to correct for this difference, the turn ratio of the output transformer must be such that the effective primary impedance of the loaded transformer will be equal to the required load impedance.

In determining the input impedance of a loaded transformer, leakage inductance can be taken into account by assuming that the inductances are made up of portions L_1 and L_2, having perfect coupling (unity coefficient of coupling), and portions L_1' and L_2', having zero coupling. The equivalent circuit, neglecting core losses, is shown in Fig. 7-13, in which r_1 and r_2 represent the primary and secondary winding resistances. Solution of the network equations for this circuit gives for the effective input impedance

$$z_i = \frac{E_1}{I_1} = r_1 + j\omega(L_1 + L_1') + \frac{M^2\omega^2}{r_2 + j\omega L_2 + j\omega L_2' + z} \tag{7-52}$$

where z is the impedance of the secondary load.

Since the coefficient of coupling between L_1 and L_2 is assumed to be unity, $M^2 = L_1L_2$. If the leakage inductance is small or is proportional to the number of turns, then $L_2 = n^2L_1$, where n is the ratio of secondary to primary turns; and so $M^2 = L_2{}^2/n^2$. Substituting this relation in Eq. (7-52) transforms it into

$$z_i = r_1 + j\omega(L_1 + L_1') + \frac{L_2{}^2\omega^2}{(r_2 + j\omega L_2 + j\omega L_2' + z)n^2} \tag{7-53}$$

Equation (7-53) may be written in the form

$$z_i = r_1 + j\omega L_1 + j\omega L_1' - \frac{j\omega L_2}{n^2} + j\,\frac{\omega L_2 r_2 + j\omega^2 L_2 L_2' + \omega L_2 z}{(r_2 + j\omega L_2 + j\omega L_2' + z)n^2} \tag{7-54}$$

In a properly designed transformer the total secondary winding reactance is so large in comparison with the secondary winding resistance, leakage reactance, and the load impedance, that only the term $j\omega L_2$ need be considered in the denominator of the last term of Eq. (7-54). Also, since $L_1 = L_2/n_2$, the second and fourth terms cancel. Equation (7-54) then simplifies to

$$z_i = r_1 + \frac{r_2}{n^2} + j\omega\left(L_1' + \frac{L_2'}{n^2}\right) + \frac{z}{n^2} \tag{7-55}$$

If the resistance and leakage reactance of the windings are small in comparison with the load, which is true at loads for which the transformer is designed, Eq. (7-55) reduces to

$$z_i = \frac{z}{n^2} \tag{7-56}$$

Equation (7-56) shows that the effective input impedance of the loaded transformer is equal to the impedance of the secondary load divided by the square of the secondary-to-primary turn ratio. It should be noted that the validity of this equation rests upon the following assumptions:

1. Negligible core loss.
2. Leakage reactance small in comparison with the total winding reactance
3. Winding resistance small in comparison with winding reactance.
4. Load impedance small in comparison with total secondary impedance, but large in comparison with winding resistances and leakage reactances.

These conditions are satisfied in output transformers and other impedance-matching transformers. The turn ratio of the output transformer should, therefore, be equal to the square root of the ratio of the effective impedance of the loud-speaker or other load, to the optimum load resistance. Thus, if the load resistance is 20 ohms and optimum load for the given power tube is 2000 ohms, the output transformer should have a 10 to 1 step-down ratio. Since the impedance and effective resistance of a loud-speaker or other load vary with frequency, it is customary in the field of communication to specify impedances at 1000 cps and to match impedances at this frequency. The resulting mismatching at frequencies above and below this value causes frequency distortion which can be prevented by the use of impedance equalizing networks, similar to that discussed on page 242, or by inverse feedback.

If the primary and secondary winding resistances are not negligible, the first two terms of Eq. (7-55) cannot be dropped and it is necessary to use the more complete expression

$$z_i = r_1 + \frac{r_2 + r}{n^2} + j\frac{x}{n^2} \tag{7-57}$$

7-22. Push-pull Power Amplifiers.[1]—In a push-pull amplifier with transformer-coupled output the instantaneous plate voltage of each tube varies not only with its own plate current but, because of coupling through the primary of the transformer, with the plate current of the other tube. The dynamic transfer characteristic of each tube is different, therefore, from the dynamic transfer characteristic obtained with only one tube. For this reason the graphical analysis of the operation of push-pull power amplifiers is not so simple as that of single-sided amplifiers. The following treatment is essentially that developed by B. J. Thompson.[2]

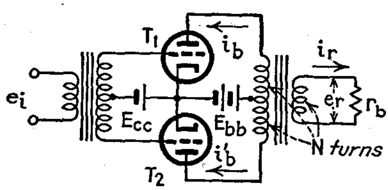


FIG. 7-14.—Push-pull amplifier with resistance load coupled through a transformer having unity turn ratio.

If the output transformer is assumed to have zero winding resistance and leakage reactance and if the small magnetizing current is neglected, the primary and secondary ampere turns are always equal in magnitude and opposite in sign, and the voltage induced across a given number of turns is proportional to the number of turns. The analysis of output transformers given in Sec. 7-21 shows, furthermore, that if the secondary load applied to such an ideal transformer is nonreactive, the voltage across the primary is in phase with the primary current and proportional to it. Well-designed output transformers approximate these ideal conditions very closely.

Figure 7-14 shows the circuit of a push-pull power amplifier with transformer-coupled resistance load r_b. In order to simplify the analysis, the secondary of the output transformer will be considered to have the

[1] The student should review Sec. 5-10.

[2] THOMPSON, B. J., *Proc. I.R.E.*, **12**, 591 (1933).

same number of turns N as each half of the primary. The effect of change of transformation ratio will be considered later.

Because the plate currents in the two halves of the primary flow in opposite directions, the net instantaneous primary ampere turns are the difference between the ampere turns of the two halves $Ni_b - Ni_b'$. Since the leakage inductance is assumed to be zero and the circuit to be symmetrical, this difference must equal Ni_r, the instantaneous value of the secondary ampere turns. The instantaneous secondary current is therefore

$$i_r = i_b - i_b' \tag{7-58}$$

The secondary voltage is

$$e_r = i_r r_b = (i_b - i_b')r_b \tag{7-59}$$

But since each half of the primary has the same number of turns as the secondary and there is no flux leakage, the voltage induced across each half of the primary is also $(i_b - i_b')r_b$. If i_b is greater than i_b', the direction of these voltages is such as to lower the plate voltage of tube 1 below E_{bb} and to raise the plate voltage of tube 2. Because the primary is assumed to have zero resistance, the operating plate voltage E_{bo} is equal to E_{bb} and the instantaneous plate voltages of the two tubes are

$$\begin{aligned} e_b &= E_{bo} - (i_b - i_b')r_b \\ e_b' &= E_{bo} + (i_b - i_b')r_b \end{aligned} \tag{7-60}$$

7-23. Composite Characteristics.—Equations (7-58) to (7-60) show that the instantaneous load current, voltage, and power, as well as the tube plate voltages, depend upon the instantaneous values of the difference between the plate currents of the two tubes and upon the load resistance. It is to be expected, therefore, that a plate diagram involving the plate-current difference should be of value in the analysis of push-pull amplifiers. Since a necessary condition for the complete suppression of even harmonics is that the amplifier shall be in all respects symmetrical, an increase of instantaneous grid voltage of either tube is accompanied by an equal decrease of instantaneous grid voltage of the other tube. For an instantaneous value of alternating grid voltage e_g, one grid has a total instantaneous voltage $e_c = E_c + e_g$, and the other $e_c' = E_c - e_g$. The sum of the instantaneous grid voltages of the two tubes is $e_c + e_c' = 2E_c$. Therefore the diagram should show the variation of e_b and e_b' with $i_b - i_b'$ under the condition that the sum of the grid voltages of the two tubes is always $2E_c$. Furthermore, since, as shown by Eqs. (7-60), the changes of plate voltage of the two tubes relative to E_{bo} are equal and opposite, the diagram should be formed in such a manner that an increase of plate voltage of one tube corresponds to an equal decrease of plate voltage of the other tube.

The method of constructing such a diagram is shown in Fig. 7-15*a*. The plate characteristics for tube 2 are plotted below those for tube 1, with a common voltage axis, but with voltage and current scales that increase in opposite directions. The relative position of the two sets of characteristics is such that the operating points O and O' of the two tubes lie in the same vertical line, and a common point on the voltage axis indicates E_{bo} for both tubes. Then curves of $i_b - i_b'$ *vs.* e_b are plotted for pairs of characteristics such that $e_c + e_c' = 2E_c$. Figure 7-15*a* shows the method of doing this for the pair of characteristics corresponding to $e_c = E_c + e_g$ and $e_c' = E_c - e_g$. Any point X on this *composite*

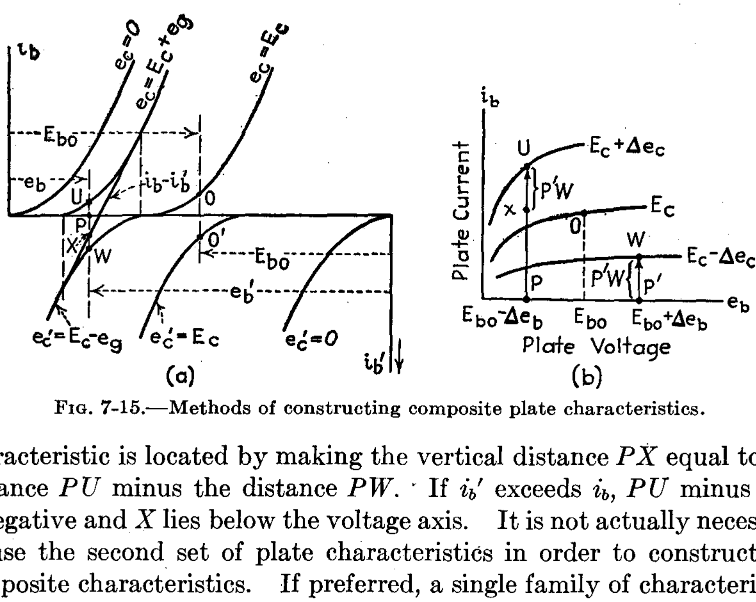


FIG. 7-15.—Methods of constructing composite plate characteristics.

characteristic is located by making the vertical distance PX equal to the distance PU minus the distance PW. If i_b' exceeds i_b, PU minus PW is negative and X lies below the voltage axis. It is not actually necessary to use the second set of plate characteristics in order to construct the composite characteristics. If preferred, a single family of characteristics may be used, as shown in Fig. 7-15*b*. At various values of Δe_b, the distance $P'W$, corresponding to the plate voltage $E_{bo} + \Delta e_b$ and the grid voltage $E_c - \Delta e_c$ is subtracted from the distance PU corresponding to the plate voltage $E_{bo} - \Delta e_b$ and the grid voltage $E_c - \Delta e_c$. Although triode characteristics are shown in Fig. 7-15*a* and pentode characteristics in Fig. 7-15*b*, either method of construction may be used for any type of tube, the procedures being independent of differences in the shapes of the characteristics.

The composite characteristics of triodes for values of E_c corresponding to Class A operation are usually nearly straight and parallel. They may be approximated by locating the point on each plate characteristic of one tube corresponding to cutoff for the other tube and joining these points with a straight line. For high values of bias the composite triode characteristics may be appreciably curved. In Fig. 7-16 are shown the composite characteristics for type 45 triodes at an operating plate

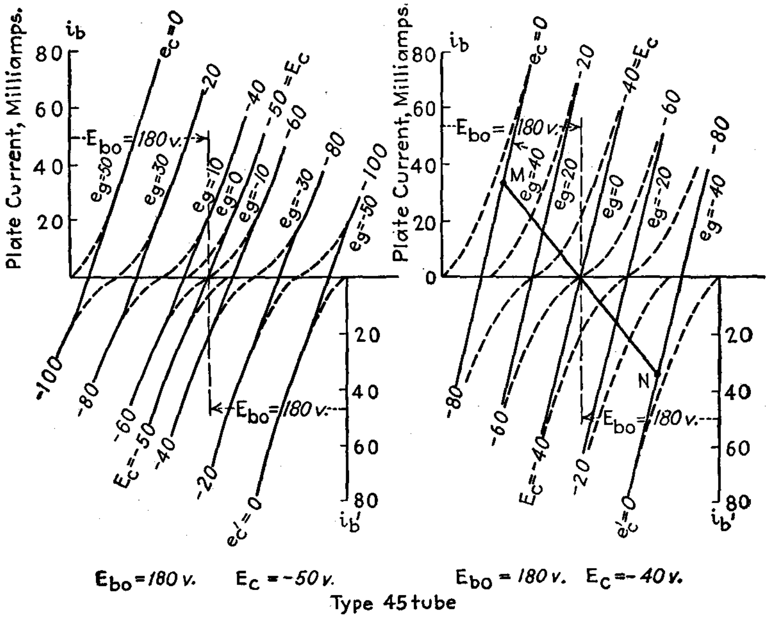


FIG. 7-16.—Composite plate characteristics for the type 45 power triode at two values of grid bias.

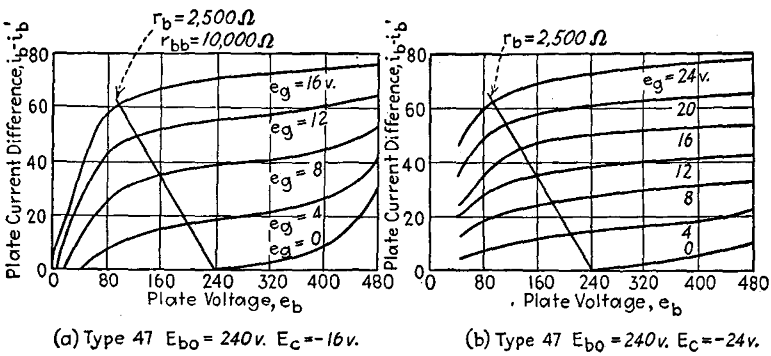


FIG. 7-17.—Composite plate characteristics for the type 47 power pentode at two values of grid bias.

voltage of 180 volts, and −50- and −40-volt grid bias. The composite characteristics of pentodes do not approximate straight lines, as may be seen from the suppressor- and beam-pentode characteristics of Figs. 7-17 and 7-18.

It is evident from the method of deriving the composite characteristics and from Fig. 7-16, that, when the two tubes are alike and the circuit is symmetrical, the composite plate diagram is symmetrical about the voltage axis. Furthermore, the load line passes through only the second and fourth quadrants. For these reasons, only the portion of the diagram lying in the second quadrant need be constructed. If the tubes are not identical, the composite plate diagram is not symmetrical about the voltage axis, and both the second and the fourth quadrants must be used.

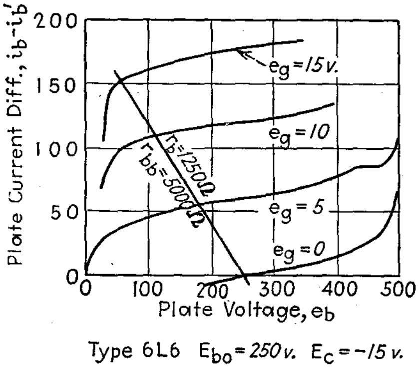


FIG. 7-18.—Composite plate characteristics for the type 6L6 beam power pentode.

7-24. Load Line.—Let $i_b - i_b'$, e_b, and e_b' be any simultaneous instantaneous values of current difference and plate voltages assumed

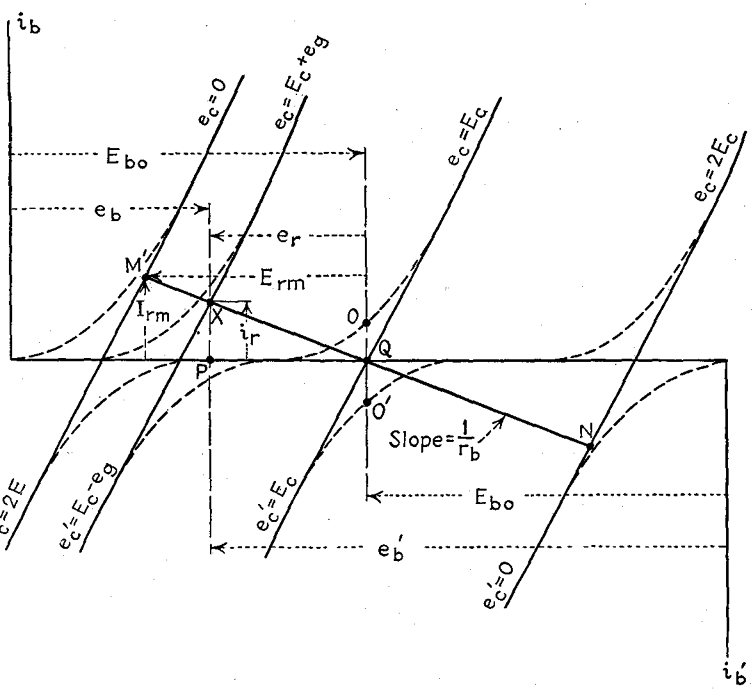


FIG. 7-19.—Composite plate diagram.

during the cycle with a given load r_b and operating plate voltage E_{bo}. The corresponding point is indicated by X in Fig. 7-19. A straight line is drawn through this point and the point Q on the voltage axis corresponding to E_{bo}. The following relations are evident from Fig. 7-19 and Eq. (7-60):

$$PX = i_b - i_b' \tag{7-61}$$

$$PQ = E_{bo} - e_b = (i_b - i_b')r_b \tag{7-62}$$

$$\frac{PX}{PQ} = \frac{1}{r_b} \tag{7-63}$$

Since $i_b - i_b'$, e_b, and e_b' are *any* simultaneous instantaneous values assumed during the cycle, it follows that the straight line through Q having a negative slope equal to $1/r_b$ represents the locus of all corresponding values of current difference and plate voltage with the given load and operating voltages. Furthermore, Eq. (7-58) shows that the current difference at any point is also equal to the load current i_r, and Eqs. (7-59) and (7-62) show that PQ, the change in voltage relative to E_{bo}, is equal to the load voltage e_r. Therefore the line MQN is also the locus of all corresponding values of instantaneous alternating grid voltage and secondary load current and voltage and is called the *load line.* These values of load current and voltage are, of course, measured relative to Q.

7-25. Path of Operation of Each Tube.—Because the instantaneous plate voltage of each tube is dependent not only upon its own plate current but also upon that of the other tube, the effective load on each tube varies throughout the cycle, and the path of operation of each tube is curved. The individual paths of operation may be obtained by plotting corresponding values of i_b and e_b and of i_b' and e_b' determined from the composite plate diagram. The method of constructing individual paths of operation is illustrated in Fig. 7-20, which shows typical triode paths. It is of interest to note that the curvature of the path of operation of each tube is such that the curvature of the dynamic transfer characteristic of each tube is greater than it would be if the tube were used by itself with constant load resistance.

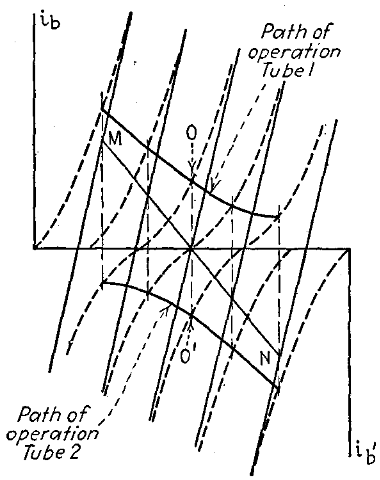


FIG. 7-20.—Triode composite plate diagram showing the individual paths of operation of the two tubes.

7-26. Graphical Determination of Harmonic Content and Power Output.—A dynamic characteristic of load current *vs.* alternating grid voltage may be derived from the composite plate diagram by plotting simultaneous values of i_r and e_g determined from the intersections of the load line with the composite characteristics. Figure 4-26 shows such a characteristic and illustrates its use in determining the wave form of the load current. The dynamic characteristic may be used for harmonic analysis of the load current. Since the characteristic is symmetrical, Eqs. (4-34), (4-35), or (4-36) may be used. These equations may also be used to determine the amplitudes of the various components of load current directly from the composite plate diagram, in the same manner as for a single tube (all currents being measured relative to the voltage axis). In Fig. 7-17*b*, for instance, $i_{1/2}$ and i_1 at full grid swing are the currents corresponding to the intersection of the load line with the composite characteristics for $e_g = 12$ volts and $e_g = 24$ volts, respectively, or 35 ma and 62 ma. Hence Eqs. (4-34) show that

$$H_1 = \tfrac{2}{3}(62 + 35) = 64.7 \text{ ma},$$

and $H_3 = \frac{1}{3}(62 - 70) = -2.7$ ma.

The fundamental power output may be computed from the graphically determined value of H_1. If the third and fifth harmonics are small, however, as is true for Class A triode amplifiers, little error is introduced by assuming that the amplitude of the fundamental component of load current is the crest instantaneous load current I_{rm} determined by the intersection of the load line with the composite characteristic corresponding to the crest value of the instantaneous alternating grid voltage. At full grid swing this is the current corresponding to the point M in Fig. 7-19. Then the power output is approximately

$$P_b = \tfrac{1}{2}I_{rm}{}^2 r_b = \tfrac{1}{2}I_{rm}E_{rm} \tag{7-64}$$

It can be seen from Fig. 7-20 that the increase of plate current of each tube when its grid swings in the positive direction is greater than the decrease when the grid swings in the negative direction. The average current, therefore, increases with excitation. The average current I_{ba} for one tube can be found by applying one of the formulas of Chap. 4 to the path of operation of one tube. (Note that $I_{bt} = I_{bo}$.) The power input to the plate circuit is

$$P_i = 2E_{bb}I_{ba} = 2E_{bo}I_{ba} \tag{7-65}$$

An approximate expression for power input can be obtained by using I_{bo} in place of I_{ba}.

7-27. Class AB Push-pull Operation.—In push-pull amplifiers with fixed grid bias there is no abrupt increase of amplitude distortion when the bias is increased beyond the limiting Class A value, as the student may show by graphical analysis of composite plate diagrams at various values of grid bias. When bias is obtained by the use of a cathode resistor, however, rectification caused by the interruption of plate current in each tube during a portion of the cycle in Class AB operation causes the average plate current to increase at high excitation. This tends to overbias the grids and raise the nonlinear distortion. Distortion is also augmented by the lowering of operating plate voltage if the B-supply regulation is high. The increased distortion is particularly noticeable with triodes. Even with self-biased triodes, however, the bias may be increased considerably beyond the limiting Class A value without raising the distortion to excessive values. At a given operating plate voltage, this increase of bias reduces the operating plate current and thus decreases the plate dissipation and increases the plate-circuit efficiency. At a given plate dissipation it allows the use of higher plate voltage and consequent development of greater power output at higher efficiency. For this reason Class AB bias is usually used in push-pull amplifiers, in preference to Class A1 bias.

7-28. Optimum Operation of Class A1 and Class AB1 Push-pull Triode Amplifier.—The shape of triode composite characteristics is such that the harmonic content in Class A1 and fixed-bias Class AB1 operation is small except at very low values of load resistance. For this reason the grid swing and power output are not limited by allowable distortion, as in single-sided amplifiers. Full grid swing is equal to the grid bias, and optimum load resistance for given operating voltages is that which gives the greatest power output at full grid swing.

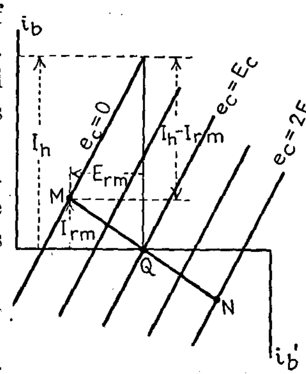


FIG. 7-21.—Idealized triode composite plate diagram used in determining theoretical optimum Class A load resistance.

The theoretical optimum load resistance for triodes in Class A1 or fixed-bias Class AB1 operation at given operating voltages can be determined from the idealized composite plate diagram of Fig. 7-21. In Fig. 7-21 the composite static plate characteristics are assumed to be straight, parallel, and equidistant. If r_{pp} is defined as the reciprocal of the slope of the composite characteristics,

$$r_{pp} = \frac{E_{rm}}{I_h - I_{rm}} \tag{7-66}$$

Substituting Eq. (7-66) in Eq. (7-64) gives

$$P_o = \tfrac{1}{2} r_{pp}(I_h - I_{rm})I_{rm} \tag{7-67}$$

P_o is maximum when $dP_o/dI_{rm} = 0$, or

$$I_h = 2I_{rm} \tag{7-68}$$

and

$$r_{pp} = \frac{E_{rm}}{I_{rm}} = r_b \tag{7-69}$$

Equation (7-69) states that the optimum load resistance is equal to the *composite plate resistance*, the latter being defined as the reciprocal of the slope of the composite characteristics. Inasmuch as the composite characteristics are not strictly straight and parallel, r_{pp} may be defined more specifically as the reciprocal of the slope of the composite characteristic for $e_g = 0$ at the point at which it crosses the voltage axis. This may be stated symbolically by the relation

$$\frac{1}{r_{pp}} = \left[\frac{\partial(i_b - i_b')}{\partial e_b}\right]_{e_b = E_{bo} \text{ and } e_g = 0} \tag{7-70}$$

Optimum power output is obtained when the slope of the load line is equal in magnitude to the slope of the composite characteristic corresponding to $e_g = 0$ ($e_c = E_c$) at their point of intersection.

7-29. Limiting Class A1 Triode Operation.[1]—Because the large variation of average current with signal amplitude in Class AB amplification increases the difficulty of maintaining constant operating voltages, it is sometimes of advantage to be able to determine the highest grid bias that can be used in Class A1 push-pull operation. At this bias the minimum instantaneous plate current of each tube is just equal to zero when the grid swing is equal to the bias. The theoretical value of this limiting bias at optimum load can be determined for triodes from Fig. 7-22 by assuming that the static characteristic corresponding to zero grid voltage follows a $\frac{3}{2}$-power law, which is approximately true for most power triodes. Under this assumption,

$$i_b = Ae_b^{3/2} \qquad \text{for} \qquad e_c = 0 \tag{7-71}$$

$$\frac{1}{r_p} = \frac{di_b}{de_b} = \tfrac{3}{2}Ae_b^{1/2} \qquad \text{for} \qquad e_c = 0 \tag{7-72}$$

Dividing Eq. (7-71) by Eq. (7-72) gives

$$i_b r_p = \tfrac{2}{3}e_b \qquad \text{for} \qquad e_c = 0 \tag{7-73}$$

At M, where $i_b = I_{rm}$ and $e_b = E_{\min}$, the composite static plate characteristic joins the static characteristic for tube 1, and the value of r_p at that point is equal to r_{pp}, the reciprocal of the slope of the composite characteristics. Hence, evaluating Eq. (7-73) at M gives

$$I_{rm} r_{pp} = \tfrac{2}{3}E_{\min} \tag{7-74}$$

[1] The author is indebted to Mr. W. R. Ferris of the RCA laboratories for the derivations in this section.

But, since $E_{\min} = E_{bo} - I_{rm}r_b$ and, for optimum output, $r_{pp} = r_b$,

$$I_{rm}r_{pp} = E_{bo} - E_{\min} \tag{7-75}$$

Combining Eqs. (7-74) and (7-75) gives

$$E_{bo} - E_{\min} = \tfrac{2}{3}E_{\min} \tag{7-76}$$

and

$$E_{\min} = 0.6E_{bo} \tag{7-77}$$

$$E_{rm} = E_{bo} - E_{\min} = 0.4E_{bo} \tag{7-78}$$

At N the plate and grid voltages of tube 1 have the values

$$e_b = E_{bo} + E_{rm} = E_{bo} + 0.4E_{bo} = 1.4E_{bo} \tag{7-79}$$

$$e_c = 2E_c \tag{7-80}$$

Since the plate current of tube 1 is zero at point N under maximum Class A1 bias, a grid voltage equal to twice the grid bias gives cutoff at a

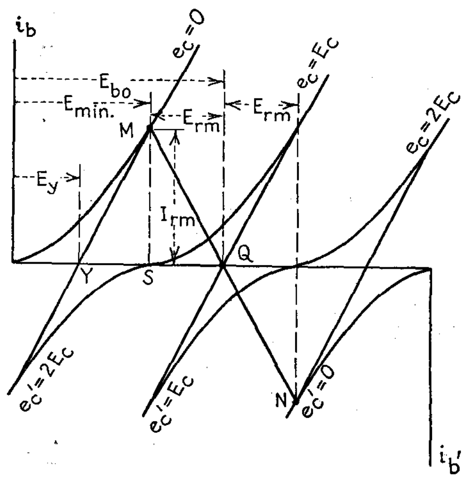


FIG. 7-22.—Idealized composite plate diagram of triodes with optimum load resistance and maximum Class A1 bias.

plate voltage $1.4E_{bo}$. Therefore the limiting value of E_c for push-pull Class A1 operation of triodes is one-half the value that reduces the plate current to zero at a plate voltage equal to $1.4E_{bo}$. Because of variation of amplification factor, the value determined in this manner is not exactly equal to the value that results in plate-current cutoff at a plate voltage of $0.7E_{bo}$. The error, however, will usually be less than 5 per cent if the limiting Class A1 bias is determined from the formula

$$E_c = -0.7\,\frac{E_{bo}}{\mu} \tag{7-81}$$

This value of bias gives the highest plate-circuit efficiency that can be obtained in optimum push-pull Class A1 operation of triodes.

Optimum power output and load resistance of triodes at the limiting value of Class A1 bias can be determined without constructing composite characteristics. By using Eq. (7-77), it can be seen from Fig. 7-22 that I_{rm} is the plate current of one tube at the plate voltage $0.6E_{bo}$ and zero grid voltage. Optimum power output may be found by substituting this value of I_{rm} in the following equation, which is derived by substituting Eq. (7-78) in Eq. (7-64):

$$\text{Optimum } P_o = 0.2I_{rm}E_{bo} \tag{7-82}$$

Optimum secondary load resistance is equal to the reciprocal of the slope of a line drawn through the point E_{bo} on the voltage axis, and the point $e_b = 0.6E_{bo}$ on the zero-grid-voltage characteristic of one tube. This is

$$\text{Optimum } r_b = \frac{0.4E_{bo}}{I_{rm}} \tag{7-83}$$

7-30. Push-pull Class A1 Triode Power Output in Terms of Operating Plate Voltage.—An expression for optimum power output of push-pull triodes in terms of operating plate voltage can be derived from the plate diagram of Fig. 7-22.

$$E_{bo} - E_y = \text{distance } YQ = YS + SQ$$

$$= I_{rm}r_{pp} + I_{rm}r_b \tag{7-84}$$

$$I_{rm} = \frac{E_{bo} - E_y}{r_{pp} + r_b} \tag{7-85}$$

Substitution of Eq. (7-85) in Eq. (7-64) gives

$$\text{Full } P_o = \frac{(E_{bo} - E_y)^2 r_b}{2(r_{pp} + r_b)^2} \tag{7-86}$$

For optimum power output, $r_b = r_{pp}$, and for maximum Class A1 bias,

$$E_y \cong E_{bo} - 0.8E_{bo} = 0.2E_{bo}.$$

$$\text{Optimum } P_o \cong 0.08 \frac{E_{bo}^2}{r_{pp}} \tag{7-87}$$

The principal value of Eq. (7-87) lies in the fact that it shows the manner in which the power output depends upon the operating plate voltage and upon plate resistance. r_{pp} decreases with increase of steepness of the static plate characteristics. Therefore, triodes used in Class A push-pull power amplification should have low plate resistance.

7-31. Suppressor and Beam Power Pentodes in Push-pull Operation. It can be seen from the composite plate diagrams of Figs. 7-17 and 7-18 that the shape of composite pentode characteristics is not such as to allow Eq. (7-69) to be applied. Distortion rapidly becomes prohibitive as the load resistance is increased beyond the value for which the load line intersects the composite characteristic for full grid swing at the knee. The characteristics fall so sharply at plate voltages lower than that of the knee that the power output also begins to decrease as the load resistance is increased beyond this value. The load lines shown in Figs. 7-17 and 7-18 represent approximately the limiting load resistances beyond which the distortion rises sharply and little or no increase in power output is obtained.

Comparison of Figs. 7-17 and 7-18 shows clearly that the minimum plate voltage of the beam power tube can be reduced to a lower percentage of the operating voltage than that of the suppressor pentode without producing excessive distortion. This makes for higher plate-circuit efficiency and higher output at a given plate dissipation. The composite characteristics of the beam pentode are also seen to be more evenly spaced than those of the suppressor pentode, so that distortion can be made much lower with the beam pentode than with the suppressor pentode. Two 6L6 tubes in Class A1 push-pull operation are capable of delivering an output of 14.5 watts with a distortion factor of only 2 per cent and a plate efficiency of 41.5 per cent.

7-32. Summary. Push-pull Operation.—A complete analysis of the operation of tubes of a given type in a push-pull amplifier with an output transformer of unity turn ratio between the secondary and each half of the primary may be made by constructing composite plate diagrams for various values of operating plate voltages. The amplitudes of the fundamental and harmonic components of the load current may be determined by applying Eqs. (4-34), (4-35), or (4-36) to the plate diagrams or to transfer characteristics of load current *vs.* instantaneous alternating grid voltage derived from the plate diagrams.

Optimum load resistance for triodes in Class A1 or Class AB1 operation is that corresponding to a load line whose slope is equal in magnitude to that of the composite characteristics. For pentodes, optimum Class A1 or Class AB1 load resistance is approximately that corresponding to a load line which intersects the composite characteristic corresponding to positive crest alternating grid voltage just to the right of the knee, the exact value being influenced by the allowable nonlinear distortion.

Optimum load resistance and power output for triodes with maximum Class A1 bias may be quickly determined without constructing composite characteristics by substituting in Eqs. (7-82) and (7-83) the plate current determined from the zero-grid-voltage plate characteristic of one tube at a plate voltage of $0.6E_{bo}$.

7-33. Turn Ratio of Output Transformer. Plate-to-plate Load.—The foregoing analysis was based upon a transformer having unity turn ratio between half the primary and the secondary. For any other turn ratio the load must be transformed into an equivalent load at unity turn ratio, and this equivalent load must be used in determining the slope of the load line and the power output. The analysis of Sec. 7-21 shows that the equivalent unity-turn-ratio load r_b is equal to r/n^2, where r is the actual secondary load and n is the turn ratio of the secondary to each half of the primary. Usually the optimum value of r_b is determined graphically or experimentally, and r is the known resistance of the loudspeaker or other load. The transformer turn ratio is then chosen so that

$$n^2 = \frac{r}{r_b} \tag{7-88}$$

It is common practice to specify the optimum *plate-to-plate load*, rather than the optimum unity-turn-ratio load. A plate-to-plate load is a load connected across the primary from plate to plate. Since the total number of primary turns is $2N$, a load connected from plate to plate has the same effect as though it were connected across a secondary having $2N$ turns. For such a transformer n is 2, and so

$$r_b = \frac{r_{bb}}{4}, \tag{7-89}$$

where r_{bb} is the plate-to-plate load. Hence the optimum plate-to-plate load is four times the optimum unity-turn-ratio load. For an optimum plate-to-plate load r_{bb} and an actual secondary load r, the turn ratio must be chosen so that

$$n^2 = \frac{4r}{r_{bb}} \tag{7-90}$$

Problems

7-1. Construct a "power-output rule" for 5 per cent second harmonic.

7-2. *a.* Neglecting the shift in operating point with excitation, locate the optimum operating point for a 45 triode operated at a plate voltage of 200 volts. The load is nonreactive and is coupled to the tube through a transformer of small primary resistance. The allowable plate dissipation is 10 watts, and the allowable second-harmonic content is 5 per cent.

b. Locate the optimum load line.

c. Determine the optimum power output graphically.

d. From the load line find the optimum load resistance.

e. Determine the plate resistance at the operating point.

f. Using the values of r_p and r_b found in (*e*) and (*d*), find the optimum power output by means of the equivalent plate circuit.

g. Check the optimum power output by means of Eq. (7-21).

h. Find the power input to the plate circuit, and the plate efficiency. Use the graphically determined value of power output.

i. Find the power sensitivity, using the graphically determined value of optimum power.

j. If the resistance of the load across the secondary of the output transformer is 40 ohms, find the correct transformer ratio.

7-3. Repeat Prob. 7-2, using an operating plate voltage of 300 volts.

7-4. Determine the optimum operating voltages, load, and power output of a type 45 triode at unspecified operating plate voltage, an allowable plate dissipation of 9 watts, and an allowable second-harmonic content of 5 per cent.

7-5. In the diagram of Fig. 7-23 are shown idealized triode and pentode characteristics for the same operating voltages and currents. Both sets of characteristics are assumed to be straight, parallel, and equally spaced, and the pentode characteristics are assumed to be parallel to the voltage axis to the right of the knees. $I_{\min}$ is assumed to be zero for both tubes. By means of this diagram show that a triode which operates at the same plate voltage as a pentode and delivers the same power

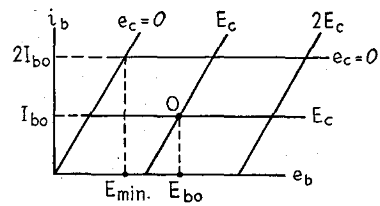


FIG. 7-23.—Diagram for Prob. 7-5.

output at the same efficiency and power sensitivity would have to have a transconductance whose value is approximately $(2E_{bo} - E_{\min})/E_{\min}$ times that of the pentode.

7-6. *a.* Construct composite characteristics for type 6F6 tubes connected as triodes, operating at a plate voltage of 250 volts and a grid bias of -30 volts.

b. Draw the load line for optimum load, and determine the optimum plate-to-plate load resistance.

c. Find the percentage H_3 and percentage H_5 at full grid swing.

d. Find the power output for two tubes.

e. Find the plate efficiency.

f. Find the required transformer turn ratio if the secondary load is 50 ohms.

g. Plot the path of operation for one tube.

Supplementary Bibliography

Class A Power Amplifiers

KELLOGG, E. W.: *J. Am. Inst. Elec. Eng.*, **44,** 490 (1925).
WARNER, J. C., and LOUGHREN, A. V.: *Proc. I.R.E.*, **14,** 735 (1926).
RADT, W. P.: *E.N.T.*, **3,** 21 (1926).
WILLIS, F. C., and MELHUISH, L. E.: *Bell System Tech. J.*, **5,** 573 (1926).
CLAVIER, A., and PODLIASKY, I.: *L'Onde élec.*, **6,** 71 (1927).
HANNA, C. R., SUTHERLIN, L., and UPP, C. B.: *Proc. I.R.E.*, **16,** 462 (1928).
BAGGALLY, W.: *Wireless Eng.*, **5,** 315 (1928).
BURKE, C. T.: *Radio. Eng.*, May, 1928, p. 18.
BARTELS, H.: *E.N.T.*, **6,** 9 (1929).
BRAIN, B. C.: *Wireless Eng.*, **6,** 119 (1929).
FORSTMANN, A.: *L'Onde élec.*, **9,** 13 (1930).
SOWERBY, A. L. M.: *Wireless World*, **26,** 548 (1930).
PIDGEON, H. A., and MCNALLY, J. O.: *Proc. I.R.E.*, **18,** 266 (1930) (with bibliography of 23 items).

Baggally, W.: *Wireless Eng.*, **7**, 430 (1930).
Forstmann, A.: *Jahrb. drahtl. Tel. u. Tel.*, **35**, 109 (1930).
Wigge, H.: *Hochfrequenztech. u. Elekrtoakus.*, **37**, 16 (1931).
Peters, J. F.: *Elec. Eng.*, **8**, 428 (1931).
Greenwood, W.: *Wireless Eng.*, **8**, 428 (1931)
Sims, L. G. A.: *Wireless Eng.*, **9**, 314 (1932).
Stromeyer, C. F.: *Radio Eng.*, February, 1932, p. 39; *Electronics*, June, 1932, p. 194; *Proc. I.R.E.*, **20**, 1149 (1932).
Nelson, J. R.: *Proc. I.R.E.*, **20**, 1763 (1932).
Tulauskas, L.: *Electronics*, May, 1933, p. 135.
Kilgour, C. E.: *Electronics*, March, 1933, p. 72.
Stromeyer, C. F.: *Radio Eng.*, August, 1933, p. 11; *Proc. I.R.E.*, **23**, 565 (1935).
Baggally, W.: *Wireless Eng.*, **12**, 496 (1935).
Burnap, R. S., and Schade, O. H.: *RCA Rev.*, **1**, 101 (1936).
Dreyer, J. F., Jr., and Schade, O. H.: *Electronics*, April, 1936, p. 18.
Harries, J. H. O.: *Electronics*, May, 1936, pp. 33, 45.
Durham, W. N.: *Radio Eng.*, June, 1936, p. 20.
Stromeyer, C. F.: *Proc. I.R.E.*, **23**, 565 (1935).
Oxman, Leon: *Electronics*, September, 1936, p. 30.
Terman, F. E.: *Electronics*, January, 1937, p. 12.
Day, J. R., and Russell, J. B.: *Electronics*, April, 1937, p. 16.
Martin, L.: *Radio Eng.*, May, 1937, p. 13.
Houghton, E. W.: *Electronics*, June, 1937, p. 18.
Fritzinger, G. H.: *Proc. I.R.E.*, **26**, 207 (1938).
Schade, O. H.: *Proc. I.R.E.*, **26**, 137 (1938).
Levy, M. L.: *Electronics*, March, 1938, p. 26.
Everitt, W. L., and Spangenberg, K.: *Proc. I.R.E.*, **26**, 612 (1938).
Chaffee, E. L.: *J. Applied Physics*, **9**, 471 (1938).
Davis, F. M.: *Electronics*, December, 1938, p. 56.
Fairweather, A., and Williams, F. C.: *Wireless Eng.*, **16**, 57 (1939).
Cook, E. G.: *Electronics*, June, 1939, p. 38.
Jonker, J. L. H.: *Wireless Eng.*, **16**, 274, 344 (1939).
van der Ven, A. J. H.: *Wireless Eng.*, **16**, 383, 444 (1939).
Block, A.: *Wireless Eng.*, **16**, 592 (1939).
Preisman, A.: *Communications*, December, 1939, p. 5.
Treviño, S. N., and Offner, F.: *Rev. Sci. Instruments*, **11**, 412 (1940).

Amplifier Distortion

Macfadyen, K. A.: *Wireless Eng.*, **15**, 310 (1938).
Scott, H. J., and Black, L. J.: *Proc. I.R.E.*, **26**, 449 (1938).
Tittle, H. C.: *R.M.A. Eng.*, May, 1938, p. 9.
Wheeler, H. A.: *Proc. I.R.E.*, **27**, 359, 384 (1939).

CHAPTER 8

CLASS B AND CLASS AB2 AMPLIFIERS

The distinguishing feature of Class AB, Class B, and Class C amplifiers as a group is that plate current in individual tubes does not flow during the whole cycle.[1] Class B and Class C operation were originally used in radio-frequency amplifiers. Later it was shown that the high power output and efficiency of Class B operation could also be attained without excessive distortion in audio-frequency amplifiers.[2] More recently Class B audio power amplifiers have to a great extent been displaced by Class AB amplifiers using beam power tubes, which give much lower amplitude distortion and operate at only slightly lower efficiency. Class B operation now finds its principal application in r-f amplifiers; in battery-operated a-f power amplifiers, where low power consumption is essential; and in high-power a-f power amplifiers used to supply signal excitation in linear plate modulation. Class C operation, as explained in Chap. 5, can be used only in r-f amplifiers.

8-1. Class B Audio-frequency Amplifiers.—Because a tube that is used in Class B amplification passes plate current during only approximately half the cycle, Class B audio amplifiers must employ two tubes in push-pull in order to prevent excessive distortion. The composite plate diagrams discussed in the preceding chapter are applicable to the analysis of Class B amplifiers. Since the static operating plate current is small, if not actually zero, the composite portions of the plate characteristics take up a relatively small portion of the whole curves, as shown in Fig. 8-1, and only at very small signal voltages does the path of operation intersect these composite curves.

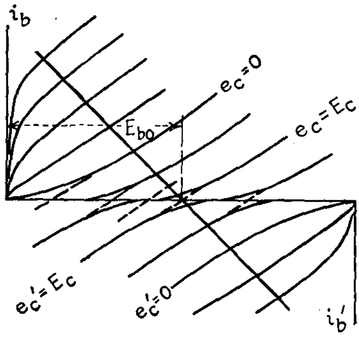


FIG. 8-1.—Composite plate diagram for triodes with grid bias slightly less than cutoff value.

The advantages of Class B operation over Class A operation in audio-frequency power amplifiers are the higher plate-circuit efficiency and the resulting higher power output that can be attained with tubes of a given

[1] The student should review Sec. 5-15.

[2] BARTON, L. E., *Proc. I.R.E.*, **19**, 1131 (1931); **20**, 1035 (1932).

size and with power supplies of a given capacity. The principal disadvantage is the much greater nonlinear distortion. Furthermore, full advantage of the capabilities of Class B amplification is obtained only when the grid swing is so great that grid current flows during a considerable portion of the cycle. The flow of grid current and the associated expenditure of power in the grid circuit are disadvantages that result in a number of design problems. These and other design problems will be discussed in detail in Sec. 8-9.

The greatest plate-circuit efficiency is obtained if the grids are actually biased to cutoff, reducing the static operating plate current and static plate power dissipation to zero. Under this condition of operation only one tube at a time carries current, and the plate-circuit diagram reduces

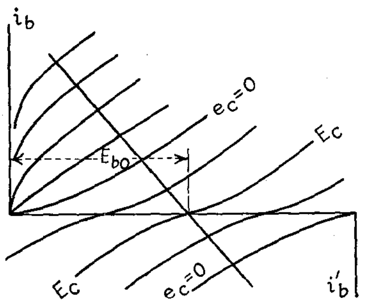


FIG. 8-2.—Composite plate diagram for triodes biased to cutoff.

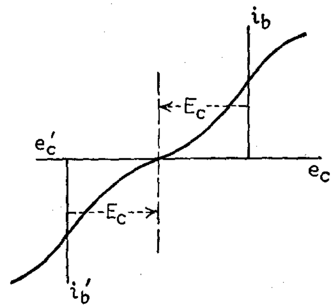


FIG. 8-3.—Dynamic transfer characteristic corresponding to the plate diagram of Fig. 8-2.

to that of Fig. 8-2. The dynamic transfer characteristic corresponding to Fig. 8-2 for the two tubes is of the form shown in Fig. 8-3. It is apparent that the relation between plate current and grid voltage is far from linear and that bad distortion will result. If, on the other hand, the grid bias is made sufficiently less negative than the cutoff value (as in Fig. 8-1) so that those portions of the dynamic transfer characteristics which are approximately straight fall on a common line, as in Fig. 8-4, then distortion will be greatly reduced.

The resultant dynamic curve of load current *vs.* alternating grid voltage for that portion of the cycle during which both tubes pass current cannot be obtained by adding the currents of the two individual dynamic characteristics of Fig. 8-4. These dynamic characteristics are drawn under the assumption that the instantaneous plate voltage of each tube depends only upon the plate current of that tube, whereas, as shown in Chap. 7, it depends upon both plate currents. The resultant dotted portion of the dynamic transfer characteristic in Fig. 8-4 must be derived by the method explained for push-pull amplifiers in Sec. 7-23. It is generally slightly S-shaped, as shown in Fig. 8-5. At high excitation this curvature is relatively unimportant, but at low excitation it results in appreciable nonlinear distortion.

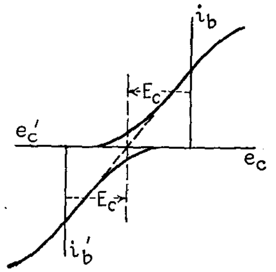


FIG. 8-4.—Dynamic transfer characteristic corresponding to the plate diagram of Fig. 8-1.

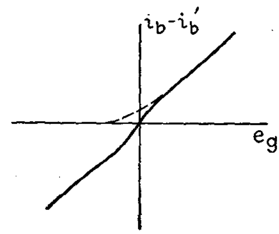


FIG. 8-5.—Dynamic transfer characteristic at small values of grid swing. (e_g is measured relative to E_c.)

8-2. Class B Characteristics.—Because the path of operation is not restricted to the negative grid-voltage region in Class B amplifiers, problems of Class B amplifier design involve not only the plate characteristics, but also the grid characteristics. In Fig. 8-6 are shown static curves of i_b and i_c *vs.* e_b for a type 46 tube.[1] The dynamic transfer charac-

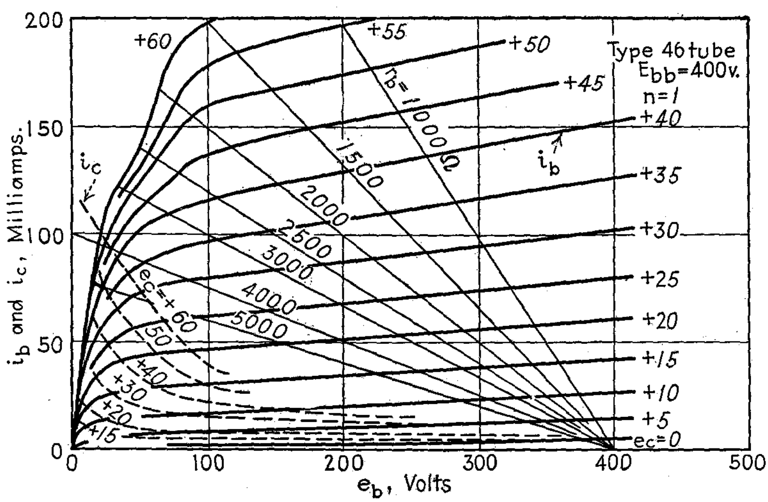


FIG. 8-6.—Static plate characteristics and plate-grid transfer characteristics for type 46 tube, connected for zero-bias Class B operation (grids tied together).

teristics and the dynamic grid characteristics, which are shown in Figs. 8-8 and 8-9, may be derived from the static characteristics by methods outlined in Chaps. 4 and 7. Figure 8-7 shows composite plate characteristics used for deriving the dynamic characteristics at low grid swings. At higher alternating grid voltage the load lines of Fig. 8-6 determine directly corresponding values of i_b and e_c for the given loads; they also

[1] The type 46 tube is designed to operate with zero bias in Class B amplifiers.

determine simultaneous values of e_b and e_c that can be used in Fig. 8-6 to pick simultaneous values of i_c and e_c for the dynamic grid characteris-

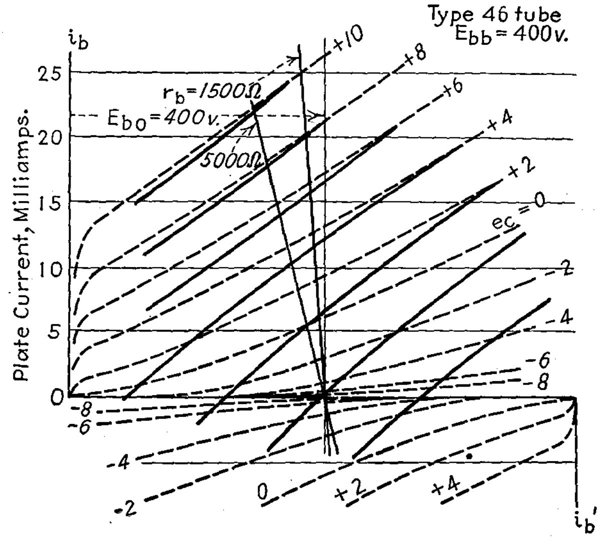


FIG. 8-7.—Composite plate characteristics at zero bias for type 46 tubes connected for Class B operation (grids tied together).

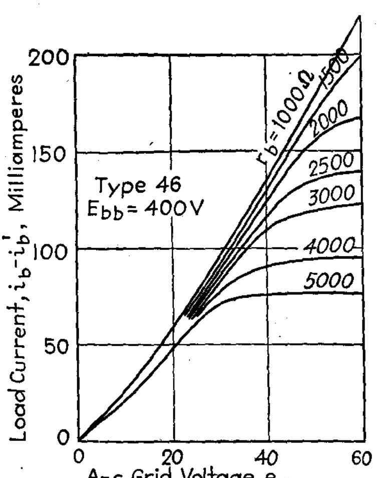


FIG. 8-8.—Dynamic transfer characteristics of type 46 tube in zero-bias Class B operation.

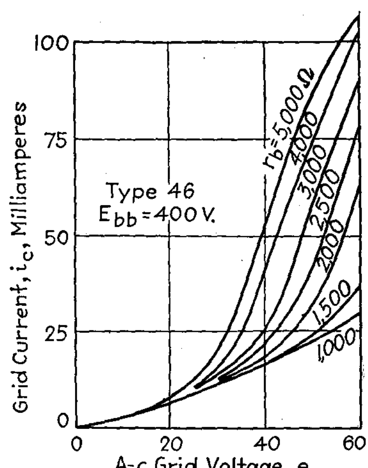


FIG. 8-9.—Dynamic grid characteristics of type 46 tube in zero-bias Class B operation.

tics of Fig. 8-9. The slope of any load line is the reciprocal of r_b, *i.e.*, of $\frac{1}{4}r_{bb}$, where r_{bb} is the effective plate-to-plate load resistance.[1]

[1] Except at low grid swings, only one tube at a time passes current and so the effective load per tube is $r_b = \frac{1}{4}r_{bb}$.

The bending over of the dynamic transfer characteristics at high positive grid voltages results from diversion of electron current from plate to grid. For a given value of grid voltage, the grid current, and hence the loss of plate current, are greater the lower the plate voltage. Since voltage drop in the load resistance lowers the instantaneous plate voltage, the diversion of current from plate to grid is greater for higher load resistances, and therefore the bending of the dynamic transfer characteristics increases with load resistance.

8-3. Harmonic Content.—Equations (4-35) or (4-36) or other more accurate methods of harmonic analysis may be used in determining the amplitudes of the fundamental and the third- and fifth-harmonic currents. It should be noted that, although the push-pull circuit prevents the

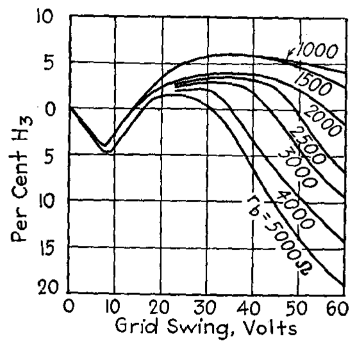


FIG. 8-10.—Per cent third harmonic as a function of grid swing. Type 46 tubes in zero-bias Class B operation.

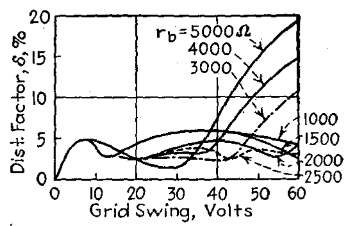


FIG. 8-11.—Distortion factor as a function of grid swing. Type 46 tubes in zero-bias Class B operation.

appearance of even harmonics or steady components in the output, the plate current of each tube contains a large steady component and that the average plate-supply current is the sum of the average currents of the two tubes. The high-order harmonic content of the plate-current wave of each tube is so great that the five- and seven-point methods of graphical analysis of Chap. 4 do not give accurate values of average plate current. The average plate current may be found by the use of more accurate selected-ordinate analyses, or by determining the average ordinate of the plate-current waves constructed from the plate diagram or dynamic transfer characteristic.

Figure 8-10 shows curves of per cent third harmonic derived from the dynamic transfer characteristics of Fig. 8-8 by the use of Eqs. (4-34), (4-35), or (4-36). As explained in Sec. 4-15, the algebraic sign of the harmonic amplitude is merely an indication of the phase relation between the fundamental and the harmonic. When the load current is represented by Eq. (4-22), the numerical value of H_3 will in general be negative if the curvature of the transfer characteristic is predominantly convex upward, and positive if the curvature is predominantly concave upward. The reduction of harmonic content with grid swing at grid swings below 30 volts and reversal to substantial negative values results from the S curvature of that part of the dynamic transfer characteristic which is shared by both tubes.

Distortion.—Figure 8-11 shows curves of distortion factor, as defined by Eq. (4-38), the third and fifth harmonics, only, being considered.

8-4. Power Output.—Accurate determinations of Class B power output must be made either experimentally or by substituting graphically determined values of fundamental plate-current amplitude in the relation

$$P_o = \tfrac{1}{2}H_1{}^2 r_b \qquad \text{(for two tubes)} \tag{8-1}$$

in which r_b is the effective load resistance per tube (one-fourth the effective plate-to-plate load resistance). Figure 8-12 shows curves of power output of two tubes as a function of grid swing. These curves were derived by substitution in Eq. (8-1) of values of H_1 determined by the application of Eqs. (4-35) to the dynamic transfer characteristics of Fig. 8-8.

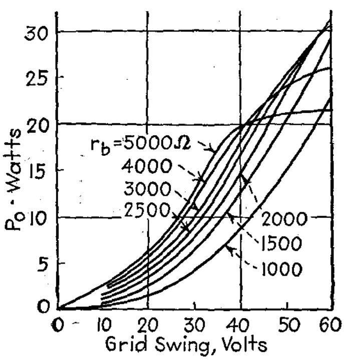


FIG. 8-12.—Power output as a function of grid swing. Type 46 tubes in zero-bias Class B operation.

Approximate expressions for power output can be derived under the assumptions that the tubes are biased to cutoff, so that I_{bo} is zero, and that the distortion is negligible. The wave of plate current of each tube then approximates a half sine wave of amplitude I_{pm}, and the two tubes combine to give sinusoidal output as shown in Fig. 5-20. The approximate power output is

$$P_o = \tfrac{1}{2}I_{pm}{}^2 r_b = \tfrac{1}{2}I_{pm}E_{pm} \qquad \text{(for two tubes)} \tag{8-2}$$

Since the current wave of each tube is assumed to be a half sine wave and since the average value of a half sine wave, averaged over the whole cycle, is the crest value divided by π, the average plate current of each tube is

$$I_{ba} = \frac{I_{pm}}{\pi} \tag{8-3}$$

Substituting Eq. (8-3) in Eq. (8-2) gives for the approximate power output

$$P_o = 4.94 I_{ba}{}^2 r_b \qquad \text{(total power in terms of average current of one tube)} \tag{8-4}$$

The value of Eq. (8-4) lies in the ease with which approximate laboratory measurements of power output of tubes in Class B amplification can be made by the use of a d-c milliammeter. In Fig. 8-13 is shown a simple circuit which may be used for this purpose. The power output is half as great as with two tubes, so the power output in the circuit of Fig. 8-13 is given by

$$P_o = 2.47 I_{ba}^2 r_b \qquad \text{(for one tube)} \tag{8-5}$$

Equations (8-2) to (8-5) cannot be applied with accuracy when the path of operation extends beyond the knee of the dynamic transfer characteristic.

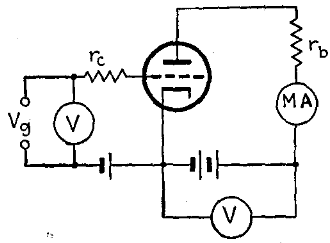


FIG. 8-13.—Circuit for the approximate measurement of power output of tubes in Class B operation.

8-5. Power Input, Plate Dissipation, and Plate-circuit Efficiency.—The power input to the plate circuit is

$$P_i = 2E_{bb}I_{ba} \tag{8-6}$$

where I_{ba} is the average plate current of one tube. An approximate expression for the power input is obtained by substituting Eq. (8-3) in Eq. (8-6), giving

$$P_i = \frac{2}{\pi} E_{bb} I_{pm} \tag{8-7}$$

The approximate plate dissipation is

$$P_p = P_i - P_o = \tfrac{1}{2} I_{pm} \left(\frac{4E_{bb}}{\pi} - E_{pm} \right) \tag{8-8}$$

The approximate plate-circuit efficiency is found by combining Eq. (8-2) and Eq. (8-7).

$$\eta_p = \frac{P_o}{P_i} = \frac{\pi}{4} \frac{E_{pm}}{E_{bb}} = \frac{0.785 E_{pm}}{E_{bb}} \tag{8-9}$$

Equation (8-9) shows that the plate-circuit efficiency of an ideal distortionless Class B amplifier increases with alternating plate-voltage amplitude and that the maximum theoretical efficiency would be obtained if the amplitude of the alternating plate voltage were equal to the plate supply voltage. This would give a value of $78\frac{1}{2}$ per cent. In order to obtain this maximum efficiency, the tubes would have to be biased to cutoff, the plate voltage would have to fall to zero at the positive crest of grid voltage, and the distortion would have to be zero. Although this condition cannot be attained in practice, practical efficiencies of 60 per cent or higher are obtained. Because of this high efficiency, plate dissipation is low, and relatively small tubes can furnish large amounts of power in Class B amplification.[1]

8-6. Use of Hyperbolas in the Determination of Approximate Output, Dissipation, and Efficiency.—For a constant value of power, Eq. (8-2) represents a rectangular hyperbola. As E_{pm} is measured in a negative direction relative to E_{bb}, the hyperbola lies in the second quadrant (with respect to E_{bb} as zero axis) and is asymptotic to the voltage axis and to the line $e_b = E_{bb}$. The hyperbola for a given power output may be constructed by plotting corresponding values of voltage, measured negatively relative to E_{bb}, and current, whose product is equal to $2P_o$, as shown in Fig. 8-14. For a constant value of plate dissipation, Eq. (8-8) also represents a rectangular hyperbola. I_{pm} is infinite when $E_{pm} = 4E_{bb}/\pi$,

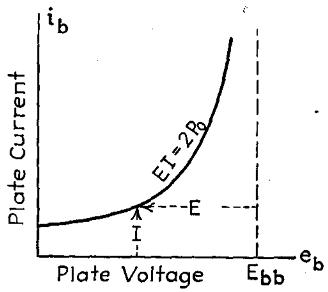


FIG. 8-14.—Class B power-output hyperbola.

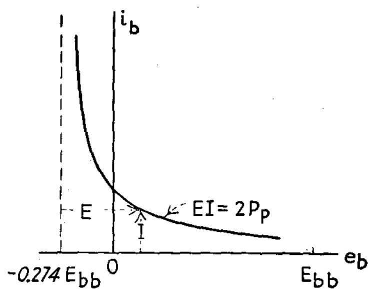


FIG. 8-15.—Class B plate-dissipation hyperbola.

showing that the hyperbola is asymptotic to the line $E_{pm} = 4E_{bb}/\pi$. As E_{pm} is measured with respect to E_{bb}, this asymptote lies at a distance $E_{bb} - (4E_{bb}/\pi) = -0.274E_{bb}$ from the origin, and the hyperbola is in the first quadrant (with respect to the line $-0.274E_{bb}$ as zero axis). The hyperbola for a given value of plate dissipation may be constructed by plotting corresponding values of voltage, measured positively relative to $e_b = -0.274E_{bb}$, and current, whose product is equal to $2P_p$, as shown in Fig. 8-15. If a number of these hyperbolas are plotted on the plate diagram, the approximate values of power output, plate dissipation, and plate-circuit efficiency can be determined at a glance for any load resistance and grid swing.[2] Output and dissipation hyperbolas have been plotted in the plate diagram of the 46 tube in Fig. 8-16. Comparison of power output read from this diagram with the more accurate values plotted in Fig. 8-12 shows that values determined in this manner are

[1] At 60 per cent efficiency the ratio of output to dissipation is 1.5; at 22 per cent efficiency it is approximately 0.28. Therefore, at the same plate voltage and with the same dissipation, one would expect to obtain roughly five times as much power in Class B as in Class A amplification.

[2] DE LA SABLONIÈRE, C. J., *Wireless Eng.*, **12**, 133 (1935).

within 6 or 7 per cent of the correct values if the path of operation does not extend into the region where the static characteristics cease to be approximately parallel. It should be noted that the shape and spacing of these hyperbolas depend only upon the current and voltage scales. They are not altered by change in tube characteristics, and change in E_{bb} merely shifts them. Therefore, if they are plotted on transparent paper they can be used for any tube and plate supply voltage providing that the static characteristics are drawn to the same scales.

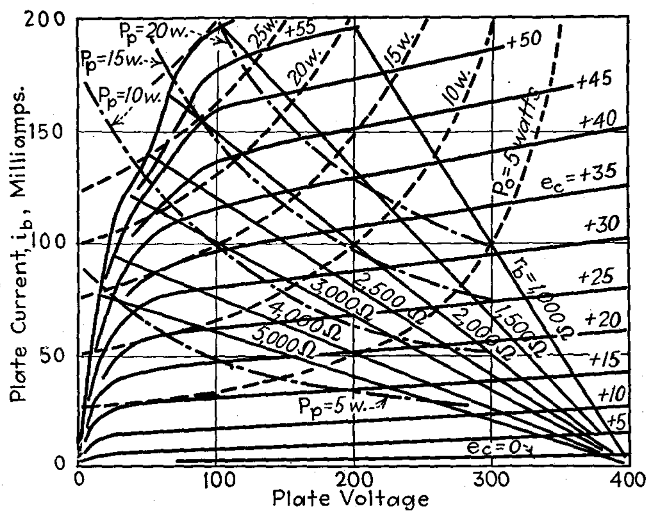


FIG. 8-16.—The use of hyperbolas in the approximate graphical determination of Class B power output and plate dissipation.

8-7. Determination of Optimum Load.[1]—The manner in which the load resistance of a Class B amplifier is chosen depends upon the amount of power available from the preceding stage, which is called the *driver*. If the available driver power is greatly in excess of the power expended in the grid circuit of the Class B tubes, then the grid driving power need not be taken into consideration, and the load resistance may be chosen so as to give the greatest power output consistent with allowable values of distortion, plate and grid dissipation, and maximum peak plate current. Thus, if the available grid driving power is unlimited,

[1] The method of analysis discussed in this section is based upon the families of i_b-e_b and i_c-e_b characteristics. An excellent method of analysis based upon the e_b-e_c characteristics at constant plate current and at constant grid current has been developed by E. L. Chaffee. At present this method is limited in its application by the fact that e_b-e_c characteristics are not ordinarily available. Those interested in the design of Class B and Class C amplifiers should refer to Dr. Chaffee's article, *J. Applied Physics*, **9**, 471 (1938). See also R. L. Sarbacher, *Electronics*, December, 1942, p. 52.

Figs. 8-8, 8-11, 8-12, and 8-16 show that with a load resistance of 2000 ohms (8000 ohms plate-to-plate) type 46 tubes in Class B amplification will give an output of approximately 30 watts without exceeding 5 per cent distortion, the maximum allowable peak plate current of 200 ma, or the maximum allowable plate dissipation of 10 watts per tube. The required average grid driving power can be determined by taking one-half the product of the peak grid voltage by the fundamental component of grid current found graphically from Fig. 8-9 at the chosen load resistance.

The first cost and economy of operation depend among other things upon the amount of power that must be supplied by the driver. When cost and economy of operation are primary considerations, it is sometimes necessary to choose the load resistance in such a manner as to give the highest power output per milliwatt of driver power, consistent with allowable distortion, dissipation, and peak plate current. Because the grid current usually rises rapidly with positive grid voltage, the peak grid power may be high even though the average grid power is small. Peak grid power, therefore, rather than average grid power, is the determining factor. Curves of optimum power output, optimum load, distortion, grid swing, plate dissipation, and efficiency as a function of peak grid power may be derived from the grid-current, power-output, distortion, and plate-dissipation curves of Figs. 8-9, 8-12, 8-11, and 8-16.

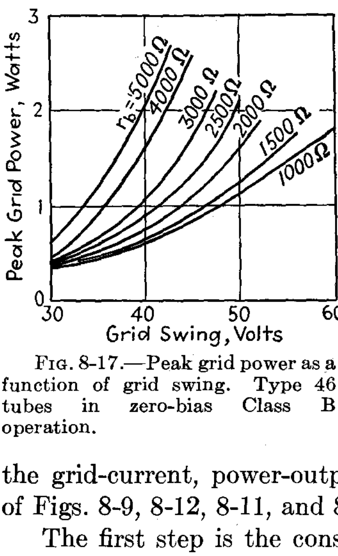


FIG. 8-17.—Peak grid power as a function of grid swing. Type 46 tubes in zero-bias Class B operation.

The first step is the construction of curves of peak grid power as a function of grid swing for different values of load resistance. These curves, shown in Fig. 8-17, are constructed by taking the products of simultaneous values of crest grid current and grid voltage read from the curves of Fig. 8-9 and plotting them against grid swing. Then points on the power-output curves of Fig. 8-12 corresponding to various constant values of peak grid power are found by means of the peak–grid-power curves of Fig. 8-17 and joined, as shown in Fig. 8-18. Thus, Fig. 8-17 shows that for 1-watt peak grid power, the grid swings for 1000-, 2000-, 3000-, and 5000-ohm loads are 48, 44, $39\frac{1}{2}$, and $33\frac{1}{2}$ volts, respectively. These and intermediate values of grid swing determine the 1-watt peak–grid-power contour of Fig. 8-18. An alternative procedure is to use Fig. 8-12 in finding points of the peak–grid-power curves of Fig. 8-17, corresponding to various constant values of power output, and to join these points. This gives the constant power-output contours of Fig.

8-19. The maxima of the constant peak-grid-power contours of Fig. 8-18 or the minima of the constant power-output contours of Fig. 8-19

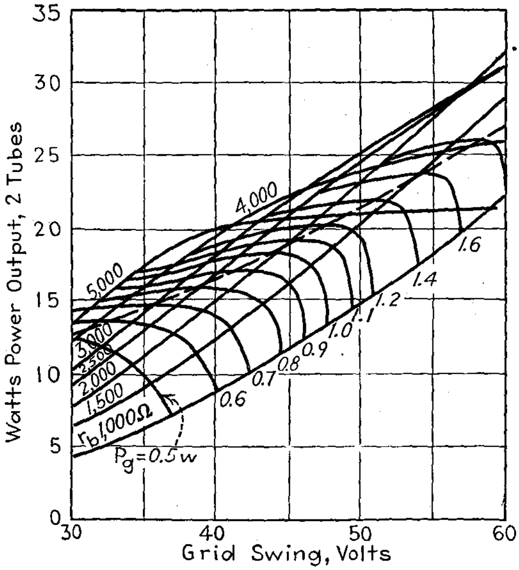


FIG. 8-18.—Power output as a function of grid swing for various values of load resistance r_b and for various values of peak grid power P_g.

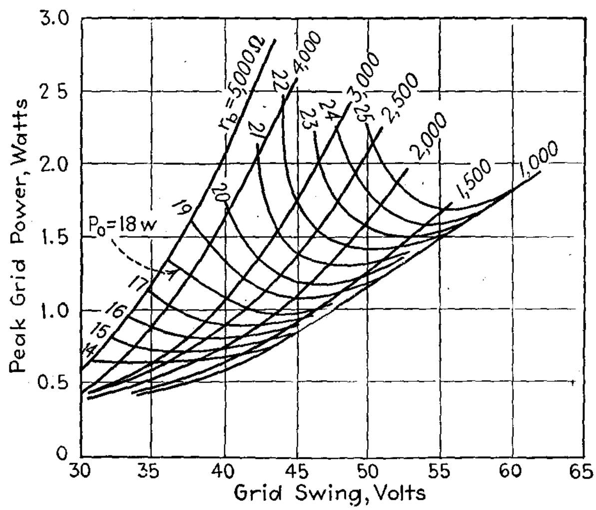


FIG. 8-19.—Peak grid power as a function of grid swing at constant load resistance r_b and at constant power output P_o.

determine simultaneous values of peak grid power, optimum power output, optimum load resistance, and grid swing which may be plotted to give the optimum-power, optimum-load, and grid-swing curves of Fig.

8-20. (The dashed curve of Fig. 8-18 passes through the maxima.) These values of grid swing may be used to determine the plate dissipation from Fig. 8-16, and percentage H_3 and distortion factor from Figs. 8-10 and 8-11. They may also be used in Fig. 8-9 to find the peak grid resistance, which is defined as the ratio of the peak grid voltage to the peak grid current.

Portions of the curves of Fig. 8-20 that lie to the right of the dotted line correspond to plate dissipation in excess of the allowable 20 watts (for two tubes) and so cannot be used. Figure 8-20 therefore shows that if it is essential to keep peak grid power to a minimum, the greatest power output that can be developed with 46 tubes is only 22 watts and is

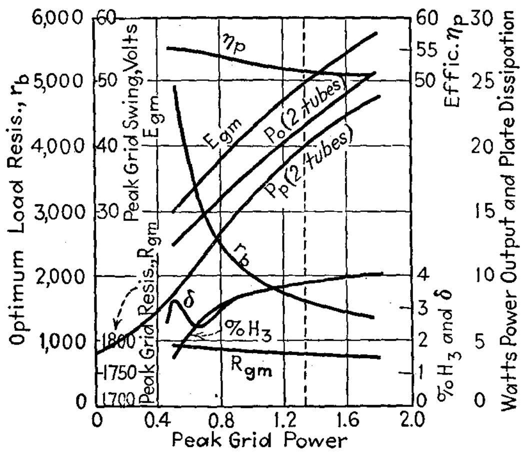


FIG. 8-20.—Optimum power output, optimum load resistance, peak grid swing, plate dissipation, percentage H_3, distortion factor, and peak grid resistance as a function of peak grid power. Type 46 tube in zero-bias Class B operation.

obtained with a load of about 1600 ohms (6400 ohms plate-to-plate). It is interesting to note, however, from Figs. 8-18, 8-16, and 8-11 that at this value of power output an increase of load resistance to 2000 ohms produces only a very small increase of peak grid power and results in an appreciable reduction in required grid swing and plate dissipation and in some reduction in harmonic content. The curves of the 46 tube are, of course, used only for the purpose of illustrating the method of analysis. The final curves for other types of tubes differ considerably and conclusions drawn from the 46 curves do not necessarily apply to other types of tubes.

8-8. Approximate Determination of Maximum Power under Specified Conditions of Operation.—A very simple method is available for the approximate determination of optimum load resistance under any specified limiting condition of operation, such as allowable peak grid power, peak grid current, grid swing, etc.[1] It is merely necessary to draw a curve on the plate diagram giving the locus of simultaneous values of I_{pm} and E_{pm} that will give the specified peak grid power, grid current, etc. The optimum load under the specified condition of operation is that corresponding to the load line whose slope is equal to minus the slope of the locus curve at their point of intersection. The proof of this is simple. In Fig. 8-21 let the curve represent the locus of all simultaneous values of I_{pm} and E_{pm} that will satisfy the specified limiting condition of operation. Under the assumptions that the grid bias is equal to the cutoff value and that distortion is negligible, the power output is

$$P_o = \tfrac{1}{2} I_{pm} E_{pm} \tag{8-2}$$

Under the specified limiting condition, the values of I_{pm} and E_{pm} must correspond to points on the locus. The power is a maximum when

$$\frac{dP_o}{dE_{pm}} = 0 \tag{8-10}$$

or

$$E_{pm}\frac{dI_{pm}}{dE_{pm}} + I_{pm} = 0 \tag{8-11}$$

$$\frac{dI_{pm}}{dE_{pm}} = -\frac{I_{pm}}{E_{pm}} \tag{8-12}$$

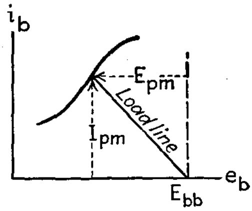


FIG. 8-21.—Class B plate diagram, showing locus of simultaneous values of I_{pm} and E_{pm} under specified limiting condition of operation.

Since I_{pm} and E_{pm} must correspond to points of the locus and must also lie on the load line, dI_{pm}/dE_{pm} is the slope of the locus at the point of intersection with the load line. I_{pm}/E_{pm} is the slope of the load line. Therefore the power output is a maximum when the slope of the locus at the point of intersection is equal to the negative of the slope of the load line. The student will find it interesting to check values determined in this manner with the more accurate determinations that have been discussed.

8-9. Problems of Class B Amplification.—A number of special problems are encountered in the design and operation of Class B amplifiers. These problems and the methods of solving them follow:

1. The variation of power input to the plate circuit with excitation amplitude necessitates the use of a power supply of low regulation. If the plate power is derived from batteries, the internal resistance of the batteries must be low. If the power is obtained from a rectifier, it is advisable to include in the power supply a voltage stabilizer of the type discussed in Sec. 14-13 and shown in Fig. 14-22. It is essential to use a rectifier transformer of low regulation, a low-resistance filter, and mer-

[1] MACFADYEN, K. A., *Wireless Eng.*, **12**, 528 (1935); MOUROMTSEFF, I. E., and KOZANOWSKI, H. N., *Proc. I.R.E.*, **23**, 1224 (1935).

cury-vapor rectifier tubes. Mercury-vapor rectifiers have a low internal voltage drop which remains practically constant over a wide range of current. The characteristics of mercury-vapor tubes and the design of filters are taken up in Chaps. 12 and 14.

2. The flow of grid current in Class B amplifiers necessitates the use of a C-voltage supply of low regulation, or of special tubes that can be operated without biasing voltage. Owing to the variation of average plate current with grid swing and the fact that the tube is biased nearly to cutoff, cathode biasing resistors cannot be used without great reduction of power output and production of excessive distortion.

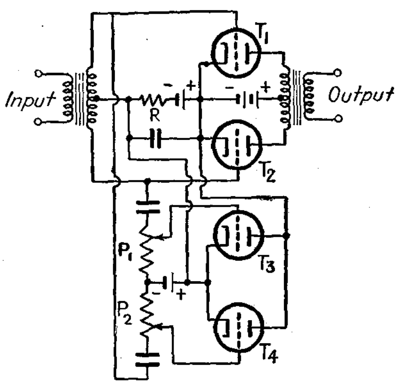


FIG. 8-22.—Circuit for preventing distortion resulting from C-supply impedance.

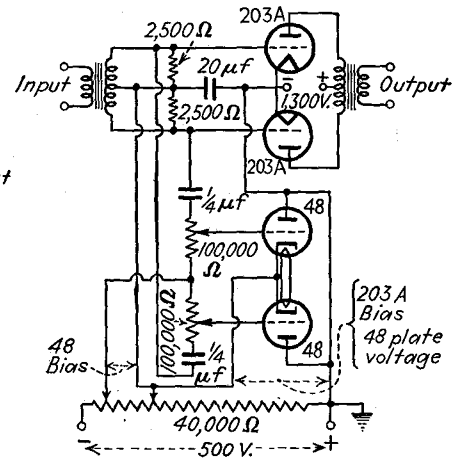


FIG. 8-23.—Practical form of the circuit of Fig. 8-22.

Because of curvature of the grid-current characteristics and the flow of grid current during only a portion of the cycle, C-supply regulation distorts the alternating grid voltage. The simplest method of overcoming this difficulty is by the use of a voltage stabilizer in the output of the C supply. A voltage-stabilizer circuit designed particularly for this type of service is discussed in Sec. 14-13 and shown in Fig. 14-24. A second method was devised by Rockwell and Platts.[1] Their basic circuit is illustrated in Fig. 8-22. T_1 and T_2 are the amplifier tubes, and T_3 and T_4 are auxiliary compensating tubes. R is the resistance of the C-supply. Without T_3 and T_4, the flow of grid current through R during the portion of the cycle in which the grids are positive would cause a voltage drop through R that would distort the alternating grid voltage. The grid bias of the auxiliary tubes T_3 and T_4 is adjusted, however, so that plate current starts flowing in one of these tubes at the same instant

[1] ROCKWELL, R. J., and PLATTS, G. F., *Proc. I.R.E.*, **24**, 553 (1936).

that grid current starts flowing in the corresponding amplifier tube. By adjustment of P_1 and P_2 the plate current of the auxiliary tubes can be made to increase in approximately the same manner as the grid current of the main tubes, so that most of the grid current of the amplifier tubes flows through the auxiliary tubes, rather than through R, and the distortion caused by the voltage drop through R is reduced. Figure 8-23 shows a practical form of this circuit. By overcompensating, so that the negative grid bias decreases as the grid swings positive, it is possible to reduce the distortion caused by the reduction of plate voltage at the current peaks because of plate supply regulation. Two sets of auxiliary tubes may be used, one to compensate for grid-supply regulation, and the other for plate supply regulation. Rockwell and Platts report 2 per cent distortion in the circuit of Fig. 8-23 at 90 watts output, as compared with 5.3 per cent distortion with uncompensated battery bias and 9 per cent distortion with uncompensated rectifier bias.

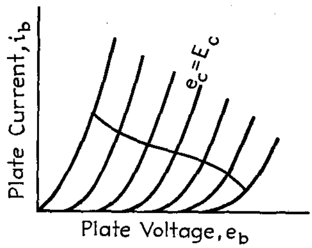


FIG. 8-24.—Curvature of path of operation of driver resulting from variation of effective load.

For relatively small power output a number of special zero-bias Class B tubes have been designed. Examples of these are the 46, the 79, the 53, and the 6N7. In order to obtain the required low value of operating plate current without grid bias, the tubes must be designed to have high amplification factors. (This follows from principles discussed in Chap. 3. High amplification factor is attained by making the shielding of the grid so great that little field from the plate penetrates to the cathode. The field at the cathode, and hence the plate current, is then small at zero grid voltage.) In the 46 tube the high amplification factor is obtained by the use of two grids, which are connected together for Class B operation. By connecting the second grid to the plate this tube can also be used as a Class A amplifier.

3. Power expended in the grid circuit of the Class B tubes requires the use of a driver stage capable of supplying, without excessive distortion, the necessary peak grid-power and the power lost in the coupling transformer. Because of variation of grid resistance of the Class B tube, evidenced by curvature of the grid-current characteristic and by the fact that grid current flows during only part of the cycle, the effective load transferred to the driver varies during the cycle. If the coupling transformer has small leakage reactance, the driver output voltage is in phase with the Class B grid voltage, and the effective driver load resistance is a minimum at the ends of the path of operation and a maximum at the middle. The path of operation of the driver is curved,

as shown in Fig. 8-24.[1] The curvature of the driver load line tends to flatten the peaks of driver output voltage and thus adds to the distortion of the Class B stage resulting from bending over of the dynamic transfer characteristics. This type of driver distortion is reduced by increasing the optimum driver power relative to the peak grid power required by the Class B tubes. For this reason it is always best, when possible, to use a driver whose power rating is sufficiently high so that it is essentially unloaded.

4. The impedance drop in the grid circuit produced by the effective grid-circuit impedance causes the instantaneous grid voltage to differ from the instantaneous applied excitation voltage. As the result of lack of proportionality between grid current and grid voltage, this distorts the grid voltage. Under certain conditions the distortion caused by resistance in the grid circuit may cancel a portion of the plate-circuit distortion, but in general grid-circuit distortion increases the total distortion. By means of a graphical analysis based upon the grid-current characteristic, McLean has shown that the presence of inductance in the grid circuit results in the production of high-order harmonics (ninth and higher) of appreciable magnitude in the grid voltage.[2] Emphasizing of the higher harmonics by the inductance is to be expected, since the inductive reactance is proportional to the frequency.

McLean's method of analysis consists of constructing a curve of grid current *vs.* electrical degrees of the excitation voltage for sinusoidal excitation of frequency f and analyzing this curve for harmonics up to the 15th. The grid-current series,

$$i_g = \Sigma I_n \sin 2\pi nf,$$

derived in this manner, is then substituted in the equation

$$e_d = r_c i_g + L_c \frac{di_g}{dt} \tag{8-13}$$

in which e_d is the distortion voltage, r_c is the effective grid-circuit resistance, and L_c is the total effective leakage inductance of the coupling transformer, referred to the secondary. r_c is approximately equal to $n^2 r_p$, where n is the secondary-to-primary turn ratio of the coupling transformer, and r_p is the plate resistance of the driver [see Eq. (7-56)]. McLean's work shows the importance of keeping the leakage reactance of the driver output transformer as small as possible. Since $r_c = n^2 r_p$,

[1] At small values of instantaneous alternating grid voltage the load on the driver is principally the primary reactance of the transformer, which tends to make this part of the path of operation elliptical. Consequently the path of operation may contain a loop about the operating point.

[2] McLEAN, T., *Proc. I.R.E.*, **24**, 487 (1936).

it is evident that the plate resistance of the driver should be as low as possible and that a step-down transformation ratio should be used to reduce the effective grid-circuit resistance. High driver output voltage (measured across the primary of the transformer) is also advantageous, since, for a given grid swing of the output tubes, the required value of n varies directly with the driver output voltage, and therefore r_c varies inversely as the square of this voltage. For this reason it is desirable to operate the driver amplifier at as high direct plate voltage as possible.

Some writers have contended that the effect of the varying grid resistance of Class B tubes can be reduced by shunting the secondary of the driver transformer with a resistance. McLean proved, however, that the effective grid-circuit impedance, and hence the distortion, is always increased by loading the transformer.[1]

The use of inverse feedback in the driver stage instead of step-down transformation ratios to reduce the effective grid-circuit resistance has been suggested.[2] As pointed out on page 207, the terminal impedance of certain inverse-feedback circuits is reduced by negative feedback in the ratio $1/(1 - A\beta)$. The use of negative feedback has the advantage that inductance-capacitance coupling can be used in place of transformer coupling, thus eliminating the objectionable effect of transformer leakage inductance.

5. Because of secondary emission effects, the form of the grid-current characteristic varies greatly in different tubes and with tube age. Secondary emission tends to increase with tube age. For this reason it is often difficult to match tubes closely. If secondary emission is so great that the grid characteristic has a region of negative slope, there is danger that oscillation of the dynatron type may take place (Sec. 10-22). This difficulty may be prevented by shunting the grids with small rectifier tubes biased so that current starts to flow at the beginning of the negative-resistance region.[3]

6. Distortion resulting from curvature of that portion of the dynamic transfer characteristic which is shared by both tubes may be high at small grid swings. At large grid swings the amplitudes of harmonics introduced by this curvature are small, and the order of the harmonics is high. Fortunately, the total distortion may be kept within allowable limits. This type of small-signal distortion can be controlled to some extent by tube design and is less in tubes that require grid bias than in zero-bias tubes such as the 46. This can be shown by comparison of the composite plate characteristics of the two types of tubes. As can be

[1] See also L. E. BARTON, *Proc. I.R.E.*, **23**, 557 (1935).

[2] BAGGALLY, W., *Wireless Eng.*, **10**, 65 (1933); MACFADYEN, K. A., *Wireless Eng.*, **12**, 642 (1935).

[3] MCLEAN, *loc. cit.*

seen from Fig. 8-7, the composite characteristics of zero-bias tubes are constructed by combining static plate characteristics of positive grid voltage with those of negative grid voltage. Because of differences of shape and spacing of the curves for positive and negative grid voltages,

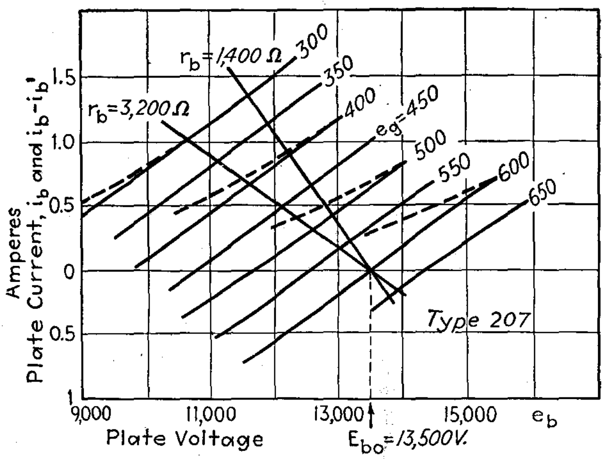


FIG. 8-25.—Composite plate characteristics of type 207 tubes at small values of grid swing. E_{bo} = 13,500 volts; E_c = −600 volts.

the composite characteristics are not straight, parallel, or equidistant, and so the dynamic transfer characteristic is curved at small-signal voltages. If the tubes operate with appreciable negative bias, on the other hand, the composite characteristics that are intersected by the load line are constructed by combining pairs of static characteristics, both of which are for negative grid voltages, and so the composite characteristics are more nearly straight, parallel, equidistant lines. Figure 8-25 shows composite characteristics of the type 207 tube at an operating plate voltage of 13,500 volts and a negative bias of 600 volts. Figure 8-26 shows percentage H_3 and distortion-factor curves at small grid swings derived from Fig. 8-25. The maximum small-signal distortion is 3 per cent as compared with 5 per cent for the type 46 tube.

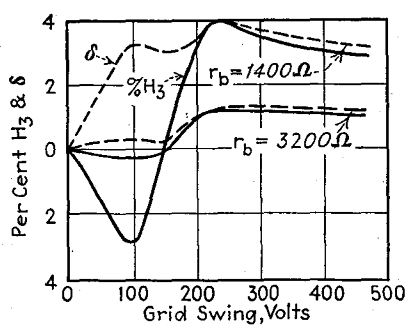


FIG. 8-26.—Curves of per cent H_3 and of distortion factor for type 207 tubes in biased Class B operation.

8-10. Driver Transformer Design.—When a Class B amplifier is designed on the basis of limited grid driving power, the design of the coupling transformer between the driver and output stages is considerably more complicated than when the driver is essentially unloaded. The limited-grid-power design has been discussed in detail in a bulletin issued by RCA.[1] When the driver is essentially unloaded, the ratio of transformation should be chosen so that grid overloading occurs simultaneously in the driver and output stages (see Sec. 5-2). The voltage output of the driver should be high enough so that a step-down transformer can be used. Leakage reactance should under all circumstances be kept as small as possible. At signal amplitudes that are so small that grid current does not flow in the output stage the secondary of the transformer is unloaded, and principles of Class A coupling transformer design apply (see Sec. 6-14). To ensure adequate response at low frequency, the primary inductance must be high. In a step-down transformer this requirement is not difficult to attain.

8-11. Class AB2 Operation.—The advantages of Class AB1 operation of push-pull amplifiers were discussed in Chap. 7. Even greater power output without excessive nonlinear distortion can be obtained by Class AB2 operation. The advantages of Class B operation are, in fact, attained in part in Class AB2 operation without some of the Class B disadvantages. The power output and plate-circuit efficiency, although less than in Class B operation, are considerably higher than in Class A operation. Small-signal distortion is avoided because the tubes operate Class A at small-signal voltage. The greatly reduced grid driving power as compared with that in Class B amplifiers simplifies the design of the driver stage and results in small grid-circuit distortion. Grid driving power need not be taken into consideration in the determination of optimum load.

Self-bias may be used in Class AB2 amplifiers, but the large increase of average plate current causes the bias to increase with signal voltage and thus reduce the optimum power output. This is clearly shown by Fig. 8-27, in which the power output of two type 45 tubes in Class AB2 push-pull operation is plotted against zero-signal grid bias, for fixed bias operation and for self-bias operation.[2] With self-bias the power output at 5 per cent distortion decreases with increase of zero-signal bias. It is therefore inadvisable to use higher bias than necessary to keep the zero-signal dissipation within the allowable limit. The cathode resistor must be shunted by an ample by-pass condenser. Increase of average plate current with signal amplitude necessitates the use of a B supply of low regulation in Class AB2 amplifiers.

An example of the large output that can be obtained with small tubes and moderate plate voltage in Class AB2 operation is given by the 6L6 beam power tube, which will deliver an output of 60 watts (two tubes)

[1] *RCA Laboratory Series UL-1.*

[2] RCA application note 40.

with 65 per cent plate-circuit efficiency (61 per cent including screen-circuit losses), and 2 per cent total harmonic, at a plate voltage of 400 and an average grid driving power of 0.35 watt. In Class AB1 operation the output is 32 watts at approximately 63 per cent efficiency (58 per cent

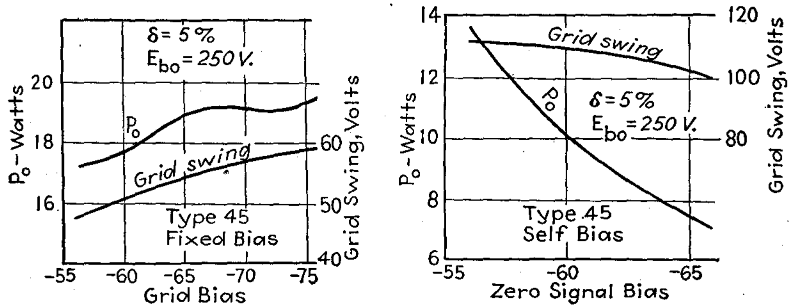


FIG. 8-27.—Power output of two type 45 tubes in push-pull operation as a function of grid bias.

including screen power, 54 per cent including screen and biasing power, and 45 per cent including also cathode heating power).

8-12. Negative-feedback Class B and Class AB Amplifiers.—The distortion in Class B and Class AB amplifiers can be greatly reduced by the use of inverse feedback. A typical Class B amplifier employing negative feedback is shown in Fig. 8-28.[1] Figure 8-29 shows the great

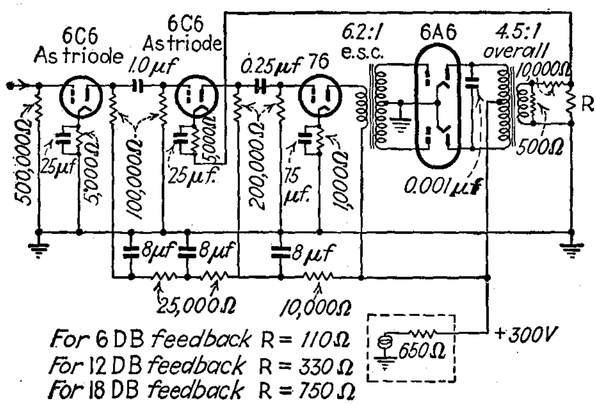


FIG. 8-28.—Class B amplifier with inverse feedback.

reduction in harmonic content that results from the use of negative feedback in a Class B amplifier.

8-13. Class C Amplification.[2]—An analysis of the theory and design of Class C amplifiers is beyond the scope of this book. The plate-circuit

[1] DAY, J. R., and RUSSELL, J. B., *Electronics*, April, 1937, p. 16.
[2] See Supplementary Bibliography.

efficiency of Class C amplifiers is higher than that of Class B amplifiers; measured efficiency of 85 per cent may be attained in practice. Because plate current flows during only a portion of the cycle, this type of ampli-

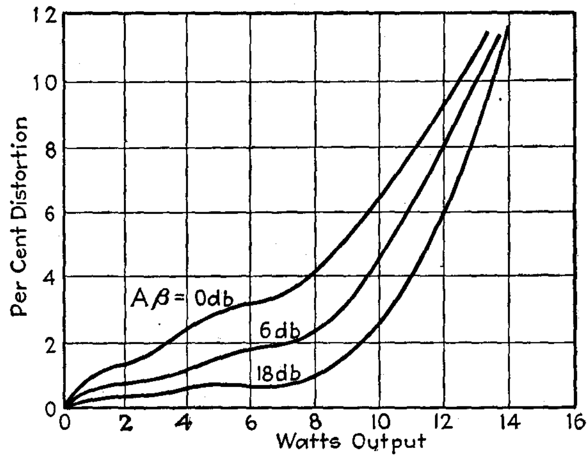


FIG. 8-29.—Amplitude distortion curves for the amplifier of Fig. 8-28, with and without feedback.

fier is employed only in radio-frequency amplification, where sharply tuned resonant circuits may be used to suppress harmonics.

Problems

8-1. *a.* Construct curves similar to those of Figs. 8-8 to 8-12 for type 53 tubes in zero-bias Class B operation. From these curves determine the optimum plate-to-plate load resistance, assuming no limitation on grid driving power.

b. From the i_c-e_c characteristic corresponding to the load found in (*a*), determine the fundamental component of grid current by graphical analysis, and from it find the average grid driving power required to deliver the full power output.

c. Find the plate-circuit efficiency at optimum output.

8-2. Prove that it is impossible to obtain an output of 40 watts in Class B amplification from two tubes that have a rated plate dissipation of 5 watts each.

Supplementary Bibliography

Class B and Class AB Amplification

SPITZER, E. E.: *Proc. I.R.E.*, **17**, 985 (1929).
ARDENNE, VON, M.: *Hochfrequenztech. u. Elektroacustik*, **34**, 161 (1929).
FARRAR, C. L.: *Electronics*, June, 1932, p. 196.
SCROGGIE, M. G.: *Wireless World*, **33**, 2 (1933).
NELSON, J. R.: *Proc. I.R.E.*, **21**, 858 (1933).
MACFADYEN, K. A.: *Wireless World*, **33**, 454 (1933).
THOMAS, H. P.: *Proc. I.R.E.*, **21**, 1134 (1933).
SAH, A. P. T.: *Proc. I.R.E.*, **21**, 1616 (1933).
MILLER, B. F.: *Proc. I.R.E.*, **23**, 496 (1935).
STAFFORD, F. R. W.: *Wireless Eng.*, **12**, 539 (1935).

KOEHLER, G.: *Electronics*, February, 1936, p. 14.
EVERITT, W. L.: *Proc. I.R.E.*, **24**, 305 (1936).
Electronics, November, 1937, p. 43.

Class C Amplification

FAY, C. E.: *Proc. I.R.E.*, **20**, 548 (1932).
OSBORN, P. H.: *Proc. I.R.E.*, **20**, 813 (1932).
EVERITT, W. L.: *Proc. I.R.E.*, **22**, 152 (1934).
TERMAN, F. E., and FERNS, J. H.: *Proc. I.R.E.*, **22**, 359 (1934).
MILLER, B. F.: *Proc. I.R.E.*: **23**, 496 (1935).
MOUROMTSEFF, I. E., and KOZANOWSKI, H. N.: *Proc. I.R.E.*, **23**, 752 (1935).
TERMAN, F. E., and KOAKE, W. C.: *Proc. I.R.E.*, **24**, 620 (1936).
REED, M.: *Wireless Eng.*, **12**, 296 (1935).
SCOTT, H. J., and BLACK, L. J.: *Proc. I.R.E.*, **26**, 449 (1938).
EVERITT, W. L.: "Communication Engineering," 2d ed., McGraw-Hill Book Company, Inc., New York, 1937.
TERMAN, F. E.: "Radio Engineering," 2d ed., McGraw-Hill Book Company, Inc., New York, 1937.
GLASGOW, R. S.: "Principles of Radio Engineering," McGraw-Hill Book Company, Inc., New York, 1936.
EASTMAN, A. V.: "Fundamentals of Vacuum Tubes," 2d ed., McGraw-Hill Book Company, Inc., New York, 1941.

CHAPTER 9

MODULATION AND DETECTION

Transmission of intelligence through space by means of electromagnetic radiation cannot be accomplished satisfactorily at audio frequencies. There are several reasons for this: (1) The radiation efficiency of antennas is very low at audio frequencies, and so the range is small. (2) Efficient radiation and reception of electromagnetic waves require the use of antennas and circuits tuned to the frequencies of the waves. The antennas required at audio frequencies would be impractical because of their great lengths, and they would not respond equally well to all frequencies in the audio range. (3) If transmission were effected at audio frequencies, all transmitters would operate over the same frequency range and so the programs of various transmitters would be heard simultaneously at the receiver. These difficulties are avoided by radiating a radio-frequency wave, the amplitude, frequency, or phase of which is varied in accordance with the audio-frequency signals that it is desired to transmit. At the receiver, the variations of amplitude, frequency, or phase of the received wave are reconverted into a-f voltages. The processes of modulation and detection, whereby radio transmission is made possible, afford two very important fields of application of vacuum tubes.

9-1. Modulation.—The process whereby some characteristic, usually amplitude, frequency, or phase, of a sinusoidally changing voltage, current, or other quantity is varied in accordance with the time variations of another voltage, current, or other quantity, is called *modulation*. The term *carrier* is applied to the quantity the characteristic of which is varied, and the term *modulation* (*signal*) to the quantity in accordance with which the variation is performed. The *carrier frequency* is the frequency of the unmodulated carrier. Usually the modulation frequency is considerably lower than the carrier frequency.

Since the sounds of the human voice and of musical instruments involve a large number of frequencies produced simultaneously, the form of the modulating wave is rarely sinusoidal. To simplify theoretical analyses, however, the assumption is generally made that the modulation is sinusoidal. It is then necessary to investigate also the effects of the interaction of two or more components of a complex modulating wave in the modulation process and in the process of reproducing these components from the modulated wave.

It will be shown that both the process of modulation and that of recovering the original modulation frequencies, which is one form of a more general process called *detection* (Sec. 9-10), involve the generation of frequencies not present in the impressed excitation. The similarity of the processes of modulation and detection by vacuum tubes is particularly marked in the case of amplitude modulation. They differ mainly as to the frequencies impressed upon the circuit and the frequencies selected from the output by some form of filter.

The production and detection of amplitude-modulated voltages is accomplished electronically by means of some form of rectifier, which is a device that gives unequal changes of current in response to changes of voltage of equal amplitude but opposite sign. Rectification is said to be complete if the change in current in response to a change of voltage of one sign is zero. Complete rectification may be obtained with a device that conducts in one direction only. Such a device is said to be a perfect rectifier. Partial rectification takes place either because of curvature of the current-voltage characteristic of the rectifier, or because of unequal resistances to flow of current in the two directions. The high-vacuum diode is an example of a device that may be used to obtain either complete or partial rectification. When rectification is accomplished by means of curvature of the characteristic of a rectifier, the characteristic curve is usually continuous over the range of operation. When this is true, a series expansion of current in terms of voltage may be used in the theoretical analysis of the modulator or detector. When the operating point is at a point of singularity, however, as must be true with perfect rectification, or if the range of operation becomes so great that it includes a singular point, then other methods of analysis must be employed. For this reason it is convenient to consider separately modulation and detection by characteristic curvature and by complete rectification.

Rectification in which the average current of an electrode is changed by application of alternating voltage to that electrode is called *simple rectification;* rectification in which the average current of an electrode is changed by the application of alternating voltage to another electrode is called *transrectification.*

9-2. Amplitude Modulation.—A quantity y varying sinusoidally with an amplitude A, an angular frequency ω_k, and a phase angle ϕ may be represented by the equation

$$y = A \sin (\omega_k t + \phi) \tag{9-1}$$

Since the phase angle ϕ is constant in amplitude modulation and is determined by the instant in the carrier cycle at which it is decided to start observing time, there is no important loss of generality in setting ϕ equal to zero.

Let the phase angle be zero and the amplitude be varied sinusoidally at modulation frequency ω_m between the limits $1 + M$ and $1 - M$, as indicated by the equation

$$A = K(1 + M \sin \omega_m t) \tag{9-2}$$

in which K and M are constants and ω_m is less than ω_k. Equation (9-1) then becomes

$$y = K(1 + M \sin \omega_m t) \sin \omega_k t \tag{9-3}$$

Equation (9-3) represents a wave the frequency of which is constant and the amplitude of which is varied at modulation frequency between the

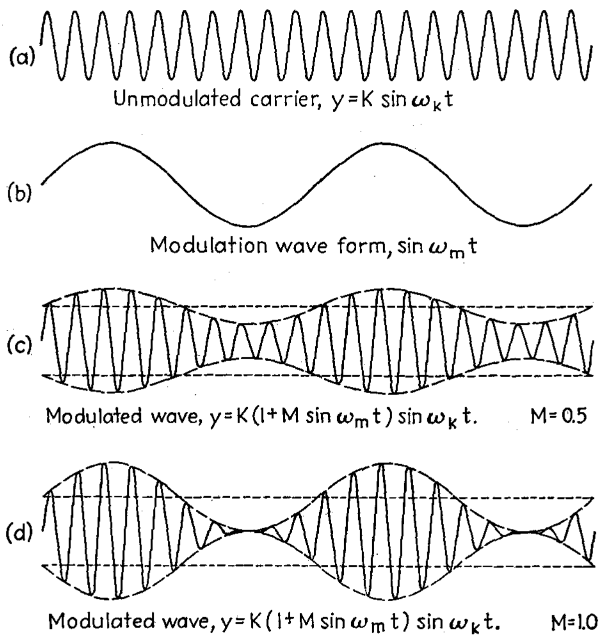


FIG. 9-1.—Amplitude modulation.

limits $1 + M$ and $1 - M$, as shown in Fig. 9-1 for several values of M. The coefficient M, which is called the *modulation factor* or *degree of modulation*, determines the extent to which the carrier is modulated. When M is zero, the carrier is unmodulated and its amplitude remains constant, as shown by curve a of Fig. 9-1. When M lies between zero and unity, the modulated wave is of the form of curve c of Fig. 9-1. When M is unity, the amplitude varies between zero and twice the unmodulated value, as shown by curve d of Fig. 9-1, and modulation is said to be *complete*.

It should be noted that there is no direct relation between the amplitude of the envelope of the modulated wave c or d and the modulation-frequency wave b in accordance with which the modulation is performed. The degree of modulation associated with modulation excitation of given amplitude depends not only upon the amplitudes of the carrier and modulation excitations, but also upon the characteristics of the circuit by means of which the carrier is modulated. In particular, the modulation factor is not necessarily unity when the carrier and modulation-excitation voltages impressed upon a modulator are equal in amplitude. The wave form and frequency of the envelope are, however, the same as those of the modulation excitation applied to the modulator unless the modulator introduces distortion in the form of undesired frequency components.

In the simple case that has been discussed, in which the wave is assumed to be sinusoidally modulated, the modulation factor is equal to the ratio of the difference between the maximum or minimum amplitude of the modulated wave and the amplitude of the unmodulated carrier wave to the amplitude of the unmodulated carrier wave. When the modulating wave is not sinusoidal, however, the maximum increase and decrease of amplitude from the unmodulated value may not be equal, and a more general definition of modulation factor is necessary. The modulation factor is defined as the maximum departure (positive or negative) of the envelope of the modulated wave from its unmodulated value, divided by its unmodulated value. The modulation factor times 100 per cent is called the *percentage modulation.* Sometimes in the discussion of an asymmetrically modulated wave it is necessary to distinguish between the modulation factors corresponding to the minimum and the maximum amplitudes of the modulated waves. These are called the *inward* and *outward* modulation factors, respectively.

9-3. Side Frequencies in Amplitude Modulation.—By trigonometrical expansion, Eq. (9-3) may be changed into the form

$$y = K \sin \omega_k t - \tfrac{1}{2}KM[\cos (\omega_k + \omega_m)t - \cos (\omega_k - \omega_m)t] \qquad (9\text{-}4)$$

Equation (9-4) shows that an amplitude-modulated wave is made up of three components having frequencies equal to the carrier frequency, the carrier frequency plus the modulation frequency, and the carrier frequency minus the modulation frequency. The process of amplitude modulation therefore involves the generation of sum and difference frequencies, called the *upper side frequency* and *the lower side frequency,* respectively. When the carrier is modulated by a band of modulation frequencies, applied individually or simultaneously, the side frequencies lie in bands, called the *upper side band* and *lower side band.* The frequency width of the side bands is the same as that of the modulation-frequency band. It is important to note that the modulation frequency is not in itself present in the modulated wave.

For reasons stated in the introduction to this chapter, carrier frequencies used in transmission by means of electromagnetic radiation lie in the radio-frequency range. It is entirely possible, however, to modulate an audio-frequency carrier by means of signals of lower audio frequency. This is done in carrier telephony over wires.[1] If telephone signals were transmitted only at their original audio frequencies, only one conversation could be transmitted at one time over a pair of wires. By modulating carrier frequencies with the voice frequencies, however, a large number of conversations can be transmitted simultaneously. Suppose, for instance, that the range from 100 to 2500 cps is adequate for the transmission of intelligible speech. Then the modulation of

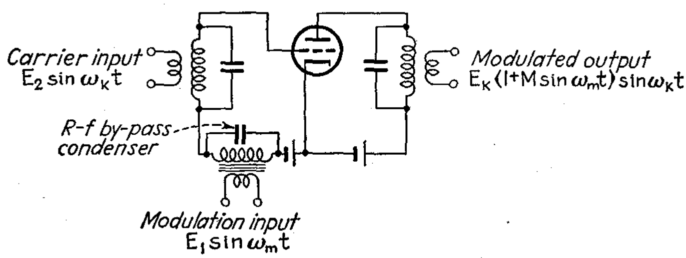


FIG. 9-2.—Van der Bijl modulator.

carriers of 3000, 6000, 9000, 12,000 cps, etc., will result in the production of upper side bands in the ranges 3100 to 5500, 6100 to 8500, 9100 to 11,500, 12,100 to 14,500 cps, etc. The carrier and lower-side-band components of the modulated waves can be removed by means of filters. The upper side bands, together with the original voice band, can then be transmitted simultaneously without interference. This is known as single-side-band transmission. At the receiving end the frequencies in the various channels are separated by band-pass filters and reconverted to the original 100- to 2500-cycle range by a similar process. In modern carrier telephony systems carrier frequencies of 20 kc or higher are used.

9-4. Amplitude Modulation by Curvature of Tube Characteristics.—Amplitude modulation by means of vacuum tubes is based upon intermodulation resulting from either curvature of an electrode or transfer characteristic or from the cutting off of an electrode current. Although, for reasons that will be explained, modulation by curvature of tube characteristics is now seldom used, the theory is nevertheless of importance. Figure 9-2 shows the circuit diagram of the van der Bijl modulator, which makes use of curvature of the transfer characteristic. In this circuit the carrier and modulation frequencies are impressed upon the grid

[1] COLPITTS, E. H., and BLACKWELL, O. B., *J. Am. Inst. Elec. Eng.*, **40**, 205 (1921).

circuit of a triode or multigrid tube and the modulated voltage is taken from a parallel-resonant plate circuit, tuned to the carrier frequency.

A graphical explanation of the operation of the van der Bijl modulator can be given with the aid of a transfer diagram. Figure 9-3 shows that, when only the carrier excitation is impressed upon the grid, the amplitude of the carrier-frequency plate current is much greater for the smaller grid bias E_c' than for the larger bias E_c''. If the bias is varied between these values, the amplitude of the carrier component of the plate current also varies. The bias is varied in effect at modulation frequency by choosing an intermediate value of bias E_c and also applying the modulation excitation voltage, as shown in Fig. 9-4. The wave a of plate current may be separated into the components b and c of Fig. 9-4. Curve b may be further resolved into a steady component, a modulation-frequency component, and modulation-harmonic components; curve c into a varying-amplitude carrier-frequency component (carrier plus side-

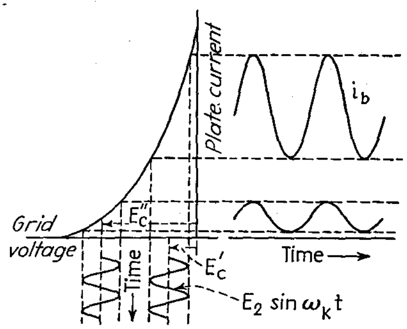


FIG. 9-3.—Variation of amplitude of carrier-frequency plate current with grid bias.

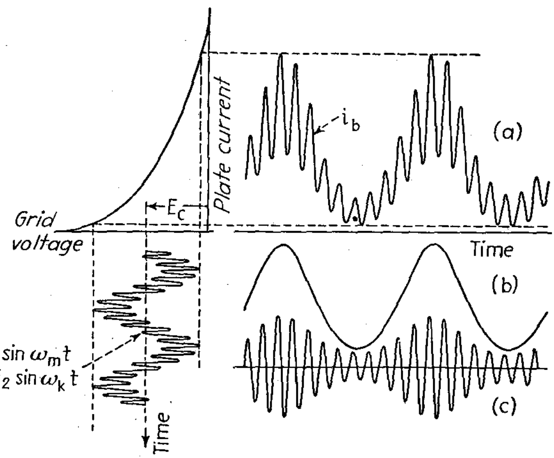


FIG. 9-4.—Amplitude modulation by curvature of the transfer characteristic. Waves (b) and (c) are two components into which the wave (a) of plate current may be resolved.

frequency components) and carrier-harmonic components. Since the plate load is tuned to the carrier frequency, the impedance of the plate circuit is high at the carrier frequency and at the side frequencies, which do not differ greatly from the carrier frequency (inasmuch as the modulation frequency is normally much lower than the carrier frequency), but negligible at the other frequencies contained in the plate current.[1] The voltage appearing across the load is therefore of the form of the varying-amplitude carrier component of curve c of Fig. 9-4, *i.e.*, the desired modulated wave.

A more rigorous analysis of modulation by characteristic curvature may be made by means of the series expansion for alternating plate current. The alternating plate current, as given by the first two terms of the series expansion, is

$$i_p = \Sigma a_1 e + \Sigma a_2 e^2 \tag{9-5}$$

where

$$e = e_g + \frac{v_p}{\mu} \tag{3-42}$$

If the excitation voltage e contains two components $E_2 \sin \omega_k t$ and $E_1 \sin \omega_m t$,

$$e = E_2 \sin \omega_k t + E_1 \sin \omega_m t \tag{9-6}$$

and the alternating plate current is

$$i_p = (a_1)_k E_2 \sin \omega_k t + (a_1)_m E_1 \sin \omega_m t + \Sigma a_2 (E_2 \sin \omega_k t + E_1 \sin \omega_m t)^2 \tag{9-7}$$

Expansion of Eq. (9-7) shows that the alternating plate current contains components whose frequencies are the carrier frequency, twice the carrier frequency, the modulation frequency, twice the modulation frequency, and the upper and lower side frequencies, as well as steady components. As pointed out in the last paragraph, since the impedance of the load is high only at the carrier and side frequencies, only these frequencies have appreciable amplitudes in the output. The output voltage is therefore equal to the sum of the products of the carrier and side-frequency currents by the corresponding values of the load impedance at these frequencies.

$$e_o = a_1 (z_b)_k E_2 \sin \omega_k t + (a_2)_{k+m} (z_b)_{k+m} E_1 E_2 \sin (\omega_k + \omega_m) t - (a_2)_{k-m} (z_b)_{k-m} E_1 E_2 \sin (\omega_k - \omega_m) t \tag{9-8}$$

If the resonance curve of the plate load is symmetrical and the circuit is

[1] The tuning of the resonant circuit must be broad enough so that the circuit can respond to the side frequencies. That the tuning should not be too sharp also follows from the fact that high selectivity is an indication of low damping. If the damping is too low, the amplitude of oscillation of the resonant circuit does not decrease in response to modulation-frequency changes of alternating plate current amplitude, and so the tuned circuit will tend to oscillate at constant amplitude. Hence the output voltage will be a constant-amplitude carrier-frequency voltage, instead of the desired amplitude-modulated voltage.

accurately tuned to the carrier frequency, the load impedance is the same at the two side frequencies and the coefficients $(a_2)_{k+m}$ and $(a_2)_{k-m}$ are equal and may be replaced by the symbol a_2. Equation (9-8) is then identical with Eq. (9-4) if the following substitutions are made in Eq. (9-4):

$$K = a_1(z_b)_k E_2 \tag{9-9}$$

and

$$M = \frac{2a_2(z_b)_{k+m}E_1}{a_1(z_b)_k} \tag{9-10}$$

This proves that the output voltage is an amplitude-modulated wave.

Because Eq. (9-5) may be used to express the alternating plate current of a diode or a multigrid tube, as well as that of a triode, it follows that any type of vacuum tube may be used for this type of modulation. The two excitation voltages may be impressed upon the same electrode circuit, or upon the circuits of any two electrodes. Since $e = e_g + v_p/\mu$ in Eq. (9-5), lower excitation voltages are required when they are impressed upon the control-grid circuit of a triode or multigrid tube than when they are impressed upon the plate circuit of a diode. Because the production of side frequencies in this type of modulation is associated with the second-order term of the series expansion, it is referred to as *parabolic*, or *square-law, modulation.*

It can be seen from Eq. (9-10) that the modulation factor in this type of modulation is proportional to the amplitude of the modulation voltage and to the ratio a_2/a_1, but is independent of the amplitude of the carrier excitation. The amplitude of the side frequencies is proportional to both the carrier excitation amplitude and the modulation excitation amplitude and to a_2. Equations (3-57) and (3-60) show that, when μ is constant or nearly constant, a_2 and a_2/a_1 increase with $\partial r_p/\partial e_b$, and thus indicate that it is desirable, in square-law modulation, to operate at a point on the plate characteristic where the curvature is high. A grid bias that is equal to approximately one-half the cutoff value is generally used.

Terms of higher order than the second, neglected in the foregoing analysis, give rise to components of plate current of other frequencies. The third-order term, for instance, is associated with the generation of frequencies equal to three times the carrier frequency, three times the modulation frequency, and twice either the carrier or the modulation frequency plus or minus the other frequency. If the modulation excitation voltage contains more than one frequency, the plate current will also contain intermodulation components having frequencies involving the sums and differences of the various modulation frequencies, the carrier frequency, and their multiples. The objectionable intermodulation components are those whose frequencies lie within the side bands and are therefore not suppressed by the tuned plate circuit. These frequencies give rise to distortion components that are superimposed upon the modulation frequencies when the modulation frequencies are again derived from the modulated wave by detection. Their presence is also evidenced by departure of the envelope of the modulated wave from the form of the impressed modulation excitation voltage. They are associated with the third- and higher-order terms of the plate-current series, and the ratio of their amplitudes to those of the desired side-frequency components is proportional to the $(n-2)$ power of the amplitudes of the modulation-excitation components that give rise to them, where n is the order of the term in the series with which they are associated. Distortion therefore increases with the amplitude of the modulation excitation, and hence also with the modulation factor, and may become objectionable at high values of modulation factor.

9-5. Balanced Modulator.—Either the carrier or the modulation frequency, as well as many of the intermodulation frequencies, may be prevented from appearing in the output by the use of the *balanced modulator* shown in Fig. 9-5. One excitation voltage e_1 is applied to the circuit in such a manner that it causes both grids to swing in the same direction, whereas the other voltage e_2 is applied so that it causes the grids to swing in opposite directions. If the grid excitation voltage of tube T is $e_1 + e_2$, then that of tube T' is $e_1 - e_2$. Under the assumption that the two tubes are identical and the circuit symmetrical, the plate currents of the two tubes are

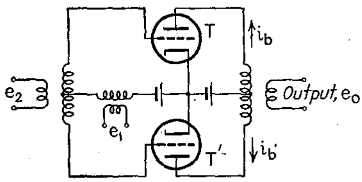


FIG. 9-5.—Balanced modulator.

$$i_p = a_1(e_1 + e_2) + a_2(e_1 + e_2)^2 + a_3(e_1 + e_2)^3 + \cdots \quad (9\text{-}11)$$
$$i_p' = a_1(e_1 - e_2) + a_2(e_1 - e_2)^2 + a_3(e_1 - e_2)^3 + \cdots \quad (9\text{-}12)$$

Because the two plate currents flow through the two halves of the primary of the output transformer in opposite directions, the voltage across the secondary is proportional to the difference between i_p and i_p'.

$$e_o = A(i_p - i_p') = 2A[a_1e_2 + 2a_2e_1e_2 + a_3(3e_1^2e_2 + e_2^3) + a_4(4e_1^3e_2 + 4e_1e_2^3) + a_5(5e_1^4e_2 + 10e_1^2e_2^3 + e_2^5) + \cdots] \quad (9\text{-}13)$$

where A is a constant of proportionality. Since a_3 and a_4 are small in comparison with a_1 and a_2 and since the ratios of the third- and higher-order terms to the first and second decrease with decrease of modulation factor, the third- and higher-order terms may be neglected at small modulation factor. Equation (9-13) then simplifies to

$$e_o = 2A(a_1e_2 + 2a_2e_2e_1) \quad (9\text{-}14)$$

If $e_1 = E_1 \sin \omega_m t$ and $e_2 = E_2 \sin \omega_k t$, Eq. (9-14) becomes, for resistance load,

$$e_o = 2Aa_1E_2\left(1 + 2\frac{a_2}{a_1} E_1 \sin \omega_m t\right) \sin \omega_k t \qquad (9\text{-}15)$$

which is of the form of Eq. (9-3). For certain applications it is advantageous to use the balanced modulator in such a manner that the carrier is eliminated from the output. This can be done by making e_1 the carrier excitation and e_2 the modulation excitation. The output then contains the modulation frequency and the side frequencies. At large values of modulation factor the second- and higher-order terms in Eq. (9-13) may not be negligible, indicating the production of undesired intermodulation frequencies. These are fewer in number than when a single tube is used.

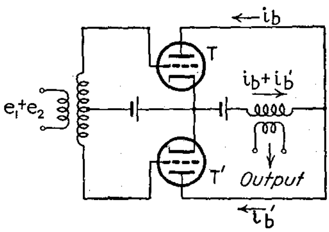


FIG. 9-6.—Balanced modulator.

Figure 9-6 shows another form of the balanced modulator circuit whose action is similar to that of the circuit of Fig. 9-5. An analysis like that given for the first circuit shows that the output voltage is

$$e_o = 2A[a_2(e_1 + e_2)^2 + a_4(e_1 + e_2)^4 + \cdots] \qquad (9\text{-}16)$$

If only the first term is considered and $e_1 = E_1 \sin \omega_m t$ and $e_2 = E_2 \sin \omega_k t$ are substituted in Eq. (9-16), Eq. (9-16) becomes

$$\begin{aligned} e_o = 2Aa_2[E_1E_2 \cos (\omega_k - \omega_m)t &- E_1E_2 \cos (\omega_k + \omega_m)t \\ &+ \tfrac{1}{2}E_1^2 + \tfrac{1}{2}E_2^2 - \tfrac{1}{2}E_1^2 \cos 2\omega_m t - \tfrac{1}{2}E_2^2 \cos 2\omega_k t] \end{aligned} \qquad (9\text{-}17)$$

The output of this circuit therefore contains neither the carrier nor the signal frequency but does contain the side frequencies.

Balanced modulator circuits have numerous applications and will be referred to in later chapters.

9-6. Amplitude Modulation by Complete Rectification.—Because the degree of modulation attainable in square-law modulation without excessive distortion is small and because the power efficiency is low, modulation by means of vacuum tubes is now accomplished almost entirely by making use of the unilateral conductivity of electron tubes, rather than of curvature of tube characteristics. Figure 9-7 shows a simple diode circuit that can be used to produce a modulated voltage. The parallel resonant circuit is tuned to the carrier frequency.

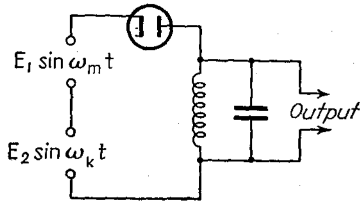


FIG. 9-7.—Diode modulator circuit.

The action of the circuit of Fig. 9-7 can be explained with the aid of the dynamic tube characteristic. Since the tube is a perfect conductor, *i.e.*, since current flows in one direction only, plate current flows only during portions of the cycle in which the plate is positive, and so the plate current consists, in general, of pulses. The manner in which the plate current varies with time when a carrier-frequency excitation voltage is applied to the circuit is shown in Fig. 9-8 for several values of fixed plate biasing voltage E_b. Because the current pulses occur at carrier-frequency intervals, the plate current contains a carrier-frequency component, the amplitude of which increases as the bias is changed in a positive direction between values equal to minus and plus the crest excitation voltage. If the bias E_b is replaced or supplemented by a modulation-frequency alternating voltage, the amplitude of the carrier component of plate current varies at modulation frequency. This is

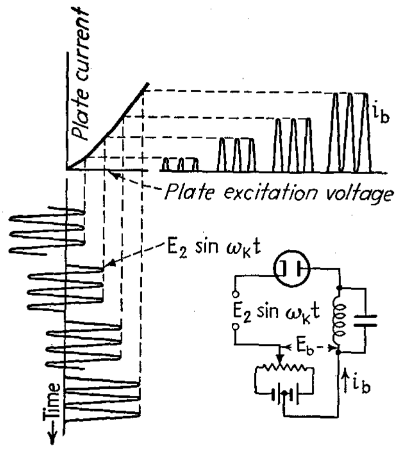


FIG. 9-8.—Variation of amplitude of alternating plate current with bias in a diode modulator circuit having carrier excitation only.

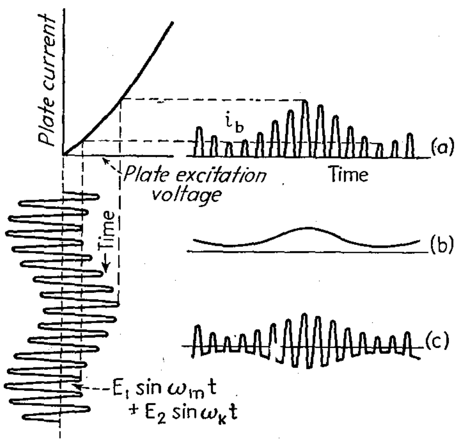


FIG. 9-9.—Amplitude modulation by complete rectification. Waves (*b*) and (*c*) are two components into which the wave (*a*) of plate current may be resolved.

accomplished by impressing the modulation excitation in series with the carrier excitation and optional bias, as in Fig. 9-7. The forms of the total excitation voltage and of the resulting plate current when the fixed bias is zero are shown in Fig. 9-9. The wave *a* of plate current may be separated into the components shown by curves b and c of Fig. 9-9. Curve b may be further resolved into a steady component and components whose frequencies are the modulation frequency and its harmonics. Curve c may be resolved into a varying-amplitude carrier-frequency component and carrier-harmonic components. Since the resonant circuit is tuned to the carrier frequency, its impedance is high only at the carrier and side frequencies, and so only the modulated-carrier component of plate current produces appreciable output voltage.

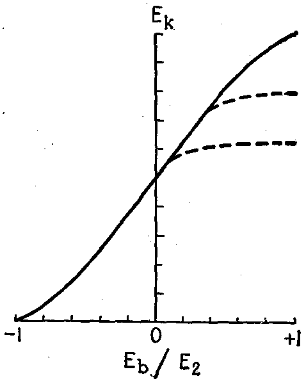


FIG. 9-10.—Ideal linear diode modulation characteristic and (dotted) typical experimentally determined characteristics.

The action of the diode modulator can be explained also with the aid of a *modulation characteristic*, which is a curve of amplitude of carrier-frequency output voltage as a function of direct voltage E_b at constant amplitude of carrier excitation E_2. A theoretical analysis indicates that if the tube characteristic were linear down to zero current and the load were a pure resistance, the fundamental component of output voltage E_k resulting from carrier excitation of amplitude E_2 would vary with E_b in the manner indicated by the solid curve of Fig. 9-10 (see Sec. A-2 of the Appendix). Experimental curves, obtained by means of the

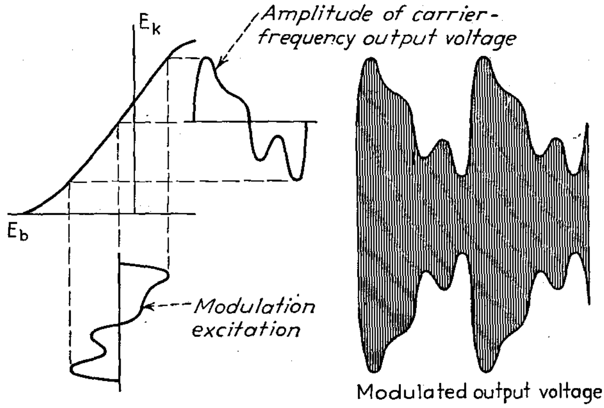


FIG. 9-11.—Essentially undistorted linear diode modulation.

circuit of Fig. 9-8, differ from this theoretical curve because of curvature of the diode characteristic and because the load is not a resistance. When the L/rC ratio of the tuned circuit is high, the curve flattens off at a positive value of E_b that is considerably less than E_2, as shown by the dotted curves of Fig. 9-10. In the circuit of Fig. 9-8, the filtering action of the resonant circuit suppresses all but the fundamental carrier frequency from the output voltage. Figure 9-11 shows the manner in which the amplitude of the carrier-frequency output voltage varies when E_b is varied periodically over an essentially linear range of the modulation characteristic in accordance with a wave of modulation excitation voltage. It can be seen that the wave of output voltage amplitude, which is the envelope of the modulated wave, has the same shape as the modulation excitation voltage. The modulated wave of Fig. 9-11 is undistorted and contains only the desired carrier and side frequencies. When the range in which E_b varies is not confined to the linear portion of the modulation characteristic, the wave form of the modulation envelope differs from that of the modulation excitation voltage and the output contains undesired intermodulation frequencies that will cause distortion when the modulation frequencies are derived from the modulated voltage wave

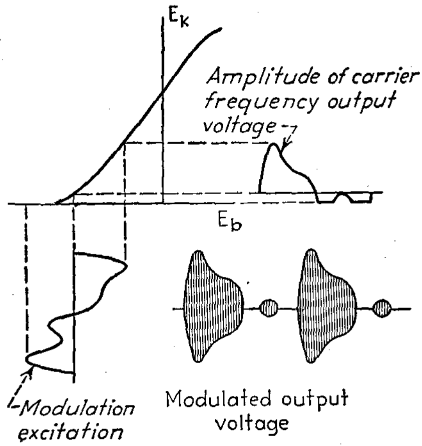


FIG. 9-12.—Modulation distortion resulting from high negative bias in a diode modulator.

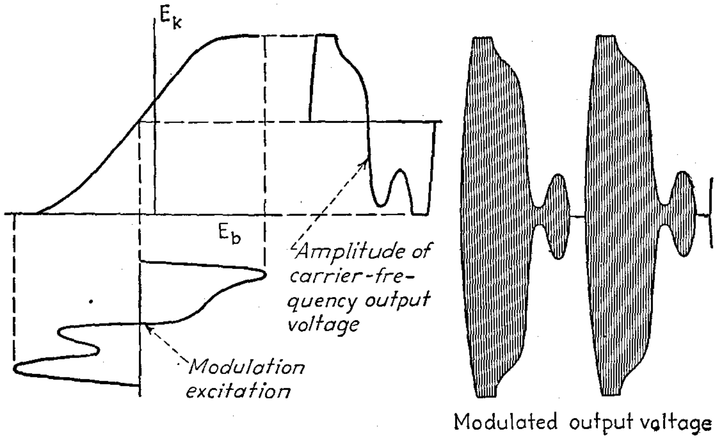


FIG. 9-13.—Modulation distortion resulting from excessive modulation excitation voltage.

by detection. Figure 9-12 shows the manner in which incorrect choice of biasing voltage E_{bo} results in distortion of the modulated wave, and Fig. 9-13 shows how the modulated wave is distorted by excessive amplitude of modulation excitation voltage.

Because the amplitude of the carrier-frequency output voltage can be made to vary essentially linearly with instantaneous modulation excitation voltage when the modulator circuit is properly designed and operated, modulation by complete rectification is called *linear modulation.* The close approach to linearity that can be achieved in practical circuits, and the resulting freedom from excessive distortion make linear modulation far superior to square-law modulation.

9-7. Linear Plate and Grid Modulation.—Although the simple diode modulator circuit of Fig. 9-7 is useful in studying and demonstrating the principle of linear modulation (see Prob. 9-10), it is not used in practice, since triode or multigrid-tube circuits require lower excitation voltages and have higher efficiency. Generally the carrier excitation is applied to the control grid of a triode or tetrode and the modulation

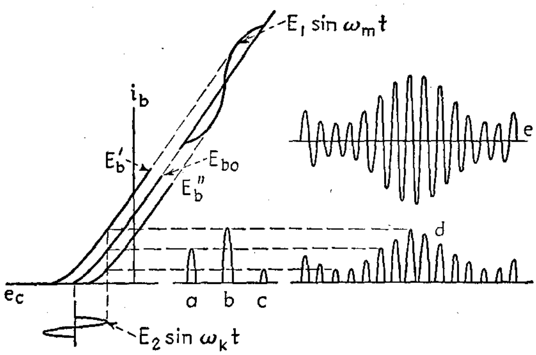


FIG. 9-14.—Graphical explanation of linear plate modulation.

excitation to the plate. Complete rectification takes place because of plate-current cutoff. This type of modulation is called *linear plate modulation* (*plate-voltage modulation*). The circuit of the linear plate modulator is similar to that of Fig. 9-2, but the modulation input transformer is in the plate circuit, instead of in the grid circuit. A graphical explanation of the operation of the circuit may be given with the aid of Fig. 9-14. E_{bo} and E_c are the operating plate voltage and grid bias. Curve *a* shows the form of the plate-current pulses produced when the carrier excitation voltage $E_2 \sin \omega_k t$ is impressed upon the grid circuit. Curves *b* and *c* show the plate-current pulses for the higher direct plate voltage E_b' and for the lower direct plate voltage E_b'', respectively. When the plate voltage is varied periodically between these limits by means of the modulation excitation voltage $E_1 \sin \omega_m t$, the plate current varies in the manner indicated by curve *d*. The tuned plate circuit converts the carrier and side-frequency components of plate current into output voltage, giving the modulated output voltage wave *e*. The modulation characteristic is similar in shape to that of the diode linear modulator, being nearly linear over a wide range of direct voltage. When the grid is biased somewhat beyond cutoff, the operation is essentially linear, and so very little distortion results even at 100 per cent modulation. Less power need be furnished by the source of carrier excitation than in a diode circuit, and the amplification produced by the tube reduces the required carrier excitation amplitude below that required with a diode.

In linear plate modulation the source of modulation voltage must furnish a large part of the power supplied to the plate circuit. The power that must be supplied by the source of modulation voltage is

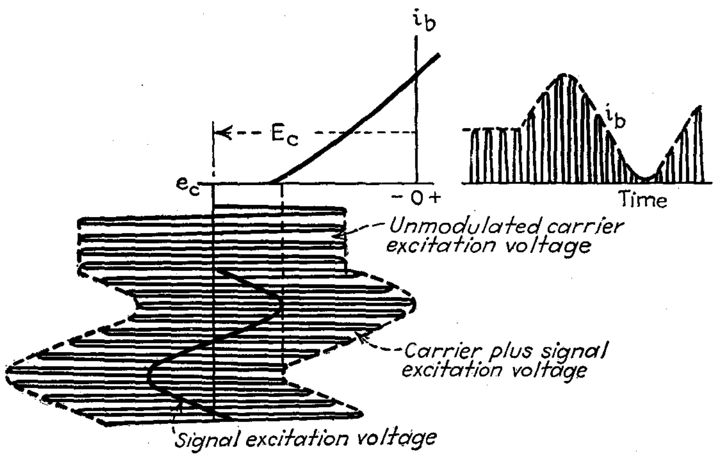


FIG. 9-15.—Linear grid modulation. A parallel resonant circuit converts the carrier and side-frequency components of plate current into output voltage.

less in *linear grid modulation* (*grid-bias modulation*) in which the modulation voltage, as well as the carrier voltage, is applied to the grid circuit.[1] Linear grid modulation has the additional advantage of requiring much lower modulation excitation voltage, but gives somewhat higher distortion and is less efficient. Linear grid modulation is illustrated graphically in Fig. 9-15. As in other modulation circuits, the tuned plate load converts the carrier and side-frequency components of plate current into output voltage.

In properly designed circuits distortion as low as 2 per cent may be attained at 100 per cent modulation in linear plate modulation and as low as 5 per cent in linear grid modulation. The distortion can be reduced below these values by the use of inverse feedback.

Several modifications of the basic linear modulation circuits have been developed in order to increase the efficiency or to reduce the required modulation-frequency exciting voltage. The modulation excitation

[1] KISHPAUGH, A. W., and CORAM, R. E., *Proc. I.R.E.*, **21**, 212 (1933).

may, for instance, be applied in series with the cathode of a triode, it may be applied simultaneously to the plate and screen of a tetrode or of a beam pentode, or it may be applied to the suppressor of a pentode.[1]

Equation (9-4) shows that the carrier-frequency component of an amplitude-modulated wave is independent of the degree of modulation. The continuous radiation of energy associated with the carrier component of the wave limits the efficiency that may be attained in a transmitter. Several systems of amplitude modulation that give higher efficiency have been proposed.[2] As will be shown in a later section, higher efficiency is also attained in frequency modulation.

9-8. Graphical Modulation Analysis.—The theoretical performance of modulators may be determined graphically by the following method:[3]

1. By applying Eq. (4-30) to the plate or transfer diagram of the tube or by a more accurate graphical method, determine the fundamental carrier-frequency component of plate current or output voltage for the given carrier excitation voltage and various values of direct voltage applied to the electrode to which the modulation excitation is to be applied. Using these values, plot the modulation characteristic (fundamental component of plate current or putput voltage *vs.* direct voltage).

2. Treating this modulation characteristic just as an ordinary dynamic transfer characteristic, determine the fundamental and harmonic components of the modulation envelope, or the distortion factor, by application of Eqs. (4-29), (4-30), or (4-31) or Table 4-I. If preferred, the modulation envelope may be derived by graphical construction as in Figs. 9-11 to 9-13. Thus, if the curve of Fig. 9-16 represents the variation of fundamental carrier output voltage with direct grid voltage for the given carrier excitation, then F_{max} and F_{min} are the maximum and minimum amplitudes of the modulated output voltage for the grid bias and wave of signal voltage shown, and F_o is the amplitude of the unmodulated carrier. Substituting F_{max}, F_o, and F_{min} in Eq. (4-37) (instead of I_{max}, I_{bt}, and I_{min}) will give the percentage second harmonic of the signal in the modulated wave. Curve *b*, constructed by projecting from curve *a* to the modulation characteristic at various instants in the cycle, represents the upper half of the envelope of the modulated wave.

It is evident that, if the three-point equations (4-29) and (4-37) are used in the second step, it is unnecessary to plot the modulation characteristic. Only the three values F_{max}, F_o, and F_{min} need be found. The method used to derive the modulation characteristic should be of higher order of accuracy than that used to determine the distortion from the modula-

[1] GREEN, C. B., *Bell Lab. Record*, **17**, 41 (1938).

[2] CHIREIX, H., *Proc. I.R.E.*, **23**, 1370 (1935); DOHERTY, W. H., *Proc. I.R.E.*, **24**, 1163 (1936); DOME, R. B., *Proc. I.R.E.*, **26**, 963 (1938); GAUDERNACK, L. F., *Proc. I.R.E.*, **26**, 983 (1938); RODER, H., *Proc. I.R.E.*, **27**, 386 (1939); VANCE, A. W., *Proc. I.R.E.*, **27**, 506 (1939).

[3] PETERSON, E., and LLEWELLYN, F. B., *Proc. I.R.E.*, **18**, 38 (1930).

tion characteristic. In using the three-point equation to find the percentage second harmonic of the signal, therefore, F_{max}, F_{min}, and F_o should be evaluated by the five-point equations (4-30) or (4-31) or the seven-point equations of Table 4-I.

A similar three-point method may be used in determining the distortion of a modulated signal by an amplifier.[1] In this case the modulation characteristic of the amplifier is a curve of fundamental carrier output *vs.* amplitude of unmodulated carrier excitation at the given operating voltages.

9-9. Secrecy Systems (Speech Scramblers).[2]—The use of radio as a means of transmitting private conversion has necessitated the

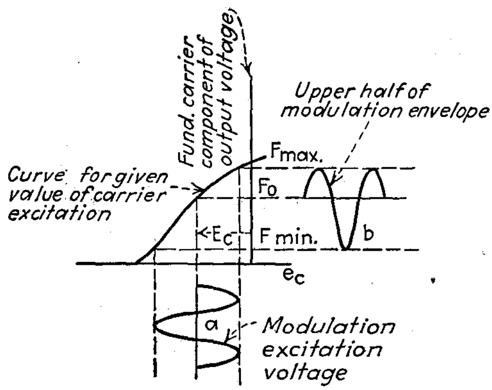


FIG. 9-16.—Use of the modulation characteristic in the graphical determination of modulation distortion.

development of methods of making transmitted speech unintelligible when received with an ordinary radio receiver. The balanced modulator affords one method of accomplishing this result. In Fig. 9-5 let e_1 be the voice input and e_2 be a 3100-cycle carrier voltage. The output will contain the carrier and the side frequencies, as well as higher-order intermodulation frequencies. The carrier, the upper side frequencies, and most of the higher-order frequencies can be removed from the output by a low-pass filter that cuts off slightly below 3100 cps. The lower side band represents a band of frequencies that is inverted with respect to the input signal band. The application of voice frequencies ranging

[1] FERRIS, W. R., *Proc. I.R.E.*, **23**, 510 (1935).

[2] FAYARD, G., *Bull. soc. franç. élec.*, **9**, 1146 (1929); **10**, 328 (1930); KUJIRAI, T., and SAKAMOTO, T., *Radio Research (Japan) Rept.*, **2**, 175 (1932) (in English); MATSUYAKI, T., *Radio Research (Japan) Rept.*, **2**, 187 (1932); NIWA, Y. and HAYASHI, T., *Radio Research (Japan) Rept.*, **2**, 195 (1932); CHIBA, S., *Radio Research (Japan) Rept.*, **3**, 267 (1933); VILLEM, R., *L'Onde élec.*, **11**, 427 (1932); GILL, A. J., *Post Office Elec. Eng. J.*, **26**, 224 (1933); *Electrician*, **110**, 801 (1933); ROBERTS, W. W., *Electronics*, October, 1943, p. 108.

from 100 up to 3000 cps produces output frequencies ranging from 3000 cps down to 100 cps, the high frequencies being converted into low and the low into high. Thus, a 100-cycle input voltage gives a 3000-cycle output voltage, whereas a 3000-cycle input voltage gives a 100-cycle output voltage. A 1550-cycle signal passes through without change of frequency. This frequency transposition converts normal speech into unintelligible gibberish, which can be reconverted into intelligible sounds by a similar frequency transposition at the receiving end. To correct for differences between the two tubes of the balanced modulator, resistance and capacitance balances must be employed in the modulator plate circuit, as shown in Fig. 9-17. For satisfactory results it is usually necessary to provide adjustment of the amplitude of the 3100-cycle carrier input to the modulator.

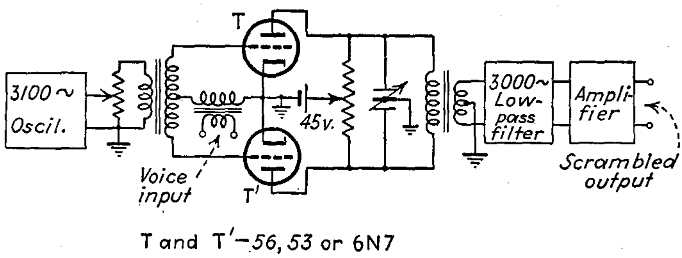


FIG. 9-17.—Simple circuit for frequency inversion (speech scrambler).

By using a higher value of carrier frequency, the signal band can be simultaneously shifted and inverted. The difficulty of "unscrambling" can be increased by breaking up the voice input into a number of small bands and inverting and shifting each of these. More complicated circuits are used in practical secrecy systems for radio telephony.

9-10. Detection of A-m Waves.—The process whereby the modulation frequency is derived from a modulated wave is a special case of a more general process called *detection*. Detection may be defined broadly as the process whereby the application of a voltage of a given frequency or frequencies to a circuit containing an asymmetrically conducting device produces currents of certain desired frequencies or desired changes in average current. The detection of a modulated wave to produce the modulation frequency is often termed *demodulation* in the United States.[1] Other examples of detection are the production of a difference or *beat* frequency by the simultaneous application of two frequencies to a detector, and the production of a steady voltage or current by the application of an alternating voltage. According to the broad definition just given, modulation may, in fact, be considered to be a special form of

[1] "Standards on Radio Receivers," p. 5, Institute of Radio Engineers, New York, 1938.

detection. As already pointed out, the processes involved in amplitude modulation are similar to those involved in detection of amplitude-modulated waves. Detection of amplitude-modulated voltages, like their production, may be accomplished either as the result of curvature of vacuum-tube characteristics or as the result of complete rectification.

9-11. Detection by Curvature of Current-voltage Characteristic.—If a wave of alternating voltage is applied to the grid or plate circuit

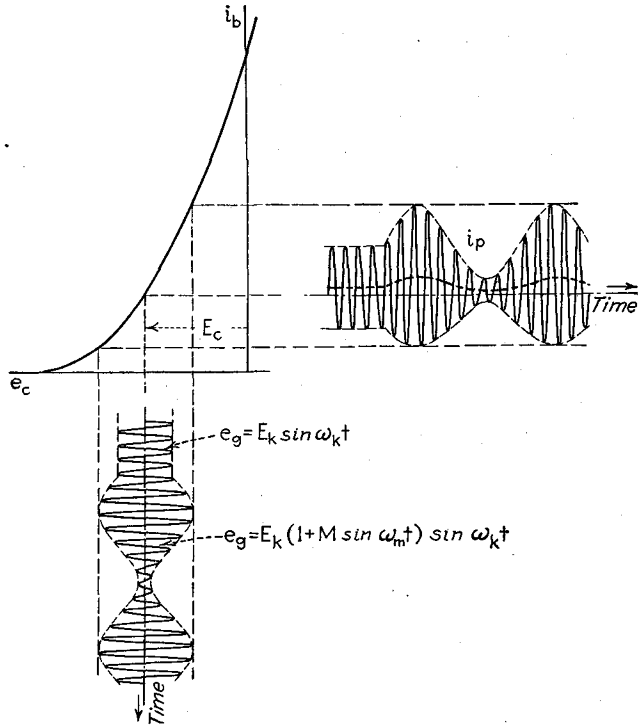


FIG. 9-18.—Detection of an a-m wave by curvature of a triode transfer characteristic.

of a diode or triode, together with proper operating voltages to bring the operating point to a point of the characteristic where the curvature is high, the wave of plate current is asymmetrical, as shown by the transfer diagrams of Figs. 3-19, 3-21, and 9-18. If the input wave is sinusoidal, the plate current contains a steady component, the applied frequency, and its harmonics. If the excitation voltage contains two or more frequencies, the plate current contains components having these frequencies, their harmonics, and the sums and differences of the impressed frequencies and their integral multiples, and a steady component. This process of frequency generation, which is identical with that of the production of a modulated wave by curvature of the characteristic, has already been discussed in Secs. 3-23 and 9-4. The production of a difference-frequency voltage by the simultaneous application of two voltages of different frequencies to a detector is of value in heterodyne radio receivers (see Sec. 9-25) and heterodyne (beat-frequency) oscillators (see Sec. 10-52).

Figure 9-18 shows the form of the asymmetrical wave of plate current for unmodulated carrier excitation and for excitation consisting of an amplitude-modulated voltage. The wave may be analyzed into a number of components, including that indicated by the dotted line, which shows the average current in each carrier-frequency cycle. For modulated excitation the dotted curve is seen to contain a steady component and an alternating component having the same frequency as the envelope of the impressed modulated voltage, *i.e.*, modulation frequency. The plate load impedance serves to convert the component of plate current of desired frequency, the modulation frequency, into output voltage and to suppress the carrier-frequency components. This is accomplished by using for the load the parallel combination of a high resistance and a condenser whose reactance is negligible at carrier frequency but high at modulation frequency.

9-12. Analysis of Square-law Detection.—The process of detection by means of curvature of vacuum-tube characteristics can be analyzed rigorously by use of the series expansion for alternating plate current. At small values of excitation voltage the plate current may be expressed approximately by the first two terms of the series.

$$i_p = a_1 e + a_2 e^2 \tag{9-18}$$

If the exciting voltage is amplitude-modulated,

$$e = E_k(1 + M \sin \omega_m t) \sin \omega_k t = E_k \sin \omega_k t + \tfrac{1}{2} M E_k[\cos (\omega_k + \omega_m)t - \cos (\omega_k - \omega_m)t] \tag{9-19}$$

If Eq. (9-19) is substituted in Eq. (9-18), the first term of Eq. (9-18) gives rise to carrier and side-frequency components. The second term gives rise to three steady components, three components of frequency double those of the carrier and side frequencies, and six components of frequencies equal to the sums and differences of the carrier and side frequencies, taken in pairs. If the impedance of the plate load is negligible at carrier and higher frequencies but high throughout the range of modulation frequency, the voltage developed across the load will contain only those frequencies which lie in the modulation-frequency range. The upper side frequency minus the carrier frequency and the carrier frequency minus the lower side frequency are both equal to the modulation frequency; the difference between the two side frequencies is equal to twice the modulation frequency. The voltage across the load is

$$e_o = \tfrac{1}{2}(a_2)_{(k+m)-k}(z_b)_m \underbrace{E_k{}^2M \cos \omega_m t}_{(\omega_k+\omega_m)-\omega_k} + \tfrac{1}{2}(a_2)_{k-(k-m)}(z_b)_m \underbrace{E_k{}^2M \cos \omega_m t}_{\omega_k-(\omega_k-\omega_m)}$$
$$+ \tfrac{1}{4}(a_2)_{(k+m)-(k-m)}(z_b)_{2m} \underbrace{E_k{}^2M^2 \cos 2\omega_m t}_{(\omega_k+\omega_m)-(\omega_k-\omega_m)} \qquad (9\text{-}20)$$

The first two terms of Eq. (9-20) have the desired modulation frequency. Because these components are derived from the second term of the plate-current series and because their amplitudes are proportional to the square of the carrier amplitude of the modulated input voltage, this type of detection is known as *square-law* detection. The third term of Eq. (9-20) represents the second harmonic of the modulation frequency. Since the input to the detector was assumed to be sinusoidally modulated, this term indicates distortion in the process of detection.

The impedances z_b and the coefficients a_2 in Eq. (9-20) are not the same in all three terms, since the frequencies are different. The values of a_2 may be determined as explained in Sec. 3-24. If μ is assumed to be constant, Eq. (3-57) gives the following values of a_2 for the three terms of Eq. (9-20):

$$\left.\begin{aligned}(a_2)_{(k+m)-k} &= \frac{-\tfrac{1}{2}\mu^2 r_p(\partial r_p/\partial e_b)}{[r_p + (z_b)_{k+m}][r_p + (\bar{z}_b)_k][r_p + (z_b)_m]}\\ (a_2)_{k-(k-m)} &= \frac{-\tfrac{1}{2}\mu^2 r_p(\partial r_p/\partial e_b)}{[r_p + (z_b)_k][r_p + (\bar{z}_b)_{k-m}][r_p + (z_b)_m]}\\ (a_2)_{(k+m)-(k-m)} &= \frac{-\tfrac{1}{2}\mu^2 r_p(\partial r_p/\partial e_b)}{[r_p + (z_b)_{k+m}][r_p + (\bar{z}_b)_{k-m}][r_p + (z_b)_{2m}]}\end{aligned}\right\} \qquad (9\text{-}21)$$

Because the load impedance is assumed to be negligible at carrier and side frequencies, $(z_b)_{k+m}$, $(z_b)_{k-m}$, and $(z_b)_k$ may be neglected in comparison with r_p in Eqs. (9-21). Equations (9-21) then reduce to

$$\left.\begin{aligned}(a_2)_{(k+m)-k} = a_{2,k-(k-m)} &= -\tfrac{1}{2}\frac{\mu^2(\partial r_p/\partial e_b)}{r_p[r_p + (z_b)_m]} = \tfrac{1}{2}\frac{\mu^2 r_p(\partial^2 i_b/\partial e_b{}^2)}{r_p + (z_b)_m}\\ (a_2)_{(k+m)-(k-m)} &= \tfrac{1}{2}\frac{\mu^2 r_p(\partial^2 i_b/\partial e_b{}^2)}{r_p + (z_b)_{2m}}\end{aligned}\right\} \qquad (9\text{-}22)$$

The output voltage is

$$e_o = \tfrac{1}{2}E_k{}^2M\mu^2 r_p \frac{\partial^2 i_b}{\partial e_b{}^2}\left[\frac{(z_b)_m}{r_p + (z_b)_m}\cos \omega_m t + \tfrac{1}{4}\frac{m(z_b)_{2m}}{r_p + (z_b)_{2m}}\cos 2\omega_m t\right] \qquad (9\text{-}23)$$

If the load impedance at twice the modulation frequency does not differ greatly from that at modulation frequency, which is normally true, the ratio of the second harmonic to the fundamental modulation-frequency output of the detector approximates $\frac{1}{4}M$, which has a value of 25 per cent for 100 per cent modulation. A more complete analysis shows that, if the carrier is modulated simultaneously by a number of modulation frequencies, the detector output contains also components whose frequencies are the sums and differences of the various modulation frequencies and whose amplitudes are proportional to the products of the corresponding modulation factors. It is these sum and difference frequencies, which are generally inharmonically related to the modulation frequencies, rather than the harmonics, that cause objectionable distortion in the reproduction of voice or music. Because the sum- and difference-frequency amplitudes are proportional to the products of the modulation factors of the modulation-frequency components that give rise to them, the largest sum- or difference-frequency amplitude cannot exceed the largest second harmonic. Therefore the factor $\frac{1}{4}M$ is a satisfactory distortion index in the detection of a modulated wave by a square-law detector. To prevent excessive distortion it is necessary to keep the degree of modulation low. Unfortunately, however, the amplitude of the fundamental output is also proportional to the modulation factor. The impossibility of maintaining high modulation without introducing excessive distortion represents one disadvantage of square-law detection. It is of interest to note that, since the third term of Eq. (9-20) arises from the difference between the upper and lower side frequencies, this term is absent when one side frequency is removed from the carrier before detection. Consideration of higher-order terms of the plate-current series shows that the detector output also contains other distortion terms, which may become of importance at high excitation voltage, as well as other modulation-frequency components. The above analysis is therefore accurate only when the excitation voltage is small.

Because the amplitude of the detector output is proportional to E_k^2, small changes of amplitude of the detector input, such as might result from fading of a received radio wave, result in relatively large changes in output voltage. This is another disadvantage of the square-law detector.

If additional carrier-frequency voltage, of amplitude E_k', in phase with the carrier component of the modulated input voltage, is supplied to the detector, the coefficient of the first term of Eq. (9-19) becomes $E_k + E_k'$, and E_k^2 in the first term of Eq. (9-23) is replaced by

$$E_k(E_k + E_k').$$

This shows that the amplitude of the detector output can be greatly increased by the superposition of additional carrier voltage upon the modulated voltage applied to a square-law detector. Because of the difficulty of synchronizing the local oscillator to the received carrier, this is not a practical method of increasing the amplitude of received radio signals. The difficulties of synchronization may, however, be eliminated by suppressing the carrier from the transmitted wave and applying it to the detector. If the carrier is removed from the modulated wave at the transmitter, the amplitude of the detector output is proportional to E_kE_k'. Another advantage of carrier suppression at the transmitter lies in the fact that the carrier component of the modulated wave requires at least twice as much power as the side-frequency components. This may be shown by computing the power supplied to a resistance by a generator whose voltage is of the form of Eq. (9-19).

By substituting Eq. (9-19) in the second term of Eq. (9-18) and applying Eqs. (3-57), the student may derive the following expression for the change in the average plate current of a square-law detector when it is excited by a modulated voltage:

$$\Delta I_{ba} = \tfrac{1}{4}E_k^2 \frac{\mu^2 r_p}{r_p + r_b}(M^2 + 2)\frac{\partial^2 i_b}{\partial e_b^2} \tag{9-24}$$

The variation of average plate current with carrier amplitude makes possible automatic volume control of radio receivers (see Sec. 9-26).

Since e in Eq. (9-18) is equal to $e_g + v_p/\mu$, it follows that the foregoing analysis holds whether the modulated voltage is applied to the grid or to the plate of a triode or to the plate of a diode. The analysis is applicable also to tubes with more than one grid if excitation voltage is applied to only one grid. For a given input amplitude, however, the output will be μ times as great if the modulated voltage is applied to the grid as if it were applied to the plate. For this reason it is distinctly advantageous to apply the modulated voltage to the grid circuit. This has the additional advantage that, if the grid is not allowed to become positive, it does not draw current.

9-12A. Frequency Converters.—It follows from the discussion of Sec. 3-23 that the simultaneous excitation of a square-law detector by two different frequencies results in the generation, among other components, of those having frequencies equal to the sum and the difference of the impressed frequencies. If the plate circuit of the detector contains or is followed by a filter tuned to either the sum or the difference frequency, the output contains only this frequency. The detector therefore serves as a frequency converter. If one of the excitation voltages is modulated, and the other is that of an oscillator of fixed frequency, the detector output contains carrier and side frequencies equal to the oscillator frequency plus or minus the original carrier and side frequencies. The effect of the detector in conjunction with the "local" oscillator is therefore to convert all input frequencies to new values that are higher or lower than the original values by the frequency of the oscillator. The filter must, of course, pass a band of frequencies of width equal to the original band width.

In practical frequency converters the two signals are usually impressed upon two grids of a multigrid tube. By the use of special five-grid tubes, called *pentagrid converters*, it is possible to make a single tube serve both as the oscillator and as the detector. The first and second grids serve as the grid and plate, respectively, of a triode oscillator. The input signal is impressed upon the fourth grid. The third and fifth grids, which are tied together within the tube, serve both as a shield for the fourth grid and as accelerating electrodes similar to the screen of a tetrode. The output is taken from the plate circuit. Frequency converters are used in superheterodyne radio receivers, which are discussed further in Sec. 9-25. When used in this manner they are called *first detectors*.

9-13. Square-law Detector Circuits.—Figure 9-19 shows the circuit of a single-tube square-law detector. The function of the condenser C' is to prevent the direct voltage and changes of direct voltage resulting from detection from appearing in the output. R_{cc} provides grid bias. A more nearly parabolic transfer characteristic may be attained by the use of a balanced push-pull circuit. Such a circuit, shown in Fig. 15-2, is occasionally of value when a true square-law detector is required for measuring purposes (see Sec. 15-2). Because of its relatively low output and its high second-harmonic distortion (10 per cent when $M = 40$ per cent, and 25 per cent when $M = 100$ per cent), the square-law detector has been largely replaced by types making use of complete rectification.

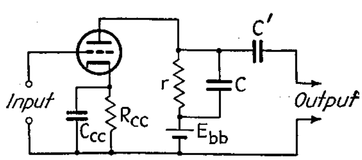


FIG. 9-19.—Square-law plate detector.

9-14. Detection by Complete Rectification.—When a wave of modulated voltage is applied to a diode in series with a pure resistance, plate current flows only during the positive half cycles of impressed voltage, as shown in Fig. 9-20. The current wave consists of pulses and is very nearly a replica of the positive portions of the applied voltage wave, as shown in Fig. 9-20. It may be resolved into the components represented by curves *a* and *b*. Curve *a*, which resembles the impressed modulated wave but is distorted, may be broken up further into components having frequencies equal to the carrier and side frequencies and their harmonics. Curve *b*, which is the average current in each carrier-frequency cycle, may be further separated into a steady component and modulation-frequency and modulation-harmonic components.

Normally the carrier frequency is so much higher than the modulation frequency that each carrier-frequency cycle of the impressed wave may be assumed to be sinusoidal. By the use of high load resistance (of the order of 500,000 ohms), the curvature of the dynamic characteristic may be made very small. Each pulse of plate current then approximates a half sine wave of amplitude $\frac{(1 + M \sin \omega_m t)E_k}{r_p + r}$, in which r is the resistance of the load, and r_p is the a-c resistance of the tube during the time it conducts.[1] The average current, averaged over one carrier-frequency cycle, is

$$I_{bk} = \frac{(1 + M \sin \omega_m t)E_k}{\pi(r_p + r)} \tag{9-25}$$

and the average voltage across the resistance is

$$E_r = \frac{E_k r}{\pi(r_p + r)} + \frac{ME_k r \sin \omega_m t}{\pi(r_p + r)} \tag{9-26}$$

The first term of Eq. (9-26) represents the direct output voltage and the second term the modulation-frequency output voltage.

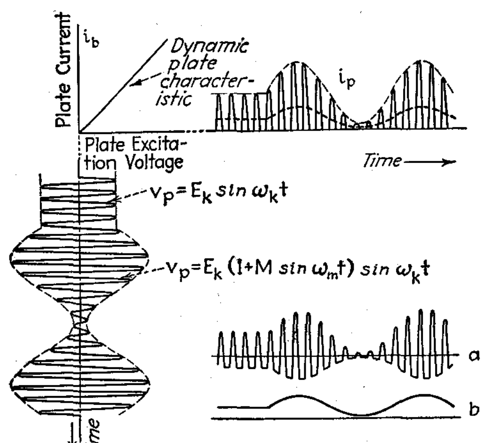


FIG. 9-20.—Graphical explanation of diode detection of a modulated wave. Waves *a* and *b* are two components into which the wave of plate current i_p can be resolved.

The operation of the diode detector is greatly improved by shunting the resistance with a condenser the reactance of which is low at carrier frequency but high at modulation frequency. The circuit is that of Fig. 9-21*a*. The condenser C increases the modulation-frequency output voltage and suppresses carrier and side-frequency voltages from the output. The action of the circuit of Fig. 9-21*a* differs from that of the circuit in which the condenser is omitted. If only the condenser

[1] Although r_p varies during the positive half cycle, r is so much larger than r_p that the variation in $r_p + r$ is small. This is another way of saying that the slope $1/(r_p + r)$ of the dynamic plate characteristic is nearly constant.

were used and the excitation consisted of an unmodulated carrier voltage, the condenser would be charged to a voltage equal to the crest impressed voltage, which is the carrier amplitude. An increase of amplitude of impressed voltage would cause an equal increase of direct condenser voltage but, because the tube is a perfect rectifier, the condenser could not discharge following a decrease of impressed voltage, and so the condenser voltage would not fall. The resistance allows the condenser to discharge when the amplitude of the impressed voltage is lowered.

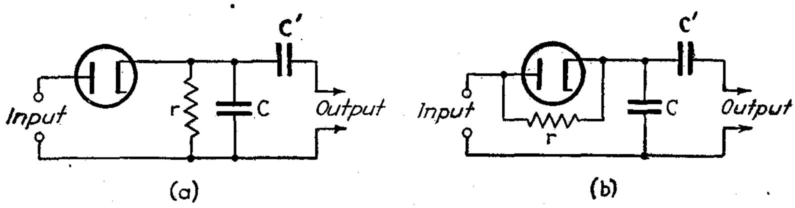


FIG. 9-21.—Diode detector circuits.

If the ratio of the resistance to the capacitance is made low enough, the condenser voltage can follow rapid reductions of amplitude. Then, when a modulated voltage is impressed upon the circuit, the condenser voltage is very nearly equal to the amplitude of the impressed voltage at all times, and so the condenser voltage is practically identical in wave form with the envelope of the impressed voltage. It therefore contains a steady component plus a modulation-frequency component.

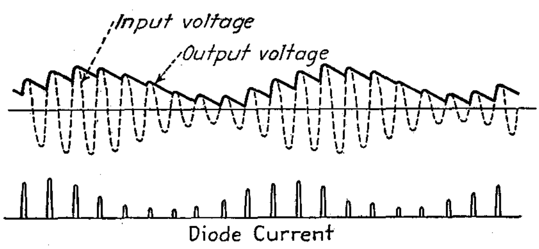


FIG. 9-22.—Wave forms of modulated input voltage, output voltage, and plate current in a diode detector circuit.

The wave forms of the output voltage and of the diode current are illustrated roughly in Fig. 9-22. (In order to show the behavior more clearly the ratio of carrier frequency to modulation frequency is made relatively low in Fig. 9-22 and the carrier-frequency ripple is made large.) It can be seen that current flows through the diode only during the relatively short portions of the positive half cycles preceding the positive crests of impressed voltage throughout which the impressed voltage exceeds the condenser voltage. During the remainder of each cycle the condenser discharges exponentially through the resistance.

The resistance r may be connected across the tube, as shown in Fig. 9-21b, instead of across the condenser. The condenser then discharges through the resistance and source in series, instead of through the resistance alone. Since the impedance of the source is low at modulation frequency, this does not affect the behavior of the circuit. Because of the low impedance of the source at modulation frequency, the modulation-frequency voltage across the tube in the circuits of Fig. 9-21 equals that across the condenser C, and so the output may be taken from across the tube instead of from across the condenser. Carrier-frequency voltage will, however, then be present in the output. The condenser C' in the circuits of Fig. 9-21 prevents the appearance of direct voltage in the output.

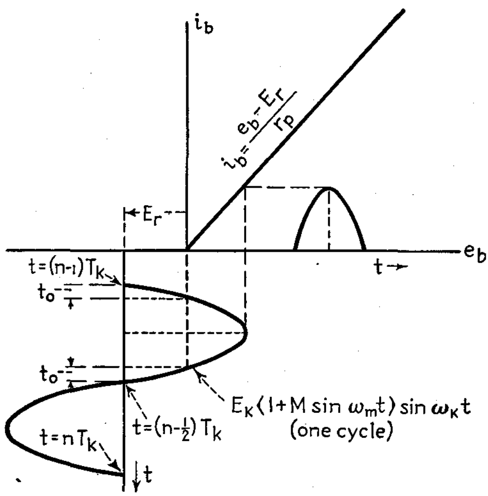


FIG. 9-23.—Diagram used in the derivation of Eqs. (9-35) and (9-37).

9-15. Analysis of Ideal Linear Diode Detection.—Detection by complete rectification of a diode can be analyzed under the assumption that the diode static characteristic is a straight line passing through the origin, as shown in Fig. 9-23. The carrier frequency is assumed to be so much higher than the modulation frequency that each carrier-frequency cycle of the impressed voltage can be considered to be sinusoidal. Only *one cycle* of the impressed modulated voltage $E_k(1 + M \sin \omega_m t) \sin \omega_k t$ is shown in Fig. 9-23. During the portion of the cycle in which plate current flows, some current flows through the resistance and the rest into the condenser. Since the condenser discharges through the resistance during the remainder of the cycle, as much charge flows through r in one cycle as is passed by the tube. The charge passed by the tube in one cycle, divided by the period of one carrier-frequency cycle, is therefore equal to the average current through the resistance. The product of this average current and the resistance is equal to the average voltage across the resistance and condenser. If the reactance of the condenser at carrier frequency is very small in comparison with the resistance, the voltage E_r across the condenser and resistance remains very nearly constant throughout the carrier-frequency cycle, and so may be assumed to be equal to the average voltage. The voltage E_r is of such polarity as to bias the tube negatively. Although E_r is practically constant throughout any carrier-frequency cycle, it varies with the amplitude of the modulated input voltage, *i.e.*, at modulation frequency, as will now be shown.

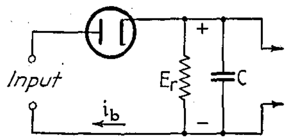


FIG. 9-24.—Diode detector circuit.

It can be seen from Fig. 9-24 that the instantaneous plate voltage is equal to the instantaneous impressed voltage minus E_r. Current flows when the instantaneous plate voltage exceeds zero, as shown in Fig. 9-23. Current starts flowing in the given carrier-frequency cycle at a time t_o after the beginning of the cycle such that

$$E_k(1 + M \sin \omega_m t) \sin \omega_k t_o - E_r = 0 \tag{9-27}$$

or

$$E_r = E_k(1 + M \sin \omega_m t) \sin \omega_k t_o \tag{9-28}$$

Although Eq. (9-28) expresses the output voltage in terms of the carrier amplitude and the modulation factor, it is necessary to express $\sin \omega_k t_o$ in terms of r and r_p before the equation can be used. This may be done as follows:

The instantaneous plate current is equal to the instantaneous plate voltage, divided by the plate resistance r_p, which has been assumed to be constant.

$$i_b = \frac{E_k(1 + M \sin \omega_m t) \sin \omega_k t - E_r}{r_p} \tag{9-29}$$

E_r may be eliminated from Eq. (9-29) by means of Eq. (9-28).

$$i_b = \frac{E_k(1 + M \sin \omega_m t)}{r_p} (\sin \omega_k t - \sin \omega_k t_o) \tag{9-30}$$

The charge that flows through the tube throughout the nth carrier-frequency cycle is

$$Q = \int_{(n-1)T_k + t_o}^{(n-\frac{1}{2})T_k - t_o} \frac{E_k(1 + M \sin \omega_m t)}{r_p} (\sin \omega_k t - \sin \omega_k t_o)\, dt \tag{9-31}$$

in which T_k is the period of the carrier-frequency cycle.

$$Q = \frac{E_k(1 + M \sin \omega_m t)}{\omega_k r_p} [2 \cos \omega_k t_o + (2\omega_k t_o - \pi) \sin \omega_k t_o] \tag{9-32}$$

The average current throughout the carrier-frequency cycle is

$$
\begin{aligned}
I_{bk} &= \frac{Q}{T_k} = \frac{Q\omega_k}{2\pi} \\
&= \frac{E_k(1 + M \sin \omega_m t)}{\pi r_p}\left[\cos \omega_k t_o + \left(\omega_k t_o - \frac{\pi}{2}\right)\sin \omega_k t_o\right] \qquad (9\text{-}33)
\end{aligned}
$$

But, since $E_r \cong I_{bk}r$, Eq. (9-27) may be written in the form

$$
\sin \omega_k t_o = \frac{I_{bk}r}{E_k(1 + M \sin \omega_m t)} \qquad (9\text{-}34)
$$

Substitution of Eq. (9-33) in Eq. (9-34) gives the following relation:

$$
\sin \omega_k t_o = \frac{r}{\pi r_p}\left[\cos \omega_k t_o + \left(\omega_k t_o - \frac{\pi}{2}\right)\sin \omega_k t_o\right] \qquad (9\text{-}35)
$$

Although Eq. (9-35) cannot readily be solved for $\sin \omega_k t_o$ as a function of r/r_p it is a simple matter to plot a curve of $\sin \omega_k t_o$ *vs.* r/r_p by substituting

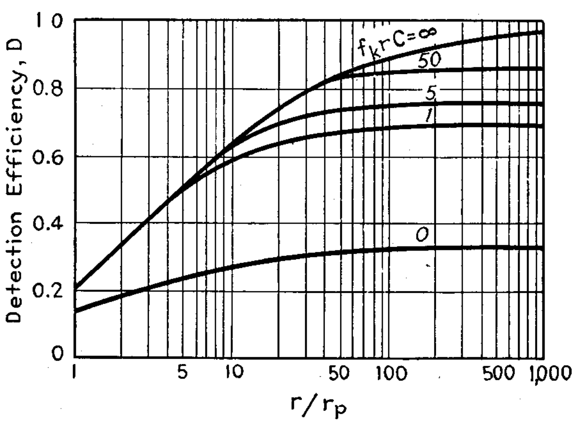


FIG. 9-25.—Variation of diode detection efficiency with ratio of load resistance to plate resistance for several values of $f_k rC$.

various values of $\omega_k t_o$ in Eq. (9-35) and finding the corresponding values of r/r_p. This gives the upper curve of Fig. 9-25. Such a curve shows that $\sin \omega_k t$ is a single-valued function of r/r_p. Since r and r_p are constant in a given circuit, it follows that $\sin \omega_k t_o$ is constant in Eq. (9-28). Hence the output of the ideal linear diode detector contains a steady component proportional to the carrier amplitude E_k, and a modulation-frequency component proportional to the carrier amplitude and to the modulation factor. Both components also increase with the ratio r/r_p, in the manner indicated by the upper curve of Fig. 9-25. In contrast with detection by curvature of tube characteristics, detection by complete rectification does not introduce harmonics of the modulation frequency if the tube characteristic is truly linear. As will be shown in Sec. 9-19, however, curvature of the diode characteristic results in some nonlinear distortion.

9-16. Detection Efficiency and Effective Input Resistance of the Ideal Linear Diode Detector.—The ratio of the amplitude of the modulation-frequency output voltage E_o to the product E_kM is called the *detection efficiency* and is represented by the symbol D. This may be stated symbolically by the relation

$$E_o = DE_kM \tag{9-36}$$

Inspection of Eqs. (9-26) and (9-28) discloses that the sum of the direct and modulation-frequency output voltages of the ideal diode detector is given by the relation

$$E_r = DE_k(1 + M \sin \omega_k t) \tag{9-37}$$

where D is equal to $\sin \omega_k t_o$ when the reactance of the detector condenser at carrier frequency is very small in comparison with r, and to $r_b/\pi(r_p + r)$ when the condenser is omitted. In Fig. 9-25 the detection efficiency is plotted against r/r_p at zero and infinite carrier-frequency condenser reactance and at several intermediate values. These curves show clearly that the detection efficiency and therefore the output voltage are increased by the condenser.

The carrier-frequency power taken from the source of detector input voltage may appreciably affect the operation of the voltage source. When the excitation voltage is taken from the output of a radio-frequency amplifier, for instance, the energy dissipation in the detector circuit reduces the sharpness of tuning of the amplifier. It is therefore often important to know how the power loss and the effective input resistance of the detector depend upon the detector plate and load resistances. The average power dissipated in the detector circuit in a given carrier-frequency cycle is equal to the integral, over the cycle, of the product of the instantaneous impressed voltage $E_k(1 + M \sin \omega_m t) \sin \omega_k t$ and the instantaneous current i_b, divided by the period of the carrier cycle. The current i_b, which is given by Eq. (9-30), flows during the time interval starting at the time t_o after the beginning of the positive half cycle and ending at the time t_o before the end of the positive half cycle, as shown in Fig. 9-23. The average power input during the nth carrier-frequency cycle is

$$P_i = \frac{1}{T_k}\,\frac{E_k{}^2(1 + M \sin \omega_m t)^2}{r_p} \int_{(n-1)T_k+t_0}^{(n-\frac{1}{2})T_k-t_0} (\sin \omega_k t - \sin \omega_k t_o) \sin \omega_k t \, dt \tag{9-38}$$

$$P_i = \frac{E_k^2(1 + M \sin \omega_m t)^2}{2\pi r_p}\left[\left(\frac{\pi}{2} - \omega_k t_o\right) - \sin \omega_k t_o \cos \omega_k t_o\right] \tag{9-39}$$

The effective input resistance may be defined as the equivalent resistance that would dissipate the same average power as the detector in each carrier-frequency cycle as the result of application of the same modulated voltage. According to this definition, the power dissipated in each cycle is equal to one-half the square of the crest applied voltage in that cycle divided by the effective input resistance.

$$P_i = \frac{\frac{1}{2}[E_k(1 + M \cos \omega_m t)]^2}{r_e} \tag{9-40}$$

Substitution of Eq. (9-39) in Eq. (9-40) gives the following expression for the effective input resistance:

$$r_e = \frac{\pi r_p}{\left(\frac{\pi}{2} - \omega_k t_o\right) - \sin \omega_k t_o \cos \omega_k t_o} \tag{9-41}$$

The values of $\omega_k t_o$, $\sin \omega_k t_o$, and $\cos \omega_k t_o$ for various values of r/r_p can be found with the aid of Eq. (9-35).

A similar analysis (see Prob. 9-6) shows that when the condenser is omitted from the detector circuit the effective input resistance is

$$r_e = 2(r_p + r) \tag{9-42}$$

In Fig. 9-26, r_e/r_p is shown as a function of r/r_p for infinite and for zero condenser reactance. In both cases $r_e = 2r_p$ when $r_b = 0$. For large values of r/r_p, r_e approaches the value $\frac{1}{2}r$ for zero reactance and $2r$ for infinite reactance. The condenser therefore lowers the effective input resistance.

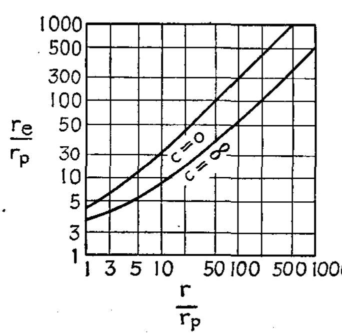


FIG. 9-26.—Effect of load resistance and capacitance upon effective input resistance of a diode detector circuit.

9-17. Ideal Linear Diode Detection of A-m Wave of General Form.—An examination of the equations of Secs. 9-14 and 9-15 shows that the integrands of the integrals evaluated in the analysis of the ideal linear diode detector do not involve the factor $E_k(1 + M \sin \omega_m t)$ which defines the form of the envelope of the modulated wave, and that this factor is therefore carried outside of the integral sign through all operations. It appears in Eqs. (9-26) and (9-28) in unmodified form. If the modulated wave is of the much more general form $E_k f(t) \sin \omega_k t$, in which $f(t)$ is any function of time, the envelope factor $E_k f(t)$ replaces the factor

$E_k(1 + M \sin \omega_m t)$ throughout the analysis and, particularly, in Eqs. (9-26) and (9-28).[1] This fact indicates that the wave form of a complex modulation voltage by means of which the carrier was modulated is not distorted in the process of detection by an ideal linear diode detector. The output voltage contains only the frequencies of the original modulation voltage. For a modulated voltage of general form, Eq. (9-37) becomes

$$E_r = DE_k f(t) \tag{9-43}$$

9-18. Ideal Linear Diode Detection of the Sum of Two Voltages of Different Frequencies.—The linear diode detector is sometimes used to produce a difference frequency as the result of the simultaneous application of two sinusoidal voltages of unequal frequency. To analyze this

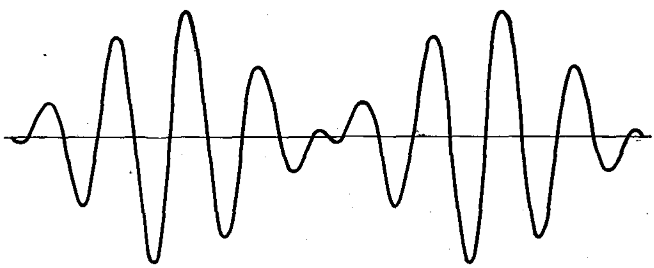


FIG. 9-27.—Resultant of two sine waves of equal amplitude and 5 to 4 frequency ratio.

process it is necessary first to study the form of the resultant wave produced by the addition of two sine waves. Figure 9-27 shows the resultant of two sine waves of equal amplitude and 5 to 4 frequency ratio. By adding waves that have both unequal frequency and unequal amplitude, the reader may show that, in general, the resultant of two sine waves of somewhat different frequency resembles an amplitude-modulated wave. It differs from such a wave in that the points at which the resultant crosses the axis are not equally spaced, indicating additional phase modulation.

The form of the envelope of the resultant wave and the manner in which the phase varies with time can be determined analytically. Let the two impressed voltages be $e_1 = E_1 \cos \omega_1 t$ and $e_2 = E_2 \cos \omega_2 t$. The resultant voltage is

$$e = E_1 \sin \omega_1 t + E_2 \sin \omega_2 t \tag{9-44}$$

$$= E_1 \sin \omega_1 t + E_2 \sin [\omega_1 t - (\omega_1 - \omega_2)t] \tag{9-45}$$

By expanding the second term of Eq. (9-45) and collecting terms in $\sin \omega_1 t$ and $\cos \omega_1 t$, Eq. (9-44) may be transformed into

$$e = E \sin (\omega_1 t - \phi) \tag{9-46}$$

[1] RODER, HANS, *Proc. I.R.E.*, **20**, 1946 (1932).

in which

$$E = E_1 \sqrt{1 + 2h \cos (\omega_1 - \omega_2)t + h^2} \tag{9-47}$$

$$\phi = \tan^{-1} \frac{h \sin (\omega_1 - \omega_2)t}{1 + h \cos (\omega_1 - \omega_2)t} \tag{9-48}$$

$$h = \frac{E_2}{E_1} \tag{9-49}$$

If e_1 represents the component of the applied voltage that has the larger amplitude, then $E_1 \geqq E_2$, and $h \leqq 1$.

Figure 9-28*a* shows how the phase angle ϕ varies throughout one cycle of difference frequency, $f_1 - f_2$, for various values of h between 0.2 and 1. When the amplitudes are equal, $h = 1$ and ϕ varies between

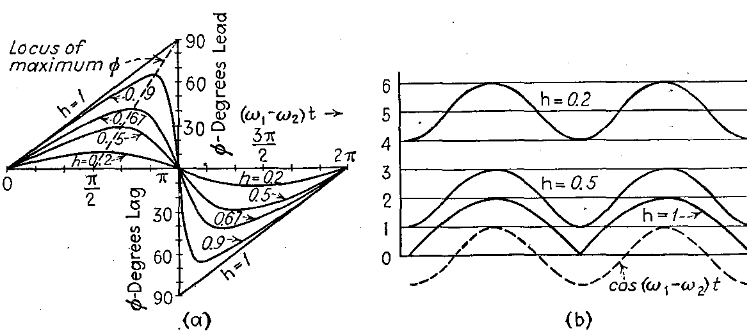


FIG. 9-28.—Effect of amplitude ratio h upon (*a*) phase modulation and (*b*) form of envelope of the resultant of two sine waves of unequal frequency.

+90 and −90 degrees, a sudden reversal of phase taking place at the instant $(\omega_1 - \omega_2)t = \pi$. As h is decreased by changing the amplitude of either component, the maximum value of ϕ rapidly decreases, being only slightly greater than $11\frac{1}{2}$ degrees when $h = 0.2$. Figure 9-28*b* shows how E varies throughout the difference-frequency cycle, for various values of h, E_2 being kept constant at unity. Comparison with the dotted curve, which represents $\cos (\omega_1 - \omega_2)t$, shows that the amplitude varies at a fundamental frequency equal to the difference frequency but also contains a steady component and harmonics of the difference frequency. E. B. Moullin has shown[1] that

$$\sqrt{1 + 2h \cos (\omega_1 - \omega_2)t + h^2} = b_o + b_1 \cos (\omega_1 - \omega_2)t \\ + b_2 \cos 2(\omega_1 - \omega_2)t + b_3 \cos 3(\omega_1 - \omega_2)t + \cdots \tag{9-50}$$

Each coefficient of this series is in itself an infinite series that converges rapidly if $h \leqq 1$. The following expressions for the first two coefficients are accurate within 1 per cent when h does not exceed unity:

[1] MOULLIN, E. B., *Wireless Eng.*, **9**, 378 (1932). See also F. M. COLEBROOK, *Wireless Eng.*, **9**, 195 (1932).

$$b_o = 1 + \frac{h^2}{2^2} + \frac{h^4}{2^6} + \frac{h^6}{2^8} + \cdots \tag{9-51}$$

$$b_1 = h\left(1 - \frac{h^2}{2^3} - \frac{h^4}{2^6} - \frac{5h^6}{2^{10}} - \cdots\right) \tag{9-52}$$

Equation (9-50) proves rigorously that E contains a steady component, a fundamental difference-frequency component, and all harmonics of the difference frequency.

This analysis shows that the resultant of two sine waves of different frequencies is equivalent to a wave whose frequency is that of the component of higher amplitude and whose amplitude and phase are modulated simultaneously. The modulation voltage with which the amplitude of this equivalent wave is modulated contains the fundamental difference frequency and all its harmonics. The phase modulation, the fundamental frequency of which is also equal to the difference frequency, may be made as small as desired by making h small. Under the assumption that h is small enough so that the phase modulation is negligible, the sum of the two sine waves is equivalent to an amplitude-modulated wave $E_1 f(t) \cos \omega_1 t$, in which $f(t)$ is given by Eq. (9-50). Substitution of Eq. (9-50) in Eq. (9-43) shows that when this voltage is applied to an ideal linear diode detector the output voltage is

$$E_r = DE_1[b_0 + b_1 \cos(\omega_1 - \omega_2)t + b_2 \cos 2(\omega_1 - \omega_2)t + \cdots] \tag{9-53}$$

The value of D in Eq. (9-53) may be found from the curves of Fig. 9-25.

Examination of Eqs. (9-51) and (9-52) and similar equations for the higher-order coefficients of Eq. (9-50) shows that the ratio of the amplitude of the nth harmonic to the amplitude of the fundamental difference-frequency output of the detector is roughly proportional to h^{n-1} and that the exactness of the proportionality increases rapidly as h is decreased. The harmonic content may, therefore, be made as small as desired by making the amplitude of one component of the input much larger than the other. The amplitude of the fundamental modulation-frequency component of detector output voltage is

$$E_o = E_1 Dh\left(1 - \frac{h^2}{2^6} - \frac{h^4}{2^6} - \frac{5h^6}{2^{10}} - \cdots\right) \tag{9-54}$$

As h is decreased, this approaches the limiting value

$$E_o = E_1 Dh = E_2 D \tag{9-55}$$

This shows that, if one input voltage is considerably larger than the other, the amplitude of the fundamental difference-frequency component of linear diode detector output depends only upon the amplitude of the smaller input voltage.[1] This is of importance in the design of super-

[1] SLAFFER, M., *Wireless Eng.*, **11**, 25 (1934).

heterodyne radio receivers and heterodyne oscillators (see Secs. 9-25 and 10-52).

9-19. Effect of Curvature of Diode Characteristic.—Equations (9-26), (9-28), and (9-43) hold only for linear characteristics. If the diode characteristic is not linear, the wave of plate current is not an exact replica of the upper half of the applied modulated voltage. This fact suggests that the output contains harmonics of the modulation voltage. In general, the relation between total instantaneous diode plate current and plate voltage may be expressed in terms of a series

$$i_b = p_1e_b + p_2e_b^2 + p_3e_b^3 + \cdots \tag{9-56}$$

Analysis by the method used in deriving Eqs. (9-28) and (9-43) shows that the voltage output across the condenser-resistance load is of the form

$$E_o = P_1E_kf(t) + P_2E_k^2f^2(t) + P_3E_k^3f^3(t) + \cdots \tag{9-57}$$

If the input voltage is a simple modulated wave, then $f(t)$ is

$$(1 + M \sin \omega_m t),$$

and the second term gives rise to steady, fundamental, and second-harmonic components, and the third term to steady, fundamental modulation-frequency, second-harmonic, and third-harmonic components. If the input wave is modulated by several modulation frequencies, the output also contains frequencies equal to the sums and differences of the modulation frequencies and their multiples. The fundamental modulation-frequency components of the output arising from the second- and higher-order terms of Eq. (9-55) destroy the linearity between modulation-frequency output voltage and carrier amplitude that is obtained when the characteristic is linear. The sum and difference frequencies produce objectionable dissonance when the detector is used in the reproduction of music. The form of the coefficients of Eq. (9-55) may be shown to be such as to indicate a reduction of distortion with increase of load resistance.

The action of practical diode detectors with high load resistance approximates that of the ideal linear detector at high excitation voltages and is intermediate between that of the ideal linear diode and a parabolic diode at low excitation voltage. In practice, detection is said to be linear if the desired output voltage is substantially proportional to the carrier voltage throughout the useful operating range of the detecting device.

9-20. Design of Practical Diode Detectors.—In order to ensure low distortion resulting from curvature of the diode characteristic, to minimize loading of the source of detector input voltage, and to attain high detection efficiency and hence output voltage, the load resistance should be high. There is no advantage, however, in making the load resistor much greater than the grid coupling resistor of the amplifier that follows the detector. Values ranging from 100,000 ohms to 1 megohm are normally used.

Care should be exercised in the choice of the condenser capacitance of a diode detector, since incorrect capacitance may cause several types of distortion. The charging current that can flow into the condenser for a given value of instantaneous impressed voltage is limited by the tube, by the impedance of the source, and by leakage through the load resistance. If the condenser is too large, many carrier-frequency cycles may be required for the condenser voltage to reach equilibrium after an increase or a decrease of amplitude of impressed alternating voltage. Then, if the impressed amplitude fluctuates rapidly, the changes in condenser voltage are smaller than the changes in amplitude. The change in condenser voltage for a given change of impressed amplitude decreases

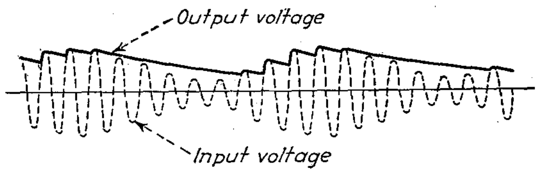


FIG. 9-29.—Nonlinear distortion in a diode detector resulting from improper choice of circuit values.

as the frequency of the amplitude variation is increased. Since the envelope of a modulated wave varies at modulation frequency, it follows that the use of too large a condenser causes frequency distortion of the detector output. With a large condenser and load resistance and low tube resistance and source resistance, the condenser voltage may be able to respond fully to an increase of amplitude of impressed voltage, but not to a decrease of amplitude, as illustrated in Fig. 9-29. Frequency distortion will then be accompanied by nonlinear distortion in the detection of a modulated wave. Both of these types of distortion may be made negligible by satisfying the relation $2\pi f_m rC \leqq \sqrt{(1/M)^2 - 1}$, in which f_m is the highest modulation frequency with which the carrier is modulated.[1]

When the detector is resistance-capacitance coupled to the grid of an amplifier tube, the detector load resistance at modulation frequency is that of r in parallel with the amplifier grid leak (the coupling capacitance being assumed negligible at modulation frequency), whereas the d-c resistance is that of r alone. To prevent the amplitude distortion dis-

[1] TERMAN, F. E., and MORGAN, N. R., *Proc. I.R.E.*, **18**, 2160 (1930); NELSON, J. R., *Proc. I.R.E.*, **19**, 489 (1931); NELSON, J. R., *Proc. I.R.E.*, **20**, 989 (1932); TERMAN, F. E., and NELSON, J. R., *Proc. I.R.E.*, **20**, 1971 (1932); COCKING, W. T., *Wireless Eng.*, **12**, 595 (1935).

cussed above, it is then necessary to satisfy the relation

$$2\pi f_m r_m C \leqq \sqrt{\left(\frac{r_m}{rM}\right)^2 - 1}$$

where f_m is the highest modulation frequency and r_m is the effective load resistance at modulation frequency.[1] The difference between the modulation-frequency and d-c resistances of the load with resistance-capacitance coupling causes another type of amplitude distortion because of complete cutting off of diode current during the inward modulation crests when the degree of modulation is high. This type of distortion, which results in *clipping* of the negative peaks of modulation output voltage, may be prevented by keeping the degree of modulation less than a limiting value given approximately by the relation $M = (r_m + 2r_s)/(r + 2r_s)$, where r_s is the effective resistance of the source at carrier frequency.[2] Peak clipping is explained graphically in Sec. 9-23 and illustrated in Fig. 9-35.

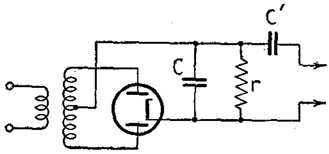


FIG. 9-30.—Full-wave diode detector circuit.

The load resistance not only allows the condenser to discharge in response to the envelope of the impressed voltage, but also allows it to discharge during the portions of the carrier-frequency cycle in which the instantaneous impressed voltage is less than the condenser voltage. This causes the condenser voltage to vary by a small amount at carrier frequency. The presence of carrier-frequency voltage in the output is the limiting factor in determining the minimum condenser capacitance that can be used. The ratio of the amplitude of the carrier-frequency voltage in the output to the amplitude of the signal-frequency voltage is approximately $1/\pi f_k r_m C$, where f_k is the carrier frequency. Since the carrier frequency is usually at least ten times the highest signal frequency, there is no difficulty in making C large enough to meet this requirement without making it too large to avoid modulation-frequency distortion. A capacitance of approximately 150 $\mu\mu$f is suitable for C in the broadcast band.

Some reduction in the required size of C results from the use of the full-wave diode circuit of Fig. 9-30. The output of this circuit does not contain the carrier frequency, and so the condenser need be only large enough to remove the second harmonic of the carrier frequency.

[1] ROBERTS, F., and WILLIAMS, F. C., *J. Inst. Elec. Eng. (London)*, **75**, 379 (1934); BENNON, S., *Proc. I.R.E.*, **25**, 1565 (1937).

[2] WHEELER, H. A., *Proc. I.R.E.*, **26**, 745 (1938); COURT, W. P. N., *Wireless Eng.*, **16**, 548 (1939); STURLEY, K. R., *Wireless Eng.*, **17**, 19 (1940); PREISMAN, A., *Communications*, August, 1940, p. 18; September, 1940, p. 8.

9-21. Linear Plate Detection (Linear Transrectification).—If a modulated wave of voltage is applied to the grid of a triode biased to cutoff, plate current flows only during the positive half cycles of the modulated voltage, as shown in Fig. 9-31. The action is similar to that of a diode except that the amplification of the tube greatly increases the

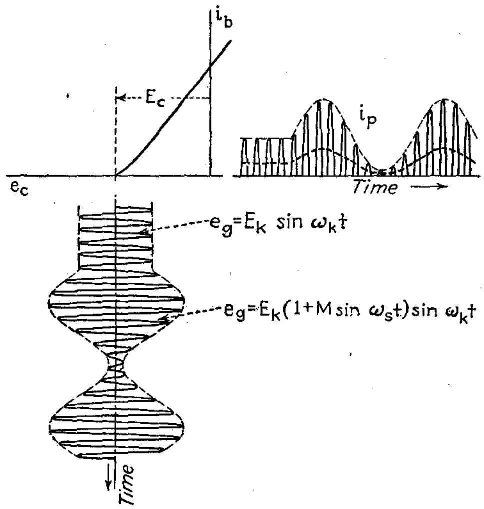


FIG. 9-31.—Linear plate detection.

sensitivity and that, if the grid is not allowed to swing positive, less load is placed on the voltage source. If the transfer characteristic were linear, there would be no signal distortion, and signal output would be proportional to carrier amplitude. Curvature of the characteristic causes distortion, which may be reduced by the use of high plate load resistance. Usually the bias is somewhat less than the cutoff value, and the action at low excitation is by curvature of the transfer characteristic. At large values of excitation voltage the action approximates that of the linear diode. The distortion of the linear plate detector is in general greater than that of the diode detector, but it has the advantages of greater sensitivity and high input impedance. Figure 9-32 shows the circuit of a typical linear plate detector. In this circuit, bias is obtained from the steady component of voltage across the load resistance.

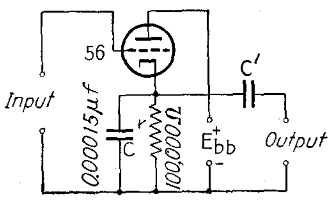


FIG. 9-32.—Linear plate detector.

9-22. Linear Grid Detection.—Figure 9-33*a* shows a detector, similar in form to the diode detector of Fig. 9-28, in which the grid and cathode of a triode are used as the rectifier and the output is taken from across the tube, as explained in Sec. 9-14. Since the modulation-frequency output voltage appears between the grid and the cathode of the triode, the plate current contains a modulation-frequency component, and amplified modulation-frequency voltage may be taken from across a plate load resistor, as shown in Fig. 9-33*b*. Because there is also carrier-frequency voltage present between the grid and the cathode, the condenser C_b must be used in order to suppress carrier voltage from the output. The grid resistor may be connected across the grid condenser, instead of between the grid and the cathode. The grid circuit is then similar in form and action to the diode circuit of Fig. 9-21. Choice of the values of r and C is governed by the same considerations as those determining the choice of the resistance and capacitance of diode circuits, as discussed in Sec. 9-20.

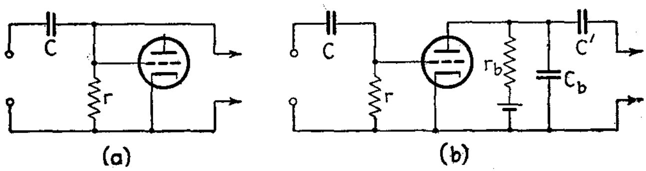


FIG. 9-33.—(*a*) Diode detector using the grid circuit of a triode; (*b*) linear grid detector.

In the linear grid detector, rectification in the grid circuit produces the biasing voltage necessary in using the triode to amplify the modulation-frequency voltage. The bias increases with the amplitude of the input voltage and at high values of input amplitude may be so great that the operating point is in a region where the transfer characteristic has high curvature and plate-current cutoff may occur. This results in two objectionable effects: The modulation-frequency voltage is distorted in amplification, and detection by transrectification is superimposed upon grid detection. Since the modulation-frequency output resulting from transrectification is opposite in phase to that resulting from grid rectification, the detection efficiency decreases at large values of excitation. Reduction of detection efficiency may, however, be prevented by the use of an inductance in place of the grid leak r. The advantage of the linear grid detector over the linear diode detector and the linear plate detector is its higher sensitivity, which results from amplification of the modulation-frequency voltage generated in the grid circuit.[1] It approaches linearity more closely than the linear plate detector. The high sensitivity of the linear grid detector can be attained without the attending disadvantages by using a diode detector and a separate stage of audio-frequency amplification. By the use of a tube that combines a diode and a triode or pentode in a single envelope, this can be accomplished without increasing the number of tubes.

[1] TERMAN, F. E., and MORGAN, N. R., *Proc. I.R.E.*, **18**, 2160 (1930).

If sufficient additional fixed positive bias is used in the circuit of Fig. 9-33*b*, grid current flows continuously and detection is the result of curvature of the grid characteristic, rather than of complete rectification. Although used in early radio receivers, this type of detection has been entirely superseded by linear diode detection in communication apparatus.

9-23. Graphical Detection Analysis.—The similarity between the variation of the average current of a detector tube with amplitude of exciting voltage and the variation of steady plate current of an amplifier tube with direct grid voltage suggests the use of detection diagrams which are analogous to amplifier plate diagrams.[1] For diode or other simple rectification, the curves of average plate current against average plate voltage for fixed values of sinusoidal exciting voltage are called *rectification characteristics.* For plate detection or other transrectification, the curves of average current to an electrode *vs.* average voltage of that

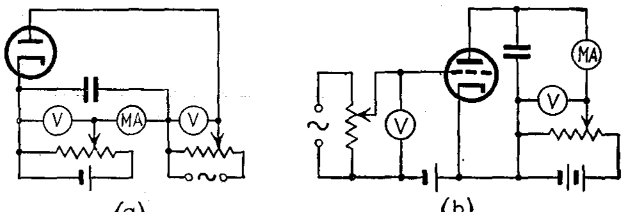


FIG. 9-34.—Circuits for the determination of (*a*) rectification characteristics and (*b*) transrectification characteristics.

electrode for fixed values of exciting voltage applied to any other electrode are called *transrectification characteristics.* Circuits for the determination of these characteristics are shown in Fig. 9-34. For the ideal linear diode the rectification characteristics obey the equation

$$I_{bk} = \frac{E}{\pi r_p}\left[\cos \omega_k t_o + \left(\omega_k t_o - \frac{\pi}{2}\right) \sin \omega_k t_o\right]$$

where $\sin \omega_k t_0 = E_b/E$, E_b is the average plate voltage, and E is the amplitude of excitation. Ideal linear diode rectification characteristics are shown in Fig. 9-35. Experimental characteristics for a type 6H6 diode are shown on page 684.

When a modulated voltage $E_k(1 + M \sin \omega_m t) \sin \omega_k t$ is applied to the detector, in place of the steady exciting voltage used in obtaining the characteristics, the amplitude of the exciting voltage varies in a manner analogous to the variation of grid voltage of an amplifier and causes an

[1] BALLANTINE, S., *Proc. I.R.E.*, **17,** 1153 (1929); NELSON, J. R., *Proc. I.R.E.*, **19,** 489 (1931); NELSON, J. R., *Electronics*, March, 1931, p. 550, July, 1931, p. 14; LUCAS, G. S. C., *Wireless Eng.*, **9,** 384 (1932); TURNER, P. K., *Wireless Eng.*, **9,** 384 (1932); KILGOUR, C. E., and GLESSNER, J. M., *Proc. I.R.E.*, **21,** 930 (1933).

analogous variation of average current. If static and dynamic load lines are drawn on the detection diagram, corresponding to the d-c resistance and the a-c resistance at modulation frequency, respectively, the fundamental and harmonic signal amplitudes may be determined graphically by the methods of Chap. 4. A typical detection diagram is illustrated in Fig. 9-35 for an ideal linear diode and a load whose resistance at modulation frequency is less than its d-c resistance (resistance-capacitance-coupled load). The "static" operating point O, which corresponds to zero modulation factor, is determined by the intersection of the static load line with the rectification characteristic for which $E = E_k$, the unmodulated carrier amplitude. The swing of E along the dynamic load line is

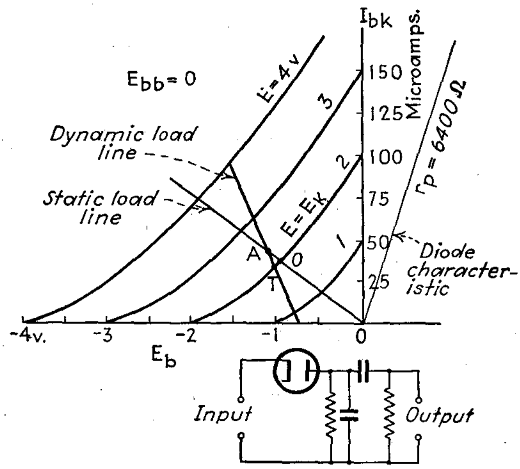


FIG. 9-35.—Detection plate diagram for ideal linear diode with load whose modulation frequency resistance is less than its d-c resistance.

equal to ME_k. It can be seen from Fig. 9-35 that, if the a-c load resistance is much lower than the d-c load resistance, cutoff may occur when the degree of modulation is high.[1] This causes peak-clipping, which has been discussed in Sec. 9-20.

9-24. Series Expansion and Equivalent Circuit for Detection.—Since the direct voltages of all electrodes except that in which the detected current flows are normally maintained constant, the average current to that electrode is a function of only the amplitude of the exciting voltage and of the average voltage of that electrode. The average plate current of a diode or plate detector may therefore be expressed by the functional equation

$$I_{bk} = f(E_b + UE) \tag{9-58}$$

[1] KILGOUR and GLESSNER, *loc. cit.* For an analysis that takes load reactance into account, see A. PREISMAN, "Graphical Constructions for Vacuum Tube Circuits," Chap. 6, McGraw-Hill Book Company, Inc., New York, 1943.

in which E_b is the direct voltage applied to the plate, E is the amplitude of the sinusoidal exciting voltage applied to the plate or other electrode, and U is a factor analogous to amplification factor. U indicates the relative effectiveness of the direct plate voltage and the excitation voltage in controlling the average plate current.

When the exciting voltage is a modulated carrier voltage, the amplitude of the exciting voltage varies at modulation frequency and causes the average current to vary at a fundamental frequency equal to the modulation frequency. The varying component of the average detector current can be expressed in terms of a series analogous to that for an amplifier with varying grid voltage, derived from Eq. (9-58) in a manner similar to that used in the derivation of Eqs. (3-40) and (3-56). The series expansion for the alternating component of average detector current is

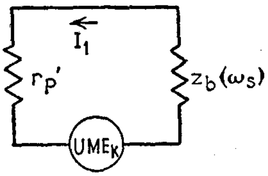


FIG. 9-36.—Equivalent plate circuit for detection.

$$i_d = \Sigma A_1 e + \Sigma A_2 e^2 + \Sigma A_3 e^3 + \cdots \tag{9-59}$$

in which e is the instantaneous amplitude of the modulated input voltage, and the coefficients are of similar form to those of Eq. (3-57). The first coefficient is given by the relation

$$A_1 = \frac{U}{r_p' + z_b(\omega_m)} \tag{9-60}$$

where r_p', the *detection plate resistance*, is defined by the equation

$$r_p' = \frac{\partial E_b}{\partial I_{bk}} \tag{9-61}$$

and U, the *detection mu-factor*, is defined by the equation

$$U = \frac{\partial E_b}{\partial E} \qquad (I_{bk} = \text{const.}) \tag{9-62}$$

For the sinusoidally modulated voltage,

$$e_m = E_k(1 + M \sin \omega_m t) \sin \omega_k t$$

$e = E_k(1 + M \sin \omega_m t)$. Substitution of this value of e and of Eq. (9-60) in the first term of Eq. (9-59) shows that the fundamental modulation-frequency component of detector current is

$$I_1 = \frac{UME_k}{r_p' + z_b(\omega_m)} \tag{9-63}$$

and that the fundamental modulation-frequency current and output voltage can, therefore, be found from the *equivalent plate circuit for* *detection* shown in Fig. 9-36. Equation (9-63) may also be written in the form[1]

$$I_1 = \frac{\partial I_{bk}}{\partial E} \cdot \frac{r_p' M E_k}{r_p' + z_b(\omega_m)} \tag{9-64}$$

$\partial I_{bk}/\partial E$, which is analogous to transconductance, is called the *rectification coefficient* (in simple rectification) or *transrectification coefficient* (in transrectification).

The detection plate resistance r_p', the rectification or transrectification coefficient $\partial I_{bk}/\partial E$, and the detection mu-factor U may be determined

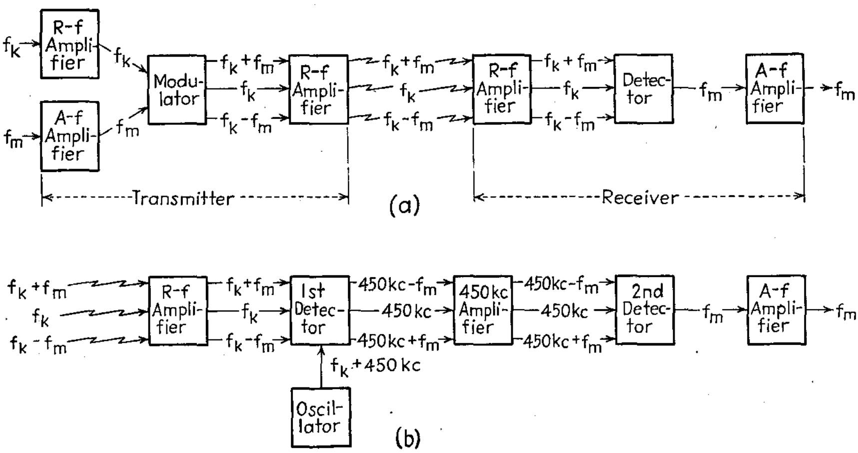


FIG. 9-37.—(*a*) Essential steps in communication by means of a-m waves. (*b*) Block diagram of the superheterodyne receiver.

from the rectification or transrectification characteristics at the operating point, in exactly the same way that the plate resistance, transconductance, and amplification factor are found from the static plate characteristics. A bridge for the dynamic measurement of r_p' is described on page 655.

Equations (9-58) to (9-64) may also be applied to grid detectors, proper changes in subscripts being made. Account must be taken, however, of the signal-voltage amplification in determining the signal voltage output in the plate circuit.

9-25. Radio Communication by Means of A-m Waves.—The block diagram of Fig. 9-37*a* shows the essential steps in communication by means of amplitude-modulated waves. The reader should note the frequencies present at various points in the system. In the *superheterodyne* type of receiver now used almost exclusively in broadcast

[1] BALLANTINE, *loc. cit.* See also E. L. CHAFFEE, and G. H. BROWNING, *Proc. I.R.E.*, **15**, 113 (1927).

reception, the carrier frequency is transformed to a value of 450 kc before the modulated signal is amplified to the level required for detection. This is accomplished by impressing the modulated signal and the output of a local oscillator, whose frequency is 450 kc higher than that of the incoming carrier, upon a frequency converter or *first detector*, as shown in Fig. 9-37*b*. The 450-kc *intermediate-frequency amplifier* selects the desired difference-frequency components from the output of the first detector. Because most of the amplification is accomplished at the fixed intermediate frequency, the superheterodyne circuit makes possible the high amplification and selectivity that can be attained with a large number of tuned amplifier stages without requiring the simultaneous variation of more than three condensers (first amplifier, oscillator, and first detector) in the operation of the receiver.

9-26. Automatic Gain (Volume) Control (A. V. C.).[1]—Because of fluctuations in the field intensity of radio waves striking the antenna, the audio-frequency output of a radio receiver also fluctuates unless the receiver is provided with some sort of automatic compensation. The simplest method of maintaining constant output is by the use of automatic control of the amplification of the radio-frequency amplifier stages. Automatic gain control is based upon the combination of two principles already discussed: (1) The amplification of an r-f amplifier varies with the transconductance of the tubes, which in turn depends upon the operating voltages; (2), the direct output voltage of a detector varies with the carrier amplitude of the modulated input voltage. The output of the amplifier is applied to a detector, the direct-voltage output of which controls the voltages of one or more grids of the amplifier tubes in such a manner as to reduce the amplification when the input of the amplifier rises.

Figure 9-38 shows a type of automatic gain control commonly applied to radio receivers in which linear diode detection is used. For simplicity, only one stage of the r-f amplifier is shown and the amplifier tube is shown as a triode, rather than as a pentode. r_1 is the detector load resistance, across which the modulation-frequency and direct output voltages appear. C_1 is the detector condenser. The d-c detector output is applied to the grid of the r-f amplifier tube through the high resistance (approximately 1 megohm, usually) r_2, which, in conjunction with the condenser C_2, filters out the modulation-frequency voltage. The modulation-frequency output of the detector is applied to the audio amplifier through the condenser C_3. As the r-f output of the amplifier increases, the direct voltage across r_1 goes up. The polarity of this voltage is such as to make the r-f amplifier grid more negative and thus reduce the r-f gain. The rapidity with which the grid bias of the r-f amplifier tube can

[1] WHEELER, H. A., *Proc. I.R.E.*, **16**, 30 (1928).

change decreases with increase of r_2 or C_2. Too rapid response tends to cause degeneration at low audio frequencies. C_2 serves the additional functions of affording a low-impedance connection between the transformer secondary and the cathode in the r-f amplifier and of preventing the r-f output of the amplifier from being applied to the grids of controlled tubes.

Similar action may be obtained with other types of detectors, but the diode detector has the advantage that the d-c output is proportional to the carrier amplitude and independent of the degree of modulation.[1]

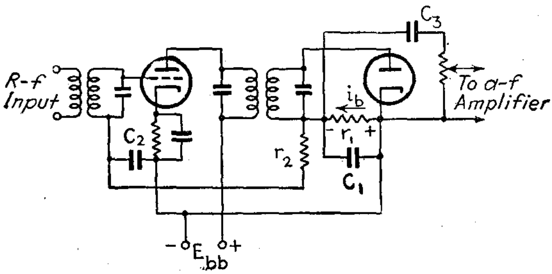


FIG. 9-38.—Basic circuit of diode automatic gain control as applied to r-f amplifier.

9-27. Frequency Modulation.[2]—Frequency modulation is the process whereby the frequency of one wave or oscillation is varied with time in accordance with the time variations of another wave or oscillation. Figure 7-18 shows a wave of a quantity y, the value of which varies periodically with time at a frequency that is a function of time. The amplitude has a constant value K. If the variation of frequency is sufficiently slow so that each cycle can be assumed to be essentially sinusoidal, the value of y at any instant t_1 is equal to sin θ, where θ is the number of degrees elapsed in the cycle in progress at the instant t_1. It can be seen from Fig. 9-39 that the value of θ, and hence of y, at that instant depends not only upon the frequency at that instant, but also upon the manner in which the frequency varied previously. If n is the number of cycles that have elapsed between the instants $t = 0$ and $t = t_1$, then the value of y at the

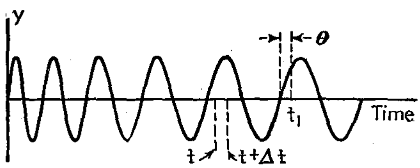


FIG. 9-39.—Frequency-modulated wave.

[1] An excellent review of various systems of automatic gain control is given by W. T. COCKING in *Wireless Eng.*, **11**, 406, 476, 483, 542 (1934); **12**, 87 (1935).

[2] CARSON, J. R., *Proc. I.R.E.*, **10**, 57 (1922); **7**, 187 (1919); VAN DER POL, B., *Proc. I.R.E.*, **18**, 1194 (1930); SMITH, C. H., *Wireless Eng.*, **7**, 609 (1930); RODER, H., *Proc. I.R.E.*, **19**, 2145 (1931) (with bibliography of 21 items); LUCK, D. G. C., and RODER, H., *Proc. I.R.E.*, **20**, 884 (1932); ARMSTRONG, E. H., *Proc. I.R.E.*, **24**, 689 (1936). See also supplementary bibliography at the end of this chapter.

instant t_1 may be expressed by the relation

$$y_1 = K \sin n(2\pi) \tag{9-65}$$

(Thus, in Fig. 9-39, $n = 5 + \theta/2\pi$ and $y_1 = K \sin (10\pi + \theta) = \sin \theta$. t_1 is equal to the number of cycles between zero and t_1 multiplied by the period of each cycle.) The portion Δn of a cycle that elapses during any small interval between t and $t + \Delta t$ is equal to the ratio of the interval Δt to the period T for that cycle. Thus

$$\Delta n = \frac{\Delta t}{T} = f\,\Delta t \tag{9-66}$$

where f is the value of the frequency at the instant t. Therefore the number of cycles in the interval 0 to t_1 is

$$n = \int_0^{t_1} f\,dt \tag{9-67}$$

If the instantaneous frequency f is varied sinusoidally at modulation frequency f_m about an unmodulated carrier value f_k in accordance with the relation

$$f = f_k(1 + B \cos 2\pi f_m t) \tag{9-68}$$

then

$$n = \int_0^{t_1} f_k(1 + B \cos 2\pi f_m t)\,dt \tag{9-69}$$

$$n = f_k\left(t_1 + \frac{B}{2\pi f_m} \sin 2\pi f_m t_1\right) \tag{9-70}$$

and

$$y_1 = K \sin\left(2\pi f_k t_1 + \frac{Bf_k}{f_m} \sin 2\pi f_m t_1\right) \tag{9-71}$$

Since t_1 may have any value, the subscript $_1$ may be omitted and Eq. (9-71) written in the form

$$y = K \sin (\omega_k t + M_f \sin \omega_m t) \tag{9-72}$$

in which $M_f = B\omega_k/\omega_m = Bf_k/f_m$. Equation (9-72) is that of a constant amplitude wave, the frequency of which is varied in accordance with Eq. (9-68). Bf_k, which is shown by Eq. (9-68) to be the maximum change in frequency from the unmodulated carrier value, is called the *frequency deviation* and will be represented by the symbol f_d. M_f, which is equal to f_d/f_m, the ratio of the frequency deviation to the modulation frequency, is called the *deviation ratio* or *frequency modulation index*. Equation (9-72) may be written

$$y = K \sin\left(\omega_k t + \frac{f_d}{f_m} \sin \omega_m t\right) \tag{9-72a}$$

Equation (9-72) may be expanded into the series[1]

$$
\begin{aligned}
y = K\{J_0(M_f) \sin \omega_k t &+ J_1(M_f)[\sin (\omega_k + \omega_m)t - \sin (\omega_k - \omega_m)t] \\
&+ J_2(M_f)[-\sin (\omega_k + 2\omega_m)t + \sin (\omega_k - 2\omega_m)t] \\
&+ J_3(M_f)[\sin (\omega_k + 3\omega_m)t - \sin (\omega_k - 3\omega_m)t] \\
&+ J_4(M_f)[-\sin (\omega_k + 4\omega_m)t + \sin (\omega_k - 4\omega_m)t] \\
&+ \cdots \}
\end{aligned}
\tag{9-73}
$$

in which the coefficients $J_n(M_f)$ are Bessel functions of the first kind. The coefficients for any given value of M_f may be evaluated by the use of published tables that list numerical values of $J_0(x)$, $J_1(x)$, $J_2(x)$, etc., for various values of x.[2]

Equation (9-73) shows that a frequency-modulated wave theoretically contains an infinite number of side frequencies, separated in frequency value by the modulation frequency. The values of $J_n(M_f)$ decrease very rapidly, however, beyond a value of n that is slightly greater than M_f. This can be seen from Fig. 9-40, which shows as a function of M_f the highest order of the side frequencies whose amplitudes are greater than 5 per cent or than 10 per cent of that of the unmodulated carrier.[3]

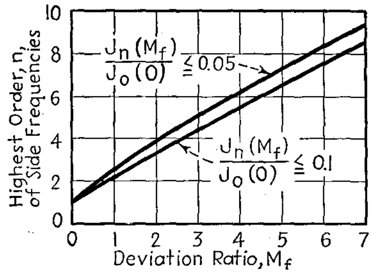


FIG. 9-40.—Highest order of side frequencies whose amplitudes are greater than 5 per cent and than 10 per cent of the unmodulated carrier, plotted as a function of the deviation ratio.

Because M_f is inversely proportional to f_m, the number of side frequencies of appreciable amplitude decreases with increase of modulation frequency and, when the deviation is large in comparison with the highest modulation frequency, the over-all band width does not vary greatly as the modulation frequency changes throughout the audio-frequency range. This fact can be shown with the aid of Fig. 9-41 by noting that the total band width of the upper and of the lower frequency bands in which the side frequencies lie is equal to nf_m, where n is the order of the highest-order side frequency component of appreciable amplitude. Inspection shows that for the 10 per cent curve n is approximately equal

[1] VAN DER POL, *loc. cit.;* RODER, *loc. cit.*

[2] JAHNKE, E., and EMDE, F., "Tables of Functions with Formulae and Curves," 3d ed., B. G. Teubner, Leipzig and Berlin (1938); GRAY, A., and MATHEWS, G. B., "A Treatise on Bessel Functions," Macmillan & Company, Ltd., London (1922). See also RODER, *op. cit.*, p. 2175. It is of interest to note that $J_0(0) = 1$ and $J_n(0) = 0$. Hence, when M_f is zero, *i.e.*, when the wave is unmodulated, Eq. (9-73) reduces to $y = K \sin \omega_k t$, as it should.

[3] RODER, *loc. cit.* In using Fig. 9-40, n is read as the next lower integer to the value indicated by the curve. Thus, for $M_f = 1$, $n = 2$.

to $1.0 + 1.1M_f$. Hence the width of the upper and of the lower band is approximately

$$\left[1.0 + 1.1\left(\frac{Bf_k}{f_m}\right)\right]f_m = f_m + 1.1f_d$$

where f_d is Bf_k, the deviation in cycles per second. The total band width is approximately $2f_m + 2.2f_d$, which is very nearly equal to $2f_d$ when f_d/f_m, *i.e.*, M_f, is large. A similar expression may be derived for the 5 per cent curve of Fig. 9-40 but, because of greater curvature of this curve, the

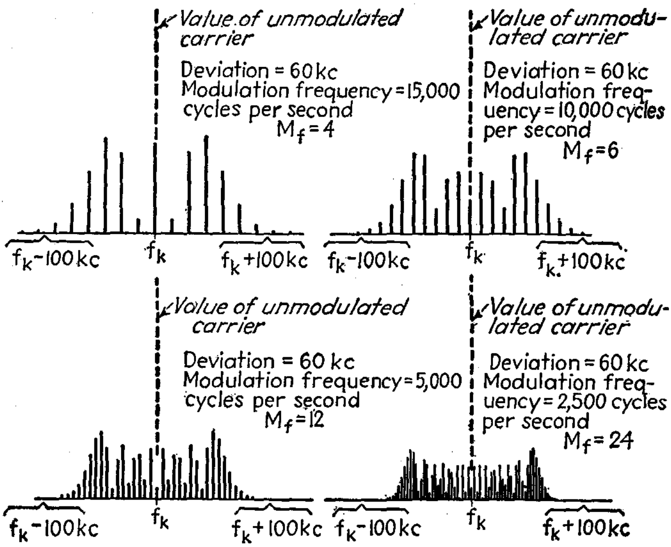


FIG. 9-41.—Amplitudes and distribution of side-frequency components of f-m wave for 60-kc deviation and four values of modulation frequency.

expression is not quite so simple. More complete analyses have been made.[1] In present frequency-modulation broadcasting systems the maximum deviation is 75,000 cps and the highest modulation frequency is 15,000 cps. M_f is therefore 5 and the band width at full deviation varies from about 165,000 to 200,000 cps as the modulation frequency varies in the a-f range. Figure 9-41 shows the manner in which the number, distribution, and amplitude of the carrier and side-frequency components change with modulation frequency for a frequency deviation of 60,000 cps.[2]

It should be noted that the amplitudes of the side-frequency components of a frequency-modulated wave depend both upon the amplitude

[1] BLACK, L. J., and SCOTT, H. J., *Electronics*, September, 1940, p. 30.
[2] EVERITT, W. L., *Trans. Am. Inst. Elec. Eng.*, **59**, 613 (1940).

of the original modulation-frequency excitation [through variation of B in Eq. (9-68)] and upon the modulation frequency, as has just been shown.

A more complete analysis than that which has just been presented shows that the side frequencies produced when the carrier is modulated simultaneously at two or more frequencies are not in general equal to the side frequencies produced when the carrier is modulated individually at these same frequencies. When the total deviation is the same, however, the side frequencies remain within a band of the same approximate width.[1]

9-28. Phase Modulation.—An analysis similar to that made in the preceding section for frequency modulation shows that when the phase angle of the carrier is varied in accordance with the relation

$$\phi = \phi_0(1 + B \sin \omega_m t) \tag{9-74}$$

the resulting phase-modulated wave is of the form[2]

$$y = K \sin (\omega_k t + M_p \sin \omega_m t) \tag{9-75}$$

in which M_p, the *phase modulation index*, is equal to $\phi_0 B$. Equation (9-75) differs from Eq. (9-72) only in that M_p replaces M_f. It may, therefore, be expanded into an equation of the same form as Eq. (9-73). Since M_p is not inversely proportional to f_m as in frequency modulation, however, the number of side frequencies of appreciable amplitude is independent of f_m and so the required band width is proportional to the modulation frequency. Because of the much larger band width required at high modulation frequency, because of the more complicated circuits required for detection, and because of sensitivity of the receiver to microphonic tube vibrations,[3] phase modulation is less satisfactory than frequency modulation in radio broadcasting and has thus far been used only experimentally.

9-29. F-m Circuits.—The most direct method of producing frequency modulation is by variation of oscillator frequency in response to the time variation of the modulation-frequency signal. Since the frequency of most oscillators is controlled by the reactance of a parallel resonant circuit (see Chap. 10), the frequency may be varied by use of one of the circuits, discussed in Sec. 6-39, in which the effective capacitive or inductive reactance is varied by means of the grid voltage of a vacuum tube. Two frequency modulators are shown in Figs. 9-42 and 9-43.[4] (Although one or both tubes in Figs. 9-42 and 9-43 may be multigrid tubes, they are shown as triodes in order to simplify the diagrams.) In the circuit of Fig.

[1] CROSBY, M. G., *RCA Rev.*, **3**, 103 (1938).

[2] RODER, *loc. cit.*

[3] CROSBY, M. G., *Proc. I.R.E.*, **27**, 126 (1939).

[4] TRAVIS, C., *Proc. I.R.E.*, **23**, 1125 (1935). See also M. G. CROSBY, *RCA Rev.*, **5**, 89 (1940).

9-42, tube T_2 and the circuit elements lying to the right of the condenser C' make up a feedback oscillator of a type that will be discussed in Sec. 10-30 and is shown in basic form in Fig. 10-49. The frequency of oscillation of this circuit is determined by the inductance of L and the total capacitance that shunts it. The main tuning capacitance C is shunted by the auxiliary capacitance C' in series with the plate resistance r_p of tube T_1. Variation of the control-grid voltage of T_1 changes the plate resistance and, hence, the effective capacitance of the series combination of r_p and C'. It follows that the frequency of oscillation may be varied by means of the control-grid voltage of T_1. Mathematical analysis shows that, if the oscillator capacitance is varied periodically over a range that is small in comparison with the average capacitance, at a frequency that is very small in comparison with the average frequency of oscillation, the instantaneous frequency varies approximately linearly with capacitance. Hence, if the circuit of the reactance tube T_1 is designed so that the

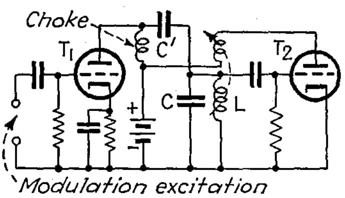


FIG. 9-42.—Frequency modulator.

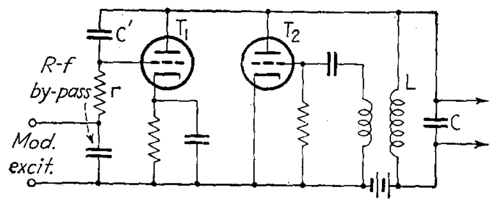


FIG. 9-43.—Frequency modulator.

effective capacitance varies linearly with control-grid voltage, the oscillator frequency will be modulated linearly with respect to modulation excitation impressed upon the control grid of T_1.

A more satisfactory circuit is that of Fig. 9-43. In this circuit T_2 and the circuit elements lying to the right of it constitute the oscillator, the frequency of which is determined by the inductance L and the total capacitance shunting L. T_1, C', and r comprise a reactance-tube circuit discussed in Sec. 6-39 and shown in Fig. 6-50. The effective capacitance of this circuit shunts the main tuning capacitance C. Equation (6-99) shows that the effective capacitance varies with the transconductance of T_1. The oscillator frequency may, therefore, be varied by means of the voltage of one or more grids of T_1. By proper choice of tube, operating voltages, and circuit constants, the frequency can be made to vary linearly with grid voltage. The frequency variation is then proportional to the instantaneous value of modulating voltage impressed upon the grid or grids. A number of modifications of these basic circuits may be made to meet specific requirements.[1]

[1] FREEMAN, R. L., *Electronics*, November, 1936, p. 20; FOSTER, D. E., and SEELEY, S. W., *Proc. I.R.E.*, **25**, 289 (1937); SHEAFFER, C. F., *Proc. I.R.E.*, **28**, 66 (1940); CROSBY, M. G., QST, June, 1940, p. 46; CROSBY, M. G., *RCA Rev.*, **5**, 89 (1940).

An entirely different principle is used in the Armstrong system of modulation.[1] If M_f is less than about 0.5, all but the first two terms of Eq. (9-73) are very small,[2] and Eq. (9-73) may be written in the approximate form

$$y = KJ_0(M_f) \sin \omega_k t + KJ_1(M_f)[\sin (\omega_k + \omega_m)t - \sin (\omega_k - \omega_m)t] \quad (9\text{-}76)$$

The amplitude-modulated wave

$$y = K'(1 + M \sin \omega_m t) \cos \omega_k t \quad (9\text{-}77)$$

may be written in the form

$$y = K' \cos \omega_k t + \tfrac{1}{2}K'M[\sin (\omega_k + \omega_m)t - \sin (\omega_k - \omega_m)t] \quad (9\text{-}78)$$

Equation (9-78) differs in form from Eq. (9-76) only in the 90-degree difference in phase of the carrier-frequency term and in the fact that the side-frequency coefficient of Eq. (9-76) decreases with increase of modulation frequency whereas that of Eq. (9-78) is independent of frequency. $J_1(M_f)$ is very nearly proportional to M_f, and hence inversely proportional to f_m, for values of M_f ranging from zero to 0.5. Hence if a 90-degree phase shift is made in the second term of Eq. (9-78), if the modulation factor M is made inversely proportional to frequency, and if MK' is of the proper value, the wave represented by Eq. (9-78) will be identical with that represented by Eq. (9-76). This is accomplished in the Armstrong system by amplitude-modulating the carrier by means of a balanced modulator [in which the carrier is suppressed from the output (see Sec. 9-5)], and combining the resulting side-frequency output with the unmodulated carrier shifted 90 degrees in phase. M is made inversely proportional to the modulation frequency by passing the modulation-frequency excitation through an "equalizer" that makes its amplitude inversely proportional to the modulation frequency before impressing it upon the modulator. The equalizer usually consists of a resistance-capacitance network.

The maximum value of approximately 0.5 that can be used for M_f directly in the Armstrong system is too small to give the benefits that can be attained by the use of frequency modulation in radio communication (see Sec. 9-31). This difficulty is avoided by performing the modulation at a much lower carrier frequency than that desired for transmission. By means of a series of frequency-multiplying circuits the frequency of the modulated voltage is then multiplied to the desired value. This process also raises the deviation ratio to the desired value.

9-30. Detection of F-m Voltage.—Prior to detecting a frequency-modulated voltage it is necessary to eliminate any amplitude variation

[1] ARMSTRONG, *loc. cit.*

[2] JAFFE, D. L., *Proc. I.R.E.*, **26**, 475 (1938).

that may have been introduced subsequent to modulation. This is accomplished by impressing the voltage upon the input of a *limiter*, which is merely a tuned radio-frequency amplifier operated at low plate voltage. The action of such a circuit may be understood by reference to the plate diagram of Fig. 9-44. Since the d-c resistance of the tuned plate circuit is small, the operating plate voltage is practically equal to the supply voltage. At resonance, however, the a-c load resistance is high. It can be seen from Fig. 9-44 that, no matter how large the grid excitation is made, the plate voltage cannot swing beyond the limits indicated by points *a* and *b*. Hence, beyond the excitation that causes the voltage to swing to these limits, increase of grid excitation has no effect upon the amplitude of the fundamental output voltage. Harmonics are suppressed by the tuned circuit. In practice, the excitation is made large enough and the

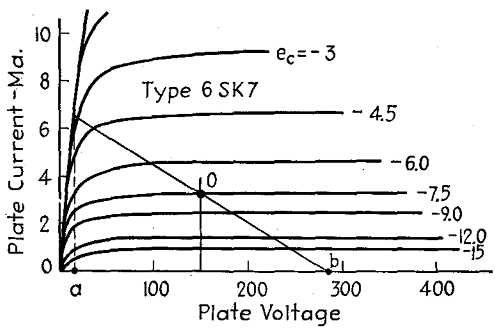


FIG. 9-44.—Limiter plate diagram.

operating voltage sufficiently low so that the output amplitude remains constant in spite of fluctuations of impressed voltage.

The detection of frequency-modulated voltages is accomplished by first causing the changes of frequency to vary the amplitude of the voltage and then detecting the resulting amplitude-modulated wave by means of a diode detector in a manner discussed in Secs. 9-14 to 9-17. The circuit used to produce amplitude modulation as the result of frequency modulation is called a *discriminator*. This term is commonly applied also to the complete detector, including the diode portion.

The simplest type of discriminator is a series resonant circuit tuned slightly off resonance with the carrier.[1] The voltage produced across either the condenser or the inductance varies approximately linearly with frequency over a limited range of frequency, and so variations of frequency produce proportional variations of amplitude. A similar action takes place in tuned-radio-frequency amplifiers tuned slightly off resonance with the carrier frequency. For this reason careful tuning is desir-

[1] ANDREW, V. J., *Proc. I.R.E.*, **20,** 835 (1932); CHAFFEE, J. G., *Proc. I.R.E.*, **23,** 517 (1935); RODER, H., *Proc. I.R.E.*, **25,** 1617 (1937).

able in amplifiers used for frequency-modulated voltages. (Note, however, that the limiter removes amplitude variations resulting from this cause.)

A more common discriminator circuit is shown in Fig. 9-45. This circuit was originally designed for use in automatically tuned radio receivers to produce a direct voltage as the result of a change of frequency.[1] The action of the circuit is as follows: The voltage E_1 impressed upon the circuit of the diode T_1 is the vector sum of E' and E'', whereas the voltage E_2 impressed upon the circuit of T_2 is the vector sum of E'

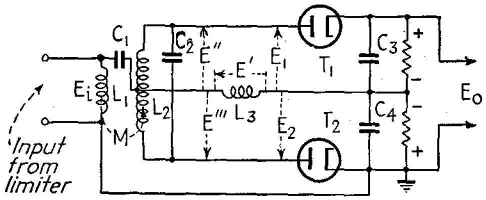


FIG. 9-45.—Discriminator circuit.

and E'''. As explained in Sec. 9-14, if the reactance of C_3 and C_4 at carrier frequency is low in comparison with the shunting resistances, then C_3 and C_4 charge in the indicated polarities to voltages nearly equal to the crest voltages E_1 and E_2 impressed in the respective circuits. The output voltage E_o is the sum of the condenser voltages.

Since the reactance of C_4 is negligible, the series combination of C_1 and L_3 is in parallel with L_1 and the voltage E' across L_3 is either in phase with or 180 degrees out of phase with the primary voltage E_i. It will be assumed that the reactance of L_3 exceeds that of C_1, so that E' is in phase with E_i. By virtue of circuit symmetry, the voltages E'' and E''' are each equal in magnitude to half the voltage E_C across C_2. E'' is of the same sign as E_C; E''' is opposite in sign. The voltage across C_2 can be determined with the aid of the equivalent circuit of Fig. 9-46, in which L_2 and r_2 are the self-inductance and resistance of the transformer secondary and M is the mutual inductance. The induced voltage E is 90 degrees out of phase with the primary current. As the primary resistance r_1 is small, E_i is also practically 90 degrees out of phase with the primary current, and so E is in phase with or 180 degrees out of phase with E_i. It will be assumed to be opposite in phase to E_i. Figure 9-47a shows the vector diagram of the secondary of the equivalent circuit at resonance.[2] E_C lags E by 90 degrees. The vector

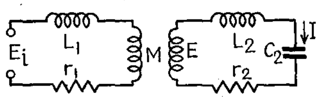


FIG. 9-46.—Equivalent circuit for the transformer of the discriminator of Fig. 9-45.

[1] TRAVIS, *loc. cit.;* ARMSTRONG, *op. cit.*, p. 699; CROSBY, M. G., *Proc. I.T.E.*, **24**, 898 (1936); FOSTER and SEELEY, *loc. cit.;* RODER, H., *Proc. I.R.E.*, **26**, 590 (1938).

[2] The voltage E_r is actually very small in comparison with E_C. It is exaggerated in Fig. 9-47 for the sake of clarity.

diagrams of Fig. 9-48*a*, showing the vector sum of E' and E'' and of E' and E''' indicate that the voltages E_1 and E_2 impressed in the two diode circuits are equal. The voltages across C_3 and C_4 are, therefore, also equal in magnitude. Because they are of opposite polarity, the output voltage E_o is zero.

Figure 9-47*b* shows the vector diagram for the equivalent secondary circuit at a frequency slightly above resonance.[1] I, E_C, and E_L are all

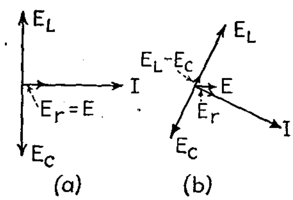


FIG. 9-47.—Vector diagrams for the secondary of the equivalent circuit of Fig. 9-46.

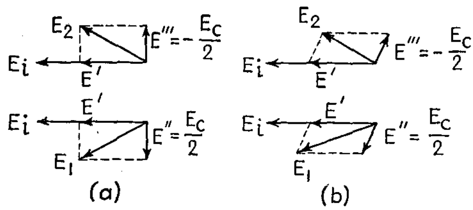


FIG. 9-48.—Vector diagrams for the circuit of Fig. 9-45.

smaller than at resonance, but the reduction of E_C is greater than that of E_L, and so E_C lags E by more than 90 degrees. Figure 9-48*b* shows that the magnitude of E_1 now exceeds that of E_2. Therefore, the voltage across C_3 exceeds that across C_4, and E_o has a positive value. Below resonance, on the other hand, E_2 exceeds E_1 in magnitude and E_o is negative. A typical curve of E_o as a function of frequency change from the resonance value f_0 is shown in Fig. 9-49. The linearity of this curve in the vicinity of resonance makes possible the conversion of the frequency-modulated input voltage into the modulation-frequency output voltage without distortion. If the circuit is tuned to the carrier frequency, the instantaneous output voltage is proportional to the instantaneous value of modulation-frequency voltage used to modulate the carrier. The output voltage is therefore of the same wave form as the original modulation voltage.

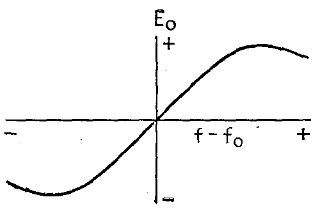


FIG. 9-49.—Detection characteristic for the discriminator of Fig. 9-45.

9-31. Interference Suppression.—The most important advantages of frequency modulation over amplitude modulation in radio communication result from the greater suppression of undesired carriers and noise. An analysis of noise suppression in amplitude-modulation and frequency-modulation systems can be made by making use of the fact, proved in Sec. 9-18, that the sum of two sinusoidal waves can be considered to be a wave modulated simultaneously in amplitude and in phase. One wave is

[1] In the vicinity of resonance, the vector $E_L - E_C$ is actually much smaller than shown.

assumed to be the desired carrier, and the other the interfering carrier or noise.

If $E_k \sin \omega_k t$ represents the desired carrier voltage impressed upon the input of the receiver, and $E_i \sin \omega_i t$ an interfering carrier voltage, both being assumed for the present to be unmodulated, the resultant voltage impressed upon the receiver is

$$e = E_k \sin \omega_k t + E_i \sin \omega_i t \tag{9-79}$$

As pointed out in Sec. 9-18, Eq. (9-79) may be transformed into

$$e = E \sin (\omega_k t - \phi) \tag{9-80}$$

in which

$$E = E_k \sqrt{1 + 2h \cos \omega_{ki} t + h^2} \tag{9-81}$$

$$h = \frac{E_i}{E_k} \tag{9-82}$$

$$\omega_{ki} = \omega_k - \omega_i \tag{9-83}$$

$$\phi = \tan^{-1} \frac{h \sin \omega_{ki} t}{1 + h \cos \omega_{ki} t} \tag{9-84}$$

Since both E and ϕ are functions of time, the impressed voltage is equivalent to a voltage modulated in both amplitude and phase. In Fig. 9-27, which shows the wave form of the sum of two sinusoidal voltages of equal amplitude and 5–4 frequency ratio, the amplitude modulation is evident at a glance. The phase modulation is indicated by the unequal spacing of the points at which the curve crosses the time axis.

9-32. Interference-to-signal Ratio in A-m Systems.—In a properly tuned amplitude-modulation (a-m) receiver, the detector responds only to amplitude modulation and not to frequency modulation. Figure 9-50 shows the manner in which E, the amplitude of the resultant input voltage, varies with time at constant amplitude of desired carrier voltage and several values of amplitude of interfering carrier voltage. The curve corresponding to $E_i = E_k$ shows that when the two carrier voltages have equal amplitude the desired carrier is fully modulated, at the difference frequency f_{ki}, by the undesired carrier. As the amplitude E_i of the interfering carrier is reduced, the degree of modulation is also reduced, and is always equal to E_i/E_k,[1] the ratio of the amplitudes of the two carriers. The ratio of the amplitude of the interference output of the detector to the desired audio-frequency output which would be obtained without the interfering carrier is equal to the ratio of the modulation factor E_i/E_k

[1] Examination of Eqs. (9-51) and (9-52) shows that the modulation factor corresponding to the *fundamental* difference-frequency component of E is less than E_i/E_k, but approaches this value as h is reduced. It is equal to approximately $0.67E_i/E_k$ when h is unity and to approximately $0.913E_i/E_k$ when h is $\frac{1}{2}$.

corresponding to the interfering carrier to the modulation factor M corresponding to the desired a-f signal. This is the interference-to-signal ratio:

$$\frac{N}{S} = \frac{E_i}{ME_k} \tag{9-85}$$

If the desired carrier is fully modulated by the desired a-f signal, this ratio reduces to E_i/E_k.

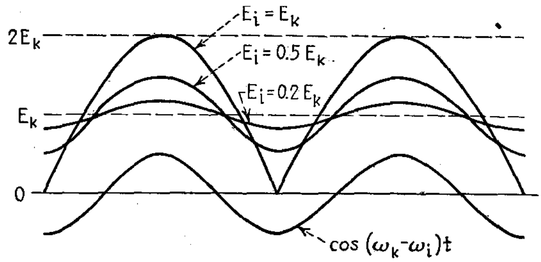


FIG. 9-50.—Wave form of the envelope of the resultant of a desired and an interfering unmodulated carrier at several ratios of carrier amplitudes.

9-33. Interference-to-signal Ratio in F-m Systems.—In a frequency-modulation receiver, the purpose of the limiter is to remove amplitude modulation from the discriminator input. Since the limiter cannot function completely when the amplitude modulation of the impressed voltage approaches 100 per cent, it can be seen from Fig. 9-50 that amplitude modulation of the desired carrier by an interfering carrier can be completely removed only when the amplitude of the interfering carrier is somewhat less than that of the desired carrier. When the amplitude modulation is completely removed, E of Eq. (9-80) is constant at the input to the discriminator. Analysis of Eq. (9-84) indicates that ϕ varies periodically with time and that the fundamental frequency at which ϕ varies is the difference between the frequencies of the two carriers. The manner in which ϕ varies during one difference-frequency cycle for several values of E_i/E_k is shown in Fig. 9-28a.

The symmetry of the curves of Fig. 9-28a is such that ϕ may be expressed as an infinite series of sine terms:

$$\phi = \phi_1 \sin \omega_{ki}t + \phi_2 \sin 2\omega_{ki}t + \phi_3 \sin 3\omega_{ki}t + \cdots \tag{9-86}$$

Equation (9-80) may therefore be written in the form

$$e = E_o \sin\left(\omega_k - \sum_{n=1}^{\infty} \phi_n \sin n\omega_{ki}t\right) \tag{9-87}$$

in which E_o is the constant amplitude of the voltage output of the limiter. Equation (9-87) is that of a phase-modulated wave modulated simultaneously at frequencies equal to the difference between the frequencies of the two carriers and at harmonics of the difference frequency. When E_i/E_k is less than 0.5, all terms of the summation of Eq. (9-87) except the first are sufficiently small so that they may be neglected in an approximate analysis. Analysis of Eq. (9-84) shows, furthermore, that ϕ_1 is approximately equal to E_i/E_k radians when E_i/E_k is less than 0.5.[1] Equation (9-87) then becomes

$$e = E_o \sin\left(\omega_k - \frac{E_i}{E_k} \sin \omega_{ki} t\right) \tag{9-88}$$

or

$$e = E_o \sin\left(\omega_k - \frac{f_d'}{f_{ki}} \sin \omega_{ki} t\right) \tag{9-89}$$

in which

$$f_d' = \frac{f_{ki} E_i}{E_k} \tag{9-90}$$

Comparison with Eq. (9-72*a*) shows that Eq. (9-89) is that of a frequency-modulated voltage of deviation f_d'. Since the amplitude of the output of the limiter has the constant value E_o, the output of the limiter when the carrier is modulated by the desired audio-frequency signal is

$$e_s = E_o \sin\left(\omega_k - \frac{f_d}{f_m} \sin \omega_m t\right) \tag{9-91}$$

The action of a properly designed discriminator is such that the output voltage is proportional to the deviation of the impressed frequency-modulated voltage. Therefore, the ratio of the undesired difference-frequency output to the desired a-f output of the discriminator is

$$\frac{N}{S} = \frac{f_d'}{f_d} = \frac{f_{ki}}{f_d} \times \frac{E_i}{E_k} \tag{9-92}$$

Comparison of Eqs. (9-85) and (9-92) shows that the interference-to-signal ratio is lower in the frequency-modulation system by the factor f_{ki}/f_dM, which equals f_{ki}/f_d when the amplitude modulation factor is unity. Equation (9-92) indicates that the strength of the heterodyne whistle caused by the interfering carrier in a frequency-modulated receiver is proportional to the difference frequency and approaches zero

[1] This may also be seen from an examination of Fig. 9-28*a*. If the second and higher terms of Eq. (9-86) are negligible, ϕ_1 is approximately equal to the maximum value of the curve of Fig. 9-28*a* corresponding to the given value of E_i/E_k. Inspection of Fig. 9-28*a* shows that values of the maxima of the curves are approximately E_i/E_k radians for values of E_i/E_k equal to and less than 0.5.

as the carrier frequencies approach equality. It is loudest when f_{ki} has its greatest value, which is the highest frequency f_h passed by the a-f amplifier that follows the discriminator. For a deviation of 75,000 cps and a 15,000-cycle a-f band width, Eq. (9-92) shows the maximum interference-to-signal ratio to be

$$\text{Max.}\ \frac{N}{S} = \frac{15{,}000}{75{,}000}\frac{E_i}{E_k} = 0.2\,\frac{E_i}{E_k} \tag{9-93}$$

Hence, under the condition least favorable to the frequency-modulation system, *i.e.*, when the interfering carrier frequency differs by 15,000 cps from the desired carrier frequency, the suppression of the undesired carrier is five times as great as in the amplitude-modulation system. Greater improvement results if the a-f pass band is narrower.

9-34. Use of Preemphasis.—Interference can be greatly reduced by the use of *preemphasis* of high frequencies. In preemphasis of high frequencies the amplitude of the modulation voltage impressed upon the frequency modulator of the transmitter is made to increase linearly with frequency above 1500 cps, and the response of the discriminator is made to vary inversely with frequency above 1500 cps. As far as the desired signal is concerned, the two effects offset each other, giving a modulation-frequency output that is independent of frequency. The interference output of the discriminator, however, is inversely proportional to f_{ki} above 1500 cps. Since the interference output is also proportional to f_d', Eq. (9-90) shows that the amplitude of the heterodyne whistle caused by an interfering carrier increases linearly with f_{ki} below 1500 cps and remains constant at the 1500-cycle value above 1500 cps. The value of the noise-to-signal ratio N/S then does not exceed the value obtained when f_{ki} is 1500 cps, *i.e.*, $0.02E_i/E_k$ in a system using 75,000-cycle deviation and 15,000-cycle audio-frequency band width. Hence in such a system the use of preemphasis increases the suppression of an undesired carrier by a factor of 10, making the suppression fifty times as great as in an amplitude-modulation system. The improvement resulting from preemphasis decreases with the a-f band width.

9-35. Interference of Modulated Carriers.—When one or both carriers are modulated, the instantaneous frequency difference f_{ki} varies. In a frequency-modulation system without preemphasis, the difference-frequency interference output has its maximum amplitude only when the difference frequency is equal to the maximum frequency passed by the audio-frequency amplifier. When preemphasis is used, the beat-frequency amplitude is constant for values of difference frequency above 1500 cps. In either case, the maximum value is obtained during only a portion of the modulation-frequency cycle. Furthermore, if both waves are modulated, the difference frequency is above the a-f amplifier range during a part of the time.[1] Hence the average interference amplitude is less than when the carriers are unmodulated, and, in a system using 75,000-cycle deviation and 15,000-cycle a-f band width, the noise-to-signal ratio is less than $0.2E_i/E_k$ without preemphasis and less than $0.02E_i/E_k$ when preemphasis is used.

In an amplitude-modulation system the amplitude of the interference output is independent of the difference frequency, and so the noise-to-signal ratio is unaffected by modulation of the carriers.

9-36. Static Interference.—Although static is not ordinarily a periodic disturbance, Fourier theory indicates that it may be considered to consist of a continuous spectrum of frequencies. The second term of Eq. (9-79) may represent any frequency component in this spectrum. Equation (9-80) then represents the instantaneous voltage impressed upon the receiver as the result of the desired carrier and the given component of the frequency spectrum.

In an amplitude-modulation receiver, the resulting output is equal to $sM'E_k$, where s is the receiver sensitivity and M' is the modulation factor corresponding to modulation of the desired carrier by the given static component. It has already been pointed out that Fig. 9-50 shows M' to be equal to E_i/E_k. The receiver output corresponding to this component is therefore sE_i. Since the output frequency f_{ki} is the difference between the frequency of the given component and that of the carrier, only those static components whose difference in frequency from the carrier lies within the audio-frequency pass band of the receiver contribute to the noise output. The total output caused by static is the sum of the outputs resulting from these components.

$$N = \int_{f_k - f_h}^{f_k + f_h} sE_i \, df \qquad (9\text{-}94)$$

in which f_h is the upper a-f limit of the receiver and E_i is the amplitude of a static component of frequency f. If E_i and s are assumed to be independent of f, Eq. (9-94) reduces to

$$N = 2sE_i f_h \qquad (9\text{-}95)$$

in which the value of E_i depends upon the form and amplitude of the noise. The desired signal output of the receiver is

$$S = sME_k \qquad (9\text{-}96)$$

[1] When the unmodulated carrier frequencies are equal, the maximum difference frequency is equal to twice the maximum frequency passed by the a-f amplifier and is obtained at an instant when one modulated carrier frequency is a maximum and the other a minimum.

Hence the noise-to-signal ratio is

$$\frac{N}{S} = \frac{2E_i f_h}{ME_k} \tag{9-97}$$

In a frequency-modulation receiver, the output resulting from a given noise component is

$$N_f = E_k s f_d' = E_k s f_{ki} \phi_1 \cong \frac{E_k s f_{ki} E_i}{E_k} = s(f - f_k)E_i \tag{9-98}$$

The total noise output from a receiver when preemphasis is not used is

$$N \cong \int_{f_k}^{f_k+f_h} E_i s(f - f_k)\, df + \int_{f_k-f_h}^{f_k} E_i s(f_k - f)\, df = E_i s f_h^2 \tag{9-99}$$

The desired a-f signal output is

$$S = E_k s f_d \tag{9-100}$$

and so

$$\frac{N}{S} = \frac{E_i}{E_k} \times \frac{f_h^2}{f_d} \tag{9-101}$$

Comparison of Eqs. (9-97) and (9-101) shows that for the same a-f band width the noise-to-signal ratio is lower in the frequency-modulation receiver in the ratio $Mf_h/2f_d$ when high-frequency preemphasis is not used. For 100 per cent amplitude modulation, a 5000-cycle a-f band width, and 75,000-cycle deviation, this improvement ratio is 1:30. For a 5000-cycle a-f band width in the amplitude-modulated receiver and a 15,000-cycle a-f band width in the frequency-modulated receiver, the improvement ratio is 1:10.

When preemphasis is used, the response of the discriminator varies inversely with frequency for difference frequencies above 1500 cps, and so N_f has the constant value $1500sE_i$ above 1500 cps and decreases below 1500 cps. If it is assumed that this value also holds below 1500 cps, the total noise output is given by the relation

$$N = 1500s \int_{f_k-f_h}^{f_k-f_h} df = 3000sf_h \tag{9-102}$$

Therefore

$$\frac{N}{S} \cong 3000 \frac{E_i}{E_k} \times \frac{f_h}{f_d} \tag{9-103}$$

Comparison of Eq. (9-103) with Eq. (9-97) shows that the noise-to-signal ratio is smaller in the frequency-modulation system than in the amplitude-modulation system in the approximate ratio $1500M/f_d$, which is equal to $\frac{1}{50}$ when M is unity and f_d equals 75,000 cps. This clearly predicts the large decrease in static interference observed in frequency-modulation receivers over amplitude-modulation receivers. Other types of noise interference may be analyzed in a like manner.

A similar analysis shows that the noise output is also smaller in the frequency-modulation than in the amplitude-modulation system when the receiver is not tuned to a carrier. This is done by letting the two voltages in Eq. (9-79) be those corresponding to any two components in the noise spectrum, one of which is, in effect, considered to be modulated by the other.

9-37. Interference Caused by Carrier or Static of High Amplitude.—In the analyses of Secs. 9-31 to 9-36 it was assumed that the amplitude of the interference input to the receiver is less than that of the desired carrier. The analyses show that the interference output is then lower in a frequency-modulation than in an amplitude-modulation system, and that the superiority of the former increases with the deviation. To analyze the behavior of the two systems when the interference input exceeds the desired carrier input, it is merely necessary to allow the first term of Eq. (9-79) to represent the interference voltage and the second term the carrier. Such an analysis indicates that the interference suppression is then greater in the amplitude-modulation than in the frequency-modulation system, and that the interference increases with the deviation in the latter. At distances remote from the transmitter, therefore, amplitude modulation may be preferable to frequency modulation, and low deviation may be preferable to high deviation. This prediction is verified by experimental observations.

9-38. Advantages and Disadvantages of Frequency Modulation in Radio Communication.—The advantages of frequency modulation over amplitude modulation in radio communication may be summarized as follows:

1. Greatly reduced interference from stations operating in the same and adjacent channels.[1] If the amplitude of one modulated carrier is greater than twice that of another of the same frequency, the interference of the weaker is negligible. For a 3 to 1 ratio of signal strengths, a maximum deviation of 75 kc, and an audio-frequency band width of 15 kc, the level of the interference is 40 db below that of the desired a-f output.

2. Greatly reduced interference from static and other noise. Unless the noise input to the receiver exceeds the desired carrier input, the ratio of the a-f signal output to the noise output is greater in a frequency-modulation system than in an amplitude-modulation system and increases with the deviation ratio.[1]

[1] ARMSTRONG, *loc. cit.*; CROSBY, M. G., *Proc. I.R.E.*, **25**, 472 (1937); RODER, H., *Electronics*, May, 1936, p. 22; WEIR, I. R., *Gen. Elec. Rev.*, **42**, 188 and 270 (1939); LEVY, M. L., *Electronics*, June, 1940, p. 26; GUY, R. F., and MORRIS, R. M., *RCA Rev.*, **5**, 190 (1940); GOLDMAN, S., *Electronics*, August, 1941, p. 37; REICH, H. J., *Communications*, August, 1942.

3. Because of reduced noise interference, a wider a-f band can be used in a frequency-modulation system without excessive noise. Greater tone fidelity may thus be obtained.

4. Increased transmitter efficiency. In frequency modulation, increase of side-frequency amplitudes with modulation is accompanied by reduction of amplitude of the carrier component and the total power furnished by the transmitter remains constant. The transmitter may, therefore, be operated in such a manner that its efficiency is high at all times.[1] This is in contrast with amplitude modulation, in which the amplitude of the carrier-frequency component of the modulated wave is independent of modulation factor and the total power furnished by the transmitter increases with modulation.

A difficulty encountered in the use of frequency modulation in broadcasting is that of maintaining the average carrier frequency constant without preventing the desired modulation. This is accomplished in the Armstrong system by the use of crystal control (see Sec. 10-47). In the reactance-tube system the frequency is kept constant by use of electronic automatic frequency control and by more complicated methods.[2]

In order to obtain the advantage of freedom from interference the deviation, and hence the band width, must be large. The great demand for broadcasting and other communication channels makes it infeasible to use wide bands except at ultrahigh frequencies. The range from 42 to 50 Mc has been set aside for frequency-modulated waves and 200 kc has been designated as the channel width. This allows a maximum deviation of approximately 75 kc for an a-f modulation range up to 15 kc. Although low level of natural static constitutes an advantage of ultrahigh-frequency transmission, it also has a number of disadvantages. These are

1. The transmission range is small. Because ultrahigh-frequency waves are not reflected by ionized layers of the earth's upper atmosphere, they move in straight paths and their range is theoretically limited to the distance of the horizon. It has been found in practice that consistent reception may be achieved over about three times this range. By locating the transmitting antennas on the tops of mountains and tall buildings, ranges of about 150 miles are now attained.

2. Buildings, hills, and other objects cast shadows and thus prevent satisfactory reception in some areas.

3. Automobile ignition systems and other electrical equipment set up serious local interference.

The Yankee network of the New England region overcomes the difficulty of limited range by the use of sufficiently close transmitter

[1] RODER, H., *Proc. I.R.E.*, **19**, 2145 (1931); CHAFFEE, J. G., *Proc. I.R.E.*, **23**, 517 (1935); EVERITT, *loc. cit.*

[2] MORRISON, J. F., *Communications*, August, 1940, p. 12; *Proc. I.R.E.*, **28**, 444 (1940).

spacing to give complete coverage. The programs that originate at one transmitter are rebroadcast at the same frequency by other transmitters of the chain without being changed to audio frequency in the process. In this manner excellent quality is maintained.[1]

Problems

9-1. By means of the third term of the series expansion for plate current, determine, for the van der Bijl modulator, the amplitudes of the components of plate current whose frequencies are the carrier frequency plus and minus twice the modulation frequency. By comparing these amplitudes with that of the side frequencies, show that distortion of the side frequencies increases with percentage modulation.

9-2. Show that at constant modulation factor the amplitude of the side-frequency components produced by a square-law modulator is proportional to the unmodulated carrier amplitude.

9-3. Determine the frequencies and the relative amplitudes of the various components of the outputs of the circuits of Figs. 9-5 and 9-6 when voltages $E_1 \cos \omega_1 t$ and $E_2 \cos \omega_2 t$ are applied to the circuits. Consider three terms of Eq. (9-13) and two terms of Eq. (9-16).

9-4. Find the power supplied to a resistance r, by a generator whose terminal voltage is of the form of Eq. (9-19), and determine what portion of the power is associated with the carrier component and with each side frequency.

9-5. Derive Eq. (9-24).

9-6. In a diode detector that does not use a condenser, if the load resistance is high enough so that the dynamic tube characteristic is essentially linear and if the ratio of the carrier frequency to the modulation frequency is high, the current that flows during the positive half of a carrier cycle may be considered to be sinusoidal in form.

a. Write an expression for the instantaneous plate current resulting from the application of an amplitude-modulated voltage to such a detector.

b. Derive Eq. (9-42).

9-7. Show that the equation of a rectification characteristic for an ideal linear diode detector using a condenser is (see Sec. 9-23)

$$I_{bk} = \frac{E}{\pi r_p}\left[\cos \omega_k t_o + \left(\omega_k t_o - \frac{\pi}{2}\right)\sin \omega_k t_o\right]$$

where $\sin \omega_k t_o = E_b/E$.

9-8. *a.* From the detection characteristics given on page 684, determine the amplitudes of the fundamental and second-harmonic components of the modulation voltage output of a type 6H6 diode used with a load having d-c and modulation-frequency resistance of 250,000 ohms and negligible carrier-frequency impedance. The unmodulated input carrier amplitude is 15 volts r.m.s., and the modulation factor is 80 per cent.

b. Repeat for 100 per cent modulation.

9-9. Repeat Prob. 9-8 for a load having a d-c resistance of 500,000 ohms and an a-c resistance of 300,000 ohms at modulation frequency.

9-10. (Suggested laboratory experiment.)

a. Using the circuit of Fig. 9-51, adjust the condenser C to resonance with the 5000-cycle impressed voltage. When the circuit is properly tuned, the deflection of the oscilloscope is a maximum, and the pattern is sinusoidal.

[1] ARMSTRONG, E. H., *Elec. Eng.*, **59**, 485 (1940).

b. Using the oscilloscope deflection as an indication of the output voltage, or with the aid of a vacuum-tube voltmeter, plot a curve of alternating output voltage as a function of biasing voltage E_b.

c. Using the modulation characteristic of part *b*, predict the shape of the envelope of the modulated wave that will be produced if 60-cycle voltage is impressed in series with the 5000-cycle voltage. Use several values of bias and of 60-cycle modulating voltage.

d. Apply to the circuit the values of bias and modulating voltage assumed in part *c* and compare the wave forms of the output voltage with those predicted in part *c*.

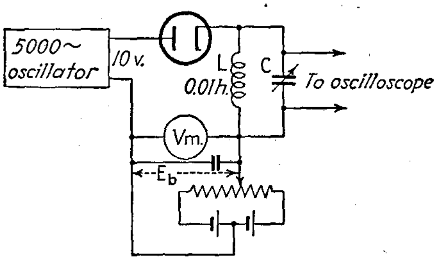


FIG. 9-51.—Circuit for determining the modulation characteristic of a diode modulator.

e. Adjust the bias and exciting voltages so as to give a modulated voltage of good wave form. Impress the output voltage upon a diode detector and observe the wave form of the output of the detector for various combinations of detector load resistance and capacitance.

Supplementary Bibliography

Amplitude Modulation and General Theory of Modulation

CARSON, J. R.: *Proc. I.R.E.*, **9**, 243 (1921); **10**, 57 (1922).
KING, R. W.: *Bell System Tech. J.*, **2** (No. 4), 31 (1923).
BOWN, R., MARTIN, D. K., and POTTER, R. K.: *Proc. I.R.E.*, **14**, 47 (1926).
PETERSON, E., and KEITH, C. R.: *Bell System Tech. J.*, **7**, 106 (1928).
JOLLIFFE, C. B.: *Proc. I.R.E.*, **17**, 660 (1929).
HECHT, N. F. S.: *Wireless Eng.*, **8**, 471 (1931).
ROBINSON, H. A.: *Proc. I.R.E.*, **9**, 131 (1932).
BULL, C. S.: *Wireless Eng.*, **10**, 83 (1933).
BENNETT, W. R.: *Bell System Tech. J.*, **12**, 228 (1933).
HAZEL, H., and RAMSEY, R. R.: *Phys. Rev.*, **44**, 321 (1933).
TYNAN, A. G.: *Proc. I.R.E.*, **21**, 1203 (1933).
GAUDERNACK, L. F.: *Proc. I.R.E.*, **14**, 819 (1934); *Wireless Eng.*, **11**, 293, 362 (1934).
BELL, D. A.: *Wireless Eng.*, **12**, 535 (1935).
BARROW, W. L.: *Proc. I.R.E.*, **22**, 964 (1934).
LAPORT, E. A., *RCA Rev.*, **1**, 26 (1937).
FRIEND, A. W.: *Proc. I.R.E.*, **26**, 786 (1938).
MOULLIN, E. B.: *Wireless Eng.*, **15**, 371 (1938).
TERMAN, F. E., and WOODYARD, J. R.: *Proc. I.R.E.*, **26**, 926 (1938).
ECKERSLEY, P. P.: *Proc. I.R.E.*, **26**, 1041 (1938).
HATHAWAY, J. L.: *Electronics*, October, 1939, p. 28.
BRAINERD, J. G.: *Proc. I.R.E.*, **28**, 136 (1940).
GRAMMER, G.: *QST*, February, 1940, p. 14.
TERMAN, F. E., and BUSS, R. R.: *Proc. I.R.E.*, **29**, 104 (1941).

Frequency Modulation

RODER, H.: *Telefunken Z.*, **53,** 48 (1929).
ECKERSLEY, T. L.: *Wireless Eng.*, **7,** 482 (1930).
RODER, H.: *Proc. I.R.E.*, **20,** 887 (1932).
BARROW, W. L.: *Proc. I.R.E.*, **21,** 1182 (1933).
EASTMAN, A. V., and SCOTT, E. D.: *Proc. I.R.E.*, **22,** 878 (1934).
RODER, H.: *Electronics*, May, 1937, p. 22.
SCOTT, E. D., and WOODYARD, J. R.: *Univ. Wash. Eng. Expt. Sta. Bull.* 68.
NOBLE, D. E.: *QST*, August, 1939, p. 11.
Editorial, *Electronics*, November, 1939, p. 20.
METSCHL, E. C.: *E.T.Z.*, November, 1939, p. 1357.
YOCUM, C. H.: *Communications*, November, 1939, p. 5; December, 1939, p. 14.
TYLER, G. W.: *R.M.A. Eng.*, **4,** 16 (1939).
SHELBY, R. E.: *Electronics*, February, 1940, p. 14.
CHAFFEE, J. G.: *Bell Lab. Record*, **18,** 177 (1940).
HUGHES, L. E. C.: *Electrician*, March 8, 1940, p. 188.
CROSBY, M. G.: *RCA Rev.*, **4,** 473 (1940).
CARSON, J. R.: *Wireless Eng.*, **17,** 477 (1940).
GUNTHER, F. A.: *Communications*, April, 1940, p. 11.
SHEA, R. F.: *Communications*, June, 1940, p. 17.
SCOTT, H. J.: *Communications*, August, 1940, p. 10.
PIERACCI, R. J.: *Proc. I.R.E.*, **28,** 374 (1940).
BROWNING, G. H.: *QST*, September, 1940, p. 19.
LAWSON, D. I.: *Wireless Eng.*, **17,** 388 (1940).
DAVID, W. R.: *Communications*, October, 1940, p. 8.
SABAROFF, S.: *Communications*, October, 1940, p. 11.
WORCESTER, J. A., JR.: *R.M.A. Tech. Bull.* **2,** Nov. 12, 1940.
WHEELER, H. A.: *Proc. I.R.E.*, **28,** 537 (1940).
LANDON, V. D.: *Electronics*, February, 1941, p. 26.
WIESS, W.: *Communications*, March, 1941, p. 16.

Additional references on the subject of frequency modulation are contained in a bibliography compiled by Miss Elizabeth Kelsey, Engineer, Zenith Radio Corp., Chicago, Ill.

Square-law Detection and General Theory of Detection

BREIT, G.: *Phys. Rev.*, **16,** 387 (1920).
SMITH, L. P.: *Proc. I.R.E.*, **14,** 649 (1926).
NELSON, J. R.: *Proc. I.R.E.*, **16,** 822 (1928).
TERMAN, F. E., and DYSART, B.: *Proc. I.R.E.*, **17,** 830 (1929).
CHAFFEE, E. L.: *Electronics*, May, 1931, p. 641.
JACKSON, W.: *Wireless Eng.*, **8,** 424 (1931).
TURNER, P. K.: *Wireless Eng.*, **9,** 384 (1932).
WHITE, E. C. L.: *Wireless Eng.*, **9,** 618 (1932).
FISHER, C. B.: *Wireless Eng.*, **10,** 548 (1933).
STRUTT, M. J. O.: *Proc. I.R.E.*, **22,** 981 (1934) (with bibliography of 89 items through 1934); **23,** 945 (1935).

Linear Detection

ROBERTS, W. VAN B.: *Proc. I.R.E.*, **15,** 793 (1927).
BALLANTINE, S.: *Proc. I.R.E.*, **16,** 593 (1928).

LANDALE, S. E. A.: *Proc. Cambridge Phil. Soc.*, **25**, 355, 482 (1929).
COCKING, W. T.: *Wireless World*, **26**, 479 (1930).
TERMAN, F. E.: *Electronics*, November, 1930, p. 386.
GREENWOOD, W.: *Wireless Eng.*, **8**, 648 (1931).
COLEBROOK, F. M.: *Wireless World*, **28**, 625 (1931).
BIEDERMANN, E. A.: *Wireless Eng.*, **10**, 123 (1933).
ROBERTS, F., and WILLIAMS, F. C.: *J. Inst. Elec. Eng.* (*London*), **75**, 379 (1934).
MARIQUE, J.: *Wireless Eng.*, **12**, 17 (1935).
BENNON, S.: *Proc. I.R.E.*, **25**, 1565 (1937).
WHEELER, H. A.: *Proc. I.R.E.*, **26**, 745 (1938).
RODER, H.: *Communications*, February, 1938, p. 15 (with bibliography of 22 items on suppression of weak by strong signal).
MOULLIN, E. B.: *Proc. of the Wireless Section, I.E.E.* (*London*), March, 1938.
LAPORT, E. A.: *RCA Rev.*, **3**, 121 (1938).
WILLIAMS, F. C., and FAIRWEATHER, A.: *Wireless Eng.*, **16**, 330 (1939).

CHAPTER 10

TRIGGER CIRCUITS, PULSE GENERATORS, AND OSCILLATORS

Among the important applications of electron tubes is their use in the generation of alternating voltage and current and of random or periodic voltage or current pulses. A closely allied application is made in vacuum-tube switching circuits, usually called *trigger* or *gate* circuits. Since certain types of pulse generators and oscillators are based upon trigger circuits, a discussion of trigger circuits serves as a logical introduction to the theory of pulse generators and oscillators.

10-1. Trigger Circuits.—A trigger circuit is one which, for fixed values of supply voltages and circuit parameters, has two states of

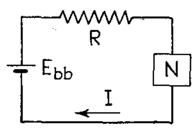


FIG. 10-1.—Basic trigger circuit.

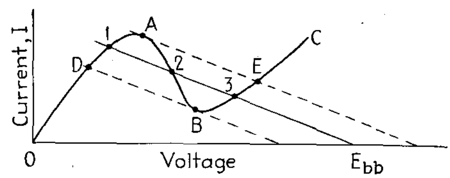


FIG. 10-2.—Current-voltage diagram for the circuit of Fig. 10-1.

equilibrium, characterized by two stable sets of circuit currents. The value of such circuits arises from the fact that they can be made to "trigger" abruptly from one state of equilibrium to the other by means of small impressed controlling voltages or small changes in circuit parameters.

Triggering may take place in a circuit of the form of Fig. 10-1 if the current-voltage characteristic of the element N has a negative slope over a portion of the operating range. It is evident from the figure that the voltage E across the element N must be related to the current I through the element by the following equation

$$E = E_{bb} - IR \tag{10-1}$$

Equation (10-1) is that of the straight line MN through the point on the voltage axis corresponding to the supply voltage E_{bb}, having a negative slope in amperes per volt equal to the reciprocal of the resistance R in series with the element. Since all corresponding values of current and voltage through the element N are also specified by the characteristic curve of the element, it follows that the only values of current that can obtain with the given supply voltage and resistance are those corresponding to the intersection of the resistance line MN with the characteristic curve. It can be seen from Fig. 10-2 that, if the characteristic curve has a portion whose slope is negative, the resistance line may intersect the characteristic curve in three points, 1, 2, and 3. This fact indicates that there are three possible equilibrium values of current.

Analysis of Fig. 10-2 shows that 2 is a point of unstable equilibrium, whereas 1 and 3 are points of stable equilibrium. If the current and voltage have values corresponding to point 2, any small increase of current due to any cause is accompanied by a decrease of voltage across the element. More voltage is thus made available to send current through the resistance, and so the current rises farther. The action is cumulative, the current rising until point 1 is reached. Any further increase of current above that corresponding to point 1 would necessarily be accompanied by an increase of voltage across the tube. The voltage across the resistance would therefore have to fall, which could be true only if the current became smaller. Hence the current would return to the value corresponding to point 1. A similar analysis shows that if the current is initially that corresponding to point 2, any small initial decrease of current becomes cumulative, and so the current falls to that of point 3.

If the applied voltage E_{bb} is gradually raised from zero, the intersection of the resistance line with the characteristic curves moves along the branch OA of the characteristic curve. When the intersection is at A, an infinitesimal further increase of voltage causes the current to fall abruptly to the value at E, *i.e.*, the circuit triggers. Further increase of supply voltage causes the intersection to move toward C. If the supply voltage is then decreased gradually, the intersection moves down the branch CB until the point B is reached, from which the current jumps abruptly to the value corresponding to point D. It can be seen that similar abrupt changes of current take place if the slope of the resistance line is varied by changing the resistance R or if the characteristic curve is displaced vertically or horizontally. In trigger circuits incorporating vacuum tubes, this displacement can be accomplished by varying one or more electrode voltages.

Since the a-c resistance of a circuit element is defined as the reciprocal of the slope of the static current-*vs.*-voltage characteristic of the element, it follows from the foregoing analysis that any circuit element that has a negative a-c resistance in a portion of its operating range may serve as the basis of a trigger circuit.[1]

[1] Lest the negative slope of the resistance line MN in Fig. 10-2 should be confusing, it is well to point out that this line shows the current through the resistance as a function of voltage *drop* in the resistance. If the current were plotted as a function of voltage applied to the resistance, the resulting curve would have a positive slope and would pass through the origin.

10-2. Tetrode Trigger Circuits.—The plate characteristics of screen-grid tetrodes are similar in form to the curve of Fig. 10-2 (see Fig. 3-9). It is to be expected, therefore, that a trigger circuit can be formed by inserting a resistance in series with the plate supply voltage of a screen-grid tetrode.[1] The slope of the negative-resistance portion of the plate

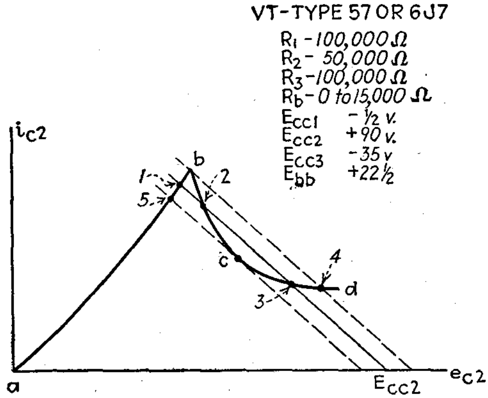


FIG. 10-3.—Current-voltage diagram for pentode trigger circuit.

characteristics of modern tetrodes is so low, however, that resistances of the order of 100,000 ohms or higher, and correspondingly large values of plate supply voltage, must be used. For this reason and because of change in shape of the tube characteristics with tube age as the result of changes in secondary emission, this type of circuit has found little or no application. The circuit is mentioned here because of the possibility of its use in future problems. The circuit can be triggered by varying any of the electrode voltages or the series resistance.

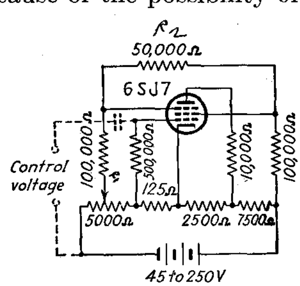


FIG. 10-4.—Pentode trigger circuit.

10-3. Pentode Trigger Circuits.—A second characteristic of the form of that of Fig. 10-2 is one relating the screen current i_{c2} of a suppressor pentode with the screen voltage e_{c2} when the suppressor grid is coupled to the screen in such a manner that a change in screen voltage is accompanied by a proportional change in suppressor voltage, the suppressor voltage being maintained negative and the plate voltage positive. A typical characteristic is illustrated in Fig. 10-3 (see also Fig. 10-43).[2] A trigger circuit can be formed by inserting a high resistance in series with the screen voltage supply, as in the circuit of Fig. 10-4.[3] In this

[1] REICH, H. J., *Electronics*, August, 1939, p. 14.
[2] HEROLD, E. W., *Proc. I.R.E.*, **23**, 1201 (1935).
[3] REICH, H. J., *Rev. Sci. Instruments*, **9**, 222 (1938).

circuit the screen and suppressor grids are coupled by means of the resistance R_2, and the suppressor grid is maintained negative by means of the resistance R_1. The circuit may be triggered by means of voltages in series with any of the electrode supply voltages, by voltage pulses impressed upon one of the grids through a condenser, as shown in Fig. 10-4, or by changes of circuit resistances. The control grid is the most sensitive triggering electrode. The values of the supply voltages are not critical, but the proper relation must be maintained between them. The circuit constants shown in Fig. 10-4 are typical. The resistance R_4 is not required but may be used for the purpose of obtaining a voltage output from the plate circuit.

The physical explanation of the operation of the circuit of Fig. 10-4 is based upon the fact that the negative suppressor prevents some electrons from reaching the plate and causes them to return toward and strike the screen. An increase of negative suppressor voltage causes more electrons to be returned to the screen and thus increases the screen current. Suppose that the screen current in the circuit of Fig. 10-4 has its lower equilibrium value and that the screen voltage is reduced. Because of the screen-suppressor coupling, the reduction of positive screen voltage is accompanied by an increase of negative suppressor voltage. Although the decrease of screen voltage tends to decrease the screen current, the associated increase of negative suppressor voltage tends to increase the screen current. In certain ranges of operating voltages the increase in screen current resulting from the increase of negative suppressor voltage exceeds the decrease resulting from the decrease of positive screen voltage, and so the net result is an increase of screen current. Since *IR* drop in the screen circuit resistance R_3 causes the increase of screen current to be accompanied by further drop in screen voltage, the action becomes cumulative and the screen current continues to rise. As the negative suppressor voltage increases, the effect of the suppressor voltage upon the screen current becomes smaller and finally, when no electrons strike the plate, is practically zero. At some value of suppressor voltage, therefore, the effect of the changing suppressor voltage becomes less than that of the changing screen voltage. As further decrease of screen voltage would then cause a decrease of screen current, the current cannot continue to rise. It then has its higher stable value. The action may be reversed by a suitable increase of screen voltage. Because an increase of screen current is accompanied by a reduction of plate current, the plate current also has two stable values, the higher of which corresponds to the lower value of screen current. By proper choice of voltages and circuit constants, the lower value of plate current may be made zero. This is important in certain applications of the circuit.

10-4. Eccles-Jordan Trigger Circuit.—The most frequently used trigger circuit, first described by Eccles and Jordan,[1] is shown in basic form in Fig. 10-5. The functioning of this circuit depends upon the fact that current flows through only one tube at a time. Assume, for the purpose of explanation, that a condition of equilibrium exists in which both tubes conduct simultaneously. Any small increase of current in either tube, due to any cause, increases the voltage drop in the corresponding coupling resistance and thus increases the negative grid voltage of the other tube and thereby decreases the plate current of the other tube. This in turn reduces the negative grid voltage of the first tube and thereby causes further increase in the plate current of the first tube. The action is cumulative and so one plate current rises while the other falls to zero.

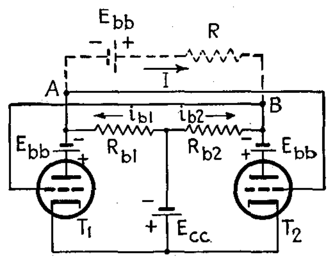


FIG. 10-5.—Basic Eccles-Jordan trigger circuit.

The action of the Eccles-Jordan circuit can also be analyzed with the aid of the current-voltage characteristic of the circuit. Figure 10-6 shows a curve of external current that flows as the result of the application of direct voltage between points A and B of the circuit of Fig. 10-5 (see also Fig. 10-46). Since this curve has a portion with negative slope, the theory of Sec. 4-1 predicts that triggering should take place if a voltage source E_{bb} in series with a resistance R is connected between points A and B, as shown by the dotted lines of Fig. 10-5. The corresponding resistance line MN may intersect the characteristic curve in three points, as shown in Fig. 10-6, if the slope of the resistance line is less than the slope of the characteristic curve at the origin, *i.e.*, if the resistance R exceeds the negative resistance at the origin. Abrupt changes of current through R and of voltage between A and B can be made to occur by varying E_{bb} or by shifting the characteristic curve by changing the operating voltages of the tubes.

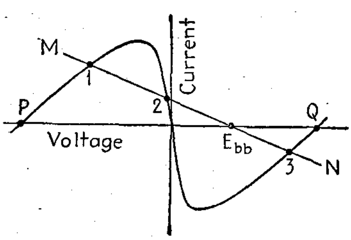


FIG. 10-6.—Current-voltage diagram for the Eccles-Jordan trigger circuit of Fig. 10-5.

If R is increased, the line MN becomes more nearly horizontal and, in the limiting case when R is infinite, coincides with the voltage axis. Making R infinite is, of course, equivalent to eliminating the external circuit made up of E_{bb} and R. The external current is then zero, but

[1] ECCLES, W. H., and JORDAN, F. W., *Radio Rev.*, 1, 143 (1919).

changes in electrode voltages can cause abrupt transfer of current from one tube to the other and a reversal of voltage between A and B. The magnitudes of the two values of this voltage are given by the intercepts P and Q of the characteristic on the voltage axis. If the circuit is in all respects symmetrical, these two values of voltage are of equal magnitude and are equal to the product of R_b and the equilibrium plate current of either tube. Ordinarily this circuit is used without the external resistance R and the voltage supply E_{bb}, the circuit being triggered by changes of grid or plate voltages of such polarities as to reduce the plate current of the conducting tube or as to cause plate current to flow in the non-conducting tube.

Figure 10-7 shows a practical form of the Eccles-Jordan trigger circuit that requires only one voltage source. In this circuit the grid of one tube is coupled to the plate of the other by means of resistance R_c and the correct operating grid voltages are maintained by means of the resistances R_c'. The circuit may be triggered by means of voltages introduced in series with the electrodes, by changes of circuit resistances, or by means of voltage pulses applied to one or more electrodes through transformers or through a condenser, as shown in Fig. 10-4.

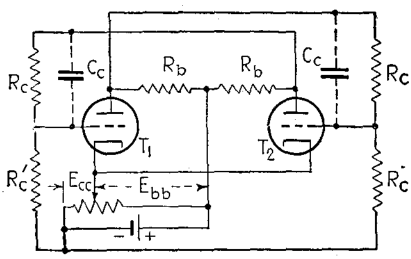


FIG. 10-7.—Practical form of the Eccles-Jordan trigger circuit.

The function of the condensers C_c, which are essential only when the circuit is triggered by means of voltage pulses of short duration, is to prevent the objectionable effects of interelectrode capacitances. Grid-cathode capacitance tends to prevent the grid voltage from changing relative to the cathode. Grid-plate capacitance couples the grid of each tube to the plate of the same tube and thus tends to prevent triggering. (If tube 1 is conducting and the plate current starts to fall, the resulting reduction in voltage drop in the resistance R_b causes the plate voltage to rise. Since the voltage across C_{gp} cannot change instantaneously, there is a corresponding rise in grid voltage, which tends to prevent the reduction of plate current.) Plate-cathode capacitance tends to prevent the plate voltage from changing relative to the cathode. The interelectrode capacitances therefore act in such a manner as to increase the time required for triggering and to reduce the reliability of operation. Condensers C_c, the capacitance of which exceeds the interelectrode capacitances, reduce the undesirable effects of the grid-cathode and grid-plate capacitances. Because the voltages across these condensers cannot change instantaneously, a change of voltage of either plate results in a nearly equal instantaneous change in voltage of the grid of the other tube. When the condensers are omitted, on the other hand, the change in grid voltage cannot exceed the change in plate voltage multiplied by the ratio $R_c'/(R_c + R_c')$, and the action of the interelectrode capacitances may reduce the change in grid voltage considerably below this value. The capacitance of the condensers C_c should not be larger than necessary, since the charging and discharging of these condensers immediately following triggering cause exponential variations of the circuit currents and voltages that are objectionable in some applications. Fifty micro-microfarads is usually a satisfactory value.

Figure 10-8 shows a modification of the basic Eccles-Jordan circuit that has desirable characteristics.[1] In this circuit the suppressor grids of pentodes serve the same function as the triode grids of the circuit of Fig. 10-7. A constant positive voltage is applied to the screen grids, as in the use of pentodes as voltage amplifiers, and the control grids are used for triggering the circuit. Because the high negative voltage of the suppressor grid of the nonconducting tube prevents the flow of plate current regardless of the voltage of the control grid, the circuit cannot be made to trigger by the application of positive voltage to the control grid of the nonconducting tube. Application of negative voltage to the control grid of the conducting tube, however, reduces its current and thus triggers the circuit. If short negative voltage pulses are applied simultaneously to both grids, the plate currents of both tubes remain zero throughout the duration of each pulse. During this time the condensers C_c cause the suppressor of the tube which has been conducting to be more negative than that of the other tube and thus cause the current to transfer to the other tube at the end of the triggering pulse.[1,2] The size of the condensers should be such that the time taken for them to charge or discharge from one equilibrium value of voltage to the other value is large in comparison with the duration of the triggering impulse, but small in comparison with the time between successive pulses. Their capacitance should also exceed the interelectrode capacitances. Condensers of 50 $\mu\mu$f capacitance are usually satisfactory. Reliable triggering necessitates the use of triggering pulses of very short duration. The purpose of the resistors R_i is to maintain the control grids at cathode potential

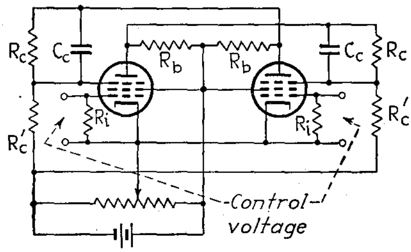


FIG. 10-8.—Modified form of Eccles-Jordan trigger circuit using pentodes.

[1] REICH, H. J., *Rev. Sci. Instruments*, **9**, 222 (1938).
[2] TOOMIM, H., *Rev. Sci. Instruments*, **10**, 191 (1939).

between triggering impulses. If the grids are allowed to float, stray fields cause random or periodic triggering.

The functions of the control and suppressor grids in the circuit of Fig. 10-8 may be interchanged, but the circuit is then sensitive to triggering voltage of either polarity. It is also possible to connect the control and screen grids of pentodes as the triode grids and plates, respectively, are connected in the circuit of Fig. 10-7, and to take the output voltage or current from the plate circuits.

10-5. Design of Eccles-Jordan Trigger Circuits.—The values of the resistances R_b, R_c, and R_c' are not critical in the circuits of Figs. 10-7 and 10-8, but it is essential that the circuit be symmetrical. Small inequalities of the resistances in the two halves of the circuit or differences in the characteristics of the two tubes tend to make one tube or the other carry current continuously. The plate currents decrease with increase of R_b, and the voltage across R_b increases. The choice of R_b is therefore governed to some extent by whether a current or voltage output is desired from the circuit. There is also a minimum value of R_b below which the circuit does not have the negative-resistance characteristic essential to triggering (see Prob. 4-5). R_c and R_c' should be at least four or five times as large as R_b and may conveniently be made equal. Typical resistance values are $R_b = 10{,}000$ ohms and $R_c = R_c' = 250{,}000$ ohms. Supply voltages of 20 volts or less may be used in these circuits, but the voltage output available across R_b increases with supply voltage. The circuits are not critical as to type of tube. The voltage divider used to provide grid bias in the circuits of Figs. 10-7 and 10-8 may be replaced by a properly chosen biasing resistor in the common cathode lead of the two tubes, as shown in Fig. 10-9. (This may be accomplished in effect by disconnecting the positive end of the voltage divider from the voltage supply in the circuits of Figs. 10-7 and 10-8).

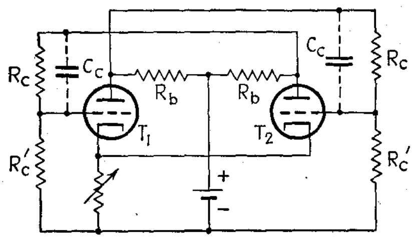


FIG. 10-9.—Self-biased Eccles-Jordan trigger circuit.

10-6. Applications of Trigger Circuits.—Any of the circuit currents of a trigger circuit may be used directly to operate a relay or other current-controlled device if the inductance or resistance of the device is not so great as to prevent triggering. It is preferable, however, to use the voltage drop across one of the resistors to control the plate current of an amplifier tube, which may in turn be used to operate the relay or other device. Although voltage output may be taken directly from one of the resistors if the resistance of the load is so high as to prevent unbalancing of the circuit, the use of a voltage amplifier between the trigger circuit and the load is usually advisable. The circuits may be triggered by means of voltages applied to one or more electrodes, as already explained. Usually the circuits are so sensitive that the change of electrode voltage resulting from touching one of the tube electrode terminals is sufficient to cause triggering. The circuit of Fig. 10-8 can be triggered by negative control-grid voltage as low as $\frac{1}{2}$ volt. Any of the Eccles-Jordan circuits may also be triggered by means of changes of illumination if phototubes are connected in series with or in place of the resistances R_c or in parallel with or in place of the resistances R_c'. Trigger circuits are used as the basis of pulse generators, relaxation oscillators, switching circuits, counting circuits, frequency meters, time and speed meters, and other important circuits. Most of these applications will be taken up in this chapter and in Chap. 15.

10-7. Circuits for Generating Triggering Pulses.—In order to ensure reliability in triggering, the triggering impulses applied to trigger circuits

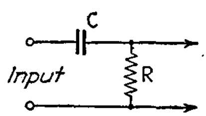


FIG. 10-10.—Pulse-sharpening circuit (differentiating circuit).

FIG. 10-11.—Form of the pulse produced by an abrupt change of direct voltage impressed upon the circuit of Fig. 10-10.

must be of very short duration. Suitable pulses may be produced by means of the simple circuit of Fig. 10-10 if R and C are small. An abrupt change of direct voltage impressed upon the input to the circuit is followed by an exponential charging or discharging of the condenser. The flow of current through the resistance R during the charging or discharging time produces a pulse of output voltage of the exponential form shown in Fig. 10-11. The duration of the pulse can be made as short as desired by making R and C small.[1] If the input voltage is a periodic wave having discontinuities, such as a rectangular or triangular wave, an output pulse is produced at each discontinuity of input voltage. Figure 10-12a shows a typical wave of output voltage produced by square-wave input voltage. The output voltage produced by a triangular

[1] The equation of the pulse of Fig. 10-11, derived under the assumption that the inductance in series with C and the capacitance shunting R are negligible, is

$$e_o = \Delta E_i \epsilon^{-t/RC},$$

in which ΔE_i is the change in direct input voltage. The series inductance and shunting capacitance tend to prevent sudden changes in output voltage and thus make it difficult to attain the rapid rise in voltage shown in Fig. 10-11 when the pulse is of very short duration.

wave of input voltage consists of periodic pulses of one polarity only, as shown in Fig. 10-12*b*.

The resistance R of Fig. 10-10 may serve also as one of the trigger-circuit coupling resistances, such as R_c' of Fig. 10-7, R_i of Fig. 10-8, or the first-grid resistor of Fig. 10-4; and C may serve as an input coupling condenser, such as that of Fig. 10-4. Figure 10-13 shows a modification of the circuit of Fig. 10-10 that makes possible triggering by the opening and closing of a switch. A pulse of one polarity is produced when the switch is closed, and of opposite polarity when it is opened. The resistance R_d of Fig. 10-13, which serves to discharge the condenser when the switch is opened, should be small in comparison with R_c' or R_i.

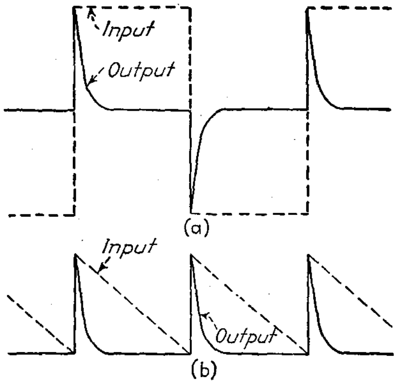


FIG. 10-12.—Wave forms of output voltages produced by the application of square and saw-tooth waves of voltage to the input of the circuit of Fig. 10-10.

Periodic triggering pulses may be derived from a sinusoidal voltage by first transforming the sinusoidal voltage into a rectangular wave and impressing the latter upon the input of the circuit of Fig. 10-10. A simple "peak-clipping" circuit by means of which a sine wave may be converted into a wave approximating rectangular form is shown in Fig. 10-14. The output voltage increases with input voltage until the instantaneous impressed voltage is approximately equal to the diode biasing voltage. Current then starts flowing in one of the diodes and the resulting voltage drop in the series resistor prevents appreciable further rise of output voltage. The solid curve of Fig. 10-15 shows the form of the output voltage, which may be amplified and applied to another peak-clipping stage in order to make the reversal of voltage more abrupt. By the use of a number of stages, the voltage may be made to approach as nearly as desired a true rectangular wave. Peak clipping may also be accomplished by means of a limiter of the type discussed in Sec. 9-30.

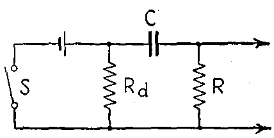


FIG. 10-13.—Circuit for generating triggering pulses by the closing and opening of a switch.

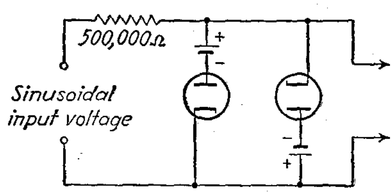


FIG. 10-14.—Peak-clipping circuit for the production of voltages of rectangular wave form.

Symmetrical voltage pulses may be produced with the circuit of Fig. 10-14 by taking the output from across the resistance, instead of across the tubes, and making the diode biasing voltage so high that current flows through the diodes only at the peaks of the voltage waves. If only one diode is used, the pulses are of one polarity only; if two diodes are used, as in Fig. 10-14, alternate pulses are of opposite polarity. Similar results are obtained by applying a sinusoidal voltage to the input of a Class C amplifier biased considerably beyond cutoff. If the amplifier is single-sided, the pulses are of one polarity only; if the amplifier is push-pull, alternate pulses are of opposite polarity.

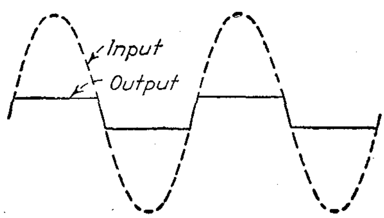


FIG. 10-15.—Input and output voltages in the circuit of Fig. 10-14.

Sharp periodic voltage pulses can also be generated by a self-biased sinusoidal oscillator in which the biasing resistor and condenser are so large that oscillation ceases after a single cycle of oscillation and does not start again until the lapse of appreciable time. The theory of such an oscillator is discussed in Sec. 10-38, and a typical circuit diagram and the form of the output voltage are shown in Figs. 10-58 and 10-59.

10-8. Rectangular-wave and Rectangular-pulse Generators.—Rectangular waves of voltage are produced across the screen and plate resistors of the circuit of Fig. 10-4 and across the plate resistors of the circuits of Figs. 10-7 and 10-8 when the circuits are triggered periodically. If alternate pulses are spaced so that the circuits remain in one equilibrium state longer than in the other, the positive and negative halves of the rectangular waves are of unequal duration. Random rectangular pulses may be generated by the use of two triggering impulses separated in time by the required duration of the pulse. The first impulse triggers the circuit in one direction, and the second in the other.

Periodic or random rectangular voltage or current pulses of controllable length, initiated by single triggering impulses, may also be generated by unbalancing a trigger circuit in such a manner that it normally remains in one equilibrium state, and returns to that state after a short time interval when triggered into the other equilibrium state. This may be accomplished in the Eccles-Jordan type of circuit by replacing one coupling resistance R_c by a condenser. A typical circuit is shown in Fig. 10-16.

The circuit of Fig. 10-16 operates as follows. The voltage dividers controlling the grid biases are adjusted so that tube T_1 normally conducts, but so that the circuit can be triggered. Application of a negative triggering impulse to the grid of T_1 stops the plate current in T_1 and starts it in T_2. The resulting voltage drop in R_b lowers the plate voltage of T_2 and hence the voltage impressed upon the series combination of R_c' and C_c. C_c therefore starts discharging. The discharge current i_i flows through R_c' in the indicated direction and thus produces a voltage drop that biases the grid of T_1 beyond cutoff. As C_c discharges, however, the discharge current falls exponentially and so the negative bias of T_1 gradually becomes smaller. After a time interval that is determined principally by the size of R_c' and C_c and to a lesser extent by the value of

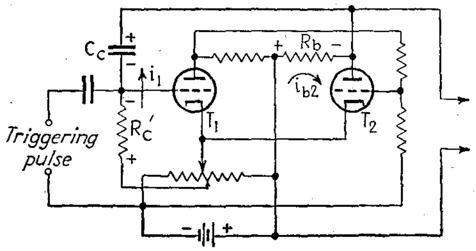


FIG. 10-16.—Circuit for the production of rectangular voltage pulses of controllable length.

R_b and of the supply voltage, the grid voltage of T_1 rises above cutoff. Current then again starts flowing in T_1, and the circuit triggers back to its original state of equilibrium, in which no current flows in T_2.[1] Because the discharging and charging of C_c immediately following triggering produces an exponentially changing component of current through R_b, the voltage across R_b rises and falls exponentially, as shown in Fig. 10-17. If the resistance R_b is sufficiently small in comparison with R_c', however, so that the condenser current does not greatly affect the voltage drop in R_b, the voltage pulse produced across R_b is approximately rectangular. It can be made truly rectangular by passing it through a circuit of the form of Fig. 10-10 or a limiter type of amplifier such as that discussed in Sec. 9-30. Periodic pulses are obtained by the use of periodic triggering pulses. The pulse length is controlled by varying R_c' or C_c.

FIG. 10-17.—Voltage pulse produced by the circuit of Fig. 10-16.

10-9. Vacuum-tube Oscillators.—Advantages of electron-tube oscillators over alternators in the generation of alternating voltage and current include their wide frequency range, the freedom from harmonics of

[1] The circuit may also be analyzed by making use of the fact that it takes time for the voltage across a condenser to change. Since the voltage across C_c cannot change instantaneously, the grid voltage of T_1 is lowered by an amount that is initially equal to the abrupt reduction of plate voltage of T_2. The high negative voltage thus applied to the grid of T_1 prevents this tube from conducting. The drop in voltage across R_b caused by the flow of i_{b2}, however, causes C_c to discharge through R_c'. The exponential reduction of voltage across C_c is accompanied by a corresponding gradual reduction of negative grid voltage of T_1 until triggering takes place.

certain types and the richness in harmonics of others, their constancy of frequency, the ease with which their frequency may be varied, their portability, and their comparatively low cost. The fact that the amplitude and frequency of the output of electron-tube oscillators may be rapidly varied by means of voltage is one of the factors that have made possible the development of radio communication and broadcasting. Electron-tube oscillators have become an indispensable part of the equipment of scientific, educational, commercial, and defense laboratories.

The following classification, although it is not complete as to modifications of fundamental types, indicates the principal kinds of vacuum-tube oscillators:

1. Relaxation:
   a. Glow-discharge and arc-discharge.
   b. Multivibrator.
   c. van der Pol (negative-transconductance).
   d. Saw-tooth–wave, using high-vacuum tubes.
2. Negative-resistance:
   a. Dynatron.
   b. Negative-transconductance (transitron).
   c. Push-pull.
   d. Negative–grid-resistance.
   e. Resistance-capacitance–tuned.
3. Feedback:
   a. Tuned-plate.
   b. Tuned-grid with inductive feedback.
   c. Tuned-grid with capacitive feedback.
   d. Hartley.
   e. Colpitts.
   f. Tuned-grid–tuned-plate and other complex types using more than one tuned circuit.
   g. Resistance-capacitance–tuned.
4. Magnetostriction and piezoelectric.
5. Heterodyne.
6. Ultrahigh-frequency:
   a. Positive-grid.
   b. Magnetron.
   c. Velocity-modulation.
7. Ionic.
8. Mechanical-electronic.

10-10. Relaxation Oscillators.—Relaxation oscillators are oscillators in which one or more currents or voltages change abruptly at one or more times in the cycle of oscillation. In certain applications, relaxation oscillators have a number of advantages over oscillators that give a sinusoidal output. Among these advantages are the number and amplitude of harmonics present in the output; the ease with which the frequency may be stabilized by the introduction into the oscillating circuit of small voltages whose frequency is a multiple or submultiple of the oscillator frequency; the wide range of frequency that can be obtained with a single oscillator; and the compactness, simplicity, and low cost of relaxation oscillators. They are of particular value in the production of saw-tooth voltages required for the operation of cathode-ray oscillographs (see Sec. 15-20). The high harmonic content of relaxation oscillators, on the other hand, makes them unsuitable for applications in which sinusoidal wave form is essential.

Glow- and arc-discharge tubes are often used in relaxation oscillators. The operation of glow and arc relaxation oscillators is so closely associated with the theory of glows and arcs that they can be discussed best in Chaps. 11 and 12, which deal with glow- and arc-discharge tubes and circuits. Relaxation oscillators using high-vacuum tubes are based upon the trigger circuits discussed in Secs. 10-2 to 10-5. Condensers are incorporated in these circuits in such a manner that triggering from one state of equilibrium to the other is followed by the charging or discharging of one or more condensers. The resulting gradual change of condenser voltages produces corresponding changes of electrode voltages. Critical values are eventually reached at which the circuit triggers to its original state of equilibrium. Triggering is again followed by gradual changing of condenser voltages to values at which the circuit again triggers and the cycle repeats.

10-11. The Multivibrator.—The most frequently used type of high-vacuum-tube relaxation oscillator is the *multivibrator*, shown in basic form in Fig. 10-18.[1] Comparison with the Eccles-Jordan trigger circuit of Fig. 10-8 shows that the multivibrator is derived from the Eccles-Jordan circuit by replacing the resistors R_c by condensers C_c. The action of the multivibrator is similar to that of the pulse generator of Fig. 10-16, explained in Sec. 10-8. Figure 10-19 shows current directions and typical voltages in the circuit immediately following transfer of current from tube T_1 to tube T_2. The cessation of current in T_1 causes the plate voltage of T_1 to rise[2] and the start of current in T_2 causes the plate voltage of T_2 to fall. Since the plate voltage of T_1 is impressed across the series

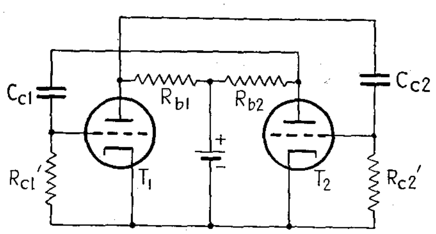


FIG. 10-18.—Multivibrator.

[1] ABRAHAM, H., and BLOCH, E., *Ann. Physik*, **12**, 237 (1919).

[2] Because of the flow of condenser current i_2 through R_{b1}, the plate voltage of T_1 does not immediately rise to the value of the plate supply voltage.

combination of R_{c2} and C_{c2}, and the plate voltage of T_2 is impressed across the series combination of R_{c1} and C_{c1}, the sudden change in plate voltages causes a condenser discharging current i_1 to flow out of C_{c1} and a charging current i_2 to flow into C_{c2}. The flow of i_1 through R_{c1}' biases the grid of T_1 negatively and prevents it from conducting; the flow of i_2 through R_{c2} and the grid of T_2 biases the grid of T_2 positively and thus keeps it conducting. As C_{c1} discharges and C_{c2} charges, however, the currents i_1 and i_2 decrease exponentially and so the biasing voltages become smaller.

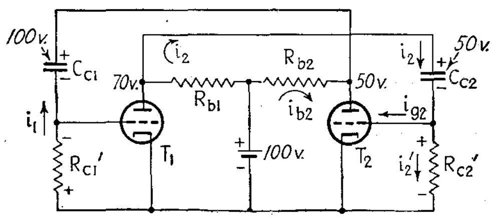


FIG. 10-19.—Instantaneous currents and voltages in the multivibrator immediately following transfer of current from tube T_1 to tube T_2.

The negative bias of T_1 eventually becomes so small that plate current starts flowing in T_1 and the circuit triggers.[1] The transfer of current from T_2 to T_1 is followed by a similar series of events, the directions of the currents i_1 and i_2 being such as to maintain the grid of T_1 positive and that of T_2 negative. As C_{c1} charges and C_{c2} discharges, the condenser currents, and hence the biasing voltages, gradually become smaller until the circuit again triggers. The cycle then repeats.

Figure 10-20 shows the wave forms of the grid voltage and the grid current of either tube. The fact that the voltage is small during the positive portion of the cycle is explained by the flow of grid current during the time the grid is positive. Because the grid current i_{g2} rises rapidly with increase of positive grid voltage and produces a voltage drop in R_{b1}, the grid voltage cannot swing far positive. The more rapid change of grid voltage during the positive portion of the cycle than during the negative portion is also caused by the flow of grid current. R_{c2}' is shunted by the grid resistance of T_2, which is small in comparison with both R_{c2}' and R_{b1} during the positive portion of the cycle. During the positive part of the cycle, therefore, the charging current of C_{c2} is limited almost

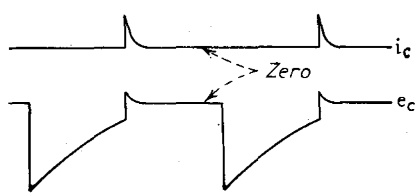


FIG. 10-20.—Wave form of grid current and grid voltage in a symmetrical multivibrator.

[1] An alternative explanation of the action of this circuit is suggested by the footnote on p. 360.

entirely by R_{b1}. During the negative portion of the cycle, however, the grid resistance is practically infinite, and so the discharging current is limited by $R_{b1} + R_{c2}'$. Since R_{c2}' is considerably larger than R_{b1}, the condenser charges much more rapidly during the positive part of the cycle than it discharges during the negative part. Because the grid voltage of the conducting tube falls to zero more rapidly than the negative grid

FIG. 10-21.—Typical wave form of condenser voltage in a symmetrical multivibrator.

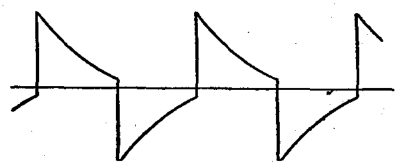


FIG. 10-22.—Typical wave form of grid-to-grid voltage in a symmetrical multivibrator.

voltage of the other tube rises to the value at which plate current starts flowing, triggering is determined by the grid voltage of the nonconducting tube, and the period of oscillation by the rate of rise of negative grid voltage. Typical wave forms of condenser voltage, grid-to-grid voltage, and plate-to-plate voltage in a symmetrical circuit are shown in Figs. 10-21 to 10-23. The plate-to-plate voltage is of the form of Fig. 10-23*a*

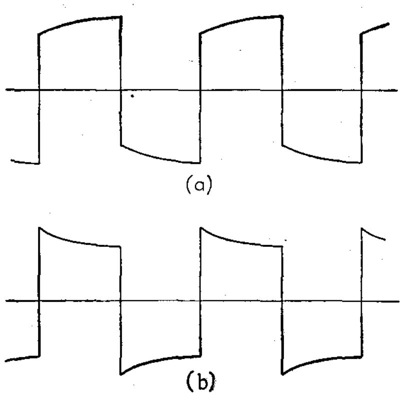


FIG. 10-23.—Typical wave forms of plate-to-plate voltage in a symmetrical multivibrator.

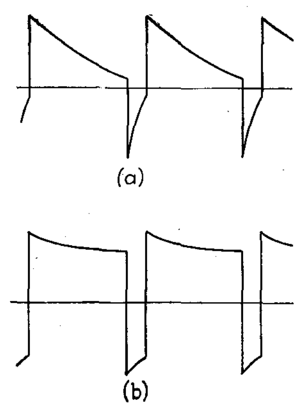


FIG. 10-24.—Typical wave forms of (*a*) grid-to-grid voltage and (*b*) plate-to-plate voltage in an asymmetrical multivibrator.

when the plate resistors are large and the grid resistors small; it is of the form of Fig. 10-23*b* when the plate resistors are small and the grid resistors large. Typical curves of grid-to-grid and plate-to-plate voltage in an asymmetrical circuit are shown in Fig. 10-24.

The multivibrator is not critical as to tube type, supply voltage, or circuit constants, provided that the resistances R_{b1} and R_{b2} are large enough to ensure that the circuit can act as a negative resistance between plates (see Prob. 4-5). The frequency of oscillation increases with decrease of resistances and capacitances. The output voltage may be taken from across any of the circuit elements. Unless the impedance of the load is very high, however, it may be advisable to connect the load to the output of an amplifier which is directly coupled to the multivibrator.

10-12. Van der Pol Relaxation Oscillator.—Figure 10-25 shows the circuit of the van der Pol relaxation oscillator,[1] which may be derived from the pentode trigger circuit of Fig. 10-4 by replacing the screen-suppressor coupling resistor by a condenser. The circuit functions as follows: Triggering from the higher to the lower value of screen current causes an abrupt rise of screen voltage as the result of decreased voltage drop in the screen resistor. Because the voltage across the condenser cannot change instantaneously, there is an initial change of suppressor voltage equal to the change in screen voltage. The decreased negative suppressor voltage maintains the lower screen current, but a charging current immediately starts flowing into the condenser through R_3 and R_1. As the voltage across the condenser rises, the suppressor voltage becomes more negative and finally reaches a critical value at which the circuit triggers back to the higher value of screen current. The resulting abrupt fall of screen and suppressor voltages is immediately followed by discharging of the condenser and exponential rise of suppressor voltage until the circuit again triggers. Figure 10-26 shows a typical wave of condenser current or suppressor or screen voltage. The wave is asymmetrical because during the charging of the condenser the screen current has its lower value and, if the suppressor swings positive, suppressor current flows. The flow of suppressor current, together with the decrease of screen current and the resulting increase in the portion of the current through R_3 that flows into C, causes the condenser voltage to change more rapidly than during the discharge of the condenser, when no suppressor current flows and the screen current is higher. The wave form of the condenser voltage is similar to the wave shown in Fig. 10-21, but the rising portion of the wave is of shorter duration than the falling portion. The frequency of oscillation increases with decrease of R_1 and C. R_3 also affects the frequency, both because it affects the rate at which the condenser charges and discharges and

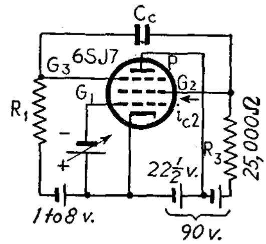


FIG. 10-25.—Van der Pol relaxation oscillator.

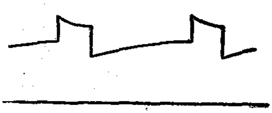


FIG. 10-26.—Typical wave form of condenser current or screen or suppressor voltage in the circuit of Fig. 10-25.

[1] VAN DER POL, B., *Phil. Mag.*, **2**, 978 (1926); PAGE, R. M., and CURTIS, W. F., *Proc. I.R.E.*, **18**, 1921 (1930); HEROLD, E. W., *Proc. I.R.E.*, **23**, 1201 (1935).

because it determines the voltages at which the circuit triggers. As R_3 is increased from a low value, the frequency first decreases, passes through a minimum, and then increases.[1]

10-13. Relaxation Oscillators for the Generation of Saw-tooth Waves.[2]—By adjustment of circuit constants and operating voltages, the condenser voltage in the circuit of Fig. 10-25 can be made to have the saw-tooth wave forms shown in Fig. 10-27. Similar waves can be obtained from across the condensers of an asymmetrical multivibrator. Waves of the form of Fig. 10-27*b* are useful in connection with the operation of cathode-ray oscillographs (see Sec. 15-20) and cathode-ray television equipment.

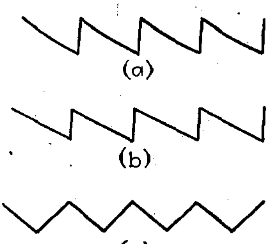


FIG. 10-27.—Typical waves of condenser voltage in the circuits of Figs. 10-25, 10-28, and 10-29.

Figure 10-28 shows another circuit, based upon the pentode trigger circuit of Fig. 10-4, with which a voltage wave of the form of Fig. 10-27*b* can be more readily produced. The action of this circuit is dependent upon the abrupt change of plate current at critical values of plate voltage. When the plate circuit is first completed, C is uncharged, and so the full plate supply voltage is applied to the plate. The plate current assumes its larger value, charging the condenser and thus lowering the plate voltage. At some value of condenser voltage, the plate voltage reaches the critical value at which the circuit triggers and the plate current falls abruptly to its lower value, which may be zero or, at most, less than the condenser discharge current through R. The condenser then discharges through R until the plate voltage rises to its upper critical value and the cycle repeats. By making the plate supply voltage much greater than the difference in the critical plate voltages, the condenser may be made to discharge very nearly linearly, giving a wave of condenser voltage of the form of curve b of Fig. 10-27.

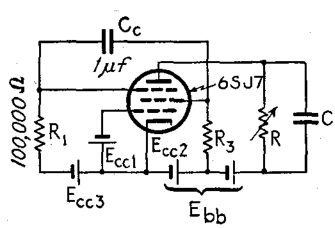


FIG. 10-28.—Relaxation oscillator for the generation of saw-tooth voltage.

The time taken for the condenser to charge decreases as the higher value of plate current is increased. By the use of screen supply voltage of 100 volts or more, positive control-grid voltage, and high values of R, the charging time may be reduced to a small fraction of the cycle. With screen supply voltage of $22\frac{1}{2}$ volts or less and negative control-grid

[1] PAGE and CURTIS, *loc. cit.*

[2] An excellent survey of saw-tooth–wave oscillators has been given by O. S. PUCKLE, *J. Inst. Elec. Eng.* (*London*), **89,** 100 (1942).

voltage, on the other hand, the two values of plate current do not differ greatly, and the wave is of the form of Fig. 10-27*c*. Decrease of R increases the charging time and decreases the discharging time. At high values of R, the net effect of decrease of R is an increase of frequency. The oscillation frequency increases with decrease of C. An upper frequency limit exists as the result of plate-cathode tube capacitance, which has the same effect as the condenser C. The function of C_c is merely to couple the screen and the suppressor. It has little or no effect upon the frequency if its reactance at oscillation frequency is small in comparison with the resistance R_1. The need of excessively high capacitance at very low frequency may be avoided by replacing the coupling condenser with a resistance, as in the basic trigger circuit of Fig. 10-4. This necessitates a readjustment of supply voltages.

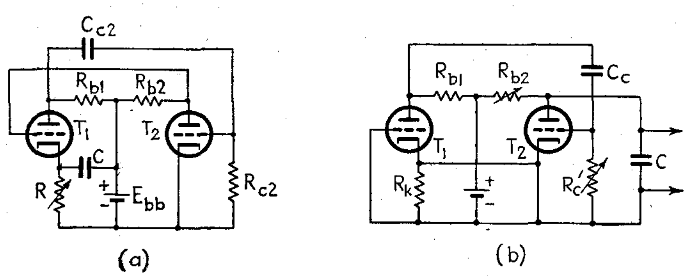


FIG. 10-29.—Relaxation oscillators for the generation of saw-tooth voltage.

A form of saw-tooth–wave oscillator derived from the Eccles-Jordan trigger circuit is shown in Fig. 10-29*a*.[1] The action is as follows: When the circuit is first closed, C is uncharged, the plate and cathode of tube 1 are at the same potential, and current flows only through tube 2. C charges through a high resistance R, and the plate voltage of tube 1 rises. At a critical value of condenser voltage, current starts flowing in tube 1, and the current abruptly ceases in tube 2. C discharges rapidly through tube 1 and its plate resistor until a lower critical voltage is reached. The current then again transfers to tube 2, and the cycle repeats. By the use of high supply voltage, or by the substitution in place of R of a pentode, the plate current of which is practically independent of plate voltage over a wide range of plate voltage, the condenser voltage wave may be made to assume the form of Fig. 10-27*b*. The frequency of oscillation increases with decrease of C and R or with increase of charging current if a pentode is used in place of R (see pages 456 and 489).

Another saw-tooth–wave oscillator based upon the Eccles-Jordan trigger circuit is shown in Fig. 10-29*b* (see Prob. 10-1).[2]

[1] PUCKLE, O. S., *J. Sci. Instruments*, **13**, 78 (1936).

[2] POTTER, J. L., *Proc. I.R.E.*, **26**, 713 (1938).

Figure 10-30 shows an interesting circuit developed by Puckle.[1] The action is as follows: When the supply voltage is first impressed, the condenser C is uncharged, and the drop in voltage in R_3 caused by the plate current of T_3 causes the grid of T_2 to be biased beyond cutoff. C charges linearly through the pentode T_1, which passes a constant current. When the voltage across C reaches a critical value, plate current starts flowing in T_2. The flow of this current through R_2 applies a negative increment of voltage to the grid of T_3. An amplified positive increment of voltage, produced across R_3 is applied to the grid of T_2, reducing its bias and thus increasing the plate current of T_2. Since the action is cumulative, C is rapidly discharged through T_2. The cycle then repeats. The rate of charging of C, and hence the frequency of oscillation, are controlled by R_1, which changes the bias of the second grid of T_1 and thus the charging current. The need for the amplifier tube T_3 is avoided in a similar

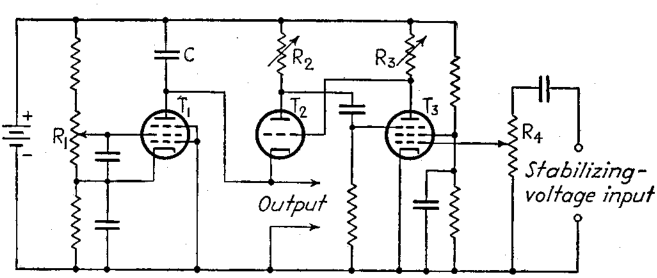


FIG. 10-30.—Relaxation oscillator for the generation of saw-tooth voltage.

circuit developed by Kobayashi, in which the plate circuit of the discharge triode is coupled to the grid by means of a transformer in such a manner that initiation of plate-current flow in the triode causes the application of a positive increment of grid voltage to the triode and thus causes the plate current to rise cumulatively.[2]

A saw-tooth wave of voltage may also be obtained from a self-biased feedback oscillator in which the biasing resistance and condenser are so chosen that the biasing voltage across the condenser increases very rapidly to such a high value that oscillation ceases after a few cycles (see Sec. 10-38). The condenser then discharges slowly through the resistance.[3] The discharge may be made linear with respect to time by substituting a pentode for the resistance and thus making the discharge current constant.

An entirely different type of high-vacuum-tube saw-tooth–wave generator is shown in Fig. 10-31. The grid of the tube is normally biased beyond cutoff, so that the condenser charges up gradually through

[1] PUCKLE, O. S., *J. Television Soc.* (*Brit.*), **2**, 147 (1936).

[2] U. S. Patent 1913449 (1929).

[3] APPLETON, WATSON, and WATT-HERD, Brit. Patent 234254.

the resistance R. Periodic positive voltage pulses of short duration, impressed upon the grid, cause the tube to conduct and thus to discharge the condenser periodically. The rate at which the condenser voltage rises and the voltage of the condenser at the instant it starts to discharge depend upon the values of the condenser capacitance C, the resistance R, and the charging voltage E. The frequency of the saw-tooth condenser voltage is that of the periodic pulses impressed upon the grid. Periodic triangular voltage pulses of the form shown in Fig. 15-51 are generated by this circuit if the grid bias is zero and the voltage pulses applied to the grid are negative.[1] High plate current then normally keeps the condenser essentially discharged. The negative pulses of grid voltage cut the plate current off for the duration of the pulses, allowing the condenser to charge through the resistance. To ensure that the condenser starts charging and discharging abruptly, the grid-voltage pulses should be rectangular. The length and frequency of the triangular pulses generated in this manner are the same as those of the grid-excitation pulses. A complete circuit, including the rectangular pulse generator that furnishes the grid excitation, is shown in Fig. 15-52. In both methods of using the circuit of Fig. 10-31, linear charging of the condenser is approximated by making the supply voltage much higher than the maximum voltage to which the condenser charges.

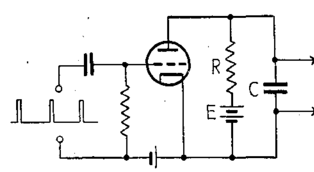


FIG. 10-31.—High-frequency saw-tooth–wave generator.

Saw-tooth–wave generators using glow- and arc-discharge tubes are discussed in Secs. 12-6 and 12-33.

10-14. Use of Relaxation Oscillators in Frequency Transformation.— Relaxation oscillators are of great value in frequency transformation. Introduction into the oscillator circuit of a small voltage equal to a multiple or submultiple of the oscillation frequency causes the relaxation oscillator to "lock in." This locking action is possible because in all relaxation oscillators the relaxations take place at certain critical values of some voltage in the circuit. If a peak of the control voltage comes just before the relaxation occurs, the added voltage is enough to trip the circuit. The control voltage may have a much higher frequency than that of the oscillator because the only peak of locking voltage which affects the oscillator is that which comes just as relaxation is about to occur. If the frequency of the control voltage is lower than that of the oscillator, then every nth relaxation of the oscillator will be controlled, where n is the ratio of the oscillator frequency to the control frequency. With reasonable care in circuit adjustment, relaxation oscillators may

[1] HAWORTH, L. J., *Rev. Sci. Instruments*, **12**, 478 (1941).

be controlled when the frequency ratio is as great as 50. The fundamental output of one controlled oscillator may be used to control a second relaxation oscillator, which may in turn control a third, etc. By this means it is possible to obtain an audio-frequency voltage of great stability from the final relaxation oscillator when the first one is controlled by a crystal-controlled radio-frequency oscillator. The a-f output may be used to drive a synchronous clock.[1] By observation of the clock over a long period of time a very accurate determination may be made of the frequency of the crystal-controlled oscillator or of any lower frequency controlled by it. By separating and amplifying the various

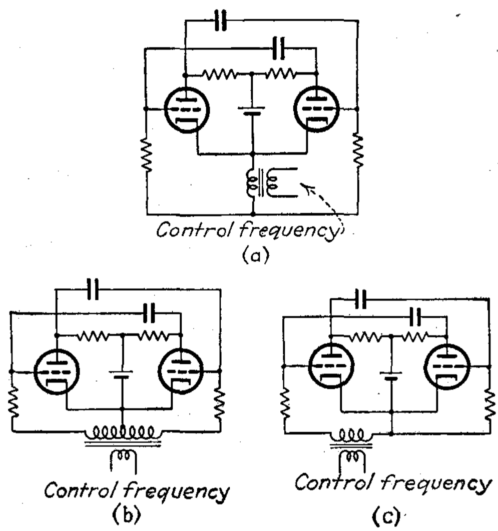


FIG. 10-32.—Multivibrator control circuits.

harmonics of the controlled relaxation oscillators, a large number of frequencies of high constancy may be obtained from one crystal oscillator.

The multivibrator has proved to be the most satisfactory type of relaxation oscillator for use in frequency conversion. Its use in this manner has been discussed by a number of investigators.[2] Figure 10-32 shows three ways in which the control voltage may be introduced into the multivibrator. In circuit *a* the control frequency must be an even integral multiple of the multivibrator frequency, in circuit *b* it must be an odd integral multiple, and in circuit *c* it may be any integral multiple. Resistance-capacitance coupling may be used in place of transformer coupling in impressing the control frequency upon the multivibrator

[1] HORTON, J. W., and MARRISON, W. A., *Proc. I.R.E.*, **16**, 137 (1928).

[2] ABRAHAM and BLOCH, *loc. cit.*; MERCIER, M., *Compt. rend.*, **174**, 448 (1922); CLAPP, J. K., *J. Opt. Soc. Am. and Rev. Sci. Instruments*, **15**, 25 (1927); MARRISON, W. A., *Proc. I.R.E.*, **17**, 1103 (1929).

circuit. An excellent analysis of the control of multivibrator frequency has been given by Hull and Clapp, and further details by Andrew.[1]

10-15. Analysis of Multivibrator Frequency Control.—As pointed out in Sec. 10-11, triggering of a multivibrator is determined by the voltage of the grid that is negative. For this reason, the action of the stabilized circuits can be analyzed by reference to the negative portion of the wave of grid voltage. Under equilibrium conditions, each cycle of oscillation of the stabilized circuit must be identical with the previous cycle, and the grid voltage at the beginning of each cycle must be the same as at the beginning of the previous cycle. Furthermore, in a symmetrical circuit, the number of cycles of control voltage per half cycle must be the same in both halves of the cycle. The instantaneous grid voltage must therefore be the same for both tubes at the beginnings of their respective half cycles. In the circuit of Fig. 10-32*a*,

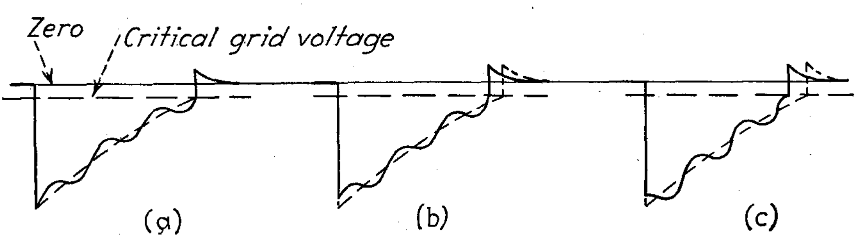


FIG. 10-33.—Control of multivibrator frequency. The wave form of the grid voltage with and without control voltage is shown by the solid and dotted waves, respectively.

in which the control voltage is applied in like phase to the grids of the two tubes, this requirement is satisfied if the instantaneous value of the control voltage of one tube at the end of the negative half cycle of that tube is the same as at the beginning of that half cycle, *i.e.*, if there is an integral number of half cycles of control voltage per half cycle of oscillation. (This assumes that the time required for triggering is of negligible duration.) Figure 10-33 shows how the frequency of oscillation is changed by an increase of control frequency at a frequency ratio of 6. The dotted curve shows the manner in which the negative grid voltage of one tube varies without control voltage. The solid curve shows the variation of grid voltage when control voltage is impressed, and the dashed line indicates the critical grid voltage at which triggering occurs. It can be seen that, as the frequency of the control voltage is increased, the beginning of the negative half-cycle shifts in phase relative to the control voltage in such a manner that there are always three cycles of control voltage per half cycle of oscillation. If the frequency is increased beyond that corresponding to curve *c*, the grid voltage curve will not

[1] HULL, L. M., and CLAPP, J. K., *Proc. I.R.E.*, **17**, 252 (1929); ANDREW, V. J., *Proc. I.R.E.*, **19**, 1911 (1931).

intersect the line of triggering voltage in three cycles, but at some time in the fourth cycle, and the frequency ratio will jump from 6 to 8. As the control frequency is increased, therefore, the oscillation frequency will also increase over a certain range, above which the frequency ratio will jump to a higher value at which control will again be attained throughout a similar frequency range. Further analysis of Fig. 10-33 shows that the range of control frequency throughout which the oscillator "locks in" increases with amplitude of control voltage.

It can be seen from Fig. 10-33 that there must always be a whole number of cycles of control voltage per half cycle of oscillation, and so the frequency ratio for the circuit of Fig. 10-32*a* must be an even integer. In the circuit of Fig. 10-32*b*, however, the control voltage is applied to the two grids in phase opposition, and so the control voltage at the end of the negative half cycle must be of equal magnitude but opposite polarity to that at the beginning of the cycle in order to make the grid voltage the same in both tubes at the beginnings of their respective negative half cycles. There must be $(n - \frac{1}{2})$ cycles of control voltage per half cycle of oscillation, and hence the frequency ratio will be $2n - 1$, where n is any positive integer. The frequency ratio is therefore odd. In the circuit of Fig. 10-32*c*, the control voltage is applied to only one tube, and it is necessary only that the voltage at the beginning of the negative half cycle of grid voltage of that tube be the same in each succeeding cycle. This is true for any integral frequency ratio.

In a multivibrator based upon the circuit of Fig. 10-8, the control frequency may be applied to one or both control grids. In the van der Pol oscillator the control frequency is best applied to the suppressor grid.[1] In the circuit of Fig. 10-28 the control frequency may be applied to any electrode, preferably the control grid or the suppressor grid.

10-16. Analysis of Sine-wave Oscillators.—A given type of oscillator can usually be analyzed in several ways. A particular oscillator may, for instance, be analyzed on the basis of negative resistance or on the basis of feedback. If made on the basis of feedback, the analysis may make use of the series expansion for electrode current or, if one is not primarily interested in studying harmonic generation, a simplified analysis may be based upon the equivalent electrode circuit. When the equivalent electrode circuit is used, the circuit equations may be written in differential form, or complex notation may be employed. Still another method of analysis is to consider an oscillator as an amplifier in which the output voltage is impressed upon the input and to determine the conditions under which the output voltage is equal in phase and magnitude to the input voltage. Choice of method in the analysis of a particular circuit is governed by the form of the circuit, by the preference

[1] PAGE and CURTIS, *loc. cit.*

of the person who makes the analysis, and by the amount of information that is desired concerning the performance of the oscillator. Usually a knowledge of the frequency of oscillation and the criterion of oscillation is all that is necessary, but it may be important to know under what conditions the oscillations are essentially sinusoidal. For purposes of illustration, various methods will be used in subsequent sections in analyzing different circuits. Except at very high radio frequencies, the circuit tuning capacitances are usually so much larger than the interelectrode capacitances that the latter may be neglected in an approximate analysis.

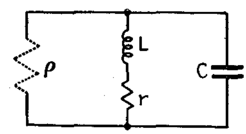


FIG. 10-34.—Basic negative-resistance oscillator circuit.

10-17. Negative-resistance Oscillators.[1]—The distinction between negative-resistance oscillators and feedback oscillators is in a sense artificial. A mathematical analysis of feedback oscillators yields as the criterion for sustained oscillators that the sum of certain parameters shall be equal to or less than zero. This criterion may always be written in such a form that one term is the positive circuit resistance. The other terms must, therefore, also have the dimensions of resistance, and at least one of them must be numerically negative in order that their sum shall be equal to or less than zero. One may say, therefore, that sustained oscillations can result if the tube in conjunction with the circuit can produce an equivalent negative resistance. Certain devices exhibit a negative a-c resistance without the action of an additional circuit, whereas the equivalent negative resistance in most feedback oscillators is dependent upon the action of the circuit. Under the first classification of the above list will be considered only those oscillators in which the negative resistance is inherent in the tube, or in the tube and associated resistors and capacitors, and does not require the presence of the tuned circuit.

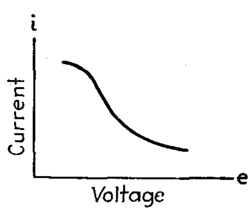


FIG. 10-35.—Current-voltage characteristic of a negative-resistance element.

Mathematical analysis shows that sustained oscillations may be set up if either a parallel or a series combination of inductance and capacitance is connected across a device that exhibits negative a-c resistance. The dotted resistance of the circuit of Fig. 10-34 represents any device that has a negative a-c resistance. L, r, and C represent the inductance, resistance, and capacitance of the resonant *tank* circuit. (Losses in the condenser are assumed to be negligible.) Although the a-c resistance, as defined by the reciprocal of the slope of the current-

[1] For a bibliography of 55 items on this subject, see E. W. HEROLD, *Proc. I.R.E.*, **23**, 1201 (1935).

voltage curve of Fig. 10-35, is not constant, it may be assumed to be constant in the range of operation if the current variations are not too large. Then the negative resistance is

$$\rho = \frac{de}{di} = \text{const.} \tag{10-2}$$

Application of Kirchhoff's laws to the circuit of Fig. 10-34 gives the following differential equation for the current through the inductance:

$$\frac{d^2i}{dt^2} + \left(\frac{r}{L} + \frac{1}{\rho C}\right)\frac{di}{dt} + \frac{r+\rho}{L\rho C}\,i = 0 \tag{10-3}$$

Solution of Eq. (10-3) gives the following expression for any of the circuit currents:[1]

$$i = A\epsilon^{-\frac{1}{2}(r/L+1/\rho C)t}\sin(\omega t + \alpha) \tag{10-4}$$

in which A and α are constants, the values of which are different for the currents in the three branches of the circuit, and the angular frequency of oscillation is

$$\omega = \sqrt{\frac{r+\rho}{\rho}\frac{1}{LC} - \frac{1}{4}\left(\frac{1}{\rho C} + \frac{r}{L}\right)^2} \tag{10-4A}$$

If ω is a real quantity, Eq. (10-4) shows that the circuit currents are sinusoidal and that their amplitude may decrease, remain constant, or increase. If the quantity $(r/L + 1/\rho C)$ is positive, the exponential factor in Eq. (10-4) decreases with time, and oscillation started in any manner eventually ceases. If the quantity $(r/L + 1/\rho C)$ is negative, on the other hand, the exponential factor increases with time and the amplitude builds up. In the critical case in which $(r/L + 1/\rho C)$ is equal to zero, the exponential factor is unity, indicating that the amplitude of oscillation remains constant. $(r/L + 1/\rho C)$ may be negative or zero if either r or ρ is negative. Hence, sustained oscillation of increasing or constant amplitude may be obtained if either r or ρ in Fig. 10-33 is a circuit element having negative a-c resistance in at least a portion of its operating range. The characteristics of negative-resistance circuit elements suitable for use in oscillators are such, however, that the equivalent oscillator circuits are necessarily of the form in which ρ is negative and r positive.

[1] COHEN, A., "Differential Equations," Chap. VII, D. C. Heath & Company, Boston, 1906. The student who is unfamiliar with the solution of differential equations can show that Eq. (10-4) is a solution of Eq. (10-3) by substituting Eqs. (10-4) and (10-4A) in Eq. (10-3) and noting that an identity results.

If ω is imaginary, Eq. (10-4) can be conveniently written in the form

$$i = \frac{A}{2j} \epsilon^{-\frac{1}{2}(r/L+1/\rho C)t} \left[\epsilon^{j(\omega t+\alpha)} - \epsilon^{-j(\omega t+\alpha)}\right] \tag{10-5}$$

or

$$i = \epsilon^{-\frac{1}{2}(r/L+1/\rho C)t}[B\epsilon^{\omega' t} + B'\epsilon^{-\omega' t}] \tag{10-6}$$

in which ω' is a real quantity.

$$\omega' = j\omega = \sqrt{\tfrac{1}{4}\left(\frac{1}{\rho C} + \frac{r}{L}\right)^2 - \frac{r+\rho}{\rho}\frac{1}{LC}} \tag{10-7}$$

The second term within the brackets of Eq. (10-6) decreases with time and eventually becomes negligible in comparison with the first, which increases with time. The bracketed factor, therefore, increases exponentially with time. If $(r/L + 1/\rho C)$ is zero or negative, the first factor is constant or increases exponentially, and so the product increases exponentially with time. Theoretically, therefore, if ρ is negative and of proper magnitude to make $(r/L + 1/\rho C)$ zero or negative, the current, once it is initiated in any manner, will continue to rise. Practically, the current cannot continue to rise indefinitely. Examination of Fig. 10-35 shows that, as the current increases, the magnitude of the negative resistance increases and that it eventually becomes positive. At some value of current $(r/L + 1/\rho C)$ becomes positive and the current either assumes an equilibrium value or starts to decrease. The action of the condenser and inductance is such as to make the current decrease and finally to reverse. The result is a nonsinusoidal sustained oscillation, such as that shown by the oscillogram of Fig. 10-36. It was proved in Sec. 10-1, however, that, if a negative and a positive resistance are connected in series, as they are in the left loop of the circuit of Fig. 10-34, triggering may occur if the magnitude of the positive resistance exceeds that of the negative resistance. The nonsinusoidal oscillation is then a relaxation oscillation of the type discussed in Secs. 10-10 to 10-15. (The analysis indicates that triggering is impossible for values of r and ρ at which sinusoidal oscillation can occur.)

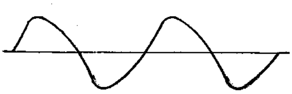


FIG. 10-36.—Typical nonsinusoidal voltage that may be generated by a negative-resistance oscillator.

If $(r/L + 1/\rho C)$ is positive, the first factor of Eq. (10-6) decreases with time. The product of the two factors first increases exponentially, passes through a maximum, and eventually approaches zero exponentially. This indicates that the current has the form of an exponential pulse. If the magnitude of the resistance of the positive-resistance element exceeds that of the negative-resistance element, the pulse may be accompanied by triggering. It may, in fact, be initiated by triggering.

The results of the foregoing analysis may be summarized as follows:

A. Sinusoidal oscillation may occur if ω is real, *i.e.*, if $|\rho| > L/(rC - 2\sqrt{LC})$.
  1. The amplitude of oscillation decreases with time if $(r/L + 1/\rho C)$ is positive, *i.e.*, if $|\rho| > L/rC$.
  2. The amplitude is constant if $|\rho| = L/rC$.
  3. The amplitude increases with time if $|\rho| < L/rC$.

B. The current is exponential in form if ω is imaginary, *i.e.*, if

$$|\rho| < \frac{L}{(rC - 2\sqrt{LC})}.$$

  1. Sustained oscillation occurs if $|\rho| \leqq L/rC$.
     a. The current wave is continuous, but nonsinusoidal if $|\rho| > r$.
     b. Triggering occurs and the wave has discontinuities, *i.e.*, relaxation oscillation takes place, if $|\rho| < r$.
  2. The current is an exponential pulse if $|\rho| > L/rC$.
     a. Triggering does not occur if $|\rho| > r$.
     b. Triggering occurs if $|\rho| < r$.

This summary is shown in graphical form in Fig. 10-37. Although similar analyses will not be made for feedback oscillators, such analyses

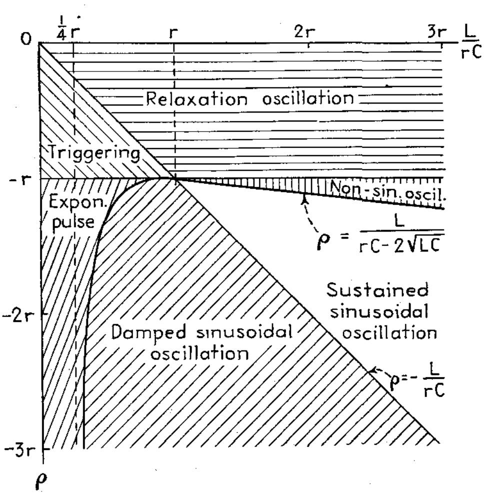


FIG. 10-37.—Theoretical diagram showing the relative ranges of circuit parameters of a negative-resistance oscillator throughout which the various types of operation should occur.

may be made. These analyses and their graphical summaries are of considerable instructional value because they show clearly that several types of operation are possible in a given circuit and that care is necessary in the design and adjustment of the circuit in order to ensure sinusoidal output from the oscillator. Figure 10-37 is checked in its general aspects by experimental investigations. It is found, for instance, that a large value of L/rC resulting from the use of insufficient capacitance causes the currents to be badly distorted from sinusoidal form. It should be noted, however, that the abrupt transitions from one type of operation to another indicated by Fig. 10-37 are not observed experimentally, since the magnitude of the negative resistance is in itself a function of the current.

10-18. Amplitude of Oscillation of Negative-resistance Oscillators.—Practical negative-resistance oscillators are of such form that r is positive and ρ negative. In order that sinusoidal oscillation may start in such a circuit, $|\rho_o|$, the static value of $|\rho|$ at the operating point, must be less than L/rC. Although Eq. (10-4) indicates that the amplitude continues to increase with time, actually an equilibrium amplitude is reached. This is true because curvature of the current-voltage characteristic eventually causes $|\bar{\rho}|$, the dynamic or average value of $|\rho|$ during the cycle, to increase with amplitude. A number of investigators have shown that equilibrium is established when $|\bar{\rho}|$ becomes equal to L/rC, if further increase of amplitude would cause $|\bar{\rho}|$ to exceed L/rC.[1] Brunetti has proved from energy relations that $\bar{\rho}$ at a given amplitude may be measured either by a bridge (see Sec. 15-34), or by taking the ratio of the amplitude of a sinusoidal applied voltage of given amplitude to the amplitude of the fundamental component of the resulting current through the negative resistance element.[2] The current produced by a given voltage may be measured, or determined graphically from the characteristic curve of the negative-resistance element by the methods discussed in Sec. 4-12.

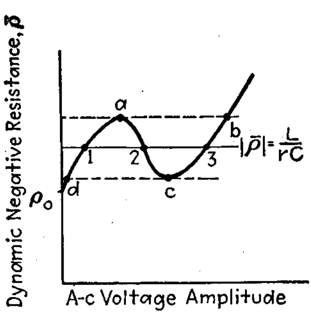


FIG. 10-38.—Typical curve of dynamic negative resistance as a function of alternating voltage amplitude.

The equilibrium amplitude of oscillation may be predicted by plotting a curve of $|\bar{\rho}|$ against amplitude of sinusoidal voltage applied to the negative resistance element at given operating voltages, as shown in Fig. 10-38.[2] Possible equilibrium amplitudes are indicated by the intersections of the curve with the line $|\bar{\rho}| = L/rC$. The values so determined are stable if further increase in amplitude results in increase of $|\bar{\rho}|$. Values corresponding to points 1 and 3 of Fig. 10-38 are seen by inspection to be stable. The value at point 2 is unstable, since increase of amplitude results in decrease of $|\bar{\rho}|$ and thus in further increase of amplitude. In the example given, oscillation starts when L/rC exceeds $|\rho_o|$. Further increase

[1] For bibliography, see C. BRUNETTI, *Proc. I.R.E.*, **25**, 1595 (1937).

[2] BRUNETTI, *loc. cit.*

of L/rC causes the amplitude to increase continuously until the intersection reaches a, from which the amplitude increases abruptly to the value corresponding to b. Decrease of L/rC then causes continuous decrease of amplitude until the intersection reaches c, from which it drops abruptly to the value corresponding to point d. Such abrupt changes of amplitude are often observed.[1]

10-19. Harmonic Content.—When curvature of the current-voltage characteristic is considered, the differential equation for the circuit of Fig. 10-33 becomes much more complicated, and the solution has the form of an infinite series, indicating the presence of harmonics in the current. Thus, although curvature of the characteristic has the desirable effect of limiting the amplitude of oscillation, it also introduces harmonics of the fundamental frequency of oscillation into the current. In order to keep the harmonic content small, it is necessary that the equilibrium

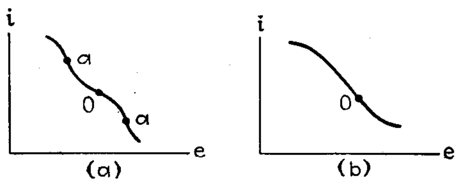


FIG. 10-39.—Two types of negative-resistance characteristics.

amplitude shall be low and that curvature of the characteristic shall be small throughout the range of operation.

The type of characteristic curve and the location of the operating point that will make it possible to satisfy these requirements can be determined from a study of the curves of Fig. 10-39. If the characteristic is of the form shown in Fig. 10-39a and the operating point is at O, increase of amplitude from a low value first causes a decrease in $|\bar{\rho}|$ (indicated by an increase of average slope over the range of operation) and hence a further rise in amplitude. When the amplitude becomes sufficiently great to extend the path of operation beyond the points of inflection at a, then $|\bar{\rho}|$ increases with further increase of amplitude and finally becomes so high that $|\bar{\rho}| = L/rC$ and the amplitude becomes constant. The equilibrium amplitude is relatively large, and the characteristic has appreciable curvature in the range of operation. Therefore, the requirements for low harmonic content are not satisfied. If the characteristic and operating point are as shown in Fig. 10-39b, on the other hand, then $|\bar{\rho}|$ increases continuously with increase of amplitude from a small value. By making L/rC only slightly greater than $|\rho_o|$, the reciprocal of the slope of the characteristic at the operating point, the amplitude of oscillation may be made as small as desired. Since the

[1] For bibliography, see BRUNETTI, *loc. cit.*

curvature is also small in the operating range, low harmonic content may be attained. Fortunately it is possible to use negative-resistance elements that have characteristic curves of the form of Fig. 10-39*b*, so that negative-resistance oscillators of very small harmonic content may be designed. The use of high-Q oscillatory circuits is also favorable to low harmonic content.

10-20. Frequency of Oscillation.—Under the threshold condition, when $|\rho| = L/rC$, Eq. (10-4*A*) reduces to

$$\omega = \sqrt{\frac{r + \rho}{\rho} \frac{1}{LC}} \tag{10-8}$$

Inasmuch as r, the a-c resistance of the inductance coil and leads, is ordinarily only a few ohms, whereas the negative resistance $|\rho|$ is seldom less than 2 or 3 thousand ohms, the frequency is practically equal to $1/2\pi\sqrt{LC}$, and small changes in ρ, such as might result from the variation of battery voltages, have a negligible effect upon the oscillation frequency if the threshold condition can be maintained.

10-21. Series Oscillatory Circuit.—For the circuit in which the negative resistance is shunted by a series combination of L, r, and C, the mathematical analysis gives the following expression for the angular frequency of oscillation:

$$\omega = \sqrt{\frac{1}{LC} - \left(\frac{r + \rho}{2L}\right)^2} \tag{10-9}$$

The oscillation is sinusoidal when $|r + \rho| < \sqrt{L/C}$. In order for sustained oscillation to be possible, ρ must be negative, and its absolute value must be equal to or greater than r. Under the threshold condition at which oscillation will just start, the magnitude of ρ is equal to that of r, and Eq. (10-9) reduces to

$$\omega = \frac{1}{\sqrt{LC}} \tag{10-10}$$

The oscillation is then necessarily sinusoidal.

Because of the necessity of providing a d-c path of low resistance, it is usually impossible to use the series type of circuit when the negative-resistance element consists wholly or in part of vacuum tubes. Although oscillation may be obtained in circuits that are presumably of the series type, careful investigation usually discloses that the circuit is actually oscillating as a parallel circuit of the form of Fig. 10-34, the distributed capacitance of the inductance coil serving as the shunting capacitance C. The series capacitance is found to have no effect upon the frequency.

10-22. The Dynatron Oscillator.—One cause of negative resistance in high-vacuum tubes is secondary emission. An excellent example of a negative-resistance characteristic resulting from secondary emission is shown by the plate characteristics of the old type 24A tube at low plate voltages. These characteristics were discussed in detail in Chap. 3 (see Fig. 3-9). The newer type of 24A tube also has negative plate resistance at plate voltages lower than the screen voltage, but the magnitude of the resistance is considerably higher than in the older tubes. The grid characteristics of many tubes also have portions with a negative slope. The possibility of using negative resistance resulting from secondary emission in producing oscillations was first shown by Hull, who termed this type of oscillator the *dynatron oscillator*.[1] The dynatron has since been studied theoretically and experimentally by many investigators, and much has been written on the subject.[2] A typical dynatron circuit is shown in Fig. 10-40. The tuned circuit can also be connected in the screen-grid lead, instead of in the plate lead.[3]

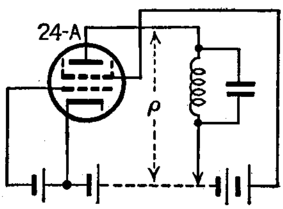


FIG. 10-40.—Dynatron-oscillator circuit.

The negative-resistance characteristics of tetrodes approximate the desirable shape shown in Fig. 10-39*b*. Therefore by proper choice of operating point the amplitude of oscillation and the curvature in the operating range may be kept small, and the harmonic content low. Other advantages of the dynatron for laboratory applications are its good frequency stability and its simplicity. Because only one inductance is required, it is a simple matter to change from one frequency band to another. The amplitude of oscillation may be controlled by means of the control-grid voltage, which varies the slope of the characteristic in the negative-resistance range, and hence the value of $\bar{\rho}$. Disadvantages of this type of oscillator result mainly from its dependence upon secondary emission. Secondary emission changes with use of the tube, and large differences are observed in the shapes of the characteristics of individual tubes of the same type.

[1] HULL, A. W., *Proc. I.R.E.*, **6**, 535 (1918).

[2] SCROGGIE, M. G., *Wireless Eng.*, **10**, 527 (1933) (with bibliography of 35 items); MOULLIN, E. B., *J. Inst. Elec. Eng. (London)*, **73**, 186 (1933); BAKER, G. B., *J. Inst. Elec. Eng. (London)*, **73**, 196 (1933); COLEBROOK, F. M., *Wireless Eng.*, **10**, 663 (1933); HAYASI, TATUO, *J. Inst. Elec. Eng. Japan*, **53**, 389 (1933); CLARKE, G. F., *Wireless Eng.*, **11**, 75 (1934); COLEBROOK, F. M., *Radio Research Special Rept.* 131 (England); GROSZKOWSKI, J., *Wireless Eng.*, **11**, 193 (1934); GAGER, F. M., *Proc. I.R.E.*, **23**, 1048 (1935); GAGER, F. M., and RUSSELL, J. B., *Proc. I.R.E.*, **23**, 1536 (1935); HOULDIN, J. E., *Wireless Eng.*, **14**, 422 (1937).

[3] HAYASI, TATUO, *Proc. I.R.E.*, **22**, 751 (1934).

10-23. Negative-transconductance Oscillator.—The objectionable features associated with secondary emission are avoided in another type of negative-resistance oscillator, the action of which depends upon the fact that negative voltage applied to the suppressor of a pentode such as the 6J7 or 6K7 causes electrons that have passed through the screen grid to be returned to the screen grid.[1] Over a certain range a positive increment of suppressor voltage (a decrease of negative voltage) allows more electrons to go to the plate and thus decreases the screen current, which means that the suppressor-screen transconductance is negative. Under proper operating conditions the screen current decreases with a positive increment of suppressor voltage even when the screen voltage is given an equal increment. By the use of the circuit of Fig. 10-41 the negative transconductance can be used to produce a negative resistance.

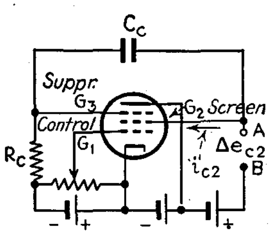


FIG. 10-41.—The use of a pentode to produce negative resistance.

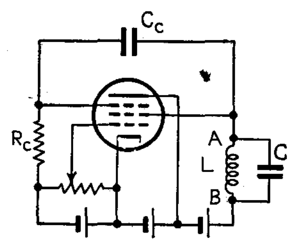


FIG. 10-42.—Transitron oscillator.

If C_c and R_c are large enough so that a change in voltage of the screen grid G_2 is accompanied by a practically equal change in voltage of the suppressor grid G_3, the action is as follows: An increase Δe_{c2} in the screen voltage is accompanied by an equal change Δe_{c3} of suppressor voltage. Δe_{c2} would by itself change i_{c2} by the amount $\Delta e_{c2}/r_{g2}$, and Δe_{c3} acting alone would change i_{c2} by the amount $\Delta e_{c3}g_{32}$. If the variation of r_{g2} with e_{c3} and of g_{32} with e_{c2} is small (*i.e.*, if the curves of i_{c2} *vs.* e_{c2} for various values of e_{c3} are practically straight, parallel, and equidistant over the range under consideration), it may be assumed that the net change of i_{c2} will be the sum of the two changes.

$$\Delta i_{c2} = \frac{\Delta e_{c2}}{r_{g2}} + \Delta e_{c3}g_{32} = \Delta e_{c2}\left(\frac{1}{r_{g2}} + g_{32}\right) \qquad (10\text{-}11)$$

Since g_{32} has been shown to be negative, it follows that, if the magnitude of g_{32} exceeds the magnitude of $1/r_{g2}$, an increase of screen voltage is accompanied by a decrease of screen current. When R_c is so large in comparison with r_{g2} that the current through R_c and C_c may be neglected,

[1] HEROLD, *loc. cit.*

this means that between points A and B the circuit will exhibit negative resistance of magnitude (see also Prob. 4-4).

$$|\rho| = \frac{r_{g2}}{1 + r_{g2}g_{32}} = \frac{r_{g2}}{1 + \mu_{23}} \tag{10-12}$$

If a parallel resonant circuit whose parameters satisfy the criterion for sustained oscillation is inserted between A and B, as in Fig. 10-42, sustained oscillations result.

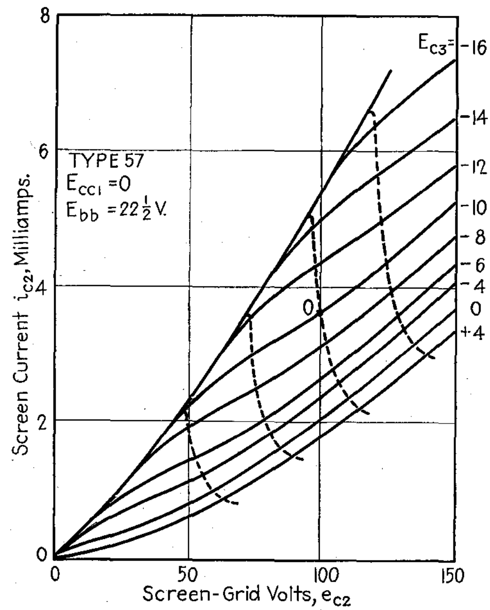


FIG. 10-43.—Static family of screen characteristics for a type 57 tube. The dotted lines show dynamic screen characteristics for the circuit of Fig. 10-41, derived under the assumption that changes of screen and suppressor voltage are equal.

Dynamic curves of i_{c2} *vs.* e_{c2} for the circuit of Fig. 10-41 may be readily derived from the static family of i_{c2}-e_{c2} characteristics.[1] Thus, in Fig. 10-43, the dotted dynamic i_{c2}-e_{c2} curves are obtained by connecting points for which e_{c2} and e_{c3} are changed by equal amounts from their static operating values. If, for instance, the operating values of e_{c2} and e_{c3} are 100 volts and $-$ 10 volts, respectively, then a 2-volt increase of e_{c2} to 102 volts is accompanied by a 2-volt increase of e_{c3} to $-$ 8 volts, and the resulting current is 3.2 ma. Other points are located in a similar manner. If C_c and R_c are so small that the reactance of C_c is appreciable in comparison with the resistance of R_c at the desired frequency of oscilla-

[1] These curves should not be confused with curves of dynamic (average) negative resistance.

tion, the voltage-dividing action of C_c and R_c causes the change of e_{c3} to be less than that of e_{c2}. The dynamic curve is less steep, indicating a higher negative resistance. For this reason C_c and R_c should be large enough so that the coupling-condenser reactance at the lowest desired frequency of oscillation is small compared with the resistance of the grid resistor.[1]

The pentode circuit of Fig. 10-42 may be modified to use a tetrode such as the 24A by the omission of the first grid G_1. Oscillation can be obtained, however, only at low cathode temperature. The reason for this is that at normal temperature the resistance of the control grid, which is used in place of G_2 of Fig. 10-41, is so low that $1/r_{g2}$ exceeds g_{32} in Eq. (10-12), and ρ is positive, even though g_{32} is negative. Use of the negative inner grid G_1 has the same effect as lowering the cathode temperature and has the additional advantage that the voltage of the first grid may be used to control the negative resistance and thus the amplitude of oscillation.

The negative-transconductance oscillator is not dependent upon secondary emission and therefore does not have the objectionable features of the dynatron oscillator. The circuit is slightly more complicated.

10-24. Push-pull Negative-resistance Oscillator.—A two-tube circuit which can be used as the basis of a negative-resistance oscillator is shown in Fig. 10-44.[2] Although the equivalent plate circuit may be readily used to prove that a negative resistance may exist between points A and B (see Prob. 4-5), it is instructive to analyze the physical action of the circuit. k represents the fraction of r_b from which voltage is applied to the grid of the opposite tube. On the assumption that the two tubes and their supply voltages are in all respects the same, the action is as follows: When the voltage Δe is zero, the points A and B are at the same potential, the two plate currents are equal, and no current flows through the external branch of the circuit. When Δe is applied, a current Δi flows in the external branch. It is composed of two components: Δi_r, flowing through the plate resistors r_b; and $\Delta i_{b2} = -\Delta i_{b1}$,

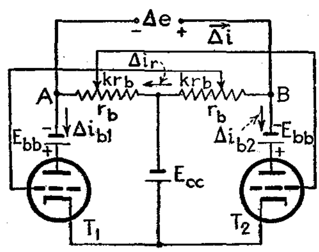


FIG. 10-44.—The use of triodes in a push-pull circuit to produce negative resistance.

[1] See also Prob. 10-3, p. 414.

[2] REICH, H. J., *Proc. I.R.E.*, **25**, 1387 (1937). See also L. B. TURNER, *Radio Rev.*, **1**, 317 (1920).

It is interesting to note that the circuit of Fig. 10-44 is identical with the Eccles-Jordan trigger circuit of Fig. 10-5, that the dynatron oscillator is based upon the same circuit as the tetrode trigger circuit mentioned in Sec. 10-2, and that the negative-transconductance oscillator is based upon the same circuit as the pentode trigger circuit of Fig. 10-4.

circulating through the tubes. The resistor current is

$$\Delta i_r = \frac{\Delta e}{2r_b} \tag{10-13}$$

Δe raises the voltage of the plate of tube 2 by $\Delta e/2$ volts and lowers the voltage of the grid of tube 2 by $k\Delta e/2$ volts; the changes of plate and grid voltage of tube 1 are equal in magnitude to those of tube 2 but opposite in sign. On the assumption that Δe is small enough so that the plate resistance and transconductance, r_p and g_m, are practically constant, the changes in plate current are

$$\Delta i_{b2} = \frac{\Delta e_b}{r_p} + g_m \Delta e_c = -\Delta i_{b1} \tag{10-14}$$

$$\Delta i_{b2} = \tfrac{1}{2}\Delta e\left(\frac{1}{r_p} - kg_m\right) \tag{10-15}$$

The total change in i is

$$\Delta i = \Delta i_r + \Delta i_{b2} = \tfrac{1}{2}\Delta e\left(\frac{1}{r_b} + \frac{1 - \mu k}{r_p}\right) \tag{10-16}$$

$$\Delta i = \tfrac{1}{2}\Delta e\,\frac{r_p + r_b(1 - \mu k)}{r_p r_b} \tag{10-17}$$

$$\rho = \frac{\Delta e}{\Delta i} = \frac{2r_p r_b}{r_p + r_b(1 - \mu k)} = \frac{2r_p}{\dfrac{r_p + r_b}{r_b} - \mu k} \tag{10-18}$$

Equation (10-18) shows that, if the magnitude of μk exceeds the magnitude of $(r_p + r_b)/r_b$, the resistance between A and B is negative. If r_b is large in comparison with r_p, the criterion for negative resistance is simply that the product μk shall be greater than unity. The negative resistance of the circuit results from the amplifying property of the tubes. If μk is sufficiently large, the change in plate current exceeds the change in current through the external resistors and is opposite in direction, causing a net external current opposite in direction to the applied voltage.

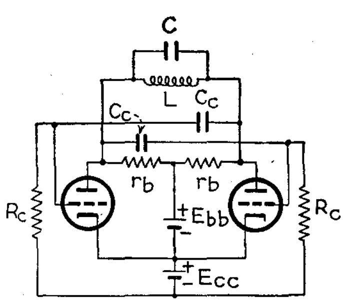


FIG. 10-45.—Push-pull negative-resistance oscillator.

Sustained oscillations are produced if a parallel resonant circuit of low resistance is connected between A and B. The necessity for two B-supply voltages can be eliminated by using condensers and grid leaks to couple the grids to the plates, as shown in Fig. 10-45. k is unity in this circuit but may be reduced for the purpose of controlling the amplitude of oscillation. The amplitude may be controlled more readily, however, by means of the grid bias. A twin triode, such as the 53 or the 6SN7GT, may be used in this circuit in place of two separate triodes.

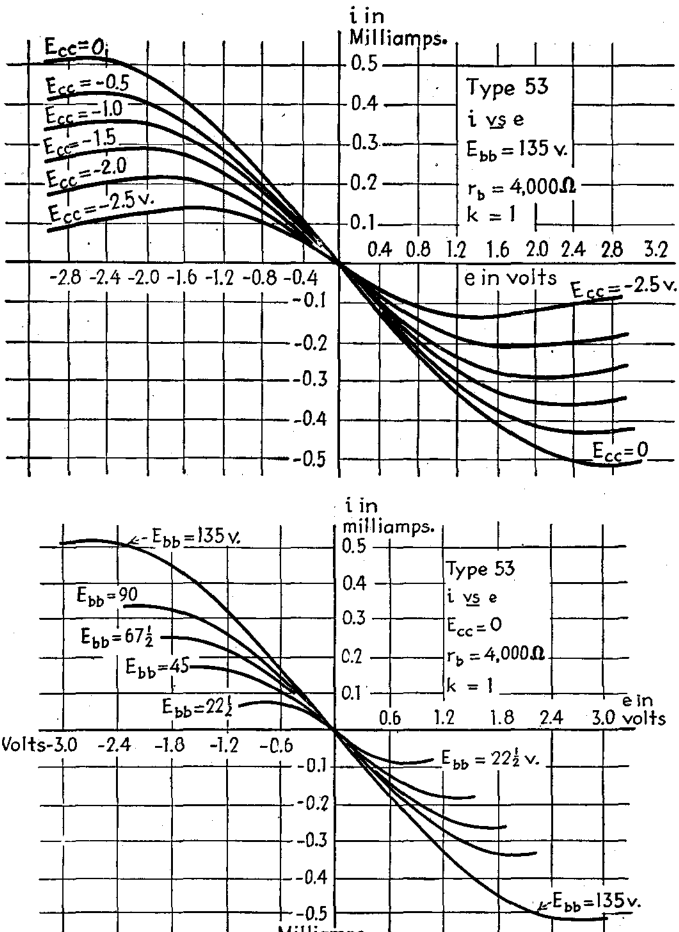


FIG. 10-46.—Current-voltage characteristics for type 53 tubes used in the circuit of Fig. 10-44.

The reactance of the coupling condensers of Fig. 10-45 should be small in comparison with the grid resistors at the lowest desired frequency of oscillation. When this is true, it does not matter whether the resonant circuit is connected between the two plates or between the two grids.

In Fig. 10-46 are shown curves of external current *i vs.* voltage *e*, between points *A* and *B* of the circuit of Fig. 10-44 for a type 53 twin triode when k is unity. It is apparent that these curves have the desirable form of the curve of Fig. 10-39*b*. The amplitude and harmonic content may, therefore, be made as small as desired by adjusting the circuit parameters or by adjusting ρ by means of k or the supply voltages. This can also be seen from Fig. 10-47, which shows a typical curve of dynamic negative resistance *vs.* voltage amplitude derived from Fig. 10-46 by use of Eqs. (4-35). The form of this curve is such that the line $|\bar{\rho}| = L/rC$ can intersect the curve at only one point, the voltage of which approaches zero as L/rC approaches $|\rho_o|$.

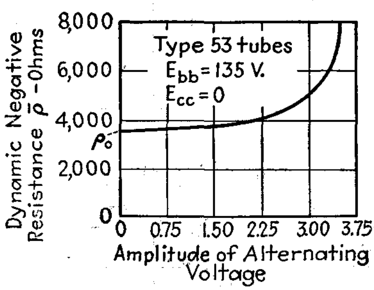


FIG. 10-47.—Typical curve of dynamic negative resistance *vs.* amplitude of alternating voltage derived from Fig. 10-47 by use of Eqs. 4-35.

The push-pull negative-resistance oscillator does not have the disadvantages of the dynatron oscillator and gives even lower harmonic content and higher frequency stability than either the dynatron or negative-transconductance oscillators. These advantages are partly offset by increased circuit complication.

10-25. Negative-grid-resistance Oscillator (Tuned-grid with Capacitive Feedback).—In Sec. 4-3 it was shown that, as the result of grid-plate electrode capacitance, the input conductance of a vacuum tube may be negative when the plate load is inductive. When this is true, sustained oscillation may be set up in a parallel resonant circuit connected between the grid and the cathode. Since the magnitude of the input conductance decreases with frequency, this type of oscillator can be used only at radio frequency. The variometer-controlled regenerative detectors which were popular in the early days of radio reception often proved to be excellent examples of this type of oscillator. The negative-grid-resistance oscillator may also be properly considered as a feedback oscillator.

10-26. Resistance-capacitance–tuned Negative-resistance Oscillators.—Sine-wave oscillators that do not require the use of inductance can be readily analyzed both as negative-resistance oscillators and as feedback oscillators. Since the feedback analysis is somewhat simpler, resistance-capacitance–tuned oscillators will be treated in Sec. 10-39.

10-27. Method of Coupling to Negative-resistance Oscillators.—In using the negative-resistance oscillators of Figs. 10-40, 10-42, and 10-45, the output voltage should be taken from across the resonant tank circuit or, preferably, from a separate coil coupled to the tank inductance L. When appreciable power is required or when the impedance of the load is such as to affect the frequency of oscillation if the load is connected across or coupled to the tuned circuit, the output should be taken from an amplifier excited by the oscillator. Such an amplifier is called a *buffer* amplifier.

10-28. Feedback Oscillators.—A feedback oscillator can be considered as a tuned feedback amplifier in which the amplitude and phase angle of the feedback voltage are such as to cause oscillation (see Sec. 6-35).[1] Suppose that a voltage e_i is applied to the input of an amplifier and that the resulting output voltage is e_o. If a portion e_f of the output voltage is applied to the input, in addition to e_i and in phase with it, this feedback voltage will act in the same manner as e_i. If the magnitude of the feedback voltage is exactly equal to e_i, it can replace e_i, and the amplifier will continue to deliver the original output e_o, if e_i is removed. In other words, the amplifier will oscillate at constant amplitude. If the feedback is increased, the amplitude will build up; if it is decreased, the amplitude will die down. Among the simpler types of feedback oscillators are the tuned-plate oscillator of Fig. 10-48 and the tuned-grid

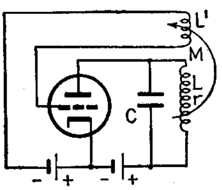


FIG. 10-48.—Tuned-plate oscillator.

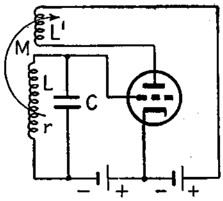


FIG. 10-49.—Tuned-grid oscillator.

oscillator of Fig. 10-49, in which a portion of the voltage developed in the plate circuit is introduced into the grid circuit by means of magnetic coupling.

Feedback oscillators are usually analyzed by use of the equivalent plate circuit, even though the amplitude of oscillation may be so high that the path of operation is far from linear. Although this method of analysis gives no indication of the production of harmonics, it yields considerable valuable information concerning the fundamental frequency of oscillation and the conditions that must be satisfied in order that sustained oscillations may be produced. In order to simplify the analysis, it is customary to neglect the effect of grid current. Although grid current cannot be neglected in a rigorous treatment, it is usually satisfactory to make use of the simpler analysis and to bear in mind that the results must be modified to take grid current into account. Losses resulting from the flow of grid current have the same effects upon the operation of the circuit as an equivalent loss in the tuned circuit. Reduction of plate current resulting from diversion of electrons to the grid increases the curvature of the dynamic transfer characteristic and thus

[1] For an analysis of phase relations in oscillators, see C. K. Jen, *Proc. I.R.E.*, **19**, 2109 (1931).

tends to increase harmonic content. Because the grid current is a nonsinusoidal pulse, the harmonic content is also increased by the flow of grid current through the grid-circuit impedance. It will be seen, however, that advantage may be taken of the flow of grid current to limit the amplitude of oscillation by automatically increasing the grid bias.

10-29. Tuned-plate Oscillator.—Figure 10-50 shows the equivalent plate circuit for the tuned-plate oscillator under the assumption that condenser losses are negligible. The alternating grid voltage e_g is induced in the grid coil by virtue of magnetic coupling to the plate coil. If M is assumed to be positive when an increase of i results in a positive induced grid voltage,

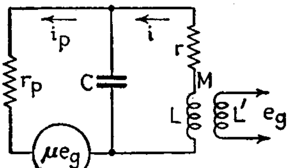


FIG. 10-50.—Equivalent plate circuit for the tuned-plate oscillator.

$$e_g = M \frac{di}{dt} \tag{10-19}$$

Application of Kirchhoff's laws to the equivalent circuit gives the following equations:

$$r_p i_p + ri + L \frac{di}{dt} = \mu e_g = \mu M \frac{di}{dt} \tag{10-20}$$

$$ri + L \frac{di}{dt} + \frac{1}{C} \int (i - i_p)\, dt = 0 \tag{10-21}$$

If Eq. (10-21) is differentiated and substituted in Eq. (10-20), the following second-order differential equation is obtained:

$$\frac{d^2 i}{dt^2} + \frac{di}{dt}\left(\frac{r_p r C + L - \mu M}{r_p L C}\right) + i\, \frac{r_p + r}{r_p L C} = 0 \tag{10-22}$$

This equation is of the form

$$\frac{d^2 i}{dt^2} + a \frac{di}{dt} + bi = 0 \tag{10-23}$$

the solution of which is[1]

$$i = A\epsilon^{-at/2} \sin\left(\sqrt{b - \tfrac{1}{4}a^2}\, t + \theta\right) \tag{10-24}$$

In order that oscillations shall not die out, the exponential factor of Eq. (10-24) must be unity or must increase with time; *i.e.*, a must be equal to or less than zero. The value of a may be obtained from Eq. (10-22). Since r_p, C, and L are all positive quantities, the criterion for sustained oscillation is

$$r_p r C + L - \mu M \leqq 0 \tag{10-25}$$

[1] COHEN, *op. cit.*, Chap. VII.

This can be satisfied only if M is positive and if

$$|\mu M| \geqq r_p r C + L \tag{10-26}$$

or

$$|g_m| \geqq \left|\frac{rC}{M} + \frac{L}{Mr_p}\right| \tag{10-27}$$

It follows from Eqs. (10-19) and (10-25) that the coupling not only must exceed a critical magnitude but must be of such sign as to result in a positive component of grid voltage when the current through the plate inductance is increasing.

Under the threshold condition, for which $r_p r C + L - \mu M = 0$, the frequency of oscillation is

$$f = \frac{1}{2\pi\sqrt{LC}}\sqrt{\frac{r + r_p}{r_p}} \tag{10-28}$$

When no power is being drawn from the oscillating circuit, r is small in comparison with r_p, and the frequency of oscillation is practically the natural frequency of the resonant circuit, $1/2\pi\sqrt{LC}$.

The ordinary complex method of solving the equivalent circuit is also conveniently used in the analysis of oscillators. The following equations may be obtained by summing the voltages in the equivalent circuit of Fig. 10-50:

$$I_p r_p + I(r + j\omega L) = \mu E_g = \mu j\omega M I \tag{10-29}$$

$$\frac{I_p}{j\omega C} - I\left(r + j\omega L + \frac{1}{j\omega C}\right) = 0 \tag{10-30}$$

The solution of these simultaneous equations gives the equation

$$j\omega(Crr_p + L - \mu M) - \omega^2 r_p LC + r_p + r = 0 \tag{10-31}$$

Equation (10-31) is an identity that can hold only if the real and the imaginary terms are individually zero. Equating the imaginary and the real terms to zero gives the criterion for oscillation [Eq. (10-25)] and the frequency of oscillation [Eq. (10-28)].

The significance of analyzing an oscillator circuit by solving the complex equations for the equivalent circuit becomes apparent when it is noted that the circuits of Figs. 10-48 and 10-50 may be considered to be those of a voltage amplifier, the output of which is taken from across L'. Since the complex equations (10-29) and (10-30) apply only to steady-state (constant-amplitude) values and since it is assumed in writing the equations that the vector input voltage E_g of the amplifier is equal to the vector output voltage $j\omega MI$, the solution of Eqs. (10-29) and (10-30) is equivalent to determining the conditions under which the output voltage is constant and equal in phase and magnitude to the input voltage. It should be noted that the complex method of analysis yields no information regarding conditions that must be satisfied in order that the oscillations shall be sinusoidal. This information may, however, be found from Eqs. (10-22) and (10-24) obtained in the differential-equation method of analysis.

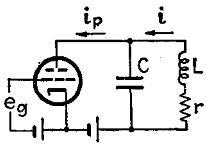


FIG. 10-51.—Plate circuit of tuned-plate oscillator.

In order to emphasize the fact that curvature of the tube characteristics results in the production of harmonics in the oscillator current, it is of interest to analyze the circuit by using the series expansion for plate current. Summation of alternating voltages in the CLr branch of the circuit of Fig. 10-51 gives the equation

$$L\frac{di}{dt} + ir + \frac{1}{C}\int (i - i_p)\,dt = 0 \tag{10-32}$$

or

$$L\frac{d^2i}{dt^2} + r\frac{di}{dt} + \frac{1}{C}(i - i_p) = 0 \tag{10-33}$$

But

$$i_p = a_1e_g + a_2e_g{}^2 + a_3e_g{}^3 + \cdots \tag{3-40}$$

$$= a_1M\frac{di}{dt} + a_2M^2\left(\frac{di}{dt}\right)^2 + a_3M^3\left(\frac{di}{dt}\right)^3 + \cdots \tag{10-34}$$

Substituting Eq. (10-34) in Eq. (10-33) gives

$$\frac{d^2i}{dt^2} + \frac{di}{dt}\left\{\frac{r}{L} - \frac{M}{LC}\left[a_1 + a_2M\frac{di}{dt} + a_3M^2\left(\frac{di}{dt}\right)^2 + \cdots\right]\right\} + \frac{i}{LC} = 0 \tag{10-35}$$

Equation (10-35) has a simple sinusoidal solution only when a_2 and coefficients of higher-order terms of the series are zero. The general solution is in the form of a series, which indicates the production of harmonics of the fundamental frequency.

10-30. Tuned-grid Oscillator.—An analysis similar to that which has been presented for the tuned-plate oscillator shows that the frequency of oscillation of the tuned-grid oscillator of Fig. 10-49 is

$$f = \frac{1}{2\pi\sqrt{C(Lr_p + L'r)/r_p}} \tag{10-36}$$

When no power is being drawn from the circuit, the term Lr_p greatly exceeds $L'r$, and the frequency is very nearly equal to $1/2\pi\sqrt{LC}$. Under this same approximation the criterion for oscillation is

$$g_m \geqq \frac{rC}{M} + \frac{M}{Lr_p} \tag{10-37}$$

The appearance of M in the denominator of one term of Eq. (10-37) and in the numerator of the other indicates that the required value of g_m is large for both high and low values of M and that there is a value of M for which the required transconductance is a minimum. Solution of Eq. (10-37) for M shows that oscillation takes place only when M lies in the range between

$$\frac{\mu L}{2} + \sqrt{\left(\frac{\mu L}{2}\right)^2 - rr_pLC} \qquad \text{and} \qquad \frac{\mu L}{2} - \sqrt{\left(\frac{\mu L}{2}\right)^2 - rr_pLC}$$

The limited range of M over which oscillation can occur is a practical disadvantage of the tuned-grid oscillator.

10-31. Hartley Oscillator.—The circuit of the Hartley oscillator is shown in Fig. 10-52. Analysis of the equivalent circuit shows that when no power is taken from the circuit the frequency of oscillation is very nearly equal to $1/2\pi\sqrt{LC}$, where L is the total inductance, $L_1 + L_2 + 2M$, of the two coils. The criterion for oscillation is (C_{gk} and C_{pk} neglected)

$$g_m \geqq \frac{\mu rCL}{(L_2 + M)[\mu(L_1 + M) - (L_2 + M)]} \tag{10-38}$$

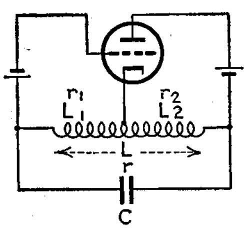


FIG. 10-52.—Hartley oscillator.

It is interesting to note that Eq. (10-38) may be satisfied when M is zero, indicating that capacitive coupling through C allows oscillation to take place even when there is no inductive coupling between the grid and plate circuits. The Hartley oscillator has long been a favorite circuit. One reason for this is that the criterion for oscillation is not critical. L_1 and L_2 are usually the two portions of a tapped coil, and the position of the tap may be used to control the amplitude of oscillation. The circuit will oscillate most readily when the ratio of L_2 to L_1 lies in the range from approximately 0.6 to 1, the ratio increasing with amplification factor. Because the tuning condenser shunts both the grid and plate coils, the Hartley oscillator gives a lower frequency for a given total inductance than either the tuned-plate or tuned-grid oscillators and is therefore particularly suitable for the production of low audio frequencies when the necessity for good wave form prevents the use of iron-core inductances.

10-32. Colpitts Oscillator.—Figure 10-53 shows the circuit of the Colpitts oscillator. The frequency of oscillation is nearly equal to

$1/2\pi\sqrt{LC}$, where $C = C_1C_2/(C_1 + C_2)$, the resultant capacitance of C_1 and C_2 in series. The capacitances of the isolating condensers, C_3 and C_4, and the inductances of the chokes are made large enough so that they do not appreciably affect the oscillation frequency. As a variable-frequency oscillator the Colpitts oscillator is less convenient to use than the tuned-plate, tuned-grid, or Hartley oscillators, owing to the necessity of varying both C_1 and C_2 in order to maintain oscillation. It has been found suitable for use in marine work, because the condensers may be sealed against moisture and the circuit tuned by means of a variometer type of variable inductance.

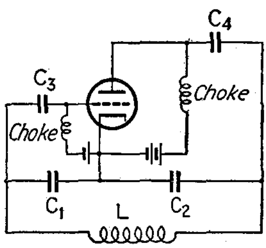


FIG. 10-53.—Colpitts oscillator.

10-33. Circuits Having More Than One Resonant Branch.—More complicated circuits than those which have been discussed are sometimes used. Among these are the Meissner circuit, shown in Fig. 10-54, and

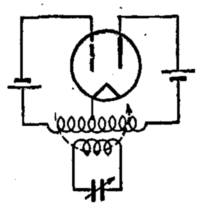


FIG. 10-54.—Meissner oscillator.

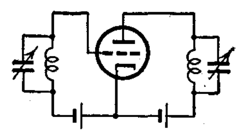


FIG. 10-55.—Tuned-grid–tuned-plate oscillator.

the tuned-grid–tuned-plate circuit of Fig. 10-55, in which feedback is through the grid-plate capacitance of the tube. The fact that both the grid and plate circuits are tuned in the latter circuit tends to improve the frequency stability and the wave form. A disadvantage of both of these circuits is that they may oscillate at two frequencies, and that the frequency may jump from one value to the other. Analysis shows that two frequencies of oscillation are to be expected when the circuit contains two resonant branches that are closely coupled. (The primary inductance of the Meissner circuit resonates with its distributed capacitance.) The two frequencies of oscillation are analogous to the two peaks of the resonance curve of a doubly tuned transformer (see Sec. 6-22).

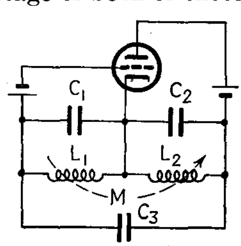


FIG. 10-56.—Generalized feedback oscillator circuit.

10-34. Generalized Feedback Oscillator Circuit.—Figure 10-56 shows a generalized form of feedback oscillator circuit from which the various types of singly and doubly tuned circuits may be derived by the elimination of one or more of the circuit elements or couplings. The interelectrode capacitances of the tube and the distributed winding capacitances may serve as the condensers of the circuit of Fig. 10-56 and those derived from it.

10-35. Push-pull Oscillators.—The basic feedback oscillator circuits may be modified to use two tubes in push-pull. As in amplifiers, the use of push-pull circuits increases the power output and decreases the harmonic content. The frequency stability of push-pull circuits is also higher than that of single-sided circuits. They are used principally at high and ultrahigh frequencies.

10-36. Figure of Merit of Oscillator Tubes.—The form of Eqs. (10-27), (10-37), and (10-38) shows clearly the importance of high transconductance in tubes used in feedback oscillators. These equations also show that, for a given transconductance, oscillation will take place more readily the higher the plate resistance. Since at a given transconductance the amplification factor is proportional to the plate resistance, it follows that the "figure of merit" of a tube for use in a feedback oscillator is the product μg_m. As high μg_m is also one requirement of a power tube, power tubes are in general good feedback oscillator tubes.

10-37. Series and Parallel Feed.—Circuits in which the plate supply voltage is connected in series with the plate inductance, as in Figs. 10-48, 10-49, 10-52, 10-54, and 10-55, are called *series-feed circuits*. In practice it may be desirable or necessary to connect the plate to the oscillating circuit through a condenser and to apply the direct plate voltage through a choke, the reactance of which is so high that it does not appreciably affect the oscillating circuit. Air-core chokes are used in radio-frequency circuits, iron-core chokes in audio-frequency circuits. Examples of this method of applying the direct plate voltage, called *parallel feed*, are given by the circuits of Figs. 10-53, 10-57, 10-69, 10-71, and 10-75. A disadvantage of parallel feed is that difficulty may be encountered in preventing parasitic oscillations in the choke circuits.

10-38. Use of Self-bias to Limit Amplitude of Oscillation.—For the purpose of simplicity, fixed biasing voltages have been indicated in the basic oscillator circuits. Fixed bias is rarely used, however, in practical oscillators. In order to prevent excessive distortion and to aid in obtaining frequency stabilization, it is necessary to limit the amplitude of oscillation. In feedback oscillators the criterion for oscillation involves the transconductance of the tube, and the amplitude builds up until the average (dynamic) transconductance (see Sec. 3-26) drops to the critical value below which oscillation cannot take place. Unfortunately, the average transconductance first increases with amplitude, and so the amplitude may increase to a high value before the average transconductance again falls sufficiently to result in equilibrium. If the circuit is adjusted to give small equilibrium amplitude, then it will not start of its own accord. As explained in Sec. 10-18, a similar difficulty may be experienced with some negative-resistance oscillators. This difficulty may be prevented by causing the grid bias to increase automatically with amplitude.

The most common method of limiting the amplitude of oscillation is the use of a grid-blocking condenser and grid leak, as shown in Fig. 10-57. The initial bias is zero but, as soon as oscillation commences, the grid is driven positive during a portion of the cycle and so electrons flow from the cathode to the grid. During the remainder of the cycle these electrons cannot return to the cathode but can only leak off the condenser and grid through the grid leak R_c. The trapped electrons make the potential of the grid negative with respect to the cathode, thus providing a bias.[1] The greater the amplitude of oscillation, the more positive the grid swings, and the greater is the average grid current. Thus the bias builds up with oscillation amplitude, causing the transconductance to fall until equilibrium is established. In this manner the amplitude may be prevented from becoming too high without making the quiescent transconductance so low as to prevent oscillation from starting spontaneously. Under equilibrium conditions, grid current flows during only a very small fraction of the cycle, and the grid bias is very nearly equal to the amplitude of the alternating grid voltage.

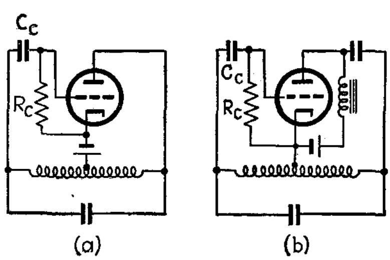


FIG. 10-57.—Self-biased Hartley oscillator with (a) series plate feed and (b) parallel plate feed.

Low power loss in the resistor, high frequency stability, and good wave form call for the use of high grid-leak resistance; but if the resistance and capacitance are too high, oscillation is not continuous. After a number of cycles of oscillation the bias becomes so high that the circuit stops oscillating, or "blocks." Because oscillation starts at a lower bias than it stops, some time elapses while the condenser discharges sufficiently to allow oscillation to recommence, and so periods of oscillation alternate with periods of rest. This phenomenon is termed *motorboating*.[2] The period of motorboating depends upon the time required for the condenser to discharge, which increases with the product of the grid condenser

[1] The bias may also be considered to result from the flow of current through the grid leak.

[2] BEATTY, R. T., and GILMOUR, A., *Phil. Mag.*, **40,** 291 (1920); RSCHEWKIN, S., and WWEDENSKY, B., *Physik. Z.*, **32,** 150 (1922); TAYLOR, L. S., *J. Opt. Soc. Am. and Rev. Sci. Instruments*, **11,** 149 (1926); *J. Franklin Inst.*, **203,** 351 (1927), **204,** 227 (1927); *Phys. Rev.*, **29,** 617 (1927).

capacitance and the grid-leak resistance. By proper choice of circuit constants, the circuit can be made to block after a single oscillation. Figure 10-58 shows one form of blocking oscillator used as a periodic-pulse generator, and Fig. 10-59 shows the form of a typical voltage pulse generated by such a circuit. If the inductances and distributed capacitances of the coils are small, the pulse is of very short duration. The repetition frequency is governed by the product *RC*.

Another factor that limits the size of the grid resistor is danger of cumulative increase of positive grid voltage and plate current as the result of primary and secondary grid emission. This action was explained on pages 157–158.

When automatic bias is produced by grid rectification, as in the circuits of Fig. 10-57, the use of a biasing battery (or cathode bias resistor)

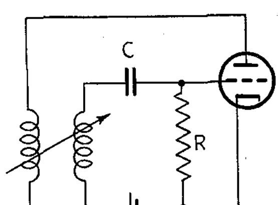


FIG. 10-58.—Blocking oscillator for the generation of periodic pulses.

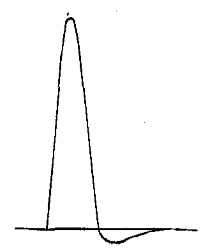


FIG. 10-59.—Single-cycle pulse generated by the circuit of Fig. 10-58.

in addition to the condenser and leak has no effect upon the amplitude of oscillation. The grid swing is approximately equal to the equilibrium value of bias. If a C battery is introduced after equilibrium is established, the total bias will be temporarily increased by the battery voltage, and so the amplitude will fall. Because of the reduced amplitude and the increased bias the grid will no longer swing positive during the positive peaks of the cycle, grid current will cease flowing and the grid condenser will discharge, reducing the negative voltage of the grid. This will continue until the negative grid voltage has fallen to the original equilibrium value, at which time the grid will again swing positive at the positive peaks, and equilibrium will be reestablished at the initial amplitude. Thus, after the oscillator has started, the only permanent effect of the addition of the C battery is the lowering of the average condenser voltage by an amount equal to the added battery voltage. The condenser cannot, however, charge in such a direction as to make its grid side positive, and so if the battery voltage is made to exceed the equilibrium value of bias voltage, oscillation will cease. An external bias somewhat smaller than the equilibrium value will prevent the oscillator from starting. Change of plate voltage changes the equilibrium bias and hence the amplitude.

By the use of a diode as the rectifier, as in the circuit of Fig. 10-60,[1] automatic amplitude control may be combined with manual control by supplementary bias. The supplementary bias produced by the cathode resistor prevents grid rectification. Since the supplementary bias is not in series with the rectifier in this circuit, increase of supplementary bias reduces the bias that the rectifier must supply in order to maintain equilibrium. Because the bias provided by the rectifier is roughly equal

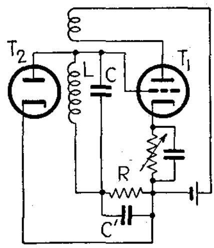


FIG. 10-60.—Circuit combining adjustable bias and diode automatic bias.

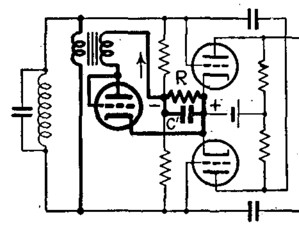


FIG. 10-61.—Application of diode amplitude control to the negative-resistance oscillator of Fig. 10-45.

to the crest amplitude, the amplitude of oscillation is thus reduced by an amount equal to the supplementary bias. The heavy lines of Fig. 10-61 show how diode amplitude control may be applied to the negative-resistance oscillator of Fig. 10-45. The primary impedance of the transformer should be high in comparison with the impedance of the tank circuit.

10-39. Resistance-capacitance–tuned Oscillators.—In Sec. 10-28 it was pointed out that if the application of a voltage E_i to the input of an amplifier produces an output voltage E_o, a portion E_f of which is impressed upon the input terminals, sustained oscillation results if E_f is equal to E_i in phase and magnitude. Since the proper phase and magnitude of feedback voltage may be attained without the use of inductance, it is possible to design oscillators that do not require inductance. The basic form of one such oscillator is shown in Fig. 10-62.[2] Inspection of Fig. 10-62 shows that the following relation must hold in order for sustained oscillation to take place:

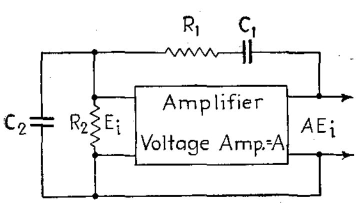


FIG. 10-62.—Basic circuit of one type of resistance-capacitance sine-wave oscillator.

[1] ARGUIMBAU, L. B., *Proc. I.R.E.*, **21,** 14 (1933); GROSZKOWSKI, J., *Proc. I.R.E.*, **22,** 145 (1934).

[2] It is of interest to note that the coupling network in this oscillator may be considered to be a special form of the Wien bridge of Fig. 15-43 in which r_3 and r_4 are zero.

$$E_i = E_f = AE_i \frac{R_2/(1 + j\omega R_2C_2)}{R_2/(1 + j\omega R_2C_2) + R_1 + 1/j\omega C_1} \tag{10-39}$$

in which A is the vector voltage amplification of the amplifier. Rationalization of Eq. (10-39) gives a complex identity, the real and imaginary parts of which must each be identities. These two identities yield the following expressions for the frequency of oscillation and the criterion for sustained oscillation:

$$f = \frac{1}{2\pi \sqrt{R_1R_2C_1C_2}} \tag{10-40}$$

$$A = 1 + \frac{R_1}{R_2} + \frac{C_2}{C_1} \tag{10-41}$$

If the two resistances are equal and the two capacitances are equal, these equations reduce to

$$f = \frac{1}{2\pi R_1C_1} \tag{10-42}$$

$$A = 3 \tag{10-43}$$

The fact that the frequency is inversely proportional to the capacitance, instead of to the square root of the capacitance, makes it possible to cover a 10-to-1 frequency range with ganged condensers of the type used in broadcast receivers. The frequency band may be readily changed by changing R_1 and R_2. Care must be taken to make the circuit constants of such values as to ensure that the oscillation is of the sinusoidal type.

Equation (10-41) shows that the vector amplification A of the amplifier must be real and positive, *i.e.*, the output voltage must be in phase with the input voltage. This is true in a two-stage resistance-capacitance–coupled amplifier in the mid-band range of amplification. The range of frequency over which the amplification is independent of frequency may be extended and the amplification made essentially independent of the feedback circuit R_1-R_2-C_1-C_2 by the use of inverse feedback within the amplifier. The effect of the feedback network R_1-R_2-C_1-C_2 upon the amplifier can be further reduced by the use of a cathode-follower output stage in addition to the two stages required to make the output voltage of the amplifier in phase with the input voltage.

Figure 10-63 shows a practical form of this circuit used in one of the most successful commercial resistance-capacitance–tuned oscillators.[1] The ballast lamp R_3, in combination with the resistance R_4, provides inverse feedback that makes the amplification and phase shift of the two-stage resistance-capacitance–coupled amplifier independent of variable circuit parameters, supply voltage, and tube characteristics, and at

[1] TERMAN, F. E., BUSS, R. R., HEWLETT, W. R., and CAHILL, F. C., *Proc. I.R.E.*, **27**, 649 (1939).

the same time affords a method of stabilizing the amplitude of oscillation. Increase of amplitude raises the current through the lamp and thus increases its resistance. This in turn increases the inverse feedback, which decreases A and hence tends to prevent increase of amplitude of oscillation.

Because the mu-factor relating the suppressor and screen voltages of a pentode is negative throughout a portion of the operating range of voltages (see Sec. 10-23), a single-stage amplifier in which the output

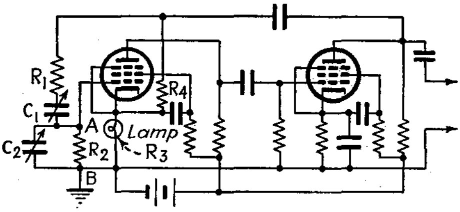


FIG. 10-63.—Practical form of the oscillator of Fig. 10-62.

voltage is in phase with the input voltage may be made by impressing the input voltage upon the suppressor of a pentode and taking the output voltage from the screen circuit. Such an amplifier may, therefore, be used in the oscillator circuit of Fig. 10-62, giving the single-tube resistance-capacitance–tuned oscillator of Fig. 10-64. The purpose of the resistance R_f and the condenser C_f is to provide inverse feedback in order to improve wave form and stability. The frequency is varied by means

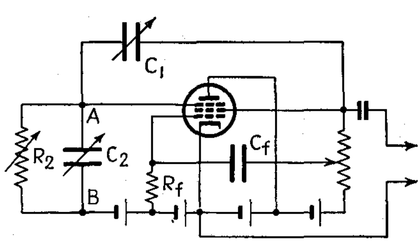


FIG. 10-64.—Single-tube version of the circuit of Fig. 10-62.

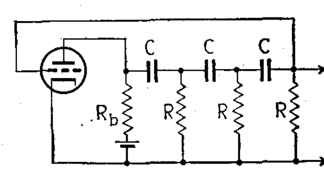


FIG. 10-65.—Basic circuit of the "phase-shift" oscillator.

of C_1, C_2, and R_2. The condenser C_2 may shunt the screen resistor, instead of R_2.[1]

A single-stage amplifier of the ordinary type may be used as the basis of a resistance-capacitance–tuned oscillator if the feedback network is capable of providing a 180-degree phase shift. The basic circuit of such an oscillator is shown in Fig. 10-65.[2] If $R_b << R$ and all capaci-

[1] DELAUP, P. S., *Electronics*, January, 1941, p. 34 (with bibliography).

[2] GINZTON, E. L., and HOLLINGSWORTH, L. M., *Proc. I.R.E.*, **29**, 43; KUNDE, W. W., *Electronics*, November, 1943, p. 132.

tances C and resistances R are alike, the frequency of oscillation is $1/2\pi\sqrt{6}\ RC$ and the required amplification of the amplifier is 29. If the tuning is varied by means of only two of the condensers, which are ganged, the frequency does not vary linearly with $1/C$, and the required amplification varies with frequency. Since the amplification must be sufficient to ensure oscillation over the entire tuning range, the amplification will then be greater than required over much of the range, and automatic amplitude control must be used in order to maintain constant amplitude and ensure good wave form. Figure 10-66 shows a practical form of this oscillator incorporating automatic amplitude control. Improvement results from the addition of a cathode-follower stage

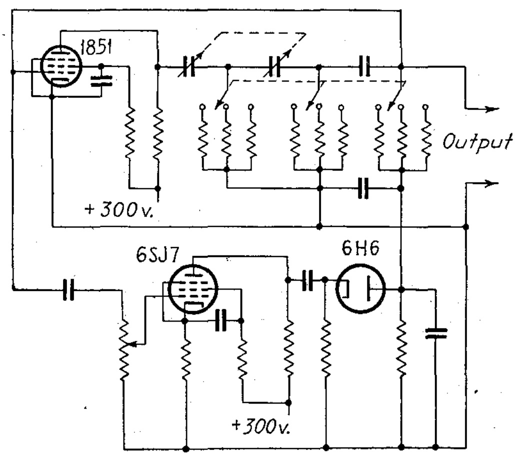


FIG. 10-66.—Practical form of the oscillator of Fig. 10-65.

between the amplifier and the phase-shifting network in order to prevent changes in the phase-shifting resistances and capacitances from affecting the amplification. Other types of feedback networks than that of Fig. 10-65 may be used.

Because the total phase shift through the amplifier and feedback network must be zero or a multiple of 2π in a feedback oscillator, any change in phase shift through the amplifier, as the result of change of supply voltage or other causes, must be accompanied by an equal change in phase shift through the feedback network. Since such a phase shift in the network results from a change in oscillator frequency, it follows that good frequency stability necessitates the use of a feedback network in which the rate of change of phase shift with frequency is high. An oscillator developed by Meacham[1] uses as the feedback network a bridge, one arm of which is a quartz crystal. The opposite arm of the

[1] MEACHAM, L. A., *Proc. I.R.E.*, **26**, 1278 (1938).

bridge is a resistance having a high positive temperature coefficient of resistance, which causes the attenuation of the bridge to build up rapidly with increase of amplitude of oscillation and thus affords automatic amplitude control. Meacham reports a frequency stability of 2 parts in 10^8. Inasmuch as the crystal frequency is not variable, this type of oscillator is suitable only for the generation of fixed frequencies. This limitation is avoided in circuits developed by Shepherd and Wise,[1] in which "bridged-T"[2] feedback networks are used. These circuits, like that of Meacham, incorporate temperature-controlled resistances (thermistors[3]) in order to obtain automatic amplitude control. Either resistance-capacitance or inductance-capacitance networks may be used. Figure 10-67 shows the oscillator stage of one of these oscillators, in which tuning is accomplished by means of ganged resistors and frequency ranges are changed by means of condensers. The oscillator stage is followed by a two-stage inverse-feedback amplifier. Shepherd and

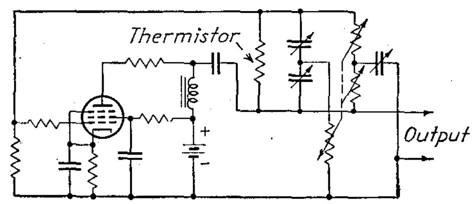


FIG. 10-67.—Resistance-capacitance-tuned oscillator using a bridged-T network.

Wise report a variation of output amplitude of less than ± 0.5 db over the frequency range from 40 cps to 50 kc.

Figure 10-68 shows the block diagram of another feedback oscillator that is tuned by means of a resistance-capacitance bridge.[4] The amplifier is designed to have constant amplification and zero phase shift over a wide range of frequency. Degenerative feedback is applied to the amplifier through an impedance bridge. Since the output of the bridge is zero at resonance, the inverse feedback is a minimum at this frequency, and the amplification of the amplifier with inverse feedback is a maximum. The regenerative-feedback network, which has flat response throughout the frequency range of the amplifier, applies positive feedback to the amplifier. If the positive feedback is high enough, sustained oscillation takes place at the frequency at which the amplification is greatest, *i.e.*, at the resonance frequency of the bridge. The phase

[1] SHEPHERD, W. G., and WISE, R. O., *Proc. I.R.E.*, **31**, 256 (1943).

[2] TUTTLE, W. N., *Proc. I.R.E.*, **28**, 23 (1940); HONNELL, P. M., *Proc. I.R.E.*, **28**, 88 (1940).

[3] PEARSON, G. L., *Phys. Rev.*, **57**, 1065 (1940).

[4] SCOTT, H. H., *Proc. I.R.E.*, **26**, 226 (1938); *Gen. Radio Expt.*, **13**, April, 1939, p. 1; **14**, January, 1940, p. 6.

inverter is necessary in order to make the feedback voltage of correct phase to give regenerative feedback. The Wien bridge, shown in Fig. 15-43, is a suitable form of degenerative-feedback bridge. If $r_3 = 2r_4$, $r_1 = r_2$, and $C_1 = C_2$, the bridge is balanced when $f = 1/2\pi r_1 C_1$.

The oscillators of Figs. 10-63, 10-66, 10-67, and 10-68 have low harmonic content, high frequency stability, and constant output voltage. They may be readily designed for wide frequency ranges, are relatively inexpensive to manufacture, and are light and compact.

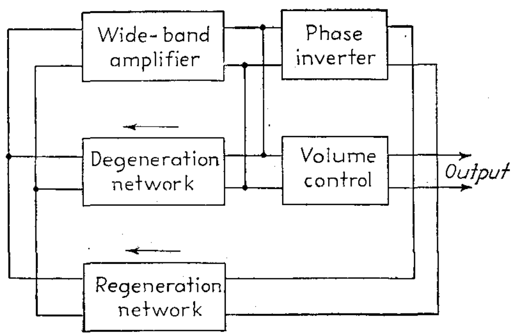


FIG. 10-68.—Block diagram of feedback R-C oscillator.

10-40. Problems of Oscillator Design.—A satisfactory oscillator for laboratory or commercial use must satisfy the following requirements: (1) The harmonic content must be small. (2) The frequency must be independent of battery voltages and load and must not drift. (3) The output voltage or power must be independent of frequency and must not change during operation of the oscillator. (4) The calibration must be accurate and constant. (5) The oscillator must be capable of furnishing the required power or voltage output.

These requirements, particularly as they affect tank-circuit types of feedback oscillators, will be discussed in the sections that follow.

10-41. Wave Form of Feedback Oscillators.—The harmonic content of the output of oscillators, like that of amplifiers, is decreased by straightening the dynamic characteristics, by limiting the amplitude of oscillation, and by the use of push-pull circuits. Although Class A operation of feedback oscillators is desirable for low distortion, the resonant circuits of oscillators greatly reduce the harmonic content and allow the use of Class B and C operation, even in single-sided circuits. Class B and C operation are particularly applicable to radio-frequency oscillators because of the ease with which harmonics may be removed by means of tuned filters. In designing small low-distortion oscillators for laboratory use, it is advisable to use Class A operation and to keep the amplitude of oscillation as low as possible. Because the current in the inductance is more nearly sinusoidal than the plate current, especially when Class A1 operation is not used, the output voltage or power should be taken from the tank circuit when good wave form is desired.

The harmonic content of feedback oscillators in Class A operation is decreased by increase of effective load impedance. The effective impedance presented by the tank circuit of a tuned-plate oscillator can be adjusted by the use of a tapped inductance. If the plate is connected to a tap and the condenser shunts the whole inductance, the effective load impedance presented to the plate is less than L/rC. If the condenser is connected to the tap, on the other hand, and the plate to the end of the inductance coil, the effective load impedance exceeds L/rC.[1] The same method is applicable to the Hartley oscillator. In the tuned-grid oscillator the load impedance may be adjusted by the inductance and coupling of the plate coil. The load may also be changed by means of the ratio of tuning inductance to capacitance.

10-42. Frequency Stability.—Undesired changes of frequency result from three major causes: changes in the mechanical arrangement of the elements of the oscillating circuit; in the values of the circuit parameters; and in the amplification factor, grid and plate resistances, and interelectrode capacitances of the tube.[2]

Changes in the mechanical arrangement of the circuit elements may be produced by vibration; by mechanical, electrostatic, or electromagnetic forces; or by temperature changes. They can be minimized by careful mechanical and electrical design and by temperature control.

Variations in the values of the circuit parameters result from changes in temperature of inductances and condensers and from variation of load, which alters the effective a-c resistance r of the tuned circuit. Changes of inductance and capacitance can be minimized by: temperature control; the use of thermally compensated inductances[3] and temperature-controlled compensating condensers;[4] and the careful choice of apparatus and the judicious location of component parts. Methods of preventing the variation of frequency with load will be considered in detail.

Tube factors are dependent upon operating voltages, upon cathode emission, and upon electrode spacing. Operating voltages can be stabilized by the use of voltage-regulating devices. Variation of cathode emission is probably the least important factor and can be reduced by the maintenance of rated cathode temperature. Electrode spacing, which depends to some extent upon tube temperature, also affects the

[1] L is the portion of the inductance shunted by C.

[2] See, for instance, F. B. LLEWELLYN, *Proc. I.R.E.*, **19**, 2063 (1931); R. GUNN, *Proc. I.R.E.*, **18**, 1560 (1930).

[3] GRIFFITHS, W. H., *Wireless Eng.*, **11**, 234 (1934).

[4] GUNN, R., *Proc. I.R.E.*, **18**, 1565 (1930).

interelectrode capacitances. The dependence of frequency upon interelectrode capacitances and upon stray circuit capacitance can be minimized by the use of a high ratio of tuning capacitance to inductance and by using circuits in which the tuning capacitance shunts the grid-plate capacitance.

10-43. Effect of Load upon Frequency. Use of Buffer Amplifier and Electron Coupling.—The exact expressions for frequency [Eqs. (10-8), (10-28), and (10-36), etc.] involve r, the effective a-c resistance of the inductance, which in turn depends upon the power delivered by the inductance. It follows that the frequency of oscillation will be affected by changes of load unless the output is taken from the oscillator in such a manner that the current in the inductance is not changed. Because of the conductive or inductive coupling between the tank inductance and the grid and plate circuits, applying the load to the grid or plate circuits is equivalent to shunting it across the inductance. Changes in load may be prevented from affecting the frequency by using the oscillator to excite a power amplifier from which the power is taken.

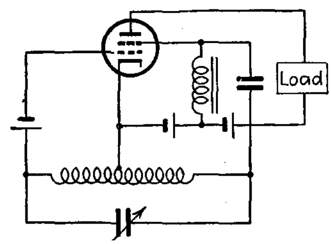


FIG. 10-69.—Hartley oscillator with electron-coupled load.

Another method of making the oscillator frequency independent of load is the use of *electron coupling* in taking the power from the oscillator.[1] A Hartley oscillator with electron-coupled load is shown in Fig. 10-69. The control grid and screen of a tetrode replace the grid and plate of a triode in the oscillating circuit. The plate circuit serves only to deliver the output. The flow of electrons through the screen grid varies with screen current, which in turn varies at the oscillation frequency. The electrons that pass through the screen are attracted to the plate by the plate field, and hence the plate current has an alternating component whose frequency is equal to the oscillation frequency. Since the field of the plate terminates almost entirely on the screen, the plate load has very little effect upon the oscillating circuit.

Because of capacitive coupling of the load to the oscillating circuit through the screen-plate capacitance, the simple electron-coupled circuit, of which Fig. 10-69 is an example, is not entirely free from the effects of reaction of load upon the oscillating circuit. This difficulty may be eliminated by the addition of a neutralizing condenser C_n, as in Fig. 10-70. The condenser C_n and the screen-plate capacitance form a voltage divider, the ends of which are connected to the control grid and screen. Since the control-grid and screen alternating voltages

[1] Dow, J. B., *Proc. I.R.E.*, **19**, 2095 (1931).

are opposite in phase, a setting of C_n may be found at which no alternating voltage is applied to the plate when the plate supply voltage is zero. The same result may be accomplished by using a pentode in place of the tetrode, the third grid, which is made positive, acting as a screen between the second grid and the plate.

If the plate and oscillator-anode voltages are properly chosen, the change in frequency resulting from a change in plate voltage is equal and opposite to that resulting from a proportional change of oscillator-anode voltage.[1] Therefore, if the supply voltage for the oscillator anode (the screen in Figs. 10-69 and 10-70) is obtained from a tap on a voltage divider across the plate supply voltage, a setting of the tap can be found such that the frequency is independent of voltage variation. Thus, the electron-coupled oscillator not only prevents reaction of load upon the oscillating circuit but also may be designed to minimize the effect of supply-voltage variation upon frequency.

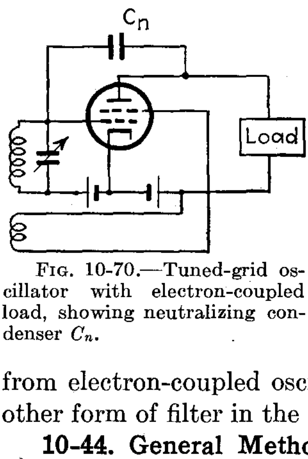


FIG. 10-70.—Tuned-grid oscillator with electron-coupled load, showing neutralizing condenser C_n.

The advantages of the electron-coupled oscillator are gained at the expense of good wave form. The high harmonic content of electron-coupled oscillators results from two causes. Inasmuch as the oscillator-anode current is in itself not sinusoidal, the space current beyond it is not likely to be. Secondly, the plate current is influenced by secondary emission effects and may even reverse. Oscillographic studies show that it is difficult to obtain sinusoidal output from electron-coupled oscillators without the use of a resonant circuit or other form of filter in the output circuit.

10-44. General Methods of Frequency Stabilization.—The general method of obtaining a high degree of frequency stability is not to attempt to prevent the variation of tube factors, but to design the oscillator in such a manner that frequency is not affected by variations of tube factors. This is done in four principal ways:

1. Proper choice of electrical parameters of the oscillating circuit:
   a. Resistance stabilization.
   b. Capacitance or inductance stabilization.
   c. Use of unity coupling.
2. Use of mechanical oscillating systems:
   a. Piezoelectric crystals.
   b. Magnetostrictive rods.
3. Use of selective filters as the oscillating circuits.
4. Elimination of harmonic currents by means of tuned filters.

[1] Dow, *loc. cit.*

Heising has shown that the frequency stability of tank-circuit feedback oscillators is favored by the use of a low L/C ratio, but that this also decreases the range of circuit adjustment over which oscillation is obtained. Q, which is a measure of sharpness of resonance would, at first thought, be expected to increase frequency stability. Heising proved, however, that high Q does not necessarily increase frequency stability but may actually decrease it. Stability is increased by the combination of high Q and low L/C.[1]

10-45. Resistance Stabilization.—One of the simplest ways of improving the frequency stability of standard oscillators is by resistance stabilization, which consists of the addition of a high resistance r_f between the plate and the oscillating circuit, as shown in the typical circuits of Figs. 10-71 and 10-72.[2] The primary function of this resistance is to make the total effective resistance in the plate circuit so high that changes in the

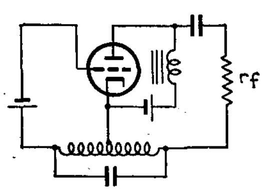


FIG. 10-71.—Resistance-stabilized Hartley oscillator.

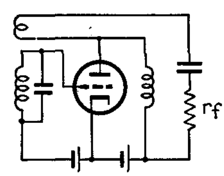


FIG. 10-72.—Resistance-stabilized tuned-grid oscillator.

plate resistance of the tube have little effect upon the oscillating circuit. A second function of the resistance is to provide a convenient means of controlling the feed back, and thus the amplitude of oscillation. The resistance should be made so high that oscillations will barely start. Then as soon as the amplitude builds up to the point where the grid starts to swing positive, the additional losses resulting from the flow of grid current prevent further increase of amplitude. In order to prevent undesirable phase shifts, the blocking condenser in series with the resistance should be so large that its impedance is negligible in comparison with the resistance.

Details of design of resistance-stabilized oscillators have been worked out by Terman,[3] who recommends the use of tubes having amplification factors ranging from 4.5 to 8, unity turn ratio of grid and plate coils, close coupling between grid and plate coils, and a feedback resistance of the order of from two to five times the plate resistance. The feedback resistance should not exceed 500,000 ohms in audio-frequency oscillators, and the resistance of the tank circuit at antiresonance[4] should lie between

[1] HEISING, R. A., *Proc. I.R.E.*, **31**, 595 (1943).

[2] HORTON, J. W., *Bell System Tech. J.*, **3**, 508 (1924).

[3] TERMAN, F. E., *Electronics*, July, 1933, p. 190.

[4] The antiresonant frequency of a parallel resonant circuit is that at which the

10,000 and 50,000 ohms. Terman gives the following formula for the value of feedback resistance at which oscillation will just start:

$$r_f = r_l(\mu - 1) - r_p$$

where r_l is the effective resistance of the resonant circuit at antiresonance. Resistance stabilization does not prevent frequency variation as the result of change of grid resistance with supply voltages.

10-46. Impedance Stabilization.—Llewellyn has shown that a more general type of stabilization, consisting of the use of capacitance or inductance in series with the grid or plate of the oscillator, or both, results in complete independence of oscillation frequency from variation

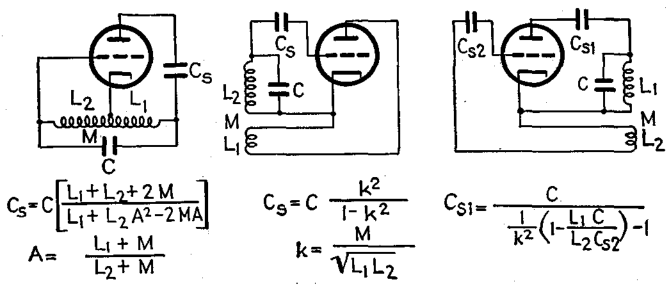


FIG. 10-73.—Impedance-stabilized Hartley, tuned-grid, and tuned-plate oscillators.

of tube factors, and hence battery voltages.[1] Llewellyn's analysis consists of setting up the equivalent circuits of feedback oscillators and determining from a general solution of the circuit equations the conditions under which the frequency of oscillation is independent of tube factors. The effect of grid current is taken into account in the analysis, but distortion resulting from nonlinearity of tube characteristics is neglected. It is assumed that losses in the oscillating circuits are so small as to be negligible, and that load is applied to the oscillator through a buffer stage which draws no power from the oscillator. In Fig. 10-73 are shown some of the methods of stabilizing Hartley, tuned-grid, and tuned-plate oscillators. Other types of oscillators may be stabilized in a similar manner.

High stability may also be attained by the use of close coupling. Llewellyn states,

> The frequency of an oscillator with unity coupling between grid and plate circuits depends only upon the inductances and capacitances in the circuit, and not at all upon the tube parameters, r_p, r_g, and μ provided, however, that losses

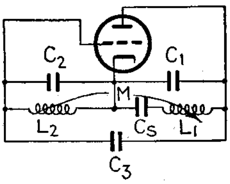
---

impedance is nonreactive. At this frequency the impedance is equal to L/rC, where r is the resistance of the inductor.

[1] LLEWELLYN, F. B., *Proc. I.R.E.*, **19**, 2063 (1931). See also G. H. STEVENSON, *Bell System Tech. J.*, **17**, 458 (1938).

in the external circuit are small, and that the harmonic voltages across the tube are small enough to allow r_p and r_g to be considered as pure resistances.

If the actual inductive coupling coefficient is not unity, an equivalent effect can be obtained by the addition of capacitance in series with the grid or plate coils, or both. Figure 10-74 shows one method of obtaining such stabilization in the generalized feedback oscillator of Fig. 10-56. The stabilizing capacitance may also be placed in series with the grid coil instead of in series with the plate coil, or stabilizing capacitance may be used in both grid and plate circuits.



$$C_S = \frac{L_2}{L_1}\left[\frac{\frac{C_1 k^2 L_1}{L_2} + C_2 + C_3\left(1+k\sqrt{\frac{L_1}{L_2}}\right)^2}{1-k^2}\right]$$

$$k = \frac{M}{\sqrt{L_1 L_2}}$$

FIG. 10-74.—Impedance-stabilized oscillator of generalized form.

In a practical stabilized circuit the stabilizing capacitances or inductances may also serve other functions in the circuit. Figure 10-75, for instance, shows a practical circuit of a doubly stabilized tuned-grid oscillator, in which the plate stabilizing condenser serves as a blocking condenser for the B-supply voltage, and the grid stabilizing condenser serves to furnish grid bias (in the manner discussed in Sec. 10-38). A resistance shunting a stabilizing condenser alters the required value of stabilizing capacitance and reduces the effectiveness of stabilization. For this reason the grid leak R_c should be large, and the exact value of stabilizing capacitance is determined most readily experimentally.

When precautions are taken to ensure a low-loss resonant circuit, the relations given for stabilization hold very accurately, remaining variations of frequency being accounted for by interelectrode capacitances and by the presence of harmonics. The effect of tube capacitances may be made unimportant by the use of circuits in which the tube capacitances form a portion of the tuned circuit. The effect of harmonics may be prevented by using low amplitudes of oscillation or by providing low-reactance paths for the harmonics so that they cannot build up reactive voltages across the tube.[1] The performance of stabilized oscillators checks the theoretical analysis. A 1-Mc oscillator showed less than 10 cps variation in frequency when the plate voltage was reduced 50 per cent and practically no change with a 50 per cent reduction of filament current.[1] Much

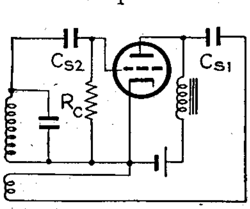


FIG. 10-75.—Practical form of doubly stabilized tuned-grid oscillator.

[1] LLEWELLYN, *loc. cit.*

greater frequency stability was found experimentally for a tightly coupled oscillator than for a loosely coupled one.

The form of the equations for stabilizing capacitance shows that for most types of oscillators it is necessary to vary the stabilizing capacitance with the tuning capacitance when adjustable frequency is required. When grid or plate stabilization alone is used, the required stabilizing capacitance is proportional to the tuning capacitance. Because the frequency is independent of grid and plate resistance when stabilization is complete, an external load may be applied either between grid and cathode or between plate and cathode without affecting the stabilization.

10-47. Crystal Oscillators.[1]—The control of frequency by means of crystals is based upon the piezoelectric effect. When certain crystals, notably quartz, are compressed or stretched in certain directions, electric charges appear on the surfaces of the crystal that are perpendicular to the axis of strain. Conversely, when such crystals are placed between two metallic surfaces between which a difference of potential exists, the crystals expand or contract. If the potential applied to the plates is alternating, the crystal is set into vibration, the amplitude being greatest at the mechanical resonance frequency of compressional oscillation of the crystal. Although special types of crystals may be used in audio-frequency oscillators,[2] crystal control is at present restricted mainly to radio-frequency oscillators.

Figure 10-76 shows one commonly used crystal oscillator, developed by J. M. Miller at the Bureau of Standards. The action of this circuit is readily understood when it is noted that the crystal and crystal mounting may be represented by the equivalent electrical circuit of Fig. 10-77.[3] When the plate load is inductive, the effective input conductance of the tube is negative (see Sec. 4-3). Oscillation may, therefore, be set up in a resonant circuit connected between the grid and the cathode. In order to maintain an inductive plate load the plate circuit must be tuned so that its resonant frequency is slightly higher than that of the crystal. The value of the crystal in controlling frequency lies in the extreme sharpness

[1] CADY, W. G., *Proc. I.R.E.*, **10,** 83 (1922); PIERCE, G. W., *Proc. Am. Acad. Arts Sci.*, **59,** 81 (1923); HULL, A. W., *Phys. Rev.*, **27,** 439 (1926); DYE, D. W., *Proc. Phys. Soc. London*, **38,** 399, 457 (1926); HUND, A., *Proc. I.R.E.*, **14,** 447 (1926), **16,** 1072 (1928); WHEELER, L. P., and BOWER, W. E., *Proc. I.R.E.*, **16,** 1035 (1928); VAN DYKE, K. S., *Proc. I.R.E.*, **16,** 742 (1928); HARRISON, J. R., *Proc. I.R.E.*, **16,** 1455 (1928); TERRY, E. M., *Proc. I.R.E.*, **16,** 1468 (1928); CADY, W. G., *Proc. I.R.E.*, **16,** 521 (1928) (with bibliography); WRIGHT, J. W., *Proc. I.R.E.*, **17,** 127 (1929); MEAHL, H. R., *Proc. I.R.E.*, **22,** 732 (1934). For other references, see A. HUND, "Phenomena in High-frequency Systems," p. 118, McGraw-Hill Book Company, Inc., New York, 1936.

[2] TYKOCINER, J. T., and WOODRUFF, M. W., *Univ. Ill. Eng. Expt. Sta. Bull.* 291, 1937.

[3] VAN DYKE, *loc. cit.*

of its resonance curve. The Q* of the equivalent circuit of a crystal is of the order of one hundred times that which can be readily attained in electrical circuits. Because of this sharpness of resonance, the crystal can oscillate over only a very narrow frequency range, and hence the frequency stability of a crystal oscillator is high. When the temperature of the crystal is maintained constant by means of a temperature-control chamber, the frequency drift may be made less than 2 parts in 10 million.

In the original Pierce[1] circuit, the crystal is connected between the grid and plate of the tube, instead of between the grid and cathode. This circuit will oscillate only when the plate load is capacitive, and therefore the natural frequency of the plate circuit must be slightly lower than that of the crystal. A crystal will also oscillate when connected to other

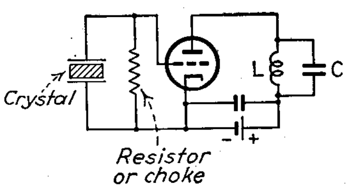


FIG. 10-76.—One type of crystal-controlled r-f oscillator.

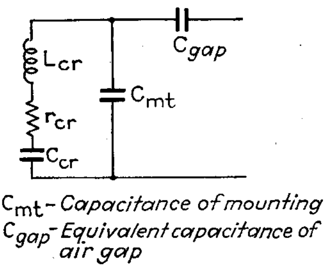


FIG. 10-77.—Equivalent circuit of crystal and crystal mounting.

types of negative-resistance elements,[2] such as were described earlier in this chapter. Crystals may be used in push-pull circuits.

Crystal-controlled r-f oscillators are essential in radio communication and broadcasting. The cutting of crystals and their use in r-f oscillators is discussed in detail in radio engineering textbooks. Because of the extremely sharp resonance of a crystal, it may also be used as a band-pass filter that passes a very narrow band of frequencies. One application of such a filter is in wave analyzers (see Sec. 15-38).

10-48. Magnetostriction Oscillators.—The phenomenon of magnetostriction is the expansion or contraction of magnetic materials as the result of magnetization, and the converse change of magnetization as the result of strain. The phenomenon may be used as the basis of an oscillator having high frequency stability.[3] The circuit diagram of a typical

* Q, the ratio of the total inductive reactance to the total effective resistance of a series oscillatory circuit, is a measure of the sharpness of resonance. See, for instance, W. L. EVERITT, "Communication Engineering," 2d ed., p. 65, McGraw-Hill Book Company, Inc., New York, 1937.

[1] PIERCE, *loc. cit.*

[2] MACKINNON, K. A., *Proc. I.R.E.*, **20**, 1689 (1932).

[3] PIERCE, G. W., *Proc. Am. Acad. Arts Sci.*, **63**, 1 (1928); *Proc. I.R.E.*, **17**, 42 (1929); SALISBURY, W. W., and PORTER, C. W., *Rev. Sci. Instruments*, **10**, 142 (1939).

magnetostriction oscillator is shown in Fig. 10-78. The operation is as follows: The steady component of plate current produces a steady strain in the rod. Any small change in plate current results in a change of magnetization of the end of the rod that is within the plate coil. This results in an elongation or contraction of this portion of the rod and causes a compressional wave to move toward the other end of the rod. When the wave reaches the portion of the rod that lies inside of the grid coil, the resulting change in magnetic field induces a voltage in the grid coil, which causes a change of plate current. This change in plate current in turn starts out another compressional wave in the bar. The compressional waves are reflected from the grid end of the rod and return to the plate end, where they are again reflected. If the polarity of the grid and plate coils is correct, the induced and reflected waves at the plate end are in phase and will reinforce one another. The tendency for the amplitude to build up is increased to the point of oscillation by tuning the electrical circuit so that it has the same natural period as the rod. The action is not dependent upon inductive coupling of the coils. In contrast with that in feedback oscillators, the orientation of the coils must be such that the inductive coupling is degenerative. The inductive coupling should preferably be made small by the use of shielding. Considerable improvement results from the use of a two-stage amplifier in place of the single tube shown in Fig. 10-78.[1]

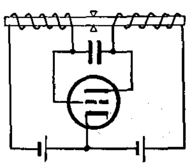


FIG. 10-78.—Single-stage magnetostriction oscillator.

Because the change in length of an initially unmagnetized rod is of the same sign for both polarities of magnetization, the rod will vibrate at twice the frequency of the electrical circuit if it is not polarized. To ensure that the rod is polarized throughout its length, it is usually magnetized permanently and placed in the coils so that the field of the plate coil increases the magnetization.

Nickel, Monel metal, Nichrome, Invar, Stoic metal, and other nickel alloys may be used to make magnetostrictive rods. Difficulties of design of suitable rods having high natural frequency of oscillation restrict the use of the magnetostriction oscillator to frequencies within the audible and supersonic ranges. Very low audio frequencies are produced by the use of rods that are loaded at the ends or that consist of an outer shell of magnetostrictive material filled with lead or other material having a low velocity of propagation of compressional waves. If the temperature of the rod is kept constant, a frequency stability of 1 part in 30,000 may be obtained.[2] Dependence of frequency upon temperature can be

[1] PIERCE, G. W., and NOYES, A., JR., *J.A.S.A.*, **9**, 185 (1938).

[2] PIERCE, *loc. cit.*

reduced by employing special alloys of low temperature coefficient[1] or rods made up of a core and a shell of two magnetostrictive metals of opposite temperature coefficients of expansion.

10-49. Tuned-filter (Multistage-feedback) Oscillator.—Excellent frequency stability can be obtained by feeding back the output of a sharply tuned multistage amplifier to the input, as shown in Fig. 10-79.[2] The principle of operation is the same as that of the common types of feedback oscillators except that amplification takes place in more than one tube. The sharper the resonance curves of the tuned circuits, the more difficult it is to cause oscillation at any but the resonance frequency. Very high frequency stability can be obtained by the use of complex filter sections or a large number of stages. Gunn reported a shift in frequency of less than 1 cps at a frequency of 1000 cps when the plate voltage of a two-stage circuit was changed by 50 per cent. At radio frequencies it is essential to prevent direct feedback in the individual stages. This requires careful shielding and the use of tetrodes or pentodes. For a properly designed radio-frequency circuit Gunn reported a 0.0003 per cent change in frequency with a 10 per cent change in plate potential, about 0.003 per cent change in frequency for an 8 per cent change of filament voltage, and a change of less than 60 parts per million per degree over a 2- or 3-degree temperature range. These results were obtained with d-c operation of the cathodes. The results were not quite so good with a-c filament operation. Although the frequency is little affected by load, the use of a Class A buffer power amplifier following the oscillator is recommended.

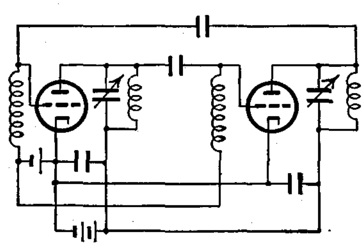


FIG. 10-79.—Tuned-filter (multistage-feedback) oscillator.

10-50. Frequency Stabilization by Harmonic Elimination.—A number of investigators have shown that frequency variation and harmonic content are interdependent.[3] The factors that ensure low harmonic content, such as a linear operating characteristic and small amplitude, therefore also tend to improve frequency stability. The frequency stability of an oscillator can be improved by the use of series filter sections tuned to the harmonic frequencies, shunted across the tank circuit, but this method cannot be conveniently used with a variable-frequency oscillator.

10-51. Power Output of Oscillators.[4]—For reasons that have already been discussed, when it is necessary to obtain large power from an oscil-

[1] IDE, J. M., *Proc. I.R.E.*, **22**, 177 (1934).

[2] GUNN, ROSS, *Proc. I.R.E.*, **18**, 1560 (1930).

[3] GROSZKOWSKI, J., *Proc. I.R.E.*, **21**, 958 (1933); ARGUIMBAU, L. B., *Proc. I.R.E.*, **21**, 14 (1933); MOULLIN, E. B., *J. Inst. Elec. Eng.* (*London*) **73**, 186 (1933).

[4] PRINCE, D. C., *Proc. I.R.E.*, **11**, 275, 405, 527 (1923).

lator it is usually expedient to use the oscillator only to excite a power amplifier, with or without intervening voltage amplifier stages. When the power is to be obtained directly from the oscillator, however, it should be borne in mind that the oscillator is in reality a self-excited amplifier and that the laws which govern the design of power amplifiers also apply to power oscillators. For maximum power output in Class A operation the load impedance is equal to the plate resistance. The effective resistance of the plate circuit can be adjusted in the manner explained in Sec. 10-41. Occasionally other considerations, such as plate-circuit efficiency and harmonic content, may be of more importance than maximum power output.

The power output that can be obtained from a Class A oscillator greatly exceeds the output from the same tube in Class A amplification. This is because the oscillator tube remains continuously excited and can normally be kept under constant load, whereas an amplifier usually has excitation of variable amplitude and may deliver zero output for long periods of time. As explained in Chap. 7, the maximum dissipation that the amplifier tube is required to withstand is equal to the plate-circuit input power. The oscillator tube dissipation under full load, on the other hand, is the power input to the plate circuit minus the power output. At 22 per cent efficiency, the dissipation is 78 per cent of the input. The Class A oscillator can therefore use about 28 per cent higher input power and so deliver 28 per cent higher output with the same maximum plate dissipation. If the circuit drops out of oscillation, the tube is likely to be damaged, both because of the increase of plate dissipation with removal of load and because of the rise in plate current resulting from the removal of automatic bias.

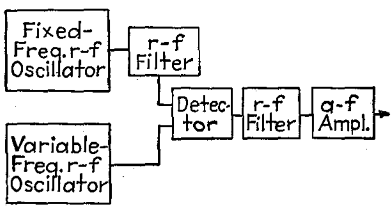


FIG. 10-80.—Schematic diagram of heterodyne oscillator.

10-52. Beat-frequency (Heterodyne) Oscillators.—In the beat-frequency oscillator, shown schematically in Fig. 10-80, the outputs of two radio-frequency oscillators of slightly different frequencies are applied simultaneously to a detector. The output of the detector contains, in addition to the impressed radio frequencies, their sum and difference. By means of a filter, the fundamental radio frequencies and their sum are removed, leaving only the difference frequency in the output, which may be suitably amplified by audio-frequency amplifiers. The popularity of the heterodyne oscillator is due principally to the fact that the whole range of audio frequencies, from 15,000 cps, or higher, down to as low as 1 cps, may be covered with a single dial. Other advantages that may be obtained with careful design include good wave form, constant output level, lightness, and compactness. By proper variable-condenser design a logarithmic frequency scale may be obtained, a considerable advantage when the oscillator is to be used in obtaining amplifier-response curves. Unless extreme care is taken in the design and construction, however, this type of oscillator is likely to have relatively poor frequency stability, which necessitates frequent setting against a standard frequency during the period of use, particularly during the time required to establish temperature equilibrium.

The design of a heterodyne oscillator involves a number of special problems, among which are the prevention of interaction between the oscillators, elimination of harmonics and other undesired frequencies, improvement of frequency stability, and the prevention of variation of output level. Interaction of the two r-f oscillators causes them to pull into synchronism when their frequency difference is small and thereby prevents the production of low audio frequency. It also tends to distort the output wave into a saw-tooth wave at output frequencies somewhat greater than that at which the oscillators pull into step. Interaction may be prevented by adequate shielding, proper location of component parts, use of chokes or decoupling resistors (see page 210) and by-pass condensers in voltage supply leads, and by correct methods of coupling the output of the r-f oscillators to the detector tube. Coupling methods include the use of buffer amplifiers between the oscillators and the detector; electron-coupled oscillators; a balanced modulator circuit (see Sec. 9-5) as in Fig. 10-81; a mixing bridge; or a multigrid mixer tube, such as the type 6L7, the oscillator outputs being applied to two grids that are shielded from each other. Of these, the first three methods are most frequently used.

If the harmonics are removed from the output of one of the r-f oscillators of a heterodyne oscillator and detector distortion is negligible, the only undesired output frequencies lie above the beat-frequency range and are removed by the r-f filter following the detector. Frequency instability in heterodyne oscillators results from the same causes as in other types of oscillators, but it is likely to be greater because a small percentage variation in the frequency of either r-f oscillator produces a much greater percentage variation of output frequency. The stability of the r-f oscillators should be as high as possible, and the two oscillators should be as nearly alike as possible in order that they will respond similarly to changes of supply voltage and of temperature. Variation of amplitude of the oscillator output results from change of amplitude of the variable oscillator over the tuning range, from the action of improperly designed r-f filters in the plate circuit of the detector, and from frequency distortion of the a-f amplifier. The first effect can be made small by using a linear detector and making the amplitude of the variable-frequency input to the detector much greater than that of the fixed-frequency input (see Sec. 9-18). Audio-frequency attenuation by the r-f filter in the detector plate circuit can be avoided most readily by the use of a radio frequency that is several times the highest output frequency. The choice of the best value of frequency of the r-f oscillators of a heterodyne oscillator is governed largely by frequency stability, ease of filtering of r-f voltages from the detector output, and desired frequency range.

The circuit diagram of a typical commercial beat-frequency oscillator is shown in Fig. 10-81.

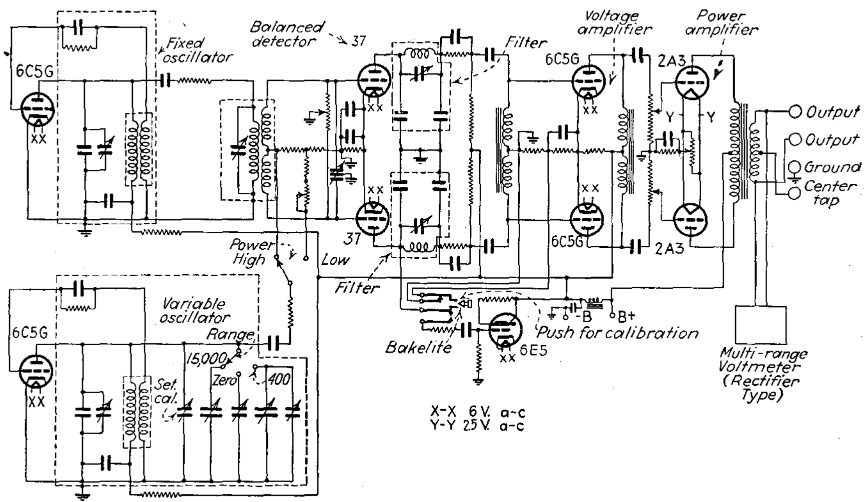


FIG. 10-81.—Circuit diagram of a heterodyne oscillator. (*Courtesy of Clough-Brengle Company.*)

10-53. Ultrahigh-frequency Oscillators.—The theory of ultrahigh-frequency oscillators such as the Barkhausen-Kurz oscillator, the magnetron oscillator, and the klystron, is beyond the scope of this book. The student is, therefore, referred to textbooks on the subject of ultrahigh frequency.[1]

Problems

10-1. Analyze the action of the circuit of Fig. 10-29*b*.

10-2. Using the complex method of analysis, derive Eq. (10-10) and determine the criterion for oscillation of a series-type negative-resistance oscillator.

10-3. The condenser C_c of Fig. 10-41 may be replaced by a resistance, as in Fig. 10-82.

a. For the circuit resistances and supply voltages shown, find the suppressor operating voltage.

[1] See, for instance, BRAINERD, J. G., KOEHLER, G.. REICH, H. J., and WOODRUFF, L. F., "Ultra-high-frequency Techniques," Chap. 10, D. Van Nostrand Company, Inc., New York, 1942.

b. Find the ratio of the change in suppressor voltage Δe_{c3} to a change of screen voltage Δe_{c2}, that causes it.

c. From Fig. 10-43 derive a dynamic i_{c2}-e_{c2} characteristic for this circuit, and from it find ρ_0.

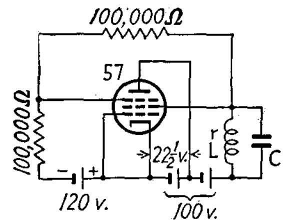


FIG. 10-82.—Negative resistance oscillator circuit for problem 10-3.

d. For an L/r ratio of 0.01 henry/ohm, find the largest condenser capacitance C at which the circuit will oscillate, and the corresponding inductance required for an oscillation frequency of 100 cps.

10-4. Derive Eq. (10-18) by means of the equivalent plate circuit (see also Prob. 4-1 *j*).

10-5. Using the equivalent plate circuit for the tuned-grid oscillator, derive Eqs. (10-36) and (10-37).

10-6. In order for oscillation of constant amplitude to take place in the circuit of Fig. 10-62, the susceptance of C_2 must be equal in magnitude and opposite in sign to the input susceptance of the amplifier, and the conductance of R_2 must be equal in magnitude and opposite in sign to the input conductance. Using Eqs. (6-107) and (6-108) to determine the input susceptance and conductance, derive Eqs. (10-40) and (10-41).

Supplementary Bibliography

The literature on the subject of oscillators is so extensive that a complete bibliography cannot be given. Pertinent references have been given throughout this chapter. The following supplementary bibliography on trigger circuits may be of value to those interested in this subject.

HARTLEY, R. V. L.: U. S. Patent 1218650.
MINORSKY, N.: *J. Franklin Inst.*, **203**, 181 (1927).
LORD, H. W., and LIVINGSTON, O. W.: *Electronics*, **7**, 1934, p. 7.
SCHMITT, O. H.: *J. Sci. Instruments*, **15**, 24 (1938).
NOTTINGHAM, W. B.: *Rev. Sci. Instruments*, **11**, 2 (1940).
KALLMANN, H. E.: *Proc. I.R.E.*, **28**, 351 (1940).

CHAPTER 11

ELECTRICAL CONDUCTION IN GASES

Preceding chapters have dealt only with electron tubes in which the gas density is so low that the probability of collision of electrons with gas molecules is small and relatively few ions are present. This chapter will treat the flow of current through regions of comparatively high gas density, in which the probability of ionizing collisions of electrons with molecules is comparatively high and many ions are consequently present in the gas. The construction, characteristics, and applications of practical gaseous conduction tubes will be taken up in the succeeding chapter. The student should review Secs. 1-1 to 1-12 of Chap. 1.

The early experiments on gas-filled tubes were performed years before the discovery of thermionic emission, but the development of useful gas-filled tubes, particularly of the grid-controlled type, lagged far behind that of high-vacuum tubes. It was not until careful experiments suggested the fundamental laws and processes governing the phenomena involved that great strides were made in the design, manufacture, and application of gaseous tubes. A knowledge of these laws is important in the understanding of gaseous tube operation. Since the formation of an arc may be preceded by a glow discharge and since the presence of glows in arc tubes may prevent their proper functioning, the theory of glow discharges is important in the study not only of glow tubes but also of arc tubes. For this reason, considerable space will be devoted to a treatment of glow discharges, even though glow tubes are now used relatively little in comparison with arc tubes, except for illumination.

11-1. Current-voltage Characteristics of Glows.—A gas or vapor containing no ions would be a perfect insulator. No current could flow as the result of a potential applied between two nonemitting electrodes immersed in such a gas. Ion-free gases are hypothetical only, for cosmic rays, radioactive materials in the walls of containers, photoelectric emission, and other ionizing agents ensure the presence of some "residual" ions. In an actual gas the application of potential between electrodes will cause a migration of the residual ions and thus produce a current.

The magnitude of the current associated with the drift of residual ions evidently depends upon the rate of production of ions and upon the rate at which ions are swept away, which in turn depends upon the applied voltage. This current is sometimes called the *dark current* because it is not accompanied by appreciable radiation. The relation between voltage and dark current under static conditions is shown graphically by the portion *oab* of the typical characteristic curve of Fig. 11-1. That the increase of voltage over the range *ab* causes no increase in the current indicates that residual ions are being swept out of the gas by the electric field at the same rate as they are being created. The dark current is usually of the order of a microampere.

As the voltage is increased, a value is reached, as at *b*, at which the current again begins to rise. If the electrode spacing is very small and the pressure sufficiently low, the current can be increased beyond *b* only by increase of voltage, and the characteristic is of the form shown by the dashed curve *bm*. With electrode spacing and pressure used in most glow-discharge tubes, on the other hand, a current is reached at *c*, called the *threshold current*, at which the current begins to rise abruptly without further increase of voltage. The threshold current is still of the order of one or two microamperes. If the external circuit resistance is low, the voltage of the discharge remains practically constant, and the current jumps to a high value (milliamperes or even amperes), corresponding to *n*. If the external circuit is such as to prevent the current from rising abruptly, then the voltage drops abruptly to some lower value, as at *d*. The value of the current at *d*, and the path along which the change takes place, appear to depend almost entirely upon the external circuit. Experimental difficulties, such as the occurrence of relaxation oscillations, have so far prevented a complete study of the characteristic immediately beyond the point *c*, but those experiments which have been performed appear to indicate that, if the terminal voltage could be reduced rapidly enough by increase of circuit resistance or decrease of applied voltage, the current at point *d* would be the same as that at point *c*.

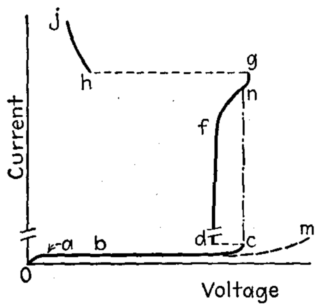


FIG. 11-1.—Current-voltage characteristic of a typical gas-filled cold-cathode diode.

The incipient condition obtaining at *c* when the current through the discharge increases without increase in voltage is called *breakdown*, and the corresponding voltage of the tube is called the *breakdown voltage*. From what has been said it follows that breakdown may not take place if the electrode spacing and pressure are small.

The increase of current above the saturation value maintained from *a* to *b* is the result of ionization of gas molecules by collision with electrons that have been accelerated by the applied electric field. Appreciable ionization by collision does not take place until the voltage considerably exceeds the ionization potential of the gas. The voltages corresponding to *c* and *d* are usually many times the ionizing potential. Ionization is accompanied by excitation of some of the gas molecules, which subsequently emit radiation as they return to a state of lower energy. It is the visible radiation that gave rise to the term *glow discharge.*

Between d and f, the current increases with very little change in voltage. The current range covered by this portion of the characteristic depends upon the area of the cathode and may extend to values of 50 ma or more. In order to indicate the difference in magnitude of the very small currents below c and the relatively large currents above d, a gap is shown in the curve.

A value of current is reached, at f, beyond which a much larger increase of voltage is required to produce a given change in current. In the vicinity of g, however, the current is so high that, if it is maintained for an appreciable time, the cathode becomes hot enough to emit electrons. The thermionic emission reduces the voltage drop through the tube in a manner that will be explained in Sec. 11-14, causing further increase of current and greater emission. The cumulative action may result in an abrupt decrease of potential to about the ionization potential of the gas, at point h, and, unless limited by circuit resistance, the current may rise to destructive values. In the region from b to g, the discharge is characterized by comparatively low current density and high voltage drop and is called a *glow discharge.* Beyond h, the discharge is characterized by high current density and low voltage drop and is called an *arc.*

The exact predetermination of the behavior of a particular glow-discharge tube is difficult, if not impossible, because a given tube does not have a single current-voltage characteristic, but an infinite number of characteristics. The shape of the characteristic depends upon gas pressure, electrode temperature, and age of the tube; upon the amount of ionization remaining from previous discharges, which in turn depends upon the time elapsed between discharges; upon the initial cathode emission, which varies with the cathode illumination; and upon the strength of other ionizing agents in the gas or container. A characteristic obtained with steady applied voltages, of which the curve of Fig. 11-1 is a typical example, is called a *static characteristic.* Characteristics obtained with varying voltages and currents are called *dynamic characteristics.*

11-2. Physical Aspects of Glow Discharge.—The appearance of the glow discharge is by no means that of a homogeneous column of light. It is made up of a number of distinct regions. The general form of a glow discharge at gas pressures of the order of a millimeter or less of mercury[1]

[1]
$$\begin{aligned} 1 \text{ atm} &= 760 \text{ mm Hg} \\ 1 \text{ micron of Hg} &= 0.001 \text{ mm Hg} \\ 1 \text{ barye} &= 1 \text{ dyne per square centimeter} \\ 1 \text{ barye} &= 0.752 \text{ micron Hg} \\ 1 \text{ micron of Hg} &= 1.333 \text{ baryes} \\ 1 \text{ atm} &= 1{,}013{,}000 \text{ baryes} \end{aligned}$$

is shown in Fig. 11-2*a*.[1] In the vicinity of the cathode is a relatively dark region, called the *cathode dark space*, or, in honor of two of the early workers in the field of glow discharges, the *Crookes* or *Hittorf dark space*. The brightness of the cathode dark space increases toward the cathode, culminating sometimes in a luminous layer very close to the cathode, called the *cathode glow*. In hydrogen and the noble gases there is a much darker layer immediately adjacent to the cathode, usually not more than

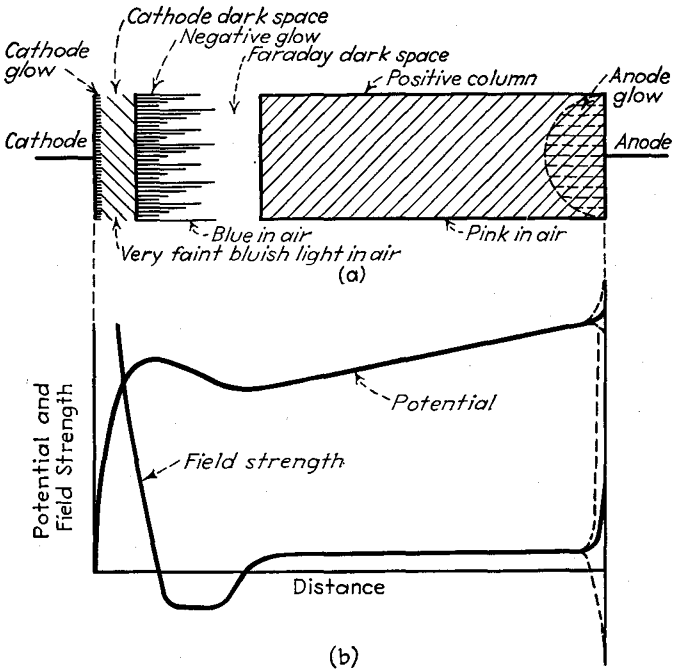


FIG. 11-2.—(*a*) Appearance of a glow discharge between plane parallel electrodes. (*b*) Potential distribution and field strength in a glow discharge between plane parallel electrodes.

a millimeter in length, called the *primary dark space*. Toward the anode, the cathode dark space terminates quite sharply and distinctly in a luminous region called the *negative glow*. The brightness of the negative glow decreases toward the anode, and the glow gradually merges into another relatively dark region, the *Faraday dark space*. Beyond the Faraday dark space is the *positive column*, a fairly luminous region which may contain striations and which extends to the anode. Sometimes the surface of the anode is covered by a luminous layer called the *anode glow*.

If the anode is moved toward the cathode while the tube is glowing, the various regions of the discharge remain fixed relative to the cathode,

[1] SLEPIAN, J., and MASON, R. C., *Elec. Eng.*, **53**, 511 (1934).

and the positive column and the Faraday dark space in turn appear to move into the anode. If the current is maintained constant during this experiment, the voltage across the tube falls slowly. As the anode enters the negative glow, the voltage begins to rise rapidly with further movement of the anode until it equals the supply voltage and the tube goes out. If the electrode spacing is kept constant, on the other hand, and the pressure is lowered, the cathode dark space increases in thickness, appearing to push away the negative glow, Faraday dark space, and positive column, which disappear into the anode in turn. As the cathode dark space expands, the outer boundary becomes less distinct. When the cathode dark space reaches the anode, the voltage rises rapidly with further change in pressure, until it equals the supply voltage and the tube goes out. At extremely low pressures, at which the dark space fills the whole interelectrode space, the walls of the tube fluoresce. This phenomenon is caused by the impact upon the walls of cathode rays, or high-speed electrons, which have been accelerated throughout the interelectrode space without suffering loss of energy by collision with gas molecules. In most practical glow-discharge tubes, the pressure is in the range from 0.1 to 50 microns.

11-3. Potential Distribution in the Glow Discharge.—The strikingly different appearance of the various parts of the glow discharge early aroused curiosity as to the electrical relationships existing in the discharge and led to many investigations. Experiments have shown that the potential distribution between cathode and anode is of the form shown in Fig. 11-2*b*. The greatest part of the applied potential appears across the cathode dark space, a region in which positive space charge predominates.[1] The fall in potential through the cathode dark space is called the *cathode fall of potential*, or *cathode drop*.

The potential maximum in the negative glow is caused by positive-ion space charge, in a manner analogous to the production of a potential minimum by electrons near the surface of a thermionic emitter in vacuum. Since the field between the potential maximum in the negative glow and the potential minimum in the Faraday dark space is such as to oppose the current, the current in this region must result from diffusion caused by ion concentration gradients (differences in ion density, page 7).

Measurements show that, if there are no striations, the field in the positive column is small and practically constant, indicating that the electron and positive-ion densities are approximately equal. The potential gradient (field) in the positive column must be great enough to supply the energy that is lost through diffusion of ions to the walls. The ratio of the diffusion current to the anode current at constant current density

[1] Positive ions move toward the cathode, electrons away. Preponderance of positive ions is accentuated by their large mass and consequent low velocity.

increases with decrease of tube radius, and hence the gradient is inversely proportional to the tube radius. This is an important consideration in the design of small-diameter tubes of the type used in neon signs. The gradient also decreases with increase of current and with decrease of pressure. When striations are present, corresponding variations in voltage are observed along the positive column, the voltage peaks corresponding to dark striations, and the voltage troughs to bright striations. The potential difference between striations is equal to an excitation or ionization potential of the gas. Stroboscopic examination of seemingly continuous positive columns sometimes shows them to contain striations that move rapidly from anode to cathode.

Adjacent to the anode there is usually a jump in potential, which may be either positive or negative, depending upon whether the current is high or low and the electrode area small or large. At some value of anode current the anode drop is zero.

11-4. Normal and Abnormal Glow.—When the glow current is limited to a sufficiently low value by means of external resistance, the glow covers only a portion of the cathode and the current is confined to the portion of the cathode that glows. An increase of current causes the discharge to spread over a greater portion of the cathode, the current density, cathode drop, and the thickness of the cathode dark space remaining constant. When the cathode is not completely covered by the glow, the discharge is said to be *normal* or *ordinary;* and the current density, cathode drop, and dark-space thickness are designated by the same terms. When the whole surface of the cathode is covered by glow, the cathode drop and the current density increase with increase of current, and the thickness of the cathode dark space decreases.[1] The discharge is then said to be *abnormal* or *extraordinary*. The terms *normal* and *abnormal*, *ordinary* and *extraordinary* are, unfortunately, misleading, inasmuch as the "abnormal" type of discharge is fully as common as the "normal," both in experimental and in practical tubes. To avoid possible ambiguity, quotation marks will be placed about these terms when they are used to differentiate between the two types of discharge. In Fig. 11-1 the "normal" range of current extends from d to f, the "abnormal" range from f to g. If the cathode is capable of dissipating considerable heat without becoming hot enough to give thermionic emission, the "abnormal" range of current may be larger than the "normal" range.

The "normal" cathode drop is a function of the gas used and the material of which the cathode is made, and is independent of gas

[1] If one makes the assumption that the current through the dark space is limited by space charge, a derivation similar to that which gives Child's law (see Sec. 2-10) shows that the dark space should decrease in thickness as the current density increases.

pressure. It is small in the noble gases and is usually smaller in gas mixtures than in pure gases. It decreases with decrease of electron affinity of the cathode material. The "normal" cathode drop usually exceeds 100 volts and may be as high as 400 volts or more but by proper choice of gases and cathode material may be made as low as 37 volts.[1] The "normal" current density (current per unit area of the cathode covered by glow) and the thickness of the dark space depend not only upon the gas and the cathode material but also upon the gas pressure.

An interesting example of the difference in the current-voltage curves for the two types of discharge is provided by the type 874 glow tube. The electrodes of this tube are a small rod surrounded by a coaxial cylinder. When the rod is the cathode, the relatively small cathode area becomes completely covered by glow at very small currents, and so the voltage rises with current over almost the entire working range of current. When the cylinder is the cathode, on the other hand, the total cathode area, comprising the two surfaces of the cylinder and the surface of the supports, is large, and the rise in voltage with current is comparatively small throughout a range of current that extends beyond the rated value. The two curves for this tube, and a sketch of the electrodes, are shown in Fig. 11-3. The terminal voltage of the tube includes not only the cathode drop but also the voltages across other portions of the discharge. The small rise of voltage with current when the cylinder is negative is probably accounted for largely by variation of voltage drop in portions of the discharge other than the cathode dark space.

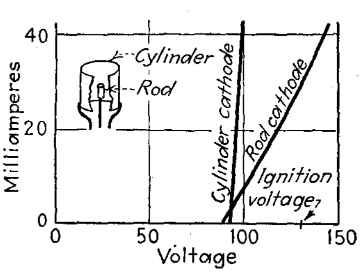


FIG. 11-3.—Current-voltage characteristics for the type 874 glow tube.

11-5. Theory of Ignition.—A theory for the increase of current with voltage as the result of cumulative ionization (at voltages higher than that corresponding to point *b* in Fig. 11-1) was first proposed by Townsend.[2] He assumed that photoelectric emission at the cathode, or some form of ionization of the gas, produces a small number of initial electrons at or adjacent to the cathode. Application of voltage to the electrodes accelerates these initial electrons toward the anode and the positive ions toward the cathode. When the applied potential exceeds the ionization potential of the gas, the initial electrons produce further ionization of the

[1] SLEPIAN, and MASON, *loc. cit.;* THOMSON, J. J., and THOMSON, G. P., "Conduction of Electricity through Gases," Vol. II, pp. 231–232, Cambridge University Press, Cambridge, 1928.

[2] TOWNSEND, J. S., "Electricity in Gases," Clarendon Press, London, 1915.

gas by collision. A derivation based upon these assumptions yields the following expression for the current:[1]

$$I = N_o \mathcal{E} \epsilon^{\beta_n d} \tag{11-1}$$

in which N_o is the number of initial electrons leaving the cathode per second; $\mathcal{E}$ is the electronic charge; β_n is the electron ionization coefficient, or number of ions formed by an electron by collision per centimeter of advance; and d is the distance between cathode and anode. Experiment shows that β_n is a function of the electric field, increasing with increase in field. To apply Eq. (11-1), β_n is assumed constant, which is true if the field is initially uniform and remains undistorted by space charge. When the current is small, so that these conditions are realized, Eq. (11-1) is checked experimentally. It should be noted that β_n is a definite function of the electric field, and the product $\beta_n d$ a measure of the voltage.

This much of the theory gives an adequate explanation for the increase of current with applied voltage in the range *bc* or *bm* in Fig. 11-1. It does not account for the increase of current without increase of voltage beyond the point *c*. Each initial electron can produce only a finite number of ion pairs. This fact implies that the current must be finite. (The position of each new pair of ions resulting from a given initial electron is closer to the anode, and so each new electron does not produce so many new ions as do the initial electrons.) If the initial source of ions were removed, the discharge would soon cease, for all ions would eventually be formed so close to the anode that the resulting electrons would not move far enough to produce other ions.

To explain the increase of current without increase of voltage, Townsend postulated ionization by collision of positive ions, as well as electrons, with gas molecules. Since positive ions travel toward the cathode, some of the electrons produced by positive ions will have the same distance of travel to the anode as did the initial electrons and will behave in the same manner as the initial electrons. If these new electrons are equal in number to the initial electrons, the discharge is self-sustaining, *i.e.*, will continue even if the source of initial ionization is removed. The advent of a self-sustaining discharge will be termed *ignition*, and the voltage at which it occurs, the *ignition potential*.[2] The ignition potential depends upon the kind and pressure of the gas and upon the electrode material and

[1] SLEPIAN, J., "Conduction of Electricity in Gases," p. 78, Westinghouse Electric & Manufacturing Co., Educational Dept., East Pittsburgh, 1933; DARROW, K. K., "Electrical Phenomena in Gases," p. 274, Williams & Wilkins Company, Baltimore, 1932; THOMSON and THOMSON, *op. cit.*, Vol. II, p. 512; COBINE, J. D., "Gaseous Conductors," Sec. 7-2, McGraw-Hill Book Company, Inc., New York, 1941.

[2] This is often called the *sparking potential*, a term that seems confusing, since no spark, in the ordinary sense, need occur. Another alternative term is *striking potential*.

geometry. It may be reduced considerably by coating the cathode with a barium compound. The ignition potential is not a constant for given gas and geometry of electrodes but depends upon the same factors, listed on page 418, that cause changes in shape of the static current-voltage characteristic.

When cumulative ionization by both electrons and positive ions occurs, the current may be expressed by the relation[1]

$$I = N_o \mathcal{E} \frac{(\beta_n - \beta_p)\epsilon^{(\beta_n - \beta_p)d}}{\beta_n - \beta_p \epsilon^{(\beta_n - \beta_p)d}} \tag{11-2}$$

if it is assumed that the field is initially uniform and remains undistorted by space charge. β_p, the ionization coefficient for positive ions (analogous to β_n), is the number of new ion pairs produced per centimeter advance of a positive ion. Experiments show that β_n increases with field strength more rapidly than does β_p. As the voltage is raised and β_n increases relative to β_p, the denominator of Eq. (11-2) decreases. As the denominator approaches zero, N_o may also approach zero without reducing the current to zero. The criterion for the advent of self-sustaining discharge is, therefore, given by the relation

$$\beta_n = \beta_p \epsilon^{(\beta_n - \beta_p)d} \tag{11-3}$$

and the voltage at which this relation is satisfied is the ignition voltage.

If Eq. (11-2) continued to hold after ignition takes place, it would indicate that ignition would be coincidental with an increase of current to a value determined only by the resistance of the external circuit. The increase of current is, however, accompanied by distortion of the field by space charge. This necessitates the use of a more general equation applying to nonuniform fields. It will be shown that under some conditions the effect of space charge is to limit the current even when the external circuit has negligible resistance, whereas under other conditions the formation of space charge favors the increase of current.

J. J. Thomson developed a similar theory based upon the assumption that the positive ions, instead of producing ionization by collision, eject secondary electrons from the cathode when they strike it. Thomson's theory yields the following equation for the current:[2]

$$I = N_o \mathcal{E} \frac{\epsilon^{\beta_n d}}{1 + k(1 - \epsilon^{\beta_n d})} \tag{11-4}$$

[1] DARROW, *op. cit.*, p. 280; SLEPIAN, *op. cit.*, p. 88; COBINE, *op. cit.*, Sec. 7-5.

[2] THOMSON and THOMSON, *op. cit.*, Vol. II, p. 518; SLEPIAN, *op. cit.*, p. 92; DARROW, *op. cit.*, p. 282.

in which k is the number of secondary electrons ejected from the cathode for each ion pair formed in the gas. Because of the increase of kinetic energy of the positive ions with increase of applied voltage, k increases with applied voltage. A self-sustaining discharge obtains when the denominator is zero, *i.e.*, when

$$\epsilon^{\beta_n d} = \frac{k+1}{k} \tag{11-5}$$

It is interesting to note that Eq. (11-5) may be shown to be equivalent to the statement that, for each electron initially released from the cathode, the ionization and secondary-emission processes result in the liberation of another electron from the cathode.[1] This is reasonable, since if each initial electron emitted from the cathode results in the emission of one secondary electron, the secondary electron can assume the role originally played by the initial electron, and the discharge will continue even if the initial emission from the cathode ceases.

Townsend's and Thomson's theories appear to be equally satisfactory, and it has not been possible to show experimentally that one alone is sufficient. The dependence of ignition voltage upon cathode material would appear to strengthen Thomson's theory. Since changes of gas and of gas mixtures, addition of metal vapors, and changes of pressure also have a great influence on the ignition voltage,[2] it seems probable that both secondary emission at the cathode and ionization by collision of positive ions, and perhaps other processes, such as increased photoelectric emission resulting from the radiation emitted by the discharge, act simultaneously. The presence of high-frequency fields in the vicinity of the gas may cause wide variations in ignition voltage. On occasion these fields may cause a 50 per cent decrease in the ignition potential. It appears probable that the high-frequency fields increase the kinetic energy of ions in the gas and hence the likelihood of ionization of other particles.

11-6. Theory of Ignition. Nonuniform Fields.—Equations (11-1) to (11-5) are based upon the assumption that the field is initially uniform and remains undistorted by space charge. In general the field is distorted by nonplanar electrodes and by space charge formed in the vicinity of the electrodes as the current builds up.

Townsend's theory for nonuniform fields leads to the following criterion for ignition:[3]

[1] DARROW, *ibid.*

[2] PENNING, F. M. and ADDINK, C. O., *Physica*, **1**, 1007 (1934); KESAEV, I., *USSR Tech. Phys.*, **4**, 1 (1937); BELJAWSKY, A., *USSR Tech. Phys.*, **4**, 493 (1937).

[3] THOMSON and THOMSON, *op. cit.*, Vol. II, p. 520. SLEPIAN, *op. cit.*, pp. 119–125.

$$\int_0^d \beta_n \epsilon^{-\int_0^s (\beta_n - \beta_p)\,du}\, ds = 1 \tag{11-6}$$[1]

in which the variables of integration, u and s, and the electrode spacing d are measured with the cathode as the origin. If the applied potential is reversed, so that the anode is at the origin, the condition for the advent of a self-maintaining discharge is[1]

$$\int_0^d \beta_p \epsilon^{-\int_0^s (\beta_p - \beta_n)\,du}\, ds = 1 \tag{11-7}$$

Since β_p and β_n are different and do not vary in the same manner with field strength, Eqs. (11-6) and (11-7) are not in general satisfied at the same value of applied voltage. This indicates that, when the field is nonuniform, the ignition voltages are different for the two polarities. A well-known example of this is the lower ignition voltage for a point-to-plane discharge when the point is negative than when the plane is negative. Another interesting example is the difference in ignition potentials for rod gaps caused in part by the distorting influence of the ground plane on the field strength.[2] In a discharge tube in which the two electrodes are of the same material but are not of the same size, the ignition potential is usually lower when the smaller electrode serves as the cathode.

Thomson's theory gives the following condition for ignition when the field is nonuniform:

$$\epsilon^{\int_0^d \beta_n ds} = \frac{1 + k}{k} \tag{11-8}$$

or

$$\epsilon^{\bar{\beta}_n d} = \frac{1 + k}{k} \tag{11-9}$$

where

$$\bar{\beta}_n = \frac{1}{d} \int_0^d \beta_n\, ds \tag{11-10}$$

Equation (11-10) defines $\bar{\beta}_n$ as the average value of β_n. A polarity effect with nonuniform fields is also predicted by Eq. (11-8), because k depends upon the field strength at the cathode. For electrodes of unequal size, the field strength will usually be greater at the cathode, and k will have the higher value, when the cathode is the smaller electrode. The value of $\bar{\beta}_n$, and hence the applied voltage, required for ignition decreases with increase of k.

[1] Note that u and s are merely variables of integration in the direction of the tube axis.

[2] BELLASCHI, B., *Trans. Am. Inst. Elec. Eng.*, **52**, 564 (1933); HAGENGUTH, J. H., *Elec. Eng.*, **56**, 67 (1937).

For the special case of uniform field, for which β_n and β_p are constant, Eqs. (11-6) and (11-7) reduce to Eq. (11-3), and Eq. (11-8) reduces to Eq. (11-5).

11-7. Time Required for Ignition.—Ignition is not an instantaneous process but requires a finite time which depends upon the applied voltage and the amount of original ionization.[1] The time lag in ignition may be attributed mainly to the mass of the positive ion, which requires some time to move to the cathode and to accelerate to a velocity sufficient to produce ionization by collision. For this reason the rapidity with which the ionizing takes place increases, and the time taken for ignition decreases, with increase of applied voltage (see Fig. 11-6).

11-8. Paschen's Law.—The ignition potential is found experimentally to be a function not only of the kind of gas and cathode material but also of the product of gas pressure and electrode spacing. It is possible to show theoretically from a consideration of Eqs. (11-3) and (11-5) that this should be true. β_n and β_p, the number of ions formed by an electron or a positive ion respectively, in moving a unit distance depend upon the gas pressure. Increase of pressure at constant temperature decreases the average distance between molecules. It therefore increases the number of collisions made by an electron or positive ion per centimeter of advance. Because of the shorter average distance moved between collisions, however, the electron or ion acquires less energy from the field between collisions and is less likely to ionize a molecule that it strikes. Since β_n and β_p depend upon gas pressure, the exponents of Eqs. (11-3) and (11-5) vary with the product of pressure and electrode spacing, $p \times d$, and the ignition voltage is a function of $p \times d$. The general form of the curve relating ignition voltage to the product of pressure and electrode spacing, first suggested by Paschen, is shown by the curve of Fig. 11-4*a*. The relation between the ignition potential and $p \times d$ shown by this curve is called *Paschen's law*.[2] Figure 11-4*b* shows an experimental curve for plane electrodes in air.[3] The increase of ignition voltage to the left of the

[1] TOROK, J. J., *Trans. Am. Inst. Elec. Eng.*, **47,** 349 (1928), **48,** 239 (1930); SCHADE, R., *Z. tech. Physik*, **17,** 391 (1936); WILSON, R. R., *Phys. Rev.*, **50,** 1082 (1936); HAGENGUTH, *loc. cit.*

[2] PASCHEN, F., *Wiedemann Annalen*, **37,** 69 (1889).

Paschen's law is only one of a number of laws grouped under the name *principle of similitude*, which states that if the relative geometry of a discharge remains the same, although the actual dimensions are changed, the current and ignition voltage remain the same. (See, for instance, SLEPIAN, *op. cit.*, pp. 139, 141, 146.) If, for example, the length of a discharge tube is doubled and the pressure decreased to one-half, the ignition voltage does not change; if all areas are tripled, the pressure remaining the same, then the current density decreases to one-third of the original value. The principle is a useful tool in the analysis of glow discharges.

[3] SLEPIAN and MASON, *loc. cit.* See also THOMSON and THOMSON, *op. cit.*, Vol. II, p. 489; COBINE, *op. cit.*, Sec. 7-8.

minima of Figs. 11-4*a* and 11-4*b* is caused by reduction with pressure and electrode spacing of the number of molecules between the electrodes and hence of the number of collisions an electron can make in moving from the cathode to the anode. The increase to the right of the minima is caused by the facts that increase of pressure decreases the spacing between molecules and so the distance through which an electron is accelerated between collisions, and that increase of electrode spacing decreases the field strength at a given potential. Increase of $p \times d$ therefore decreases the kinetic energy acquired by an electron between collisions, and hence the likelihood of ionization by collision.

It can be seen from Fig. 11-4*a* that at any pressure there is an electrode spacing that will result in minimum ignition voltage. Moreover, if the

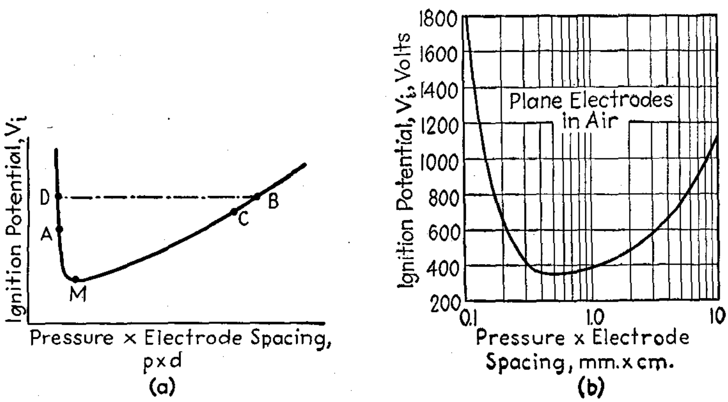


FIG. 11-4.—(*a*) General form of a curve of ignition potential as a function of the product of pressure and electrode spacing. (*b*) Experimentally determined curve of ignition potential as a function of the product of pressure and electrode spacing, for plane electrodes in air.

electrode spacing is made very small, the ignition voltage may be very high. This fact may be applied in preventing discharge between two electrodes in a gas-filled tube. If the greatest distance between points on the two electrodes is considerably less than that corresponding to point M of the ignition potential curve, the voltage required for ignition may be very high. This is sometimes called the *Hittorf principle* or, because the discharge does not always take place between the closest points of two electrodes if the distance between these points corresponds to a point to the left of the minimum of the ignition potential curve, the *short-path principle.*

Another interesting example of the application of Paschen's law has been given by Penning.[1] Using coaxial cylindrical electrodes and an axial magnetic field, Penning found that the ignition potential was varied

[1] PENNING, F. M., *Physica*, **3**, 873 (1936).

by changing the strength of the magnetic field. The effect of the axial magnetic field is to lengthen the paths of electrons moving to the anode and thus to increase the effective length of the tube. In Penning's tube the anode-to-cathode distance was made so small that the product of pressure by electrode spacing corresponded to a point to the left of the minimum of the ignition potential curve, such as A in Fig. 11-4a. The apparent tube length could be increased up to 1000 times, and the ignition voltage reduced to 0.01 of its original value.

11-9. Breakdown.—As soon as appreciable current flows, formation of space charge in the vicinity of the cathode causes almost the entire applied voltage, which was initially distributed uniformly over the whole cathode-to-anode distance, to become concentrated in the cathode dark space. The field strength in the remainder of the discharge is small, and the values of β_n and β_p are correspondingly small outside of the dark space.

The integration of Eqs. (11-6) and (11-8) may be done in two parts: over the region of the cathode dark space, and over the region between the dark space and the anode. Outside of the dark space the values of β_n and β_p are so small that the second part of the integration contributes very little to the total result. Hence, in order that a self-maintaining discharge shall continue after the field has become distorted by space charge, it is necessary and sufficient that the criterion for ignition be satisfied throughout the cathode dark space. The result is the same as though the effective length of the tube were reduced from its original value to a much smaller value equal to the length of the cathode dark space. Under the assumption that the ignition potential curve of Fig. 11-4, which applies to a uniform field, is applicable at least in its general aspects to the nonuniform field in the cathode dark space, it is possible to predict from this curve the behavior of the discharge subsequent to ignition.

Assume first a combination of electrode spacing and constant pressure such that the product $p \times d$ corresponds to a point to the left of the minimum ignition potential, such as A in Fig. 11-4a. Increase of current and formation of the dark space subsequent to ignition would reduce the effective length of the discharge. For the point A, however, a reduction in length must be accompanied by an increase of ignition voltage. It follows that the current cannot increase without an increase in applied voltage. The theoretical behavior of tubes with a low product of pressure by electrode spacing is verified by experimental curves, which are of the form of the dashed curve bm of Fig. 11-1. Unless the cathode becomes hot enough to emit thermionic electrons, the discharge does not require ballast resistance to limit the current, and breakdown, as defined on page 417, does not occur.

Analysis of the behavior of the discharge for a value of $p \times d$ to the right of the minimum ignition potential, say at B of Fig. 11-4a, must account for the fact that, when the current is kept small by means of external resistance, only a portion of the cathode surface may be covered by glow; *i.e.*, the discharge may be "normal." Suppose, first, that when ignition takes place the discharge is uniformly distributed over the cathode and that the resulting current density following ignition gives a cathode dark space which reduces the effective length of the discharge to a value corresponding to point C. If, for any reason, the discharge shrinks, so as to cover only part of the cathode, and if the current does not decrease, the current density must increase. It is predicted theoretically and verified experimentally that increase of current density is accompanied by reduction in length of the dark space.[1] But for point C, reduction of effective length of the discharge lowers the voltage necessary to maintain the discharge. Thus reduction of the area covered by discharge is accompanied by reduction of the terminal voltage, if the current does not fall. If the battery voltage is assumed to remain constant, the voltage across the resistance must rise. This can take place only if the current rises, increasing the current density still further. The initial shrinking may be caused by slight differences of field over the surface of the cathode, and the action immediately becomes cumulative. As soon as ignition occurs, the discharge area shrinks, and the effective length decreases until the voltage becomes equal to the minimum ignition potential corresponding to point M of Fig. 11-4a, beyond which further shrinking would be accompanied by rise of voltage. The shrinking takes place very rapidly and is accompanied by an increase of current to a value determined by the external circuit, and by decrease of voltage across the tube. The abrupt increase of current and decrease of tube voltage have already been defined as *breakdown.*

If the discharge after breakdown is "normal," an increase of current resulting from a change of battery voltage or circuit resistance might be accompanied either by increase of current density at constant area of discharge or by increase of area at constant current density. The former would involve decrease in length of cathode dark space, and hence increase of voltage; the latter does not. Hence in the "normal" current range the portion of the cathode covered by the discharge increases, the voltage remaining practically constant. This is the portion df of the current-voltage curve of Fig. 11-1. Any slight variation in the observed voltage may be attributed largely to changes in potential drop in the positive column, negative glow, and Faraday dark space. These potential drops were neglected in the analysis by confining the integration of Eq. (11-6) or Eq. (11-8) to the dark space, the contribution of the remainder of the discharge to the integral being considered negligible. Small

[1] SLEPIAN, *op. cit.*, p. 139; THOMSON and THOMSON, *op. cit.*, Vol. II, p. 423.

changes in voltage may also result from minor variations of cathode drop caused by irregularities over the cathode surface.

If the current is allowed to become so large that the cathode is completely covered, further increase of current must be accompanied by increased current density, decreased length of dark space, and increased voltage, as shown by the "abnormal" range *fg* of the current-voltage curve of Fig. 11-1. On the ignition voltage curve this corresponds to points to the left of the minimum.

If the external resistance is low, the current will, of course, jump to "abnormal" values as soon as ignition and breakdown take place. As previously explained, if the external resistance is negligible the current will rise to the value *n* in Fig. 11-1, corresponding to point *D* in Fig. 11-4*a*. In practical tubes this current is usually so high that, if it is maintained, the cathode becomes hot enough to emit electrons, and the discharge changes abruptly into an arc.

11-10. Distinction between Ignition and Breakdown Voltage.—It is necessary in general to distinguish between ignition and breakdown. With plane electrodes and values of $p \times d$ corresponding to points to the left of the minimum of the ignition potential curve of Fig. 11-4*a*, ignition is not accompanied or followed by breakdown, even when the voltage is raised to very high values; for values of $p \times d$ corresponding to points to the right of the minimum, on the other hand, ignition is accompanied by breakdown. In certain types of discharge, notably between needle points, the change in field distribution with current is such that with small currents the effective length of discharge first increases and then decreases with increase of current. Initially the field strength is very high in the immediate vicinity of the electrode points and the criterion for ignition must be satisfied only in a small volume adjacent to the cathode point. At very small currents, space charge weakens the field at the cathode and increases the distance from the cathode throughout which the criterion for ignition must be satisfied. The effective length therefore increases with current. At atmospheric pressure the product of pressure and electrode spacing corresponds to a point to the right of the minimum of the ignition potential curve and increase of effective length with current raises the voltage required to maintain the discharge. For a limited range of current, therefore, the voltage necessary to maintain discharge between needle points at atmospheric pressure rises with current and the discharge is stable, being observed as *corona*. At somewhat higher current, however, the space charge increases the potential gradient near the cathode and thus again decreases the distance throughout which the criterion for ignition must be satisfied. The voltage required to maintain the discharge then decreases with increase of current and breakdown immediately occurs. The current-voltage curve is of the form shown in Fig. 11-5. The breakdown voltage for needle points is thus higher than the ignition potential. A similar phenomenon is observed at high pressure when a wire of small diameter is used as the cathode, as in high-voltage power transmission.

Breakdown is not an instantaneous process, the total time required for breakdown being the sum of that required for ignition and that for the current to rise to its final value after ignition. Breakdown time, like ignition time, varies with applied voltage. Figure 11-6 shows the variation of breakdown time with crest alternating voltage of a 10-in. rod gap in air at atmospheric pressure.[1]

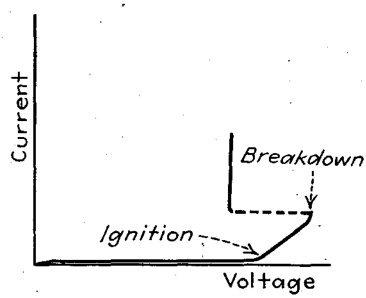


FIG. 11-5.—Current-voltage curve for discharge between needle points.

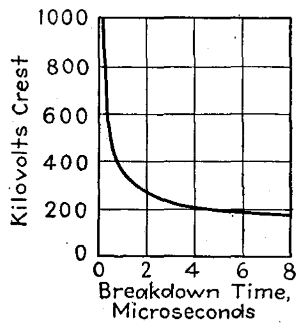


FIG. 11-6.—Variation of breakdown time with crest alternating voltage; 10-in. rod gap in air at atmospheric pressure.

11-11. Application of the Current-voltage Diagram.—The manner in which the current and voltage of a glow tube vary with supply voltage and circuit resistance under static conditions can be determined from the static current-voltage characteristic by the use of a resistance line similar to those used in high-vacuum-tube studies. The slope of the resistance line is the reciprocal of the external resistance, and the line passes through the voltage axis at a point corresponding to the battery voltage. The intersection of the load line with the static characteristic determines the current and the tube voltage. If a series of battery voltages is used, such as A, B, . . . J, corresponding points 1, 2, . . . 10 are determined, as shown in Fig. 11-7a. From 1 to 4 the current is in the dark-current range and is usually less than a microampere. (For clarity, the dark current is exaggerated in Fig. 11-7.) At 4, ignition and breakdown occur, the current immediately jumping to 4′, which would ordinarily correspond to several milliamperes or more. As the battery voltage is decreased from F to J, the characteristic is followed from 6 to 10. At 10 a further decrease in current requires an increase of voltage, while the voltage is actually reduced under the conditions of the experiment. The current

[1] HAGENGUTH, *loc. cit.*

therefore falls abruptly to the value 10′ in the dark range. The corresponding curve of current *vs.* battery voltage is shown in Fig. 11-7*b*. A similar analysis may be made when the applied voltage is maintained constant at a value exceeding the ignition potential and the resistance is varied. It is apparent that a glow tube and high resistance form a

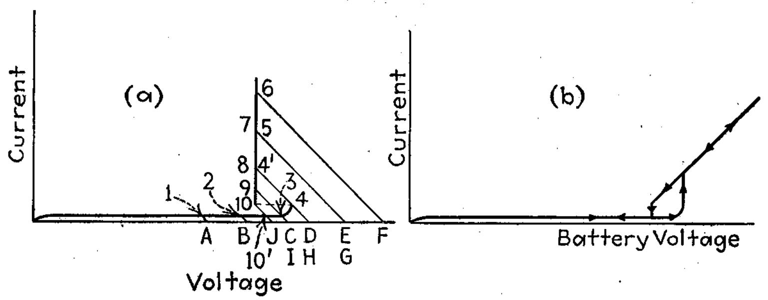


FIG. 11-7.—(*a*) Current-voltage diagram of a glow tube; (*b*) corresponding curve of current *vs.* battery voltage.

trigger circuit similar to the high-vacuum-tube circuits discussed in Chap. 10.

The voltage corresponding to point 10, at which the current falls to a low value and the tube "extinguishes," is called the *extinction potential* or *breakoff voltage.*

11-12. Dynamic Current-voltage Characteristics.—The dynamic current-voltage characteristics of a glow-discharge tube may be studied by

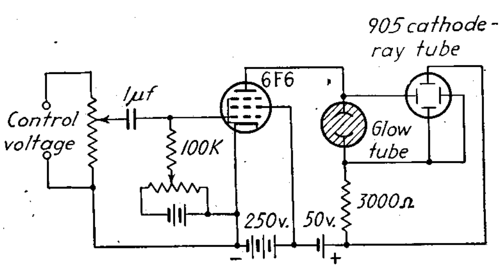


FIG. 11-8.—Circuit for the determination of dynamic current-voltage characteristics of glow tubes.

causing the current through the tube to vary in a predetermined way and observing the current and voltage by means of a cathode-ray oscillograph.[1] The wave form of the current is controlled by means of a high-vacuum pentode, to the grid of which is applied a voltage wave of the desired form and frequency, as shown in Fig. 11-8. If the voltage across

[1] REICH, H. J., and DEPP, W. A., *Phys. Rev.*, **52**, 245 (1937) (abstr.); *J. Applied Physics*, **9**, 421 (1938).

the discharge tube is applied to one pair of deflecting plates of the oscillograph (see Sec. 15-17) and the voltage across a resistance in series with

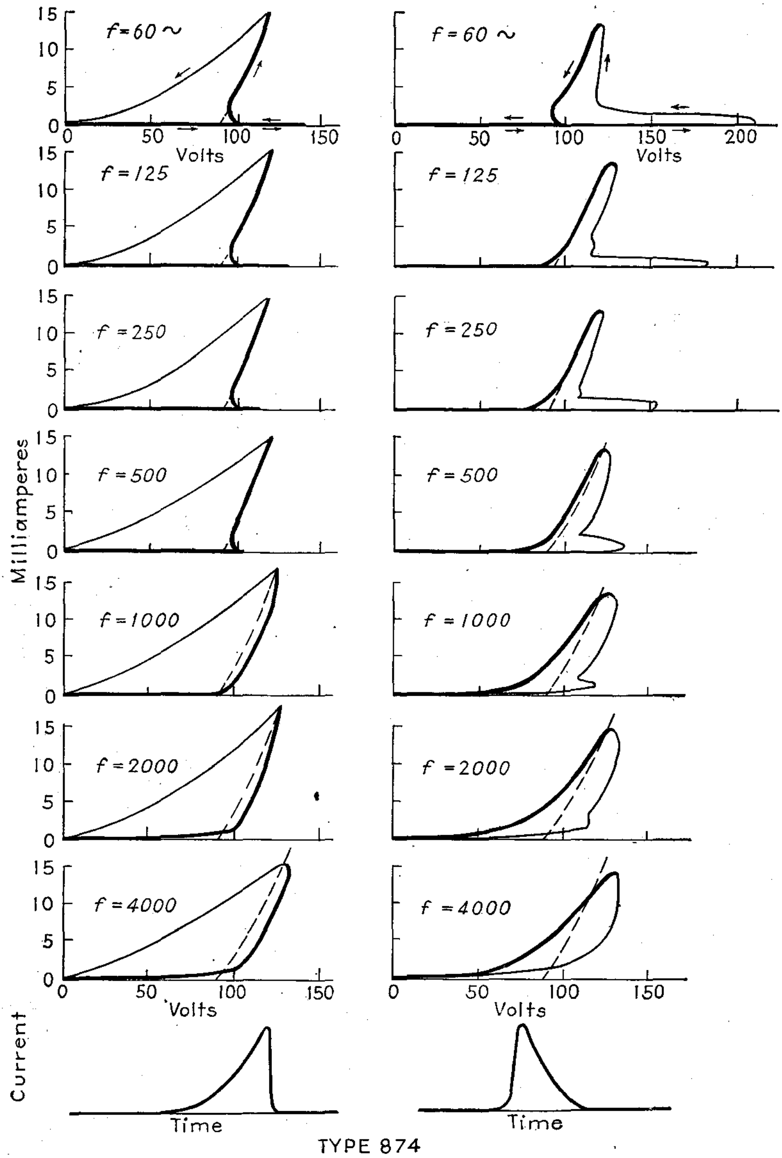


FIG. 11-9.—Dynamic current-voltage characteristics of type 874 glow tube for seven frequencies and two current wave forms. Light lines show rapid variation, heavy lines slow variation, of current with time.

the tube to the other pair, the current-voltage curve is observed on the screen of the cathode-ray tube.

Figure 11-9 shows a series of dynamic characteristics for an old-type (argon-filled) 874 tube, obtained with the rod as cathode (see Fig. 11-3) and the current in the "abnormal" range. The manner in which the current is varied is shown by the bottom oscillogram in each column. The static characteristic is indicated by the dashed curve. Similar characteristics are obtained with other types of glow tubes.

Space does not allow a complete analysis of these curves, but the following facts and their probable explanations are of interest:

1. When the current rises abruptly, the ignition potential is considerably higher than the static ignition potential. This results from the time required for ignition to take place. The voltage across the tube increases so rapidly that it exceeds the static ignition potential before sufficient ionization can build up to result in a self-sustaining discharge. Another way of stating this is to say that the voltage required to cause the current to build up at the rate determined by the control tube exceeds the static ignition potential.

2. The ignition potential falls with increase of frequency of the current wave and with increase of amplitude, the fraction of the cycle during which the current rises being maintained constant. (Dependence of ignition potential upon current amplitude is shown by a set of characteristics not included in Fig. 11-9.) If the current and frequency are made sufficiently high, the ignition voltage may be lower than the static extinction voltage. The reduction of ignition potential is undoubtedly explained by the presence of ions that have not had adequate time to recombine or to leave the interelectrode space. If the current is high and little time intervenes between removal and reapplication of voltage, the ion density may be high enough to allow the formation of a self-sustaining discharge at voltages considerably lower than the static ignition voltage.

3. When the current is decreased abruptly, it does not follow the static characteristic but, instead, a slightly curved path to the origin. This is explained if it is assumed that the time required for deionization greatly exceeds the time taken for the current to fall to zero. The increase in tube resistance accompanying deionization is evidenced by the decrease in slope of the characteristic as the current falls.

4. When the current rises abruptly or, in any case, when the frequency is high, the portion of the characteristic corresponding to increasing current lies to the right of the static characteristic. This is probably caused by the time taken for the ionization to build up beyond the value requisite to ignition. While the current is increasing, the value corresponding to a given voltage is lower than the current that would be obtained after the current had risen to the equilibrium (static) value for that voltage.

5. The characteristics obtained with gradually increasing current appear to indicate that, if the current could be held constant after ignition at the value prevailing at the time of ignition (the threshold value), the voltage would fall abruptly to a value not greatly in excess of the static extinction potential.

Many other interesting dynamic characteristics may be obtained by the use of other types of current waves. Figure 11-10 shows characteristics, for instance, that are obtained when the current is not reduced to zero.

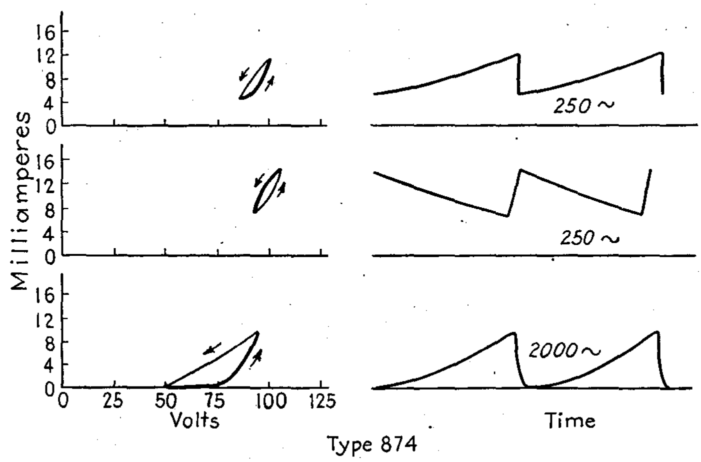


FIG. 11-10.—Three dynamic current-voltage characteristics obtained when the glow-tube current is not reduced to zero.

11-13. Grid Control of Glow Discharge.—The voltage at which breakdown occurs can be controlled by means of a grid placed between the cathode and the anode, as illustrated in principle by the sketch of Fig. 11-11. This tube is assumed to be of such construction that the entire cathode surface is shielded from the anode by the grid before breakdown. The grid is so close to the anode (about 1 mm) that the grid-anode ignition potential greatly exceeds the grid-cathode ignition potential, and the grid-anode discharge, if it does form, is of the stable type corresponding to the dotted curve *bm* of Fig. 11-1.

In the use of a grid-controlled glow-discharge tube the anode supply voltage E_{bb} exceeds the cathode-anode ignition potential. If the grid voltage E_c is zero, no discharge can take place between grid and cathode. The anode is positive with respect to the grid by the entire supply voltage E_{bb}, but breakdown between anode and grid is prevented by their close spacing. Breakdown does not occur between anode and cathode because the field of the anode terminates on the grid, rather than on the cathode. If the grid voltage E_c is gradually raised, ignition and breakdown occur between grid and cathode when the grid voltage equals the grid-cathode ignition potential. If the anode is sufficiently more positive than the grid, the discharge at once transfers to the anode. The anode voltage at which the transfer occurs decreases with increase of grid-cathode current. Resistance in the cathode lead limits the grid and anode currents to their rated values.

After breakdown the grid has no further control over the discharge. If the grid is made negative with respect to the region of the discharge in which it is situated, positive ions are drawn toward the grid wires, and electrons are repelled from them. In the immediate vicinity of the grid wires, therefore, positive-ion space charge predominates. The resulting positive-ion *sheath*, illustrated in Fig. 11-12, builds up to such a thickness that at its outer boundary the negative field caused by the

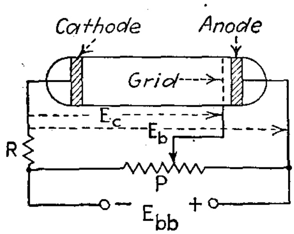


FIG. 11-11.—Diagram illustrating the principle of the grid-controlled glow-discharge tube.

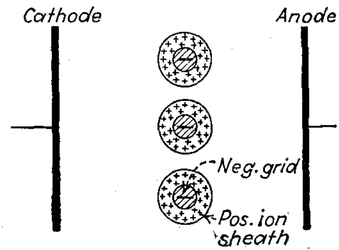


FIG. 11-12.—Sketch depicting the formation of a positive-ion sheath around a negative grid in an ionized gas.

voltage applied to the grid is just balanced by the positive field from the sheath.[1] Beyond the boundary of the sheath, in the main part of the discharge, the discharge is completely shielded from the grid by the sheath, so that electrons and positive ions are not influenced by the grid. If the grid is positive with respect to the discharge in its vicinity, electrons are drawn toward the grid, and positive ions are repelled from it. An electron sheath of sufficient thickness to shield the main body of the discharge from the positive grid is thus formed. Although the thickness of the sheath is constant at a given grid voltage, the ions or electrons within the sheath are constantly drifting toward the grid and being replaced by others that enter across the boundary. As the result of the movement of charges to the grid, a grid current flows in the grid circuit.

If the spacing between the grid wires can be made sufficiently small so that the sheaths surrounding adjacent wires overlap, the anode will be completely shielded from the cathode by the sheath, and the dis-

[1] LANGMUIR I., *Science*, **58**, 290 (1923); LANGMUIR, I., and MOTT-SMITH, H. M., JR., *Gen. Elec. Rev.*, **27**, 449, 538, 616, 762, 810 (1924); TONKS, L., MOTT-SMITH, H. M., JR., and LANGMUIR, I., *Phys. Rev.*, **28**, 104, 727 (1926); TONKS, L., and LANGMUIR, I., *Phys. Rev.*, **34**, 876 (1929); LANGMUIR, I., *Phys. Rev.*, **33**, 964 (1928).

charge will stop.[1] Unfortunately, however, the thickness of the sheath, like that of the cathode dark space, decreases with increase of current density of the discharge. At values of anode current density that are high enough to be of practical value, the sheath thickness is usually too small to make complete grid control possible. The only external effect of variation of grid voltage after the tube has broken down is change in the grid current. Because electrons are more readily accelerated than the more massive positive ions, the electron grid current that flows when the grid is positive is considerably greater than the positive-ion grid current that flows when the grid is negative.

The grid also prevents the tube from breaking down if it is unconnected, or "floats." If electrons and positive ions are in temperature equilibrium, the electrons have a much higher average velocity than the more massive positive ions. When an insulated electrode is immersed in an ionized gas, many more electrons than positive ions will therefore strike it per second. The positive ions recombine with electrons at the surface, but the excess electrons, which cannot leak off the insulated electrode, charge it negatively. If no electrons or positive ions were present, or if the electrode were struck by equal numbers of electrons and positive ions, it would assume a potential equal to the space potential at that point, which is intermediate between the potential of the cathode and the anode. The negative charge acquired by the electrode lowers its voltage relative to the space until the negative field which it produces attracts additional positive ions and repels electrons in such numbers that as many positive ions as electrons strike the electrode per second. If the grid of the tube of Fig. 11-11 is left unconnected, it assumes a potential that may be only slightly positive with respect to the cathode. The small dark current, which must flow between anode and cathode before ignition can occur, results in sufficient ionization to charge the grid in this manner and thus prevent ignition and breakdown. When this charge is allowed to leak off the grid so rapidly that the resulting grid current is at least equal to the threshold grid-cathode current, grid-to-cathode ignition and breakdown take place, and the current may transfer to the anode if the anode voltage is somewhat higher than the grid voltage after breakdown.

Since a grid-controlled glow-discharge tube has some of the characteristics of the high-vacuum trigger circuits discussed in Chap. 10, it is classified as a *trigger tube*.

11-14. Arc Discharges.—So far in this chapter the discussion has related to glow discharges, which are characterized by comparatively small currents and by voltages considerably greater than the ionization potentials of the gases. The high voltage drop in the glow discharge is

[1] LÜBCKE, E., *Z. tech. Physik*, **8**, 445 (1927).

accounted for mainly by the fall in potential across the cathode dark space. According to Thomson's theory, the high cathode drop is necessary to maintain the current flow, for in this region the positive ions are given the requisite energy to eject electrons from the cathode by bombardment. The increase of cathode drop with current in the "abnormal" range of current may be explained by the greater energy required to eject the additional electrons involved in the increase of current.

That this explanation of the cathode drop in potential is valid appears to be proved by the use of a cathode which can be heated to a temperature high enough to cause thermionic emission. It is found that, when the cathode is heated, the voltage drop across the discharge falls. If the thermionic emission is sufficient to supply all the electrons required for the current, the potential across the tube drops to a value that is usually approximately equal to the ionization potential of the gas or vapor. Because of cumulative ionization (see Sec. 1-5) it may be as low as the first excitation potential.

Discharges in which current density is high and the voltage drop is of the order of the ionization potential are commonly called *arcs*. Distinction between glows and arcs on the basis of voltage drop and current density is not a fundamental one. The essential distinction lies in the copious emission of electrons at the cathode of an arc by some process other than positive-ion bombardment. Possible emission processes include thermionic emission or emission resulting from the intense electric field caused by positive-ion space charge at the cathode. Glow currents are normally limited to low densities because of heating of the cathode and subsequent change to an arc when the density is high. High-current-density glows may be maintained for time intervals that are too short to allow the cathode to heat to the temperature required for emission (as in the discharge of a large condenser through a glow tube). Likewise, low-current-density arcs may be maintained if the cathode is heated by external means.

The transformation of a glow into an arc does not afford a satisfactory method of forming an arc in cold-cathode tubes with ordinary solid metallic cathodes because the resulting high temperature required to give the necessary emission will usually destroy the cathode. By the use of a special cathode made in the form of a cup containing a cesium compound, enough emission can be obtained at low cathode temperature to allow an arc to be maintained without damage to the cathode (see Sec. 12-61).[1] Although the peak current that can be handled by such a cathode may be of the order of several hundred amperes, the average current is usually a fraction of an ampere. More commonly used is the mercury pool type of cold-cathode arc, in which electron emission from

[1] GERMESHAUSEN, K. J., and EDGERTON, H. E., *Elec. Eng.*, **55**, 790 (1936).

the cathode is probably obtained because of high positive-ion space-charge field at the surface of the mercury cathode. The arc may be started by making and breaking contact between the anode and the mercury cathode, or between an auxiliary anode and the cathode, or by other means. After the tube fires, discharge takes place between the anode and a luminous area on the mercury surface, called the *cathode spot*, which furnishes electrons for the discharge. The area of the cathode spot is very small, and the current density at the spot is of the order of 1000 to 5000 amp/cm^2.* Positive-ion bombardment depresses the surface of the mercury in the vicinity of the cathode spot. The tendency of the spot to climb the sides of the depression, and consequent displacement of the depression, causes the spot to skim over the surface unless means are provided to fix its position.[1] Very large currents can be carried by mercury pool tubes. A third type of practical arc-discharge tube is that in which copious electron emission is obtained by the use of a cathode of high emission efficiency, heated by a separate electrical source.

11-15. Arcs with Separately Heated Cathodes.—Usually practically the entire voltage impressed between the electrodes of an arc is accounted for by the cathode drop, which is approximately equal to the ionization potential. In arcs with separately heated cathodes, cumulative ionization may result, so that the cathode drop need only be equal to the first excitation potential. Furthermore, the formation of space charge in regions other than the cathode dark space may result in fields that are in opposition to the applied field, a large concentration gradient carrying the electrons to the anode by diffusion against the opposing field. When this is so and the electrode spacing is small, the applied voltage may be less than the ionization potential and the cathode fall still be equal to the ionization potential.[2] (A similar action may be observed in glow discharges. Figure 11-2 shows that the voltage across the negative glow opposes the cathode drop. If the electrode spacing were only half as great, the voltage across the tube could be less than the cathode drop.) Arcs have been maintained with an applied voltage of only 0.5 volt in sodium vapor, 1.7 volts in mercury vapor, and 3.5 volts in helium. In practical tubes, however, the voltage does not ordinarily fall below the first ionization potential. It varies somewhat with current and with gas pressure or vapor temperature.

* There appears to be evidence that, when the current is large, it may divide among a large number of spots, each of which carries from 5 to 15 amp. See H. W. LORD, *Electronics*, May, 1936, p. 11.

[1] WAGNER, C. F., and LUDWIG, L. R., *Elec. Eng.*, **53**, 1384 (1934).

[2] COMPTON, K. T., and ECKART, C., *Phys. Rev.*, **25**, 139 (1925); DARROW, *op. cit.*, p. 384.

The low cathode drop that distinguishes an arc from a glow is maintained only as long as an emission current equal to the anode current is obtained without the need of a higher drop. When the arc current in a tube with a separately heated cathode exceeds the thermionic emission current from the cathode, a positive-ion space charge develops near the cathode that raises the cathode drop sufficiently to supply the required additional electrons by positive-ion bombardment. Although the secondary emission from the cathode or the increased thermionic emission resulting from the rise in cathode temperature may not be harmful to the cathode, oxide-coated cathodes disintegrate under positive-ion bombardment when the ion velocity exceeds a critical value. Under severe conditions the coating may be completely removed from the core. The value of the cathode drop above which the positive-ion velocity becomes so high as to deactivate or destroy coated cathodes is called the *disintegration voltage*. For the inert gases and mercury vapor the disintegration voltage lies between 20 and 25 volts.[1] The ionization potentials of some of the noble gases and of mercury vapor are sufficiently lower than the disintegration voltage to make it possible to design hot-cathode arc-discharge tubes in which the cathode does not disintegrate if the average anode current is kept below the value corresponding to the electron emission from the cathode. Such tubes will pass currents considerably in excess of the static emission current for periods of time that are too short to allow much rise in temperature of the cathode.

In arc tubes the anode voltage is usually nearly independent of anode current in the normal operating range of current. For this reason the current is determined practically entirely by the applied voltage and the circuit parameters. The supply voltage and circuit parameters of hot-cathode arc tubes must always be chosen so that the average anode current does not exceed the emission current.

11-16. Grid Control of Arcs.—The most commonly used method of controlling the firing of hot-cathode arc tubes is by means of a grid. The grid acts primarily as a shield between the anode and the cathode, preventing electrons emitted by the cathode from moving into the accelerating field between the grid and the anode. Inasmuch as only a few electrons in the grid-anode space are enough to start the discharge, the effectiveness of the grid depends upon how completely the grid shields the cathode. Hence, the largest hole in the grid structure is the determining factor, rather than the average mesh as in high-vacuum tubes. By proper design of electrode shape and spacing, the grid may be made to shield the cathode so completely that a positive grid voltage is required to start the arc. It is sometimes convenient to indicate the effectiveness

[1] HULL, A. W., *Trans. Am. Inst. Elec. Eng.*, **47**, 753 (1928); *Gen. Elec. Rev.*, **32**, 213 (1929).

of the grid of a negative-grid tube by means of the *grid-control ratio*, which is defined as the ratio of the anode voltage to the negative grid voltage that will just prevent firing. Because the grid-control ratio is not constant, it is usually necessary to use, instead, a curve of critical grid voltage *vs.* anode voltage. Such a curve is called a *grid-control characteristic*. Grid-control characteristics for typical tubes will be shown in Chap. 12 (see Figs. 12-25, 12-27, 12-29, and 12-31).

The critical grid voltage depends not only upon the electrode structure and the anode voltage, but also upon the pressure, which, in mercury vapor tubes, is a function of the temperature of the condensed mercury. The vapor pressure not only affects the critical grid voltage directly, but at high temperature the grid-anode ignition potential may be so low that a glow may form between the grid and the anode at relatively low anode voltage, and the grid thus lose control. At low temperature, on the other hand, the vapor density may be too low to make possible sufficient ionization for the neutralization of negative space charge near the cathode. The voltage therefore cannot fall to the low value that is characteristic of an arc, and the current is limited by space charge.

Grid control can also be used in mercury pool tubes provided with an auxiliary anode, called the *keep-alive electrode*. An auxiliary arc between the keep-alive electrode and the cathode, maintained by direct voltage, provides a cathode spot that supplies the electrons necessary for breakdown to the main anode. The disadvantage of a keep-alive electrode is that the continuous source of ions reduces the peak negative voltage that can be applied to the anode without danger of breakdown and formation of an arc.

Experiments have shown that the time required for complete breakdown of grid-controlled arc tubes is of the order of a few microseconds.[1] No study appears to have been made of the mechanism of breakdown. The following is a possible theory: Electrons that pass beyond the grid are accelerated by the field between the grid and the plate and cause ionization in the grid-anode region, the resulting positive ions being drawn to the grid. Because the positive ions are greater in number than the initial electrons that pass through the grid, a positive-ion sheath is formed adjacent to the grid. As the grid is made less negative, or more positive, more initial electrons enter the grid-anode space, and the ionization increases. The resulting increase of charge density in the sheath strengthens the field in the vicinity of the grid. At some critical grid voltage the field strength becomes great enough to cause glow ignition. This is immediately followed by arc breakdown to the cathode.

[1] NOTTINGHAM, W. B., *J. Franklin Inst.*, **211**, 271 (1931); HULL, A. W., and SNODDY, L. B., *Phys. Rev.*, **37**, 1691 (1931) (abstr.); SNODDY, L. B., *Physics*, **4**, 366 (1933).

In negative-grid tubes another possible factor in the breakdown process is the shielding of the grid by positive-ion space charge, which causes the grid to become ineffective in holding back electrons from the cathode. Transfer of the discharge from the grid to the cathode is probably caused by the high voltage drop through the dark space that forms at the grid when glow ignition occurs, the number of electrons from the cathode that pass beyond the grid being insufficient to neutralize the positive-ion space charge. In positive-grid tubes the transfer is favored by the greater voltage that exists between the cathode and the anode than between the grid and the anode. In some tubes the critical grid voltage at low anode voltage may be so positive that the initial breakdown is between grid and cathode. In order that the discharge may transfer to the anode the grid current must then exceed a minimum value which decreases with increase of anode voltage.

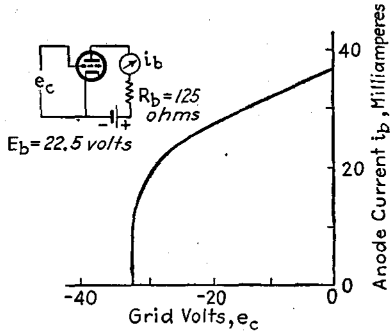


FIG. 11-13.—Transfer characteristic for type 884 arc-discharge tube, showing the complete control of anode current by the grid.

11-17. Completeness of Grid Control.—It is usually stated that the formation of an ion sheath about the grid causes the grid to lose control of an arc discharge when the arc starts. If the positive-ion sheath can be made so thick, however, that it extends completely across the openings in the grid, the arc may be stopped. Since the sheath thickness decreases with increase of anode current (ion density), complete grid control can be achieved only at small currents, or by the use of very small openings in the grid.[1] In the type 884 argon-filled tube, small anode currents can be interrupted at low anode voltage by a negative grid voltage of about the same magnitude as the anode voltage.[2] A typical curve of anode current *vs.* negative-grid voltage for this tube is shown in Fig. 11-13.

Up to the present time the construction of arc tubes in which the grid can interrupt or maintain complete control of large anode currents has not proceeded beyond the experimental stage. Kobel[3] was able to stop a 100-amp arc by making the grid several hundred volts negative relative

[1] LÜBCKE, E., *loc. cit.;* LE VAN, J. D., and WEEKS, P. T., *Proc. I.R.E.*, **24**, 180 (1936); BOUMEESTER, H., and DRÜYVESTEYN, M. K., *Philips Tech. Rev.*, **1**, 367 (1936); *Electronics*, May, 1937, p. 66 (abstr.).

[2] REICH, H. J., *Elec. Eng.*, **55**, 1314 (1936); LION, K. S., *Helvetica phys. Acta*, **12**, 50 (1939).

[3] KOBEL, A., *Schweiz. elektrotech. Verein Bull.* **24**, 41 (1933).

to the cathode. The experiments of Lüdi[1] in stopping arcs show that the maximum diameter of the holes in the grid should be less than 2 mm for extinction control. These tests also show that the magnitude of the negative grid voltage required to stop the arc increases with increase of pressure and with increase in diameter of the holes in the grid. The grid was able to stop arcs in 10^{-4} sec, indicating that deionization was very rapid. A combination of magnetic field and high negative-grid voltage has also been used to stop the anode current.[2]

11-18. External Grid Control.—An electrode on the outside of the tube, in contact with the glass, can induce a charge on the inner surface of the wall or change the charge that collects from the ionized gas. If the electrode structure is such that negative charge on the walls of the envelope can prevent emitted electrons from leaving the cathode, then such an external grid can be used to control firing of a hot-cathode arc.[3] Because charge induced on the inside of the glass wall rapidly becomes neutralized by charge of opposite sign from the ionized gas, the external grid voltage is effective only for short time intervals, and the external grid is of principal value in a-c operation.

An external electrode may also be used to initiate breakdown of a mercury pool tube containing a small amount of inert gas. The sudden application of high voltage to the external electrode causes the formation of a glow, with subsequent transition into an arc between the anode and the cathode. The objection to this type of control lies in the lowering of the glow ignition voltage by the use of gas.

11-19. Control by Magnetic Field.—Another type of control of breakdown of hot-cathode arcs is by means of magnetic fields.[4] The electrode structure of certain types of arc-discharge tubes provided with grids is such (see Figs. 12-24, 12-26, and 12-28), that a field at right angles to the tube axis deflects the electrons in a manner to cause them to strike the grid rather than to enter the grid-anode region. A field parallel to the axis, on the other hand, tends to focus the electrons into a beam directed at the anode (see Sec. 1-21) and thus increases the number entering the grid-anode region. Strong radio-frequency fields, especially at very high frequencies, also affect the firing of grid-controlled arcs.[5]

[1] LÜDI, F., *Helvetica phys. Acta*, **9**, 655 (1936).

[2] RISCH, R., *Schweiz. elektrotech. Verein Bull.*, **26**, 507 (1935).

[3] CRAIG, PALMER H., *Eletronics*, March, 1933, p. 70; WATANABE, Y., and TAKANO, T., *J. Inst. Elec. Eng. Japan*, **54**, 131 (1934); CRAIG, P. H., and SANFORD, F. E., *Elec. Eng.*, **54**, 166 (1935).

[4] SAVAGNONE, R., *Elettrotecnica*, **19**, 689 (1932); REICH, H. J., *Rev. Sci. Instruments*, **3**, 580 (1932), *Electronics*, February, 1933, p. 48; WATANABE, Y., and TAKANO, T., *J. Inst. Elec. Eng. Japan*, **53**, 62 (1933); KANO, I., and TAKAHASHI, R., *J. Inst. Elec. Eng. Japan*, **35**, 83 (1933); MCARTHUR, E. D., *Electronics*, January, 1935, p. 12; PENNING, *loc. cit.*; JURRIAANSE, T., *Physica*, **4**, 23 (1937).

[5] REICH, *loc. cit.*

11-20. Igniter Control.—A fourth method of controlling the firing of a mercury arc is by means of an igniter. This method is used exclusively with mercury pool tubes. The igniter is a high-resistance refractory rod which dips into the mercury pool, as shown in Figs. 11-14, 12-92 and 12-93. If the potential applied to the rod is such that the gradient along the rod exceeds a critical value of about 100 volts per centimeter, an arc forms between the rod and the mercury pool.[1] This arc constitutes a source of electrons for the anode-cathode system, and the main arc then strikes between anode and cathode. It has been found by experiment that the main arc strikes within a few microseconds of the igniter arc.[2]

The igniter rod is made of a crystalline material that is not wet by mercury, such as silicon carbide. Consequently, contact with the mercury is made at a large number of small points. The small area of these points, together with the relatively high resistance of the rod material, results in such a high potential gradient at the points of contact that either electrons are drawn from the mercury or thermal explosion occurs, resulting in the formation of tiny arcs. The action is similar to the formation of an arc between separating contacts. The currents in these minute arcs rapidly rise. If the resistivity of the rod is not too great, the resulting electron emission from the cathode spots becomes adequate to allow the formation of the main arc between the mercury and the anode. Igniter-controlled tubes will be discussed further in Chap. 12.

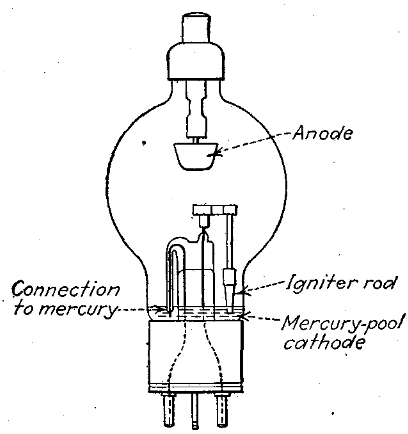


FIG. 11-14.—Construction of typical glass ignitron. (*Courtesy Westinghouse Electric & Manufacturing Co.*)

In practice the igniter arc may be formed by discharging a condenser through the igniter and by other methods which will be discussed in Chap. 12.

11-21. Arc Initiation by Auxiliary Glow.—Arcs can be initiated in cold-cathode tubes containing an inert gas by establishing a glow discharge between two auxiliary electrodes or between an auxiliary electrode and either the anode or the cathode.[3] In a tube of proper design, glow ignition is immediately followed by arc breakdown between the main electrodes if the anode power supply is capable of supplying sufficient

[1] SLEPIAN, J., and LUDWIG, L. R., *Trans. Am. Inst. Elec. Eng.*, **52,** 693 (1933).
[2] DOW, W. G., and POWERS, W. H., *Elec. Eng.*, **54,** 942 (1935).
[3] GERMESHAUSEN and EDGERTON, *loc. cit.*

current to maintain an arc and if the auxiliary glow current exceeds a minimum value.

11-22. Deionization.—In both glow and arc discharges the problem of the deionization of the gas or vapor is of great importance. The time taken for deionization after interruption of the anode current limits the application of discharge tubes, particularly of the grid-controlled type. In order that the grid shall be able to prevent firing when the anode voltage is reapplied it is necessary that the ions remaining in the interelectrode space shall be so few that anode-to-grid ignition cannot take place and positive-ion sheaths cannot shield the grid. The completeness of deionization necessary to reestablish grid control increases with anode voltage. For this reason the time required for the grid to regain control

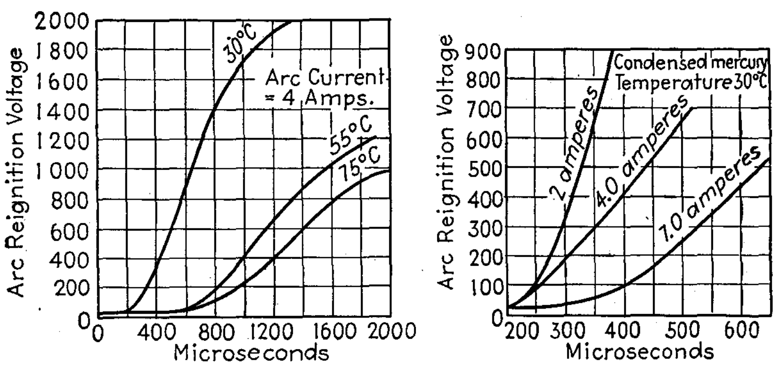


FIG. 11-15.—Typical curves of arc reignition voltage of grid-controlled hot-cathode arc tubes as a function of time after extinction.

increases with anode voltage, and a single nominal value of deionization time may not adequately indicate the performance of the tube. The behavior of the tube can be specified completely by means of curves of voltage necessary to cause reignition as a function of time after extinction, for given values of grid voltage, temperature or pressure, and current previous to extinction. Typical curves of reignition voltage *vs.* time are shown in Fig. 11-15.[1]

The most effective method of deionization in discharge tubes is by recombination at electrodes and walls (see Sec. 1-11), and so deionization is favored by reducing the distance that ions must travel to walls and electrodes. Experimental curves such as those of Fig. 11-15 show that pressure, magnitude of anode current, and potential of the grid relative to the surrounding space also affect the rate of deionization.[2] Because

[1] BERKEY, W. E., and HALLER, C. E., *Elec. J.*, **31**, 483 (1934).

[2] HULL, A. W., *Gen. Elec. Rev.*, **32**, 222 (1929); BERKEY and HALLER, *loc. cit.* See also S. S. MACKEOWN, J. D. COBINE, and F. W. BOWDEN, *Elec. Eng.*, **35**, 1081 (1934); J. D. COBINE, *Physics*, **7**, 137 (1936); J. D. COBINE, and R. B. POWER, *J. Applied Physics*, **8**, 287 (1937).

of increase of ion density, the deionization time increases with increase of current previous to extinction. Negative voltage impressed upon the grid or anode causes the positive ions to be pulled out of the interelectrode space, and so speeds up deionization. As the result of flow of positive-ion grid current, resistance in the grid circuit causes a lowering of negative grid voltage and thus slows down deionization. Increase of gas pressure retards diffusion to the walls, increasing the deionization time. In mercury vapor tubes the vapor pressure increases with condensed mercury temperature. The rate of deionization therefore decreases with increase of condensed mercury temperature.

The effect of deionizing time upon the dynamic characteristics of glow tubes has already been discussed. The manner in which the operation of glow and arc tubes in particular circuits is affected by deionization will be discussed in Chap 12.

11-23. Peak Forward Voltage and Peak Inverse Voltage.—Because of the possibility of the formation of a glow discharge between the anode and the grid and resultant loss of grid control, there is a limiting value of positive anode voltage above which a grid cannot prevent an arc tube from firing. This is called the *peak forward voltage*. Likewise, there is a limiting negative anode voltage, called the *peak inverse voltage*, above which a glow discharge may form between the anode and the grid or cathode. There is then the possibility of *arcback*, which is the flow of arc current in the reverse direction to the normal flow. When arcback occurs in an alternating-voltage circuit, the tube ceases to rectify. Both the peak forward voltage and the peak inverse voltage depend upon the material of which the electrodes are made (or with which they become coated), upon the kind of gas or vapor, and upon the gas or vapor pressure. In vapor tubes, the large change of vapor pressure with temperature makes these voltages dependent upon tube temperature. The peak forward voltage is greatly lowered if the grid becomes hot enough to give appreciable thermionic emission. Similarly, the peak inverse voltage is lowered by undue heating of the anode. In periodic operation of the tube, the peak forward and peak inverse voltages are also dependent upon the time intervening between extinction of the tube and reapplication of anode voltage and upon the magnitude of the anode current previous to extinction.

Supplementary Bibliography

KINGDON, K. H.: *Phys. Rev.*, **21**, 408 (1923).
See also references at the end of Chap. 1.

CHAPTER 12

GLOW- AND ARC-DISCHARGE TUBES AND CIRCUITS

The development of practical glow- and arc-discharge tubes based upon the theoretical principles presented in Chap. 11 has made possible new apparatus and instruments which are of great value in industry and in the laboratory. The design, characteristics, and applications of these tubes will be treated in this chapter.

12-1. Essential Facts concerning Glow and Arc Discharge.—The student will find it advantageous to review the following facts, summarized from Chap. 11:

1. A glow discharge is a discharge in which the electrons are released mainly by secondary emission at the cathode and by ionization by collision. It is characterized by cathode drop and anode voltage at least several times as great as the first ionization potential of the gas or vapor and usually, but not necessarily, by low current density.
2. The formation of a glow discharge is dependent upon the existence of a source of some initial electrons or positive ions. Probable sources are photoelectric emission, radioactivity, and cosmic radiation.
3. Ignition is the advent of a self-sustaining discharge, *i.e.*, one that will continue even if the initial source of electrons at the cathode is removed. Ignition takes place when the ionization becomes great enough so that one electron is released at the cathode for each initial electron that leaves the cathode. The anode voltage at which ignition takes place is called the *ignition voltage*.
4. Ignition may be accompanied or followed by breakdown, which is the incipient condition in which the current rises with no increase of applied voltage. Unless limited by external resistance, the current may rise to destructive values. The anode voltage at which breakdown takes place is called the *breakdown voltage*. The anode voltage at which the glow discharge ceases and the current falls to the dark-current range is called the *extinction voltage*.
5. If the cathode is only partly covered with glow, the discharge is said to be "normal." The "normal" cathode drop is constant and usually the anode voltage is also essentially constant. Increase of current in the "normal" range is accompanied by increase of the portion of the cathode covered by glow. If the cathode is entirely covered by glow, the discharge is said to be "abnormal." In the "abnormal" range the cathode drop and anode voltage rise with current.
6. Glow ignition and breakdown may be prevented by a grid, placed so close to the anode that grid-anode breakdown cannot take place. The grid loses control when breakdown occurs and therefore cannot be used to extinguish the anode-cathode current.

7. An arc discharge is one in which there is some source of copious electron emission at the cathode other than secondary emission. The usual sources of emission are thermionic or field emission. An arc discharge is characterized by a cathode drop of the order of the first ionization potential of the gas or vapor and often, though not necessarily, by high current density.

8. If the anode current of a hot-cathode arc exceeds the emission current, the cathode drop rises sufficiently to produce the necessary additional electrons at the cathode by secondary emission. The resulting positive-ion bombardment is likely to deactivate or destroy oxide-coated emitters.

9. The anode voltage of an arc is essentially constant, or may fall, in the working range of current. Hence the current must be limited by means of external resistance to prevent its rise to destructive values.

10. The formation of an arc discharge may be controlled by an internal grid, an external grid, magnetic field, radio-frequency electromagnetic field, an igniter rod, and by the initiation of an auxiliary glow discharge. Once the arc has formed, a grid or other control mechanism loses control of the discharge, and the arc must be extinguished and deionization take place before control is regained. The time required for deionization to be sufficiently complete to enable the grid to regain control depends upon the tube and electrode structure, the kind and pressure of the gas or vapor, the current flowing at the time of extinction, and the electrode potentials during deionization.

12-2. Applications of Glow Tubes.—It will be noted from the discussion which follows that for certain applications either glow or arc tubes may be used, although there are usually factors which make one or the other preferable. Applications of glow tubes include the following:

1. Production of light (modulated or for display).
2. Voltage regulation.
3. Rectification.
4. Oscillation.
5. Control of current or power.
6. Protection of apparatus or circuits.
7. Amplification.

For most of these applications special types of tubes have been developed.

Figure 12-1 shows symbols that will be used for glow-discharge tubes. Symbol (*a*) represents a two-element tube in which the electrodes are of equal area, symbol (*b*) one in which the area of one electrode exceeds that of the other, and (*c*) a grid-controlled glow tube. In such a tube the larger electrode ordinarily serves as cathode. The shading indicates that the tube contains gas or vapor in sufficient quantity to allow the formation of a glow or an arc.

12-3. Glow-discharge Tube as a Light Source.—A discussion of the glow-discharge tube in the form of the familiar "neon" sign is beyond the scope of this book. The principles outlined in Chap. 11 are applicable, but the particular requirements make this field a specialty.[1] A more pertinent application is the use of a glow-discharge tube as a source of modulated light. In the early development of television, the glow tube was used in this manner in the formation of the received image;[2] but, since the application of the cathode-ray tube to television reception, the glow tube as a source of modulated light has been of value principally in the recording of sound on motion-picture film. The value of the glow tube as a modulated-light source lies in the facts that over the proper range of current the illumination which it produces is directly proportional to the current and that response to changes in current is rapid.

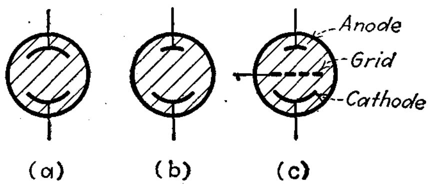


FIG. 12-1.—Symbols for glow-discharge tubes.

Tubes for the production of modulated light have been made in a variety of forms. The simplest, shown in Fig. 12-2*a*, consists of two parallel plane electrodes, placed so close together that the discharge takes place between the outer surfaces rather than the inner ones (see page 428).

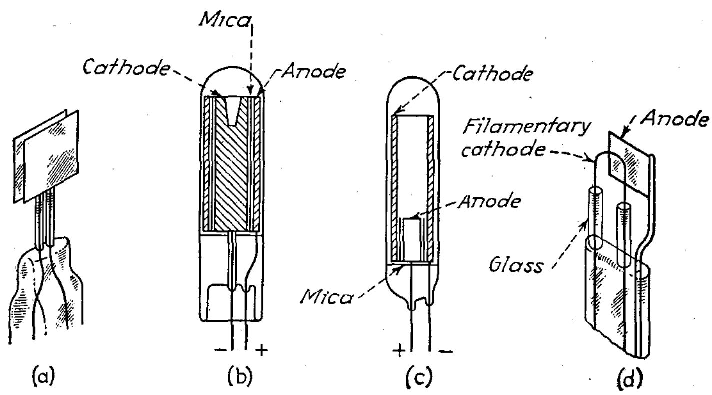


FIG. 12-2.—Tubes for the production of modulated light.

If the electrode spacing and gas pressure are correct, the outer surface of the cathode is uniformly covered with glow. Another type of tube is the *crater lamp* illustrated in Fig. 12-2*b*. The cathode consists of a solid cylinder which fits within a thin cylindrical anode and is separated from the anode by mica, and in the outer end of which is a cone-shaped depression. The design is such that the discharge takes place between the

[1] See, for instance, S. C. MILLER and D. G. FINK, "Neon Signs," McGraw-Hill Book Company, Inc., New York, 1935.

[2] See, for instance, E. H. FELIX, "Television—Its Methods and Uses," McGraw-Hill Book Company, Inc., New York, 1931.

cone-shaped crater and the outside of the anode. The high current density at the center of the cone results in a concentrated source of light. A modified form of the crater lamp is shown in Fig. 12-2c. In this tube the discharge takes place between the anode and the inside of the cathode. The *aeolight lamp* shown in Fig. 12-2d was developed for sound recording on motion-picture film.[1] The hot filament furnishes a weak source of electrons which reduce the cathode drop. The number of electrons supplied by the hot filament is not high enough to produce an arc.

In using glow-discharge tubes to obtain modulated light from modulated current, the voltage across the tube should at all times be high enough to prevent the tube from going out. If the voltage falls below the extinction potential, the tube goes out, and reignition will not occur until the voltage has risen to the higher ignition potential. The interruption of the light during this interval results in distortion. The

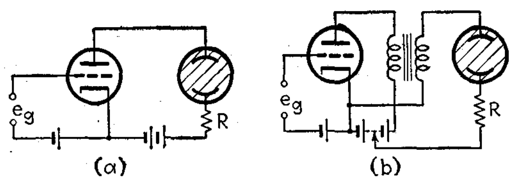


FIG. 12-3.—Circuits for the production of modulated light by means of glow tubes.

current should preferably be sufficiently high so that the cathode is completely covered with glow.

One of the simplest ways to use a glow tube as a source of modulated light is to connect it directly into the plate circuit of a power amplifier tube, as shown in Fig. 12-3a. Since the a-c resistance of the glow tube is small, especially in the "normal" range of current, a resistance should be used in series with the glow tube in order to reduce nonlinear distortion (see Sec. 3-25). The plate supply voltage must be adequate to take care of the drop through the glow tube and the resistance. A circuit that requires lower battery voltage and that has the additional advantage of allowing independent adjustment of the operating currents of the amplifier and glow tubes is shown in Fig. 12-3b. The resistance R acts both as a load for the power tube, through the transformer, and as a means of limiting and adjusting the average glow current. The transformer should be connected so that the primary and secondary d-c fluxes oppose each other.

The principle of constant current density in the "normal" glow may be used as the basis of an oscillograph. Two rod electrodes are mounted in the same plane, but not quite parallel, as shown in Fig. 12-4. The gas pressure is chosen so that the glow strikes first at the base of the elec-

[1] "Recording Sound for Motion Pictures," p. 71, edited by Lester Cowan, McGraw-Hill Book Company, Inc., New York, 1931.

trodes, and the tube is operated in the "normal" range of current. As the current is changed, the discharge covers a greater or smaller portion of the electrodes. The height of the column of light then indicates the magnitude of the current. In combination with a rotating mirror to provide a time axis, the tube makes a simple and inexpensive oscilloscope.[1] In order to prevent extinction of the discharge if the current is temporarily interrupted, a high-frequency field is sometimes applied to the lower part of the tube to act as a keep-alive device. The ignition potential is then close to the extinction potential, and distortion is decreased. The high-frequency field also reduces the voltage required by the glow tube. Another application of glow tubes is as tuning indicators in radio receivers.[2]

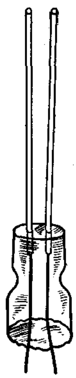


FIG. 12-4. Electrode structure of glow tube used as an oscillograph element.

12-4. The Glow-discharge Tube as a Voltage Stabilizer. Because a large increase of current in the "normal" range is accompanied by only a small increase of tube voltage, a glow tube may be used as a simple voltage stabilizer. The circuit is that of Fig. 12-5.[3] The resistance R is high enough to limit the tube current to the "normal" range when the load current is zero. Load current increases the voltage drop in the resistance and so reduces the tube voltage. A small reduction of tube voltage in the "normal" current range, however, results in a comparatively large decrease in tube current, which tends to decrease the IR drop. Small variations of load current, therefore, cause compensating variations of tube current, the voltage across the tube remaining essentially constant. A tube designed for this purpose should have large cathode area in order to give a large range of "normal" current. The 874 neon or (in an earlier model) argon-filled tube, which has a "normal" voltage of approximately 90 volts, an ignition voltage of approximately 120 volts, and a maximum current rating of 50 ma, was designed for this service. Other tubes, designed to operate at different useful values of voltage (indicated by the number in the type designation) are the VR-75, the VR-105, and the VR-150.

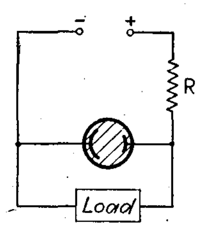


FIG. 12-5.—Circuit for using a glow tube as a voltage stabilizer.

[1] SUNDT, E. V., *International Projectionist*, July and August, 1937. Also private communication from Littelfuse Laboratories, Chicago.

[2] DREYER, J. F., JR., *Electronics*, February, 1933, p. 40; HEINZE, W., and POHLE, W., *E.T.Z.*, **56**, 917 (1935); MIRAM, P., *Funkt. Mon.*, **10**, 373 (1935); see also *Electronics*, January, 1936, p. 42 (abstr.).

[3] Glow tubes designed for purposes of illumination ordinarily contain a current-limiting resistance in the base. This resistance must be removed if the tube is to be used in the circuit of Fig. 12-5.

By the use of a number of tubes in series[1] or of the *stabilivolt*, which has a number of additional electrodes between the anode and cathode, as shown in Fig. 12-6, it is possible to obtain several stabilized voltages from a single source.[2]

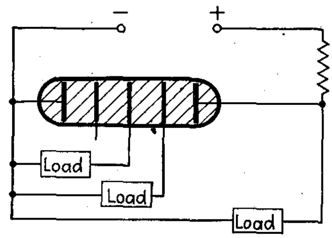


FIG. 12-6.—Multielectrode glow tube used to furnish two or more stabilized voltages.

12-5. The Glow-discharge Tube as a Rectifier.—Because of their high voltage drop and consequent low efficiency, as well as their small current capacity, glow tubes are now seldom used for rectification. Nevertheless, the principle of rectification by glow tubes is still of interest. A typical tube developed for this purpose is illustrated in Fig. 12-7*a*. The operation of the tube is based upon the manner in which the form of the static characteristic curve is affected by the cathode area. Curve *a* of Fig. 12-8 is for a cathode of small area. The "normal" current range is very small, so that at relatively small currents the voltage rises rapidly with current. Curve *b* is for a cathode of large area, for which the "normal" current range is large. Intersection of the load line of Fig. 12-8 with the two curves shows that for the same load resistance R and the

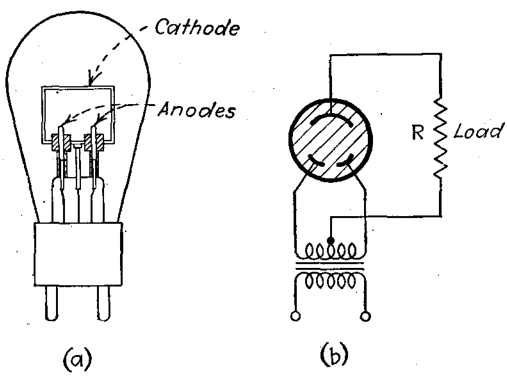


FIG. 12-7.—(*a*) Electrode structure of a full-wave glow-discharge rectifier; (*b*) circuit for its use.

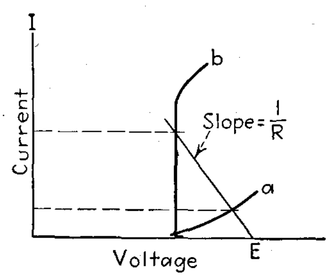


FIG. 12-8.—Current-voltage diagram of a glow tube with two electrodes of unequal area, showing how rectification is obtained.

same applied voltage E the current will be much higher for the large cathode area than for the small cathode area. If a tube having a small and a large electrode is connected to an alternating voltage through a load resistance, current will flow from the time in the cycle at which the sup-

[1] KÖRÖS, L., *E.T.Z.*, **50**, 786 (1929); *Wireless Eng.*, **6**, 460 (1929) (abstr.).

[2] NOACK, F., *Z. Ver. deut. Ing.*, **74**, 548 (1930); *Wireless Eng.*, **7**, 408 (1930) (abstr.); GOÇKEL, H., *Physik. Z.*, **38**, 65 (1937).

ply voltage exceeds the ignition voltage of the tube (not necessarily the same for both polarities) until it falls below the extinction voltage. For a given instantaneous applied voltage the current is much greater when the large electrode acts as cathode than when the small electrode acts as cathode, and so rectification takes place. The tube illustrated in Fig. 12-7*a* contains two small electrodes and one large one, so that full-wave rectification is obtained by the use of the circuit of Fig. 12-7*b*.

Rectification may also be accomplished by making use of the difference in ignition potentials for the two polarities of glow tubes having dissimilar electrodes.

12-6. The Glow-tube Relaxation Oscillator.—A useful application of the glow-discharge tube is made in the glow-tube relaxation oscillator, the simplest circuit of which is shown in Fig. 12-9. A glow tube, shunted by a condenser C, is connected to a d-c supply through a resistance R that is sufficiently high to limit the charging current to the dark-current range (usually not less than 1 megohm). When the circuit is first closed,

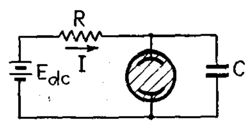


FIG. 12-9.—Basic circuit of the glow-tube relaxation oscillator.

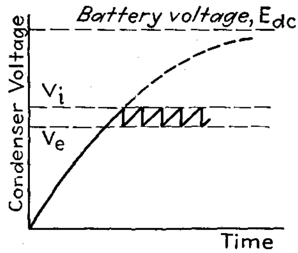


FIG. 12-10.—Variation of condenser voltage in the glow-tube relaxation oscillator of Fig. 12-9.

the condenser charges through the high resistance. The condenser voltage rises exponentially, as indicated by the relation

$$V = E_{dc}(1 - \epsilon^{-t/RC}) \tag{12-1}$$

The curve of condenser voltage is shown in Fig. 12-10. When the condenser voltage becomes equal to the ignition (breakdown) voltage V_i of the tube, the tube breaks down, and the condenser discharges through the tube. Since no resistance is usually used between the condenser and the tube, the discharge current is high, being limited by the small resistance and inductance of the connecting wires and by the tube characteristic. (The time during which high current flows is ordinarily too short to allow the electrode temperature to rise enough to result in thermionic emission, and so the current does not exceed that corresponding to point n of Fig. 11-1.) The condenser discharges very quickly to a voltage equal to the extinction potential V_e, at which the tube goes out, and the condenser begins recharging. The condenser voltage varies periodically between the voltages V_e and V_i at a frequency determined by these voltages and by the circuit constants.

The time of discharge is so short that the frequency of oscillation depends almost entirely upon the time taken to charge the condenser from the extinction potential to the ignition potential. This time increases with the difference between the ignition and extinction voltages, with the capacitance of the condenser, and with the resistance. It decreases with increase of applied voltage. By substituting V_i and V_e for V in Eq. (12-1) and solving for the two corresponding values of t, the following approximate expression for the oscillation frequency may be derived:[1]

$$f \cong \frac{1}{RC \log_\epsilon \dfrac{E_{dc} - V_e}{E_{dc} - V_i}} \tag{12-2}$$

Equation (12-2) is of value only because it shows the manner in which the various factors influence the frequency of oscillation; it cannot be used to predict exact values of frequency from known circuit constants, because the ignition and extinction voltages are themselves complicated functions of the frequency, the tube temperature, illumination of the electrodes, and the magnitude of the discharge current (as shown by Fig. 11-8).

Frequencies as low as 1 cycle in 15 min. may be readily obtained. The chief difficulty encountered in obtaining such a low frequency results from the increase of condenser leakage with condenser capacitance. The upper frequency limit usually does not exceed 10,000 cps and is obtained without the use of a condenser, the electrodes and leads furnishing the required capacitance. The upper frequency limit results because there is a lower limit to the resistance, below which the tube glows permanently. A complete analysis of the criterion for oscillation must be based upon the dynamic characteristics of the tube and is beyond the scope of this book.

Because the ignition and breakdown of glow tubes are dependent upon the presence of some initial ions, irregularity of oscillation is usually observed, particularly at very low frequencies, unless care is taken to provide a constant source of initial ions. If no light is allowed to strike the tube, the initial ionization is probably produced largely by cosmic radiation and is not constant. A small amount of light on the cathode is usually sufficient to supply enough photoelectrons to give stable opera-

[1] RIGHI, A., *Rend. accad. Bologna*, **6**, 188 (1902); PEARSON, S. O., and ANSON, H. S., *Proc. Phys. Soc. London*, **34**, 175, 104 (1922). For bibliography, see F. BEDELL and H. J. REICH, *J. Am. Inst. Elec. Eng.*, **46**, 563 (1927).

tion. Some change in oscillation frequency with illumination of the cathode is generally noted.[1]

Like other types of relaxation oscillators discussed in Chap. 10, the glow-tube oscillator can be locked into step with a frequency equal, or very nearly equal, to the oscillator frequency or a multiple of that frequency. The control voltage is best introduced through a transformer in series with the terminal of the tube that does not connect to the resistance, as shown in Fig. 12-11, and should preferably have a steep wave front.

The wave form of the condenser voltage is shown by Fig. 12-10. The wave shape is virtually that of a right triangle, except that the hypotenuse is curved because of the exponential rise of voltage. For given ignition and extinction voltages the curvature decreases with increase of battery voltage and may be made very small. A true saw-tooth wave of voltage in which the hypotenuse of the triangle is straight is obtained if the charging current is maintained constant.[2] During the charging period the condenser charging current is equal to the supply current I. If I is constant, the charge on the condenser at any time t after extinction of the tube is

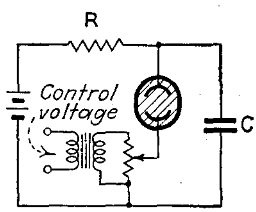


FIG. 12-11.—Glow-tube oscillator stabilized by control voltage.

$$Q = Q_0 + It \tag{12-3}$$

where Q_0 represents the charge on the condenser at the instant of extinction. The voltage across the condenser and the glow tube is

$$V = \frac{Q}{C} = \frac{Q_0}{C} + \frac{It}{C} = V_e + \frac{It}{C} \tag{12-4}$$

Equation (12-4) shows that the condenser voltage rises linearly during the charging period. If it is assumed that the discharge is instantaneous, the following expression for the oscillation frequency may be found by substituting V_e and V_i in Eq. (12-4) and solving for the difference between the two values of t:

$$f = \frac{I}{(V_i - V_e)C} \tag{12-5}$$

Evidently the use of a constant charging current makes the frequency independent of supply voltage, as would be expected. Since V_i and V_e are functions of frequency, it is not possible to use Eq. (12-5) in the exact determination of frequency.

[1] OSCHWALD, U. A., and TARRANT, A. G., *Proc. Phys. Soc. London*, **36**, 241 (1924); REICH, H. J., *J. Opt. Soc. Am.*, **17**, 271 (1928).

[2] See also Sec. 15-20.

Although nearly constant supply current can be obtained by replacing the resistance R of Fig. 12-9 by a voltage-saturated diode or a triode with positive grid voltage and low filament temperature, the best method is by the use of a voltage pentode, as shown in Fig. 12-12. At the very small currents (microamperes) required for charging the condenser, the plate current of a pentode is practically independent of plate voltage over a large range of voltage. An additional advantage of the pentode is the ease with which the plate current, and hence the frequency, can be controlled by means of the grid and screen voltages.

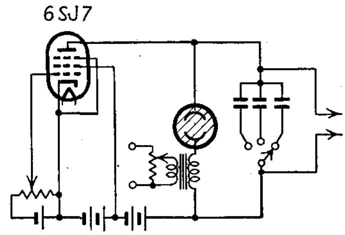


FIG. 12-12.—Form of glow-tube oscillator in which the condenser voltage rises linearly with time.

12-7. Applications of the Glow-tube Oscillator.—The glow-tube oscillator can serve as a convenient generator of audio-frequency voltage for applications in which high harmonic content is desirable, or at least not objectionable.[1] Since resistance across the glow tube changes the frequency, distorts the wave form, and may even stop oscillation, the oscillator must be followed by a capacitance-coupled or, preferably, direct-coupled amplifier if it is necessary to vary the output voltage or if current or power output is required. A typical circuit is shown in Fig. 12-13. Figure 12-14 shows an interesting modification of

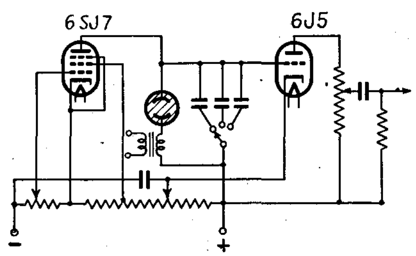


FIG. 12-13.—Glow-tube oscillator and direct-coupled amplifier.

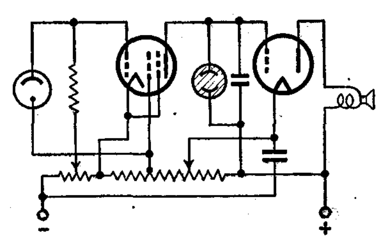


FIG. 12-14.—Circuit in which oscillator frequency is controlled by light.

this circuit in which the addition of a phototube makes possible the control of oscillator frequency by light.

Another modification of the circuit of Fig. 12-13 is obtained by connecting a relay in the plate circuit of the amplifier tube.[1] The relay opens and closes as the condenser charges and discharges, the frequency of operation being controlled by the condenser capacitance and the bias of the pentode. The fraction of the cycle during which the relay remains

[1] KOCK, W. E., *Physics*, **4**, 359 (1933); *Electronics*, March, 1935, p. 92; *Radio Eng.*, May, 1936, p. 17.

[2] REICH, H. J., *Rev. Sci. Instruments*, **2**, 164, 234 (1931).

energized can be controlled by varying the bias of the amplifier tube (see Prob. 12-1).

Other applications of the glow-tube oscillator to the measurement of resistance and capacitance and to the cathode-ray oscillograph are discussed in Chap. 15 (see Secs. 15-13 and 15-20).

12-8. Control of Power by Glow-discharge Tubes. Grid-glow Tube. Because of the lower voltage drop and higher current capacity of arc tubes, glow tubes are now used relatively little in the control of power. The advantage of the glow tube is that no cathode heating power is required.

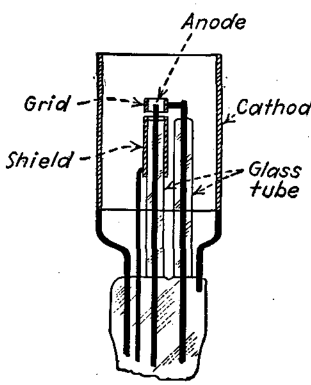


FIG. 12-15.—Electrode structure of the grid-glow tube.

A practical grid-controlled glow-discharge tube, the *grid-glow* tube,[1] is illustrated in Fig. 12-15. The cathode consists of a cylinder, approximately 1 in. in diameter and $1\frac{1}{4}$ in. long. The anode is a small wire which projects through the end of a glass tube and is surrounded at its base by a small cylindrical shield. The grid is also a short cylindrical electrode which surrounds the tip of the anode.

The use of a small anode accomplishes two results. As in the rectifier tube of Fig. 12-7*a*, very little current flows when the small electrode is negative, and rectification results when alternating voltage is used in the anode circuit. Secondly, the small size of the anode and grid accentuates the negative charging of the grid when it is free. During the flow of dark current between anode and cathode before breakdown occurs, the large number of electrons formed in the large volume outside of the grid are drawn toward the anode, and many of these strike the grid. The only positive ions that can strike the grid, on the other hand, are the relatively few formed in the small volume of gas between the grid and the anode. Thus the number of electrons that strike the grid exceeds the number of positive ions both because of the greater speed of the electrons and because the grid is placed at a point where large numbers of electrons are converging. Hence, when the grid is free, it is always only slightly positive relative to the cathode and, as explained in Sec. 11-13, grid-to-cathode breakdown cannot occur. The close spacing of the grid and anode prevents grid-to-anode breakdown.

The grid-cathode ignition potential of the grid-glow tube is about 270 volts, and the extinction potential about 170 volts. In order for the discharge to transfer from grid to anode after grid ignition, it is necessary

[1] KNOWLES, D. D., *Elec. J.*, **27**, 232 (1930); KNOWLES, D. D., and SASHOFF, S. P., *Electronics*, July, 1930, p. 183.

that the anode voltage should be higher than the grid voltage by an amount that depends upon the current flowing in the grid circuit. The manner in which the required difference between grid and anode voltages varies with grid current is shown in Fig. 12-16.

When the grid-glow tube is used with direct voltages, the anode current can be stopped only by reducing the anode voltage below the extinction value. In order that the grid may regain control, the grid voltage must also be reduced below the extinction potential, at least momentarily, and kept below the ignition potential. Since the grid-glow tube is a rectifier and the anode current is cut off at the time in each cycle when the anode voltage drops below the extinction potential, the tube may be conveniently used with alternating voltage. If the grid voltage is reduced below the ignition value, the grid regains control at the instant in the cycle when the tube goes out, and it prevents further flow of anode current until the grid voltage is again raised above the ignition value.

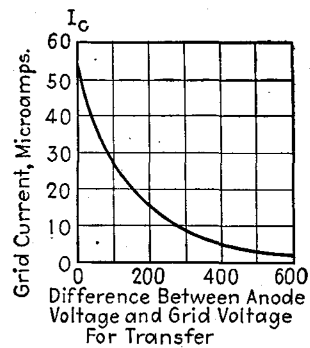


FIG. 12-16.—Curve relating grid current and difference between grid and anode voltages necessary to cause transfer of discharge from grid to anode of the grid-glow tube.

12-9. Grid-glow Tube Circuits.—The basic circuit for the grid-glow tube is that of Fig. 12-17, in which the grid potential is controlled by the potentiometer setting. Subject to the condition that the grid current and anode voltage are adequate to ensure transfer of the discharge to the anode, anode current flows when the grid potential equals the grid-cathode ignition potential. The tube can also be fired by increasing the applied voltage.

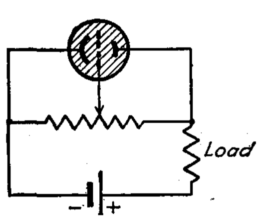


FIG. 12-17.—Basic d-c circuit for grid-glow tube.

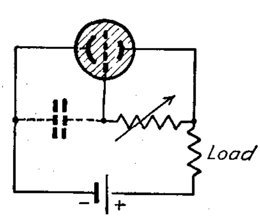


FIG. 12-18.—Modified form of d-c circuit for the grid-glow tube.

Omission of one portion of the voltage divider of the circuit of Fig. 12-17 simplifies the circuit to that of Fig. 12-18. If the variable resistance is sufficiently high, the dark current which flows in the grid circuit preliminary to breakdown reduces the grid voltage to a value less than the ignition potential. Reduction of the resistance raises the grid voltage. Breakdown requires not only that the grid voltage equal the ignition potential but also that the grid current be made large enough to allow transfer of the discharge from the grid to the anode at the prevailing anode voltage. The addition of a condenser between the grid and the cathode, as shown by the dotted lines, provides a relatively large surge of grid current when the grid voltage equals the ignition value; it thus ensures transfer of the discharge to the anode at this grid voltage. Two

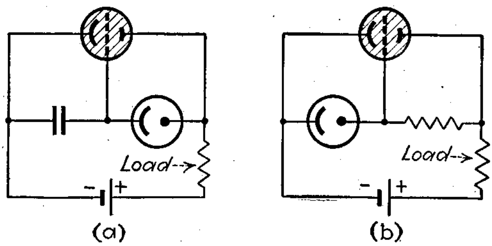


FIG. 12-19.—Two grid-glow tube circuits that incorporate phototubes.

circuits by means of which it is possible to use a phototube to control breakdown are shown in Fig. 12-19. With circuit *a*, increase of illumination causes breakdown; with circuit *b*, decrease of illumination causes breakdown (see also Sec. 13-11).

The circuits of Figs. 12-17 to 12-19 may be used with alternating as well as with direct voltage. When alternating voltage is used, variation of potentiometer setting or grid-circuit resistance changes the time in the cycle at which breakdown occurs and hence the portion of the cycle during which current flows. In this manner the average rectified anode current is under control, the grid in effect having complete control over the average anode current.

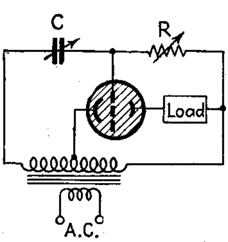


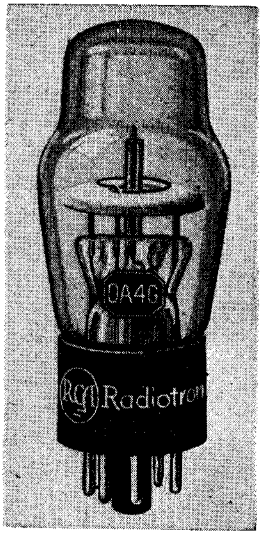
FIG. 12-20.—Grid-glow-tube circuit in which change of resistance or capacitance changes the average anode current.

In another type of control, known as *phase control*, the average anode current is regulated by the phase relation between the anode and grid voltages. In Fig. 12-20 a phase-control circuit is shown in which change of the resistance R or of the capacitance of the condenser C controls the average anode current. The average anode current increases with decrease of R or C. A complete explanation of phase control will be given in conjunction with the treatment of grid-controlled arc-discharge tubes. This explanation also applies, in its essential details, to grid-glow-tube phase control (see Sec. 12-43).

It should be noted that no voltage amplification is obtained with the grid-glow tube, the required grid-control voltage being of the same order of magnitude as the anode voltage. The value of the grid-glow tube lies in its current amplification. Currents of 8 or 10 ma can be controlled by means of changes of grid voltage with the flow of only a few microamperes of grid current and the expenditure of a correspondingly small amount of grid power.

The shield of the grid-glow tube may be maintained at a fixed potential, or used to control breakdown in the same manner as the grid.

12-10. Starter-anode Glow Tubes.—A different method of controlling anode-cathode breakdown is used in the OA4 "gas triode," shown in Fig. 12-21. This tube uses a *starter anode* in place of a grid.[1] The tube is designed so that breakdown does not occur directly between the main anode and the cathode unless the anode voltage is at least 225 volts. Breakdown occurs between the starter anode and the cathode, on the



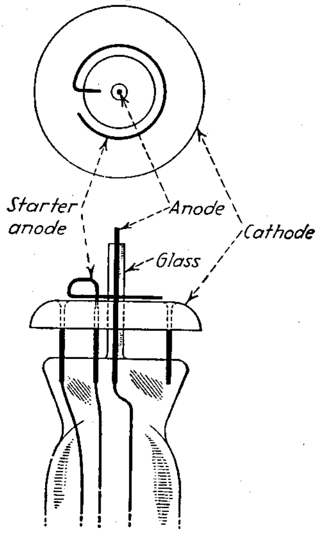


FIG. 12-21.—Starter-anode glow tube.

other hand, when the starter anode voltage is 90 volts or less. Starter-anode breakdown is followed by breakdown to the main anode if the starter-anode current and main-anode voltage are sufficiently great. The required value of starter-anode current increases from zero to approximately 250 ma with decrease of main-anode voltage from its breakdown value to a value equal to the starter-anode breakdown voltage. Since the main anode is very small in comparison with the cathode, the tube rectifies and may be used on alternating supply voltage. The tube can carry 25 ma continuously, which is ample for the operation of relays. Because of the very small current required in the starter-anode circuit, the tube may be controlled by voltages in tuned radio-frequency circuits.[1] It has other applications as a control device.

[1] BAHLS, W. E., and THOMAS, C. H., *Electronics*, May, 1938, p. 14; INGRAM, S. B., *Trans. A.I.E.E.*, **58**, 342 (1939).

12-11. Glow-tube Protective Devices.—Another interesting application of the glow tube is in the protection of electrical apparatus against surge voltages.[1] Connected across the apparatus, the glow tube serves as a by-pass for surge voltages above the ignition potential of the tube. It is evident that the extinction potential should be higher than the normal voltage of the apparatus which it protects, in order that the tube may extinguish when the surge is over. Heating of the electrodes may cause the glow to change into an arc. For this reason a current-limiting resistance must be provided if the circuit is capable of supplying sufficient steady current to maintain the arc. Glow tubes have been used to dissipate the energy released by the opening of a circuit containing a large inductance, such as an electromagnet or the field coil of a generator.

12-12. Glow-discharge Amplifier Tubes.—A number of glow-discharge tubes have been designed for the purpose of voltage and power amplification. In the most successful of these tubes a small glow discharge between auxiliary electrodes serves as a source of electrons which are controlled by the main electrodes in the same manner as in thermionic amplifier tubes.[2] In another type, complete control of the anode current over a limited current range is attained by the use of a special type of grid.[3] The disadvantages of these tubes outweigh their advantages, and so far they have been largely of theoretical interest.[4]

12-13. Arc-discharge Tubes.—In Chap. 11 it was pointed out that arc discharges are characterized by low voltage drop and high current density and that the difference between the characteristics of the glow and the arc results from the presence at the cathode of the arc of a copious source of emission other than secondary emission. Arc-discharge tubes may be either of the hot-cathode or the cold-cathode type. In the former the source of electrons at the cathode is thermionic emission; in the latter it is probably emission caused by very high fields at the cathode, resulting from space charge. Hot-cathode arc diodes (rectifiers) are of two types: high-pressure, which carries the trade name *tungar* or *rectigon;* and low-pressure, designated by the trade name *phanatron.* Grid-controlled hot-cathode arc tubes are known as *thyratrons.* This was originally also a trade name but has now been released for general use.

[1] Private communication from Littelfuse Laboratories, Chicago.

[2] Hund, August, "Phenomena in High-frequency Systems," pp. 271–277, McGraw-Hill Book Company, Inc., New York, 1936; Guntherschulze, A., and Keller, F., *Z. Physik*, **12**, 1, 8, 28 (1931); *Electronics*, December, 1931, p. 242. See also *Electronics*, January, 1933, p. 6.

[3] Reich, H. J., and Hesselberth, W. M., *Electronics*, October, 1933, p. 660.

[4] See also E. Lübcke and W. Schottky, *Wiss. Veröffentlich. Siemens-Konzern*, March, 1930, p. 9; F. Schröter, *Electronics*, April, 1935, p. 131 (abstr.), *Telefunken Röhre*, **1**, 103 (1935).

12-14. Advantages of Arc-discharge Tubes.—One advantage of the arc tube over both the glow tube and the high-vacuum tube is the high anode-circuit efficiency. Since the tube dissipation at a given anode current is proportional to the anode voltage, it follows that the dissipation in an arc-discharge tube having an anode voltage of from 10 to 20 volts under load is considerably smaller than in a high-vacuum tube which passes the same current at a plate voltage of from 250 to several thousand volts. In a mercury arc tube designed to operate from a peak supply voltage of 2000 volts and having an anode voltage drop of 15 volts under load, the anode circuit efficiency may approach the value

$$\frac{(2000 - 15)}{2000} = 99.25 \text{ per cent.}$$[1]

The advantage of the arc tube lies not only in the low value of the tube drop but also in its constancy. Because the tube drop does not vary with load, the voltage regulation of a rectifier using a mercury vapor arc tube is much better than that of a similar circuit using a high-vacuum rectifier tube. In high-vacuum tubes the high voltage drop results from electron space charge. In arc tubes there are practically equal numbers of electrons and positive ions throughout the tube, and, as explained in Chap. 11, the voltage drop in the rated current range is only that required to produce ionization. Another advantage of the hot-cathode arc tube over the high-vacuum tube is the higher cathode emission efficiency attainable. This will be explained later in this chapter.

The high anode and emission efficiencies of hot-cathode arc tubes make possible the control of large current by means of relatively small tubes. The FG-27 thyratron, for instance, has a maximum average current rating of $2\frac{1}{2}$ amp and a peak anode voltage rating of 1000 volts and can control 2500 watts of power. This tube is only slightly larger than the type 50 high-vacuum triode, which has a maximum operating plate current of 55 ma and a power output that cannot exceed about 20 watts, even in Class C operation. An important disadvantage of arc-discharge tubes over high-vacuum tubes is the incompleteness of grid control, the current after firing being determined by the applied voltage and the load impedance.

12-15. Tungar Rectifier.—The first type of hot-cathode arc rectifier to be developed was the tungar.[2] One of the principal functions of the argon gas or mercury vapor used in this tube is to protect the oxide-coated cathode against evaporation of the barium so that the cathode can be operated at high temperature and hence high emission efficiency. To accomplish this the pressure of the gas or mercury vapor must be about

[1] In a-c operation the efficiency may approach about 98 per cent.

[2] HULL, A. W., *Trans. Am. Inst. Elec. Eng.*, **47**, 753 (1928).

1 mm. Glow ignition voltage at this relatively high pressure is so low that the tubes cannot be used at high voltage, and satisfactory grid control cannot be attained. The principal application of tungar rectifiers has been in battery charging.

12-16. Cathode Structure of Low-pressure Hot-cathode Arc Tubes.[1] At low gas or vapor pressures, such as must be used in grid-controlled arc tubes and high-voltage arc rectifiers, the gas does not exert a protective action upon the cathode, and positive-ion bombardment may result in cathode disintegration. Since the cathode cannot be operated at high temperature, as in the tungar, high cathode efficiency must be attained by the use of special cathodes that have low heat loss. This is made possible by the characteristics of the arc discharge. In high-vacuum tubes the electrons that leave the cathode must follow the electric field, and any emitting surface to be of value must be placed so that some lines of force from the anode will terminate upon it. In the arc tube, on the other hand, as soon as the tube fires, the whole tube is filled with large, and practically equal, numbers of positive ions and electrons. The flow of current consists of a relatively slow drift of electrons toward the anode and of positive ions toward the cathode. Electrons in the immediate vicinity of the anode are drawn toward it, leaving excess positive charge behind them which immediately causes the advance of other electrons that are nearer the cathode. At the cathode, electrons that drift away are replaced by those emitted from the cathode. This drift is superimposed upon the high random velocity due to the temperature of the gas and need not take place along the initial field from cathode to anode. In effect, the presence of the ionized gas enables the electrons to drift in curved paths which do not coincide with the field, so that portions of the cathode surface that are not subject to the initial field set up by the anode voltage are effective in supplying electrons as soon as the tube is allowed to break down. Hence, cathodes may be made up in forms that greatly reduce the loss of heat. High emission efficiency is attained at the expense of rapid heating, however, since the greater heat capacity and reduced heating energy increase the time taken to establish thermal equilibrium. In practice a compromise must be made between high efficiency and short heating time. The heating time of commercial hot-cathode arc-discharge tubes ranges from 5 sec to 1 hr.

In Fig. 12-22*a* is shown the type of filamentary cathode originally used in the FG-27 thyratron. Heat that is radiated by any of the inner turns

[1] HULL, *ibid.* HULL, A. W., *Gen. Elec. Rev.*, **32**, 213 (1929); HULL, A. W., and LANGMUIR, I., *Proc. Nat. Acad. Sci.*, **15**, 218 (1929); HULL, A. W., *Physics*, **2**, 409 (1932), **4**, 66 (1933); LOWRY, E. F., *Electronics*, October, 1933, p. 280; December, 1935, p. 26; *Elec. J.*, April, 1936, p. 187; KNOWLES, D. D., LOWRY, E. F., and GESSFORD, R. K., *Electronics*, November, 1936, p. 27; PIKE, O. W., *Communications*, October, 1941, p. 5.

of the spiral filament is absorbed by adjacent turns. Radiation from the outer surface is reduced by the bright nickel cap that surrounds the filament. Heat loss may also be reduced by crimping the ribbon, as in

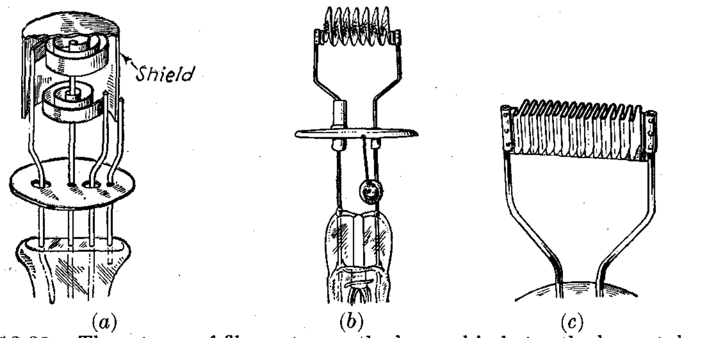


FIG. 12-22.—Three types of filamentary cathodes used in hot-cathode arc tubes.

the filament of Fig. 12-22*b*, or by folding it, as in the filament of Fig. 12-22*c*.[1] Figure 12-23*a* shows the construction of the indirectly heated cathode used in the FG-67 and other types of thyratrons. Heat loss is greatly reduced by reflection from the outer concentric cylinders. The active coating covers the outside of the inner cylinder which surrounds the heater, the inner surface of the first outer cylinder, and the con-

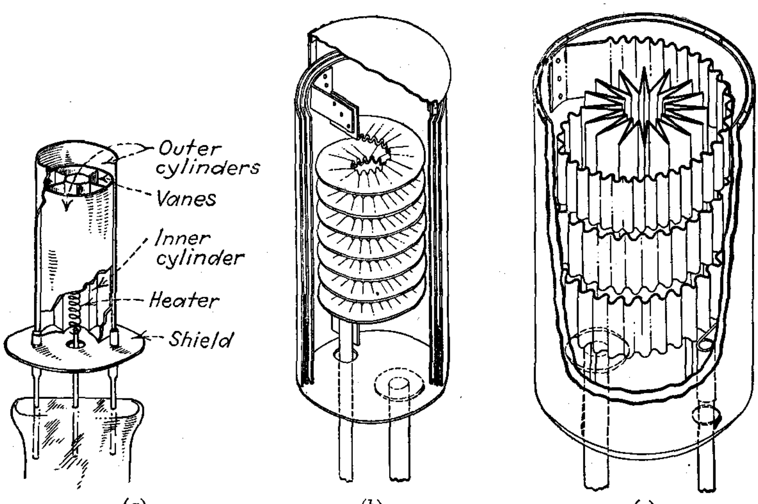


FIG. 12-23.—Three types of indirectly heated cathodes used in hot-cathode arc tubes.

necting vanes. Figures 12-23*b* and 12-23*c* show two cathodes that combine direct and indirect heating.[1] These cathodes have the oxide coating on the surfaces of the corrugated heating element and on the

[1] LOWRY, *loc. cit.*

inner surface of the inner cylinder. The cathode of Fig. 12-22*a* has an average emission current rating of $2\frac{1}{2}$ amp at a power consumption of 35 watts, that of Fig. 12-22*c* a peak current rating of 65 amp at 200 watts, and that of Fig. 12-23*a* an average emission current rating of $2\frac{1}{2}$ amp at $22\frac{1}{2}$ watts.

Two requirements must be satisfied in the design of cathodes for grid-controlled hot-cathode arc tubes. (1) The difference in potential between the ends of the filament or heater must be less than the ionization potential of the gas or vapor. Otherwise discharge takes place and produces a continuous supply of ions that prevent the grid from controlling the anode current. (2) In filamentary cathodes the crest potential drop through the filament must be less than the difference between the ionization and disintegration potentials (see Sec. 11-16). If the filament voltage is higher than this, the anode potential relative to the negative end of the filament will exceed the disintegration potential when the anode voltage relative to the positive end is equal to the ionization potential. Special care must be taken in preventing discharge between the ends of 110-volt heaters. Five-volt cathode heaters or filaments are generally used. In designing filamentary cathodes it is necessary to take into account the heating effect of the anode current in passing through the filament. This necessitates limitation of peak anode current to the order of magnitude of the filament current and of average anode current to considerably smaller values, even though the saturation emission current may be higher.[1]

12-17. Choice of Gas or Vapor.—Either mercury vapor or the inert gases may be used in hot-cathode arc-discharge tubes. Because the vapor pressure of mercury varies greatly with temperature, the characteristics of mercury vapor tubes are very susceptible to small changes of tube temperature such as may result from air currents or changes of cathode temperature. The disadvantages of the rare gases are their higher ionization potentials, which result in higher tube drop than with mercury; their very much lower ignition voltages, which greatly reduce the maximum alternating voltage that can be applied to the anode without danger of arcback and without loss of control by the grid; and the tendency of gases to "clean up," or combine with the electrode metals to form stable compounds and thus to change pressure. The tube drop under load is about 10 or 11 volts in mercury vapor tubes and about 16 volts in argon tubes. Mercury vapor is used in most arc tubes. Only a small quantity of mercury, usually not more than a few drops, is sufficient to supply ample vapor. The vapor pressure is controlled by the temperature of the tube, and usually lies in the range from 1 to 50 microns (20° to 70°C.). The lower limit of pressure is that at which

[1] LOWRY, *loc. cit.*

there are insufficient positive ions to neutralize electron space charge near the cathode, and the upper limit that at which the glow ignition potential is too low or at which the diffusion of positive ions is too slow to give sufficiently rapid deionization.

12-18. Tube Ratings.[1]—The amount of power that can be controlled by means of a hot-cathode arc tube depends upon the voltage and current that the tube can safely withstand. Voltage limitations exist because of the possibility of the production of a glow discharge between the grid and the anode (see Sec. 11-23). The glow results in loss of control by the grid because of space-charge sheaths formed about the grid. Current limitations exist because of danger of excessive anode heating and of destructive positive-ion bombardment of the cathode (see Sec. 11-15).

These terms are commonly used in rating arc rectifiers:

The *maximum peak inverse voltage* is the maximum instantaneous negative anode voltage that can be applied to the tube without *arcback*, breakdown in the reverse direction from that in which current normally flows. It depends upon electrode material; upon pressure of the gas or vapor and hence temperature in mercury vapor tubes; and upon residual ionization, which depends, in turn, upon the frequency of the anode supply voltage and the load current.

The *maximum peak forward voltage* is the maximum instantaneous positive anode voltage at which the grid can prevent firing. At higher voltage a glow forms between anode and grid, the grid loses control, and an arc forms between anode and cathode.

The *maximum instantaneous anode current* is the highest instantaneous periodic current that the tube can stand under normal operating conditions without damage to the anode because of overheating or to the cathode because of positive-ion bombardment. The length of time during which a given tube will stand this instantaneous current or the frequency with which it will stand an instantaneous current surge of a given duration depends upon tube heating.

The *maximum surge current* rating is a measure of the ability of a tube to stand extremely high transient currents. This rating is intended to form a basis for circuit design in limiting the abnormal currents that occur during short-circuit conditions. The tube cannot, however, be subjected to repeated short circuits without the probability of a corresponding reduction in life and the possibility of failure.

[1] Some of the definitions that follow are taken, with minor modifications, from the General Electric Co. bulletin GET-426. See also H. C. STEINER, A. C. GABLE, and H. T. MASER, *Elec. Eng.*, **51**, 312 (1932); O. W. PIKE, and D. ULREY, *Elec. Eng.*, **53**, 1577 (1934); Standards on Electronics, p. 4, Institute of Radio Engineers, New York, 1938.

The *maximum average anode current* is a rating based upon tube heating. It represents the highest average anode current that can be carried continuously through the tube. In the case of a rapidly repeating cycle of operation, this may be measured on a d-c meter. Otherwise, it is necessary to calculate the average current over a period not to exceed a definite interval of time which is specified for each design of tube. For instance, a tube with a maximum instantaneous anode current rating of 15 amp, a maximum average anode current rating of 2.5 amp, and an integration period of 15 sec could carry 15 amp for 2.5 sec out of each 15 sec or 7.5 amp for 5 sec out of every 15 sec.

The grid current ratings are given in terms of the *maximum instantaneous grid current* and the *maximum average grid current*, and the integration period is the same as for the anode current.

The *tube voltage drop* is the anode-to-cathode voltage under normal anode current flow. It will be represented by the symbol E_a. Tube drop varies with temperature in mercury vapor tubes.

The *deionization time*, specified in microseconds, is the time required under normal conditions to bring about sufficient deionization to regain grid control. The time is based on maximum average anode current and (in mercury vapor tubes) a condensed mercury temperature of 40°C. It decreases with decrease of temperature and of anode current. This factor is very important in many circuits, but, inasmuch as it depends upon tube temperature, grid-circuit regulation, and other factors, the specified figure should be considered as only relative (see Sec. 11-22).

The *grid-control characteristic* (approximate starting characteristic) is a graph relating the anode voltage with the critical grid voltage at which the tube fires. The grid-control characteristic is affected by tube temperature.

12-19. Breakdown Time of Thyratrons.[1]—The breakdown time of thyratrons is so small that it need not ordinarily be taken into consideration in the design and operation of thyratron circuits. The investigations that have been made indicate that the ionization time is of the order of a few microseconds.

12-20. Grid Current Previous to Firing.—One very important factor in the application of thyratrons is the grid current that flows prior to firing. Grid current is particularly objectionable in tubes used in control circuits in which the grid current must flow through a very high resistance such as a phototube or a coupling resistance. The resulting voltage drop may prevent firing of the tube.[2] Grid emission may also cause grid-anode breakdown and thus result in loss of grid control.

Grid current results from five principal causes: electrons attracted to the grid from the cathode, positive ions attracted to the grid, electrons

[1] SNODDY, L. B., *Physics*, **4**, 366 (1933).

[2] FRENCH, H. W., *J. Franklin Inst.*, **221**, 83 (1936).

emitted by the grid, capacitance between the grid and other electrodes, and internal or external leakage between the grid and other electrodes. Electron flow from cathode to grid is obtained only when the grid is positive or insufficiently negative to counteract the initial velocities of electrons emitted by the cathode. When the grid is negative, the grid current results both from positive ions drawn to the grid and from electrons emitted by the grid as the result of emitting material deposited upon the grid. The closer the grid is to the cathode, the greater is the amount of emitting material that is condensed upon the grid, and the higher is the grid temperature during operation. The grid current resulting from grid emission therefore increases with decrease of spacing between the grid and the cathode. It also increases with the grid surface available for the depositing of emitting material and for subsequent electron emission. Because positive-ion bombardment tends to remove active material deposited upon the grid, grid emission is less with tubes designed to operate with a negative grid.[1] Current caused by capacitance between the grid and other electrodes increases with decrease of electrode separation.

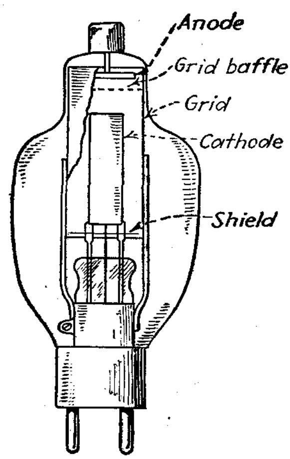


FIG. 12-24.—Electrode structure of the type FG-67 thyratron.

12-21. Electrode Structure and Characteristics of Thyratrons.—In most applications in which thyratrons are operated with direct anode voltage, or in which the frequency of the alternating voltage is high, it is necessary to use tubes that deionize as rapidly as possible. Figure 12-24 shows the construction of a typical tube of this type, the FG-67. Rapid deionization is achieved by the use of close spacing and small volume between the grid "baffle" and the anode and cathode, small distance between the back of the anode and the glass bulb, close spacing between the glass bulb and the upper part of the grid cylinder, and small holes in the grid baffle. The rated deionization time of the FG-67 thyratron is 100 μsec. The grid-control characteristics are shown in Fig. 12-25. Because of the high shielding by the grid and the closeness of the grid to the cathode, the grid current previous to breakdown is high, and the tube is not suitable for use in circuits in which the grid current must flow through a high resistance.

In three-electrode thyratrons designed for control circuits with high grid-circuit resistance the grid current is kept small by the use of adequate spacing between electrodes; by designing and placing the grid so as to

[1] MORACK, M. M., *Gen. Elec. Rev.*, **37**, 288 (1934).

reduce the depositing of emitting material evaporated from the cathode, as well as heating of the grid by the cathode; and by making the shielding

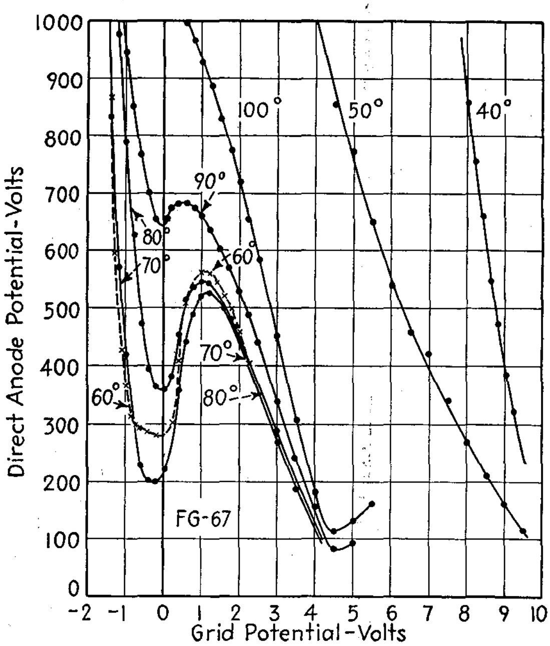


FIG. 12-25.—Grid-control characteristics of the type FG-67 thyratron.

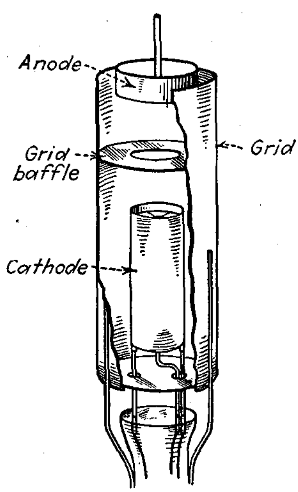


FIG. 12-26.—Electrode structure of the type FG-57 negative-grid thyratron.

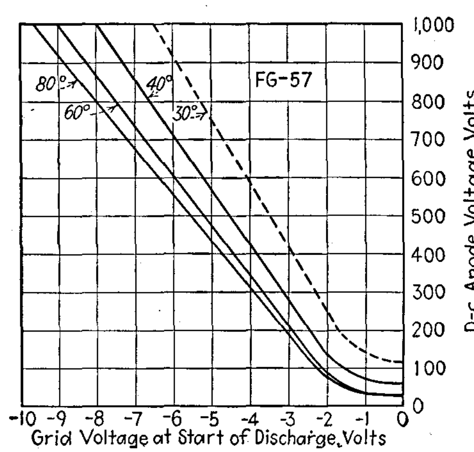


FIG. 12-27.—Grid-control characteristics of the type FG-57 thyratron.

action of the grid small enough so that a negative voltage must be used on the grid to prevent firing. Figure 12-26 shows the electrode structure of a typical heater-cathode thyratron triode of the negative-grid type, the FG-57. The grid-control characteristics of this tube are shown in Fig. 12-27. The rated deionization time is 1000 μsec.

In some control circuits it is advantageous or necessary to use a tube in which a positive grid voltage is required to start the arc. Positive grid control can be attained at the expense of increased grid current by making the shielding action of the grid so effective that the field from the anode cannot penetrate to the cathode. This is accomplished by the use of three baffles in the grid structure, in place of the single baffle used in nega-

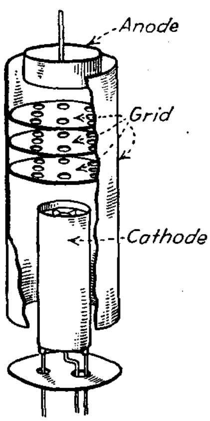


FIG. 12-28.—Electrode structure of the type FG-33 positive-grid thyratron.

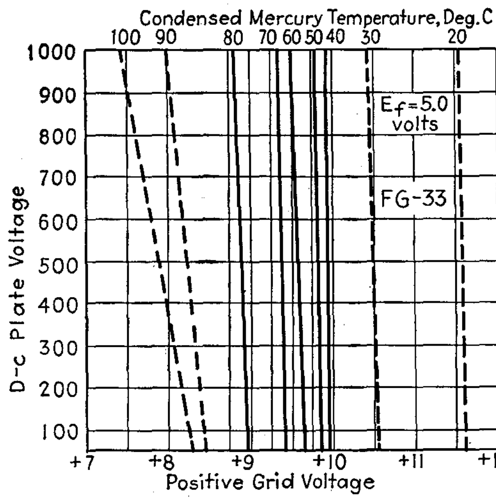


FIG. 12-29.—Grid-control characteristics of the type FG-33 thyratron.

tive-grid tubes. The holes in the grid baffles are also much smaller than the single hole used in the grids of negative-grid tubes. Figures 12-28 and 12-29 show the electrode structure and grid-control characteristics of the type FG-33 positive-grid thyratron, which has a rated deionization time of 1000 μsec.

Even with the most careful design, the grid current of three-electrode thyratrons previous to firing may be so great as to give difficulty when the grid circuit contains very high resistance. Furthermore, the changes in grid current as the tube warms up or ages may necessitate frequent readjustment of the circuit constants or voltages. These difficulties may be avoided by the addition of a fourth electrode, which acts as a shield.[1]

[1] LIVINGSTON, O. W., and MASER, H. T., *Electronics*, April, 1934, p. 114; JACOBI, W., and KNIEPKAMP, H., *E.T.Z.*, **58**, 1233 (1937).

The construction of a typical shield-grid thyratron, the FG-98, is shown in Fig. 12-30. In this tube the shield grid consists of a structure intermediate in form between the grids of negative- and positive-grid three-electrode thyratrons, two baffles being used. The control grid consists of a short cylinder whose diameter is somewhat larger than the holes in the grid baffles, and the grid is placed between the two baffles. It can be seen that the form and position of the control grid are such as to prevent the depositing of appreciable amounts of emitting material from the cathode or the absorption of much heat radiated by the cathode or the

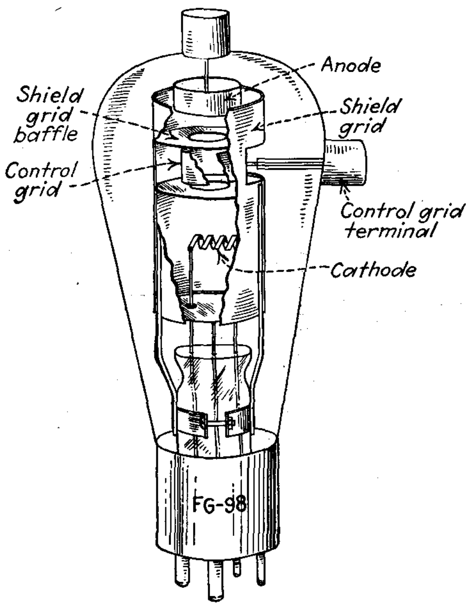


FIG. 12-30.—Electrode structure of the type FG-98 shield-grid thyratron.

anode. Moreover, this grid is out of the direct path of the arc, so that the grid current when the tube conducts is also reduced, and the grid receives little heat from the arc. The small size of the control grid minimizes the grid current resulting from electrons or ions drawn to the grid or reaching the grid from the cathode as the result of initial velocities. The small size of the grid and the grid lead, in conjunction with the shielding afforded by the shield grid, also results in low capacitance between the grid and the anode and cathode. The short and direct grid lead minimizes both the internal and the external leakage.

Another very important advantage of the shield-grid thyratron is that the tube can be made to have either positive or negative control characteristics by adjusting the voltage of the shield grid. This is indicated by the characteristics of the type FG-95 thyratron, shown in Fig. 12-31. The shield also makes it possible to adjust the tube characteristics in replacing tubes in circuits that are susceptible to changes in characteristics.

It is of interest to note that the design of the grids in the tubes illustrated in Figs. 12-24, 12-26, 12-28, and 12-30 is such as to shield the other electrodes from charges that collect on the inner surface of the glass envelope. If this were not so, the control action would be erratic.

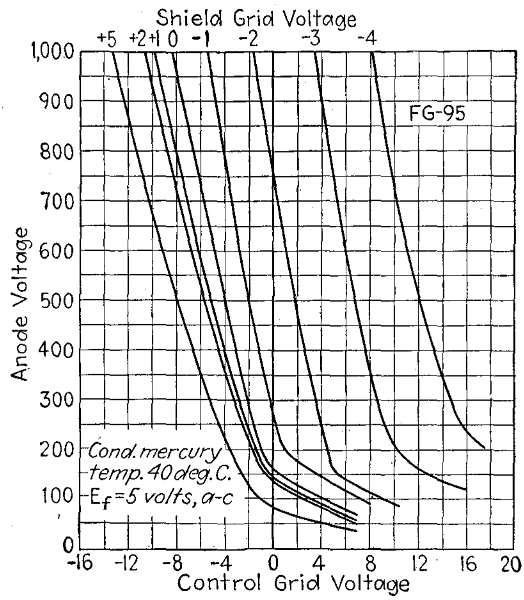


FIG. 12-31.—Grid-control characteristics of the type FG-95 shield-grid thyratron.

It can be seen from the grid-control characteristics of Figs. 12-25, 12-27, 12-29, and 12-31 that the thyratron, unlike the grid-glow tube, has a high grid-control ratio. A grid voltage of only a few volts suffices to prevent starting of the arc when the anode voltage is as high as 1000 volts.

12-22. Comparison of Thyratrons and High-vacuum Tubes.—Thyratrons differ from high-vacuum tubes in more respects than they resemble them. The similarity consists mainly of the dependence of current in both types of tubes upon thermionic emission and upon movement of electrons from cathode to anode. The differences between the two types, in addition to those involving the structure of the electrodes and envelopes, the presence of gas or vapor and of positive ions in thyratrons, and the emission of light from thyratrons, are shown in Table 12-I.

The high anode and emission efficiencies of hot-cathode arc tubes make possible the control of large currents and high power by means of relatively small tubes. The FG-57 thyratron, for instance, has a maximum average anode current rating of $2\frac{1}{2}$ amp and a peak anode voltage rating of 1000 volts and can, therefore, control 2500 watts of power. This tube is only slightly larger than the type 50 high-vacuum triode,

TABLE 12-I.—COMPARISON OF THYRATRONS AND HIGH-VACUUM TUBES

Thyratron	*High-vacuum*
Grid loses control after the tube fires.	Grid has complete control of anode current.
Anode and grid currents are determined by the applied voltage and the circuit resistance. Circuit resistance is essential.	Anode and grid currents are determined by the electrode voltages. No circuit resistance needed to limit the current.
Essentially a high-current, low-voltage tube. Because of low tube drop, anode-circuit efficiencies may exceed 98 per cent.	Essentially a low-current, high-voltage tube. Plate-circuit efficiency may approach 85 per cent in class C operation, but is usually less than 50 per cent.
Anode voltage is constant after the tube fires.	Anode voltage may vary.
Cathode emission efficiency may be high.	Cathode emission efficiency is relatively low.
Cathode heating time is high in types using high-efficiency heater-type cathodes. Tube heating time may also be high in mercury vapor tubes.	Except in large power tubes, heating time is short.
Cathode may be damaged if operated at low temperature or if the anode current exceeds the thermionic emission current for any other reason.	Cathode temperature may be reduced without damage to the tube. The cathode is not ordinarily damaged by currents exceeding rated values.
Characteristics of mercury vapor tubes are affected by tube temperature.	Characteristics are independent of tube temperature.
Grid current flows even when the grid is negative.	Negligible grid current flows when the grid is more than approximately $\frac{1}{2}$ volt negative.
Deionization time limits frequency of operation to several thousand cycles or less.	The frequency of operation may exceed 1000 Mc.
Arcback may occur.	Tube conducts in one direction only.

which has a maximum operating plate current of 55 ma and a power output that cannot exceed about 20 watts, even in class C operation. An important disadvantage of thyratrons over high-vacuum triodes and multi-grid tubes is the incompleteness of grid control, the current after firing being determined by the applied voltage and the circuit impedance.

Hot-cathode arc tubes have been built in sizes ranging from 10 watts to 100 kw (load power). Current ratings range up to 100-amp average current, and voltage ratings up to 1500 volts. There is a corresponding variation in tube structure and design, ranging from small glass-enclosed argon-filled tubes to large water-cooled metal-enclosed mercury vapor tubes. Operating data for typical thyratrons are listed on pages 686–687.

12-23. Special Precautions in the Use of Arc Tubes.—Before proceeding to a discussion of circuits for hot-cathode arc tubes it seems advisable to list special precautions that must be observed in the use of these tubes. They are as follows:

1. The cathode should be brought to normal operating temperature before anode voltage is applied. Failure to observe this precaution results in excessive cathode drop and in cathode disintegration.
2. Mercury vapor tubes should be heated for a sufficient time to allow the vapor pressure to assume its normal value. Greater heating time is required if the mercury has been scattered about the tube walls and electrodes than is otherwise needed.
3. Enough anode resistance must be used to limit the maximum instantaneous and average anode currents to their rated values. Failure to observe this precaution results in cathode disintegration and overheating of the anode.
4. Sufficient resistance must be used in the grid circuit to limit the maximum instantaneous and average grid currents to their rated values. Excessive grid current may cause damage because of heating of the grid or the grid-lead wires. If the grid resistance is omitted and an arc forms to the grid, the current may rise to such a high value that the grid wires may melt or the heating of the grid lead shatter the seal and allow air to enter the tube.
5. The use of fuses in series with the anode and grid voltage supplies is advisable.

12-24. Applications of Hot-cathode Arc-discharge Tubes.—Applications of hot-cathode arc tubes include light production, rectification, current and power control, oscillation (inversion), commutation, and amplification. Some of these applications will be discussed in the following sections.

Hot-cathode Arc Tube as a Light Source.—The use of arcs as light sources was one of the early practical applications of electricity. Arc-discharge tubes such as the mercury vapor and sodium vapor lamps are efficient sources of light, for they have the advantage of producing visible radiation in the manner of the glow tube and have a much smaller voltage drop.[1] Recently the fluorescent vapor lamp has taken its place among practical high-efficiency light sources. Except during starting, the filamentary cathodes in this type of tube are heated by positive-ion bombardment. The light emanates mainly from a fluorescent coating on the inner surface of the glass tube, a great variety of colors being obtained by the use of different fluorescent materials. For details of

[1] DUSHMAN, S., *Elec. Eng.*, **53**, 1283 (1934).

the design and operation of arc lamps the reader will find it instructive to refer to the technical literature on the subject.[1]

Thyratrons may be used as a stroboscopic light source. Usually, however, the production of light of sufficient intensity for the observation of rapidly moving objects requires such high values of peak current that cold-cathode tubes stand up better. A second advantage of cold-cathode arc tubes for this purpose is the more rapid deionization, which makes possible a light flash of extremely short duration. Stroboscopes will, therefore, be discussed in a later section dealing with cold-cathode arc tubes.

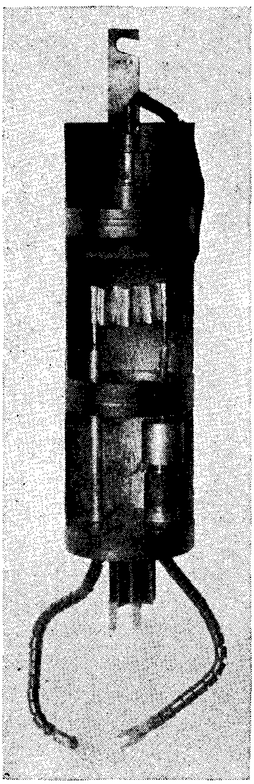


FIG. 12-32.—Cut-away view of the type FG-166 phanatron rectifier. (*Courtesy of General Electric Co.*)

12-25. Hot-cathode Arc Diode as a Rectifier. Like its high-vacuum counterpart, the hot-cathode arc tube is essentially a rectifier. At voltages for which the tube is designed, current flows in only one direction. There are, however, two essential differences, which have been noted before. The tube voltage drop is only from 10 to 20 volts, and the voltage is almost independent of the current passed by the tube. Hence the arc tube acts like a switch of zero resistance in series with a counter-e.m.f. of from 10 to 20 volts. Figure 12-32 shows the construction of the type FG-166 phanatron.

Rectifier circuits using hot-cathode arc tubes are the same as those using high-vacuum tubes, but provision must be made to ensure that the tube is brought to rated temperature before the anode voltage is applied. The advantages of arc tubes over high-vacuum tubes in rectification will be discussed in Sec. 14-4 (see also Sec. 12-22). Circuits and filters for rectifiers will be taken up in Chap. 14.

12-26. Arc Tube as a Control Device. D-c Operation.—Since the tube drop in an arc-discharge tube is practically independent of anode current and the anode current is unaffected by grid voltage, the tube is equivalent, so far as its action in a d-c circuit is concerned, to a switch in series with a counter-e.m.f. equal to the tube drop E_a. It is convenient in the analysis of some d-c thyratron circuits to use an equivalent circuit

[1] MAILEY, R. D., *Elec. Eng.*, **53,** 1446 (1934); BUTTOLPH, L. J., *Elec. Eng.*, **55,** 1174 (1936); MCKENNA, A. B., *Elec. J.*, **33,** 439 (1936); HAWKINS, L. A., *Trans. Am. Illuminating Eng. Soc.*, **32,** 95 (1937).

in which the tube is replaced by a switch and a battery of zero resistance and of such polarity as to oppose the flow of anode current. In most applications of thyratrons to the control of direct current and power it is essential not only to control the firing of the tube but also to stop the anode current at a subsequent time. Since the grid usually has little or no effect upon the anode current after the tube fires, it is necessary to find some other means of stopping the current. The simplest method is, of course, to open the circuit, but this is usually not feasible, especially when the anode current is large. Furthermore, convenience and flexibility of control demand that the current should be interrupted by electrical, rather than mechanical, means. Four circuits, called the *parallel*, *series*, *relaxation*, and *counter-e.m.f.* circuits, have been devised for this purpose.

12-27. Parallel Control.—The basic *parallel* circuit for the control of direct current is shown in Fig. 12-33.[1] Closing switch S_1 reduces the negative bias on the grid (or applies a positive bias) and causes the tube to fire. The grid no longer has control, and the switch S_1 may be opened without affecting the anode current. The potential drop across the load R_L charges the condenser C through R_s to a voltage equal to the supply voltage less the tube drop. The polarity is such that terminal b of the condenser is positive relative to terminal a. If the switch S_2 is closed, the positive terminal b of the condenser is connected to the cathode, and, since the negative terminal is already connected to the anode, the anode is made negative with respect to the cathode. The arc current stops, but the voltage E_{dc}, added to that across C, causes current to continue flowing through R_L. This current discharges C and charges it in opposite polarity. As the condenser charges, the anode voltage rises. If the grid is negative and the tube deionizes so rapidly that the anode voltage never becomes equal to the tube reignition voltage (see Sec. 11-22), the arc stays out and the grid regains control. This is usually expressed somewhat less precisely by stating that the grid regains control if the deionization time is shorter than the time required for the anode voltage to reach the value corresponding to the normal tube drop. The rate of rise of anode voltage decreases with increase of condenser capacitance and of load resistance. The rapidity of deionization, on the other hand, decreases with increase of load current. The capacitance necessary to ensure that the grid shall regain control therefore increases with load current. The purpose of the resistance R_s is to prevent a short circuit of the battery when S_2 is closed.

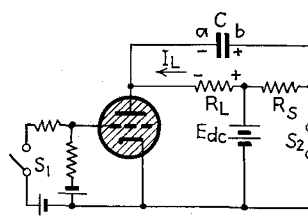


FIG. 12-33.—Basic parallel d-c control circuit.

[1] HULL, A. W., *Gen. Elec. Rev.*, **32**, 390 (1929).

A modification of the circuit of Fig. 12-33 that is occasionally useful is obtained by replacing S_2 with a glow-discharge tube.[1] The voltage drop through R_s falls with the charging current of the condenser, and the voltage applied to the glow tube goes up. The glow tube fires when its voltage becomes equal to its ignition voltage. Firing of the glow tube is equivalent to closing S_2; but, because of the voltage drop through the glow tube, the negative voltage applied to the anode of the thyratron is less than when the switch is used. For this reason the modified circuit requires higher supply voltage and greater capacitance than the basic circuit. The applied voltage must be at least as great as the ignition

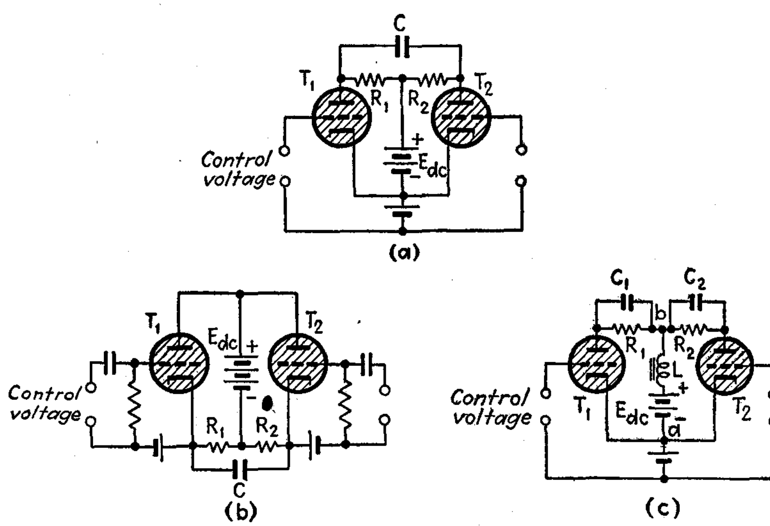


FIG. 12-34.—Three forms of the parallel d-c control circuit.

voltage of the glow tube, and the difference between the ignition and extinction voltages of the glow tube should preferably be large. R_s must be high enough so that the glow tube does not conduct steadily, but oscillates. The anode current is extinguished the first time the glow tube fires after the thyratron grid is made negative.

A much more useful circuit is obtained by replacing S_2 in Fig. 12-33 by a second thyratron, as shown in Fig. 12-34*a*.[2] When either tube is conducting, the condenser charges to such polarity that the terminal connected to the anode of that tube is negative with respect to the other terminal. Firing of the other tube by a positive grid impulse applies a negative voltage to the anode of the first tube, causing it to be extinguished. Similar action is obtained if the resistors are adjacent to the

[1] REICH, H. J., *Electronics*, December, 1931, p. 240.
[2] HULL, A. W., *Gen. Elec. Rev.*, **32**, 399 (1929).

cathodes and the condenser is connected between the cathodes, as in Fig. 12-34*b*.

A somewhat different form of parallel circuit is that of Fig. 12-34*c*.[1] The action of this circuit is as follows: Assume that tube 1 is conducting. The condenser C_1 charges to a voltage equal to the applied voltage less the tube drop, the condenser terminal adjacent to the anode being negative. Prior to the firing of tube 2 there is no voltage across C_2. Since C_2 cannot charge instantaneously, the firing of tube 2 reduces the voltage V_{ab} between a and b to the tube drop of tube 2, or approximately 15 volts. (The remainder of the line voltage is taken up by the inductance L.) The anode voltage of tube 1 is therefore made negative by an amount equal to the difference between the voltage of C_1 and the voltage V_{ab}, or to the supply voltage less twice the tube voltage drop. If the deionizing time is less than the time taken for the condenser C_1 to discharge through R_1, and for C_2 to charge to such voltages that the anode voltage of tube 1 exceeds the ionization potential, the grid of tube 1 will regain control. The resistances and condensers may also be adjacent to the cathodes, instead of to the anodes (see Fig. 12-68).

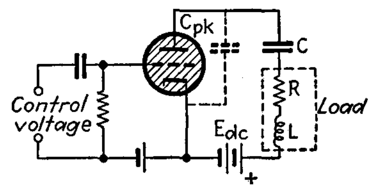


FIG. 12-35.—Basic series control circuit.

The firing impulses may be applied to the grids directly, as in Fig. 12-34, through one or more transformers, or through resistance-condenser networks. Examples of these circuits will be shown. The circuits of Fig. 12-34 are the basis of the parallel inverter, several high-speed counting circuits, and other interesting and useful devices, which will be discussed in later sections.

12-28. Series Control.—In the *series* type of control circuit, a condenser (or a condenser and an inductance) is connected in series with the load and the anode, as shown in Fig. 12-35.[2] Before the tube fires, the anode circuit is made up of L, C, R, and C_{pk}, the anode-to-cathode capacitance of the tube. Because C_{pk} is very much smaller than C, practically the entire applied voltage appears across C_{pk}. A control impulse applied to the grid will cause the tube to fire. Current can flow, however, only while C charges. Since the tube passes current in only one direction, anode current ceases as soon as the condenser is fully charged. The tube cannot be fired again until the charged condenser has been discharged. The presence of inductance in series with the condenser causes the condenser to charge to a voltage which, if the load resistance R is small, is nearly equal to twice the applied voltage

[1] HERSKIND, C. C., *Elec. Eng.*, **53**, 926 (1934); **56**, 1372 (1937).
[2] HULL, *op. cit.*, p. 390.

minus the tube voltage drop.[1] If the load resistance is high, on the other hand, the condenser charges to approximately the applied voltage minus the tube drop. The condenser may be discharged through a second thyratron, as in Fig. 12-36.

If the second thyratron of Fig. 12-36 should fire before current has ceased flowing in the first one, a short circuit would result, and the tubes would be subjected to destructive currents. The likelihood of this difficulty is reduced by the use of a center-tapped inductance, as in Fig. 12-37. When tube 2 fires, current starts flowing from *c* to *b*. A voltage is induced between *c* and *b* which is of such polarity as to tend to prevent this flow of current. Since the two halves of the transformer are wound in the same direction, an equal voltage is induced between *a* and *b*, *a* being negative relative to *b*. A voltage of opposite polarity

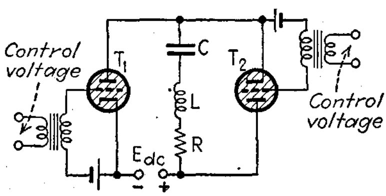


FIG. 12-36.—Series control circuit.

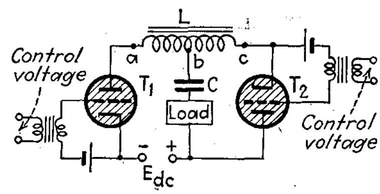


FIG. 12-37.—Series control circuit with center-tapped commutating inductance.

is also induced between *a* and *b* by the decay of current through tube 1, but this becomes smaller as the condenser approaches maximum voltage. Therefore, if tube 2 fires when the condenser is nearly charged to maximum voltage, both the net induced voltage between *a* and *b* and the condenser voltage are in the direction to make the anode of tube 1 negative and extinguish this tube.

The circuits of Figs. 12-35 to 12-37 differ from those of Figs. 12-33 and 12-34 in that current flows in the former only during the charging and discharging of the condensers. They are the basis of the series inverter, counting circuits, and other devices.

12-29. Relaxation Control.—The third, or *relaxation*, method of interrupting the anode current electrically makes use of a series combination of inductance and capacitance shunted across the tube, as in Fig. 12-38.[2] The action of the circuit is most readily explained with the aid of equivalent circuits of the type discussed in Sec. 12-26. Before the

[1] PIERCE, G. W., "Electric Oscillations and Electric Waves," Chap. II, McGraw-Hill Book Company, Inc., New York, 1920; KURTZ, E. B., and CORCORAN, G. F., "Electric Transients," pp. 42–44, John Wiley & Sons, Inc., New York, 1935; SKILLING, H. H., "Transient Electric Currents," pp. 93–98, McGraw-Hill Book Company, Inc., New York, 1937.

[2] REICH, H. J., *Rev. Sci. Instruments*, **4**, 147 (1933); *Elec. Eng.*, **52**, 817 (1933).

tube fires, the equivalent circuit is that of Fig. 12-39*a*. The condenser charges through the load and the inductance L_1. If the ratio of inductance to resistance in the circuit is high and the condenser is initially uncharged, the condenser may charge nearly sinusoidally to a maximum voltage only slightly less than twice the applied direct voltage.[1] If the ratio of inductance to resistance is small, on the other hand, the condenser charges exponentially toward a value equal to the applied voltage. Whether or not the condenser actually attains maximum voltage depends upon the firing voltage of the tube.

The equivalent circuit during the time in which the tube conducts is that of Fig. 12-39*b*. Because the tube voltage drop is practically constant in the operating range of anode current, the load current I_L and the condenser current I_1 are independent of each other. At the

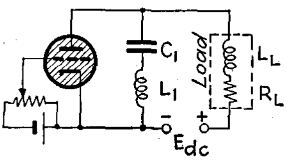


FIG. 12-38.—Relaxation control circuit.

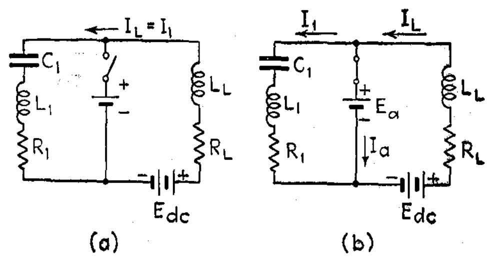


FIG. 12-39.—Equivalent circuits for the relaxation control circuit of Fig. 12-37; (*a*) before the tube fires; (*b*) while the tube is conducting.

instant of firing, I_L and I_1 are equal and may be either positive, zero, or negative, depending upon whether the tube fires before the condenser voltage reaches its maximum value, at the instant of maximum condenser voltage, or after the time of maximum condenser voltage. After the tube fires, the load current increases exponentially toward the value $(E_{dc} - E_a)/R_L$ where E_{dc} is the applied direct voltage, E_a is the tube voltage drop, and R_L is the d-c resistance of the load. The wave of condenser current in the equivalent circuit is a damped sine wave. The condenser current does not stop when the condenser voltage has fallen to zero but, because of voltage induced in the inductance L_1, continues to flow until the condenser is charged in the opposite direction to a voltage which is slightly less than $E_o - E_a$, where E_o is the condenser voltage at the instant of firing. The condenser current then reverses and charges the condenser in the initially positive direction. The curves of load current, condenser current, and condenser voltage that would be obtained if the tube were in every respect equivalent to a

[1] See footnote [1], p. 480.

switch and counter-e.m.f. and could pass current in the reverse direction, are plotted in Fig. 12-40.[1]

Actually, at some time t_e shortly after the condenser current has reversed, the reverse condenser current becomes equal to the load current. At this instant the anode current is zero, and the tube goes out. As soon as anode current stops flowing, the anode voltage becomes nearly equal to the condenser voltage,[2] which is negative at the instant of extinction. After extinction, the condenser begins to discharge through the d-c supply and subsequently charges to a positive voltage, in a manner determined by the constants of the circuit of Fig. 12-39*a*. If the time taken for the condenser voltage to reach a positive value equal to the ionization potential exceeds the deionization time of the tube, the grid will regain control.

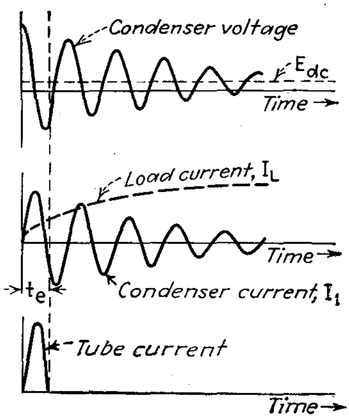


FIG. 12-40.—Curves of condenser current, load current, tube current, and condenser voltage in the circuit of Fig. 12-39*b*.

In order that it shall be possible for the negative condenser current to become equal to the supply current and thus make the net anode current zero, the rate at which the condenser current increases must be greater than that at which the supply current increases. This necessitates high C_1/L_1 and L_1/R_1 ratios and high load inductance or resistance. The allowable C_1/L_1 ratio is limited, however, by the rated maximum instantaneous anode current. For a high L_1/R_1 ratio the peak anode current is approximately equal to $V_i\sqrt{C_1/L_1}$, where V_i is the ignition voltage.[3] A high L_1/R_1 ratio may be obtained by the use of an air-core coil wound with heavy wire and having optimum inductance shape.[4] Suitable values of L_1 and R_1 are 1.5 mh and 0.15 ohm, respectively.

Because the condenser voltage is negative at the instant of extinction, the maximum positive voltage to which it can charge is greater than it was during the first charging from zero initial voltage. If the grid bias is insufficient to prevent the tube from firing at the maximum voltage to which the condenser can charge from zero initial voltage, relaxation oscillations will occur, the condenser periodically charging from the supply and discharging to a negative voltage through the inductance L_1. The load

[1] See references in footnote [1], p. 480.

[2] The voltage across L_1 is small, since L_1 is small.

[3] PIERCE, *op. cit.*, p. 17; KURTZ and CORCORAN, *op. cit.*, p. 66; SKILLING, *op. cit.*, p. 121.

[4] BROOKS, M., and TURNER, H. M., *Univ. Ill. Eng. Expt. Sta. Bull.* 53, 1912.

current and the wave form and frequency of the condenser voltage depend upon the circuit constants and upon the firing voltage, which in turn depends upon the grid bias. The flow of current through the load can be interrupted by increasing the bias to a value sufficient to prevent the tube from firing. Because thyratrons that have short deionization time oscillate even when the supply current is large, this makes a useful way of obtaining complete grid control of anode current in d-c operation when the periodic variation of load current is not objectionable. The

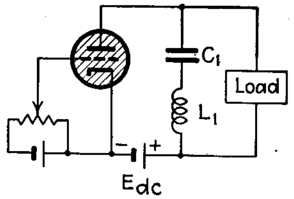


FIG. 12-41.—Variant of the relaxation control circuit of Fig. 12-38.

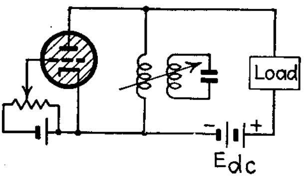


FIG. 12-42.—Variant of the relaxation control circuit of Fig. 12-38.

circuit of Fig. 12-38 is the basis of the relaxation inverter, sweep circuits for cathode-ray oscillograph and television tubes, underload d-c circuit breakers, audio oscillators, and other devices.

A variant of the circuit of Fig. 12-38 is shown in Fig. 12-41.[1] In this circuit the condenser charges through the tube and discharges through the load. The action of the inductance causes the condenser to charge to nearly twice line voltage, after which it starts discharging. When the resulting reverse current becomes equal to the supply current through the load, the net anode current is zero, and the tube goes out. Another modification of the circuit of Fig. 12-38 is that of Fig. 12-42, in which the condenser is shunted across the load through a transformer.[2] It may be shown from transient theory that the coefficient of coupling between the primary and secondary of the transformer must exceed 70 per cent.

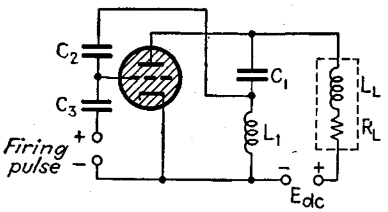


FIG. 12-43.—Relaxation circuit in which the application of a firing pulse causes a single cycle of condenser charging and discharging.

In Fig. 12-43 is shown a form of the circuit of Fig. 12-38 in which the application of a voltage pulse to the grid circuit causes a single cycle of condenser discharge through the tube and recharge through the load. The voltage induced in L_1 during the discharge of C_1 is applied to the grid through the condenser C_2. Toward the end of the discharge the rate of change of current is negative and so the induced voltage is of such polarity

[1] LIVINGSTON, O. W., and LORD, H. W., *Electronics*, April, 1933, p. 96.

[2] REICH, *loc. cit.*

as to make the grid positive and thus cause the flow of electrons to the grid. After extinction these electrons are trapped on the grid and the condensers C_2 and C_3, making the grid voltage highly negative. A positive pulse of voltage applied to the input reduces the negative grid voltage for an instant and allows the tube to fire. If desired, the condenser C_3 may be omitted and the tube fired by discharging C_2 in any manner or by applying an impulse to the grid through a transformer in series with C_2. This circuit is useful in the production of voltage and current surges of various shapes.[1] The wave form is dependent upon the portion of the circuit from which the voltage or current is taken and upon the circuit parameters.

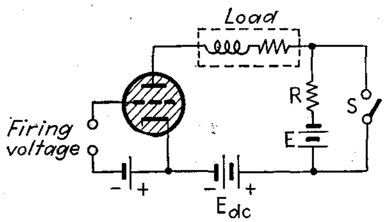


FIG. 12-44.—Basic counter-e.m.f. control circuit.

12-30. Counter-e.m.f. Control.—The anode current of an arc tube may be stopped by the insertion into the anode circuit of an e.m.f. of proper polarity to oppose the flow of anode current. If this counter-e.m.f. is greater than the supply voltage minus the tube drop and is maintained long enough for tube deionization, the anode current will be stopped and grid control reestablished. Thus, if the switch S of Fig. 12-44 is opened, the battery E is added to the anode circuit, and the arc may be extinguished. The resistance R prevents the counter-e.m.f. battery from being short-circuited when S is closed. Although the basic

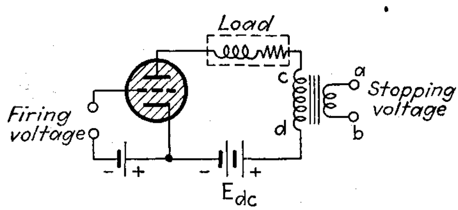


FIG. 12-45.—Practical form of the counter-e.m.f. control circuit.

circuit of Fig. 12-44 has no great practical value, the counter-e.m.f. principle is used in a number of applications of thyratrons discussed in later sections.

By use of the circuit of Fig. 12-45, the counter-e.m.f. may be applied in the form of a pulse. A voltage impulse applied to the primary terminals of the transformer, of such polarity as to make the terminal c negative relative to the terminal d, may be used to stop the anode current.[2] Evidently the applied impulse must be of sufficient duration to

[1] REICH, H. J., *Elec. Eng.*, **55**, 1314 (1936).
[2] WILLIS, C. H., *Gen. Elec. Rev.*, **35**, 632 (1932).

allow the tube to deionize. The impulse may result from the discharge of a condenser through the primary of the transformer, from the application of a battery voltage to the primary, or from the operation of an associated thyratron connected to the primary.

The counter-e.m.f. method of stopping anode currents can be applied to the commutation of the arc current from one tube to another. Assume that tube 1 of Fig. 12-46 is conducting and that the grid voltage of tube 1 is such as to allow it to regain control. An impulse of proper polarity applied to the terminals *a-b* of the transformer will cause tube 1 to extinguish and tube 2 to fire if the grid voltage of the latter is simultaneously brought to the firing point. It is not necessary that the counter-e.m.f. be equal to the supply voltage minus the tube drop. The direction of the counter-e.m.f. makes the net positive voltage on the

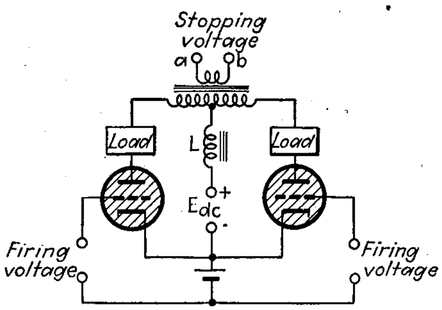


FIG. 12-46.—Circuit in which counter-e.m.f. extinction is used in commutating current between two thyratrons.

anode of tube 2 greater than that of tube 1. Since the tubes may be considered to have negligible resistance when conducting and since the inductance in the d-c supply tends to hold the supply current constant, the anode current will be diverted from tube 1 to tube 2.[1] A succeeding impulse, of opposite polarity, applied to terminals *a-b*, will fire tube 1 and extinguish tube 2, provided that the grid voltages are simultaneously readjusted.

The counter-e.m.f. type of control is the basis of the counter-e.m.f. inverter and the d-c thyratron motor, both of which will be discussed.

12-31. Choice of Tubes for D-c Control Circuits.—From the theory of operation of the parallel and relaxation types of control circuits it follows that the time available for deionization of the tubes is usually small in these circuits. This is also true in some applications of the series and counter-e.m.f. circuits. In general, therefore, tubes used in these circuits should have short deionization time. This is especially important when the circuits are used as the basis of oscillators. The

[1] DALLENBACH, W., and GERECKE, E., *Arch. Elektrotech.*, **14**, 171 (1924).

FG-67 and FG-43 thyratrons and other types having equal or shorter deionizing time are suitable for this type of service.

12-32. Applications of Basic D-c Control Circuits.—The modifications and applications of the basic d-c thyratron control circuits are too numerous to make it possible to include them all in this text.[1] Because of their frequent use and because they emphasize fundamental principles, two thyratron applications that are based upon d-c control circuits will be discussed in detail. These are switching and counting circuits and oscillators.

Parallel Switching Circuits. High-speed Counting Circuits.—The parallel type of d-c control circuit is essentially a trigger circuit, and as such may be used as an electronic switch. In the two-tube circuits of

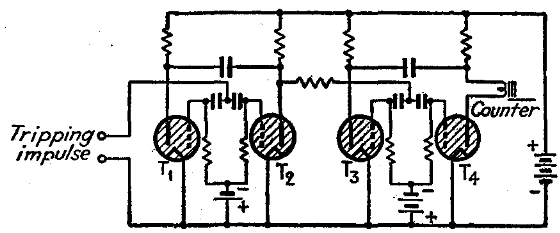


FIG. 12-47.—Scale-of-eight type of counting circuit.

Fig. 12-34 the current of either tube may be used to operate electrical apparatus, or the voltage drop through the anode (or cathode) load resistors may be applied to the grids of other tubes. In conjunction with two high-vacuum amplifier tubes, which are alternately overbiased by the voltage drop through the anode resistors of the switching circuit, the thyratron switch has proved useful in the simultaneous observation of two voltages with a single cathode-ray oscillograph. This application will be discussed in further detail in Sec. 15-22. The parallel switching circuit is also the basis of several high-speed counting circuits.

The frequency at which a mechanical counter can respond is limited by the inertia of moving parts. For high-speed counting it is necessary to design circuits in which the relay operates once in a relatively large number of impulses. If a mechanical counter that is capable of operating at a maximum frequency of 20 per second is called upon to count only each tenth impulse, for instance, the device will be capable of counting a maximum of 200 impulses per second. Two interesting electron-tube counting circuits developed by Wynn-Williams[2] are based upon the parallel thyratron switch. A typical example of one of these, which is

[1] See, for instance, K. HENNEY, "Electron Tubes in Industry," 2d ed., McGraw-Hill Book Company, Inc., New York, 1937; F. H. GULLIKSEN, and E. H. VEDDER, "Industrial Electronics," John Wiley & Sons, Inc., New York, 1935.

[2] WYNN-WILLIAMS, C. E., *Proc. Roy. Soc.* (*London*), **136**, 312 (1932).

usually called the *scale-of-eight counter*, is shown in Fig. 12-47. It consists of two or more parallel circuits in tandem. The impulses to be counted are applied simultaneously to the grids of the first parallel circuit, either through a transformer or through a combination of resistance and capacitance. The anode voltage of one tube of this first parallel circuit is applied to the grids of the following parallel circuit, which has a counting relay in series with one anode. The grids are biased sufficiently to prevent firing without the application of an impulse to the input. Because of slight differences between tubes or because one anode circuit is temporarily opened, the first impulse will cause one of the tubes in the first stage, say tube 2, to fire. The next impulse fires tube 1 and extinguishes tube 2. Extinguishing tube 2 reduces the negative grid voltage of tubes 3 and 4, causing one of them, say tube 3, to fire. The third impulse fires tube 2 and extinguishes tube 1. The fourth impulse fires tubes 1 and 4, closing

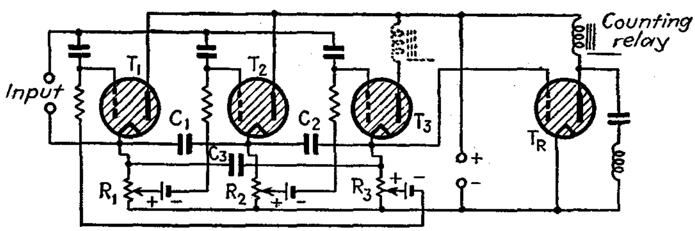


FIG. 12-48.—Ring type of counting circuit.

the relay, and extinguishes tubes 2 and 3. The addition of another stage allows the relay to count every eighth impulse. With n stages of two tubes each, the relay need operate only each 2^nth impulse.

The second type of counter developed by Wynn-Williams, known as the *ring circuit*, is shown in Fig. 12-48.[1] It consists of three or more grid-controlled arc-discharge tubes connected in a manner similar to that of the parallel d-c control circuit of Fig. 12-34*b*. The impulses to be counted are applied simultaneously to all the grids. Suppose that one tube, say tube 1, has fired. The IR drop resulting from the flow of current through the cathode resistor R_1 accomplishes two results. It reduces the negative grid bias of tube 2, so that the next impulse applied to the grids will cause tube 2 to fire; and it charges the condenser C_1, so that, when tube 2 fires, the anode of tube 1 is made negative relative to the cathode, causing tube 1 to extinguish. In a similar manner, the voltage drop in R_2, caused by the anode current of tube 2, "primes" the grid of tube 3 and charges condenser C_2 so that the next impulse fires tube 3 and extinguishes tube 2. Firing of tube 3 "primes" tube 1 and

[1] WYNN-WILLIAMS, C. E., *Proc. Roy. Soc. (London)*, **132**, 295 (1931); HULL, A. W., *Physics*, **2**, 409 (1932); SHUMARD, C. C., *Elec. Eng.*, **57**, 209 (1938).

charges C_3 so that the third impulse will again fire tube 1 and extinguish tube 3. As many tubes as desired may be connected in this manner. The counting relay is operated either directly by the anode current of one tube or by an auxiliary high-vacuum or arc tube T_r, which is controlled by the voltage drop in one of the cathode resistors of the ring circuit. The relaxation type of d-c control circuit makes a convenient means of interrupting the anode current of the auxiliary tube if an arc tube is used to operate the counting relay. This is shown in Fig. 12-48. When four or more tubes are used, the scale-of-eight circuit gives a higher counting ratio than the ring circuit and requires fewer voltage supplies. Counters based upon the series and relaxation d-c control circuits have also been devised.[1]

12-33. The Thyratron as an Oscillator.—Thyratron oscillators may be divided into two types: those designed primarily to furnish only voltage output, usually of saw-tooth or rectangular wave form, and those designed to convert large amounts of d-c power into a-c power. Arc-tube oscillators capable of converting appreciable d-c power into a-c power at commercial frequencies are called *inverters*.

Saw-tooth-wave Generator.—As generators of saw-tooth voltage waves, thyratrons have a number of advantages over glow tubes. Dependence of ignition voltage upon grid voltage makes possible simple control of amplitude. Because of the amplifying property of the grid, smaller voltage is required to synchronize the oscillator to a control frequency, and there is less likelihood of distortion of the voltage wave than when the control voltage is introduced in series with a glow tube. The lower tube drop allows the use of smaller supply voltage and, in conjunction with the use of negative grid bias to obtain high ignition voltage, results in large amplitude of oscillation. The high instantaneous current that the tube will pass results in shorter discharge time than with glow tubes and makes possible the increase of amplitude of oscillation by the use of inductance in series with the condenser. Shorter deionizing time allows oscillation to take place at much higher frequency. Finally, the presence of an ample supply of electrons at the cathode makes the ignition potential independent of cathode illumination and of cosmic radiation and other random ionizing agents. The frequency is therefore much more constant than with glow tubes. The dependence of firing voltage of mercury vapor tubes upon tube temperature makes it advisable to use gas-filled tubes when frequency stability is important.

The circuit of Fig. 12-38 oscillates when the grid voltage is insufficiently negative (or too positive) to prevent the tube from firing at the maximum voltage to which the condenser charges. If the load consists of resistance only, the condenser charges exponentially, and the condenser

[1] LORD, H. W., and LIVINGSTON, O. W., *Electronics*, January, 1934, p. 7.

voltage is of the form shown in Fig. 12-49. Because of the action of the inductance L_1, the condenser discharge is nearly sinusoidal, and the condenser polarity reverses during discharge. The total change in condenser voltage is approximately equal to twice the ignition voltage minus the tube drop, E_a. The inductance should be large enough to limit the discharge current to the rated maximum instantaneous anode current. By the method used to derive Eq. (12-2), the frequency of oscillation may be shown to be approximately

$$f = \frac{1}{R_L C_1 \log_\epsilon \dfrac{E_{dc} + V_i - E_a}{E_{dc} - V_i}} \tag{12-6}$$

The circuit also oscillates when L_1 is replaced by a resistance R_1, the behavior being similar to that of the glow-tube oscillator of Fig. 12-9. The change in condenser voltage is equal to the difference between the ignition and extinction voltages, and the expression for the frequency of oscillation is of the same form as Eq. (12-2), R representing the sum of the resistances R_1 and R_L.

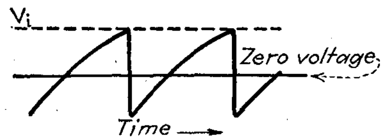


FIG. 12-49.—Wave form of the condenser voltage in the circuit of Fig. 12-38 used as a relaxation oscillator.

If the resistance R_L is replaced by a voltage pentode operated on the flat portion of the plate characteristic, the charging current is nearly constant, and the condenser charges practically linearly, as in the glow-tube oscillator of Fig. 12-12. The type 884 argon-filled tube is commonly used in this circuit. The extinction voltage of this tube is only about 15 volts, and the firing voltage may be as high as 300 volts (with -30 volts on the grid). It is evident that the amplitude of the condenser voltage is much greater than in the glow-tube oscillator. When the inductance L_1 is used, the condenser voltage may vary between the maximum limits of approximately minus 300 and plus 300 volts. The frequency of oscillation is approximately

$$f = \frac{I}{(2V_i - E_a)C} \tag{12-7}$$

where I is the constant charging current, V_i is the ignition voltage, E_a is the tube drop, and C is the total capacitance, including that of the tube and wires. This assumes that the inductance L_1 has negligible resistance and that the condenser discharge is instantaneous. When L_1 is replaced by a current-limiting resistance R_1, the frequency is given approximately by Eq. (12-5).

Since V_i and V_e are independent of frequency in arc tubes, Eqs. (12-2) and (12-5) to (12-7) may be used in predetermining the frequency of oscillation or in designing an oscillator for a desired frequency range. As in glow-tube oscillators, very low frequency of oscillation may be attained by the use of large condensers. Because of the greater difference between ignition and extinction voltages, the condenser capacitance required for a given frequency is lower in the arc-tube oscillator than in the glow-tube oscillator. The maximum frequency of an arc-tube oscillator is determined in part by the deionization time. When the period of oscillation approaches the deionization time, the tube fails to extinguish. Frequencies of the order of 50 to 100 kc per second may be attained.

In order to protect the tube against damage as the result of excessive peak current, resistance or inductance must be used in series with the

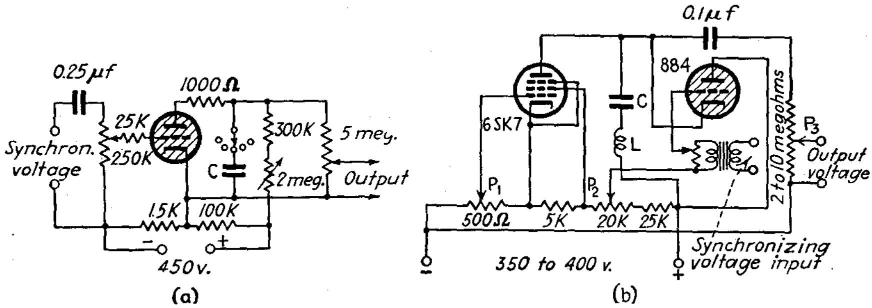


FIG. 12-50.—Thyratron relaxation oscillator circuits in which the condenser charging current is controlled (*a*) by a high resistance and (*b*) by a pentode.

tube. For a maximum condenser voltage of 300 volts, the type 884 tube requires a current-limiting resistance of 1000 ohms, or an inductance equal in henrys to the condenser capacitance in microfarads.[1]

An arc-tube saw-tooth-wave generator may be readily stabilized at a frequency equal to, or nearly equal to, the natural frequency of oscillation or a multiple thereof by introducing the control voltage into the grid circuit.[2] The control voltage varies the firing voltage at control frequency. The control action may be explained by diagrams similar to those of Fig. 10-33, the control-voltage ripple being superimposed upon the firing-voltage line V_i of Fig. 12-49, instead of upon the saw-tooth voltage as in Fig. 10-33. Although the oscillator can also be stabilized by a control voltage having a frequency that is a submultiple

[1] When a condenser of capacitance C, initially charged to a voltage E_0, is discharged through an inductance L, and the ratio of inductance to resistance is high, the peak discharge current is approximately equal to $E_0 \sqrt{C/L}$. Thus, if the inductance in henrys is equal to the capacitance in microfarads, the peak current is approximately $E_0/1000$ amp, which is the same as would be obtained if the condenser were discharged through a 1000-ohm resistance.

[2] See, for instance, W. R. MACLEAN, *Communications*, March, 1943, p. 23.

of the oscillation frequency, the condenser voltage is then amplitude-modulated at control frequency. The wave of control voltage should preferably have a steep wave front for accurate stabilization.

Figure 12-50*a* shows a typical circuit in which the condenser charging current is controlled by a resistance, and Fig. 12-50*b* one in which a pentode is used to control the charging current. P_1 in circuit *b* adjusts the charging current by varying the control-grid voltage of the pentode and thus controls the frequency. P_2 adjusts the 884 bias and thus the ignition voltage. Since the ignition voltage affects both the amplitude and the frequency of oscillation, P_2 changes the amplitude and frequency simultaneously. P_3 varies the amplitude of the output voltage. The principal application of this and similar circuits is in providing a linear time axis for cathode-ray oscillographs and in television "scanning."[1] Modifications of the circuit of Fig. 12-50 will be discussed in Sec. 15-20.

An amplifier tube, the plate circuit of which contains a relay, may be directly coupled to a thyratron relaxation oscillator in such a manner that the amplifier current rises and falls with the condenser voltage. The periodic contactor thus formed is more reliable in operation than the glow-tube contactor mentioned in Sec. 12-7. For some purposes the relay may be replaced by an 884 thyratron (gas-discharge tube), use being made of the ability of the grid to interrupt small anode currents. Such a circuit is shown in Fig. 15-61.

12-34. Thyratron Inverters.—Certain problems that arise in the distribution of load in a-c systems and in the interconnection of a-c systems of different frequencies or of a-c and d-c systems, can be solved by the use of grid-controlled rectifiers[2] to convert from alternating current to direct current and of thyratron inverters to invert from direct current to alternating current.[3] The use of tube rectifiers, d-c ties, and thyratron inverters to interconnect a-c systems and to feed a-c power to substations would eliminate the difficulties arising from synchronization and power factor. Since the frequency of the inverter output is independent of the generator frequency, the machines at the generating and receiving ends of the line could each be operated at the most efficient frequency. Many

[1] See, for instance, E. H. FELIX, "Television—Its Methods and Uses," McGraw-Hill Book Company, Inc., New York, 1931; V. K. ZWORYKIN, and E. D. WILSON, "Photocells and Their Applications," John Wiley & Sons, Inc., New York, 1931; H. H. SHELDON and E. N. GRISEWOOD, "Television," D. Van Nostrand Company, Inc., New York, 1929; D. G. FINK, "Principles of Television Engineering," McGraw-Hill Book Company, Inc., New York, 1940; V. K. ZWORYKIN, "Television; the Electronics of Image Transmission," John Wiley & Sons, Inc., New York, 1940.

[2] See footnote [2], page 519.

[3] WILLIS, C. H., BEDFORD, B. D., and ELDER, F. R., *Elec. Eng.*, **54**, 102 (1935); BEDFORD, B. D., ELDER, F. R., and WILLIS, C. H., *Gen. Elec. Rev.*, **39**, 220 (1936); ALEXANDERSON, E. F. W., *Electronics*, June, 1936, p. 25.

problems involving hunting, transients, and other types of instability would also be solved in this manner. The use of direct current in place of alternating current in the transmission of power has a number of advantages, the most obvious of which is that direct current involves only voltage and polarity, whereas alternating current involves voltage, frequency, phase, phase sequence, and power factor. From an economic point of view the advantage of d-c transmission arises from the higher voltage that can be used without danger of corona. The maximum r-m-s alternating voltage that can be transmitted is less than the maximum direct voltage because the crest voltage determines the magnitude of corona effects. Other advantages of d-c transmission are elimination of dielectric losses and line reactance and reduction of insulation stresses at a given effective voltage. Disadvantages of d-c transmission are the greater likelihood of insulator flashover and resultant power arc, and the greater difficulty of interrupting an arc when it does occur.

Because direct voltage cannot be transformed, the possibility of using direct current in distribution and high-voltage transmission is contingent upon the development of reliable inverters for changing high-voltage direct current into high-voltage alternating current of satisfactory wave form. Four types of single-phase inverters have been developed, based upon the parallel, series, relaxation, and counter-e.m.f. control circuits.

Another field of application of inverters may be developed in the speed control of synchronous and induction motors. Since the frequency of the a-c output of the inverter is easily controlled by the grid excitation, speed control of motors connected to the inverter may thus be obtained.

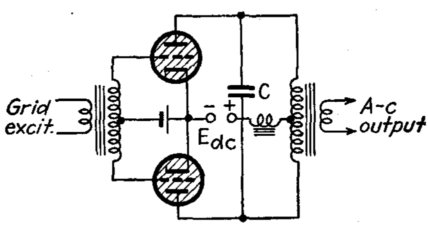


FIG. 12-51.—Separately excited parallel inverter.

12-35. The Parallel Inverter.—The parallel type of inverter is essentially the parallel d-c control circuit of Fig. 12-34a, with the anode resistors replaced by a center-tapped transformer.[1] In the simplest form of parallel inverter the grids are excited from an external source of voltage of desired frequency through a center-tapped grid transformer, as in Fig. 12-51. Each tube conducts during one-half of the cycle of grid excitation voltage. Commutation is obtained by use of the commutating condenser C, as explained in Sec. 12-27. The output voltage wave form is dependent upon the magnitude of the load current, the power factor of the load, the leakage inductance of the transformer and the inductance in series with the d-c supply, and upon the capacitance of the condenser. If the coupling of the transformer is close, as is usually

[1] PRINCE, D. C., *Gen. Elec. Rev.*, **31**, 347 (1928); HULL, *Gen. Elec. Rev.*, *loc. cit.*

true, the wave form of the output voltage is approximately that of the condenser voltage.

An examination of the operation of the parallel-inverter circuit shows that the wave form across the condenser is essentially that across the condenser of the equivalent circuit of Fig. 12-52 when the commutator is rotated at the frequency of the inverter grid excitation. Hence the inverter wave form may be determined by analyzing the transients set up by the switching of the equivalent circuit. Typical wave forms are shown in Fig. 12-53.[1] The wave form obtained in a particular circuit depends upon the values of commutating and load capacitance, commutating and load inductance, and circuit and load resistance. Waves of the forms of a and b are likely to be obtained with medium and heavy nonreactive loads. If the load is reactive or light, however, the condenser may charge nearly sinusoidally. If commutation then takes place at the instants when the condenser voltage is a maximum, the output voltage approaches sinusoidal form, as in Fig. 12-53c.

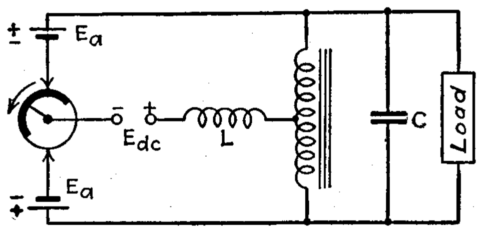


FIG. 12-52.—Equivalent circuit of the parallel inverter.

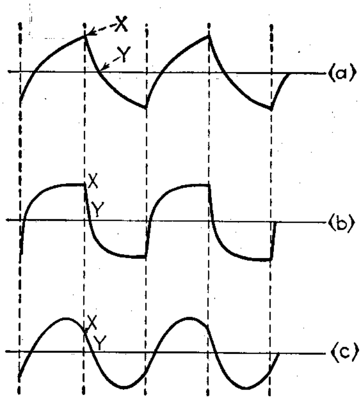


FIG. 12-53.—Typical voltage wave forms of the parallel inverter.

In the oscillograms of Fig. 12-53 the point X indicates the instant at which commutation takes place. At point Y, where the condenser voltage is zero, the potentials of the anodes are equal. Beyond Y the potential of the nonconducting tube again becomes greater than the ionizing potential. The tube must deionize in the interval between X and Y. It is seen that the time available for deionization is short with a large noninductive load. A study of the circuit shows that the voltage induced across one-half of the primary by the start of current in

[1] TOMPKINS, F. N., *Trans. Am. Inst. Elec. Eng.*, **51**, 707 (1932); SCHILLING, W., *Arch. Elektrotech.*, **27**, 22 (1933); RUNGE, I., and BECKENBACH, H., *Z. tech. Physik* **14**, 377 (1933); WAGNER, C. F., *Elec. Eng.*, **54**, 1227 (1935), **55**, 970 (1936).

the other half is in the direction to raise the anode voltage of the tube being extinguished. Hence this factor enters into the determination of the available deionizing time. A more complete analysis of commutation in the parallel inverter indicates that the size of the condenser must be carefully determined. The type and magnitude of the load, as well as the constants of the inverter circuit itself, must be taken into consideration. The condenser must also be designed to withstand the transient voltages to which it is subjected.

The grids of a parallel inverter may be excited from the output or from the anode circuit of one of the tubes. When this is done, it is essential that the exciting voltage have the proper phase relation to the output voltage. The phase of the grid voltage may be controlled by means of combinations of capacitance and resistance. Figure

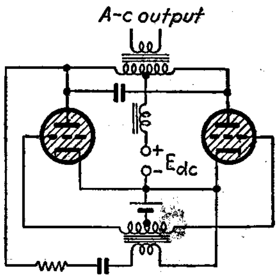


FIG. 12-54.—Self-excited parallel inverter.

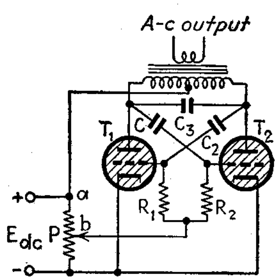


FIG. 12-55.—Self-excited parallel inverter.

12-54 shows one form of self-excited circuit.[1] The frequency is controlled by the constants of the grid and anode circuits. The frequency varies with d-c line voltage and with load current and power factor. To prevent excessive frequency variation and to ensure that incorrect phase relation between grid and output voltages will not cause the operation to become unstable and both tubes to fire simultaneously, considerable care is necessary in the design of the grid circuit.

A type of self-excited parallel inverter the operation of which depends upon grid rectification is shown in Fig. 12-55.[2] The action is as follows: Assume that tube 1 is conducting. Then the anode voltage of tube 2 is equal to d-c line voltage plus the voltage induced in one-half of the transformer secondary, and the condenser C_2 charges through R_1 and the grid of T_1 to a voltage equal to the induced transformer voltage V_t plus the voltage V_{ab} between points a and b of the voltage divider. When tube 2 fires, the commutating condenser C_3 extinguishes tube 1. The firing of tube 2 lowers the voltage of its anode to about 15 volts, the

[1] BAKER, W. R. G., FITZGERALD, A. S., and WHITNEY, C. F., *Electronics*, April, 1931, p. 581.

[2] STANSBURY, C. C., *Elec. Eng.*, **52**, 190 (1933).

normal tube drop, and thus makes the grid of tube 1 negative by an amount equal to the voltage of C_2 minus the anode voltage of tube 2, or $V_{ab} - 15$ volts. As C_2 discharges through R_1, the grid voltage of tube 1 rises until it becomes equal to the critical voltage, at which tube 1 again fires, and tube 2 is extinguished. In a similar manner, the condenser C_1 applies a negative voltage to the grid of tube 2 when tube 1 fires. The frequency of oscillation depends upon the size of the grid condensers and

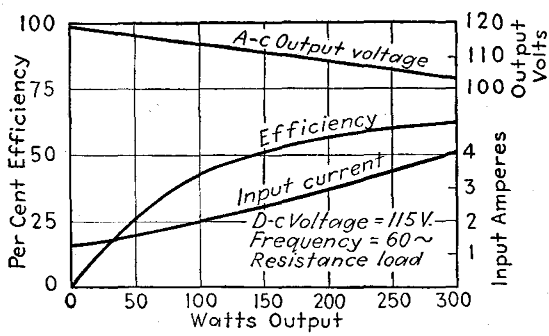


FIG. 12-56.—Performance curves for a parallel inverter.

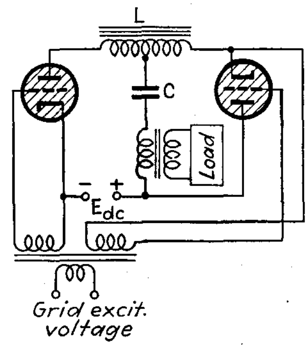


FIG. 12-57.—Separately excited series inverter.

resistors and upon the setting of the voltage divider P. Parallel inverters have also been designed that are based upon the circuit of Fig. 12-34c.[1]

Figure 12-56 shows curves of output voltage and efficiency for a parallel inverter operating from a 115-volt d-c. supply.[2] The low efficiency is caused by the comparatively small direct supply voltage and

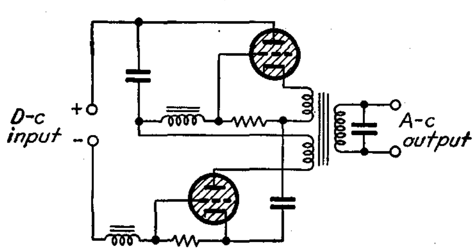


FIG. 12-58.—Self-excited series inverter using two condensers in the oscillating circuit.

the consequent high ratio of the tube drop to the supply voltage. Both the efficiency and the output are increased by raising the direct voltage.

12-36. Series Inverter.—The series inverter is based upon the series d-c circuit of Fig. 12-37. The output may be taken from a transformer with a center-tapped primary, connected in place of the center-tapped inductance L of Fig. 12-37, but commutation is affected adversely by the load because of the reduction of effective primary inductance. A

[1] HERSKIND, *loc. cit.*

[2] BAKER, W. R. G., and CORNELL, J. I., *Electronics*, October, 1931, p. 152.

better circuit is obtained by connecting the output transformer in series with the condenser, as shown in Fig. 12-57. Figures 12-58 and 12-59 show circuits in which each tube simultaneously charges one condenser and discharges another.[1] That of Fig. 12-58 is self-excited.

The wave form of the output voltage of the series inverter can be predicted from an analysis of the equivalent circuit of Fig. 12-60.[2] When

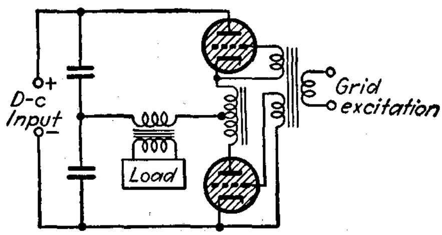


FIG. 12-59.—Separately excited series inverter using two condensers.

the load is sufficiently great so that the exciting current of the transformer is small in comparison with the load current, and the load is nonreactive, the load voltage has the same wave form as the condenser current. If the inductance-to-resistance ratio is high, the condenser charges and discharges nearly sinusoidally. If commutation then occurs at the instants when the condenser voltage reaches crest value, the output voltage will be a close approximation to a sine wave, as shown by oscillogram *a* of Fig. 12-61 (traced from the screen of a cathode-ray oscillograph). In order to make this possible, the design must be such that the natural frequency of oscillation of the series combination of condenser, choke, and loaded transformer is equal to the frequency of the exciting voltage. The phase of the exciting voltage relative to the output voltage must also be such as to cause commutation to take place at the crests of condenser voltage. The adjustment is critical, a small increase of effective inductance, such as results from a decrease in load, causing the circuit to stop oscillating. A decrease in inductance or capacitance, on the other hand, makes the natural frequency of the circuit greater than the excitation frequency, and each tube stops conducting before the other fires. The load voltage is then of the form shown by oscillogram *b*

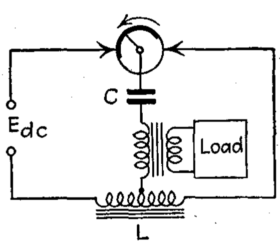


FIG. 12-60.—Equivalent circuit of the series inverter of Fig. 12-57.

[1] SABBAH, C. A., *Gen. Elec. Rev.*, **34**, 288, 580, 738 (1931); *Modern Radio*, February, 1933, p. 26.

[2] Mechanical inverters based upon the equivalent circuits of Figs. 12-52 and 12-60 may be readily constructed. Particularly good results are obtained with the series type. See H. J. REICH, *Electronics*, September, 1933, p. 252.

of Fig. 12-61. The regulation, stability, and output are better when the condenser charges and discharges exponentially, rather than sinusoidally. The output voltage then has a wave form similar to one of the oscillograms *c*, *d*, or *e* of Fig. 12-61. Oscillogram *c* is for the circuit adjustment that makes the condenser charging time[1] equal to half the excitation period. For the adjustment corresponding to oscillogram *d* the charging time is less than half the excitation period, whereas, for oscillogram *e*, the time taken to charge the condenser fully is greater than half the excitation period.

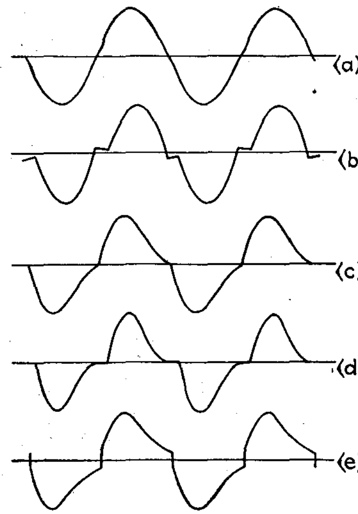


FIG. 12-61.—Typical wave forms of the output voltage of the series inverter shown in Fig. 12-57.

In Fig. 12-62 are shown curves of efficiency, frequency, and alternating voltage for a self-excited series inverter similar to that of Fig. 12-58.[2] The change of frequency with load shown in Fig. 12-62 would be prevented by the use of external excitation of fixed frequency.

By means of more complicated inverters based upon the fundamental series circuits it is possible to obtain output voltage that is very nearly sinusoidal. This is accomplished by the use of a large number of tubes, which are fired and extinguished in proper sequence so that the

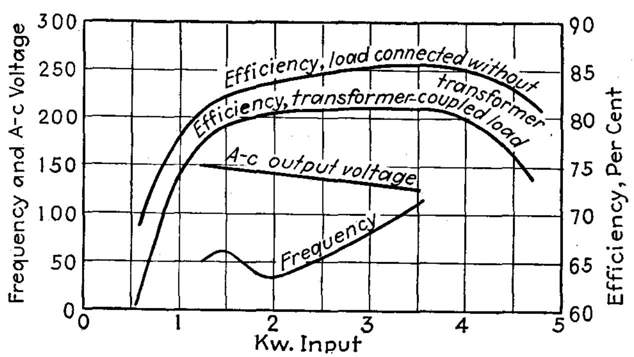


FIG. 12-62.—Performance curves for a parallel inverter.

resultant wave is approximately sinusoidal, and of inductance to smooth it out.[2]

[1] Although an infinite time is theoretically required to charge a condenser exponentially, actually the charging current stops when it becomes too small to maintain the arc.

[2] SABBAH, *loc. cit.*

12-37. The Relaxation Inverter.—The relaxation inverter is based upon the relaxation d-c circuit of Fig. 12-38, the load usually being connected through an output transformer and the grid circuit being supplied with suitable excitation.[1] The simplest type of separately excited circuit is that of Fig. 12-63. The a-c output may also be obtained by coupling to L_1, or through a transformer in series with L_1 and C_1, but the resulting damping tends to prevent extinction of the arc and may stop oscillation at high load. The excitation frequency may be adjusted so that the circuit oscillates once or several times for each cycle of grid exciting voltage, a single oscillation per cycle giving the best wave form. The circuit may be made self-exciting by obtaining the grid excitation from the output or in a number of other ways.

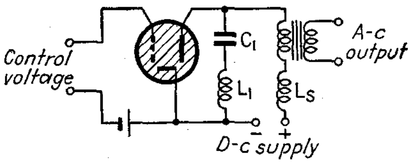
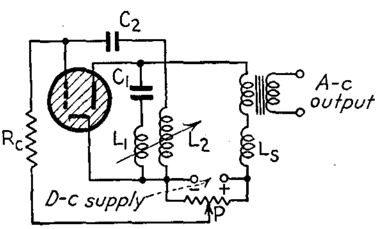


FIG. 12-63.—Separately excited relaxation inverter.

A reliable self-excited circuit is shown in Fig. 12-64. The action of this circuit is as follows: Assume that C_1 is charged. When the tube fires, the condenser discharges through the tube, the inductance L_1 causing the condenser polarity to reverse and extinguishing the tube, in a manner explained in Sec. 12-29. The condenser-discharge current flowing through L_1 induces a voltage pulse in L_2 which charges the grid condenser C_2 in the direction that makes the grid terminal negative. Because of the rectifying action of the grid, C_2 cannot discharge through the tube after the arc is extinguished. It discharges relatively slowly through the grid resistor R_c, during the time that the condenser C_1 is recharging from the d-c supply. Thus the negative grid voltage decreases (and may become positive) at the same time that the anode voltage rises, and at some time in the cycle the tube again fires, allowing the cycle to repeat. The frequency of oscillation is readily changed by varying the grid resistor R_c, the coupling between L_1 and L_2, and the setting of the potentiometer P. It is also affected by the size of C_1 and C_2, by the load, and by the direct voltage. L_2 may be omitted and C_2 connected to the junction of L_1 and C_1. Making the slider of the potentiometer P more positive increases the rate of change of grid voltage at the instant of firing



$C_1 = 10$ to $40\mu f$
$L_1 = 1.3$ mh ($R_1 = 0.13\Omega$)
$C_2 = 0.1\mu f$
$R_c = 50{,}000\Omega$

FIG. 12-64.—Self-excited relaxation inverter.

[1] REICH, H. J., *Rev. Sci. Instruments*, **3**, 580 (1932); *Rev. Sci. Instruments*, **4**, 147 (1933); *Elec. Eng.*, **52**, 817 (1933).

and therefore improves the constancy of frequency. The potentiometer may be omitted and R_c connected to the positive side of the line. L_1 should have low a-c resistance and so should preferably not contain iron. Good results are obtained with an air-core coil having an inductance of about 1.5 mh and a resistance of 0.15 ohm or less. In order to supply sufficient voltage to the grid when air coupling is used, L_2 should have many more turns than L_1.

At a given frequency the wave form is affected by the circuit constants and by the load. The wave shape may be predicted from an analysis of the equivalent circuits of Fig. 12-39. If the ratio of the effective inductance to the effective resistance in the supply branch of the circuit is small, which is true under heavy nonreactive load when no inductance

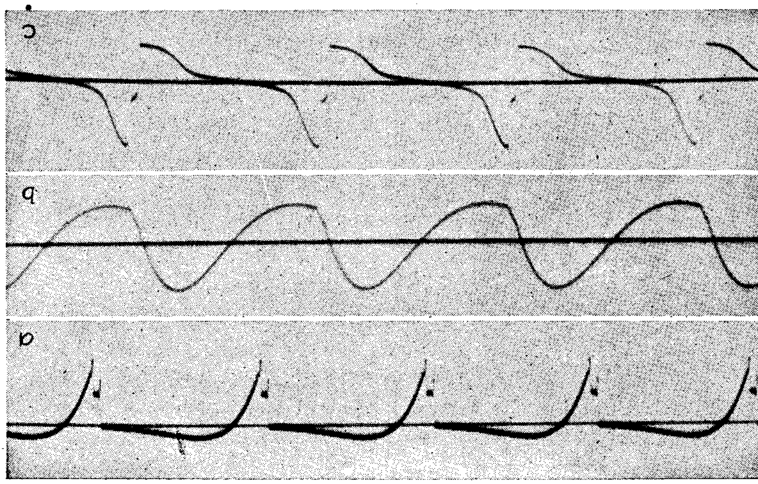


FIG. 12-65.—Oscillograms of output voltage of relaxation inverter: (*a*) nonreactive load without L_s; (*b*) heavy nonreactive load with L_s; (*c*) light nonreactive load with L_s.

L_s is used in series with the d-c supply, the condenser C_1 charges exponentially. The output voltage approximates saw-tooth form, as shown by the oscillogram of Fig. 12-65*a*. If the ratio of inductance to resistance is high, however, as may be true if inductance L_s is used in series with the d-c supply, the condenser charges very nearly sinusoidally. Then, if the tube fires at the peak of condenser voltage, the curve of charging current approximates a half of a sine wave. Under this condition the output voltage resembles a sine wave in form when the load is great enough so that the exciting current of the transformer is negligible. A typical oscillogram is shown in Fig. 12-65*b*. Under light load the secondary voltage is proportional to the rate of change of primary current, and the abrupt changes in rate of change of primary current when the tube fires distort the output voltage badly, as shown by the oscillogram of Fig. 12-65*c*.

Curves of output voltage, power, and efficiency at 60 cps as a function of load current for the circuit of Fig. 12-64 are shown in Fig. 12-66. The maximum power output increases with frequency and is proportional to the square of the direct voltage. The curves of Fig. 12-66 were obtained without inductance in series with the primary of the output transformer. Some reduction in efficiency results from the use of additional inductance, and the terminal voltage falls more rapidly at high load current.

It is interesting to note that, if deionization is sufficiently rapid, the circuit of Fig. 12-38 will oscillate with arc tubes which do not have grids. The circuit is then essentially that of the Poulsen arc oscillator. Power

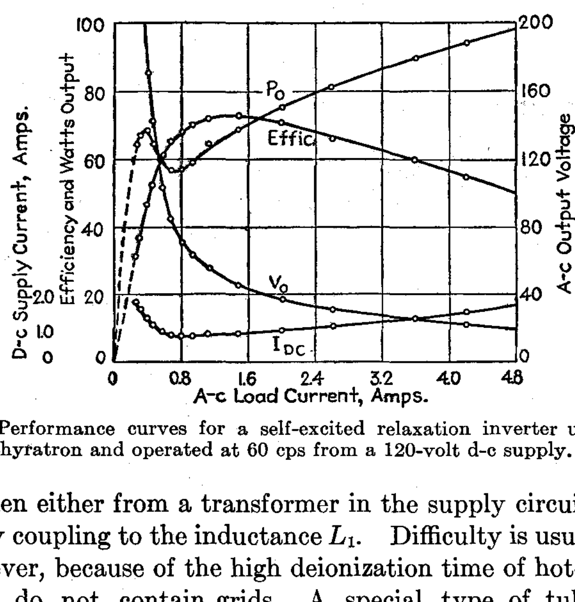


FIG. 12-66.—Performance curves for a self-excited relaxation inverter using an FG-67 thyratron and operated at 60 cps from a 120-volt d-c supply.

may be taken either from a transformer in the supply circuit, as in Fig. 12-63, or by coupling to the inductance L_1. Difficulty is usually encountered, however, because of the high deionization time of hot-cathode arc tubes that do not contain grids. A special type of tube has been developed for this purpose by means of which outputs as high as 500 watts have been obtained at frequencies in the kilocycle range.[1] The tube contains inert gas at relatively high pressure. The rapid deionization that is obtained probably results from the close electrode spacing used in this tube. Because of their short deionization time, mercury pool tubes may be successfully used in the relaxation inverter circuit. even when power is drawn from the condenser branch.[2]

12-38. Counter-e.m.f. Inverter.—A fourth type of inverter is based upon the counter-e.m.f. d-c circuit of Fig. 12-46. Figure 12-67 shows such an inverter. The commutating voltage impulses applied to ter-

[1] MILES, L. D., *Elec. Eng.*, **54**, 305 (1935).
[2] WATANABE, Y., and AOYAMA, R., *J. Inst. Elec. Eng. Japan*, **56**, 983 (1936).

minals *a-b* must occur at the proper instants in the cycle of grid excitation. Both the commutating and the excitation voltages can be obtained from the *a-c* output. The circuit is in some respects similar to that of the parallel inverter. It becomes practically equivalent to the parallel inverter circuit when the commutating voltage is obtained from the charge and discharge of a condenser connected across the primary of the commutating transformer *a-b*.[1] This principle has been used in the development of d-c power transmission.[2]

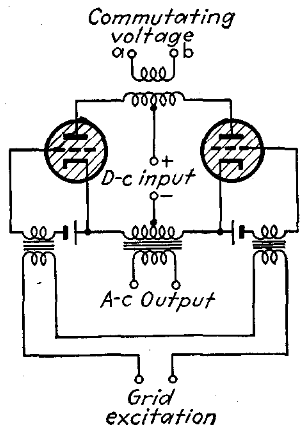


FIG. 12-67.—Separately excited counter-e.m.f. inverter.

12-39. D-c Transformer.—The problem of d-c power transmission has resulted in the development of *d-c-transformer* circuits which combine inverters with rectifiers. One type of circuit is shown in simplified form in Fig. 12-68.[3] The tubes T_1 and T_2, together with the transformer, comprise an inverter of the parallel type based upon the circuit of Fig. 12-34*c*. The discharge of the condensers C_1 and C_2 through the transformer primaries induces voltages in the secondaries which are rectified by the tubes T_3 and T_4. Tubes 3 and 4 are fired at, or shortly after, the times of extinction of tubes 1 and 2, respectively. Adjustment of the firing times of tubes 3 and 4 relative to the firing times of tubes 1 and 2 controls the output current. The voltage output is determined in part by the transformer turn ratio.

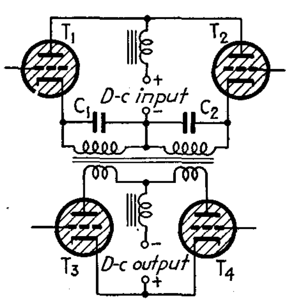


FIG. 12-68.—D-c transformer based upon the parallel inverter.

The relaxation inverter circuit of Fig. 12-64 may also serve as the basis of a d-c transformer. If the ratio of the inductance to the resistance in the supply branch is high, the condenser C_1 can charge to a maximum voltage approximately equal to twice the line voltage the first time it charges. If the grid-circuit constants are chosen so that the tube fires when this maximum voltage is reached, the condenser discharges to a negative voltage that is somewhat less than twice line voltage minus the tube drop. The second time the condenser charges from the line, it has this initial negative voltage and will charge to an equal positive

[1] WILLIS, C. H., *Trans. Am. Inst. Elec. Eng.*, **52**, 701 (1933).

[2] WILLIS, C. H., BEDFORD, B. D., and ELDER, F. R., *Elec. Eng.*, **54**, 882 (1935); *Electronics*, February, 1935, p. 56.

[3] HERSKIND, *loc. cit.*

voltage plus approximately twice line voltage.[1] If there were no losses in the circuit, the condenser voltage would increase by a little less than twice line voltage in each succeeding cycle until condenser or tube failure occurred. Because of losses, the increase of voltage becomes less each cycle until equilibrium is established, or until arcback occurs. The maximum voltage of C_1 may without difficulty be made equal to five or six times line voltage. The addition of a second rectifier tube and a filter, as in Fig. 12-69, gives a circuit that will supply direct current at a voltage several times the supply voltage. Such a d-c transformer, using an FG-67 tube and operating on a 110-volt d-c supply, will deliver 80 to 100 ma at 400 volts.

Other circuits have been devised for obtaining constant-potential direct current from constant-potential a-c or d-c supplies, and constant direct current from constant-potential alternating current.[2]

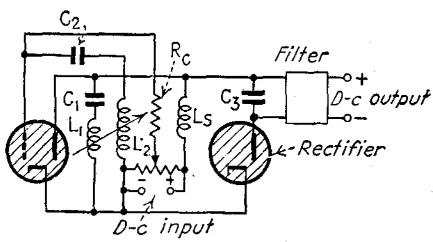


FIG. 12-69.—D-c transformer based upon the relaxation inverter.

12-40. Problems of Inverter Design.—The power output and efficiency of inverters increase with the applied direct voltage. In the parallel, series, and relaxation inverters, the power output also goes up with the size of the commutating condensers. The difficulty of designing large high-voltage condensers of reasonable cost is one of the major problems encountered in inversion. The counter-e.m.f. inverter does not require commutating condensers, but the magnitude of the commutating pulse must be increased with load. Other difficulties encountered in the design of inverters are poor wave form, high voltage regulation, and the need of low-voltage heater power. It is difficult to make 110-volt heaters, and they are correspondingly expensive. Low-voltage heaters may be heated initially by batteries, or from the d-c line through series resistance, and subsequently transferred to the a-c output of the inverter. Adequate protection must be provided to avoid damage to cathodes in case of failure of heater power or temporary interruption of d-c supply voltage and against high current surges if the circuit stops oscillating. Difficulties involving heaters are, of course, avoided by the use of igniter-type tubes, which will be discussed in a later section.

Since the time available for commutation is small, tubes used in inverters must have short deionization time. The FG-67 and FG-43 are among the tubes designed for inverter service.

[1] See footnote [1], p. 480.

[2] HERSKIND, C. C., *Elec. Eng.*, **56,** 1377 (1937).

12-41. Thyratron Motors.—An interesting and useful application of counter-e.m.f. extinction of thyratrons is made in thyratron commutation of motors. The counter-e.m.f. induced in the armature coils is used to extinguish the thyratrons and thus makes possible the use of tubes and a low-current, low-voltage, rotary grid distributor in place

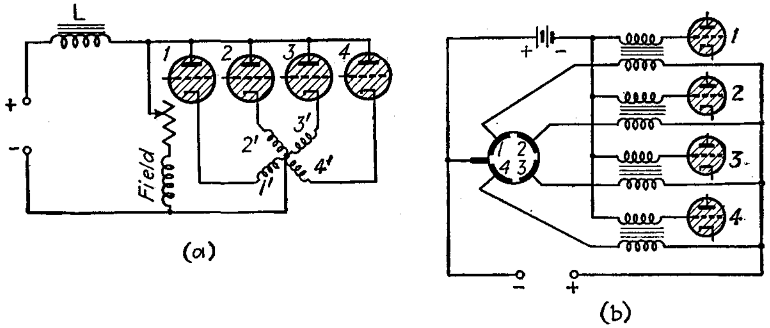


FIG. 12-70.—Circuit of d-c thyratron motor: (*a*) anode circuit; (*b*) grid commutating circuit.

of the usual commutator.[1] Because the grid voltages and currents are small, sparking is avoided and insulation problems are simplified.

Figure 12-70*a* is a simplified diagram of the thyratron d-c shunt motor. The associated grid circuit is shown in Fig. 12-70*b*. Assume that tube 1 conducts at such a time that the conductors of the armature coil 1′ exert torque in the desired direction. As the machine rotates, a counter-e.m.f., the magnitude of which may approach the impressed voltage, is induced in the coil 1′ and opposes the flow of current. The flux distribution of the air gap is designed so that the counter-e.m.f. generated in a coil is much larger under the trailing tip of the field pole than at any other point. When coil 1′ reaches the trailing pole tip, its counter-e.m.f. is greater than that of coil 2′, and commutation of current from tube 1 to tube 2 results if the grid voltages are correct. The position of the grid commutator on the shaft must be such as to fire tube 2 at or shortly after the instant at which the counter-e.m.f. in coil 1′ becomes greater than that in coil 2′.

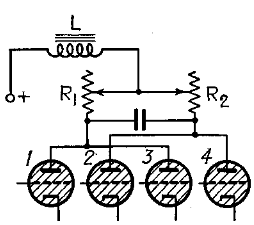


FIG. 12-71.—Starting circuit for d-c thyratron motor.

It will be noted that in the arrangement of Fig. 12-70 the armature coils are not prudently used, since each tube conducts only one-fourth of the time. Other circuits have been developed in which the coils are used more economically. It is also seen that commutation could not be

[1] WILLIS, C. H., *Gen. Elec. Rev.*, **36**, 76 (1933). See also W. M. GOODHUE, *Trans. Am. Inst. Elec. Eng.*, **51**, 714 (1932).

effective at low speed, since the counter-e.m.f. is very small. To remedy this difficulty, parallel commutation is used in conjunction with the starting resistance, as shown in Fig. 12-71. Cutting out the resistances as the motor comes up to speed also short-circuits the commutating condenser.

Thyratrons are also used with an a-c supply to operate series motors, the tubes acting both as rectifiers and as a commutator.[1] Figure 12-72 shows curves obtained from data taken on tests of a 400-hp, 2300-volt

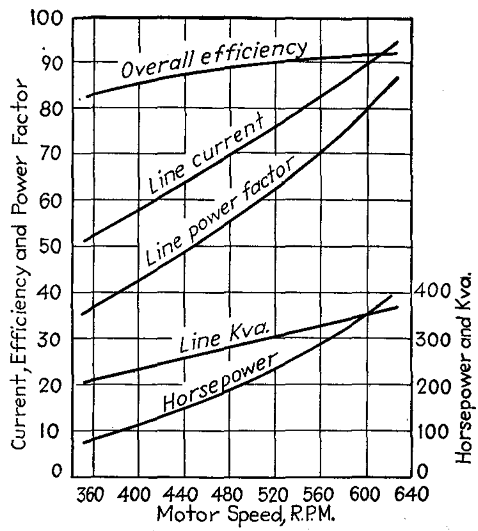


FIG. 12-72.—Performance curves for 400-hp., 2300-volt a-c thyratron motor.

motor with 18 12.5-amp (type FG-118) thyratrons acting as the commutator and as rectifiers of a three-phase a-c supply.[2] The high efficiency and low power factor at light loads are of particular interest. The data for these curves were obtained by changing the position of the brushes on the grid distributor, and thus the time in the cycle at which the tubes fire.

The advantages of the thyratron commutator are twofold. Since the commutator is stationary, rather than rotating as in conventional machines, it is possible to commutate current in the coils regardless of

[1] MARTI, O. K., *Trans. Am. Inst. Elec. Eng.*, **51**, 659 (1932); WILLIS, *loc. cit.*; SCHENKEL, M., and VON ISSENDORF, J., *Siemens Z.*, **13**, 289 (1933); STÖHR, M., *Elektrotech. u. Maschinenbau*, **51**, 584 (1933); ALEXANDERSON, E. F., and MITTAG, A. H., *Elec. Eng.*, **53**, 1517 (1934); **54**, 750 (1935); ALEXANDERSON, E. F., EDWARDS, M. A., and WILLIS, C. H., *Trans. Am. Inst. Elec. Eng.*, **57**, 343 (1938); GARMAN, G. W., *Trans. Am. Inst. Elec. Eng.*, **57**, 335 (1938).

[2] BEILER, A. H., *Trans. Am. Inst. Elec. Eng.*, **57**, 19 (1938); see also *Elec. World*, July 6, 1935.

the motor speed. Secondly, by controlling the portion of the cycle during which the tubes conduct, a simple and efficient method is made available for changing the speed of the motor over wide limits. The first advantage permits the use of high-voltage stationary d-c armatures and results in increase of efficiency. (The field rotates, as in synchronous machines, connection being made through slip rings.) The ease with which speed can be controlled eliminates the use of complicated systems of speed reduction. The speed is independent of supply frequency. The disadvantages, as determined from limited operating experience, are the high cost in comparison with other types of motors, the presumable fragility and lack of reliability of arc-discharge tubes as compared with other types of electrical and mechanical equipment, the higher voltage for which the motor must be insulated to protect against surges within the motor, and increased possibility of trouble because of the required auxiliary equipment. Just how important a place this type of motor will take among electrical machines is still to be determined.

An interesting tube-controlled motor of the notched-rotor type used in electric clocks has been described by P. B. King, Jr.[1] This motor has three field coils, two of which serve as driving coils and the other as a control coil for generating an exciting voltage which is applied to the grids. One form of the circuit uses a parallel inverter to supply power to the driving coils, the voltage induced in the grid coils serving to commutate the tubes. The speed is controlled by the capacitance of the commutating condenser. Other forms of the circuit, using high-vacuum tubes, have also been devised.

12-42. The Use of Arc Tubes as Amplifiers.—Loss of control of plate current by the grid prevents the use of ordinary thyratrons in amplifiers or feedback oscillators. Special tubes have been designed, however, in which the use of gas or vapor at relatively low pressure causes sufficient ionization to result in the partial neutralization of electron space charge.[2] This greatly reduces the voltage drop through the tube without causing loss of grid control. The advantages of such a tube lie in the low anode voltage and the accompanying low plate resistance and high transconductance, and in increased cathode efficiency. A serious disadvantage is the flow of grid current at negative grid voltages. Tubes of this type have been successfully used in amplifiers and oscillators, both at audio and at radio frequency.

12-43. A-c Operation of Thyratrons. Phase Control.—When a thyratron is operated on alternating voltage, the tube extinguishes near

[1] KING, P. B., JR., *Electronics*, January, 1936, p. 14.

[2] NELSON, J. R., and LE VAN, J. D., *QST*, June, 1935, p. 23; LE VAN, J. D., and WEEKS, P. T., *Proc. I.R.E.*, **24**, 180 (1936); BOUMEESTER, H., and DRUYVESTEYN, M. K., *Philips Tech. Rev.*, **1**, 367 (1936); see also (abstr.) *Electronics*, May, 1937, p. 66.

the end of the positive half cycle, and the grid regains control without the use of special circuits for interrupting the anode current. By controlling the time in the cycle at which the tube fires, the grid can vary the average anode current. This type of control may be attained by varying the phase of the grid voltage with respect to the anode voltage.[1] The principle of operation can be explained by the aid of the diagrams of Fig.

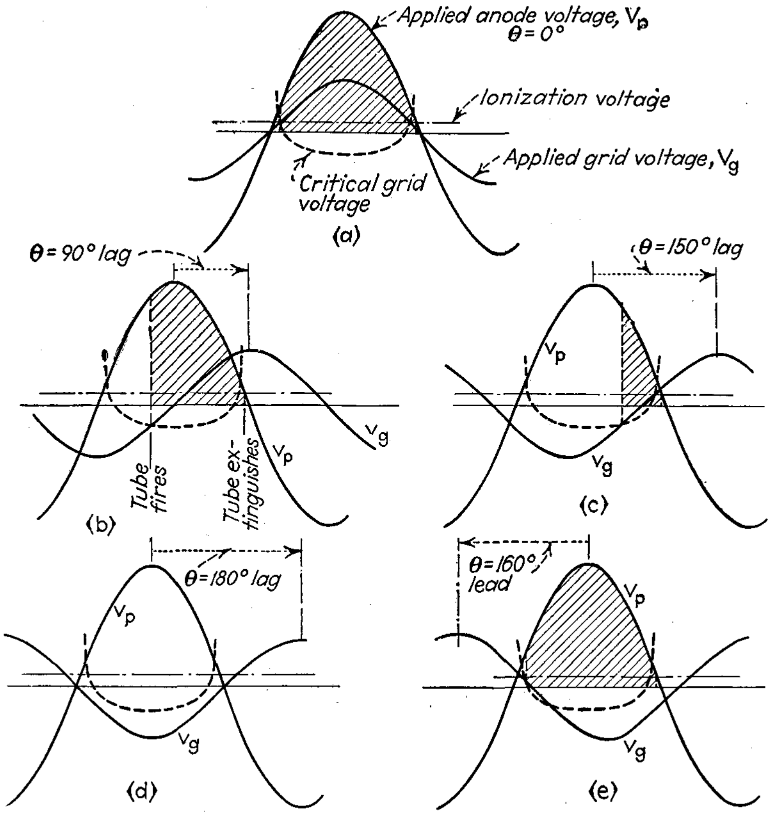


FIG. 12-73.—Variation of firing period of thyratron by change of phase angle between grid voltage and anode voltage.

12-73, which apply to resistance load. The solid curve v_p represents the sinusoidal applied voltage in the anode circuit, which is equal to the instantaneous anode voltage e_p up to the time of firing. The other solid curve v_g represents the instantaneous applied voltage in the grid circuit. Under the assumption that no grid current flows previous to firing, the grid voltage e_g is equal to v_g up to the instant of firing. The dotted curve indicates the critical grid voltage at which the tube would fire at the instantaneous anode voltage given by the curve of v_p. The

[1] TOULON, P., U. S. Patent 1654949.

critical-grid-voltage curve can be derived from the grid-control characteristic by plotting values of critical grid voltage corresponding to the anode voltage at various instants of the cycle.

The tube fires at the instant in the cycle when the grid voltage becomes less negative (or more positive) than the critical grid voltage. The anode voltage then becomes equal to the normal tube drop and, with resistance load, the anode current rises abruptly to a value equal to the applied anode voltage less the tube drop, divided by the load resistance. Current continues to flow until the anode supply voltage falls below the ionization potential toward the end of the positive half cycle. The wave of tube current during the conducting portion of the cycle has nearly the same shape as the wave of applied anode voltage. To avoid needless complication of the diagrams, the current waves are not shown, but the portion of the cycle during which current flows is indicated by shading. The instant in the cycle at which the tube fires is determined by the intersection of the applied-grid-voltage curve with the curve of critical grid voltage. Diagrams *a*, *b*, *c*, and *d* correspond respectively to 0, 90, 150, and 180 degrees lag of grid voltage relative to anode voltage. Inspection of these diagrams shows that the instant of firing is retarded by increase of angle of lag θ and that the portion of the cycle during which current flows is decreased, becoming zero when θ is slightly less than 180 degrees. The average anode current therefore decreases progressively with increase of angle of lag.

The reader can readily show by means of similar diagrams that, if the grid excitation voltage leads the anode excitation voltage by any angle up to about 170 degrees, the tube fires as soon as the anode voltage exceeds the ionization potential, near the beginning of the cycle. Since the grid has no control after the tube fires, current flows during practically the entire positive half cycle. Diagram *e* corresponds to a 160-degree angle of lead. When the angle θ approaches 180 degrees lead, however, a value is reached at which the grid voltage is more negative (or less positive) than the critical value throughout the positive half cycle. For values of θ greater than this, no current flows. Thus a small change of θ in the vicinity of 170 degrees lead causes an abrupt change of anode current from full average value to zero. This will be termed *on-off* control, in contrast to the *gradual* or *progressive* control obtained when the grid voltage lags the anode voltage.

When the load contains inductance, the anode current wave differs in shape from that of the applied anode voltage. The current rises gradually after firing and continues to flow after the anode supply voltage falls below the ionization potential, the shape of the wave depending upon the ratio of inductance to resistance in the load. The flow of anode current during a portion of the negative half cycle reduces the time available for deionization and may thus cause the grid to lose control.

The critical grid voltage of a positive-grid tube such as the FG-33 is very nearly constant over the whole range of anode voltage, as can be seen from Fig. 12-29. Therefore, the critical–grid-voltage curve of such a tube is very nearly a horizontal straight line lying above the axis. The critical grid voltage of a negative-grid tube, such as the FG-57, on the other hand, varies considerably, as shown by Fig. 12-27. If the anode voltage amplitude is high, the critical–grid-voltage curve of a negative-grid tube is approximately sinusoidal except at the beginning and end of the half cycle. If the anode voltage amplitude is small, the nonlinearity of the grid-control characteristic causes the curve of critical grid voltage to depart markedly from sinusoidal form. Because a positive grid voltage is required to fire most negative-grid tubes at low

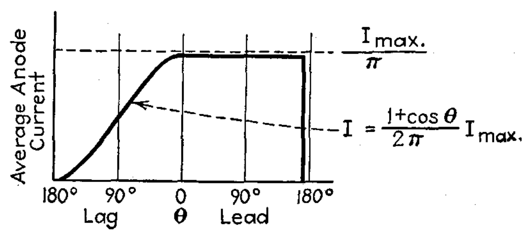


FIG. 12-74.—Theoretical curve of average anode current as a function of the phase angle θ of the grid voltage relative to the anode voltage.

anode voltage, the curve lies above the axis at the beginning and end of the half cycle, as in Fig. 12-73.

If the amplitude of the impressed grid voltage is much greater than the maximum critical grid voltage, the curves intersect at a sharp angle, and the shape of the critical–grid-voltage curve has little effect upon the instant at which firing occurs at a given phase angle. When the grid-voltage amplitude is small, however, the average current at a given value of θ is appreciably affected by the form of the grid-voltage curve. A series of diagrams for phase angles lying between 150 and 180 degrees lag shows that one result of the rapid rise of critical grid voltage to positive values at small anode voltage is to prevent continuous control at very low average anode current when the grid-voltage amplitude is small. The current decreases progressively with increase of phase angle down to a certain value and then falls abruptly to zero. Smooth control necessitates the use of adequate grid voltage.

Figure 12-74 shows the manner in which the average anode current varies through a resistance load as the phase angle θ is varied from 180 degrees lag to 180 degrees lead. This diagram is derived under the assumption that the amplitude of the applied grid voltage greatly exceeds the maximum critical grid voltage. The fairly wide range of phase angle in the vicinity of 180 degrees throughout which the anode current is zero results from failure of the tube to conduct at anode voltages lower than the ionization potential. Because of reduction of the fraction of the cycle during which the anode voltage is less than the ionization potential, the portion of the control characteristic of Fig. 12-74 over which the current is zero decreases with increase of anode voltage amplitude.

A similar type of control is effected by keeping the phase of the grid voltage constant and varying the amplitude. If v_g lags v_p by slightly less than 180 degrees, as shown in Fig. 12-75, the variation of anode current is gradual. If the angle of lag is 180 degrees, on-off control is obtained. Phase control can also be used with tubes controlled by external grids[1] or magnetic fields[2] or with the ignitron (see Sec. 12-59).

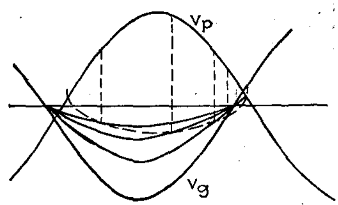


FIG. 12-75.—Wave diagram showing the variation of firing period by change of amplitude of grid voltage.

The anode current can be varied gradually over a smaller range when alternating voltage is used in the anode circuit and direct voltage in the grid circuit. If the grid voltage is insufficiently negative (or too positive) to prevent the tube from firing at the beginning of the cycle, anode current flows during nearly the whole positive half cycle. Increase of negative grid voltage (or decrease of positive voltage, in the case of a positive-grid tube) makes the tube fire later in the cycle. When the grid voltage is only slightly less than that required to prevent firing at the crest of the applied anode voltage, the tube conducts for approximately a quarter of the cycle. Further increase of negative grid voltage prevents the tube from firing at all. The current can be decreased continuously from a value corresponding to half-wave rectification to half this value. One disadvantage of this type of control is that small fluctuations of anode supply voltage or of grid voltage cause relatively large changes in average anode current. Large changes of current are also produced by small changes in tube characteristics. This type of control is not applicable to external grid tubes but can be used with magnetic control.

12-44. Phase-control Circuits.—Any convenient method can be used to shift the phase of the grid voltage. A simple method is the use of a phase shifter (a transformer having a single-phase secondary winding which can be turned through 360 electrical degrees within a polyphase winding arranged to produce a rotating field) as in Fig. 12-84*a*. A series combination of resistance and reactance is the basis of a number of prac-

[1] See footnote [3], p. 444.
[2] See footnote [4], p. 444.

tical phase-control circuits, of which Fig. 12-76 is a typical example. The phase of the grid voltage is varied by changing R or C. The function of R' in series with the grid is to limit the grid current to a safe value after the tube fires and in case of short circuit of C. The required value of R' depends upon the minimum values of the two phase-shifting impedances. It may be omitted if both impedances are at all times large enough to limit the grid current to the maximum average value. With tubes in which appreciable grid current flows previous to firing, the voltage drop through this resistor may affect the phase control adversely. For this reason, this resistor should, in general, not be much larger than necessary to limit the grid current after firing. It has little effect, however, with shield-grid thyratrons. For the same reason, the values of Z_1 and Z_2 should not be larger than necessary to limit the current through these impedances to allowable values.

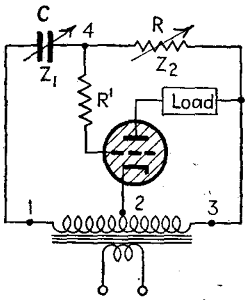


FIG. 12-76.—Simple type of practical phase-control circuit.

The vector diagram representing the voltages of the circuit of Fig. 12-76 up to the time that the tube fires is shown in Fig. 12-77. The current I_{21} is used as reference, and the grid current is assumed to be negligible previous to firing. The voltage V_{43} across the resistance is in phase with the current, whereas the voltage V_{14} across the condenser lags the current by 90 degrees. The secondary voltage V_{13} is the vector sum of V_{14} and V_{43}. The grid voltage V_g is equal to V_{24}, which is the vector sum of V_{14} and V_{21}, or the vector difference of V_{14} and V_{12}. The impressed anode voltage V_p is equal to V_{23}.

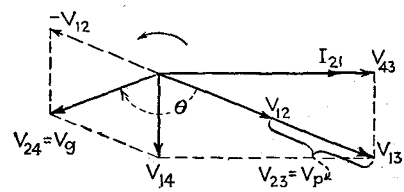


FIG. 12-77.—Vector diagram of the circuit of Fig. 12-76 previous to firing.

A vector diagram that is more susceptible to analysis may be derived from that of Fig. 12-77 by application of the theorem that the significance of a vector diagram is unchanged by displacing the vectors without altering their magnitudes or the angles between them. The resulting diagram for the special case in which the secondary is center-tapped is that of Fig. 12-78.[1] Since V_{14} and V_{43} are perpendicular and their sum is always equal to the secondary voltage V_{13}, which is assumed to be constant, it follows from a theorem of plane geometry that their intersection must lie on the circumference of a circle, the diameter of which is V_{13}. It can be readily seen from this diagram that decreasing R, or increasing the capacitive reactance by decreasing C, decreases θ, the

[1] WEINBACH, M. P., "Alternating Current Circuits," p. 98, The Macmillan Company, New York, 1933.

angle of lag of the impressed grid voltage relative to the impressed anode voltage, and thus produces a gradual increase of average anode current.

When the resistance and condenser are interchanged, the vector diagram becomes that of Fig. 12-79.[1] The angle θ is now measured in the opposite direction, showing that the grid voltage leads the anode voltage by the angle θ. If θ is made to approach 180 degrees by the use of very small resistance or capacitance, the current is zero, as shown by Fig. 12-74. A slight increase in either R or C will retard the phase of the grid voltage and cause the average current to increase suddenly to full value. This circuit evidently gives on-off control.

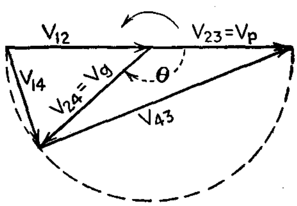


FIG. 12-78.—Circle vector diagram equivalent to the vector diagram of Fig. 12-77.

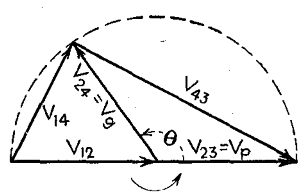


FIG. 12-79.—Circle vector diagram for the circuit obtained by interchanging C and R in Fig. 12-76.

An inductance can be used in place of the condenser in the circuit of Fig. 12-76. Since it is impossible, however, to construct an inductance coil that does not have appreciable resistance, the voltage across the inductance will not be 90 degrees out of phase with the current, and the vectors representing the voltages across the inductance and the resistance will not be perpendicular. The general behavior can, however, be determined from a diagram in which the vectors are assumed to be perpendicular. The reader will find it instructive to construct these diagrams for the circuit obtained by replacing C in Fig. 12-76 by an inductance and for the circuit in which the inductance and resistance are transposed. Table 12-II is convenient in determining the circuit to be used for a particular application. [In the table, z_2 refers to the impedance adjacent

TABLE 12-II

z_2 (adjacent to load)	z_1	Manner of varying resistance	Manner of varying reactance	Manner in which current increases
Resistance	Capacitance	Decrease	Increase	Gradually
Capacitance	Resistance	Increase	Decrease	Suddenly
Resistance	Inductance	Decrease	Increase	Suddenly
Inductance	Resistance	Increase	Decrease	Gradually

[1] Note that transposing C and R is equivalent to changing the connection of the anode circuit from terminal 3 to terminal 1 of the transformer.

to the load, and z_1 to the impedance between the grid and the side of the transformer to which the load is not connected (see Fig. 12-76).]

It is a simple matter to determine, without reference to Table 12-II, the correct circuit to produce a desired result. Suppose, for instance, that it is desired to increase the current suddenly by an increase of resistance. Figure 12-74 shows that sudden increase of average current requires that the angle of lead of the grid voltage relative to the anode voltage should be reduced by a small amount from a value somewhat less than 180 degrees. The grid voltage must lie in the second quadrant of the circle vector diagram, and so the diagram must be that of Fig. 12-80. If increase of resistance is to decrease the lead of V_g, V_{14} must be the voltage across the resistance, and V_{43}, which lags V_{14} by 90 degrees, must be the voltage across a condenser. Thus the impedance adjacent to the load must be a capacitive reactance, and the other must be the resistance.

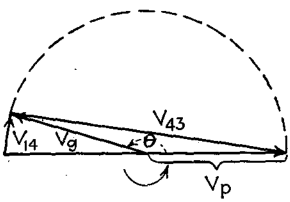


FIG. 12-80.—Circle vector diagram for phase-control circuit in which increase of resistance causes an abrupt increase of average anode current.

The analysis that has been made for thyratron phase control can also be applied to grid-glow–tube control if account is taken of the fact that, in order for cathode-to-anode breakdown to occur, the anode voltage must exceed the grid voltage by an amount determined by the grid current that flows subsequent to grid-cathode ignition (see Fig. 12-16).

When the line voltage is high enough to give the required load current, some reduction in cost may be effected by drawing the load current

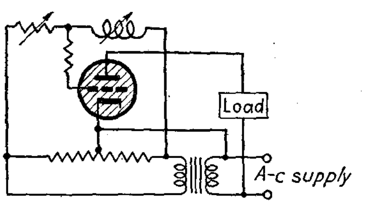


FIG. 12-81.—Variant of the circuit of Fig. 12-76 in which the load current is taken directly from the line.

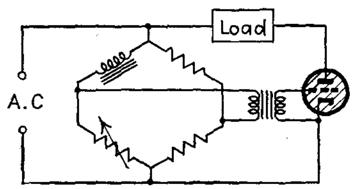


FIG. 12-82.—Form of phase-control circuit in which the load current is taken directly from the line.

directly from the line. The center-tapped transformer may then be replaced by an untapped transformer shunted by a center-tapped resistance as in Fig. 12-81. This transformer need have only sufficient current capacity to carry the voltage-divider current and the grid current after firing. The secondary voltage rating may usually be less than the anode voltage. To ensure the proper phase relations of the voltages, the resistance of the voltage divider should be small in comparison with the phase-shifting impedances. A somewhat similar circuit is shown in Fig. 12-82. It should be noted that the circuits of Figs. 12-76, 12-81, and 12-82 are suitable only for use with loads that will operate on direct or pulsating current.

Another modification of the circuit of Fig. 12-76 that is particularly useful when the load current or voltage exceeds the rated value for the tube is shown in Fig. 12-83.[1] When the phase of the grid voltage is adjusted so that the tube does not conduct, the current through the load cannot exceed the normal exciting current of the transformer. As the average tube current is increased, the primary current also increases, both because of the alternating component of secondary current and because of the saturating effect of the direct component of secondary current. In order that the minimum load current shall be small, the transformer primary impedance must be much larger than the load impedance. If necessary, the minimum load current may be reduced by shunting the load with a resistance *R*. If a step-up transformer is used, the load current exceeds the tube current. In this manner a tube can be made to control load current many times as great as the rated maximum average anode current. The turn ratio, and hence the load current, are limited because the secondary voltage must not exceed the rated maximum peak inverse or forward voltage. If a step-down transformer is used, on the other hand, the tube can be used with a supply voltage greatly in excess of the rated maximum peak voltage, the turn ratio and voltage being limited by the allowable average tube current. This circuit is suitable only for loads that operate on alternating current.

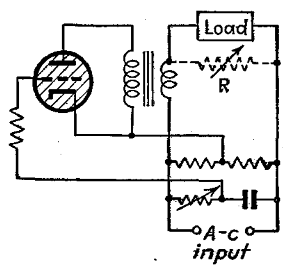


FIG. 12-83.—Phase-control circuit that is useful when the current or voltage of the load exceeds the maximum value for which the tube is rated.

The load current can be doubled by the use of two thyratrons in a full-wave circuit such as that of Fig. 12-84*a*.[2] This circuit also illustrates the use of a phase shifter, which may be replaced by a combination of resistance and reactance as in Fig. 12-84*b*.[2]

A high-vacuum amplifier tube may be used in place of the resistance in phase-control circuits. By this means a very small direct voltage in the grid circuit of the high-vacuum tube can control large amounts of current and power. Figure 12-85 shows circuits based upon that of Fig. 12-76.[3]

[1] CRAIG, P. H., *Electronics*, **6**, 71 (1933), **10**, 26 (1937); LOWRY, E. F., and GESSFORD, R. K., *Electronics*, **9**, 27 (1936).

[2] BAKER, W. R. G., FITZGERALD, A. S., and WHITNEY, C. F., *Electronics*, January, 1931, p. 467.

[3] BAKER, FITZGERALD, and WHITNEY, *ibid.;* CORONITI, S. C., *Proc. I.R.E.*, **31**, 653 (1943).

Similar circuits using inductance in place of capacitance have discontinuities in the curve of average current *vs.* control voltage. Oscillographic

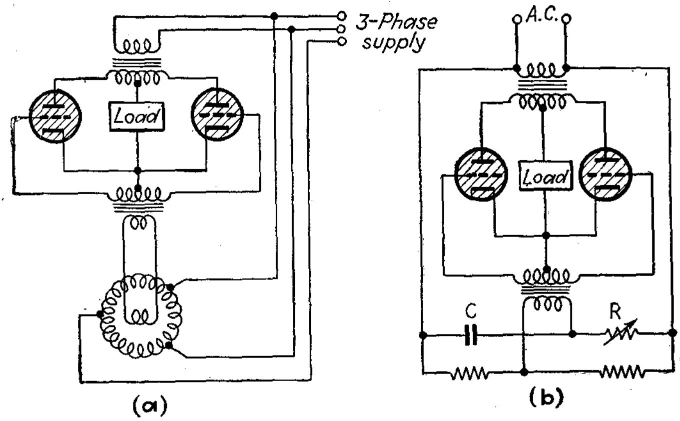


FIG. 12-84.—Full-wave phase-control circuits with (*a*) phase shifter and (*b*) phase-shifting network.

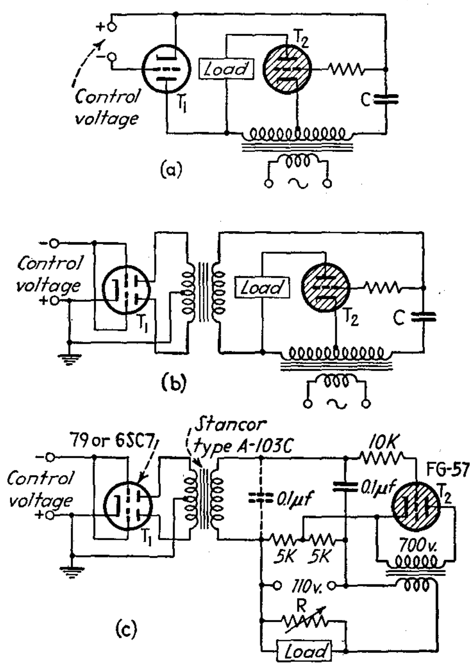


FIG. 12-85.—Phase-control circuits in which the use of vacuum-tubes as resistance makes possible control by direct voltage.

studies show that this difficulty is the result of transient oscillations set up periodically in the inductance when the high-vacuum-tube plate current cuts off. An objection to the circuit of Fig. 12-85*a* is the fact that the cathode of T_1 connects to the grid of T_2. Since the cathode of T_2 is capacitively grounded through the heater transformer, the cathode of T_1 cannot be grounded without affecting the control action adversely. When the cathode of T_1 is not grounded, however, the control is affected by body capacitance to T_1. This objectionable characteristic is avoided in the circuits of Figs. 12-85*b* and *c*. The function of the resistance R in the circuit of Fig. 12-84*b* is to adjust the minimum load current. (Reliable operation of this circuit necessitates that the phase relation between the grid and plate voltages of the thyratron be correct. If the action is erratic, or if the circuit fails to function, the connections to the primary or secondary of the thyratron anode transformer should be reversed.) Both circuits give gradual control of load current. By the use of a diode rectifier in the grid circuit of the vacuum tube T_1, the control can be made sensitive to small alternating voltages of any frequency. A very sensitive light-operated control is obtained by using a phototube to change the grid voltage of the high-vacuum tube (see Sec. 13-8).

The field of application of most phase-control circuits in which a very high resistance may be used as the controlling element can be greatly extended by replacing the resistance by a phototube. This is discussed in Chap. 13 (see Sec. 13-10).

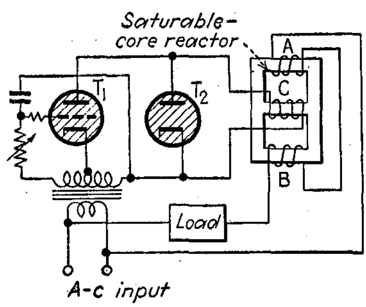


FIG. 12-86.—Phase-control circuit incorporating a saturable reactor.

12-45. Use of Saturable Reactor in Phase-control Circuits.—Sometimes it is necessary to apply phase control to loads that require alternating current. Although this can be accomplished by means of the circuits of Fig. 12-84, the fact that the anode current in general consists of periodic pulses lasting only a portion of the cycle causes the wave form of the load current to be poor. This difficulty can be avoided by using the thyratron current to saturate the core of a reactor, the reactance of which controls the load current. This method has the additional advantage of making possible the control of larger load currents at higher voltages.

A typical control circuit incorporating a saturable reactor is shown in Fig. 12-86. The load current is limited by the reactance of the coils A and B. These coils are of equal dimensions and are connected so that the fluxes that they produce add in the outer arms of the core but cancel in the center arm, which holds the winding C. For this reason no alternating voltage is induced in coil C. The anode current of the thyratron T_1 flows through coil C, producing a d-c flux which divides equally between the outer arms of the core and thus reduces the reactance of A and B. When the average anode current is zero, the load current cannot exceed the exciting current of the reactor. If the maximum average anode current is sufficient to saturate the core, the resulting reactance will be so small as to have very little effect upon the load current. The function of the rectifier T_2 is to provide a path for the direct current in coil C during the portion of the cycle when T_1 does not

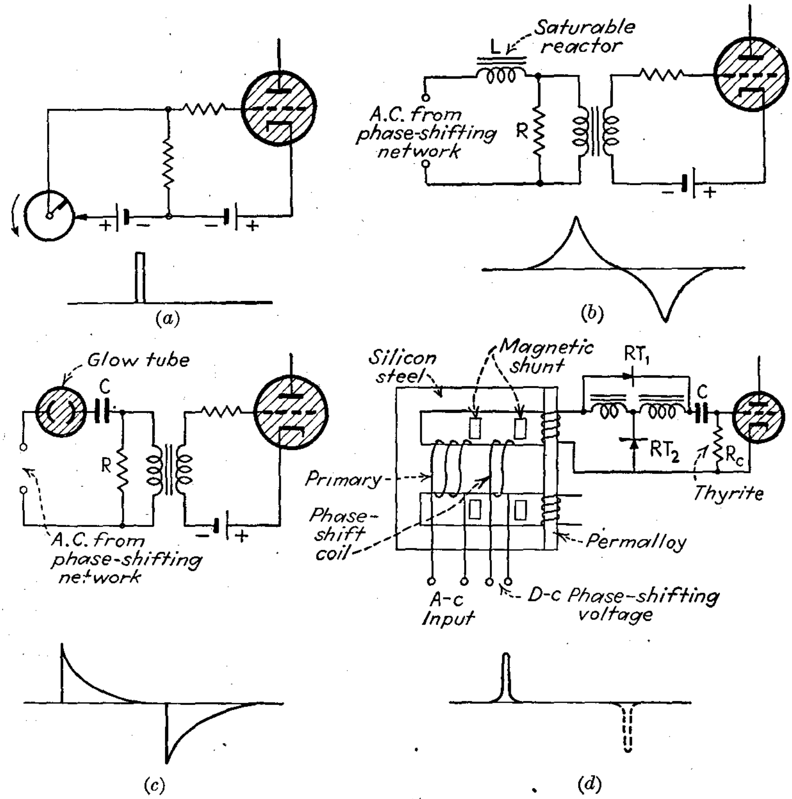


FIG. 12-87.—Circuits for producing voltage pulses for thyratron control.

conduct.[1] If T_2 were omitted, the voltage induced in coil C would tend to maintain current flow in T_1 during the whole cycle and thus prevent the variation of saturating current by phase control of T_1.

12-46. Voltage Impulses for Grid Control.—When it is very important that the thyratron fire at exactly the same point in the cycle in succeeding cycles, it is sometimes desirable to use a fixed negative grid bias and to fire the tube by means of a positive pulse of grid voltage.[2] This type of control has the additional advantages that the negative

[1] BABAT, G., *Proc. I.R.E.*, **22**, 314 (1934).

[2] MORACK, M. M., *Gen. Elec. Rev.*, **37**, 288 (1934).

grid voltage aids deionization and that the positive-ion bombardment of the grid obtained with negative bias helps to remove emitting material deposited upon the grid.

Figure 12-87 shows four methods by which voltage impulses can be produced and the shape of the resulting voltages. In (*a*) the pulse is produced by a synchronous contactor which applies an additional positive grid battery to the grid circuit for an instant. Circuit (*b*) uses a saturable reactor in series with a resistor. The reactor saturates near zero of the applied voltage and produces, across the resistor, a peaked voltage wave that is applied to the grid through a transformer. In (*c*), when the glow tube breaks down, the condenser current starts to flow abruptly, giving a secondary voltage with a steep wave front. In circuits (*b*) and (*c*) the firing time is changed by shifting the phase of the input voltage to the pulse circuit.

A more satisfactory method of producing control pulses is by the use of an impulse transformer of the type shown in (*d*).[1] The legs of the core holding the grid-voltage coils are of small cross section and are made of permalloy, which has high permeability up to a certain flux density, above which it saturates rapidly. The secondary voltage is very peaked and of short duration. The positive pulse is applied to the thyratron grid through the rectifier RT_1. The negative pulse is by-passed through the rectifier RT_2. *Thyrite* is used for R_c because its resistance increases with decreasing voltage, thus helping to maintain a small negative grid voltage caused by electrons trapped on C during the time the grid is positive, but allowing rapid discharge when the voltage is high. The time in the cycle at which the voltage pulse occurs is readily changed by means of direct saturating current in the phase-shifting winding. Current in one direction advances the phase of the secondary voltage; current in the other direction retards the phase. By using both secondary coils it is possible to control two tubes which fire during opposite halves of the cycle of supply voltage, giving full-wave operation. Other types of pulse generators have been discussed in Secs. 10-7 and 10-8.

12-47. Tubes for Phase-control Circuits.—In most phase-control circuits the grid-circuit resistance or impedance is high. For this reason these circuits function best with low-grid-current tubes, preferably of the shield-grid type.

Applications of A.c.-operated Control Circuits.—Phase-control circuits are used for the control of illumination, for the control and stabilization of alternating load current and voltage, for the control and stabilization of the voltage and frequency of alternators, for temperature control and stabilization, for the control of rectifier output voltage, and for

[1] KILTIE, O., *Elec. Eng.*, **51**, 802 (1932).

many other purposes. The advantages of thyratron control of current over the use of series resistance are the large saving in power and the ease with which the current can be varied by means of small low-current variable resistors, reactors, or condensers. As a switch, the a.c.-operated thyratron has found its greatest application in the control of spot welders. The following applications are of special interest because of the principles involved. For a comprehensive treatment of commercial and laboratory applications of thyratrons the reader should refer to other texts.[1]

12-48. Illumination Control.—Any of the phase-control circuits that have been discussed can be used for the purpose of controlling illumination. The simple phase-control circuits of Figs. 12-76 and 12-81 to 12-86 control the load current but not the load voltage. These circuits are consequently not entirely satisfactory for illumination control, since the brightness of each lamp depends not only upon the setting of the controller but also upon the number of lamps in the circuit. This difficulty can be avoided by the use of a circuit that controls the lamp voltage, rather than the current. Such a circuit, developed by the General Electric Company, is shown in Fig. 12-88.[2]

12-49. Alternating Load-voltage Control.—The circuit of Fig. 12-88 is useful not only in the control of illumination but also in any application in which it is desired to control the voltage of an a-c load and to stabilize the voltage against variations of load. The operation of this circuit is based upon the reactor-saturation control of Fig. 12-86, control of the thyratron T_1 being obtained by variation of direct grid voltage. This grid voltage is made up of two parts, V_L and V_s. The former is derived by rectifying the voltage between a and b, which varies with load voltage; the latter is derived by rectifying the portion of the line voltage between c and d. When alternating voltage is first applied, the reactor has maximum reactance, the load voltage is low, and so V_s exceeds V_L. This causes the grid of T_1 to be positive and allows the tube to conduct during the entire positive half cycle. The resulting high rectified current through the d-c winding of the reactor reduces its inductance and raises the load voltage. As the load voltage rises, V_L also rises, eventually making the grid negative and reducing the conduction time of T_1 and hence the saturating current. Equilibrium is established at some value of load voltage that depends upon the setting of P_1. If the voltage across c-d is lowered, V_s goes down, and the thyratron grid becomes so negative that the thyratron stops conducting. The direct current through the reactor falls, lowering the lamp voltage. This in turn

[1] HENNEY, K., "Electron Tubes in Industry," 2d ed., McGraw-Hill Book Company, Inc., New York, 1937; GULLIKSEN, F. H., and VEDDER, E. H., "Industrial Electronics," John Wiley & Sons, Inc., New York, 1935.

[2] CHAMBERS, D. E., *Elec. Eng.*, **54**, 82 (1935).

reduces the voltage V_L and hence reduces the negative grid voltage. Equilibrium is again established at a lower value of anode current and load voltage.

Since the action of this circuit depends upon balancing the voltage between *c* and *d* against that between *a* and *b*, and not upon the load current, the load voltage is practically independent of load current. The load voltage is also practically independent of tube and reactor characteristics and of frequency. The speed of response depends principally upon the rapidity with which the steady reactor flux can change. It should be noted that the voltage between *c* and *d* varies with alternating line voltage and that this device will not, therefore, correct for fluctuations of line voltage.

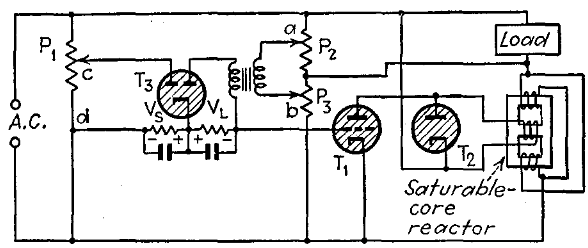


FIG. 12-88.—Thyratron control circuit in which the load voltage is independent of load resistance.

12-50. Voltage, Speed, and Frequency Regulators.—Since the development of thyratrons, many circuits have been devised for the regulation of voltage, speed, and frequency. The principles involved in these circuits are in general those which have been discussed. Those particularly interested in this field will find it instructive to refer to the technical literature.

12-51. Temperature Control of Ovens.—A number of thyratron circuits have been developed for controlling and stabilizing the temperature of electric furnaces or ovens. The most sensitive that has been devised uses a phototube in place of the resistance in the phase-control circuit of Fig. 12-76[1] (see also Fig. 13-19). The anode current supplies part or all of the heat. The phototube is controlled by a beam of light reflected from the mirror of a galvanometer which is energized by the current from a thermojunction in the oven. Increase of temperature deflects the beam, reducing the illumination of the phototube and increasing its resistance, and thus decreases the average anode current. With this circuit the temperature may readily be held constant within a tenth of a degree at temperatures in the vicinity of 1000°C.

[1] LA PIERRE, C. W., *Gen. Elec. Rev.*, **35**, 403 (1932); ZABEL, R. M., and HANCOX, R. R., *Rev. Sci. Instruments*, **5**, 28 (1934).

12-52. Grid-controlled Rectifiers.[1]—Grid-controlled rectifiers are advantageously used in the conversion of alternating into direct current. Among the advantages are the ease with which the direct voltage can be controlled by the use of phase-control circuits, the rapidity with which the current can be interrupted in case of overload or short circuit, and the possibility of using inverter action to feed back to the a-c line energy stored in filter condensers and chokes which would otherwise have to be dissipated in an arc when the circuit is interrupted.[2] The use of grids also reduces the danger of arcback.

12-53. Welding Control.[3]—In electric welding it has been found to be advantageous to use a series of equally spaced spot welds, rather than

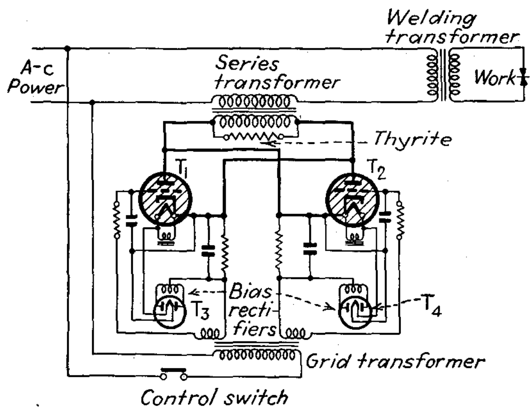


FIG. 12-89.—Thyratron welding control circuit.

a continuous weld. In order to accomplish this, the welder must be switched on and off periodically. Because of the heavy currents, mechanical switches are unsatisfactory. The value of the thyratron in this service arises from the ability of the grid to control large currents without in itself requiring large current, from the consistency with which the grid can control the welding time, and from the absence of large current transients when tube control is used.

Figure 12-89 shows one form of thyratron switching circuit for welding control.[4] The operation is in some respects similar to that of the circuit of Fig. 12-83. The thyratron grids are biased negatively by the auxiliary

[1] BROWN, H. D., *Gen. Elec. Rev.*, **35**, 439 (1932); FOOS, C. B., *Elec. Eng.*, **35**, 568 (1934); KIME, R. M., *Electronics*, August, 1933, p. 219; JOURNEAUX, D., *Elec. Eng.*, **53**, 976 (1934); DURAND, S. R., and KELLER, O., *Proc. I.R.E.*, **25**, 570 (1937).

[2] DURAND and KELLER, *ibid.*

[3] GRIFFITH, R. C., *Gen. Elec. Rev.*, **33**, 511 (1930); MARTIN, S., JR., *Welding*, **3**, 293, 361 (1932); LORD, H. W., and LIVINGSTON, O. W., *Electronics*, July, 1933, p. 186; CHAMBERS, D. E., *Elec. Eng.*, **54**, 82 (1935;) PALMER, H. L., *Gen. Elec. Rev.*, **40**, 229, 321 (1937); CLARK, S. A., *Electronics*, August, 1942, p. 36, September, 1942, p. 55, October, 1942, p. 62, November, 1942, p. 65. See also footnote 3, p. 527.

[4] CHAMBERS, D. E., *ibid.;* PALMER, H. L., *ibid.*

rectifier tubes, so that no anode current flows when the control switch is open. The current through the welding transformer then cannot exceed the exciting current of the series transformer. When the switch is closed, there is applied to the thyratron grids an alternating voltage that is in phase with the anode voltage and thus causes the flow of full average anode current. When the thyratrons conduct, the effective primary

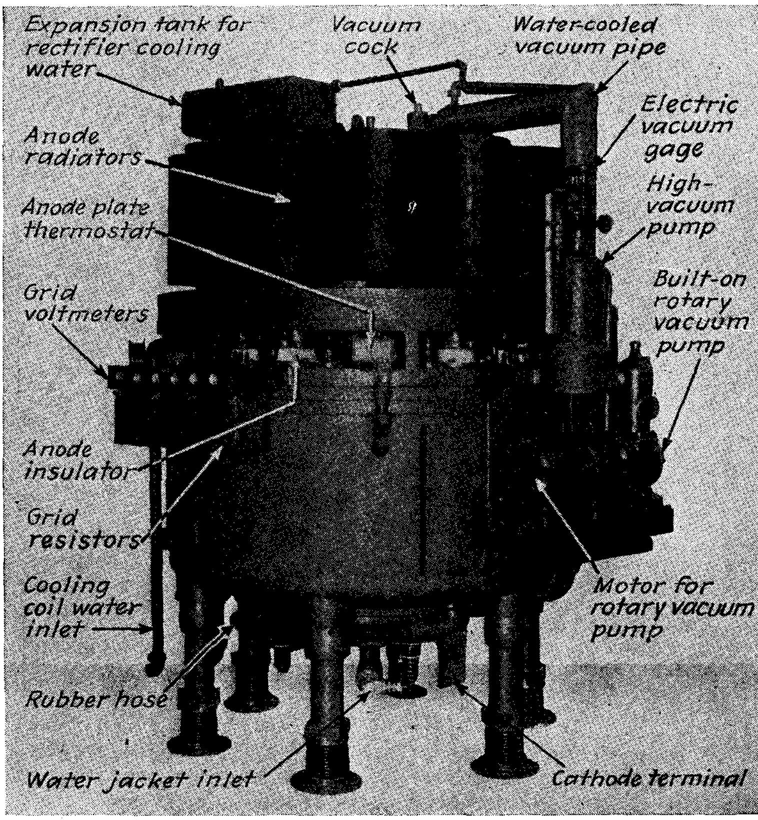


FIG. 12-90.—Typical grid-controlled multi-anode tank rectifier. This rectifier is approximately 9 ft. high and can handle 2750 kw at 625 volts. (*Courtesy of Allis-Chalmers Manufacturing Co.*)

impedance of the series transformer is reduced to a very low value, and practically full line voltage is applied to the primary of the welding transformer. The thyrite resistor across the primary of the series transformer prevents dangerous surges of voltage. The control switch of the circuit of Fig. 12-89 is usually replaced by an intermediate control circuit, which is in turn operated by the voltage from a thyratron relaxation oscillator of the type discussed in Sec. 12-33. Adjustment of the circuit constants of this oscillator and of the bias applied to the intermediate control circuit makes it possible to vary the frequency of the welding cycle and the portion of the cycle during which welding current flows.

Because the oscillator is controlled by 60-cycle synchronizing voltage (see Secs. 10-14 and 10-15), the timing is very accurate.

In another type of circuit the welding time is limited to half a cycle or less.[1] The portion of the cycle during which current flows is controlled by the use of phase control.

12-54. Relay Control.—A promising field of application of thyratrons, either alone or in combination with high-vacuum tubes, is in the operation of relays as the result of changes of relative magnitude or phase of two or more voltages or currents.[2]

12-55. Mercury Pool Arc-discharge Tubes.—Mercury pool arc-discharge tubes may, in general, be used in the circuits and applications treated under hot-cathode arc tubes. Mercury pool tubes are much more rugged than hot-cathode arc tubes. The fact that the cathode need not be heated results in four other advantages over the thyratron: (1) the over-all efficiency is higher, (2) the tube is instantly available for operation, (3) freedom from danger of cathode deactivation or destruction gives increased reliability in industrial service, and (4) the anode current is not limited by available cathode thermionic emission current, but only by heating of the anode or anodes and of the tube as a whole. By the use of water cooling, very high current capacity may be attained. Mercury pool tubes are designed for types of service requiring currents measured in tens and hundreds or even, instantaneously, thousands of amperes.

As rectifiers, mercury pool tubes have been used for many years, first in glass envelopes, later in metal tanks continuously evacuated, and finally, in small sizes, in sealed metal envelopes. For a detailed discussion of constructional and operational features of these rectifier tubes the reader should refer to books on the subject of commercial power rectification.[3]

Grid-controlled mercury pool rectifiers provided with keep-alive anodes are used as power rectifiers, phase control affording a simple method of varying or regulating the output voltage (see Sec. 14-3). The grids, which surround the anodes, also serve to prevent mercury from depositing on the anode and causing arcback. A typical water-cooled high-power, grid-controlled tank rectifier is shown in Figs. 12-90 and 12-91. Tubes of this type have relatively long deionizing time and are not convenient for use in the control of small amounts of power.

[1] GRIFFITH, *loc. cit.;* LORD and LIVINGSTON, *loc. cit.*

[2] WIDEROË, R., *Trans. Am. Inst. Elec. Eng.*, **53**, 1346 (1934).

[3] PRINCE, D. C., and VOGDES, F. B., "Principles of Mercury Arc Rectifiers and Their Circuits," McGraw-Hill Book Company, Inc., New York, 1927; MARTI, O. K., and WINOGRAD, H., "Mercury Arc Rectifiers—Theory and Practice," McGraw-Hill Book Company, Inc., New York, 1930; JOLLEY, L. B. W., "Alternating Current Rectification," John Wiley & Sons, Inc., New York, 1931. See also bibliography at the end of Chap. 14.

The field of application of mercury pool arcs has been greatly extended by the development of igniter-controlled tubes.[1]

12-56. Comparison of Igniter and Grid Control.—There are two fundamental differences between the hot-cathode thyratron and the cold-cathode ignitron (see Sec. 11-20). The first is the mechanism of electron emission at the cathode. In the cold-cathode arc, emission probably results from the formation by space charge of very high fields at the surface of the cathode, rather than from thermionic emission, as

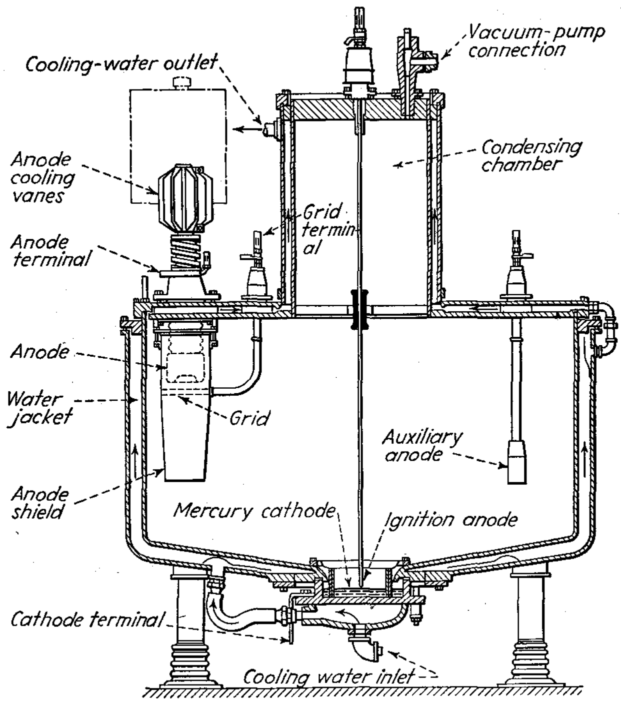


FIG. 12-91.—Cross section through typical Allis Chalmers mercury-arc power rectifier of the multianode type. The main anode, at the left, is shown air-cooled. The broken line indicates an optional radiator for water-cooling.

in the hot-cathode arc. The second difference, the method of controlling the formation of the arc, is in reality the result of the first. In a hot-cathode arc tube the continuous ample emission of electrons at the cathode makes it possible for the arc to form when the anode voltage exceeds the ionization potential of the gas or vapor. The function of the grid in a thyratron is to prevent formation of the arc. Before the grid can regain control after arc extinction, it is necessary that deionization shall be so complete that the arc cannot form between the grid and the cathode and that the grid is not surrounded by an ion sheath which

[1] SLEPIAN, J., and LUDWIG, L. R., *Trans. Am. Inst. Elec. Eng.*, **52**, 693 (1933); KNOWLES, D. D., *Electronics*, June, 1933, p. 164.

prevents it from holding back electrons emitted from the cathode. This may require as much as 100 μsec. In the cold-cathode arc tube, on the other hand, the emission mechanism is such that high current density and high ion density in the vicinity of the cathode are necessary for its maintenance. Within a few microseconds after extinction, the deionization in the vicinity of the cathode is great enough so that the required

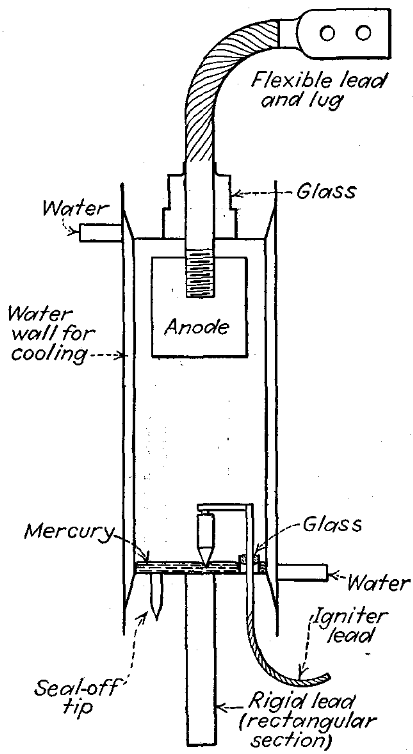


FIG. 12-92.—Construction of the type WL-651 water-cooled metal ignitron. (*Courtesy of Westinghouse Electric and Manufacturing Co.*)

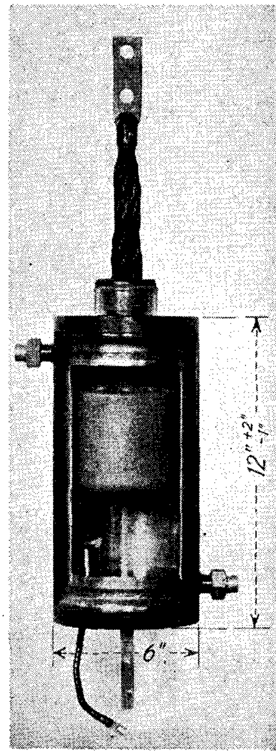


FIG. 12-93.—Cut-away view of type FG-258A water-cooled metal ignitron. (*Courtesy of General Electric Company.*)

emission is not reestablished by reapplication of anode voltage.[1] Furthermore, even if the anode voltage is high enough to cause a glow, breakdown into an arc occurs only rarely and at random intervals. The function of the control electrode in a mercury pool tube is not, therefore, to prevent formation of the arc, but to initiate the arc. This can be done by producing ionization of density comparable with that caused by the arc. In the ignitron the required ion density is obtained by means of an auxiliary rod, the igniter, which touches the surface of the mercury pool. A heavy current from igniter to pool causes the formation of an

[1] KNOWLES, D. D., and BANGRATZ, E. G., *Elec. J.*, **30**, 501 (1933).

auxiliary arc which is followed within a microsecond by the formation of the main arc.[1]

12-57. Ignitron Construction.—A typical air-cooled glass ignitron, the KU-637, was shown in Fig. 11-14. The use of metal construction and water cooling has resulted in tubes of greatly reduced size and with smaller probability of arcback. It has been found that tendency toward arcback in mercury pool tubes is favored by the positive-ion *back* current that flows during the negative half cycle before deionization is complete.[2] Since the deionizing time is reduced by lowering the temperature (see Sec. 11-22), the efficient cooling attainable in water-jacketed metal ignitrons is conducive to freedom from arcback.

The important features of construction of water-cooled ignitrons may be discerned from Figs. 12-92 and 12-93, which show a drawing of the type WL-651 ignitron and a cutaway view of the type FG-258A ignitron. The shell is not insulated from the mercury cathode and is, therefore, at cathode potential. The igniter and anode leads are insulated from the envelope by means of glass bushings. If the conducting time of the tube exceeds the time in which the spot can move to the edge of the pool, it may anchor to the envelope and cause material to be sputtered from the metal wall. For this reason, some ignitrons are provided with a *spot fixer*.[3] This is a metal strip at the surface of the mercury pool, which prevents the cathode spot from moving about.

Choice of material for the igniter rod is an important consideration in the design of ignitrons. It has been found that a semiconductor, such as boron carbide or silicon carbide, is satisfactory. The compound is finely ground and formed into a rod with the aid of a binder. An irregular surface, resulting from the projection of numerous sharp edges or points of crystals, is essential. In order to limit the energy required to initiate the arc, the rod diameter and the shape of the end of the rod must be properly chosen.[4] The starting current increases with rod diameter.

12-58. Comparison of Thyratron and Ignitron.—Because of the great rapidity with which control is reestablished, the certainty of control is greater for the ignitron than for the thyratron. Arcbacks may occur occasionally, but it has been shown that arcback is essentially a random phenomenon and that the probability of its occurrence can be reduced to as low as one arcback in a month of use.[5] Other advantages of the

[1] DOW, W. G., and POWERS, W. H., *Elec. Eng.*, **54,** 942 (1935).

[2] LUDWIG, L. R., MAXFIELD, F. A., and TOEPFER, A. H., *Elec. Eng.*, **53,** 75 (1943).

[3] WAGNER, C. F., and LUDWIG, L. R., *Elec. Eng.*, **53,** 1384 (1934); TONKS, L., *Physics*, **6,** 294 (1935).

[4] SLEPIAN, J., and LUDWIG, L. R., *Trans. Am. Inst. Elec. Eng.*, **52,** 693 (1933); CAGE, J. M., *Gen. Elec. Rev.*, **38,** 464 (1935); TOEPFER, A. H., *Elec. Eng.*, **56,** 810 (1937).

[5] SLEPIAN, J., and LUDWIG, L. R., *Trans. Am. Inst. Elec. Eng.*, **51,** 92 (1932). See also D. C. PRINCE, *J. Am. Inst. Elec. Eng.*, **46,** 667 (1927).

ignitron were discussed in Sec. 12-55. A serious disadvantage of the ignitron is the necessity for a large pulse of energy to initiate the arc. An igniter current of the order of several amperes for a time of about 100 μsec may be required. During this time the anode voltage must remain about 25 volts above the normal tube drop of 10 volts in order that the tube shall fire. In contrast to the hot-cathode arc tube, which will operate at small anode currents, the mercury pool type of arc tube requires an anode current of the order of 1 or 2 amp in order to maintain the arc. This is sometimes a disadvantage but may, on the other hand, assist in the extinction of the arc in d-c operation. The maximum peak voltages at which ignitrons can be operated are in general lower than those of thyratrons. Operating data for typical ignitrons are listed on p. 688.

12-59. Ignitron Control Circuits.—The methods used in stopping the anode current of thyratrons are applicable to ignitrons. The arc

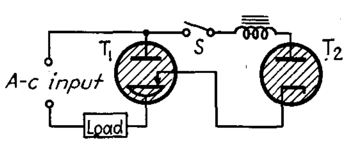


FIG. 12-94.—Ignitron firing circuit.

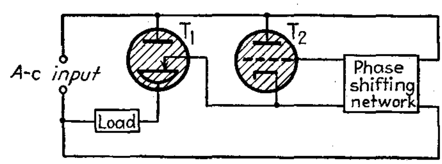


FIG. 12-95.—Ignitron phase-control circuit incorporating a thyratron.

may be extinguished by reducing the current for an instant below the value necessary to maintain the cathode spot. Because of the relatively high energy required to fire the ignitron, the simple phase-control circuits discussed earlier in the chapter cannot be applied directly to ignitron control. A phase shifter capable of supplying the necessary power may be used. Several circuits have been developed that allow the tube to be fired by energy supplied by the line.[1]

Figure 12-94 shows the basic form of one type of ignitron firing circuit. The rectifier tube T_2 serves the dual function of preventing the flow of igniter current during the negative half of the cycle of applied voltage and of interrupting the igniter current after the anode fires in the positive half cycle. After the anode fires, the voltage applied to the igniter circuit is the ignitron tube drop, which is approximately equal to the rectifier tube drop. The igniter current therefore falls to zero. Substitution of a phase-controlled thyratron in place of the rectifier, as in Fig. 12-95, makes possible the control of the time in the cycle at which the ignitron is fired.

In the circuits of Figs. 12-94 and 12-95 the igniter current flows through the load. For this reason there is likely to be some variation of the time in the cycle at which the anode fires. This difficulty is

[1] KNOWLES and BANGRATZ, *loc. cit.*

avoided in the circuit of Fig. 12-96, in which the condenser C is charged through the rectifier T_3 and discharged through the thyratron T_2 at a time determined by the phase-shifting network. Oscillograms show that with this circuit the ignitron fires with regularity at a predetermined point of the cycle of alternating supply voltage.[1] In order that the

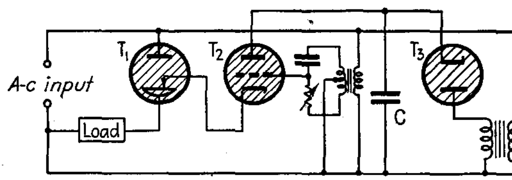


FIG. 12-96.—Phase-control circuit in which the ignitron is fired by energy stored in the condenser C.

thyratron T_2 will extinguish, the rectifier circuit is designed so that the charging current of the condenser is less than the current required to maintain the arc of T_2.

Because of the high peak currents to which the thyratrons are subjected in the circuits of Figs. 12-95 and 12-96, particularly if anode

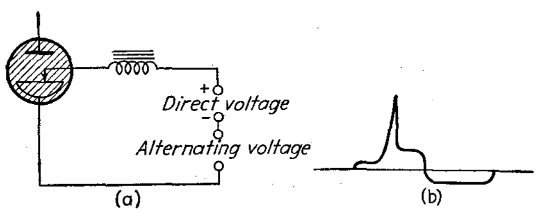


FIG. 12-97.—(a) Ignitron firing circuit. (b) Wave form of ignitor current.

breakdown fails to occur, the life of the thyratrons is relatively short. The need for thyratrons is avoided in the circuit of Fig. 12-97a, in which direct and alternating voltages are impressed simultaneously in the igniter circuit, which contains an iron-core choke.[1] As the result of core saturation, a high pulse of igniter current flows during the positive crest of impressed alternating voltage, as shown in Fig. 12-97b. The average igniter current and the reverse current during the negative half cycle are, however, small. Phase control may be obtained by shifting the phase of the impressed alternating igniter voltage relative to the anode voltage.

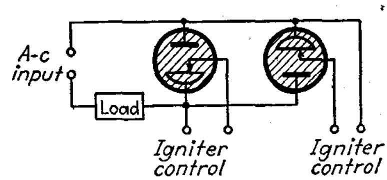


FIG. 12-98.—Full-wave ignitron circuit.

The circuits of Figs. 12-95 to 12-97 can be readily converted into full-wave circuits. Figure 12-98 shows a full-wave ignitron circuit that

[1] KLEMPERER, HANS, *Electronics*, December, 1939, p. 12.

permits current to start flowing at a predetermined point in each half cycle.[1]

Ignitron Applications.—The applications of the ignitron are in general the same as those of the thyratron. They include spot-welding control, inversion and frequency transformation, illumination control, and motor commutation.[2] At present spot-welding control is by far the most important application of ignitrons.

12-60. Externally Controlled Mercury Pool Arc Tube.—Breakdown of a glass mercury pool tube containing a small amount of inert gas may be initiated by means of an external electrode in contact with the glass wall. The sudden application of high voltage between the external electrode and the cathode causes the formation of a glow, with subsequent formation of an arc between the anode and the cathode. The objection to this type of control lies in the lowering of the anode-cathode glow breakdown voltage by the use of gas, but it has found application in stroboscopes.

12-61. The Strobotron.—Another type of cold-cathode arc tube has been developed by Germeshausen and Edgerton.[3] The electrode structure of this tube is illustrated in Fig. 12-99. The cathode consists of a cup containing a cesium compound that liberates free cesium under the action of the cathode spot. In this manner high thermionic emission is obtained at relatively low temperature. The cathode is surrounded by a ceramic insulator which concentrates the discharge on the active portion of the cathode and serves as a support for the inner grid. This grid is a wire-mesh screen directly above the cathode. During the operation of the tube, the inner grid becomes coated with cesium, which lowers the breakdown voltage when this grid acts as a cathode. A second or outer grid, a graphite ring, is mounted above the ceramic insulator. Graphite is used because the cesium which condenses on it does not lower the breakdown voltage between it and the cathode. The anode is about an inch above the grid, the anode support wire being insulated by means of a glass shield in order to prevent the discharge from taking the

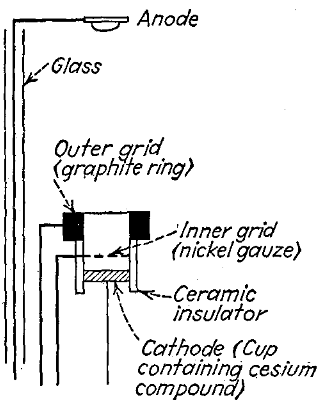


FIG. 12-99.—Strobotron electrode structure.

[1] STODDARD, R. N., *Elec. Eng.*, **53**, 1366 (1934).

[2] WAGNER, C. F., and LUDWIG, L. R., *Elec. Eng.*, **53**, 1384 (1934); LUDWIG, L. R., MAXFIELD, F. A., and TOEPFER, A. H., *Elec. Eng.*, **53**, 75 (1934); SILVERMAN, D., and COX, J. H., *Elec. Eng.*, **53**, 1380 (1934); DAWSON, J. W., *Elec. Eng.*, **55**, 1371 (1936); PACKARD, D., and HUTCHINGS, J. H., *Elec. Eng.*, **56**, 37, 875 (1937).

[3] GERMESHAUSEN, K. J., and EDGERTON, H. E., *Elec. Eng.*, **55**, 790 (1936).

shorter path to the support. One of the noble gases, usually neon or argon, is used in the tube.

Breakdown is initiated as a glow discharge between two of the electrodes, usually the two grids. If the glow current exceeds a certain value and the anode source is capable of supplying sufficient current to maintain an arc, a cathode spot immediately forms, and the discharge changes into an arc between the anode and the cathode. Because either grid may be made either positive or negative relative to the cathode, there is considerable latitude in the operating voltages that may be used.

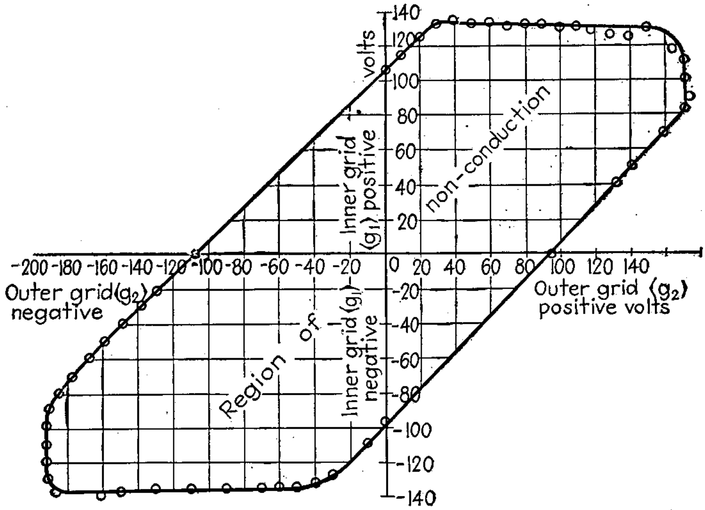


FIG. 12-100.—Strobotron control characteristic.

A typical control characteristic is shown in Fig. 12-100. By proper choice of grid voltages, the glow current required to initiate arc breakdown may be made as small as 2×10^{-9} amp.[1] The tube will carry a peak current of 200 to 300 amp and an average current of 50 ma.

Developed primarily as a stroboscopic light source, the strobotron has other applications.[2] It may, for instance, be used to fire an externally-controlled mercury pool arc tube. Its advantages as a stroboscopic light source result in part from the large peak current that can be carried without damage to the cathode.

12-62. Stroboscopes.—An important industrial and scientific application of the arc tube is as a stroboscope. If an intermittent source of light is synchronized to the frequency of rotation or vibration of a

[1] WHITE, A. B., NOTTINGHAM, W. B., EDGERTON, H. E., and GERMESHAUSEN, K. J., *Electronics*, March, 1937, p. 18.

[2] GERMESHAUSEN, K. J., and EDGERTON, H. E., *Electronics*, February, 1937, p. 12; GRAY, T. S., and NOTTINGHAM, W. B., *Rev. Sci. Instruments*, **8**, 65 (1937).

mechanical device or to a multiple or submultiple of the frequency, the moving object will appear to be stationary when examined by the light. The requirements of such a light source are (1) easily controlled frequency of illumination, to allow synchronization of light and motion over a large frequency range, (2) extremely short duration of light flash, so that the moving part under observation will not appear to move while it is illuminated, and (3) light of great brightness, so that the moving part will be clearly seen despite the short time of illumination. By acting as a shutter of extremely high speed, the stroboscopic light source also makes possible the photographing of rapidly moving objects.

One type of stroboscopic tube, the strobotron, has already been discussed. Where small illumination suffices, thyratrons may be used as the light source. Ordinary thyratrons are not designed to be efficient light sources, however, and the production of intense illumination necessitates the use of such large peak currents that the life of a hot-cathode tube is shortened. Mercury pool tubes are therefore generally used in preference to thyratrons. The tubes may be constricted where the discharge takes place. The constriction not only concentrates the light-emitting region, but the proximity of the walls also speeds deionization of the gas and thus prevents serious afterglow.

To produce a sharply defined flash of light, a condenser is discharged through the stroboscopic tube. Discharge may be initiated by a voltage impulse applied to an external electrode or an internal grid or by firing an auxiliary thyratron which is in series with the tube and a source of limited current, so connected that the condenser discharges only through the main tube. The inert gas which is essential to the use of external arc initiation is not objectionable in stroboscopic tubes because the relatively small diameter and large electrode spacing prevent the glow breakdown voltage from being too low. The complete circuit of a stroboscope is complicated by the necessity of tripping the tube by an impulse of steep wave front in order to ensure accurate adjustment and synchronization, especially in the observation of rapidly moving objects.

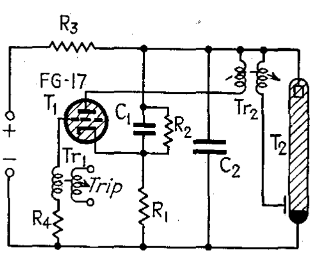


FIG. 12-101.—Stroboscope circuit in which the discharge of a condenser through a cold-cathode arc tube is initiated by a thyratron.

Figure 12-101 shows a typical circuit in which a thyratron controls the tripping impulse applied to the external grid of a mercury pool arc tube.[1] The large condenser C_2 is charged through the resistance R_3,

[1] EDGERTON, H. E., and GERMESHAUSEN, K. J., *Rev. Sci. Instruments*, **3**, 535 (1932). See also H. E. EDGERTON, *Elec. Eng.*, **50**, 327 (1931), *Electronics*, July, 1932 p. 220; R. C. HITCHCOCK, *Elec. J.*, **32**, 529 (1935).

and the condenser C_1 is charged through R_1 and R_3. Application of a tripping impulse to the primary of the transformer Tr_1 fires the thyratron T_1, which discharges the condenser C_1 through the primary of the step-up transformer Tr_2. The induced voltage in the secondary of Tr_2 is applied between the anode and the external grid of the mercury pool tube T_2, causing the latter to fire. Since the impedance between C_2 and the mercury pool tube is very low, the discharge current is high and of extremely short duration. The light emitted by T_2 is therefore intense

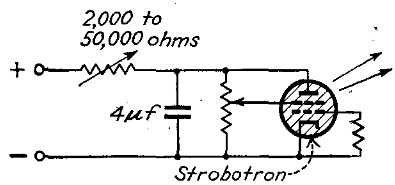


FIG. 12-102.—Self-excited strobotron stroboscope.

and of short duration. Figure 12-102 shows a simple self-excited strobotron circuit which is essentially an arc-tube relaxation oscillator of the type discussed in Sec. 12-33.[1] Figure 12-103 shows a modification of the circuit of Fig. 12-102 in which the firing of the strobotron is controlled by the output voltage of a multivibrator.[1]

Because of the very short duration of the condenser discharge in the circuits of Figs. 12-43 and 12-64, these circuits may be used as the basis of a simple and reliable stroboscope. This is accomplished by connecting the light source across the secondary of an induction coil, the primary

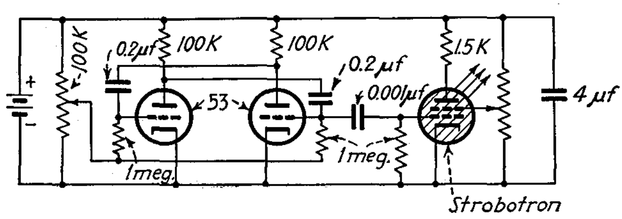


FIG. 12-103.—Strobotron stroboscope controlled by a multivibrator.

of which shunts L_1 and which contains no vibrator. The circuit of Fig. 12-64 is self-exciting but may be synchronized to the frequency of the system under observation by means of voltage applied to the grid circuit through a transformer in series with R_c. The circuit of Fig. 12-43, which must be tripped by a voltage pulse synchronized to the frequency of the object under observation, may be used over a wide frequency range without readjustment of circuit constants. The synchronizing voltage may be applied through a transformer in series with C_2, C_3 being

[1] GERMEHAUSEN and EDGERTON, *Elec. Eng.*, **55**, 790 (1936).

omitted. In either circuit, only a resistance need be used as "load" in series with the supply voltage. An Edison-type neon lamp, with the ballast resistance removed from the base, or, preferably, a small neon sign, makes a suitable light source. The flash is of longer duration than that obtained with a strobotron or a mercury pool tube, but excellent results are obtained in many applications. Both the duration of the flash and the intensity are decreased with decrease of L_1. For an FG-67 thyratron, operated from a 110-volt direct-current supply, C_1 should have a capacitance of from 20 to 40 μf, and L_1 an inductance of from 0.2 to 0.7 mh. L_1 should have as low resistance as possible.

Problems

12-1. *a.* Design a circuit in which a glow-tube oscillator causes the periodic opening and closing of a relay. The frequency of operation and the fraction of the cycle during which the relay is energized are to be adjustable. All voltages except for cathode heating should be derived from a single 220-volt d-c supply. The relay closes on 7 ma and opens on 6 ma. Specify all tubes, circuit constants, and operating voltages.

b. By constructing a curve of plate current *vs.* time, on which the critical relay currents are shown, prove that the percentage of the cycle during which the relay is energized cannot be varied progressively from zero to 100 per cent.

12-2. Derive Eqs. (12-2), (12-6), and (12-7).

12-3. By means of a circle vector diagram, determine which of the impedances of Fig. 12-76 should be a resistance and whether the other should be inductive or capacitive in order that increase of resistance may produce progressive increase of average anode current.

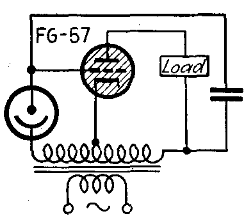


Fig. 12-104.—Circuit diagram for Prob. 12-5.

12-4. *a.* Draw waves of anode voltage *vs.* electrical degrees for sinusoidal anode voltages of 200 volts and 800 volts crest value; and, by the use of the grid-control characteristics of Fig. 12-27, construct corresponding waves of critical grid voltage for a type FG-57 thyratron.

b. Use the curves of (*a*) to derive curves of average anode current *vs.* phase angle, θ, between grid and anode voltages, for crest grid voltage equal to twice and to ten times the maximum critical grid voltage.

12-5. The operation of the phase-control circuit of Fig. 12-104 is erratic, regardless of the size of the condenser used. The phototube is found to be good, and all tube voltages are correct. Suggest a probable cause of this difficulty and a remedy.

12-6. In the course of the tests on the circuit of Fig. 12-104, the grid suddenly becomes completely and permanently ineffective. Suggest a probable cause, and state how the tube might have been protected against damage.

CHAPTER 13

LIGHT-SENSITIVE TUBES AND CELLS

In all tubes discussed in preceding chapters, the flow of current is controlled by means of voltages applied to the electrodes. There is another very important class of tubes or cells, in which the current and the output voltage are controlled by incident radiant energy and the operation of which depends upon some form of the photoelectric effect. These devices have made possible sound pictures, television, and many kinds of commercial and laboratory control and measuring equipment.

13-1. Types of Photoelectric Phenomena.—The control of current by light may be accomplished by the application of three types of photoelectric phenomena: the *photomissive* effect, or emission of electrons from metallic surfaces as the result of incident radiation; the *photoconductive* effect, or change of resistance of semiconductors by the action of incident radiation; and the *photovoltaic effect*, or production of a potential difference across the boundary between two substances that are in close contact, by illumination of the boundary. The best known of these phenomena, the photoemissive effect, appears in practical form in the *phototube* (*photocell*). The phototube is an electron tube in which one of the electrodes is irradiated for the purpose of causing electron emission.

13-2. Historical Survey.—The photoemissive effect was discovered by Heinrich Hertz in 1887. In the course of his classical experiments on electric oscillations and electric waves, Hertz noticed that the maximum length of a spark that could be made to jump across a small spark gap was increased when light from another spark was allowed to fall upon it. By means of a series of experiments he showed conclusively that the effect was caused by ultraviolet light and that it was greatest when the light fell upon the negative electrode of the spark gap. Subsequent contributions to knowledge on the photoemissive effect were made by Wiedemann and Ebert, Hallwachs, Righi, Stoletow, Elster and Geitel, J. J. Thomson, Lenard, and others.[1]

Hallwachs proved that the photoemissive effect involved the loss of negative electricity from the illuminated surface. Righi connected two electrodes to an electrometer, by means of which he observed the flow of current when the surface of one electrode was illuminated. Stoletow added an external source of voltage and measured the current

[1] See references at end of chapter.

by means of a galvanometer. Elster and Geitel predicted and proved that the alkali metals should be the most sensitive photoelectric emitters of all the metals and found that sodium and potassium respond not only to ultraviolet radiation, but also to visible light. Lenard measured the ratio of the charge to the mass of the particles that carry the photoelectric current and showed them to be the same particles as those which Thomson had previously proved to carry the current in a beam of cathode rays and which were called electrons.

13-3. Laws of Photoelectric Emission.—The early experiments resulted in the formulation of two laws relating photoelectric emission with the rate at which radiant energy strikes the emitting surface and with the wave length of the incident energy. The rate at which the radiant energy strikes is specified in terms of *radiant flux* (intensity of radiation), which is defined as the time rate of flow of radiant energy and is usually expressed in ergs per second or watts. The two laws of photoemission are

1. The number of electrons released in unit time at a photoelectric surface by radiation of constant spectral distribution is directly proportional to the incident radiant flux.

2. The maximum energy of electrons released at a photoelectric surface is independent of the amount of radiant flux incident upon the surface but varies linearly with the frequency of the radiation.

The first law was one of the factors that led to the formulation of the quantum theory of radiation. This law is of practical importance because of the desirability of a linear relation between anode current and incident flux in applications of phototubes.

The second law was explained in 1905 by Einstein, who assumed that incident radiant energy could be transferred to the electrons only in quanta of magnitude $h\nu$ and that a portion of this energy was used in removing the electrons from the emitter, the remainder appearing as kinetic energy of the emitted electrons. This is stated mathematically by Einstein's equation

$$h\nu = w + \tfrac{1}{2}mv^2 \tag{13-1}$$

in which ν is the frequency of the incident radiation, w is the electron affinity of the emitter, v is the maximum velocity of the emitted electrons, h is Planck's constant, and m is the mass of an electron. The equation holds only for the fastest electrons, which are removed from the atoms at the surface of the emitter. Some radiation penetrates through the outer layers of atoms, liberating electrons within the emitter. These electrons may lose additional energy in moving to the surface, so that their velocities after emission will be less than that indicated by Einstein's equation. It can be seen from Eq. (13-1) that there is a minimum value of ν below which the energy of the incident photon is less than the electron affinity and no emission can take place. This limiting frequency is called the *threshold frequency* and is indicated by the symbol ν_0. Substituting $v = 0$ and $\nu = \nu_0$ in Eq. (13-1) shows that

$$\nu_0 = \frac{w}{h} \tag{13-2}$$

The wave length corresponding to the threshold frequency is termed the *long-wave limit*, λ_0. Einstein's equation was verified experimentally in 1916 by Millikan,[1] who determined the maximum velocities of emitted electrons for various frequencies of incident radiation by measuring the retarding potential necessary to reduce the anode current to zero. Einstein's equation is of practical importance because it indicates that the electron affinity of the emitter must be small in order for the emitter to be sensitive to visible radiation.

13-4. Definitions.—Before proceeding farther it is necessary to list a number of definitions.[2]

Light is radiant energy evaluated according to its capacity to produce visual sensation.

Luminous flux is the time rate of flow of light.[3] It is represented by the symbol F.

The *lumen* is the unit of luminous flux. It is equal to the flux through a unit solid angle from a uniform point source of one candle, or to the flux on a unit surface, all points of which are at unit distance from a uniform point source of one candle. It follows that the total luminous flux over the surface of a sphere is equal to 4π times the number of candles of an enclosed source.

Luminous intensity of a source of light, in a given direction, is the solid-angular flux density in the direction in question. Hence, it is the luminous flux on a small surface normal to that direction, divided by the solid angle (in steradians) which the surface subtends at the source of light.

The *candle* is the unit of luminous intensity. The unit used in the United States is a specified fraction of the average horizontal candlepower of a group of 45 carbon-filament lamps (preserved at the Bureau of Standards) when the lamps are operated at specified voltages.

Candlepower is luminous intensity expressed in candles.

[1] MILLIKAN, R. A., *Phys. Rev.*, **7**, 355 (1916).

[2] See "Illuminating Engineering Nomenclature and Photometric Standards," Illuminating Engineering Society, New York, 1932.

[3] Luminous flux is radiant flux of particular energy distribution, lying entirely in the visible range.

Illumination is the density of the luminous flux on a surface; it is the quotient of the flux by the area of the surface when the latter is uniformly illuminated.

Foot-candle is the unit of illumination when the foot is taken as the unit of length. It is the illumination on a surface one square foot in area on which there is a uniformly distributed flux of one lumen, or the illumination produced at a surface all points of which are at a distance of one foot from a uniform point source of one candle. When the flux density is uniform, the illumination in foot-candles equals the flux in lumens divided by the area in square feet.

The lumen and the foot-candle are defined with respect to the response of the normal eye. Methods of measuring luminous flux and illumination that are ordinarily used, such as the photometer, also involve the response of the normal eye. Therefore the number of lumens from a source of given total radiant flux will depend upon the energy distribution of the radiation. Since the manner in which phototubes respond to radiant flux of different wave lengths is not the same as that of the normal eye, the validity of using the lumen and the foot-candle in making measurements of the sensitivity of phototubes may be questioned. In accordance with current practice, however, the lumen and the foot-candle will be used with the characteristic curves of phototubes and in the definitions of sensitivity.

13-5. Current–wave-length Characteristics.—Photoelectric tubes and cells, like the human eye, are not equally responsive to equal amounts of radiant flux of different wave lengths. For this reason, the response of a phototube to a given amount of radiant flux depends upon the manner in which the energy of the incident radiation is distributed in regard to wave length, which in turn depends upon the source of the radiation. In order to specify fully the manner in which a phototube behaves, it is necessary to indicate not only the response to radiation of particular energy distribution, but also the manner in which the current varies with wave length at constant incident energy. Tube manufacturers therefore usually furnish the characteristic curves for illumination from a tungsten-filament lamp operated at a temperature of 2870°K and, in addition, a *current–wave-length characteristic,* which is a graph showing the relation between direct anode current per unit energy of incident radiant flux, and the wave length of the incident constant radiant flux.

Current–wave-length characteristics for the alkali metals are shown in Fig. 13-1. Comparison of these curves with the dotted curve, which shows the visual sensitivity of the human eye, clearly indicates that cesium is by far the most satisfactory of the alkali metals for use with visible radiation.

Equation (13-2) suggests that the long-wave limit can be pushed farther toward the red end of the visible spectrum by reduction of the electron affinity. The electron affinity of photoelectric emitters, like that of thermionic emitters, can be lowered by the use of composite films made up of consecutive layers of different metals, of metals on

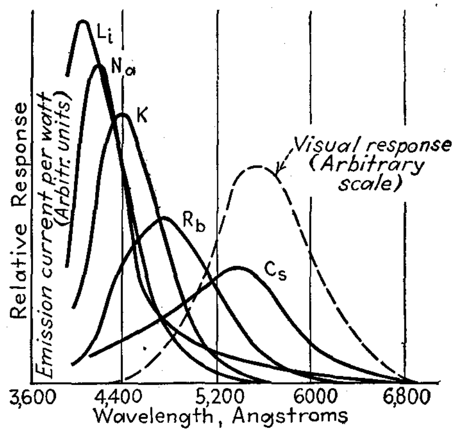


FIG. 13-1.—Current-wave-length characteristics for the alkali metals.

oxides, and of metals on a monatomic layer of oxygen. The use of composite emitters results not only in response at much longer wave lengths, but also in a marked increase of sensitivity over the whole visible range of wave length. At present the most satisfactory type of cathode for use with visible radiation consists of a thin layer of cesium deposited upon a layer of cesium oxide formed upon a silver surface. The sensitivity of cesium oxide tubes extends well into the infrared band of radiation. This is shown, in Fig. 13-2, by the current-wave-length characteristic of a typical cesium oxide tube, the type 868. The sharp cutoff at the ultraviolet end of the spectrum is caused by absorption by the glass envelope of the shorter wave lengths of the radiation. High sensitivity in the ultraviolet region may be obtained

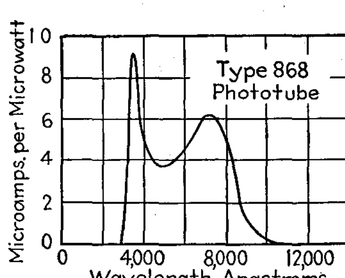


FIG. 13-2.—Current-wave-length characteristic for the type 868 phototube.

by use of a pure sodium cathode enclosed in an envelope of quartz or special glass such as *nonex*, which transmits ultraviolet radiation. By the proper choice of cathode materials, the use of color filters, and the combination of two or more tubes of different color response, it is possible to adjust the resultant characteristic curve to meet special requirements. The double peak in the characteristic of Fig. 13-2 may probably be explained by the fact that emission may take place from more than one substance in the composite film.

13-6. Types of Phototubes.—The electrodes and envelopes of phototubes may assume a variety of forms, depending upon the purposes for which the tubes are designed. In the most common form of tube, illustrated in Fig. 13-3, the cathode consists of a cylinder of silver or silver-plated copper covered with a composite emitting film. The shape of the anode is not critical, the most important requirement being that it should intercept as little light as possible. A common form of anode in tubes with cylindrical cathodes is a wire located at the axis of the cylinder, as shown in Fig. 13-3. In such a tube the initial velocities of only relatively few electrons are such as to carry them to the anode. Voltage must, therefore, be applied in order to draw all emitted electrons to the anode.

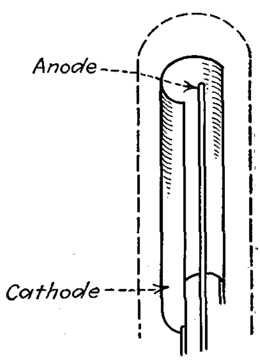


FIG. 13-3.—Electrode structure of one type of phototube.

Phototubes are classified as *vacuum phototubes* and *gas phototubes*. A *vacuum phototube* is one that is evacuated to such a degree that its electrical characteristics are essentially unaffected by gaseous ionization. A *gas phototube* is one into which a quantity of gas has been introduced, usually for the purpose of increasing its sensitivity.

13-7. Characteristics of Vacuum Phototubes. Static current-voltage curves for a typical cesium oxide vacuum phototube, the 917, are shown in Fig. 13-4, and static curves of anode current *vs.* luminous flux in Fig. 13-5. The solid curve of Fig. 13-5 applies to all anode voltages above saturation; the dotted curves, which are ordinarily of no great practical interest, are for anode voltages below saturation. Curvature of the dotted curves is probably the result of space charge. It should be emphasized that the curves of Figs. 13-4 and 13-5 apply only to light from a tungsten filament at a color temperature of 2870°K. Curves obtained with radiation of different energy distribution are of similar form, but the current range may be different. Such curves may be derived by means of the following procedure:

1. Multiply ordinates of the current–wave-length characteristic by the corresponding ordinates of the curve of Fig. 13-6, which shows the relative energy distribution of radiation from a tungsten filament at 2870°K, and plot these products as a function of wave length.

2. Plot a similar curve of the products of ordinates of the current–wave-length characteristic and those of the energy-distribution curve for the given source.

3. Multiply the current scale of the current-voltage characteristics by the ratio of the area under the second product curve to the area under the first

product curve. For approximate results, the ratio of the sums of 15 or 20 uniformly spaced ordinates of the two product curves may be used in place of the ratio of the areas, and so the product curves need not be constructed.

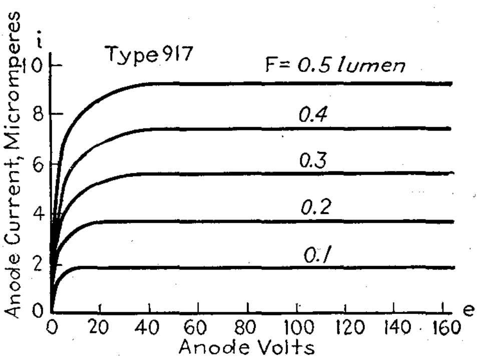


FIG. 13-4.—Current-voltage characteristics for the type 917 vacuum phototube.

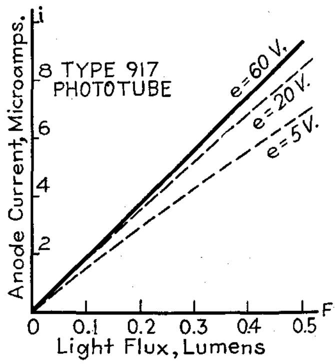


FIG. 13-5.—Static curves of anode current *vs.* luminous flux for the type 917 vacuum phototube.

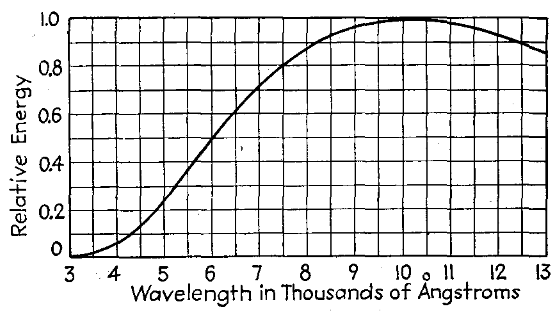


FIG. 13-6.—Relative energy distribution for radiation from a tungsten filament at a color temperature of 2870°K. (*From data by W. E. Forsythe and E. Q. Adams, Denison University Bulletin, J. of the Scientific Laboratories,* **32,** 70 (1937).)

13-7A. Phototube Sensitivity.[1]—*Static sensitivity* of a phototube is the ratio of the direct anode current at a specified steady anode voltage to the incident constant radiant flux of specified value. It is usually expressed in microamperes per microwatt. When the radiant flux is confined to a single frequency or a very narrow band of frequencies, the ratio is called the *monochromatic sensitivity.*

Luminous sensitivity of a phototube is the ratio of the direct anode current at a specified steady anode voltage to the total steady incident luminous flux in lumens. Usually the luminous sensitivity is determined by means of a tungsten-filament lamp at a color temperature of 2870°K.

[1] These definitions are essentially those given in the 1933 and 1938 reports of the Standards Committee of the Institute of Radio Engineers.

The 2870 *tungsten sensitivity* of a phototube is the ratio of the anode current at a specified steady anode voltage to the total luminous flux in lumens entering the tube from a tungsten-filament lamp operated at a color temperature of 2870°K.

Variational sensitivity of a phototube is the ratio of the change in anode current at a specified steady anode voltage to the change in the total luminous flux entering the tube. As most precisely used, the term refers to infinitesimal changes, as indicated by the defining equation

$$s = \frac{\partial i}{\partial F} \tag{13-3}$$

It is evident that the variational sensitivity is the slope of the curve of current *vs.* flux at given values of anode voltage and flux. It may also be found approximately from the family of current-voltage characteristics by the use of small increments of current and flux.

13-8. The Gas Phototube.—It can be seen from the curves of Figs. 13-4 and 13-5 that the currents that are obtained from a vacuum phototube are very small even with cathodes of high sensitivity. The sensitivity of the tube can be greatly increased by introducing a small quantity of gas. The increase of current results mainly from ionization of gas by the emitted electrons in moving to the anode. Ionization by collision takes place in a gas phototube when the anode voltage approaches the ionization potential of the gas. The new electrons thus released pass to the anode, and the positive ions are drawn to the cathode. The current is increased at every point between the cathode and anode by an amount corresponding to the number of ionizing collisions per second between the cathode and anode. Increase of anode voltage increases the number of primary electrons that make ionizing collisions. When the voltage becomes sufficiently high, then, after a primary electron makes one ionizing collision, both it and the new electron may make one or more other ionizing collisions in moving toward the anode. Some increase of current also results from secondary emission produced when positive ions strike the cathode. A tenfold increase of anode current may be attained by the use of gas, the current increasing rapidly as the voltage is raised above the ionization potential. Figures 13-7 and 13-8 show static curves of anode current *vs.* anode voltage and of anode current *vs.* luminous flux for the type 868 gas phototube. Below 15 volts the current-voltage curves are of the same general form as those of a vacuum phototube, since little ionization takes place below this voltage.

The gas that is used in phototubes must meet a number of requirements. It must not react with the electrodes or be absorbed by the electrodes or the tube walls, it should have a low ionization potential, and it should have a low molecular weight in order that the positive ions may be quickly accelerated. The only gases that do not react with the alkali metals used in the cathodes are the inert gases, helium, neon, argon, krypton, and xenon. The ionization potential of these gases decreases in the order in which they are listed, whereas the atomic weight increases. Argon gives a satisfactory compromise as regards these

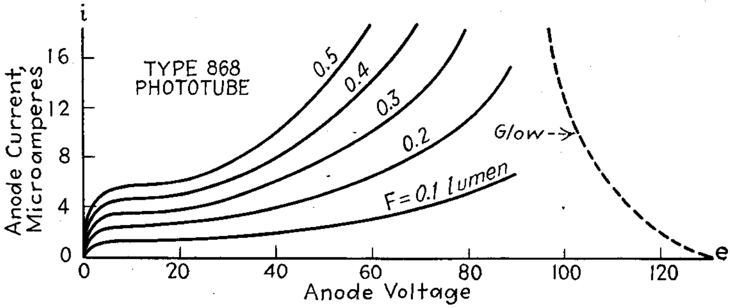


FIG. 13-7.—Current-voltage characteristics for the type 868 gas phototube.

factors and is relatively inexpensive. Neon and helium are used in some tubes, but the use of argon is more common.

The ratio of the sensitivity of a gas phototube at a given voltage and flux to the sensitivity obtainable at the same voltage and flux without ionization resulting from presence of gas is called the *gas amplification factor*. This factor depends upon the gas pressure. If the pressure is too low, the spacing between the gas molecules is relatively so great that there is slight probability that an emitted electron will strike a gas molecule in moving from the cathode to the anode. If the pressure is too high, on the other hand, the probability is great that an electron will be stopped by a gas molecule before it has been accelerated far enough to have acquired sufficient energy to cause ionization. As the pressure is increased from a very low value, the number of collisions made by an electron in moving from the cathode to the anode increases, but the kinetic energy acquired between collisions with molecules decreases. As the result of these opposing effects, there is an optimum gas pressure, which is found to be about 0.2 mm Hg in the noble gases.

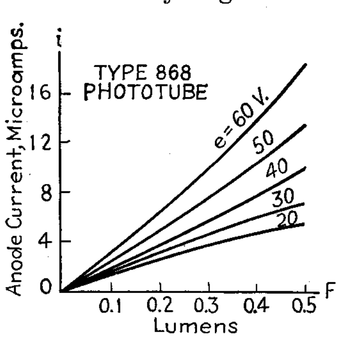


FIG. 13-8.—Static curves of anode current *vs.* luminous flux for the type 868 gas phototube.

The gain in sensitivity resulting from the use of gas is partly offset by a number of disadvantages. One, shown by Fig. 13-8, is the loss of linearity between anode current and flux. A second disadvantage is that a glow discharge may take place at voltages above about 90. The breakdown voltage falls with increase of flux, in a manner indicated roughly by the dotted curve in Fig. 13-7. Glow is objectionable not only because it causes the current to be independent of illumination, but also because it may permanently reduce the sensitivity of the tube. The anode voltage must, therefore, be kept below the value at which glow breakdown can occur. Limitation of supply voltage to 90 volts makes the maximum output voltage that can be obtained less than that attainable with vacuum phototubes. To prevent destructive rise of anode current in the event that glow breakdown does occur, a resistance of 100,000 ohms or more must always be used in series with a gas phototube. Since a high resistance is also necessary in order to produce high voltage output, the need of such a resistance is not a disadvantage.

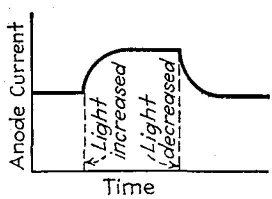


FIG. 13-9.—Lag in response of a gas phototube.

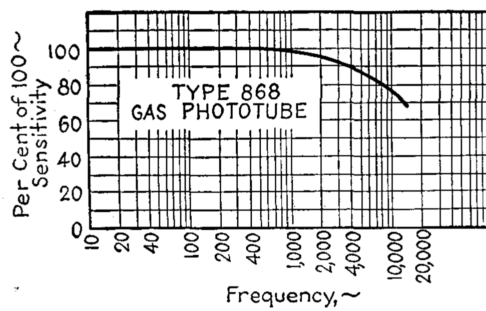


FIG. 13-10.—Dynamic response curve for type 868 gas phototube.

The third disadvantage resulting from the use of gas in phototubes is caused by the relatively large mass of the positive ions. Because the time taken for a positive ion to pass from the vicinity of the anode to the cathode is appreciable, there is a perceptible time lag in the response to a change in flux. The manner in which a gas phototube responds to an abrupt increase and decrease of flux is shown in Fig. 13-9. The effect of the time lag is similar to that of inductance in the phototube circuit, as it tends to prevent changes of anode current and produces a phase difference between sinusoidal periodic light fluctuations and the resulting alternating component of anode current. It also causes the variational sensitivity of the tube to decrease with increase of modulation frequency of the light.[1] This is illustrated by the *dynamic response curve* of Fig. 13-10, which shows the relative variational sensitivity of a type 868 gas phototube as a function of modulation frequency. When a gas phototube is used in conjunction with an amplifier to convert changes of light into

[1] Modulation frequency, the frequency at which the illumination varies, should not be confused with the radiation frequency of the incident light.

sound, the effect of a drooping dynamic response curve can be offset by the use of an amplifier that has a rising frequency characteristic.

13-9. Phototube Circuits.—The standard phototube symbol is shown in Fig. 13-11. Phototube currents are so small that a galvanometer is the only current-operated device that they can operate directly. One or more stages of amplification must be used in practical applications of the phototube. Since vacuum-tube amplifiers are voltage-operated devices, the changes of phototube current must be converted into voltage changes by means of impedance in series with the tube. Because the currents are very small, the impedance must be high, usually from 1 to 25 megohms. Although transformers have been designed for use with phototubes, the difficulty of obtaining adequate primary reactance at the lower audio frequencies makes it simpler and cheaper to use resistance or resistance-capacitance coupling between the phototube and the amplifier.

In order to simplify diagrams of circuits discussed in this chapter, all tubes are shown as triodes, and all direct voltage sources as batteries.

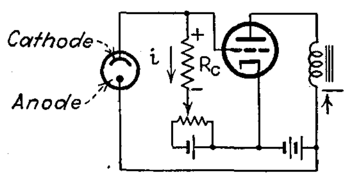


FIG. 13-11.—Forward d-c phototube circuit.

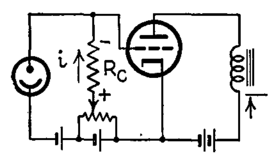


FIG. 13-12.—Reverse d-c phototube circuit.

Tetrodes and pentodes of proper characteristics (Sec. 13-14) may be used in place of triodes, and batteries may be replaced by other types of power supplies.

When a relay or other current-operated electrical device is to be controlled by changes of steady or average flux, a direct-coupled amplifier must be used. A circuit in which increase of illumination causes an increase in plate current of the amplifier tube is termed a *forward circuit;* one in which increase of illumination causes a decrease of plate current is called a *reverse circuit.* Figure 13-11 shows a simple forward circuit in which increase of illumination causes a relay to be energized. Figure 13-12 shows a similar reverse circuit. It can be seen that, for the same tubes, the forward circuit requires less total supply voltage than the reverse circuit.

By the use of a voltage divider, the various B and C batteries of Figs. 13-11 and 13-12 may be replaced by one voltage source, as in Fig. 13-13. Degenerative feedback resulting from the flow of plate current through R_1 reduces the sensitivity of the circuit. Because the amplifier must respond to changes of direct voltage, a by-pass condenser does not remedy this difficulty and R_1 and R_2 should be no larger than necessary to limit the dissipation in these resistances to a reasonable value. Reduction of sensitivity as the result of degenerative feedback also makes it inadvisable to use cathode self-biasing resistors in the amplifier stages of phototube circuits designed to respond to changes of steady illumination. If the resistances R_1 and R_2 are of proper size to give the correct filament or heater current, the filament or heater may be inserted between these resistances. Only a single power supply is then required. If the relay is shunted by a condenser in order to by-pass the alternating component of plate current, or if a slow-acting relay is employed, the circuit of Fig. 13-13 can also be used on an alternating voltage supply. The phototube and amplifier then pass current during only one-half of the cycle.

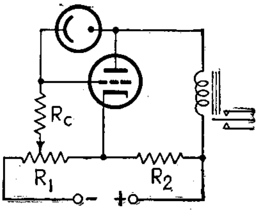


FIG. 13-13.—Forward phototube circuit using a single source of grid and plate voltage.

Improved forward and reverse a-c–operated circuits are shown in Figs. 13-14 and 13-15.[1] The purpose of the grid condensers is to eliminate the difference in phase between grid and anode voltages resulting from the capacitance of the phototube and amplifier electrodes. The circuit of Fig. 13-15 also functions if the polarity of the phototube is reversed. The grid condenser then charges during the half cycle in which the

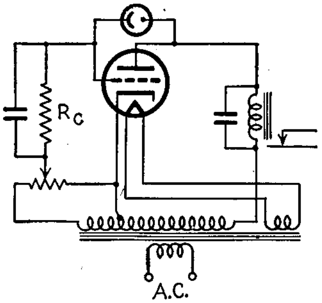


FIG. 13-14.—A-c–operated forward phototube circuit.

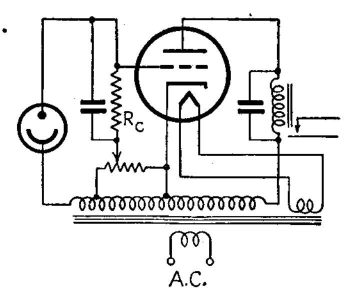


FIG. 13-15.—A-c–operated reverse phototube circuit.

amplifier plate is negative and discharges through R_c during the positive half cycle. The average plate current of the amplifier tube depends upon the average condenser voltage, which in turn depends upon the charging current and therefore the illumination. The sensitivity of these a-c circuits is less than half that of the corresponding d-c circuits at the same voltages. The illumination at which the relay closes is adjusted in

[1] Although the phototube is shunted between the amplifier grid and plate in Figs. 13-13 and 13-14, the only actual difference in the phototube connection from that in the circuit of Fig. 13-11 is that the phototube current flows through the relay. The voltage drop through the relay is small and so does not affect the operation.

the circuits of Figs. 13-11 to 13-15 by variation of grid bias of the amplifier tube.

Another a-c–operated circuit is that of Fig. 13-16. The action is as follows: During the half of the cycle in which the anode and the amplifier plate are negative, a positive voltage is applied to the grid through the condenser C. The flow of grid current charges the condenser to a voltage approximately equal to the peak voltage between A and B, the side of the condenser adjacent to the grid being negative. During the half of the cycle that makes the plate and the anode positive, the grid is negative with respect to the cathode of the amplifier by an amount equal to the sum of the condenser voltage and the instantaneous voltage between A and B. If the phototube is dark, the charge on the condenser cannot leak off; but if the phototube is illuminated, the charge flows off through the phototube, lowering the negative voltage of the grid and increasing the plate current. Adjustment of the potentiometer varies the average grid bias and thus changes the illumination at which the relay closes. It should be noted that the circuit is in effect a bridge, the voltage of the grid during the negative half of the cycle depending upon the relative size of C and the sum of the phototube and amplifier grid-to-plate capacitances. If C is too small, no charge is stored during the negative half cycle. If C is too large, on the other hand, the reduction of condenser voltage during the positive half cycle, caused by the phototube current, is too small and the sensitivity is low. A 0.0005-μf condenser gives satisfactory results with a type 6J5 amplifier tube and a 0.00025-μf condenser with a type 6SK7 amplifier. The over-all sensitivity is considerably higher with the 6SK7 than with the 6J5 amplifier tube.

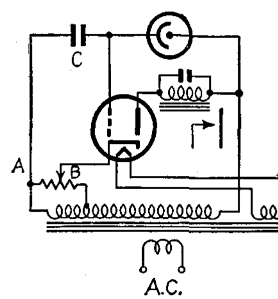


FIG. 13-16.—A-c–operated forward phototube circuit.

13-10. Circuits for Measurement of Illumination.[1]—The foregoing circuits may be used for measuring illumination if the relay is replaced by a meter. Because the variation of plate current may be only a small fraction of the operating plate current of the amplifier, it is usually necessary to balance out the steady component of plate current in order to obtain accurate readings. This may be accomplished by the circuit of Fig. 13-17. The voltage divider P is adjusted so that there is no difference of potential across the meter and all the plate current flows through the resistance. Any change in plate current results in the flow of current

[1] For more complete treatments of this subject, see W. E. FORSYTHE (editor) "Measurement of Radiant Energy," McGraw-Hill Book Company, Inc., New York, 1937; also KEITH HENNEY, "Electron Tubes in Industry," McGraw-Hill Book Company, Inc., New York, 1937.

through the meter. For high sensitivity the resistance R must be large as compared with the meter resistance, but increase of resistance necessitates an increase of plate supply voltage. Steady current can also be eliminated from the meter by the use of the two-tube balanced circuit of Fig. 13-18.[1] This circuit is unaffected by fluctuations of battery voltage and reads linearly over a wide range of illumination. The meter in the circuits of Figs. 10-17 and 10-18 may be replaced by other types of current-operated devices.

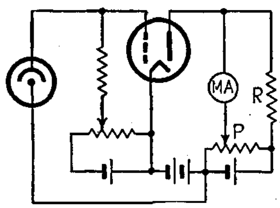


FIG. 13-17.—Phototube circuit for the measurement of illumination.

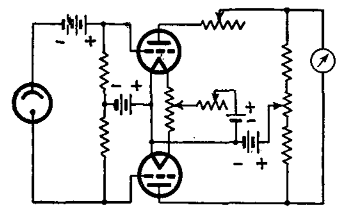


FIG. 13-18.—Balanced circuit for the measurement of illumination.

13-11. Use of Phototubes in Phase-control and Trigger Circuits.—Phototubes can be incorporated in the thyratron phase-control circuits of Figs. 12-76 to 12-86, thus making possible the control of load current by light. The phototubes are used in place of the variable resistance elements in the phase-shifting circuits, being connected so that the phototube and thyratron anodes are positive during the same half of the cycle.

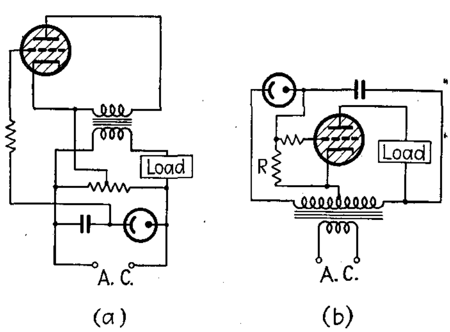


FIG. 13-19.—Thyratron phase-control circuits incorporating phototubes, by means of which large currents may be controlled by illumination.

Because of the very high resistance of phototubes and the difficulty of obtaining comparable reactance at commercial frequencies by means of an inductance, a condenser must be used as the reactive element in the phase-shifting circuit. Two typical circuits are shown in Fig. 13-19. The function of the resistance R in circuit b is to provide a leakage path for electrons that collect on the grid side of the condenser.

[1] EGLIN, J. M., *J. Opt. Soc. Am.*, **18**, 393 (1929); KOLLER, L. R., *J. Western Soc. Eng.*, **36**, 15 (1931).

Phototubes may also be used to control the trigger circuits of Secs. 10-1 to 10-5. Phototubes in series with or in place of the resistances R_c of Fig. 10-7 allow the current transfer to be initiated by the interruption of light beams. Phototubes in parallel with or in place of the resistances R_c' allow the circuit to be tripped by increase of illumination.

13-12. Limitations of Direct-coupled Circuits.—The circuits of Figs. 13-11 to 13-19 respond both to periodically repeated changes of illumination and to sustained changes of steady illumination. Greater sensitivity may be obtained by the use of an additional stage of amplification, directly coupled to the first, but it is then difficult to adjust the circuit to operate at different illumination levels, and small variations in the illumination level alter the bias of the second stage. The difficulty of obtaining high amplification in direct-coupled circuits imposes a lower limit to the change in illumination that can be used to operate circuits of this type. In addition to having relatively low sensitivity, these circuits are unsatisfactory for the amplification of modulated light because they are thrown out of adjustment by changes of average illumination.

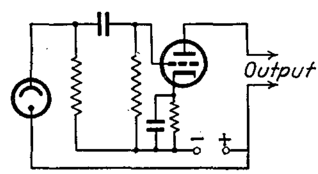


FIG. 13-20.—Circuit for converting modulated light into alternating current or voltage.

13-13. Circuits for Use with Modulated Light.—To amplify only the varying component of anode current produced by fluctuating or modulated light, as in the reproduction of sound recorded on film, the phototube is usually coupled to the first amplifier tube by means of resistance-capacitance coupling, as shown in Fig. 13-20. The coupling condenser prevents the variation of amplifier bias with changes of average illumination. After the first stage, any type of amplifier having the desired characteristics may be used.

Circuits in which the phototube is coupled to the amplifier through a condenser can also be made to respond to sustained changes in average or steady illumination by interrupting the light periodically by means of a revolving disk containing evenly spaced holes or by interrupting the anode current. One ingenious form of the latter method involves the use of an additional phototube in series with the main one. The light from an oscillating glow-discharge tube is allowed to fall upon the auxiliary phototube, which therefore passes a pulsating current, the amplitude of which varies with the illumination of the first tube. The frequency of interruption is known as the *carrier frequency*. Other schemes make use of magnetic fields which deflect the photoelectrons, of the Kerr cell, and of various types of oscillators.[1] By these means much

[1] RICHTER, W., *Electronics*, August, 1935, p. 245; ZWORYKIN, V. K., and WILSON, E. D., "Photocells and Their Applications," p. 206, John Wiley & Sons, Inc., New York, 1934.

greater sensitivity to sustained changes of steady illumination may be attained than by the circuits of Figs. 13-11 to 13-18.

13-14. Sensitivity of Phototube and Amplifier.—The theoretical sensitivity, or change of plate current per lumen change in light flux, of the circuits of Figs. 13-11 to 13-13 can be readily determined. For a vacuum phototube the change in voltage across the coupling resistor per lumen change of flux is equal to the product of the tube sensitivity in microamperes per lumen by the coupling resistance in megohms, or

$$\frac{\Delta e_c}{\Delta F} = sR_c \text{ volts per lumen} \tag{13-4}$$

If the change of voltage is not too great, the resulting change of plate current is

$$\Delta i_b = \frac{\mu \Delta e_c}{r_p + R_b} \text{ amp} \tag{13-5}$$

and

$$\frac{\Delta i_b}{\Delta F} = \frac{s\mu R_c}{r_p + R_b} \text{ amp per lumen} \tag{13-6}$$

in which the resistances r_p and R_b are measured in ohms and R_c in megohms. The resistance of milliammeters and of relays suitable for use in the plate circuit of an amplifier tube is usually negligible in comparison to the plate resistance of the tube, so that Eq. (13-6) may be simplified to

$$\frac{\Delta i_b}{\Delta F} \cong sg_mR_c \times 10^{-3} \text{ ma per lumen} \tag{13-7}$$

in which s is the phototube sensitivity in microamperes per lumen, g_m is the amplifier transconductance in micromhos, and R_c is the coupling resistance in megohms.[1] If the cathode is uniformly illuminated or if the light is concentrated upon the cathode by means of a uniformly illuminated lens, the sensitivity can be expressed also in milliamperes per foot-candle change in illumination and is equal to $sg_mR_cA \times 10^{-3}$ ma/ft-candle, where A is the effective cathode area or the area of the lens. It is important to note that the sensitivity is proportional to the transconductance of the amplifier tube.

If the changes of flux are very large, the value of the transconductance of the amplifier tube may vary considerably over the range of operation. A more accurate determination of the sensitivity of the circuit can then be made by using in place of g_m the ratio of the change in plate current to the change in grid voltage, as determined from the plate diagram or the transfer characteristic of the amplifier tube. If a gas phototube is used in place of a vacuum phototube, the phototube sensitivity is not constant, and the ratio of the actual change of anode current to the change in flux should be used in place of s. This may be determined from the

[1] Note that r_p, μ, g_m, and s must be determined for the given operating points.

anode-circuit diagram or from a curve of output voltage *vs.* flux derived from it.

13-15. The Anode Diagram.—Anode diagrams similar to the plate diagrams of amplifier tubes are often of value in determining the performance of phototubes. An anode diagram for a vacuum phototube is illustrated in Fig. 13-21. The load line passes through the point on the voltage axis corresponding to the anode supply voltage and makes an angle with the voltage axis whose tangent is equal to the reciprocal of the load resistance (expressed in volts per ampere). Above saturation the characteristic curves of Fig. 13-21 corresponding to equal increments of flux are practically horizontal, equidistant, straight lines and therefore intersect the load line at equidistant points. This shows that if the anode voltage of a vacuum phototube is always sufficiently high to give saturation current, the current and the voltage across the load resistor are essentially proportional to the flux.

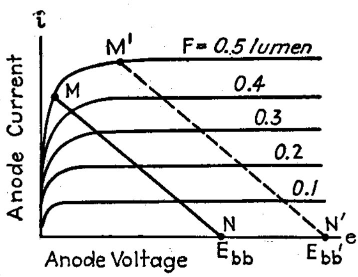


FIG. 13-21.—Anode diagram for a vacuum phototube.

Except over a very small range of voltage, the characteristics of gas phototubes are not equidistant, parallel, straight lines, and so the anode current and output voltage are not proportional to flux. The anode

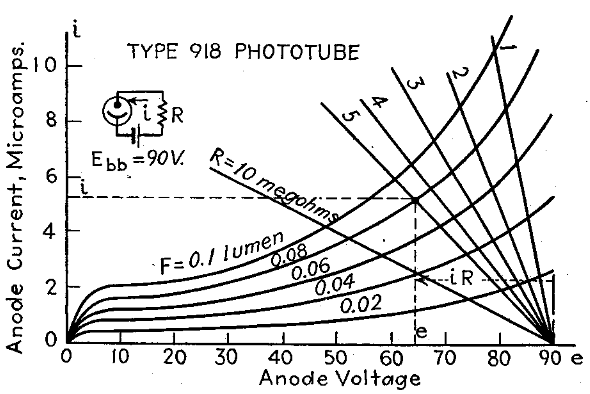


FIG. 13-22.—Anode diagram for the type 918 gas phototube.

current may be expanded in an infinite series similar to the plate-current series for an amplifier tube. It is difficult, however, to evaluate the coefficients of such an expansion in the solution of problems. It is more practical to determine the output voltage and harmonic content by means of the anode diagram. Figures 13-22 and 13-23 show the anode-circuit diagram of the type 918 phototube and the curves of output voltage *vs.*

luminous flux which are derived from it. Harmonic content can be determined from the anode diagram of Fig. 13-22 or the voltage-flux curves of Fig. 13-23 by the use of equations given in Chap. 4, in the same manner as for amplifier tubes.

It is important to bear in mind that the anode-circuit diagram and the curves derived from it do not take into account the lag in response of a gas phototube and are strictly applicable therefore only at frequencies below 100 or 200 cps. At higher frequencies accurate determinations can be made only by laboratory measurements. It should be noted that, when the a-c resistance of the coupling circuit differs from the d-c resistance, as in Fig. 13-20, both the static and the dynamic load line must be used in the anode diagram (see Secs. 4-9 and 4-10).

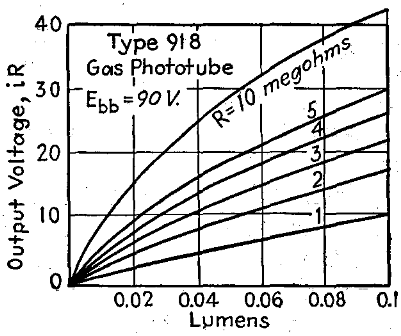


FIG. 13-23.—Curves of output voltage *vs.* luminous flux for the type 918 gas phototube with load, derived from the anode diagram of Fig. 13-22.

13-16. Design of Phototube Circuits.—The fact that the anode current of a vacuum phototube is practically independent of anode voltage above saturation means that the anode current is also independent of coupling resistance if the anode voltage is always above saturation. The output voltage therefore increases linearly with coupling resistance. If the coupling resistance is made very large, however, the anode voltage may fall below saturation at high values of flux, destroying the linearity between current and flux, as shown by the load line *MN* in Fig. 13-21. This difficulty can be remedied by increasing the anode supply voltage, as indicated by the load line *M'-N'*. The size of the coupling resistance that can be satisfactorily used at a given maximum flux is limited by the maximum allowable anode voltage (usually not less than 250 volts in vacuum phototubes), by tube and circuit leakage conductance, and by resistor "noise." With special care in the elimination of circuit and tube leakage it is possible to use properly designed coupling resistors as large as 100 megohms, but ordinarily it is not advantageous to exceed 25 megohms. When extreme sensitivity is essential, as in the measurement of minute quantities of light, leakage in the phototube and phototube socket should be minimized by the use of a phototube in which the anode terminal is at the top of the tube, and of low anode voltage. If the anode voltage is kept below 20 volts, the danger of ionization of residual gas in vacuum phototubes is avoided. Grid leakage in the amplifier tube is reduced by the use of low screen and plate voltage and low cathode temperature. High sensitivity may be attained without the need of high anode supply voltage by using a pentode[1] (see Fig. 6-6) or a second phototube[1] in place of the coupling resistor. The illumination of the second phototube may be maintained constant or may be decreased when that of the first is increased.

When gas phototubes are used in the conversion of modulated light into alternating voltage or current, the harmonic content of the output increases with coupling resistance. Allowable amplitude distortion therefore usually limits the coupling resistance to 1 or 2 megohms. In control circuits, however, large values of coupling resistance can also be used with gas phototubes.

As explained on page 157, danger of cumulative increase of plate current as the result of primary and secondary grid emission and of ionization of residual gas imposes a limit on the grid-circuit resistance that can be used with many amplifier tubes at rated plate voltages. To make use of the large values of coupling resistance that are desirable in order to give high sensitivity, it is necessary, therefore, to use low values of plate and screen operating voltage. This is normally not a disadvantage, since enough plate current to operate small relays may be readily obtained with most tubes at plate voltages of 100 or less. A second amplifier stage, operated at normal voltages, may be used if necessary. Care must also be taken to ensure that the allowable amplifier plate dissipation is not exceeded for any value of light flux at which the phototube is operated. Since *IR* drop in the grid resistor prevents the application of appreciable positive voltage to the grid, ample protection is usually afforded by the use of sufficiently low plate and screen supply voltages so that the allowable plate dissipation is not exceeded when the grid voltage is zero.

13-17. Secondary-emission Multipliers.—The high amplification necessary for the conversion of modulated light into voltage and current in the reproduction of sound recorded on film, and in television, may be obtained not only by the use of conventional amplifiers, but also by making use of secondary emission.[2] The basic construction of one type of secondary-emission multiplier is shown in Fig. 13-24.[3] Photoelectrons emitted by the cathode are drawn to the first anode by means of electrostatic field. The secondary electrons from this anode, and reflected primary electrons, are drawn to the second anode where secondary emission also takes place. The process is repeated at successive anodes, each of which is at a higher potential than the preceding one. Since

[1] SHEPARD, F. H., JR., *RCA Rev.*, **2,** 149 (1937); BULL, H. S., and LAFFERTY, J. M., *Electronics*, November, 1940, p. 31 and December, 1940, p. 71.

[2] IAMS, H., and SALZBERG, B., *Proc. I.R.E.*, **23,** 55 (1935).

[3] ZWORYKIN, V. K., MORTON, G. A., and MALTER, L., *Proc. I.R.E.*, **24,** 351 (1936). See also supplementary bibliography on pp. 561–563.

each primary electron may knock out 5 or more secondary electrons, very high current amplification is obtainable by the use of 8 or 10 anodes.

Focusing of the electron beam between anodes may be accomplished either by electrostatic field alone or by a combination of crossed electrostatic and magnetic fields. The general arrangement of a multiplier based upon the use of crossed magnetic and electrostatic fields is shown

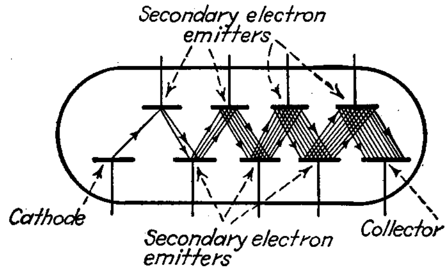


FIG. 13-24.—Diagram showing the basic principle of operation of a secondary-emission amplifier tube.

in Fig. 13-25. It consists of two rows of electrodes, the bottom row being cesium-treated secondary emitters, while the upper row serves solely to maintain a transverse electrostatic field between the two sets of elements. Each target in the bottom row is made positive with respect to the preceding one so that it will produce secondary electrons when struck by electrons originating from the latter. Each upper field electrode is maintained positive relative to the target just below it by an internal

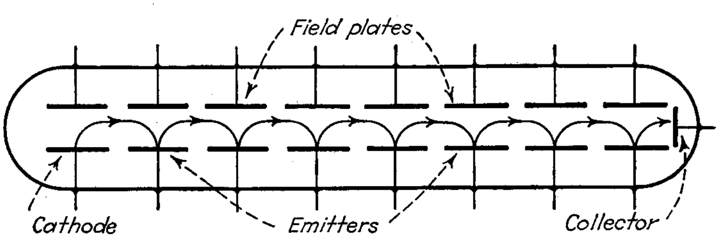


FIG. 13-25.—Arrangement of electrodes in secondary-emission current multiplier using crossed magnetic and electric fields.

connection to the next target. A magnetic field is established in the tube at right angles to its axis and to the electric field between the two rows of plates. Electrons leaving any of the targets are deflected by the combined fields in such a way that they strike the next target, giving rise to new secondary electrons which are in turn deflected onto another target and so on, through the tube. The voltages necessary for the various anodes may be obtained from a single voltage source by the use of voltage dividers. A voltage divider for a number of the earlier stages, which need carry only small currents, may be incorporated in the tube.

In another type of electron multiplier the electrons are caused to move back and forth between two electrodes by means of a high-frequency electric field.[1]

By means of electron multipliers an amplification of several million and a sensitivity of ten or more amperes per lumen may be obtained with a single tube which is not much larger than a receiving tube. The saving in bulk and the gain in simplicity are obvious advantages. A more important advantage of such a tube over a conventional amplifier and phototube is a great reduction of noise-to-signal ratio. In addition, the multipliers are very stable, are insensitive to external interference, and have an excellent frequency characteristic. The power required for their operation is comparable with that required for a conventional amplifier. The type 931A is a typical electron-multiplier tube.

13-18. Special Photoemissive Tubes.—A number of special tubes have been devised that incorporate phototubes and glow tubes or phototubes and amplifiers in a single envelope. An example of the former is the *photoglow tube,* developed by Knowles, in which the glow is initiated by light falling upon the cathode, which is designed to give high photoelectric emission.[2] Some receiving tubes show appreciable sensitivity to light, the plate current changing when light falls upon a free grid.[3] By special treatment of the grid the sensitivity can be greatly enhanced.[3] Ordinarily it is more satisfactory to use separate tubes to perform the functions of the phototube and the glow tube or amplifier.

An interesting and valuable application of photoemission is made in the iconoscope, which is used in television transmission and in electron telescopes.[4] An understanding of the operation of the iconoscope requires a knowledge of the principle of television "scanning," which is beyond the scope of this book. The interested reader should refer to the literature or to books on television.

13-19. Photoconductive Cells.—The photoconductive effect was discovered in 1871 and discussed in 1873 by Willoughby Smith, who reported a change in the resistance of the crystalline form of selenium with illumination. Although a number of semiconducting substances, such as thalium oxysulphide and molybdenite, are now known to be photo-

[1] FARNSWORTH, P. T., *J. Franklin Inst.*, **218**, 411 (1934); *Electronics*, August, 1934, p. 242, November, 1935, p. 31.

[2] ZWORYKIN and WILSON, *op. cit.*, p. 214.

[3] ASADA, T., and HAGITA, K., *J. Inst. Elec. Eng. Japan*, **51**, 8 (1931); KOECHEL, W. P., *Electronics*, December, 1932, p. 372; MCILVAINE, H. A., *Electronics*, August, 1933, p. 224.

[4] ZWORYKIN, V. K., *Proc. I.R.E.*, **22**, 16 (1934).

conductive,[1] selenium is still used in most photoconductive cells. Early selenium cells were made by covering two closely spaced conductors with a thin coating of amorphous selenium, which was subsequently converted into crystalline selenium by heat-treatment. The conductors were commonly made in the form of interlocking metallic combs or of two wires wound close together on a nonconducting frame, as shown in Fig. 13-26. Modern cells are made by condensing selenium vapor in a thin film upon a double grid of gold or platinum fused to or sputtered upon a glass plate. The glass is maintained at the proper temperature so that the film is deposited in the crystalline form.

The photoconductive effect is thought to be caused by the emission of electrons within the material. In their motion under the influence of applied voltage, these photoelectrons produce other ionization. Equilibrium is established when the rate of recombination becomes equal to the rate of ion formation. This theory is in agreement with experimentally observed laws. It is found experimentally that (1) for a short time after the light is applied, before much displacement of positive ions or appreciable secondary ionization takes place, the change in current resulting from illumination is proportional to the incident flux; (2) for exposures of sufficient length to enable the current to reach equilibrium, the current is proportional to the square root of the incident flux; (3) with rapidly fluctuating light the alternating current is proportional to the incident flux and inversely proportional to the frequency of fluctuation. The lag in the response of a selenium cell is caused mainly by the low velocity of the relatively heavy positive ions. The dependence of current upon time of illumination and the related falling dynamic response curve make the photoconductive cell unsatisfactory for the undistorted conversion of modulated light into voltage or current.

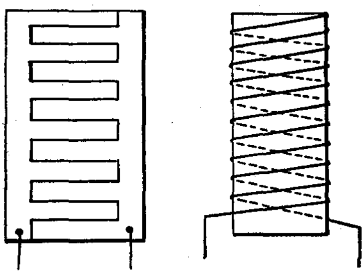


FIG. 13-26.—Two forms of grids used in selenium cells.

Unlike the phototube, the photoconductive cell passes appreciable current when it is dark, the *dark resistance* of commercial cells ranging from about 100,000 ohms to 25 megohms. The ratio of *light* to *dark* currents may be as high as 25 but is more commonly about 8 or 10. The current capacity does not usually exceed a few milliamperes, but by proper design may be increased to as much as $\frac{1}{4}$ amp. The sensitivity is directly proportional to the voltage across the cell, and consequently cells are usually operated at as high voltage as possible without overheating and subsequent breakdown. In Figs. 13-27 and 13-28 are

[1] PFUND, A. H., *Phys. Rev.* (2), **7**, 551 (1917).

shown the current–wave-length characteristic curve and the dynamic response curves of a typical selenium cell, the General Electric type FJ-31. The peak of the spectral sensitivity curve is in the red, which

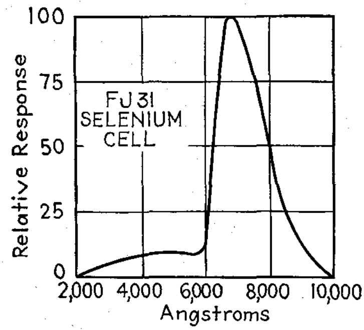


FIG. 13-27.—Relative response of the type FJ-31 selenium cell in per cent of maximum response as a function of wave length.

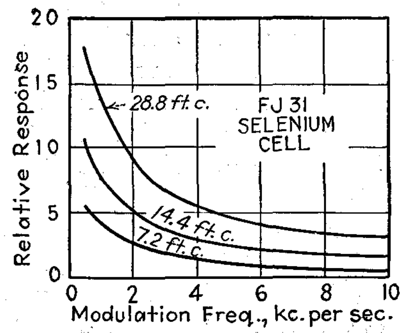


FIG. 13-28.—Dynamic response curves for the type FJ-31 selenium cell.

makes the cell suitable for use with artificial illumination and infrared radiation.

Expressions for the voltage sensitivity of selenium cells and of the ampere-turn sensitivity of selenium cells used to operate relays may be

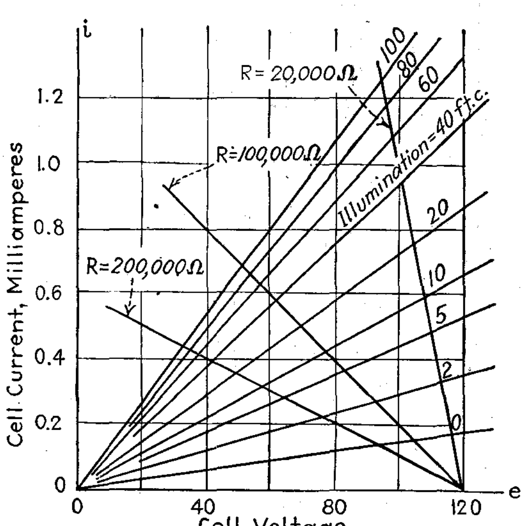


FIG. 13-29.—Typical current-voltage diagram for a selenium cell.

derived.[1] It is generally simpler, however, to make use of a current-voltage diagram in the same manner as with phototubes and amplifiers.

[1] ZWORYKIN, V. R., and WILSON, E. D., "Photocells and Their Applications," 2d ed., John Wiley & Sons, Inc., New York, 1934, pp. 175–180.

Because the cell resistance at a given value of illumination is independent of cell voltage, the current-voltage curves are straight lines and may be readily derived from a single static curve of current *vs.* illumination at any voltage. A family of current-voltage characteristics, together with a number of load lines for 120-volt supply voltage, are illustrated in Fig. 13-29. It can be seen from this diagram that the change in current per foot-candle change in illumination, and hence the current sensitivity, are a maximum for zero load resistance and decrease rapidly as the load resistance is increased. The change in voltage across the load resistance per foot-candle change in illumination, and hence the voltage sensitivity, pass through a maximum as the load resistance is increased. With small changes of illumination, maximum voltage sensitivity is obtained when the load resistance is approximately equal to the average resistance of the cell in the range of operation.

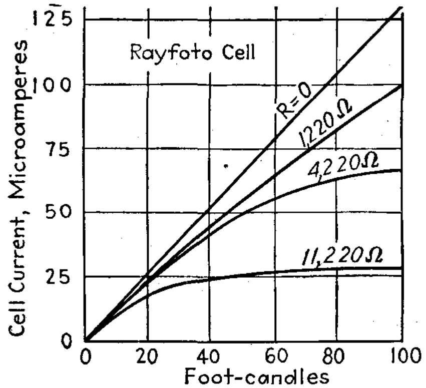


FIG. 13-30.—Curves of current *vs.* illumination for the Rayfoto cell.

13-20. Photovoltaic Cells (Barrier-layer Cells).—The photovoltaic effect was first observed by Becquerel,[1] who discovered that an e.m.f. is set up between two electrodes immersed in an electrolyte when one of the electrodes is illuminated. Two cells based upon the Becquerel effect were the Rayfoto cell, which employed a cathode of cuprous oxide on copper and an anode of lead in an electrolyte of lead nitrate; and the Photolytic cell,[2] which was similar in construction but which employed cuprous oxide on copper for both electrodes, thus eliminating voltage when the cell was dark. Curves for the Rayfoto type of cell, taken from an article by Fink and Alpern,[3] are shown in Figs. 13-30 and 13-31. It can be seen that the current is proportional to the illumination when the cell is short-circuited, but that external circuit resistance destroys the linearity and also results in considerable reduction of current sensitivity. Bending of the curves at high values of external resistance results from internal leakage conductance, which increases with illumination and thus reduces the effective load resistance. The response of this cell to fluctuating light falls rapidly with increase of light-modulation frequency, so that it is not satisfactory for converting modulated light into sound.

[1] BECQUEREL, E., *Compt. rend.*, **9**, 144, 561 (1839).

[2] Although these cells are no longer on the market, they are still of theoretical interest.

[3] FINK, G., and ALPERN, D. K., *Trans. Am. Electrochem. Soc.*, **58**, 275 (1930).

A more recent type of photovoltaic cell, the *barrier-layer* cell, is made of a layer of cuprous oxide in very close contact with a metal, usually copper, gold, or platinum, or of a thin layer of iron selenide on iron. It is found that illumination of the boundary or *barrier plane* between the compound and the metal sets up an e.m.f. and that current flows if the compound and the metal are connected through an external circuit. The phenomenon, which is now known to be caused by the emission of electrons from the compound into the metal, was first observed by Fritts in 1884 in the course of experiments on selenium cells.[1] Improvements in design and manufacture have resulted in selenium[2] and cuprous oxide[3] barrier-layer cells capable of producing currents of 100 μa or more per lumen.

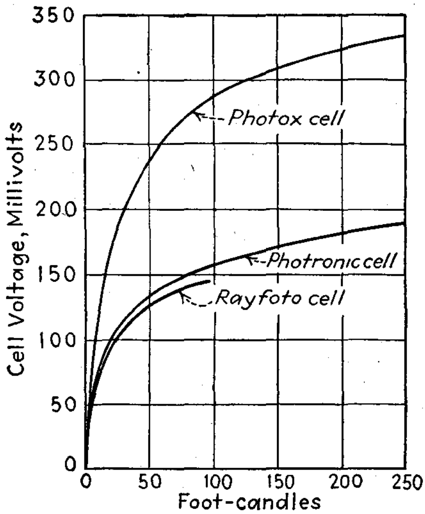


FIG. 13-31.—Curves of generated voltage *vs.* illumination for three types of photovoltaic cells.

There are two forms of barrier-layer cells. In the first, called the *back-effect* cell, the cuprous oxide or iron selenide is deposited upon the metal in the form of a thin layer. The light passes through the compound to the metal, as shown in Fig. 10-32*a*, the photoelectric action taking place at the boundary between them. Electrons flow across the boundary from the oxide or selenide to the metal, and the conventional direction of current in the external circuit is from the compound to the

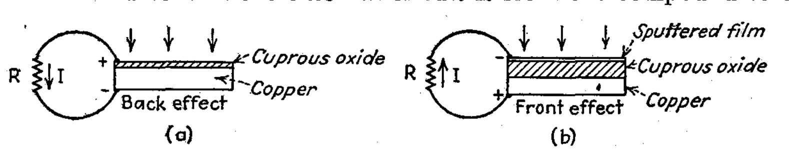


FIG. 13-32.—Arrangement of electrodes in back-effect and front-effect barrier-layer cells.

metal. Contact to the compound is made by means of a conductor pressed against the compound or by means of a thin layer of metal sputtered on the compound near the outer edge of the surface. Because of absorption of the light of shorter wave lengths in the oxide, the current–wave-length characteristic of a cuprous oxide back-effect cell shows a maximum at the red end of the spectrum.

[1] FRITTS, C. F., *Proc. Amer. Asso. Adv. Science*, **33**, 97 (1884); ADAMS, W. G., and DAY, R. E., *Proc. Roy. Soc. (London)*, **25**, 113 (1877).

[2] BERGMANN, *Physik. Z.*, **32**, 286 (1931).

[3] LANGE, B., *Physik. Z.*, **31**, 139 (1930).

The second form of cell, called the *front-effect cell,* is made by sputtering a thin film of silver, gold, or platinum on the exposed surface of the cuprous oxide or selenide layer, which is formed on the metal as in the back-effect cell. Light enters the cell through the sputtered film, as shown in Fig. 13-32*b*. Emission of electrons takes place at both surfaces of the oxide or selenide but is greater at the front surface, causing the sputtered film to become negatively charged with respect to the metal and resulting in current in the conventional sense from the metal to the sputtered film through the external circuit. Absorption of light by the sputtered film is small, and the peak of the spectral sensitivity curve is at the blue end of the visible spectrum.

The short-circuit current of a barrier-layer cell, like that of an electrolytic cell, is directly proportional to the total flux incident upon the surface and is independent of the area illuminated. The introduction of external resistance destroys the linearity between current and flux and reduces the current sensitivity, the curves of current *vs.* illumination being similar to those of Fig. 13-30. The internal series resistance of the cell results almost entirely from the fact that the emitted electrons must pass along the poorly conducting layer of cuprous oxide or iron selenide from the point of emission to the external contact. The response of barrier-layer cells to fluctuating light falls rapidly with increase of modulation frequency because of the relatively high capacitance between the two surfaces.

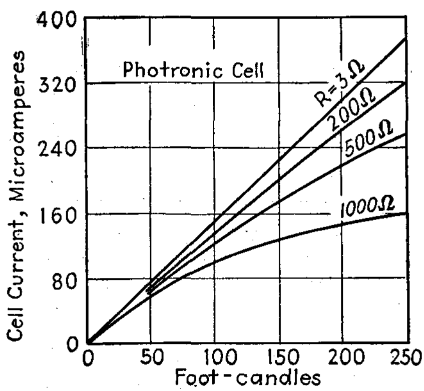


FIG. 13-33.—Curves of current *vs.* illumination for the Photronic cell.

The Photox cell is an example of a front-effect cuprous oxide cell. It has a current sensitivity of 100 μa/lumen, a cathode area of 0.02 sq ft, and produces a voltage of 0.2 volt with a 50-ft-candle illumination. A curve between cell voltage and illumination is shown in Fig. 13-31.

The Photronic cell is an iron selenide cell. The characteristics of the Photronic cell are shown in Figs. 13-31, 13-33, 13-34, and 13-35. The manufacturers state that the response of the cell does not vary over long periods of time. It does, however, show a certain amount of fatigue, as shown by Fig. 13-35, and is sensitive to temperature changes. Because of high interelectrode capacitance, the response to modulated light is poor. For 100 per cent response at 60 cps, the response is 58 per cent at 120 cps, 30 per cent at 240 cps, and 6.4 per cent at 1000 cps. The cell is unsuited, therefore, to the conversion of modulated light into alternating current. The internal resistance of the Photronic cell differs from that of the Photox cell, necessitating differences in the design of circuits.[1]

The chief advantage of photovoltatic cells lies in the fact that they require no external voltage supply.[2] Relays have been designed to

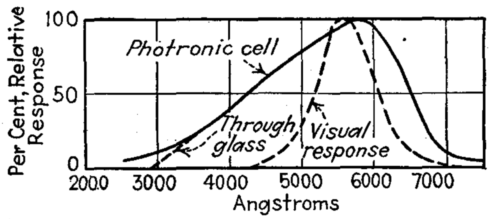


FIG. 13-34.—Curve of relative response *vs.* wave length for the Photronic cell.

operate directly from photovoltaic cells, but, because the resistance of the relay reduces the current sensitivity, it is difficult to obtain high armature torque. The generated voltage is so small that an amplifier of high current sensitivity is required if an amplifier is used between the cell and the relay. A single-stage amplifier is ordinarily inadequate. Even for small values of flux, for which the voltage sensitivity is high, the Photronic cell has a sensitivity of only about 25 mv/lumen, and the Photox cell about 50 mv/lumen. This is very small as compared with the 500-volt/lumen sensitivity obtainable even with the vacuum type of phototube. The voltage sensitivity of photovoltaic cells may be increased by series operation of units, and the current sensitivity by parallel operation. The photovoltaic cell finds one of its principal applications in photometry. In combination with a low-resistance microammeter it makes a simple and rugged foot-candle meter.

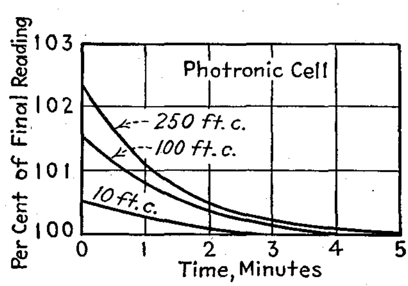


FIG. 13-35.—Curves showing fatigue in the Photronic cell.

13-21. Comparison of Phototubes, Photoconductive Cells, and Photovoltaic Cells.—From the point of view of linearity of response, absence of time lag, and freedom from fatigue and temperature effects, the phototube is far superior to either the photoconductive or the photovoltaic cell. It gives the highest voltage output of the three types but has low current output and low current sensitivity. Because of its high

[1] BARTLETT, C. H., *Rev. Sci. Instruments*, **3**, 543 (1932).
[2] LAMB, A. H., *Instruments*, **5**, 230 (1932).

voltage output, however, it may be efficiently used with an amplifier in the operation of current-controlled apparatus. The poor dynamic response of photoconductive and photovoltaic cells makes them unsuitable for the conversion of modulated light into sound, but they have the advantage of ruggedness and relatively high current output, which makes possible the direct operation of relays or meters if relatively large changes of illumination are available. The photovoltaic cell has the additional advantage of not requiring an external voltage source. Since the voltage output of a photovoltaic cell is small and independent of load resistance, the current sensitivity of a single-stage amplifier is insufficient to increase the current change available to operate a relay. Because of the inconvenience and expense involved in the use of a two- or three-stage d-c amplifier, it is usually not advantageous to use an amplifier in conjunction with a photovoltaic cell. Low-resistance relays and comparatively large changes of illumination are consequently required. The over-all voltage or current sensitivity of a photoconductive-cell circuit may be increased by the use of an amplifier.

Problems

13-1. Criticize the circuit of Fig. 13-36. (Six errors.)

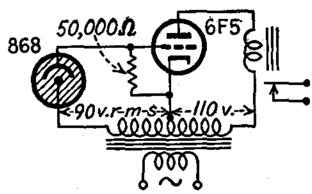


FIG. 13-36.—Circuit diagram for Prob. 13-1.

13-2. The following data apply to a photoelectrically controlled relay:

Phototube sensitivity	7 μa/lumen
Phototube cathode area	1 sq in.
Amplifier tube	Type 6J5
Coupling resistor	8 megohms
Relay operating currents	8 ma, closing; 7 ma, opening
Relay resistance	1000 ohms

a. Draw the diagram of a circuit in which increase of illumination will cause the relay to be opened. All phototube and amplifier voltages are to be derived from a single 220-volt d-c supply.

b. Specify values of resistance that will provide the correct filament current and suitable values of plate voltage and adjustable biasing voltage.

c. Specify the value and polarity of the C-supply voltage that will just open the relay at a phototube illumination of 72 ft-candles. The source of light is a tungsten filament operated at 2870°K.

d. How much must the illumination be decreased to cause the relay to close?

13-3. The following data apply to a phototube-amplifier-relay circuit:

Phototube	Type 868 (see Fig. 13-7)
Phototube cathode area	0.9 sq in.
Amplifier tube	Type 6J5
Coupling resistance	1 megohm
Relay resistance	1000 ohms
Relay closing current	6 ma
Relay opening current	4 ma
Phototube anode supply voltage (including biasing voltage)	90 volts
Amplifier plate supply voltage	90 volts
Light source	Tungsten filament at 2870°K

a. Show the circuit diagram for a circuit in which increase of illumination causes the relay to close.

b. From the plate characteristics of the type 6J5 tube find the grid voltages at which the relay will open and close.

c. From the characteristic curves of the type 868 phototube (Fig. 13-7), plot a curve of voltage across the coupling resistor as a function of light flux.

d. From the curve of (*c*) determine the grid supply voltage necessary to open the relay when the incident flux is 0.2 lumen.

e. Determine the change in illumination, in foot-candles, necessary to close the relay after the circuit is adjusted as in (*d*).

13-4. *a.* Plot curves of per cent second and third harmonic in the voltage output of a type 918 phototube operated from an 80-volt supply with a 7-megohm load, as a function of incident light flux for 80 per cent and 100 per cent modulation of incident light.

b. Repeat for a load having a d-c resistance of 10 megohms and an a-c resistance of 7 megohms at signal frequency.

13-5. Determine the factor by which the current scale of the current-voltage characteristics of a type 868 gas phototube must be multiplied in order that the characteristics shall hold for a light source whose energy distribuion is given by the following table:

Wave length, angstroms	Relative energy	Wave length, angstroms	Relative energy	Wave length, angstroms	Relative energy
3000	0	5500	42	8000	15
3500	1	6000	65	8500	6
4000	6	6500	93	9000	2
4500	14	7000	100	9250	0
5000	25	7500	55		

Bibliography

The literature on the scientific aspects of photoelectricity is very extensive. More complete treatments of the subject than that given in this chapter will be found in the following books and articles:

Hughes, A. L., and DuBridge, L. A.: "Photoelectric Phenomena," McGraw-Hill Book Company, Inc., New York, 1932.

ZWORYKIN, V. K., and WILSON, E. D.: "Photocells and Their Application," John Wiley & Sons, Inc., New York, 1934.

WALKER, R. C., and LANCE, T. M. C.: "Photo-electric Cell Applications," 3d ed., Sir Isaac Pitman & Sons, Ltd., London, 1939.

FIELDING, T. J.: "Photo-electric and Selenium Cells," Chapman & Hall, Ltd., London, 1935.

KOLLER, L. R.: "The Physics of Electron Tubes," McGraw-Hill Book Company, Inc., New York, 1937.

ALLEN, H. S.: "Photoelectricity," Longmans, Green & Company, New York, 1913.

HUGHES, A. L.: "Photo-Electricity," Cambridge University Press, London, 1914.

CAMPBELL, N. R., and RITCHIE, D.: "Photoelectric Cells," Sir Isaac Pitman & Sons, Ltd., London, 1934.

HUGHES, A. L.: "Report on Photoelectricity," National Research Council.

BARNARD, G. P.: "The Selenium Cell," Archibald Constable & Co., Ltd., London, 1930.

BENSON, T. W.: "Selenium Cells," Spon & Chamberlain, New York, 1919.

LANGE, B.: "Photoelements and Their Application," Reinhold Publishing Corporation, New York, 1938.

HUGHES, A. L.: Fundamental Laws of Photoelectricity, *Elec. Eng.*, **53**, 1149 (1934).

NIX, F. C.: Photo-conductivity, *Rev. Modern Phys.*, **4**, 723 (1932).

GRONDAHL, L. O.: The Copper Cuprous Oxide Rectifier and Photoelectric Cell, *Rev. Modern Phys.*, **5**, 141 (1933).

The following partial list of books and articles covers applications of light-sensitive tubes and cells:

HENNEY, KEITH: "Electron Tubes in Industry," McGraw-Hill Book Company, Inc., New York, 1937.

GULLIKSEN, F. H., and VEDDER, E. H.: "Industrial Electronics," John Wiley & Sons, Inc., New York, 1935.

BREISKY, J. V., and ERIKSON, E. O.: *J. Am. Inst. Elec. Eng.*, **48**, 118 (1929).

VEDDER, E. H.: *Elec. J.*, **27**, 152 (1930).

ROMAIN, P. B.: *Rev. Sci. Instruments*, **4**, 83 (1933).

CHAMBERS, D. E.: *Elec. Eng.*, **54**, 82 (1935).

GULLIKSEN, F. H., and STODDARD, R. N.: *Elec. Eng.*, **54**, 40 (1935).

WILSON, E. D.: *Electronics*, April, 1935, p. 118.

SHEPARD, F. H., JR.: *Electronics*, February, 1935, p. 59; June, 1936, p. 34; *Proc. I.R.E.*, **24**, 1573 (1936); *RCA Rev.*, **2**, 149 (1937).

POWERS, R. A.: *Electronics*, September, 1935, p. 11; June, 1936, p. 22.

VEDDER, E. H.: *Elec. J.*, **32**, 425 (1935).

LAMB, A. H.: *Elec. Eng.*, **54**, 1186 (1935).

HOLMES, R. M.: *Electronics*, April, 1937, p. 33.

Secondary Multiplier Phototubes

IAMS, H., and SALZBERG, B.: *Proc. I.R.E.*, **23**, 55–64 (1935) (early tubes).

WEISS, G.: *Electronics*, September, 1936, p. 60 (abstract on German "mesh" tube).

ZWORYKIN, V. K., MORTON, G. A., and MALTER, L.: *Proc. I.R.E.*, **24**, 351–75 (1936) (early tubes, especially electromagnetically focussed).

KUBETSKY, L. A.: *Proc. I.R.E.*, **25**, 421–33 (1937) (Russian tubes).

Electronics, August, 1934, p. 242 and November, 1935, p. 31 (Farnsworth's multipactor and image dissector tubes).

FINKE, H. A.: *Proc. I.R.E.*, **27**, 144–7 (1939) (television pickup tube).

LARSON, C. C., and SALINGER, H.: *Rev. Sci. Instruments*, **11**, 226–9 (1940) (Farnsworth's commercial multiplier phototubes).

RAJCHMANN, J. A., and SNYDER, R. L.: *Electronics*, December, 1940, p. 20 (RCA's commercial multiplier phototube and its operation).

WINANS, R. C., and PIERCE, J. R.: *Rev. Sci. Instruments*, **12**, 269–77 (1941) (Western Electric commercial phototube multiplier and operation).

Multipliers with Thermionic Cathodes, Noise Theory, Etc.

JONKER, J. L. H., and v. OVERBEEK, A. J. W. M.: *Wireless Eng.*, **15**, 150–6 (1938) (single-stage grid-controlled multiplier).

SCHOCKLEY, W., and PIERCE, J. R.: *Proc. I.R.E.*, **26**, 321–32 (1938) (theory of noise in multipliers).

ZWORYKIN, V. K., and RAJCHMANN, J. A.: *Proc. I.R.E.*, **27**, 558–66 (1939) (design of electrostatic multipliers).

FLECHSIG, W., and SANDHAGEN, M.: *Wireless Eng.*, **18**, 379–80 (1941) (abstract of German article on grid-controlled multipliers, etc.).

SCHNITGER, H.: *Wireless Eng.*, **19**, 269–70 (1942) (abstract of German article on static multipliers).

MALTER, L.: *Proc. I.R.E.*, **29**, 587–98 (1941) (variation of transconductance with frequency; limiting frequency).

WAGNER, H. M., and FERRIS, W. R.: *Proc. I.R.E.*, **29**, 598–602 (1941) (single-stage orbital-beam tube for high g_m at ultrahigh frequency).

SCHNITGER, H.: *Wireless Eng.*, **19**, 79–80 (1942) (abstract of German article on design calculations).

THOMPSON, B. J.: *Proc. I.R.E.*, **29**, 583–7 (1941) (gain limitations in voltage-controlled multipliers).

CHAPTER 14

POWER SUPPLIES

The practically universal use at the present time of alternating current in the transmission and distribution of electric power necessitates the use of some means for converting to direct current in certain applications of electric power. Although this may be accomplished by rotary converters or motor-generator sets, it is often simpler, cheaper, and more efficient to use electronic rectifiers. Other advantages of electronic rectifiers include quiet operation, ease and speed of starting and, in mercury pool types, long life, high momentary overload capacity, and immunity to damage from short circuits. Outstanding applications of electronic rectifiers are made in supplying direct voltages for electric railways, for electrolytic plants, and for the operation of electron tubes.

Since the output of a rectifier is pulsating, and the voltage used in most tube circuits must be nearly free of pulsation, rectifiers that supply direct voltages to tubes must be followed by smoothing filters. In order to prevent interference in adjacent communication lines, it is also important to keep the ripple small in the distribution of rectified currents. Because of the many factors involved in rectification and filtering, only the principal aspects of the subject are presented in this chapter, and emphasis is placed upon rectifiers and filters for use with vacuum-tube amplifiers and other vacuum-tube circuits.

14-1. Definitions.—A *rectifier* is a device having an asymmetrical conduction characteristic which is used for the conversion of an alternating current into a current having a unidirectional component. Such devices include vacuum-tube rectifiers, gaseous rectifiers, oxide (barrier-layer) rectifiers, and electrolytic rectifiers. A *half-wave rectifier* is a rectifier that changes alternating current into pulsating current, utilizing only one-half of each cycle. A *full-wave rectifier* is a double-element rectifier arranged so that current is allowed to pass to the load circuit in the same direction during each half cycle of the alternating supply voltage, one element functioning during one half cycle and another during the next half cycle.

The alternating component of unidirectional voltage from a rectifier or generator used as a source of d-c power is called *ripple voltage*. The ripple voltage is not generally sinusoidal and may, therefore, be analyzed into fundamental and harmonic components.

The importance of the ripple voltage relative to the d-c component of the rectifier or filter output voltage is often specified by means of the *per cent ripple*, which is defined by the Standards Committee of the Institute of Radio Engineers[1] as the ratio of the r-m-s value of the ripple voltage to the algebraic average value of the total voltage, expressed in percentage. The effectiveness of a smoothing filter increases, however, with the frequency of the voltage input to the filter. Furthermore, in all but the half-wave single-phase rectifier, the amplitudes of the ripple harmonics decrease rapidly with increase of the order of the harmonics. For these reasons, if a filter is designed to produce adequate filtering at the fundamental ripple frequency, the ripple harmonics will usually be reduced to negligible values, and so only the fundamental component of ripple voltage need be considered in designing the filter. The ratio of the effective value of the fundamental component of ripple voltage to the average (direct) value of the total voltage will be called the *ripple factor*.

The effectiveness of a smoothing filter in removing a component of ripple voltage of any frequency may be indicated by the effect that it would have upon a sinusoidal voltage of that frequency. The ratio of the amplitude of a sinusoidal voltage of given frequency impressed upon the input of a filter to the amplitude of the resulting sinusoidal output voltage will be termed the *smoothing factor* at that frequency.

The following symbols will be used:

Supply frequency	f
Fundamental ripple frequency	f_r
Amplitude of fundamental component of ripple voltage at filter input	E_{r1}
Average value of total output voltage	E_{dc}
Ripple factor at input to filter, E_{r1}/E_{dc}	ρ_1
Smoothing factor at any frequency	α
Smoothing factor at fundamental ripple frequency	α_1

Primed symbols will be used to indicate values at the output of the filter (across the load).

14-2. Basic Rectifier Circuits.—Two of the simplest and most commonly used rectifier circuits, the single-phase half-wave and single-phase full-wave circuits are shown in Fig. 14-1. The half-wave circuit of Fig. 14-1*a* makes use of only one-half of each cycle of the alternating voltage. In the full-wave circuit of Fig. 14-1*b*, on the other hand, current flows in the two plates (or two diodes) during alternate half cycles. In both circuits diode rectification causes current to flow through the load in only one direction. The duration of plate-current flow, the magnitude and regulation of the direct output voltage, the ripple factor, and other

[1] 1933 Report of I.R.E. Standards Committee.

characteristics of the circuits depend upon the load, which in Fig. 14-1 is assumed to include the smoothing filter, if one is used.

In the circuits of Fig. 14-1 the direct output voltage cannot exceed the crest value of the alternating voltage impressed upon one plate circuit. In many applications, however, particularly when it is desired to reduce cost and weight by eliminating the input transformer, it is advantageous to obtain output voltage in excess of the crest alternating

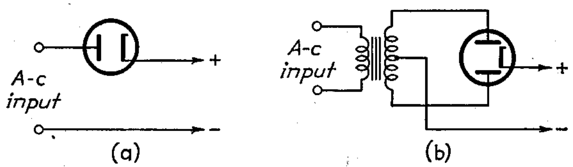


FIG. 14-1.—(a) Half-wave rectifier. (b) Full-wave rectifier.

voltage. This may be accomplished by means of voltage multiplier circuits. The full-wave voltage doubler, shown in Fig. 14-2*a*, consists essentially of two half-wave rectifiers. The two condensers are charged in alternate half cycles through two rectifier tubes. Since the load voltage is the sum of the average condenser voltages, the load voltage approaches twice the crest alternating voltage at light loads (high load resistance). In the half-wave voltage doubler of Fig. 14-2*b*, the condenser C_1 is charged to crest alternating voltage through T_1 during one half cycle. During the following half cycle the voltage of this condenser is added to the impressed voltage in the circuit of T_2, causing the con-

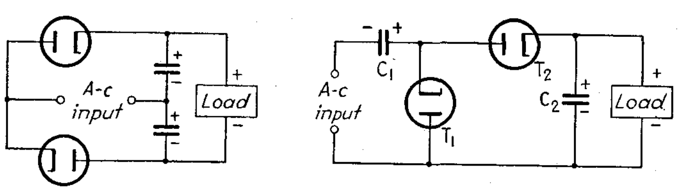


FIG. 14-2.—(a) Full-wave voltage doubler. (b) Half-wave voltage doubler.

denser C_2 to be charged to approximately twice the crest alternating voltage. The circuit of Fig. 14-2*a* has the advantages over that of Fig. 14-2*b* of lower voltage regulation, higher ripple frequency, which is easier to filter, and lower condenser voltage. The circuit of Fig. 14-2*b*, on the other hand, has the advantage of a common input and output terminal.

Figure 14-3*a* shows a voltage quadrupler in which two half-wave voltage doublers are connected in series.[1] By leaving out the tube T_1'

[1] GARSTANG, W. W., *Electronics*, February, 1932, p. 50; HONNELL, M. A., *Communications*, January, 1940, p. 14. See also "The Radio Amateur's Handbook," 1939 ed., p. 357.

and shortcircuiting the condenser C_1', this circuit can readily be converted into a voltage tripler.[1] Another voltage quadrupler based upon the half-wave doubler is shown in Fig. 14-3*b*.[2] In this circuit the alternating input voltage is impressed upon the upper voltage doubler (indicated by primed symbols) through the condenser C_2. This circuit has the advantage over that of Fig. 14-3*a* of having a common input and output terminal, but the condenser C_1' is subjected to a maximum voltage equal to three times the crest alternating voltage. Voltage multiplication of any desired ratio may be attained in this circuit by the use of additional voltage-doubler stages. The common input and output

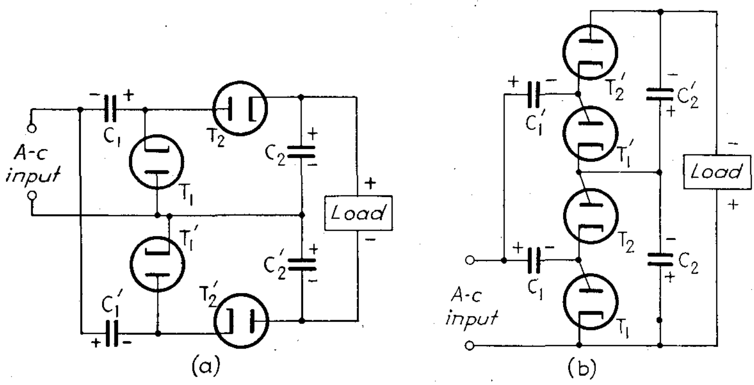


FIG. 14-3.—Voltage quadrupling circuits.

terminal can be changed from positive to negative by reversing all tubes.

Because of the differences of potential between the various cathodes in voltage multiplier circuits, it is necessary to use either separate filament transformer windings or heater-type rectifier tubes with adequate insulation between the cathodes and heaters. Special rectifier tubes having two diodes heated by a single heater coil have been developed for voltage-doubler service. One such tube is the 25Z5.

Polyphase rectifier circuits are shown in Table 14-I, page 571.

14-3. Choice of Rectifier Tubes and Circuit.—The choice of rectifier tubes and circuit involves a number of factors, of which the following merit special consideration:

1. *Type of Rectifier Tubes.*—The low voltage drop of mercury vapor tubes results in higher efficiency of operation than with high-vacuum tubes and, because the drop is practically constant throughout the working range of current, the voltage regulation is better than with high-

[1] WAIDELICH, D. L., *Electronics*, May, 1941, p. 28.

[2] COCKCROFT, J. D., and WALTON, E. T. S., *Proc. Roy. Soc. (London)*, **136**, 619 (1932).

vacuum tubes. For a tube of given size and given filament power the mercury vapor tube will pass higher current than the high-vacuum tube. The peak voltage of the mercury vapor rectifier is limited, however, because of danger of glow and subsequent breakdown into an arc during the half of the cycle when the anode is negative. The hot-cathode mercury vapor rectifier is much more subject to damage due to overload or low cathode temperature, and adequate provision must be made to ensure that the anode voltage will not be applied before the cathode has reached operating temperature. Another disadvantage of mercury vapor rectifiers is that anode current does not flow unless the applied voltage exceeds the normal tube drop. The resulting abrupt changes of current at the beginning and end of the firing period produce objectionable transient and high-frequency disturbances. Provision must be made to prevent these disturbances from affecting other apparatus or from being transferred to the a-c line or being radiated.

Grid-controlled arc rectifiers are advantageously used in large power supplies for radio transmitters.[1] The use of phase control affords a simple method of reducing the direct voltage during starting, in order to prevent large charging surges in the filter condensers and damage to the transmitter tubes. The grids may be used to provide high-speed automatic regulation of the direct voltage. Grid control also makes possible interruption of the anode current within one cycle of the supply frequency when a short circuit or other overload occurs. The circuit may be designed so that the energy stored in the filter is returned to the line by inverter action, instead of being dissipated in a flashover arc. The likelihood of arcback (current flow in the reverse direction) in mercury arc rectifiers is reduced by the use of grids.

Grid control affords a convenient means of keying in radio telegraph transmitters.

2. *Transformer Design.*—The design of the transformer or transformers is dependent upon the type of circuit and tubes. Factors that must be taken into consideration are the primary and secondary volt-amperes, the secondary voltages, and insulation required between windings.

3. *Efficiency.*—The efficiency of operation depends upon the transformer and tube efficiencies. Tube efficiency, as already stated, is greater for mercury arc rectifiers than for high-vacuum rectifiers. The efficiency of both the tubes and the transformer increases with the percentage of the cycle during which anode current flows, so that for high efficiency it is advisable to design the filter so that current flows in each tube during its maximum firing period. In some rectifier circuits, cur-

[1] DURAND, S. R., and KELLER, O., *Proc. I.R.E.*, **25**, 570 (1937). See also footnote 2, p. 519.

rent flows in only one direction in each leg of the transformer. The resulting d-c component of flux saturates the core, reducing the efficiency by increasing the magnetizing current and hysteresis loss and by introducing harmonics into the secondary voltage.

4. *Filament Supply.*—In certain rectifier circuits the cathodes are at the same potential and may be operated from the same heating transformer; in others, on the other hand, it is necessary to use either separate filament supplies or heater-type tubes with adequate insulation between cathodes and heaters.

5. *Ripple Frequency and Amplitude.*—The size and cost of a filter capable of reducing the ripple factor to a given value at the load decrease with increase of ripple frequency and with decrease of ripple amplitude at the filter input. The allowable ripple factor at the load depends upon the type of service for which the power supply is designed. In the microphone circuit of a radio transmitter the ripple factor must be less than 0.003 per cent. In audio-frequency amplifiers not followed by high-gain radio-frequency amplifiers, ripple factors may range from 0.005 to 0.05 per cent without causing objectionable hum, and even 1 per cent may sometimes be tolerated. About 1 per cent represents the upper limit in voltages used for cathode-ray oscillographs. In control devices of the type discussed in Chaps. 12 and 13, the allowable ripple may sometimes be so high that the smoothing filter may be omitted.

In general, the ripple frequency increases and the ripple amplitude decreases with increase of the number of phases. The wave form of the rectified voltage for different numbers of phases is shown by the solid lines in Fig. 14-6.

6. *Peak Inverse Tube Voltage.*—Since the attainable output voltage of a rectifier is limited by the peak inverse voltage that the tubes can withstand, the magnitude of the peak inverse voltage to which the tubes are subjected is an important consideration in the choice of both tubes and rectifier circuits.

7. *Peak and Average Tube Currents.*—The life of rectifier tubes is dependent upon both the peak and average tube currents. For this reason peak and average tube currents are factors that affect the choice of tubes and rectifier circuits.

8. *Voltage Regulation.*—The voltage regulation of a power supply is dependent not only upon the type of rectifier tube used but also upon the type of circuit and upon the resistance of the various circuit elements.

14-4. Characteristics of Rectifier Circuits.—Table 14-I lists important information concerning some of the more common types of rectifier circuits, with inductive load.[1] It should be noted that all voltage, current, and volt-ampere values in this table are given with respect to the

[1] Armstrong, R. W., *Proc. I.R.E.*, **19**, 78 (1931).

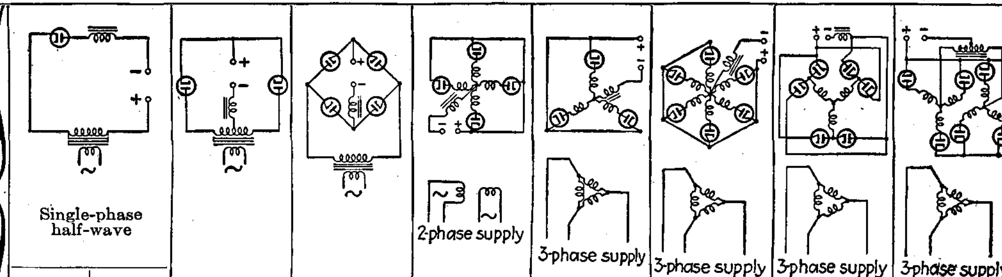
average values of direct voltage, direct current, and d-c power, respectively. The values listed are derived under the assumption that tube and transformer voltage drops are negligible and that the transformer is perfect. In deriving the current relations it is furthermore assumed that each tube conducts during the entire time in which the induced voltage in its phase is positive and exceeds the positive voltages of the other phases. An examination of this table discloses the advantages of certain circuits over others.

The single-phase half-wave rectifier has as its chief merits its simplicity and low cost. These advantages are offset, however, by the high ripple amplitude and low ripple frequency, which necessitate the use of a comparatively expensive smoothing filter, by low voltage output, and by low transformer efficiency. The low transformer efficiency results partly from the high effective alternating voltages and currents and partly because the rectified direct current flows through the secondary and saturates the core. Core saturation increases the magnetizing-current and the hysteresis loss and introduces harmonics into the secondary voltage. It is impractical to use sufficient filter inductance to cause current to flow during the whole cycle. In fact, in order to raise the output voltage and assist in removing ripple, a condenser is usually shunted across the output of the rectifier. This reduces still further the time during which secondary current flows, and so decreases the transformer efficiency. The single-phase half-wave rectifier is now used comparatively little.

The single-phase full-wave rectifier is the most commonly used type of rectifier at voltages below 1000 volts and for currents of 1 amp or less. The ripple frequency is twice that of the a-c supply, and the ripple amplitude is sufficiently small so that smoothing is not difficult. The circuit is simple and requires only one full-wave rectifier tube. The single-phase bridge circuit subjects the tubes to a lower peak inverse voltage than the single-phase full-wave circuit and requires smaller transformer primary volt-ampere rating. It has the disadvantages that three separate filament transformers must be used, that the circuit requires three tubes (one full-wave and two half-wave rectifiers) instead of one, and that the efficiency is somewhat lower because of increased tube loss.

The ripple amplitude is much lower in polyphase rectifiers, and the ripple frequency higher than in single-phase rectifiers. The three-phase star connection has the disadvantage that current flows in each leg of the secondary in one direction only, causing core saturation. This difficulty does not occur in the other polyphase circuits that are listed in Table 14-I. The three-phase star full-wave connection requires low transformer volt-ampere rating and subjects the tubes to low peak inverse

TABLE 14-I.—CHARACTERISTICS OF RECTIFIER CIRCUITS USED WITH INDUCTANCE LOAD
(See first paragraph of Sec. 14-4 for explanation of items listed.)



Type of circuit	Single-phase half-wave		Single-phase full-wave	Single-phase bridge	Two-phase full-wave (four-phase star)	Three-phase star	Three-phase double star (six-phase star full-wave)	Three-phase star full-wave	Double three phase with balance coil
	Inductance in series with load	Resistance load, without choke							
Secondary volts per leg ÷ direct voltage	*	2.22	1.11	1.11	0.785	0.855	0.740	0.428	0.855
Primary volts per leg ÷ direct voltage	*	2.22	1.11	1.11	0.785	0.855	0.740	0.428	0.855
Primary current per leg ÷ direct current	*	1.57	1.0	1.0	0.707	0.471	0.577	0.816	0.408
Secondary current per leg ÷ direct current	*	1.57	0.707	1.0	0.500	0.577	0.408	0.816	0.289
Secondary kva ÷ d-c watts	*	3.49	1.57	1.11	1.57	1.48	1.81	1.05	1.48
Primary kva ÷ d-c watts	*	3.49	1.11	1.11	1.11	1.21	1.28	1.05	1.05
Peak inverse tube voltage ÷ direct voltage	*	3.14	3.14	1.57	2.22	2.09	2.09	1.05	2.42
R-m-s tube current ÷ direct current	*	1.57	0.707	0.707	0.50	0.577	0.408	0.577	0.289
Peak tube current ÷ direct current	*	π	1.00	1.00	1.00	1.00	1.00	1.00	0.50
Fundamental ripple frequency f_r in terms of supply frequency f	f	f	$2f$	$2f$	$4f$	$3f$	$6f$	$6f$	$6f$
Ripple factor ρ_1	9.18*	1.11	0.472	0.472	0.106	0.177	0.040	0.040	0.040
Second-harmonic ripple factor ρ_2	1.25*	0.472	0.094	0.094	0.021	0.040	0.01	0.01	0.01
Third-harmonic ripple factor ρ_3	1.09*	0	0.040	0.040	0.011	0.018	0.004	0.004	0.004

* Values change with ratio of load resistance to series inductance.

voltage but requires four filament transformers. In the double three-phase rectifier with balance coil, induced voltages in the balance coil cause equal currents to flow simultaneously in two phases, one in each wye. This results in low average and peak tube currents, but the peak inverse tube voltage is comparatively high. Many other connections are possible, some of which give increased efficiency or improved voltage regulation.[1]

14-5. Current Wave Form.—In Fig. 14-4 are shown curves of positive anode voltage and of load current in a single-phase half-wave rectifier, in

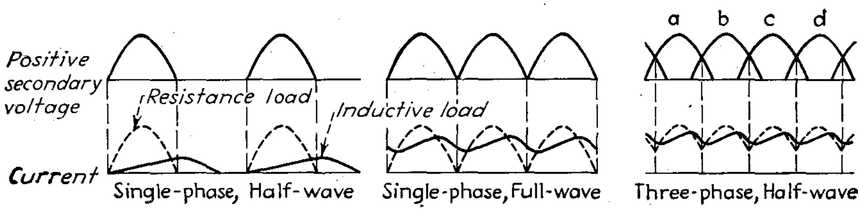


FIG. 14-4.—Curves of positive secondary voltage and load current for single-phase and three-phase rectifiers.

a single-phase full-wave rectifier, and in a three-phase half-wave rectifier, with resistance and inductance loads, under the assumption that tube and transformer voltage drops are zero. When the load[2] is nonreactive, current flows in each phase only when that phase has the highest induced voltage, and the instantaneous current is proportional to the instantaneous voltage. The current waves are shown by the dotted curves. (The effect of the constant voltage drop in an arc rectifier tube is to prevent the flow of current when the induced voltage is less than the normal tube drop.) The presence of inductance in the load tends to prevent the current from building up and dying down, causing the current waves to have the forms shown by the solid curves. If the inductance is made sufficiently large, the load current becomes practically constant, and so the current in each phase is very nearly a square-topped pulse.

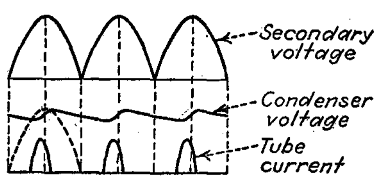


FIG. 14-5.—Approximate wave form of load voltage and tube current in a full-wave single-phase rectifier with resistance-capacitance load.

When a resistance load is shunted by a condenser, current flows through the rectifier only when the induced voltage exceeds the condenser voltage. If the condenser is large and the resistance high, current flows just for a short time at the peaks of the induced voltage. The approximate forms of typical voltage and current waves with resistance-capacitance load are shown in Fig. 14-5.

[1] ARMSTRONG, *loc. cit.*

[2] Load on the rectifier is construed here to include the filter.

When the load is such that current flows in each phase during the entire time in which its induced voltage exceeds those of the other phases (inductive or nonreactive load), the voltage wave across the rectifier load consists of the portions of the phase voltage waves lying between the intersections of waves of the different phases. The voltage

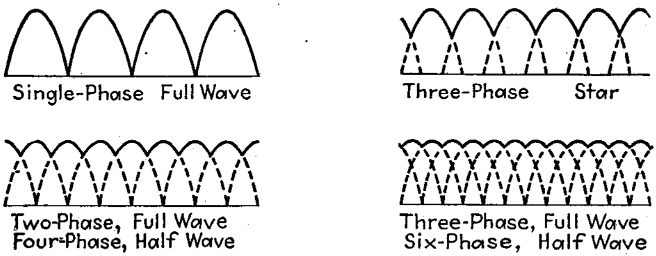


FIG. 14-6.—Wave form of load voltage for resistance or inductance load.

waves for different numbers of phases are shown by the solid curves of Fig. 14-6. The form of the corresponding current wave can be determined by analyzing the voltage wave into its Fourier components, dividing each component by the impedance of the load (including the filter) at the corresponding frequency, and adding the component currents obtained in this manner.[1] If the presence of capacitance in the rectifier

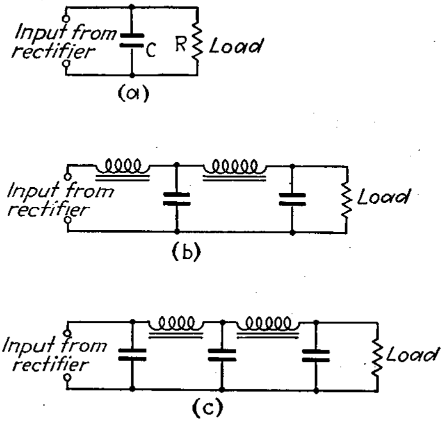


FIG. 14-7.—Three types of smoothing filters: (*a*) condenser; (*b*) choke-condenser (L-section); (*c*) choke-condenser with condenser input (pi-section).

load causes each tube to pass current during only a portion of the possible firing period, then the load voltage during the nonconducting portion of the cycle depends upon the parameters of the load. Under these conditions the determination of the current wave form becomes difficult and necessitates the use of successive approximations[1] or other involved

[1] STOUT, M. B., *Elec. Eng.*, **54**, 977 (1935).

analyses, the results of some of which will be presented in usable form in Sec. 14-7.

Because the characteristics of rectifier circuits depend upon the type of smoothing filter by which they are followed, it is best to discuss the characteristics and design of smoothing filters before discussing further the choice of rectifier circuits for power supplies.

14-6. Smoothing Filters.—Three types of filters are commonly used in power supplies. These are (1) a condenser shunted across the rectifier output, as in Fig. 14-7*a*; (2) one or more sections, each of which consists of a series inductance followed by a shunt capacitance, as in Fig. 14-7*b*; and (3) a combination of 1 and 2, as in Fig. 14-7*c*. The use of a single condenser across the rectifier output is of principal value when the load current is small, as in voltage supplies for cathode-ray oscillographs. The multisection inductance-condenser filter is in general the most satisfactory, since it results in good voltage regulation, low peak anode current, and high tube and transformer efficiency. The multisection inductance-condenser filter with condenser input is used mainly in small single-phase power supplies. It has a higher smoothing factor than a multisection filter without input condenser that has the same total inductance and capacitance, and it results in higher direct voltage. The direct voltage at light load approximates the crest alternating secondary voltage E_m, as compared with $0.318E_m$ in a half-wave single-phase rectifier without the input condenser and $0.636E_m$ in a full-wave single-phase rectifier without the input condenser (see Table 11-I). The voltage regulation is greater, however, than when the condenser is omitted. Other disadvantages resulting from the use of a condenser in the input to the filter are increased peak anode current, particularly in mercury vapor tubes, and higher peak inverse voltage. The peak inverse voltage is the crest secondary voltage plus the condenser voltage at the instant of crest secondary voltage. Since the condenser voltage approximates the crest secondary voltage under light load, the peak inverse voltage may approach twice the crest secondary voltage.

The chokes of the circuits of Figs. 14-7*b* and *c* may be replaced by resistances. The d-c resistance of resistance-capacitance filters is much higher than that of inductance-capacitance filters using the same size condensers and having the same smoothing factor. Hence the voltage regulation, the loss in direct voltage, and the heat dissipation are greater. Furthermore, resistance-capacitance filters result in higher peak anode current and lower transformer efficiency than do inductance-capacitance filters. The development of small, low-cost electrolytic condensers of high capacitance has, however, made feasible the use of resistance-condenser filters when the requirements of low cost, lightness, and small size outweigh the desirability of high quality.

14-7. Design of Condenser Filters.—The currents and the condenser voltage in a rectifier with capacitive load consist of periodic transients of such form that the currents and voltage at the end of each transient are the same as at the beginning of the transient. Circuits of this type have been treated analytically by Lewis, Roberts, and Waidelich[1] and graphically by Schade.[2] Both methods of analysis are of necessity so long and involved as to make undesirable their inclusion in this text. The generalized curves obtained by means of these analyses are, however, of great value in the design of rectifiers with filters of the form of Fig. 14-7*a* and *c*. For circuits using high-vacuum tubes, the analyses are based upon the assumption, justified experimentally, that a high-vacuum rectifier tube may be replaced in an equivalent circuit by a fixed resistance in series with a synchronous switch. The value of the equivalent resistance depends upon whether the peak, average, or effective values of circuit voltages and currents are to be found. The peak resistance $\hat{r}_p$ of a rectifier diode is defined as the ratio of the peak value of the diode plate voltage to the corresponding peak value of the diode plate current. The average resistance $\bar{r}_p$ of a diode is defined as the ratio of the average plate voltage during the conduction period to the average plate current during the conduction period. The effective resistance $|r_p|$ of a diode is defined as the ratio of the diode plate dissipation to the square of the effective plate current. For high-vacuum diodes the ratios of these three values of resistance are essentially independent of current amplitude and circuit constants, and are given to an accuracy of 5 per cent by the following relations:

$$|r_p| = 1.075\hat{r}_p \qquad \bar{r}_p = 1.14\hat{r}_p \tag{14-1}$$

Schade's curves for ripple factor and r-m-s plate current involve the effective equivalent resistance $|r_p|$. The error resulting from the use of $\hat{r}_p$ in place of $|r_p|$, however, does not exceed the probable error in reading the curves. For the sake of simplification of design procedure, therefore, the crest resistance $\hat{r}_p$ will be used in place of $|r_p|$ in the remainder of this section. Figure 14-8 shows curves from which the crest resistance of high-vacuum rectifier tubes may be found.

For circuits using arc rectifier tubes, the analyses are based upon an equivalent circuit in which the diode is replaced by a small fixed resistance and a counter-e.m.f. equal to the firing voltage of the tube, in series with a synchronous switch. The three values of equivalent resistance of arc

[1] LEWIS, W. B., *Wireless Eng.*, **9**, 487 (1932); ROBERTS, N. H., *Wireless Eng.*, **31**, 351 and 423 (1936); WAIDELICH, D. L., *Trans. Am. Inst. Elec. Eng.*, **61**, 1161 and 1389 (1941); *Proc. I.R.E.*, **29**, 554 (1941) and **30**, 535 (1942).

[2] SCHADE, O. H., *Proc. I.R.E.*, **31**, 341 (1943).

rectifier tubes are identical and are equal to approximately 4 ohms in small rectifier tubes. The counter-e.m.f., which has a value of approximately 10 volts in small rectifier tubes, may be taken into account in rectifier design by subtracting this voltage from the direct output voltage computed under the assumption that the counter-e.m.f. is zero. The resulting error is small when the crest alternating voltage exceeds 50 volts.

Figure 14-9 shows curves of the ratio of direct output voltage E_{dc} to the crest alternating voltage E_m for the half-wave single-phase rectifier

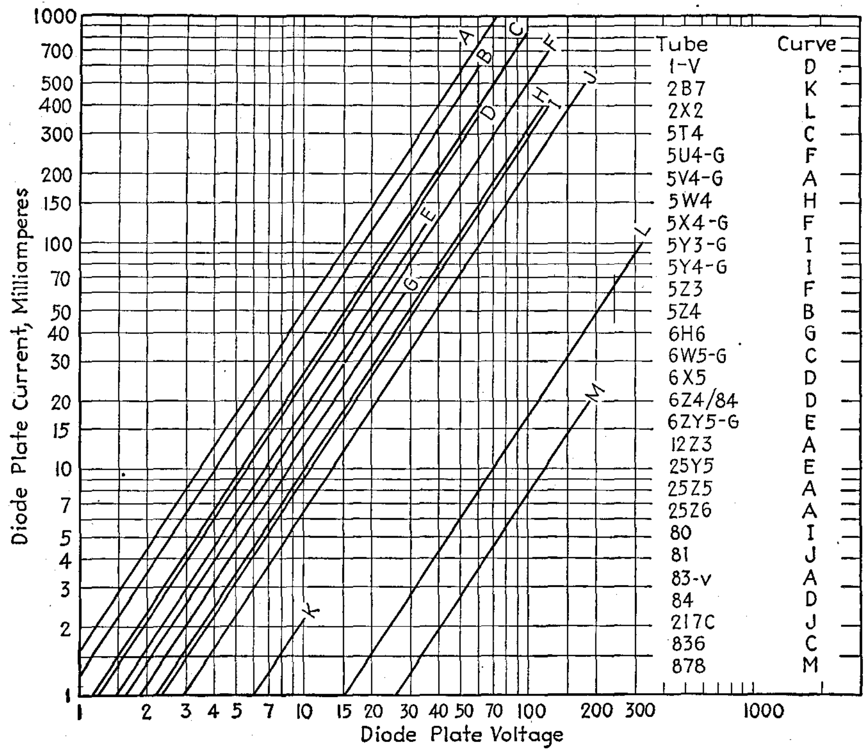


FIG. 14-8.—Average static plate characteristics of high-vacuum power rectifier tubes.

with condenser filter.[1] Figures 14-10 and 14-11 show similar curves for the full-wave single-phase rectifier and the full-wave voltage doubler.[2] Curves of the ratios of effective and peak plate currents to the direct load current are given in Fig. 14-12 and curves of ripple factor in Fig. 14-13. In Figs. 14-9 to 14-13, R is the load resistance, ω is 2π times the supply frequency, $\hat{R}_s = \hat{r}_p + R_s$, and $\bar{R}_s = \bar{r}_p + R_s$, where R_s is the sum

[1] SCHADE, *loc. cit.*

[2] Curves for the half-wave voltage doubler have been given by D. L. Waidelich, *Proc. I.R.E.*, **30**, 535 (1942).

of all circuit resistances in series with each plate, including line resistance, transformer winding resistances, and series resistance added to limit the peak diode plate current.

Since the desired output voltage can ordinarily be obtained by the use of an input transformer, it is usually logical to base the rectifier

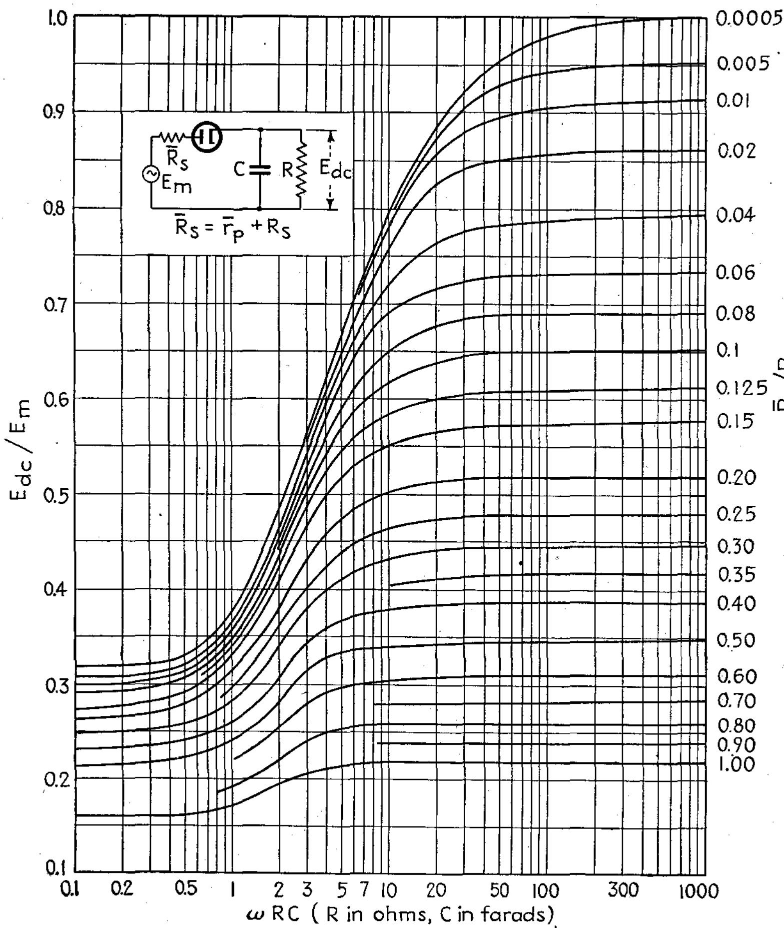


FIG. 14-9.—Ratio of direct output voltage to crest alternating voltage, as a function of ωRC, for the half-wave single-phase rectifier with condenser filter.

design upon allowable ripple factor, rather than upon a specified value of the voltage ratio E_{dc}/E_m. If no input transformer is used, as is often true in voltage doublers, some increase of C above the value necessary to give the required filtering may have to be made in order to obtain the desired output voltage. In order to ensure good voltage regulation, ωRC at full load must lie to the right of the knee of the appropriate voltage-ratio curve of Fig. 14-9, 14-10, or 14-11. The following design procedure is convenient:[1]

1. From the given load current, the average diode current per plate is estimated. In the half-wave rectifier and the full-wave voltage doubler, the average diode current per plate is equal to the direct load current. In the full-wave rectifier it is equal to half the load current.

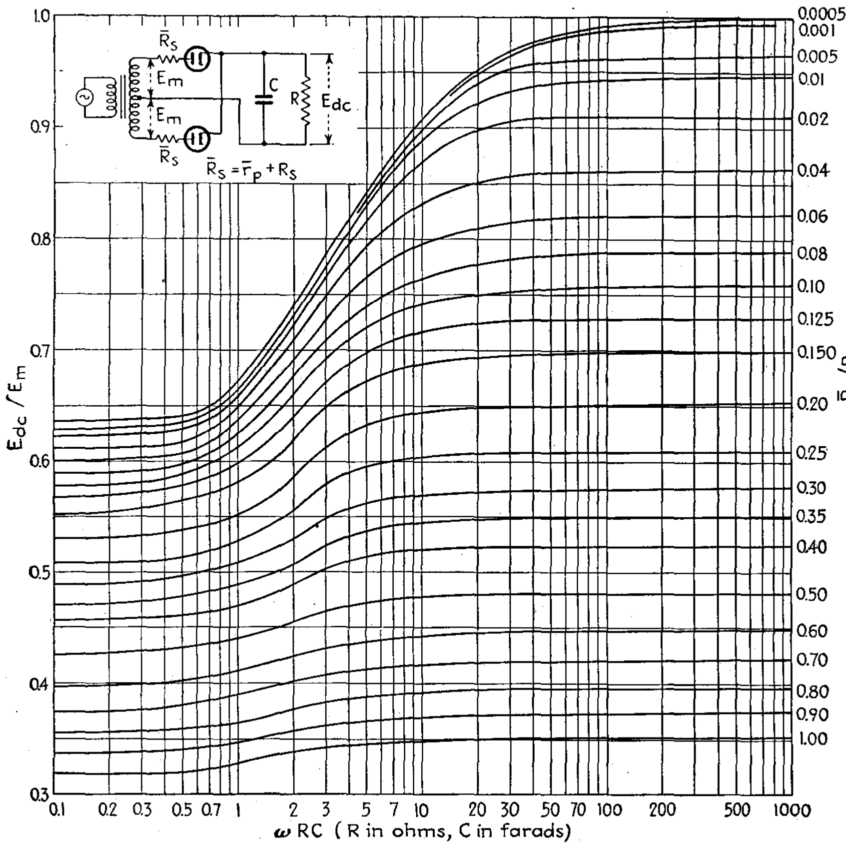


FIG. 14-10.—Ratio of direct output voltage to crest alternating voltage, as a function of ωRC, for the full-wave single-phase rectifier with condenser filter.

2. The peak diode current is assumed to be four times the average diode current.

3. (*a*) When a high-vacuum diode is used, the peak diode voltage corresponding to the assumed peak diode current is determined from the curve of Fig. 14-8 corresponding to the type of tube used. The ratio of the peak diode voltage to the peak diode current gives the peak diode resistance, $\hat{r}_p$. $\bar{r}_p$ is found by the use of Eq. (14-1). The estimated series resistance of the transformer windings

[1] This is essentially the procedure suggested by Schade.

(usually about 125 ohms for each half of the secondary) plus any additional series resistance is added to $\hat{r}_p$ and $\bar{r}_p$ to obtain $\hat{R}_s$ and $\bar{R}_s$. The ratios $\hat{R}_s/R$ and $\bar{R}_s/R$ are then found.

(b) When an arc rectifier tube is used, $\hat{R}_s$ and $\bar{R}_s$ are assumed to be equal to $R_s + 4$ ohms in finding the ratios $\hat{R}_s/R$ and $\bar{R}_s/R$.

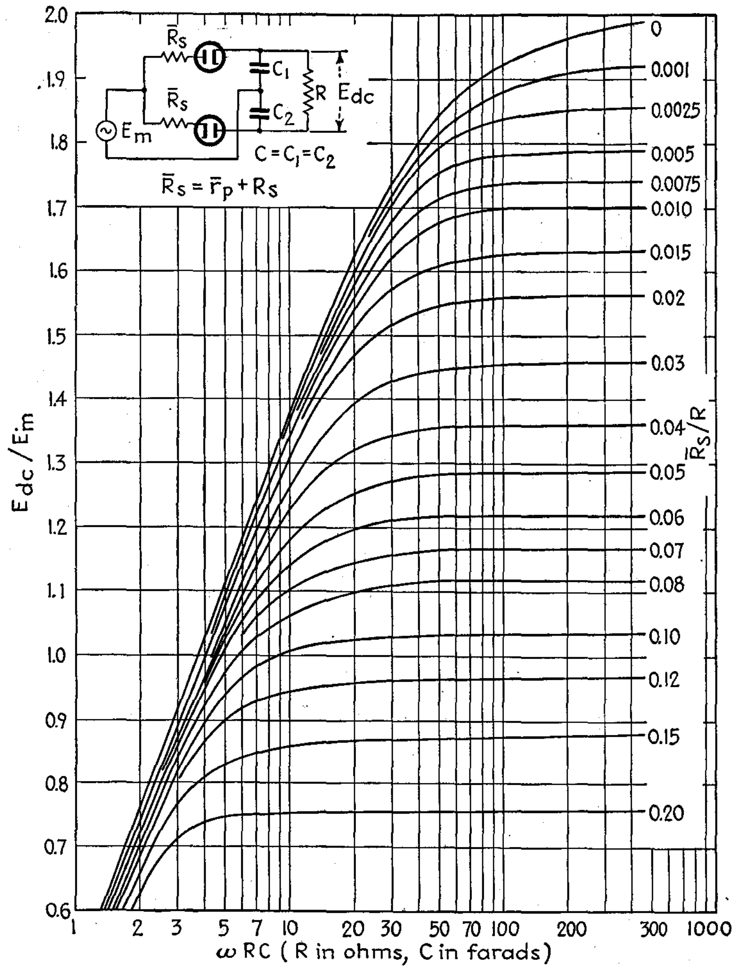


FIG. 14-11.—Ratio of direct output voltage to crest alternating voltage, as a function of ωRC, for the full-wave voltage doubler.

4. From the curve of Fig. 14-13 corresponding to the value of R_s/R found above, a value of ωRC may be found that will reduce the ripple factor to a value equal to or less than the allowable value.

5. When a high-vacuum rectifier tube is used, a more accurate tentative value of peak plate current is now found from Fig. 14-12, and steps 3 and 4 are repeated. Figure 13-12 will then ordinarily give a value of peak plate current that does not

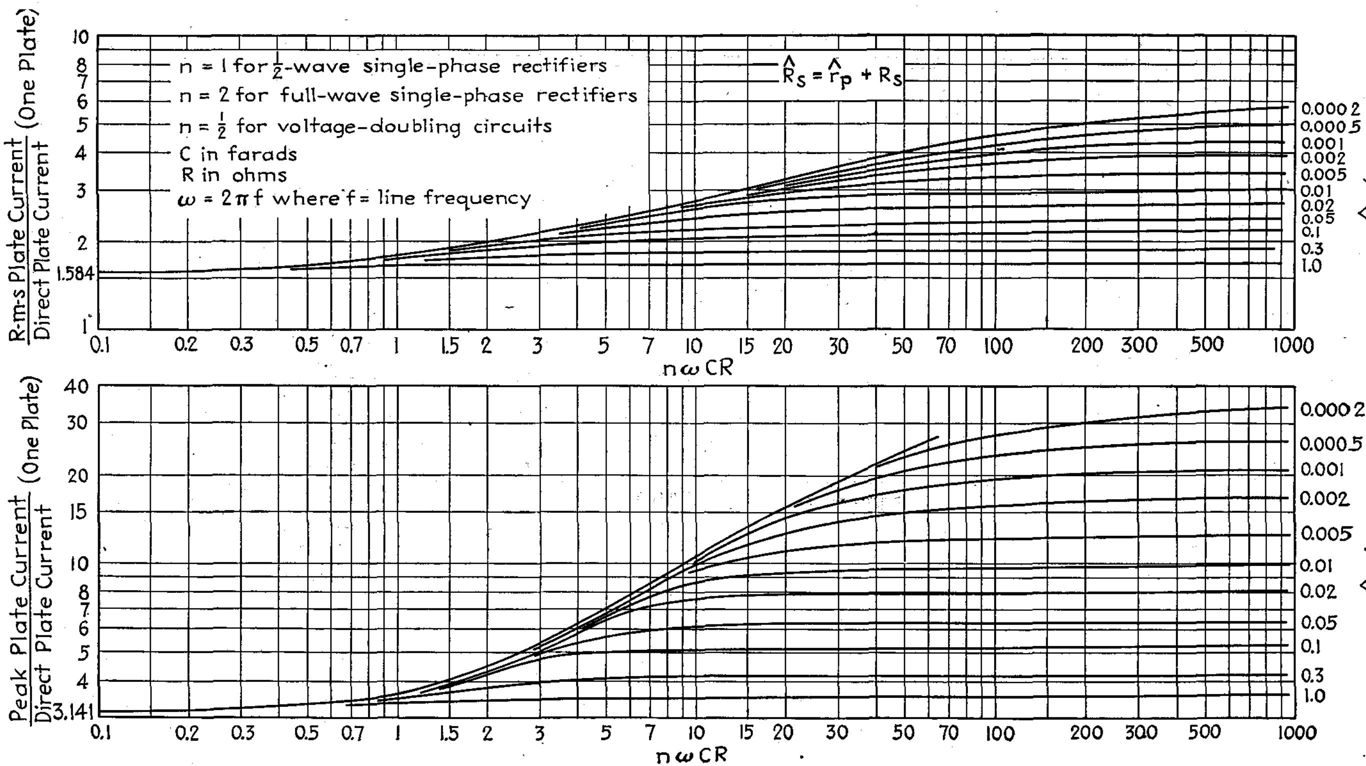


FIG. 14-12.—Ratio of r-m-s plate current to direct plate current and of peak plate current to direct plate current as a function of $n\omega CR$.

differ greatly from the second tentative value. If desired, a third approximation may be made, but this is ordinarily unnecessary.

6. The ratio E_{dc}/E_m may now be found from the appropriate curve of Figs. 14-9, 14-10, or 14-11. If necessary, the value of ωRC may be raised in order to bring it within the range corresponding to the flat portion of the voltage-ratio curve. When no input transformer is used, E_m is fixed by the supply voltage, and E_{dc} may be found from the voltage ratio. When an arc rectifier tube is used, this value of E_{dc} should be reduced by 10 volts. If E_{dc} is specified, it is usually

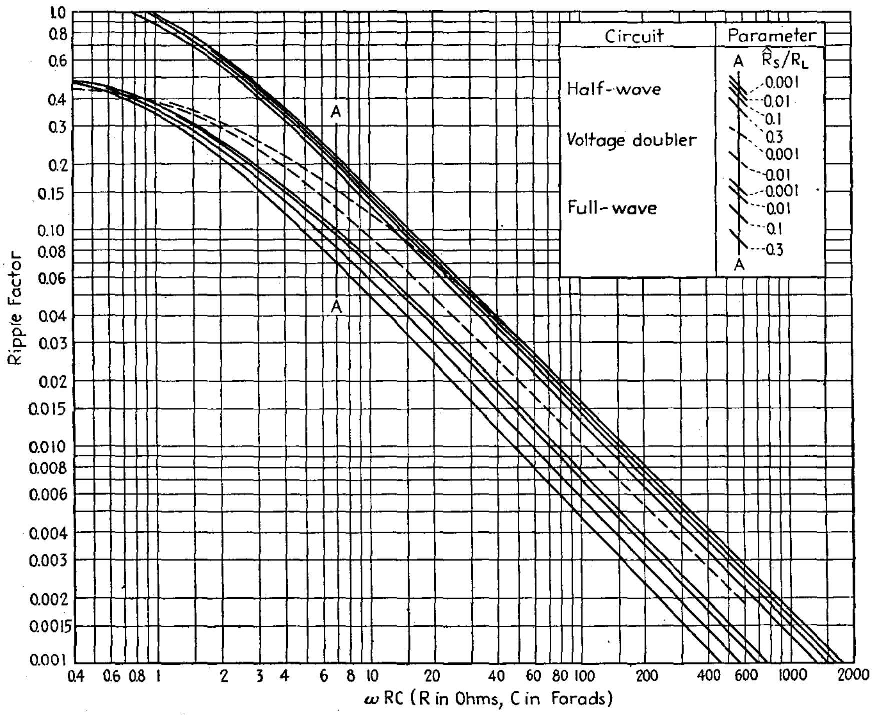


FIG. 14-13.—Curves for the determination of ripple factor of rectifiers with condenser filter.

necessary to use an input transformer of the correct turn ratio to give the desired output voltage.

7. After the input transformer has been chosen, the winding resistance may be determined. If the value of R_s differs considerably from the assumed value, steps 4 to 6 should be repeated.

8. The peak and average plate currents, found from Fig. 14-22, and the peak inverse voltage should be compared with the maximum rated values for the tube in order to make sure that they are not excessive. The peak inverse voltage approaches twice the crest alternating voltage E_m at light load, and should be assumed to have this value.

14-8. Illustrative Problem.—The following problem illustrates the use of the foregoing procedure. The numbering of the various steps of the analysis corresponds to that used in Sec. 14-7. Assume that the desired rectifier is to have an output voltage of 500 volts at full-load current of 50 ma, with a ripple factor of not to exceed 0.01, and that it is to be operated on a 120-volt, 60-cycle line. Use a full-wave circuit with a 5V4-G tube.

1. Assume that R_s consists entirely of transformer winding resistance, and is equal to 125 ohms. $R = 500/0.05 = 10{,}000$ ohms. Average plate current per plate $= 0.05/2 = 0.025$ amp.

2. Assumed peak plate current $= 4 \times 0.025 = 0.1$ amp.

3. From Fig. 14-8 the peak plate voltage is found to be 16 volts. $\hat{r}_p = 16/0.1 = 160$ ohms. $\hat{R}_s = 125 + 160 = 285$ ohms. $\hat{R}_s/R = 285/10{,}000 = 0.0285$.

4. For a ripple factor of 0.01 and this resistance ratio, Fig. 14-13 gives a value of approximately 66 for ωRC.

5. Figure 14-12 shows that the ratio of peak plate current to average plate current is approximately 9 (note that n is 2). A second tentative value of peak plate current is therefore $9 \times 0.025 = 0.225$ amp. Figure 14-8 shows the corresponding peak plate voltage to be 28 volts. This makes $\hat{r}_p$ equal to 124 ohms, and $\hat{R}_s$ equal to 249 ohms. The new resistance ratio $\hat{R}_s/R$ is 0.0249, which does not differ sufficiently from the value first assumed to make an appreciable difference in the values of ωRC and the current ratio read from the curves.

$$\bar{R}_s/R = (125 + 124 \times 1.14)/10{,}000 = 0.0266.$$

6. Figure 14-10 shows that, when $\bar{R}_s/R$ is 0.0266 and ωRC is 66, the voltage ratio E_{dc}/E_m has the value 0.895 and that this value of ωRC corresponds to a point on the flat portion of the curve, ensuring good voltage regulation.[1] The r-m-s voltage of each half of the transformer secondary should be $0.707 \times 500/0.895 = 395$ volts. The filter condenser should have the value $66/(120\pi \times 10{,}000) = 17.5\ \mu$f.

7. The peak and average plate currents and the peak inverse voltage lie within the allowable values for the tube.

14-9. Choke-condenser (L-section) Filters.—Examination of Fig. 14-13 shows that, unless the load current is small, the capacitance required to give adequate filtering with the simple condenser filter at values of ripple factor less than 0.01 is likely to be so large as to be impractical. Filters of the forms of those of Fig. 14-7*b* or *c* must then be used.

From Fig. 14-14 it can be seen that the smoothing factor for a single-section choke-condenser filter is equal to the total input impedance of the filter divided by the impedance of the parallel combination of the condenser and the load resistance. With values of C that are sufficiently

[1] Unless R_s is very small, little error results from the use of $\hat{R}_s$ in place of $\bar{R}_s$ in finding E_{dc}/E_m.

large to give adequate filtering, the vector sum of the condenser and load admittance is approximately equal to the condenser admittance alone. The resistance of the choke, which is small in comparison with the inductive reactance, may be neglected. The smoothing factor is approximately

$$\alpha = \frac{x_l - x_c}{x_c} = \omega^2 LC - 1 \tag{14-2}$$

The required value of the product LC to give a smoothing factor of value α is

$$LC = \frac{\alpha + 1}{\omega^2} \tag{14-3}$$

The derivation of an exact expression for the smoothing factor of a multisection filter, such as that of Fig. 14-15, is complicated, but the

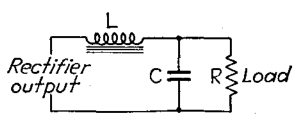


FIG. 14-14.—Single-section choke-condenser (L-section) filter.

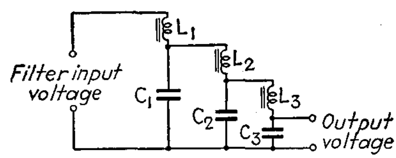


FIG. 14-15.—Three-section choke-condenser filter.

fact that adequate filtering calls for relatively large values of L and C makes possible a simplified approximate derivation.[1] With typical values of L and C the ratio of x_l to x_c is about 20. Thus, although the condensers in all but the last section are shunted by a combination of other inductances and capacitances, the impedances of these shunting combinations will be from ten to twenty times the condenser reactances and may, with little error, be assumed to be infinite. As in the single-stage filter, the effect of the load on the action of the last section may also be neglected without great error. Under these assumptions, the smoothing factor for each section of a filter of the type shown in Fig. 14-15 is given by Eq. (14-2), and the smoothing factor for a filter of n similar sections is

$$\alpha = \left(\frac{x_l}{x_c} - 1\right)^n = (\omega^2 LC - 1)^n \tag{14-4}$$

The approximations made in the derivation cause the value of α given by Eq. (14-4) to be somewhat low. The error resulting from the use of this approximate expression therefore makes the design conservative. If the different sections of the filter are dissimilar, the smoothing factor is

$$\alpha = (\omega^2 L_1C_1 - 1)(\omega^2 L_2C_2 - 1)(\omega^2 L_3C_3 - 1) \cdots \tag{14-4a}$$

[1] LEE, REUBEN, *Elec. J.*, **29**, 186 (1932).

It may be shown, however, that for a given total capacitance and inductance the highest smoothing factor is obtained when all sections are alike.[1] Hence Eq. (14-4) may generally be used. Equation (14-4) may be converted into the form

$$LC = \frac{\sqrt[n]{\alpha_1} + 1}{\omega_r^2} = 0.0253\,\frac{\sqrt[n]{\alpha_1} + 1}{f_r^2} \tag{14-5}$$

where ω_r is the angular ripple frequency and α_1 is the smoothing factor at fundamental ripple frequency.

The required value of α_1 may be found from the relation

$$\alpha_1 = \frac{\rho_1}{\rho_1'} \tag{14-6}$$

in which ρ_1 has the value given in Table 14-I and ρ_1' is the allowable ripple factor at the load. By means of Eq. (14-5) it is then possible to determine the product LC for a filter of any number of sections. It should be noted that the value of L given by Eq. (14-5) is the actual inductance under load and not the nominal value listed by the manufacturer of the choke.

14-10. Relative Values of Inductance and Capacitance.—The choice of the relative values of L and C, the inductance and capacitance of each stage, is affected by a number of considerations. In the first place, the inductance in the first stage must be adequate to ensure that the rectifier passes current throughout one cycle of ripple frequency. The limiting value of L for which the plate current just falls to zero can be determined as follows. Figure 14-16 shows that the current just touches zero if the amplitude of the alternating component of the filter input current (rectifier ripple-current amplitude) is equal to the average (direct) current. The average current is

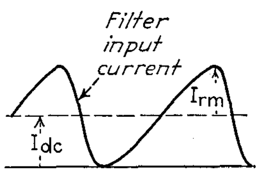


FIG. 14-16.—Filter input current (rectifier current) wave and direct output current when the input current just falls to zero.

$$I_{dc} = \frac{E_{dc}}{R_t} \tag{14-7}$$

in which R_t is the resistance of the load plus that of the chokes. If ripple harmonics are assumed to be negligible, the ripple amplitude I_{rm} is equal to the amplitude of the fundamental component of ripple current I_{r1}. If the additional assumption is made that the condenser reactance is negligible in comparison with the reactance of the first choke, the ripple-current amplitude is

$$I_{rm} = I_{r1} = \frac{E_{r1}}{\omega_r L_1} \tag{14-8}$$

[1] LEE, *loc. cit.*

Combining Eqs. (14-7) and (14-8) gives for the theoretical limiting value of first choke inductance, below which current flows in each anode during less than the entire ripple period

$$L_1 = \frac{E_{r1}R_t}{\omega_r E_{dc}} = 0.159\,\frac{\rho_1 R_t}{f_r} \tag{14-9}$$

For the single-phase full-wave rectifier operating on a 60-cycle source, Eq. (14-9) reduces to

$$L_0 = 0.159\,\frac{0.667\,R_t}{120} = \frac{R_t}{1132} \tag{14-10}$$

The interruption of current when the inductance of the first filter section is insufficient to maintain constant current through the rectifier occurs because the filter condensers keep the voltage across the input to the filter higher than the instantaneous secondary voltage during the portion of the cycle in which the current is zero. Therefore the filter input voltage is higher during this portion of the cycle than it would be if current were flowing, and so the average output voltage is higher. The output voltage rises with increase in the portion of the cycle during which current does not flow in the rectifier. Hence, if good voltage regulation is essential, the first-stage filter inductance must exceed the critical value given by Eq. (14-10) over the whole range of load current. Curve a of Fig. 14-17 shows the load voltage of a single-phase full-wave rectifier as a function of inductance of the first filter section. It is taken from a paper by Dellenbaugh and Quinby,[1] who found that for the inductance below which the voltage rises greatly or "soars," the relation between the first-stage filter inductance and the load resistance is given by the equation

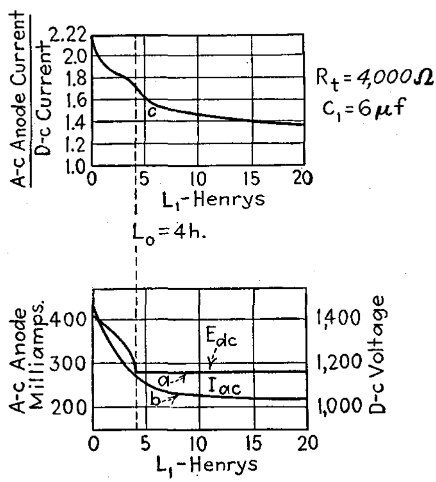


FIG. 14-17.—Experimentally determined performance curves for a single-phase full-wave rectifier with choke-condenser filter.

$$L_0 = \frac{R_t}{1000} \tag{14-11}$$

The agreement between the theoretical Eq. (14-10) and the experimental Eq. (14-11) is as close as could be expected under the assumptions

[1] DELLENBAUGH, F. S., JR., and QUINBY, R. S., *QST*, February, 1932, p. 14; March, 1932, p. 27; April, 1932, p. 33. For a discussion of critical inductance in filters for grid-controlled rectifiers, see W. P. OVERBECK, *Proc. I.R.E.*, **27**, 655 (1939).

made in the derivation of Eq. (14-10), the greatest error probably resulting from failure to take ripple harmonics into account. For conservative design the value indicated by Eq. (14-11) should be used for a full-wave single-phase rectifier. Under the assumption that the relative error is not dependent upon the number of phases, the value of first choke inductance indicated by Eq. (14-9) should be increased by about 15 per cent.

The value of L_0 given by Eq. (14-11) was termed by Dellenbaugh and Quinby the *critical* value of first choke inductance. Their investigation showed that as L_1 is increased beyond L_0, the ratio of the alternating to the direct current continues to decrease fairly rapidly until $L_1 = 2L_0$. To ensure that the peak rectifier current is not excessive at full load,

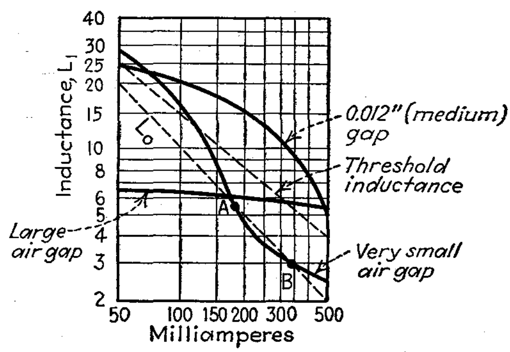


FIG. 14-18.—Variation of choke inductance with direct current.

therefore, the inductance of the first choke at full load should not be less than $2L_0$.

Because of the saturating effect of the direct current, the inductance of the filter chokes varies with load.[1] It is important that they should be designed so that the first inductance is sufficiently large at all loads. At minimum load the rectifier peak current is relatively small, and so it is satisfactory to make L_1 somewhat larger than $L_0 = R_{max}/1000$, where R_{max} is the total load resistance at minimum load, including the resistance of the chokes. At full load, L_1 should exceed $2L_0 \equiv R_{min}/500$, where R_{min} is the total resistance at full load. It may be assumed that between minimum and full load the inductance should vary linearly with load resistance. The inductance of the choke can be made to vary in the desired manner by proper adjustment of the air gap.[2] The required

[1] For a discussion of variation of choke inductance and the method of measuring the inductance, see E. H. MEIER and D. L. WAIDELICH, *Communications*, November, 1941, p. 5.

[2] DELLENBAUGH and QUINBY, *loc. cit.* See also E. GLAZER, *QST*, **15**, 13 (October, 1931).

value of inductance at any load may be termed the *threshold inductance.* Figure 14-18, taken from Dellenbaugh and Quinby's paper, shows how the inductance may be made to exceed the required threshold inductance by gap adjustment.[1] When the inductance falls below the critical value over a certain range of load, as from A to B in Fig. 14-18, violent oscillation takes place.[2] This can be prevented by increasing the air gap.

Other factors also affect the choice of L and C. It is important to choose L and C so that no section of the filter will resonate at the fundamental ripple frequency or within the frequency range covered by an amplifier with which it is used. Although the filter is a complicated circuit which may resonate at a number of frequencies, usually it is sufficient to make sure only that the highest resonant frequency of the individual sections is less than the fundamental ripple frequency and less than the lowest frequency limit for which the amplifier is designed.

To reduce the likelihood of amplifier regeneration caused by impedance of the power supply (see Sec. 6-38), it is important that the impedance presented to the amplifier by the supply should be small at the low-frequency limit of amplification. Since the reactance of the final condenser is small in comparison with the impedance in parallel with it, the internal power-supply impedance may be assumed to be the condenser reactance. The specification of a maximum allowable internal impedance at a particular frequency therefore determines a minimum value of last condenser capacitance.

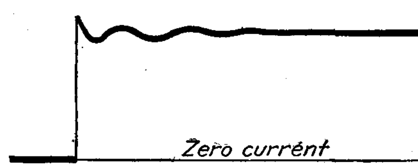


FIG. 14-19.—Oscillogram of transient oscillation of load current caused by filter inductance and capacitance.

Another factor that must occasionally be taken into consideration is transient oscillation of the filter when the load current is suddenly changed. Figure 14-19 shows a transient oscillation of load current resulting from the sudden decrease of load resistance across a power supply.[3] It is evident that such an oscillation may cause serious difficulties when the power supply is used with an amplifier, particularly in the amplification of abrupt changes of current or voltage.

The best values of L and C also depend upon space requirements and upon the relative cost of chokes and condensers. The physical size of chokes and condensers varies with their ability to store energy. For this reason it is sometimes desirable to make the total energy storage a

[1] Under the assumed linear variation of threshold inductance with load resistance, the threshold inductance line in Fig. 14-18 should be slightly curved. The curvature is so small, however, that the line is assumed to be straight.

[2] DELLENBAUGH and QUINBY, *loc. cit.*

[3] REICH, H. J., and MARVIN, G. S., *Rev. Sci. Instruments*, **2**, 814 (1931).

minimum. The total energy stored in one filter section is

$$\text{Stored energy} = \tfrac{1}{2}LI^2 + \tfrac{1}{2}CE^2 \tag{14-12}$$

By substituting in Eq. (14-12) the values of I_{dc} and L given by Eqs. (14-7) and (14-5) and differentiating with regard to C, the following expressions may be obtained for the optimum size of condenser and choke for a filter of n similar stages.

$$C = \frac{\sqrt{1 + \sqrt[n]{\alpha_1}}}{2\pi f_r R_t} \tag{14-13}$$

$$L = \frac{R_t \sqrt{1 + \sqrt[n]{\alpha_1}}}{2\pi f_r} \tag{14-14}$$

It is better, in general, to use a condenser that is larger than the optimum size rather than one that is smaller. In practice it is likely that the balance between choke and condenser size will be determined to a great extent by the size and cost of available standard chokes and condensers. At moderate voltages the use of electrolytic condensers of

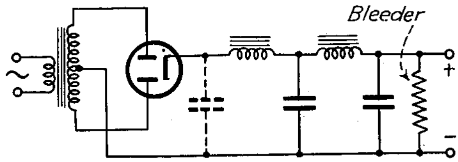


FIG. 14-20.—Circuit diagram of a full-wave single-phase power supply.

high capacitance makes possible the design of a filter of small size and low cost. It is important to note that account must always be taken of the difference between nominal and actual choke inductance.

Figure 14-20 shows the complete circuit of a single-phase, full-wave power supply with a two-section choke-condenser filter. One function of the "bleeder" resistance is to prevent excessive transient voltages in the filter condensers when the power supply is turned on under light load. It may also serve as a voltage divider. The chokes should always be in the ungrounded side of the filter. If the chokes are in the grounded side of the filter and the transformer case is grounded, capacitance between the core and the secondary winding shunts the chokes and by-passes some of the ripple around the chokes, thus reducing the filtering.[1] Since the negative terminal is usually grounded, the chokes are ordinarily in the positive side of the filter.

It is good practice to place and orient the chokes and power transformer in such a manner as to minimize coupling between them resulting

[1] TERMAN, F. E., and PICKELS, S. B., *Proc. I.R.E.*, **22**, 1040 (1934).

from leakage flux. The arrangement that gives minimum ripple can often be determined by trial.

14-11. Illustrative Problem.—The method of designing a rectifier with a choke-input filter can be best explained by means of an illustrative example. Suppose that the required power supply is to be single-phase and full-wave and must be able to furnish a d-c output of 200 ma at 350 volts with a ripple factor that does not exceed 0.07 per cent.

Examination of the operating characteristics of the 5U4-G high-vacuum rectifier (see page 683) shows that this tube can carry the required current and voltage. Assume that the output will be shunted by a 20,000-ohm voltage divider (bleeder), which will draw $17\frac{1}{2}$ ma at 350 volts. Table 14-I gives 120 cps and 0.472 for the values of f_r and ρ_1. Since the allowable value of ρ_1' is 0.0007, $\alpha_1 = 0.472/0.0007 = 675$. Substitution of these values of f_r and α_1 in Eq. (14-5) gives the following values of LC:

Single-section filter	1190×10^{-6} henry $\times$ farad
Two-section filter	47.5×10^{-6} henry $\times$ farad
Three-section filter	17.2×10^{-6} henry $\times$ farad

At full load L_1 should not be smaller than $R_{\text{min}}/500$. Choke resistance neglected, R_{min} is equal to the full-load voltage divided by the total full-load current, or $350/(0.200 + 0.0175) = 1610$ ohms. Therefore L_1 should not be less than 3.22 henrys. At minimum load L_1 should exceed $R_{\text{max}}/1000 = 20{,}000/1000 = 20$ henrys. A first choke having an inductance of 25 henrys at a current of $17\frac{1}{2}$ ma should satisfy this requirement. If the inductance falls to half this value at full load, the required capacitance for each filter section is

Single-section filter	95.2 μf
Two-section filter	3.8 μf
Three-section filter	1.37 μf

A two-section filter using two 4-μf condensers, or a three-section filter using three 1.5- or 2-μf condensers, will give the required filtering. For the two-section filter the resonant frequency of one section is $1/2\pi\sqrt{12.5 \times 4 \times 10^{-6}} = 22.5$ cps at full load. The resonant frequency of one section of the three-section filter is 31.8 cps. Both of these values are far below the fundamental ripple frequency. The optimum full-load inductance indicated by Eq. (14-14) is 11 henrys for the two-section filter, which is not far from the value chosen. Because of its lower resistance, the two-section filter gives better voltage regulation than the three-section filter.

Although an approximate value of the required transformer voltage, under the assumption that the tube drop is negligible, could be found from Table 14-I, the required voltage can be found more accurately from the operating characteristics of the 5U4-G rectifier. If each choke has 75 ohms resistance, the transformer secondary voltage should be 515 volts each side of center to give the 350-volt output at a total current of 217.5 ma.

The direct voltage will rise to approximately 450 volts when the external load is removed. Much improved regulation could be obtained by the use of the type 83 mercury vapor tube. Although improved regulation would also be obtained with the type 83v high-vacuum rectifier, this tube cannot deliver the terminal voltage required in this problem, since the maximum rated r-m-s plate voltage is only 400 volts.

The values of L and C actually chosen in both the two- and three-section filter are enough larger than the values that will just give the maximum allowable ripple to make it unnecessary to recheck the computations, taking into account the resistance of the chokes.

It is interesting to compute the ripple factor that would be obtained if the above two-section filter were used with a three-phase rectifier (see Table 14-I). For a three-phase rectifier $f_r = 180$ cps. Substituting $f_r = 180$ cps, $C = 4 \times 10^{-6}$ μf and $L = 12.5$ henrys in Eq. (14-12) gives 3960 for α_1. Since $\rho_1 = 0.177$ for the three-phase rectifier, $\rho_1' = 0.0045$ per cent. This clearly shows the improvement in filtering resulting from an increase in the number of phases.

The student will find it instructive to compute the second- and third-harmonic ripple factors at the load by substituting in the following equations values of ρ_1 and ρ_2 listed in Table 14-I.

$$\alpha_k = [(k\omega_f)^2LC - 1]^n \tag{14-4b}$$

$$\rho_k' = \frac{\rho_k}{\alpha_k} \tag{14-6a}$$

14-12. Choke-condenser Filter with Condenser Input.—Power supplies with filters of the form of Fig. 14-7c can be readily designed by a combination of the methods discussed in Secs. 14-7 and 14-10. Since the output ripple factor can be readily reduced to the desired value by the L-section portion of the filter, the input condenser may be chosen on the basis of voltage regulation and allowable peak plate current. If possible, $\omega R_t C_1$ should be large enough to bring it within the range of the flat portion of the voltage-ratio curve of Fig. 14-9, 14-10, or 14-11. It should not be so large, however, that the peak plate current exceeds the allowable value.

14-13. Voltage Stabilizers.—In many applications of electron tubes, notably in oscillators and measuring instruments, it is necessary to provide direct voltage that remains essentially constant in spite of fluctuations of line voltage. Direct voltages of almost any desired degree of stability can be obtained by the use of electronic voltage stabilizers.[1] Since a stabilizer that reduces random changes of direct voltage will, in general, also reduce periodic changes of voltage, voltage stabilizers used in the output of a power supply produce the additional beneficial effect of reducing ripple voltage. Types of stabilizers that reduce the variation of voltage with load current also reduce the internal terminal impedance of the power supply.

The simplest type of voltage stabilizer is the glow-tube stabilizer discussed in Sec. 12-4. Glow-tube stabilizers are usually incorporated in stabilizer circuits using high-vacuum tubes, in order to maintain constant operating voltages. Three types of high-vacuum-tube stabilizers are shown in Fig. 14-21. In the degenerative circuit of Fig. 14-21*a*,

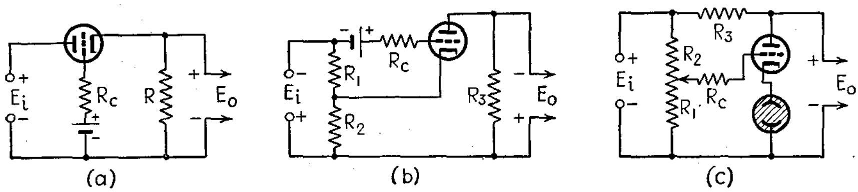


FIG. 14-21.—Basic voltage stabilizer circuits: (*a*) degenerative type; (*b*) amplification-factor bridge type; (*c*) transconductance-bridge type using glow-tube bias.

increase of output voltage E_o as the result of either increase of load resistance or increase of input voltage E_i increases the negative voltage of the grid and thus tends to reduce the plate current. The reduction of plate current in turn partly eliminates the change in voltage across R, which may be the load. The function of R_c is to prevent the flow of high grid current in the event of removal of E_i. Analysis of the equivalent plate circuit shows that best stabilization is obtained with tubes having high amplification factor. Loss of output voltage as the result of tube drop is minimized by the use of a high-current power tube. Since tubes having high amplification factor in general have low plate current, this circuit is useful only in applications where the load current is small.

In circuit *b*, the ratio of R_2 to R_1 is made equal to the amplification factor. When the input voltage changes, the change of plate voltage is then μ times the change of grid voltage, and so the plate current and the voltage across R_3, which may be the load, remain unchanged. Good stabilization is obtained over a range of input voltage throughout which the amplification factor remains approximately constant. The output

[1] GRAMMER, G., *QST*, August, 1937, p. 14; BOUSQUET, A. G., *Electronics*, July, 1938, p. 26; WALTZ, W. W., *Electronics*, December, 1938, p. 34; RCA Application Note 96; HUNT, F. V., and HICKMAN, R. W., *Rev. Sci. Instruments*, **10**, 6 (1939) (with 13 references); NEHER, H. V., and PICKERING, W. H., *Rev. Sci. Instruments*, **10**, 53 (1939) (with bibliography); COOMBS, J. N., and NIMS, P. T., *Electronics*, January, 1940, p. 40.

voltage may be made to increase with decrease of input voltage by making R_2/R_1 less than μ. R_c serves the same function as in circuit a. The output voltage in this circuit is less than the input voltage by the sum of the voltage across R_1 and the drop through the tube. Since $R_1 = R_2/\mu$ for balance, the voltage loss in R_1 may be minimized by the use of a high-mu tube. Because the plate current of high-mu tubes is small, however, this circuit, like that of Fig. 14-21a, is best suited to applications in which the load current is small.

In circuit c the glow tube, which may be replaced by a battery, maintains the cathode of the triode at a constant positive voltage relative to the negative side of the input. The negative bias of the triode therefore decreases with rise of input voltage and thus increases the plate current. By proper choice of circuit constants, the resulting increase of voltage drop across R_3 can be made equal to the increase of input voltage, so that the output voltage remains constant. Analysis of the equivalent plate circuit shows that when $R_3 = (R_1 + R_2)R_1g_m$ the output voltage is independent of input voltage throughout a range of input voltage and plate current in which g_m is essentially constant. Since R_3 should be small in order to prevent unnecessary loss of voltage, a tube used in this circuit should have high transconductance.

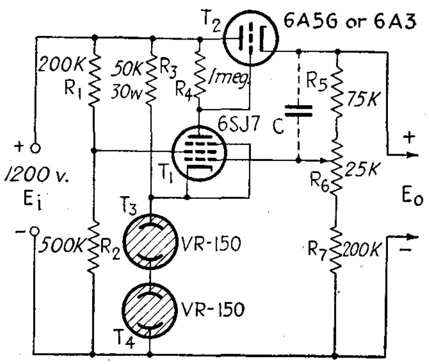


FIG. 14-22.—Two-stage degenerative voltage stabilizer.

In some applications of voltage compensators, internal resistance (change in output voltage with change in load current) is objectionable. It is reduced in the circuits of Fig. 14-21 by reduction of plate resistance but is relatively high even when power triodes are used. Improved circuits, incorporating amplifiers, have been developed that have low internal resistance and give high stabilization over wide ranges of input voltage and load current. The circuit most commonly used, shown in Fig. 14-22, is essentially a two-stage inverse-feedback circuit. Increase of E_o decreases the magnitude of the negative grid voltage of T_1 and thus increases the plate current of T_1. The resulting increase of voltage across R_4 makes the grid of T_2 more negative, thus reducing the plate current of T_2 and hence tending to reduce the output voltage. The output voltage may be varied by means of R_6.

Because of the voltage-dividing action of the resistance combination R_5-R_6-R_7, only a portion of the change of output voltage is applied to the grid of T_1. Addition of the condenser C increases the fraction of the voltage change impressed upon the grid of T_1 when the change occurs rapidly. If the reactance of this condenser at ripple frequency is small in comparison with the resistance which it shunts, the condenser appreciably increases the ability of the circuit to reduce ripple. If the condenser is too large, however, it may lead to undesirable transient oscillations of output voltage because the initial compensation as the result of sudden change of output voltage exceeds the final value obtained after the condenser voltage reaches equilibrium. Although the screen voltage for T_1 may be taken from the output side of the circuit, some improvement of stabilization against change of input voltage is obtained when the screen voltage is taken from the input side, as in Fig. 14-22.

The stabilization against changes of input voltage can be made complete by taking the excitation voltage for the first grid of T_1 partly from the input, as in the circuit of Fig. 14-23.[1] The slider of R_1 is set so that it is at approximately the same direct voltage as that of R_6. R_8 is adjusted so that there is no change in E_o when E_i is varied. The improvement in stabilization against changes of input voltage is accompanied by reduction of stabilization against changes of load.

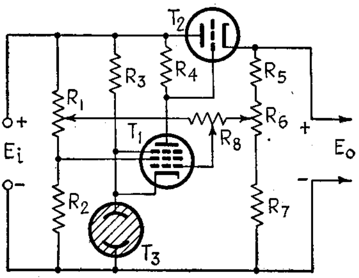


FIG. 14-23.—Variant of the circuit of Fig. 14-22 which stabilizes completely against changes of input voltage.

T_1 should be a high-mu tube and T_2 a high-transconductance tube. The coupling resistance R_4 should be large in comparison with the plate resistance of T_1. R_3 should be chosen so as to limit the current through the glow tubes to the range for which they are designed. Because of the large difference of potential between the cathodes of T_1 and T_2, it is necessary to use separate transformer windings for the two heaters, insulated to withstand the high voltage. A detailed discussion of the design of this type of voltage stabilizer has been given by Bereskin.[2]

When a stabilized voltage supply is used to supply the grid bias for a power amplifier in which the grids are driven positive during a portion of the cycle of excitation voltage, the grid current flows in opposition to the supply voltage and thus reduces the supply current. The grid current cannot increase beyond the value at which the supply current is reduced to zero. In order to use the circuit of Fig. 14-22 for this purpose, therefore, it is necessary to shunt the output with a resistance which is low enough to ensure that current through the resistance exceeds the highest grid current. Since the peak grid current may be high, the

[1] GOODWIN, C. W., *Yale J. Biol. and Med.*, **14**, 101 (1941).

[2] BERESKIN, A. B., *Proc. I.R.E.*, **31**, 47 (1943).

high current that must flow through the resistor results in considerable loss of power. It is also difficult to design the circuit of Fig. 14-22 to give good stabilization at output voltages below 100 volts. These difficulties are avoided in the circuit of Fig. 14-24, which is derived from the basic circuit of Fig. 14-21*c* by the addition of a stage of voltage amplification.[1] Since the voltages across T_1 and T_3 are essentially constant, reduction of output voltage E_o results in an equal reduction of grid bias of T_2. This increases the plate current of T_2 and hence the drop in voltage across R_2, increasing the grid bias of T_4. The resulting reduction of plate current of T_4 reduces the voltage drop in R_3 and thus tends to counteract the change in output voltage. Detailed design procedure for this circuit has been given by Pihl.[1]

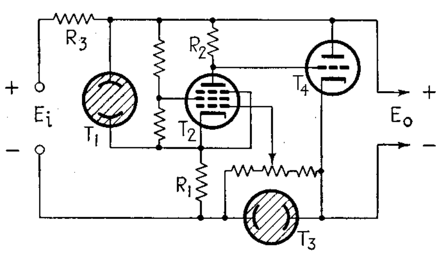


FIG. 14-24.—Voltage stabilizer for use in a *C* supply for power amplifiers in which grid current flows.

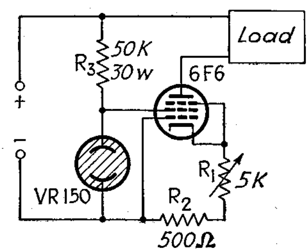


FIG. 14-25.—Current stabilizer.

Stabilizers of the type of Figs. 14-22 and 14-23 cannot be readily designed to deliver voltages lower than about 100 volts. Low stabilized voltages can, however, be obtained by the use of two stabilized voltage supplies connected in opposition.

14-14. Current Stabilizer.—Figure 14-25 shows a circuit by means of which current may be maintained essentially constant in spite of variations of supply voltage or load resistance. This circuit makes use of the fact that, within the proper ranges of voltages and plate current, the plate current of a pentode varies comparatively little with plate voltage. Additional stabilization is attained by the use of a cathode biasing resistance, which causes the grid bias to increase with increase of plate current. The load current may be varied by means of R_1. R_3 must be large enough to prevent the flow of excessive screen current when the plate current is reduced to a low value, but small enough to ensure that the glow tube conducts. The circuit values of Fig. 14-25 are suitable for an input voltage ranging from 200 volts to 1200 volts. The cathode resistor varies the current from approximately 5 to 30 ma. In the vicinity of 20 ma the plate current varies by about $1\frac{1}{2}$ ma when the input

[1] PIHL, G. E., *Bull.* 10, College of Engineering, Department of Electrical Eng., Northeastern Univ. (1943).

voltage is varied from 300 to 1000 volts or when the load is varied from zero to 50,000 ohms.

Problems

14-1. Design a B supply with condenser filter to meet the following specifications:

Output voltage	Not less than 280 volts
Maximum load current	20 ma
Ripple factor	Not to exceed 0.01
Supply	110-volt, 60-cycle, single-phase

The rectifier should operate directly from the line (no input transformer).

14-2. Design a B supply with choke-input filter to meet the following specifications:

Output voltage	300 volts
Maximum load current	100 ma
Ripple factor	Not to exceed 0.0005
Supply	120-volt, 60-cycle, single-phase

Impedance presented to the load at 50 cps must not exceed 1000 ohms. Voltage regulation should be kept as small as possible.

14-3. Design a B supply with a condenser-input choke-condenser filter to meet the following specifications:

Output voltage	400 volts
Maximum load current	100 ma
Ripple factor	Not to exceed 0.0001
Supply	110-volt, 60-cycle, single-phase

14-4. Design a single-phase full-wave rectifier that will furnish a direct voltage of 1500 volts at a load current of 250 ma and a ripple factor not to exceed 0.1 per cent. Minimum load current is 50 ma.

14-5. By means of the equivalent plate circuit, show that the effectiveness of the circuit of Fig. 14-21*a* increases with μ.

14-6. By means of the equivalent plate circuit, show that the circuit of Fig. 14-21*b* should be adjusted so that $R_2/R_1 = \mu$.

14-7. By means of the equivalent plate circuit, show that the circuit of Fig. 14-21*c* should be adjusted so that $R_3 = (R_1 + R_2)/R_1 g_m$.

Supplementary Bibliography

Prince, D. C., and Vogdes, F. B.: "Principles of Mercury Arc Rectifiers and Their Circuits," McGraw-Hill Book Company, Inc., New York, 1927.

Jolley, L. B. W.: "Alternating-current Rectification," John Wiley & Sons, Inc., New York, 1931.

Marti, O. K., and Winograd, H.: "Mercury Arc Power Rectifiers," McGraw-Hill Book Company, Inc., New York, 1930.

Record, F., and Holmes, I.: *Wireless World*, **21,** 841 (1927).

Kuhlman, J., and Barton, J.: *J. Am. Inst. Elec. Eng.*, **47,** 17 (1928).

Vincer-Minter, N. P.: *Wireless World*, **22,** 184, 226, 323 (1928).

Crouse, G. B.: *Radio Eng.*, May, 1928, pp. 19, 42.

Warren, A. G.: *Wireless Eng.*, **5,** 601 (1928).

Shaw, A. E.: *Proc. I.R.E.*, **17,** 849 (1929).

Reudy, R.: *Can. J. Research*, **2,** 101 (1930).

STEINER, H. C., and MASER, H. T.: *Proc. I.R.E.*, **18**, 67 (1930).
MASER, H. T., and SAXTON, H. L.: *QST*, March, 1931, p. 21.
MCDONALD, D.: *Wireless Eng.*, **8**, 522 (1931).
CLARK, H. T., and KOHLHAGEN, W.: *Phys. Rev.*, **39**, 184 (1932).
STREET, J. C., and JOHNSON, T. J.: *J. Franklin Inst.*, **214**, 155 (1932).
CALLENDAR, M. V.: *Wireless Eng.*, **9**, 24 (1932).
SCROGGIE, M. G.: *Wireless World*, **30**, 558 (1932).
DAVIS, R. L., and TROUANT, V. E.: *Proc. I.R.E.*, **20**, 921 (1932); *QST*, June, 1933, p. 21.
KIMP, R.: *Electronics*, August, 1933, p. 219.
TANASESCU, T.: *Wireless Eng.*, **11**, 68 (1934).
DUNHAM, C. R.: *J. Inst. Elec. Eng.*, **75**, 278 (1934).
EVANS, R. D.: *Rev. Sci. Instruments*, **5**, 371 (1934).
FOOS, C. B.: *Elec. Eng.*, **53**, 568 (1934).
WALTZ, W. W.: *Electronics*, December, 1935, p. 30.
MASLIN, A. J.: *Electronics*, December, 1935, p. 29.
FOOS, C. B., and LATTEMANN, W.: *Proc. I.R.E.*, **24**, 977 (1936).
MASLIN, A. J.: *Electronics*, December, 1936, p. 28.
KILGORE, L. A., and COX, J. H.: *Trans. Am. Inst. Elec. Eng.*, **56**, 1134 (1937); **57**, 168 (1938).
MCDONALD, G. R.: *Trans. Am. Inst. Elec. Eng.*, **58**, 563 (1939).
LEE, REUBEN: *Electronics*, April, 1938, p. 39.
DEELEY, P. M.: *Communications*, April, 1938, p. 19.
SCOTT, H. J.: *Electronics*, June, 1938, p. 28.
TRUCKSESS, D. E.: *Bell Lab. Rec.*, **17**, 24 (1938).
WALLIS, C. M.: *Electronics*, October, 1938, p. 12.
HUNTOON, R. D.: *Rev. Sci. Instruments*, **10**, 176 (1939).
SCOTT, H. H.: *Electronics*, August, 1939, p. 42.
STOUT, M. B.: *Electronics*, September, 1939, p. 32.
MERRILL, F. D., JR.: *Electronics*, November, 1939, p. 16.
HONNELL, M. A.: *Communications*, January, 1940, p. 14.
WALLIS, C. M.: *Electronics*, March, 1940, p. 19.
WITTING, E. G.: *Rev. Sci. Instruments*, **11**, 182 (1940).
RICHTER, W.: *Electronics*, June, 1940, p. 20.
PANOFSKY, W. K. H., and ROBINSON, C. F.: *Electronics*, April, 1941, p. 42.
FROMMER, J. C.: *Proc. I.R.E.*, **29**, 481 (1941).
WARE, L. A.: *Proc. I.R.E.*, **30**, 500 (1942).
TILLOTSON, L. C., and WALLIS, C. M.: *Electronics*, April, 1941, p. 42.

CHAPTER 15

ELECTRON-TUBE INSTRUMENTS AND MEASUREMENTS

The development of measuring instruments using electron tubes has resulted in great improvements in the field of measurement. This final chapter will deal with the theory and characteristics of electronic measuring instruments and with the methods of making measurements on vacuum tubes and vacuum-tube circuits. Instruments incorporating electron tubes include voltmeters, ammeters, wattmeters, ohmmeters, time and speed meters, frequency meters, and oscillographs. Measurements on tubes and tube circuits include measurement of tube factors, negative resistance, harmonic content, voltage amplification, and power output.

15-1. Vacuum-tube Voltmeters.—One of the most useful of electronic measuring instruments is the vacuum-tube voltmeter.[1] The advantages of vacuum-tube voltmeters over electromagnetic meters are their high input impedance, wide frequency range, and high sensitivity. Although vacuum-tube measuring instruments usually embody electromagnetic meters, the required sensitivity of the latter is generally small enough so that the cost of a complete electronic instrument is considerably less than that of an electromagnetic or electrostatic meter of comparable sensitivity. The vacuum-tube voltmeter is of particular value in the measurement of voltage in circuits in which the flow of meter current would appreciably alter the voltage, as in radio-frequency circuit measurements, the measurement of amplifier output voltage, or the measurement of the voltage of a charged condenser.

Vacuum-tube voltmeters are of many types. According to principle of operation they may be classified as follows:

1. Transfer characteristic.
2. Plate-detection (transrectification):
   a. Low-bias.
   b. High-bias.
3. Diode.
4. Grid-detection.
5. Exponential.

[1] For a general discussion of vacuum-tube voltmeters covering developments through 1932, see T. P. Hoar, *Wireless Eng.*, **10**, 19 (1933). See also AUGUST HUND, "High-frequency Measurements," pp. 136–161, McGraw-Hill Book Company, Inc., New York, 1933.

6. Amplifier.
7. Slide-back.
8. Inverted.

Classification can also be made according to whether the instruments measure direct, r-m-s, crest, or average voltage. This will be discussed in connection with each of the types listed above.

The design of vacuum-tube voltmeters involves a number of factors, which include the following:

1. Input impedance.
2. Sensitivity.
3. Type of scale (direct, r-m-s, crest, or average voltage).
4. Stability and permanence of calibration.
5. Voltage range.
6. Effect of harmonics upon readings.
7. Portability and cost.

15-2. Use of Transfer Characteristics in Voltage Measurement.—The most direct application of vacuum tubes to the measurement of voltage lies in the use of a transfer characteristic as a calibration curve. Unknown direct voltages less in magnitude than the cutoff grid voltage may be applied to the grid circuit and measured by means of a milliammeter in the plate circuit. If the voltage measured is applied with such polarity as to make the grid negative, negligible current is drawn by the grid.

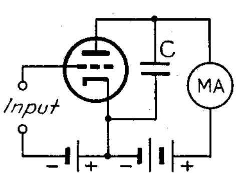


FIG. 15-1.—Basic circuit of plate-detection vacuum-tube voltmeter.

Large direct voltages applied to the plate may also be measured by relatively small voltages applied to the grid, the grid voltage being adjusted so as to reduce the plate current to zero.[1]

15-3. Plate-detection (Transrectification) Voltmeters.[2]—One of the most frequently used types of vacuum-tube voltmeters makes use of plate detection. The basic circuit of this type of meter is shown in Fig. 15-1. Alternating voltage applied to the grid cicruit is measured by means of the change in reading of the d-c milliammeter in the plate circuit. Because of curvature of the transfer characteristic, the amplitude of the positive half cycle of plate current exceeds the amplitude of the negative half cycle. The average plate current therefore exceeds the quiescent plate current, as shown in Fig. 3-19 (see Secs. 3-15, 3-19, and 9-11), the change in current being a function of the impressed

[1] VAN DER BIJL, H. J., "The Thermionic Vacuum, Tube and Its Applications," p. 369, McGraw-Hill Book Company, Inc., New York, 1920.

[2] MOULLIN, E. B., and TURNER, L. B., *J. Inst. Elec. Eng.* (*London*), **60,** 706 (1922); MOULLIN, E. B., *J. Inst. Elec. Eng.* (*London*), **61,** 295 (1923), *Wireless World*, **10,** 1 (1922).

alternating voltage. The condenser, whose reactance is small at the lowest input frequency, prevents variation of load impedance with frequency and thus eliminates a possible cause of dependence of calibration upon frequency. The instrument may be calibrated by the application of known sinusoidal voltages. This type of vacuum-tube voltmeter may be further classified according to the magnitude of the grid bias and the form of the transfer characteristic.

1. *Low-bias Operation.*—If the shape of the transfer characteristic is assumed to be approximately parabolic, then for Class A operation the alternating plate current may be adequately expressed by the first two terms of the series expansion

$$i_p = a_1 e + a_2 e^2 \tag{15-1}$$

Since a periodic wave of any shape may be broken up into a series of harmonically related frequency components, the application of a periodic voltage wave of any form is equivalent to the simultaneous application of a number of periodic voltages of proper frequency and amplitude. In general the applied voltage may then be expressed as

$$e = e_g = E_1 \cos \omega t + E_2 \cos 2\omega t + E_3 \cos 3\omega t + \cdots \tag{15-2}$$

Substitution of Eq. (15-2) in Eq. (15-1) shows that the change in average plate current is

$$H_0 = \tfrac{1}{2} a_2 (E_1^2 + E_2^2 + E_3^2 + \cdots) \tag{15-3}$$

But the mean-square signal voltage is

$$\frac{\omega}{2\pi} \int_0^{2\pi/\omega} (E_1 \cos \omega t + E_2 \cos 2\omega t + E_3 \cos 3\omega t + \cdots)^2 dt = \tfrac{1}{2}(E_1^2 + E_2^2 + E_3^2 + \cdots) \tag{15-4}$$

The change of average plate current of a tube that has a parabolic transfer characteristic and is used in Class A operation is therefore proportional to the mean-square signal voltage and the meter may be calibrated to read r-m-s signal voltage.[1] In practice it is difficult to obtain with a single tube a transfer characteristic that approaches closely to parabolic form. The change in average plate current is therefore not strictly proportional to mean-square signal voltage. The change in current will depend to some extent upon the wave form of the signal voltage.

A close approach to a parabolic characteristic can be attained by the use of the balanced circuit of Fig. 15-2.[2] The plate current of T_1 is

$$i_p = \Sigma a_n (e_g)^n \tag{15-5}$$

[1] JANSKY, C. M., and FELDMAN, C. B., *J. Am. Inst. Elec. Eng.*, **47**, 126 (1928).
[2] See, for instance, W. C. MICHELS, *Rev. Sci. Instruments*, **9**, 10 (1938).

That of T_2 is

$$i_p' = \Sigma a_n(e_g')^n = \Sigma a_n(-e_g)^n \tag{15-6}$$

Since all the odd-order terms of i_p' are negative, whereas all terms of i_p are positive, the sum of i_p and i_p' contains only even-order terms. With triodes the value of a_4 and higher-order coefficients is small in comparison with a_2, and so the total plate current is

$$i = i_p + i_p' \cong 2a_2(e_g)^2 = \tfrac{1}{2}a_2e^2 \tag{15-7}$$

The objection to this circuit is that it draws a small amount of current from the circuit whose voltage is being measured.

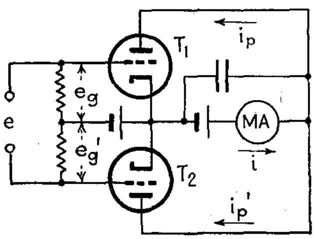


FIG. 15-2.—Basic circuit of balanced plate-detection vacuum-tube voltmeter.

2. *High-bias Operation.*—If the tube of Fig. 15-1 is biased to cutoff, plate current will flow only during the positive half of the signal-voltage cycle. The average plate current will depend upon the shape of the dynamic transfer characteristic which may be made to approach linearity by the use of an unby-passed high resistance in series with the plate. Under this condition the instantaneous plate current is nearly proportional to the instantaneous signal voltage during the positive half cycle and zero during the negative half cycle. The average plate current, under the assumption that the characteristic is linear, is

$$\begin{aligned} I_{ba} &= \frac{\omega}{2\pi}\int_0^{\pi/\omega} i_b\,dt = \frac{\omega}{2\pi}\int_0^{\pi/\omega} a_1e_g\,dt = \\ &\frac{a_1\omega}{2\pi}\int_0^{\pi/\omega}(E_1\sin\omega t + E_2\sin 2\omega t + E_3\sin 3\omega t + \cdots)\,dt \\ &= \tfrac{1}{2}a_1 \times (\text{average signal voltage during positive half cycle}) \end{aligned} \tag{15-8}$$

Thus the average plate current of a tube that has a linear characteristic and is biased to cutoff is proportional to the average signal voltage during the positive half cycle. Since in practice it is not possible to make the characteristic linear, this relation is only approximately true. The curvature is greatest in the vicinity of cutoff, and so the departure from the theoretical linear reading increases as the signal amplitude goes down. If the load resistance is zero or small, the characteristic in the vicinity of cutoff is more nearly parabolic than linear. Then the average plate current is

$$I_{ba} = \frac{\omega}{2\pi}\int_0^{\pi/\omega} i_b\,dt \cong \frac{\omega}{2\pi}\int_0^{\pi/\omega} a_2 e_g^2\,dt$$
$$= \frac{a_2\omega}{2\pi}\int_0^{\pi/\omega} (E_1 \sin \omega t + E_2 \sin 2\omega t + E_3 \sin 3\omega t + \cdot\cdot\cdot)^2\,dt \quad (15\text{-}9)$$
$$= \frac{a_2}{4}(E_1^2 + E_2^2 + E_3^2 + \cdot\cdot\cdot) \quad (15\text{-}10)$$

The average plate current for a tube with a parabolic characteristic, biased to cutoff, is, therefore, approximately proportional to mean-square positive signal voltage, and the meter may be calibrated to read the r-m-s value of the voltage during the positive half cycle.

Since high input impedance can be obtained only if the grid is negative throughout the cycle of signal voltage, voltages of amplitude greater than the cutoff grid voltage can be read only if the tube is biased beyond cutoff. If Class C bias is used, current flows during only a portion of the positive half cycle. The average plate current cannot be affected by the form of the wave during the portion of the cycle in which current does not flow, and so the average plate current is not necessarily proportional to either the average voltage or the mean-square voltage, and large errors in reading are likely to result if the voltage departs materially from sinusoidal form. Class C bias is sometimes used in multirange meters.

At high radio frequencies the reading of plate-detection voltmeters may be affected by impedance drop caused by inductance of the grid input leads and by resonance of this inductance with the capacitance of the grid circuit. Input capacitance also causes the input admittance to increase with frequency. For these reasons the instrument should be designed so as to minimize input circuit capacitance and inductance, and at high radio frequencies it is advisable to use specially designed tubes having low input capacitance and inductance, such as acorn tubes (Fig. 2-16). It is also good practice to mount the tube at the end of a shielded flexible cable, so that the input voltage may be applied directly to grid terminal of the tube.[1]

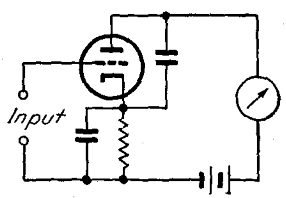


FIG. 15-3.—Self-biased plate-detection vacuum-tube voltmeter.

The calibration curve may be straightened by the use of self-bias, as shown in Fig. 15-3.[2] In addition to reducing curvature of the calibration curve, self-bias increases the voltage range of the instrument. The bias increases with input amplitude, and so the amplitude at which grid current starts to flow greatly exceeds the zero-signal bias.

[1] BOYLE, H. G., *Electronics*, August, 1936, p. 32. RCA Application Note 47, 1935.

[2] MEDLAM, W. B., and OSCHWALD, U. A., *Wireless Eng.*, **3**, 589, 664 (1926), **5**, 56 (1928).

15-4. Balanced Plate Circuits.—In any circuit in which change of plate current of a vacuum tube is used as a measure of input to the circuit, as in the measurement of direct voltages by use of transfer characteristics or of alternating voltages by means of detection, it is desirable to balance out the static operating plate current in order that the whole scale of the milliammeter or galvanometer may be used. One method of accomplish-

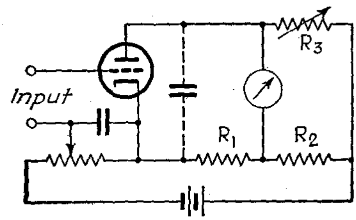


FIG. 15-4.—Vacuum-tube voltmeter with zero balance.

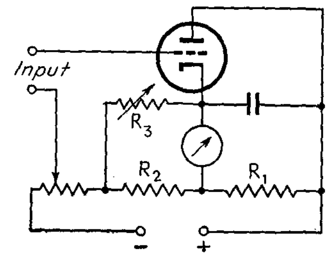


FIG. 15-5.—Vacuum-tube voltmeter with zero balance.

ing this is shown by the circuits of Figs. 15-4 and 15-5, in which the tube forms one arm of a Wheatstone bridge.[1] The circuit may be balanced by varying any of the resistances. To avoid loss in sensitivity, R_3 must be large in comparison with the meter resistance. The required B-supply

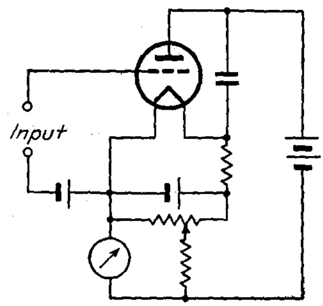


FIG. 15-6.—Circuit in which the filament supply voltage is used to balance out the zero-signal plate current.

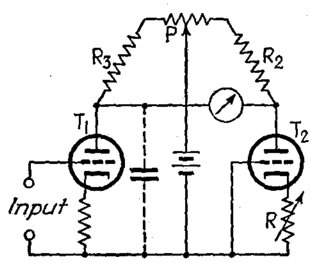


FIG. 15-7.—Circuit in which the zero-signal plate current is balanced out by the use of an auxiliary tube.

voltage increases, however, with R_3. When the cathode is heated by means of a battery, the circuit of Fig. 15-6 may be used. A modified form of the bridge circuit of Fig. 15-4, shown in Fig. 15-7, uses a second tube in place of the resistance R_1.[2] The advantage of the use of two

[1] HOARE, S. C., *Trans. Am. Inst. Elec. Eng.*, **46**, 541 (1927); HAYMAN, W. G., *Wireless Eng.*, **7**, 556 (1930). See also footnote 1, p. 613, and footnote 2, p. 614.

[2] WOLD, P. I., U. S. Patent 1232879 (1916–1917); WYNN-WILLIAMS, C. E., *Proc. Cambridge Phil. Soc.*, **23**, 810 (1927), *Phil. Mag.*, **6**, 324 (1928); EGLIN, J. M., *J. Opt. Soc. Am. and Rev. Sci. Instruments*, **18**, 393 (1929); WINCH, G. T., *J. Sci. Instruments*, **6**, 376 (1929); NOTTINGHAM, W. B., *J. Franklin Inst.*, **209**, 287 (1930); DUBRIDGE, L. A., *Phys. Rev.*, **37**, 392 (1931).

tubes (or a double triode) is that the tubes respond similarly to changes of plate supply voltage and cathode temperature, so that the bridge does not so readily become unbalanced. It is possible to find settings of P and R such that small variations of plate supply voltage have little or no effect upon the zero balance. The circuit of Fig. 15-7 may be used with alternating plate supply voltage.[1]

Changes of calibration with operating voltages can be reduced by using a common source to supply cathode power and grid and plate voltages. A typical circuit is shown in Fig. 15-8. Adjustment of the heater or filament current to a definite value also brings the grid and plate voltages to the proper values. If a heater type of tube is used, constant operating conditions are approximated by obtaining the plate and grid

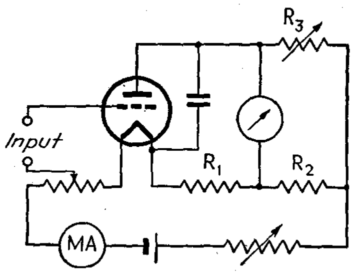


FIG. 15-8.—Plate-detection vacuum-tube voltmeter operated from a single source of voltage.

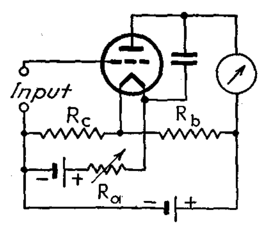


FIG. 15-9.—Circuit that is self-compensating against variations of supply voltage.

voltages from a common supply, the voltage of which is adjusted to a definite value by the aid of a voltmeter or is kept constant by means of a glow tube or other type of voltage stabilizer (see Secs. 12-4 and 14-13). Small changes of heater voltage in the neighborhood of normal value have comparatively small effect upon calibration. The problem of maintaining a fixed operating point may also be simplified by the use of a cathode resistor to provide self-bias, as in Fig. 15-3. When self-bias is used, the static operating point is fixed by the plate supply voltage, which may be kept constant with the aid of a voltmeter or a voltage stabilizer. A number of self-compensating single-tube circuits have also been developed.[2] One such circuit is shown in Fig. 15-9. The meter reading is independent of filament supply voltage when $R_c = \frac{1}{g_m} \cdot \frac{\partial i_b}{\partial i_f}$, and independent of plate supply voltage when $R_b = \mu R_c R_a'/R_c + R_a'$, where R_a' is the sum of R_a and the filament resistance.[3]

[1] THURSTON, J. N., *Electronics*, October, 1943, p. 102.

[2] RAZEK, J., and MULDER, P. J., *J. Opt. Soc. Am. and Rev. Sci. Instruments*, **18**, 460 (1929); DEARLE, R. C., and MATHESON, L. A., *Rev. Sci. Instruments*, **1**, 215 (1930); TURNER, L. A., and SIEGELIN, C. O., *Rev. Sci. Instruments*, **4**, 429 (1933); TURNER, L. A., *Rev. Sci. Instruments*, **4**, 665 (1933). See also p. 615.

[3] TURNER, *ibid.*

Changes of calibration with tube age are prevented to a certain extent by keeping the static operating current constant. It is necessary, however, to check the meter calibration from time to time regardless of the circuit used. The life of tubes of standard makes is sufficiently long so that tube deterioration is ordinarily an unimportant factor.

The condensers shown in Figs. 15-4 to 15-8 need be used only if the circuits are used in the measurement of alternating voltages by plate detection.

15-5. Diode Voltmeters.—Diode rectification can be used as the basis of vacuum-tube voltmeters that read average voltage.[1] The circuit, which is shown in Fig. 15-10, is one of the simplest of all vacuum-tube voltmeter circuits. If it is assumed that the rectifier resistance is zero and that no current flows when the applied voltage is zero, the instantaneous current is equal to the instantaneous voltage divided by the load resistance during the positive half cycle. The average current is the average voltage of the positive half cycle, averaged over the entire cycle, divided by the resistance. For a sinusoidal input voltage $E_{max} \sin \omega t$, the average plate current is

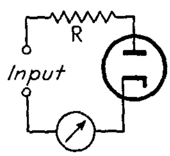


FIG. 15-10.—Simple diode voltmeter.

$$I_{ba} = \frac{E_{max}}{\pi R} = \frac{E_{rms}}{2.22R} \tag{15-11}$$

For a direct input voltage E, the plate current is

$$I_b = \frac{E}{R} \tag{15-12}$$

Thus, if the diode were a perfect rectifier having zero resistance, the circuit could be calibrated on direct voltage and used to measure either direct or sinusoidal alternating voltage. Actually, curvature of the diode static characteristic and flow of plate current when the plate voltage is zero, resulting from initial velocity of electrons, cause the ratio of the alternating to the direct voltage reading to depart somewhat from the theoretical value 2.22. The error, which is most pronounced at low voltages, can be made negligible by the use of load resistances exceeding 100,000 ohms. High load resistance is also desirable in order to keep the current small. Since the average input voltage depends upon wave form, the reading of this type of meter changes with harmonic content of the impressed voltage. Although it draws a small current, this

[1] SHARP, C. H., and DOYLE, E. D., *Trans. Am. Inst. Elec. Eng.*, **35**, 99 (1916); CHUBB, L. W., *Trans. Am. Inst. Elec. Eng.*, **35**, 109 (1916); TAYLOR, JAMES, *J. Sci. Instruments*, **3**, 113 (1925); *Wireless Eng.*, **10**, 310 (1933). DAVIS, R., BOWDLER, G. W., and STANDRING, W. G., *J. Inst. Elec. Eng. (London)*, **68**, 1222 (1930).

circuit makes possible the use of a sensitive d-c meter in the measurement of alternating voltages.

Figure 15-11 shows the direct- and alternating-voltage calibration curves of the diode sections of a type 6B7 tube used with a 100,000-ohm

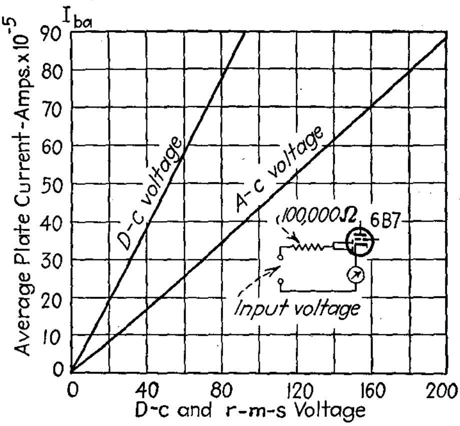


FIG. 15-11.—Direct- and alternating-voltage calibrations of the diode sections of a 6B7 with a 100,000-ohm load.

load. Similar curves for a 500,000-ohm load are practically straight lines through the origin, having slopes indicated by Eqs. (15-11) and (15-12).

Another type of diode voltmeter, the basic circuit of which is shown in Fig. 15-12, consists of two diodes, a condenser, and an electromagnetic d-c meter.[1] The condenser current flows through the meter during only half of the cycle, and so the meter current is unidirectional. With sinusoidal input, the condenser voltage varies between the values E_m and $-E_m$, where E_m is the crest input voltage. The amount of charge that flows into the condenser during half a cycle is $2CE_m$, and the average current through the meter is equal to the product of this charge by the frequency, or $2fCE_m$. Therefore

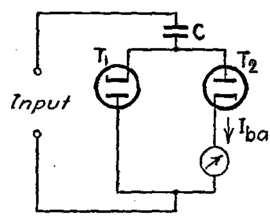


FIG. 15-12.—Diode voltmeter which measures the input voltage by means of the charging current of a condenser.

$$E_m = \frac{I_{ba}}{2fC} \tag{15-13}$$

To prevent circulating current through the meter as the result of emission velocity of electrons, a small counter-e.m.f. may be used in series with

[1] STARR, A. T., *Proc. Phys. Soc. London*, **46,** 45 (1934). See also L. W. CHUBB and C. L. FORTESCUE, *Trans. Am. Inst. Elec. Eng.*, **32,** 739 (1913).

the tube T_2, or a negatively biased triode may be substituted for the diode. Since the reading of the meter depends upon the algebraic difference between the positive and negative crest voltages, the reading will be affected by harmonics that change the crest voltages, unless one crest voltage is increased by the same amount that the other is decreased.

These simple diode voltmeters are of particular value in the measurement of very high voltages when some flow of current is not objectionable. With large values of R, the calibration curves of Fig. 15-11 are sufficiently straight, and the intercepts small enough, so that in the measurement of high voltage they may be considered to be straight lines passing through the origin and having slopes indicated by Eqs. (15-11) and (15-12). Experimental calibration is then unnecessary.

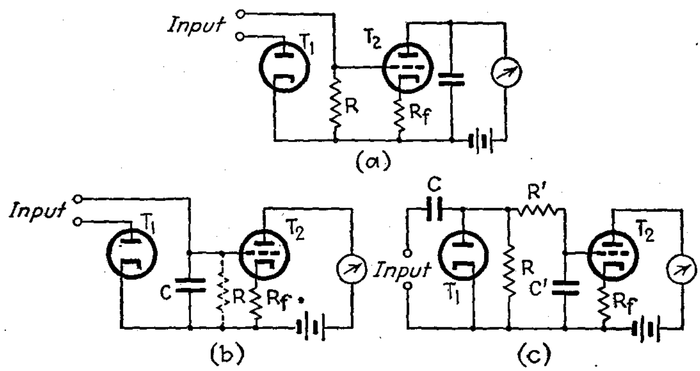


FIG. 15-13.—Diode voltmeters incorporating a stage of d-c amplification.

For the measurement of small voltages the diode circuit of Fig. 15-10 may be followed by a d-c amplifier, as shown by Fig. 15-13*a*. Since the average plate current of T_2 depends upon the average voltage across R, the reading of this meter, like that of Fig. 15-10, is dependent upon the average value of the positive half cycle of the applied voltage. If R is replaced by a condenser C, as in Fig. 15-13*b*, the condenser charges to crest positive signal voltage. The reading is therefore then a measure of the crest signal voltage. Because C cannot discharge through either tube, a high resistance R may be shunted across C in order to allow C to discharge when the input voltage is reduced or removed. To prevent the dependence of calibration upon frequency, the product fRC should be at least of the order of magnitude of 100 at the lowest frequency f of the input voltage.[1] A variant of the circuit of Fig. 15-13*b* is that of Fig. 15-13*c*, in which R' and C' act as a filter to keep the alternating voltage from being impressed upon the amplifier grid. The amplifier plate current in the circuits of Fig. 15-13 decreases with increase of

[1] MARIQUE, JEAN, *Wireless Eng.*, **12,** 17 (1935).

input voltage, and so the meter cannot be damaged by excessive input voltage. If the zero-signal current is balanced out, the meter current flows in the opposite direction and cannot exceed the zero-signal plate current. The circuits of Figs. 15-13*a* and *b* may also be used to measure direct voltage.

Inspection of the diagrams of Fig. 15-13 shows that the self-biasing resistor R_f results in degenerative feedback, the circuit of the amplifier stage being identical with the feedback circuit of Fig. 6-40*c*. Degeneration in the amplifier has four beneficial effects: (1) If R_f is large, the meter reading is practically proportional to the direct grid voltage, and hence to the positive crest voltage applied to the voltmeter input. (2) Because of the opposing voltage developed across the resistor, the grid circuit can

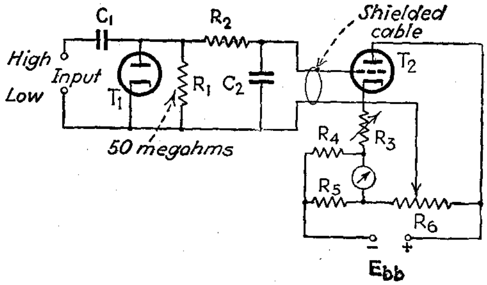


FIG. 15-14.—General Radio diode voltmeter.

handle much larger direct voltages than the static cutoff grid voltage. Thus, much larger input voltages may be read. (3) The voltage range can be conveniently varied by changing the resistance R_f and the static biasing voltage. (4) The meter reading is practically independent of tube constants.

These facts were employed by W. N. Tuttle in the design of a multi-range voltmeter.[1] The circuit of this meter, shown in Fig. 15-14, incorporates the zero-balance bridge of Fig. 15-5. The sensitivity is adjusted by means of R_3, the correct amplifier bias being maintained by simultaneous adjustment of R_6. A type 955 acorn tube is used as the diode rectifier and is mounted at the end of a shielded flexible cable for high-frequency measurements. A regulated B supply prevents change of calibration with alternating line voltage.

Diode voltmeters are more sensitive than plate-detection voltmeters but have the disadvantage of lower input impedance.

15-6. Grid-detection Voltmeter.—The functions of the two tubes of the circuits of Fig. 15-13 can also be performed by a single triode grid-leak detector, as shown in Fig. 15-15. The disadvantage of such a

[1] TUTTLE, W. N., *Gen. Radio Expt.*, **11**, 1 (May, 1937); THURSTON, J. N., *Electronics*, October, 1943, p. 102.

circuit over the two-tube circuits is that the desirable features resulting from the use of a cathode resistor cannot be attained, since the resulting bias would prevent the flow of grid current.

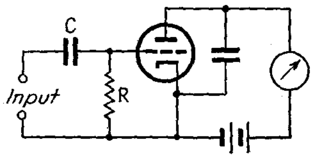


FIG. 15-15.—Grid-detection voltmeter.

15-7. Logarithmic Voltmeter.—Ballantine has shown that an amplifier with automatic gain control may be used as a voltmeter with a logarithmic scale.[1] A block diagram of the fundamental circuit is shown in Fig. 15-16. The amplifier uses variable-mu tubes, in which the amplification factor is an exponential function of the grid voltage. The output is maintained constant by rectifying and filtering the output voltage and using the resulting direct voltage to bias the amplifier tubes. If the

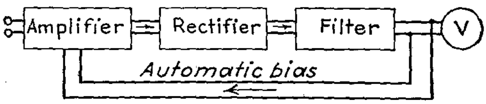


FIG. 15-16.—Schematic diagram of logarithmic voltmeter using *a-v-c* circuit.

amplification factor of the amplifier tube is an exponential function of the grid bias, the ratio of amplifier output voltage to input voltage is

$$\frac{e_2}{e_1} = A\epsilon^{naE_c} \tag{15-14}$$

where n is the number of identical stages and a and A are constants. If e_2 is constant,

$$E_c = K - \frac{1}{na}\log_\epsilon e_1 \tag{15-15}$$

But

$$db = 20\log_{10}\frac{e_2}{e_1} = K' - 8.69\log_\epsilon e_1 \tag{15-16}$$

Therefore

$$db = K' - 8.69na(K - E_c) \tag{15-17}$$
$$db = k + mE_c \tag{15-18}$$

where K, K', k, and m are constants. This type of voltmeter should, therefore, theoretically have a linear scale when calibrated in decibels. Actually some departure from linearity will result because e_2 does not remain entirely constant and because the relation between amplification and bias may not be quite exponential.

Another type of logarithmic voltmeter is shown in Fig. 15-17.[2] The input voltage is impressed upon the grid of a variable-mu tube such as a

[1] BALLANTINE, S., *Electronics*, January, 1931, p. 472.
[2] HUNT, F. V., *Rev. Sci. Instruments*, **4**, 672 (1933).

35, 58, or 6K7. The voltage developed across the load resistor is applied to a diode rectifier circuit. During the positive half cycle of output voltage, current flows through the diode circuit. Because of the exponential variation of amplification factor of the pentode with grid voltage, the average diode current varies logarithmically with pentode input voltage over a large range of input voltage. With the single-stage circuit of Fig. 15-17, Hunt obtained a linear calibration curve of microamperes average diode current against decibels from about 3 db below 1 volt to 20 db above 1 volt. In the two-stage circuit of Fig. 15-18 the rectified current is obtained principally from the second stage at low voltages, and from the first stage at voltages that make the second stage input so great as to reduce the plate current to zero during almost half the cycle. For the two-stage circuit the calibration is linear from − 40 db to + 20 db; for a three-stage circuit from − 60 db to + 20 db (relative to 1 volt).

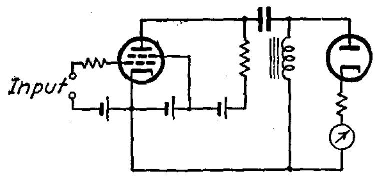


FIG. 15-17.—Logarithmic vacuum-tube voltmeter using variable-mu tube and diode rectifier.

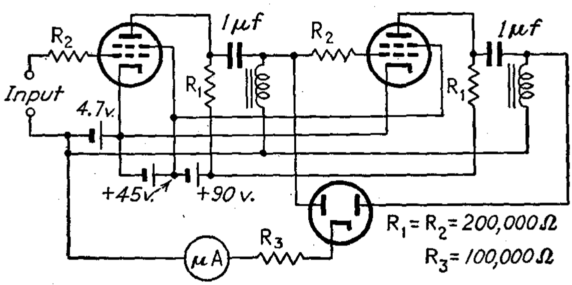


FIG. 15-18.—Two-stage logarithmic vacuum-tube voltmeter.

15-8. Voltmeters Based on Voltage or Current Amplifiers.—Any of the voltmeters that have been described may be preceded by amplifiers in the measurement of very small voltages. A high-gain amplifier may be used directly to measure small voltages, the output voltage or current of the amplifier being read by means of an electromagnetic meter of comparatively low sensitivity. Johnson and Neitzert[1] have described a five-stage pentode amplifier capable of reading down to 10^{-6}, or to 10^{-8} volt with an auxiliary tuned circuit. The amplifier uses four type 38 voltage pentodes and one power pentode. The voltage amplifier stages have 300,000-ohm coupling resistors and use an operating plate voltage of only 12 volts. The screen voltage is 6 volts. A matching transformer is used between the output stage and the meter.

[1] JOHNSON, E. A., and NEITZERT, C., *Rev. Sci. Instruments*, **5**, 196 (1934).

A single-stage amplifier-type meter, described by R. M. Somers,[1] is shown in Fig. 15-19. In this type of meter the current output is a measure of the alternating input voltage. Its sensitivity is the current sensitivity of the circuit, which depends upon the transconductance of the tube.

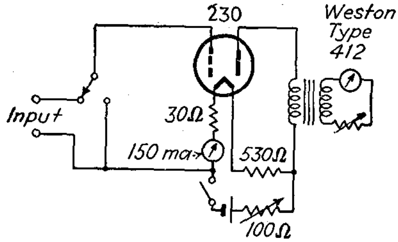


Fig. 15-19.—Vacuum-tube voltmeter based upon current amplifier.

15-9. Slide-back Voltmeters.—A very useful type of meter for the measurement of crest or direct voltages is the slide-back type. The original form of this type of meter is based upon the cutting off of triode plate current by negative grid voltage.[2] Just sufficient negative grid bias is applied to a triode to reduce the plate current to zero. The addition of signal voltage in the grid circuit results in the flow of current during the positive half cycles of the signal voltage and, in order to prevent the flow of plate current at any time during the signal-voltage cycle, the bias must be increased by an amount equal to the positive crest signal. In operation the bias is adjusted to reduce the plate current to zero with and without signal voltage, the difference between the two values of bias equaling the crest signal voltage of the positive half cycle. To read direct voltages, the bias is adjusted to give any convenient reading of plate current. The change in bias required to return the plate current to this value when input is applied is equal to the input voltage.

In making a-c readings, difficulty arises from the fact that cutoff is not sharp. Because the transfer characteristic approaches the grid-voltage axis exponentially, the change in current corresponding to a small change in grid bias decreases with increased sensitivity of the plate-current meter. The accuracy obtainable with a triode is rather low. Much more satisfactory results can be obtained with a tetrode or any multi-element tube in space-charge tetrode connection.[3] The i_b-e_{c2} curve for a positive voltage of 1 or 2 volts on the first grid approaches the voltage axis much more sharply than do the transfer characteristics of triodes. Figure 15-20 shows the diagram of a meter that will read direct voltages to within 0.05 volt and alternating voltages to within $\frac{1}{2}$ per cent. To simplify operation, separate voltages are used for zero adjustment and for balancing. The d-c meter then reads the crest voltage directly. For measuring alternating voltages the balancing voltage must always be negative. For measuring direct voltages it is advisable to apply the

[1] Somers, R. M., *Proc. I.R.E.*, **21**, 56 (1933).
[2] Heising, R. A., U. S. Patent 1232919.
[3] Reich, H. J., Marvin, G. S., and Stoll, K., *Electronics*, September, 1931, p. 109.

unknown voltage in such a direction as to make the grid negative, and so the balancing voltage must be positive.

The inaccuracy resulting from lack of sharpness of cutoff can also be overcome by working at a small value of plate current, instead of at cutoff.[1] Because of plate detection, the change in bias required to keep the plate current constant when signal voltage is applied is not exactly equal to the crest signal voltage. The meter must therefore be calibrated.[2] The calibration curve of balancing voltage *vs.* crest signal voltage approaches a straight line of unit slope through the origin as the operating plate current is reduced. Although the calibration is practically independent of plate and heater voltages and tube age, absence of the direct comparison feature of the cutoff type of meter is often a disadvantage of the small-current type.

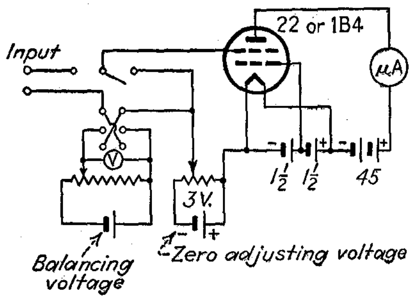


FIG. 15-20.—Slide-back voltmeter using tetrode with space-charge connection.

The use of a grid-controlled arc-discharge tube in place of a vacuum triode affords a third means of increasing the accuracy of the slide-back type of meter.[3] If care is taken to maintain the operating temperature constant, the grid voltage that prevents the firing of an arc-discharge tube at a given anode voltage is very definite. The difference between the grid bias voltages at which the tube fires with and without signal voltage in the grid circuit is equal to the crest signal voltage, or an unknown direct voltage. A circuit that uses an 885 thyratron tube is shown in Fig. 15-21. R_1 and R_2 are current-limiting resistors. No anode-current meter is necessary, since firing is indicated by glow.

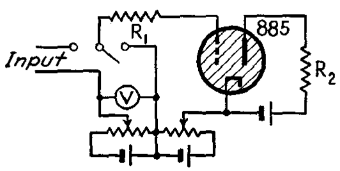


FIG. 15-21.—Slide-back voltmeter using a type 885 thyratron.

A. T. Starr[4] has described a similar voltmeter in which the arc tube is replaced by a trigger circuit. A diagram of the circuit is shown in Fig. 15-22. To operate the meter the switch S is first opened momentarily in order to stop the plate current of T_1. The balancing voltage is adjusted until the current transfers from T_2 to T_1. The operation

[1] MEDLAM and OSCHWALD, *loc. cit.*

[2] For a discussion of the error in reading of a slide-back voltmeter resulting from plate rectification and the method of constructing a calibration curve, see C. B. AIKEN and L. C. BIRDSALL, *Trans. Am. Inst. Elec. Eng.*, **57**, 173 (1938).

[3] HUGHES, E., *J. Sci. Instruments*, **10**, 180 (1933); RUIZ, J. J., *Rev. Sci. Instruments*, **6**, 169 (1935).

[4] STARR, A. T., *Wireless Eng.*, **12**, 601 (1935).

is then repeated with signal voltage applied. The difference in balancing voltages equals the direct or crest alternating voltage. The ease of operation may be increased by the use of separate zero-adjustment and balancing voltages, as in Figs. 15-20 and 15-21.

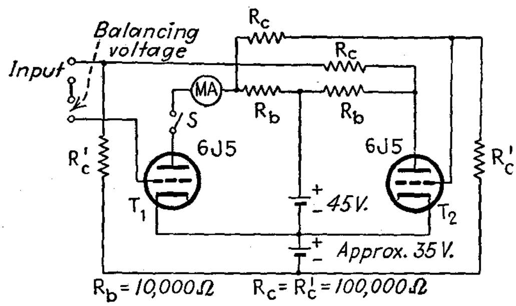


FIG. 15-22.—Slide-back voltmeter based upon the Eccles-Jordan trigger circuit.

15-10. Inverted Voltmeter.—A very useful meter for the measurement of high voltages has been described by Terman.[1] This meter is called the *inverted voltmeter* because the usual functions of the grid and the plate are interchanged. Its action depends upon the fact that the grid current which flows when a positive voltage is applied to the grid of a triode may be controlled by negative plate voltage. Figure 15-23 shows the method of measuring high direct voltages. The amplification factor of the tube for inverted operation is very nearly the reciprocal of the ordinary amplification factor and is therefore less than unity. The plate-grid transconductance, rate of change of grid current with plate voltage, is also small. Therefore large plate voltages may be read by means of small grid currents, and the grid operating voltage may be small. Circuits analogous to the ordinary plate- and grid-detection voltmeters may also be used with the inverted tube. Since their action is similar to that of the common circuits, they need not be discussed further. In addition to the large voltage range, the inverted voltmeter has the advantage of high input resistance and low effective input capacitance.

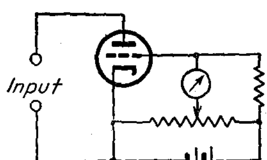


FIG. 15-23.—Inverted vacuum-tube voltmeter.

15-11. Vacuum-tube Ammeters.—The use of vacuum tubes in the measurement of large currents usually offers no particular advantage over the use of ordinary meters unless it is necessary to measure crest values of current. For the latter purpose some form of crest vacuum-tube voltmeter may be used to measure the voltage drop across a resistor through which the current flows. Because of the difficulty of making

[1] TERMAN, F. E., *Proc. I.R.E.*, **16**, 447 (1928).

sensitive dynamometer-type instruments for the measurement of very small currents, the vacuum tube is of great value in this field of measurement. A number of methods are available for the measurement of small currents by means of vacuum tubes.[1]

One of the most sensitive methods of measuring small direct currents consists in determining the change in potential of a small condenser of negligible leakage which is charged by means of the unknown current. The sensitivity of this method increases with decrease in condenser size, and for extreme sensitivity the condenser consists merely of the grid-to-cathode capacitance of a vacuum tube. The charge that collects on the grid and associated conductors determines the grid potential and hence the plate current. This method depends for its success on the possibility of reducing all grid leakage to a negligible amount. Grid leakage current is caused by (1) leakage over the surface of grid supports within the tube, over the surface of the glass outside of the tube, and from external conductors connected to the grid; (2) initial velocity of emitted electrons, which causes them to go from cathode to grid when negative grid voltage is small; (3) primary electrons emitted by the grid and attracted to other electrodes of higher potential; (4) secondary electrons emitted by the grid; (5) ionization of residual gas. Tubes designed to prevent grid leakage were first described by Nelson[2] and by Metcalf and Thompson.[3] A good example of this type of tube is the General Electric FP-54 vacuum-tube electrometer,[4] which is capable of measuring currents as small as 10^{-17} amp. Leakage in this tube is prevented by mounting the grid on quartz supports, by thorough evacuation, and by choice of electrode structure and operating voltages that prevent primary and secondary electron emission from the grid and ionization of residual gas. The FP-54 is not suitable for voltage amplification, since its amplification factor is unity.

P. A. MacDonald and E. M. Campbell[5] have shown that ordinary amplifier tubes such as the type 22 may be used with floating grid to read currents down to 2×10^{-16} amp. Under the action of primary electrons from the cathode, primary and secondary electrons from the grid, and leakage from the grid over the surface of the glass supports, a floating grid assumes an equilibrium value of voltage, which is usually negative with respect to the cathode. If a small current from an external

[1] An excellent review of this subject is given by P. A. MacDonald in *Physics*, **7**, 265 (1936). A bibliography of 29 items is included.

See also F. Müller and W. Dürichen, *Z. Elektrotech.*, **42**, 31 (1936) (with bibliography of 54 items).

[2] Nelson, H., *Rev. Sci. Instruments*, **1**, 281 (1930).

[3] Metcalf, G. F., and Thompson, B. J., *Phys. Rev.*, **36**, 1489 (1930).

[4] Similar tubes are made by the Westinghouse and Western Electric Co.

[5] MacDonald, P. A., and Campbell, E. M., *Physics*, **6**, 211 (1935).

source is allowed to flow into the grid, the equilibrium grid voltage is changed, and therefore also the plate current.

A third method of measuring small currents is to measure by means of a vacuum tube the voltage drop produced when the current flows through a very high resistance. The size of the resistor that can be used, and hence the sensitivity that can be attained, are limited by the

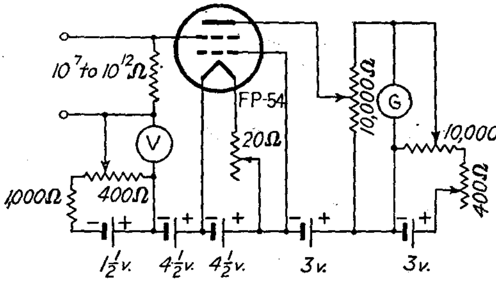
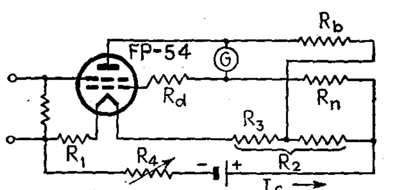


FIG. 15-24.—Circuit for the measurement of small currents with the FP-54 vacuum-tube electrometer.

leakage resistance of the grid. The FP-54 is suitable for use in such a circuit. In Fig. 15-24 is shown a simple circuit employing one FP-54 tube. It may be followed by a current amplifier if a milliammeter is to be used in place of the plate galvanometer. The sensitivity of this simple circuit is limited by fluctuations of battery voltages and by difficulties in balancing the steady component of plate current. Numerous papers have been published on single-tube bridge circuits that are designed to make readings independent of small fluctuations of battery voltage. A circuit that incorporates the advantages of a number of single-tube bridge circuits is shown in Fig. 15-25.[1] Essential features of balanced single-tube circuits are the use of a single battery to supply all voltages and the choice of such values of operating voltages that small changes of electrode voltages compensate one another. The circuit constants given in Fig. 15-25 are average values suitable for use with the Western Electric type D-96475 electrometer tube. Figure 15-26 shows a balanced two-tube circuit.[2]



$R_b = 15{,}000\,\Omega$ $R_1 = 11.1\,\Omega$ $I_f = 270$ mils
$R_d = 0$ $R_2 = 27.8\,\Omega$
$R_n = 9{,}700\,\Omega$ $R_3 = 15.9\,\Omega$

FIG. 15-25.—Electrometer circuit that is compensated against changes of battery voltage. Circuit constants listed are for Western Electric type D-96475 tube.

A very useful microammeter that is capable of reading down to 10^{-8} amp and that uses ordinary amplifier tubes, has been described by A. W.

[1] PENICK, D. B., *Rev. Sci. Instruments*, **6**, 113 (1935).

[2] MOLES, F. J., *Gen. Elec. Rev.*, **36**, 156 (1933); see also DUBRIDGE, L. A., *Phys. Rev.*, **37**, 392 (1931).

Vance.[1] This meter, which is essentially a d-c amplifier with inverse feedback, is accurate and sturdy, is not damaged by overload, and can be designed with as many direct-reading ranges as desired. The basic

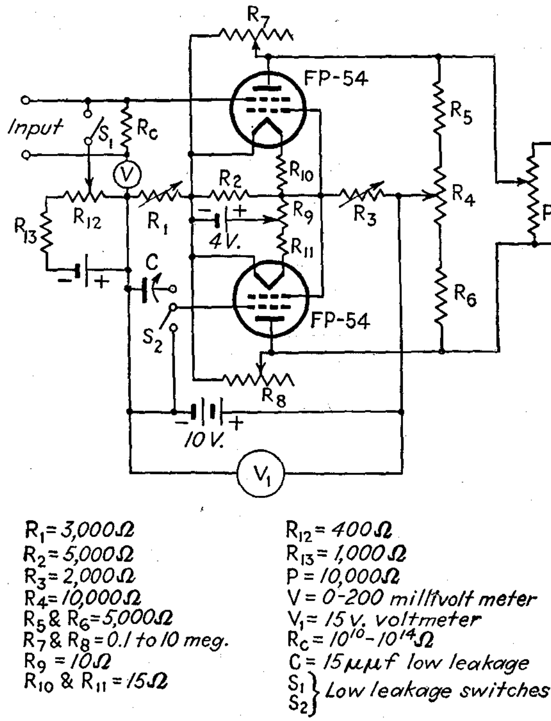


FIG. 15-26.—Balanced two-tube electrometer circuit.

diagram is shown in Fig. 15-27. If the voltage amplification of the amplifier without feedback is A, the output voltage is

$$e_o = Ae_g \tag{15-19}$$

With inverse feedback the voltage actually applied to the grid of the amplifier is

$$e_g = e_i - e_o \tag{15-20}$$

where e_i is the total voltage impressed upon the input. Therefore,

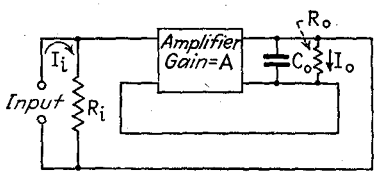


FIG. 15-27.—Feed-back amplifier that may be used as the basis of a vacuum-tube microammeter.

$$e_o = e_i \frac{A}{A + 1} \tag{15-21}$$

Equation (15-21) shows that, if A is very large, e_o is to a first approxima-

[1] VANCE, A. W., *Rev. Sci. Instruments*, **7**, 489 (1936).

tion equal to e_i, and

$$\frac{i_o}{i_i} = \frac{R_i}{R_o} \tag{15-22}$$

If the amplifier is direct-coupled, it will respond to changes of direct current. Equation (15-22) then becomes

$$\frac{\Delta I_o}{I_i} = \frac{R_i}{R_o} \tag{15-23}$$

in which ΔI_o is the change in output plate current when a direct current I_i is sent through the input resistance. By making the ratio of R_i to R_o large, small values of I_i may be read by means of a milliammeter in series with R_o or a voltmeter in shunt with R_o. The function of C_o is to prevent oscillation of the amplifier.

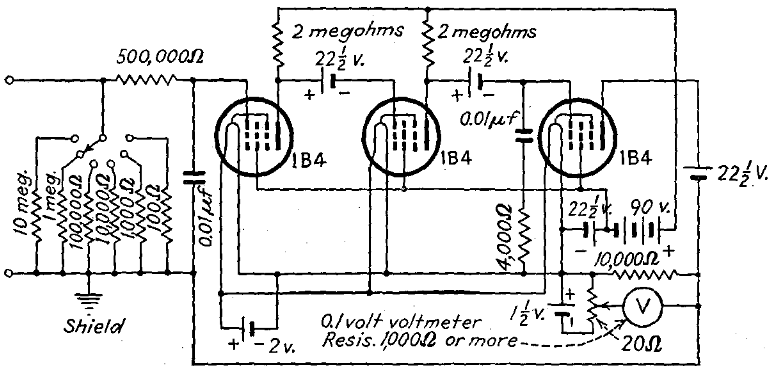


FIG. 15-28.—Circuit diagram of microammeter based upon inverse feed-back amplifier.

The circuit diagram of a practical meter of this type is shown in Fig. 15-28. The circuit employs a three-stage d-c amplifier having an amplification of approximately 5000. With no input to the meter the circuit assumes an equilibrium state in which a definite current flows through the output resistance. To make the meter direct reading, this equilibrium current is balanced out. The size of input resistor required for any range can be readily determined by making use of the fact that the input and output voltages are practically equal. Thus the required resistance is equal to the full-scale reading of the output voltmeter divided by the desired full-scale current reading. Amplifier oscillation is prevented by the resistance-capacitance filter in the input to the last tube. The function of the input filter is to by-pass alternating current. The form of Eq. (15-21) shows that if A is large, small changes of A, such as might result from changes of battery voltage, have little effect upon the reading. A rugged, accurate voltmeter of low sensitivity may be used in the output. Danger of damage to the meter in case of excessive input current can be prevented by designing the output stage so that the maximum attainable plate current will not injure the voltmeter.

Figure 15-29 shows the circuit of an interesting current amplifier that employs phototubes and a triode amplifier in a bridge circuit.[1] The two phototubes determine the grid potential of the triode, which forms the adjustable arm of the bridge. A null galvanometer, operated by the unbalance current, is optically connected to the phototubes so that deflection causes a differential change in conductance, preferably an increase in one and a decrease in the other. Unbalance current deflects the galvanometer, changing the grid potential, and thus the plate resistance, of the triode. The deflection is just sufficient to restore balance and thus reduce the galvanometer current to zero. The insertion of a small potential E into the galvanometer circuit at the point X causes the galvanometer to deflect and thus unbalance the bridge to such an extent that the unbalance current I causes a voltage drop in the resistance that is equal and opposite to E. The current I may be used as a measure of the voltage E. Since the unbalance current necessary to restore

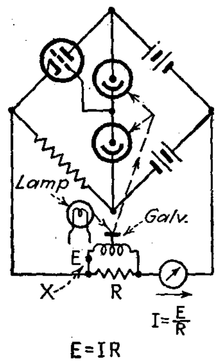


FIG. 15-29.—Direct-current amplifier that may be used for the measurement of small voltages or currents.

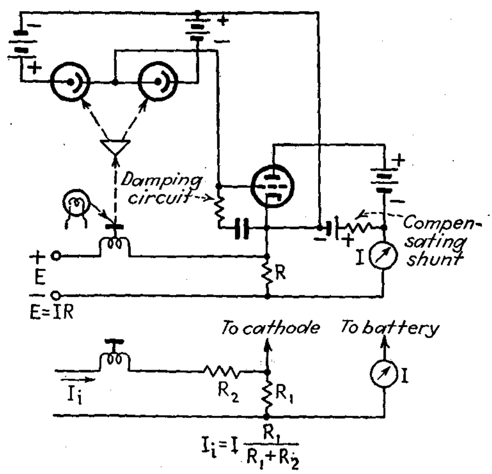


FIG. 15-30.—Practical forms of the circuit of Fig. 15-29 for the measurement of voltage and of current.

equilibrium is inversely proportional to R, it is evident that the sensitivity may be made high by using a small value of R. No spring is required in the galvanometer, and damping may be accomplished by the

[1] GILBERT, R. W., *Proc. I.R.E.*, **24**, 1239 (1936).

use of resistance and capacitance connected in series between the grid and the cathode of the amplifier tube.

Figure 15-30 shows practical forms of this circuit for the measurement of voltage and of current. These circuits have been used to regulate the voltage of a d-c line and to maintain constant temperature. The voltage was held constant to within one part in 10,000 over a load-current range of 0 to 3 amp, and temperature was held constant to within 0.005°C.

15-12. Vacuum-tube Wattmeter.—The balanced modulator may be used as the basis of a vacuum-tube a-c wattmeter.[1] A d-c microammeter is used in the output to measure the difference between the plate currents of the two tubes, as shown in Fig. 15-31. The resistances r_2 are equal and of sufficiently small value so that voltage drop across them does not appreciably reduce the load voltage. Hence e_1 may be considered to be proportional to the load voltage. The resistance r_1 is high in order to prevent appreciable power loss. The voltage applied to the grid of tube 1 is $e_1 + e_2$; that applied to the grid of tube 2 is $e_1 - e_2$. If only the first two terms of the series expansion are considered, and the circuit is assumed to be symmetrical, the plate currents are

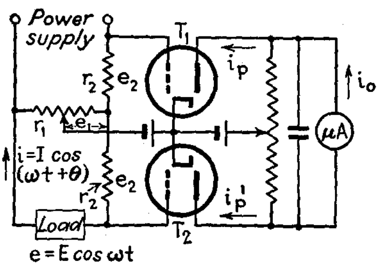


FIG. 15-31.—Vacuum-tube wattmeter circuit.

$$i_p = a_1(e_1 + e_2) + a_2(e_1 + e_2)^2 \tag{15-24}$$
$$i_p' = a_1(e_1 - e_2) + a_2(e_1 - e_2)^2 \tag{15-25}$$

The output meter reads the difference between i_p and i_p'.

$$i_o \propto i_p - i_p' = 2a_1e_2 + 4a_2e_1e_2 \tag{15-26}$$

But

$$e_1 = ke = kE \cos \omega t \tag{15-27}$$

where k is the fraction of r_1 from which e_1 is taken; and

$$e_2 = r_2 i = r_2 I \cos (\omega t + \theta) \tag{15-28}$$

If the output meter is a d-c instrument, it reads the average value of i_o. The average value of i_o is

$$I_o = \frac{\omega}{2\pi}\int_0^{2\pi/\omega} i_o\, dt \propto \frac{2a_1\omega}{2\pi}\int_0^{2\pi/\omega} r_2 I \cos (\omega t + \theta)\, dt + \frac{4a_2\omega}{2\pi}\int_0^{2\pi/\omega} (kE \cos \omega t)[r_2 I \cos (\omega t + \theta)]\, dt \tag{15-29}$$

$$I_o \propto 0 + 2a_2 k r_2 EI \cos \theta \tag{15-30}$$

[1] PETERSON, E., U. S. Patent 1586553, 1926; TURNER, H. M., and MCNAMARA, F. T., *Proc. I.R.E.*, **18**, 1743 (1930). See also T. B. WAGNER, *Elec. Eng.*, **53**, 1621 (1934) and J. R. PIERCE, *Proc. I.R.E.*, **24**, 577 (1936).

Equation (15-30) shows that the reading of the d-c meter in the plate circuit is proportional to the power expended in the load. A more complete analysis shows that, if the load current contains harmonics, the reading of the d-c meter is proportional to the total load power. The effect of higher-order terms of the plate-current series modifies the expression given in Eq. (15-30), and at large amplitudes the meter reading is not strictly proportional to the load power. Since the meter must be calibrated, this is a disadvantage only in that it prevents the use of a linear scale. The effect may be minimized by keeping the input to the tubes small.

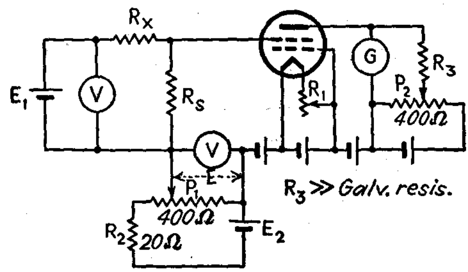


FIG. 15-32.—Vacuum-tube electrometer circuit for the measurement of high resistance.

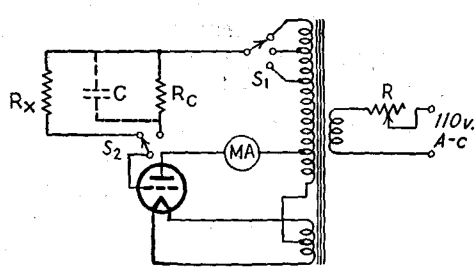


FIG. 15-33.—Vacuum-tube ohmmeter that makes use of grid current.

15-13. Vacuum-tube Ohmmeters.—Circuits capable of measuring small currents can be readily modified to measure high resistances. Figure 15-32 shows an FP-54 circuit for measuring high resistances.[1] An unknown resistance R_x is measured by applying the voltages E_1 and E and adjusting the voltage divider P_1 until the galvanometer reading is the same as when E_1 and E are equal to zero. Then E is equal to the voltage across the standard resistor R_s. $R_x = R_s(E_1/E - 1)$. Standard resistors of values 10^6, 10^9, 10^{12}, and 10^{15} ohms are recommended. The value of E_1 is optional; that recommended for E_2 is approximately $E_1/2$.

Another type of ohmmeter, which may be calibrated, is shown in Fig. 15-33. It depends for its action upon the flow of grid current.[2]

[1] ROSE, G. M., JR., *Rev. Sci. Instruments*, **2**, 810 (1931).

[2] PREISMAN, A., *Electronics*, July, 1935, p. 214.

Preisman found that, if the meter reads correctly at two points of its range, it will read correctly over the whole scale. It is adjusted by first setting the milliammeter needle to zero mechanically when the grid is floating and then adjusting S_1 and R so that the reading is correct for the calibrating resistance R_s. The function of the condenser C is to minimize the effect of capacitance in the resistance that is being measured. It may ordinarily be omitted.

The glow-tube relaxation oscillator can be used to measure resistance and capacitance with a high degree of accuracy by comparison with known values.[1] The variable standard resistance or capacitance of Fig. 15-34 is adjusted so that the frequency of oscillation is the same as with the unknown resistance or capacitance. If one of the parameters is fixed, the circuit may be calibrated to read the other in terms of frequency. The oscillation frequency may be made low enough so that the frequency can be determined by counting flashes of the tube during a measured time interval. The accuracy of this instrument in the measurement of condenser capacitance is affected by leakage conductance of the condenser, since the frequency of oscillation depends not only upon the capacitance but also the conductance.

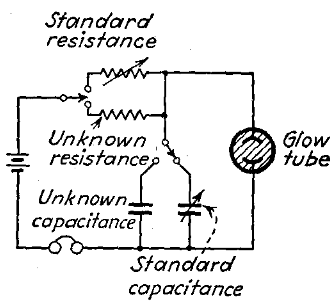


FIG. 15-34.—Glow-tube circuit for the measurement of resistance or capacitance.

15-14. Time and Speed Meters.—Time intervals may be accurately measured with the aid of electron tubes.[2] In most electronic time and speed meters, tubes are used to control the charging or discharging of a condenser and to read the change in condenser voltage, which is proportional to the elapsed time. To use the charging of a condenser in the measurement of time, means must be provided for starting and stopping the charging current instantaneously at the beginning and end of the time interval. This may be accomplished by the parallel d-c thyratron circuits of Fig. 12-34 or by the high-vacuum-tube trigger circuits of Figs. 10-6 and 10-8.

Figure 15-35 shows the circuit diagram of a timer based on this principle.[3] The action of this circuit is as follows: S_2 is closed for an

[1] TAYLOR, J., and CLARKSON, W., *J. Sci. Instruments,* **1,** 173 (1924).

[2] SPEAKMAN, E. A., *Rev. Sci. Instruments,* **2,** 297 (1931); ROBERDS, W. M., *Rev. Sci. Instruments,* **2,** 519 (1931); SIXTUS, K. J., and TONKS, L., *Phys. Rev.,* **37,** 930 (1931); DU BOIS, R., and LABOUREUR, L., *Compt. rend.,* **194,** 1639 (1932); PARTRIDGE, H. M., *Electronics,* August, 1932, p. 262; LORD, H. W., *Electronics,* October, 1932, p. 309; STEENBECK, M., and STRIGEL, R., *Arch. Elektrotech.,* **26,** 831; *Wireless Eng.,* **10,** 225 (1933) (abstr.); SHEPARD, F. H., JR., *Electronics,* February, 1935, p. 59.

[3] REICH, H. J., and TOOMIM, H., *Rev. Sci. Instruments,* **8,** 502 (1937).

instant to ensure that plate current flows in T_2 and not in T_1. C is then discharged by closing S_3 for an instant. The flow of the plate current of T_2 through R_2 produces a voltage drop which biases the control grid of T_3 beyond cutoff. The closing of S_1 for an instant then causes plate current to transfer from T_2 to T_1 and thus reduces the current in R_2 to such a low value that T_3 is practically unbiased, and plate current flows through T_3. The plate current of T_3, which is nearly constant, charges C. The subsequent closing of S_2 causes plate current to flow again through T_2, thus stopping the charging current. The change in the reading of the plate milliammeter of T_4 caused by the change in voltage of the condenser gives a measure of the time between the closing

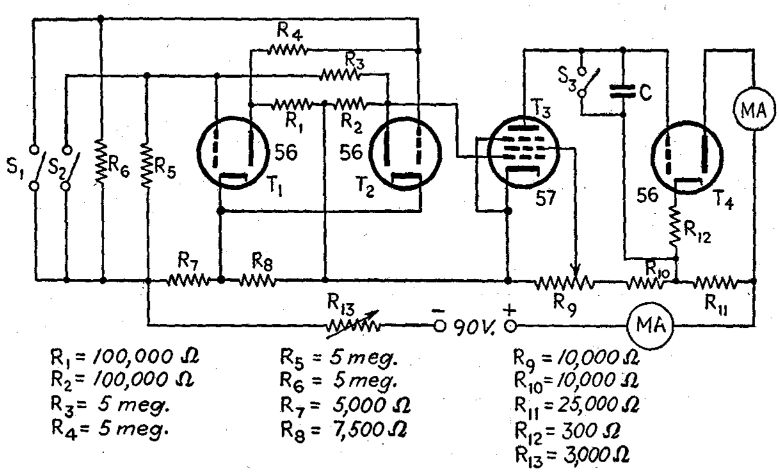


FIG. 15-35.—Circuit for the measurement of time intervals.

of S_1 and of S_2. If phototubes are connected in series with R_3 and R_4, the circuit may be controlled by the interruption of light beams, instead of by keys.

A variant of this circuit can also be used to measure the average speed of moving objects by means of two switches or light beams. Instantaneous speeds may be approximated by placing the switches or beams so close together that there is little change in velocity over the interval.

These circuits may be calibrated on the basis of known circuit constants and measured current of T_3, but better results are obtained by experimental calibration. The circuit of Fig. 15-35 may be calibrated by means of a contactor which closes S_2 after S_1. The modified form using phototubes may be calibrated by interrupting the light beams by means of a pendulum or a falling mass. The time range may be changed by adjusting the screen voltage of T_3 and the size of C. Changes in battery voltage are compensated by adjustments of R_{13}.

Figure 15-36 shows a time-interval meter based upon the thyratron switching circuit of Fig. 12-33.[1] Instantaneous closing of switch S_1 fires the thyratron and initiates the flow of current through the ballistic galvanometer. Subsequent closing of switch S_2 extinguishes the thyratron and stops the flow of current through the galvanometer. The

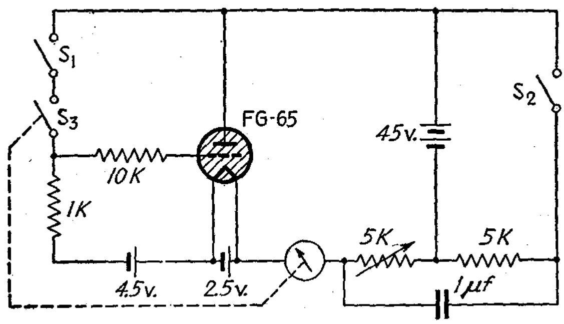


FIG. 15-36.—Circuit for the measurement of time intervals.

deflection of the meter is proportional to the time interval that elapses between the closing of the two switches. Switch S_3, which is opened by movement of the galvanometer, inactivates switch S_1 and thus ensures extinction of the tube when S_2 is closed, even though S_1 may still be closed.

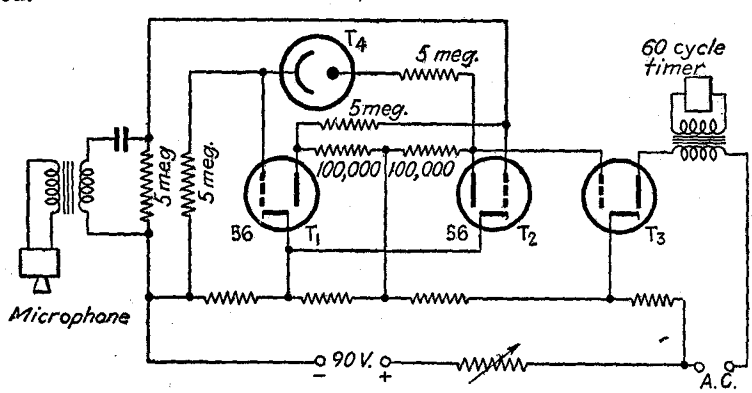


FIG. 15-37.—Electronic race timer.

If a-c power is available, time may be measured by means of an a-c–operated meter that measures time in terms of 1/60-sec intervals.[2] A race timer that uses such a meter, started and stopped by means of a trigger circuit, was devised by W. M. Roberds.[3] A simplified form of Roberds's circuit, requiring only one B supply, is shown in Fig. 15-37.

[1] REICH, H. J., and TOOMIM, H., *Rev. Sci. Instruments*, **12**, 96 (1941).

[2] KLOPSTEG, P. E., *J. Opt. Soc. Am. and Rev. Sci. Instruments*, **19**, 345 (1929).

[3] ROBERDS, *loc. cit.*

Sound from the starting gun, striking the microphone, stops the flow of current in T_2. This reduces the grid bias of T_3 and starts the operation of the timer. Interruption of light falling on T_4 starts the flow of current in T_2, thus biasing T_3 beyond cutoff and stopping the timer.

15-15. Counting Circuits.—Thyratron circuits for "scaling down" the rate of occurrence of random voltage pulses in order to make possible

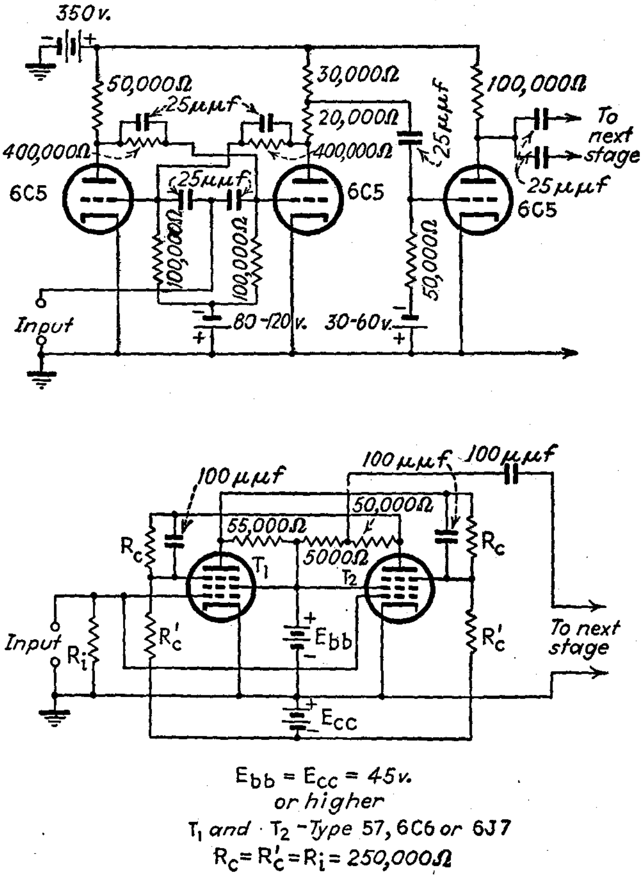


FIG. 15-38.—Counting circuits based upon high-vacuum-tube trigger circuits.

their counting by means of relatively slow-acting electromagnetic counters were discussed in Sec. 12-32. A number of circuits have also been devised in which high-vacuum-tube trigger circuits have been adapted to this purpose.[1] Single stages of two such circuits are

[1] LEWIS, W. B., *Proc. Cambridge Phil. Soc.*, **33**, 549 (1937); STEVENSON, E. C., and GETTING, I. A., *Rev. Sci. Instruments*, **8**, 414 (1937); LIFSCHUTZ, H., and LAWSON, J. L., *Rev. Sci. Instruments*, **9**, 83 (1938); REICH, H. J., *Rev. Sci. Instruments*, **9**, 222 (1938). See also supplementary bibliography on trigger circuits, p. 415.

shown in Fig. 15-38.[1] In these circuits the voltage pulses are applied simultaneously, and in the same polarity, to both tubes of an Eccles-Jordan trigger circuit. Each pulse transfers the current from one tube to the other. The abrupt change in current in either plate resistor is used to produce a voltage pulse which is applied to the grids of the next stage. By the use of n similar stages, the final one of which controls an electromagnetic counter, the rate at which the counter must operate is reduced in the ratio 2^n. Circuits of this type are more rapid in operation than thyratron circuits but require input pulses of very steep wave front.

15-16. Frequency Meters.—Figure 15-39 shows the circuit of an electronic frequency meter.[2] Frequency is measured in this meter by

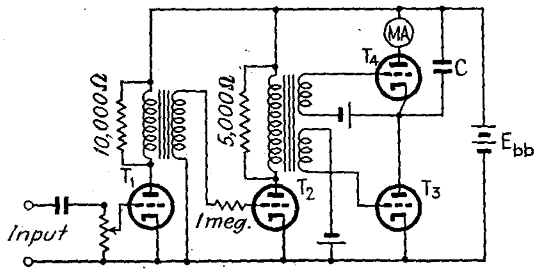


FIG. 15-39.—Electronic frequency meter.

means of the charging current of a condenser which is periodically charged through the plate of one triode and discharged through the plate of another. The action is as follows: The bias of T_3 and T_4 is great enough to prevent the flow of plate current when no signal voltage is applied to the input of T_1. The excitation voltages of T_3 and T_4 are in phase opposition, so that when voltage is applied to the input of the meter, plate current can flow alternately in T_3 and T_4. During one-half of the cycle of applied voltage, C charges through T_3, and during the other half cycle it discharges through T_4 and the milliammeter. If the amplitudes of the exciting voltages of T_3 and T_4 are high, and C is not too large, charge and discharge of the condenser are practically complete. The average condenser discharge current is equal to

$$I = CE_{bb}f \tag{15-31}$$

It may, therefore, be used as a direct measure of the frequency of the input voltage. The functions of T_1 and T_2 are to amplify the input voltage and to increase the input impedance of the meter. The reading is independent of input amplitude provided that the amplitude is great enough to ensure "complete" charging and discharging. Because of the falling of amplification of transformer-coupled amplifiers at very low

[1] LIFSCHUTZ and LAWSON, *ibid.*; REICH, *ibid.*

[2] GUARNASCHELLI, F., and VECCHIACCI, F., *Proc. I.R.E.*, **19**, 659 (1931).

and very high frequencies, the minimum input voltage that will give a correct frequency reading varies with frequency. A wider frequency range can be attained by the use of resistance-capacitance coupling and a phase inverter in place of transformer coupling. For high sensitivity to small changes of frequency a portion of the meter current may be balanced out by a circuit similar to that of Fig. 15-4. Since the reading is dependent upon the plate supply voltage E_{bb}, the voltage should be held constant by means of a glow tube or other type of voltage stabilizer, or provision should be made for adjusting the voltage to a reference value.

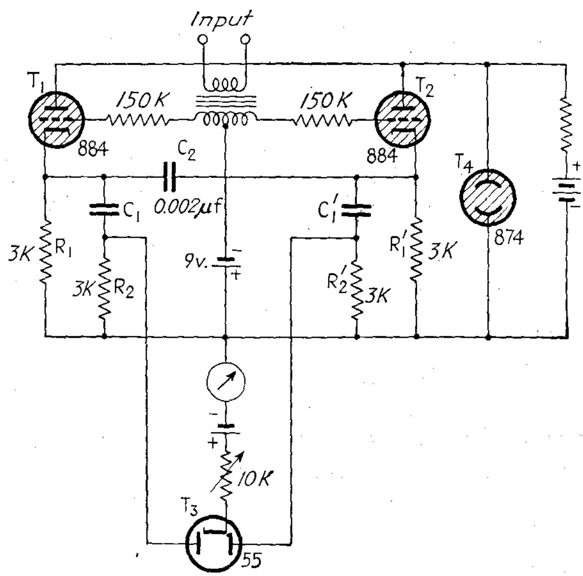


FIG. 15-40.—Frequency meter based upon the thyratron parallel switching circuit of Fig. 12-34*b*.

Figure 15-40 shows a frequency meter based upon the thyratron switching circuit of Fig. 12-34*b*.[1] The input voltage causes periodic switching of plate current from tube T_1 to tube T_2. When current starts flowing in T_1, the resulting voltage across R_1 causes the condenser C_1 to charge through the parallel paths afforded by R_2 and one plate of the diode rectifier T_3. When the current is switched to T_2, C_1 discharges through R_1 and R_2 and at the same time C_1' charges through R_2' and the other plate of T_3. The average rectified diode current is indicated by the meter in series with the diode cathode. Throughout the frequency range in which C_1 and C_1' charge "fully" in less than the time of one half cycle of impressed voltage, the average diode current, and hence the reading of the meter, are proportional to frequency. For the circuit

[1] HUNT, F. V., *Rev. Sci. Instruments*, **6**, 43 (1935).

values shown in Fig. 15-40 the calibration of the instrument is linear up to 7000 cps and approximately so at higher frequencies.

One of the simplest methods of measuring audio frequencies is by means of some form of impedance bridge. Although they do not ordinarily incorporate vacuum tubes, a brief discussion of impedance bridges will be included in this section because of their frequent use in the study and application of vacuum tubes and vacuum-tube circuits.

The fundamental criterion for the balance of a four-arm impedance bridge is that the product of one pair of opposite impedances shall be equal in magnitude and phase to the product of the other pair. Since the voltages must be balanced for both magnitude and phase, it is necessary to provide the bridge with both a variable resistance and a variable reactance.

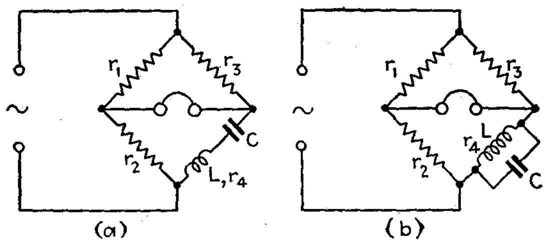


FIG. 15-41.—Belfils resonance bridges: (*a*) series; (*b*) parallel.

Simple forms of impedance bridges are the resonance bridges of Fig. 15-41.[1] Balance is obtained by varying either L or C and either r_2 or r_3. The student may readily show that the bridge of Fig. 15-41*a* is balanced when

$$r_2 = \frac{r_1}{r_3} r_4 \tag{15-32}$$

and

$$\omega L = \frac{1}{\omega C} \tag{15-33}$$

or

$$f = \frac{1}{2\pi \sqrt{LC}} \tag{15-34}$$

The bridge of Fig. 15-41*b* is balanced when

$$r_2 = \frac{r_1}{r_3} \cdot \frac{L}{r_4 C} \tag{15-35}$$

and

$$f = \frac{1}{2\pi} \sqrt{\frac{1}{LC} - \frac{r_4{}^2}{L^2}} \tag{15-36}$$

[1] BELFILS, G., *Rev. gén. élec.*, **19**, 523 (1926).

The series resonance bridge of Fig. 15-41*a* is evidently easier to use than the parallel resonance bridge of Fig. 15-41*b*, for in the former the resistance balance is not affected by the reactance balance and the expression for the frequency is simpler.

Two other forms of impedance bridge are of particular value in the measurement of frequency. The first of these is a modified form of Hay's bridge,[1] shown in Fig. 15-42.[2] For this bridge the general criterion for balance becomes

$$r_2r_3 = \left(r_4 - \frac{j}{\omega C}\right)(r_1 + j\omega L) \tag{15-37}$$

in which r_1 includes the resistance of the inductor L.

Equating the real terms of Eq. (15-37) gives

$$r_4 = \frac{r_2r_3 - L/C}{r_1} \tag{15-38}$$

Equating the imaginary terms of Eq. (15-37) gives

$$\omega^2 = \frac{r_1}{r_4LC} \tag{15-39}$$

Substitution of Eq. (15-38) in Eq. (15-39) yields the relation

$$f = \frac{\omega}{2\pi} = \frac{r_1}{2\pi\sqrt{L(Cr_2r_3 - L)}} \tag{15-40}$$

Thus, if the bridge is balanced by means of r_1 and r_4, the frequency is directly proportional to r_1. r_1 may be made numerically equal to the frequency by choosing the other bridge parameters so that the radical of Eq. (15-40) is equal to $1/2\pi$. For maximum sensitivity at 1000 cps and direct-reading scale, the parameters should have the following values:

$L = 1/2\pi$ henry, $C = 1/2\pi$ μf, and $r_2 = r_3 = 1414$ ohms.[2]

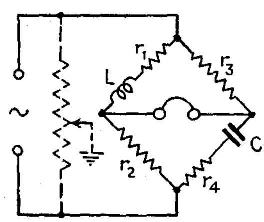


FIG. 15-42.—Modified form of Hay's bridge.

The inductor resistance must be added to the reading of the decade box used for r_1. The bridge should preferably be ungrounded, but, if a ground cannot be avoided, least error will result if the ground is applied at the junction of L and r_2, or through a potentiometer connected across the a-c input (Wagner ground), as shown by the dotted lines in Fig. 15-42. This type of bridge has two disadvantages. At the higher frequencies the distributed capacitance

[1] HAY, C. E., *Inst. P.O. Elec. Eng., Papers*, November, 1912.

[2] SOUCY, C. I., and BAYLY, B. DE F., *Proc. I.R.E.*, **17**, 834 (1929). See also K. KUROKAWA and T. HOASHI, *J. Inst. Elec. Eng., Japan*, **437**, 1132 (1924).

of the inductor destroys the linearity of reading, and, since r_1 must always be inversely proportional to r_4 for balance, the bridge cannot be conveniently balanced by means of a single control.

The disadvantages of the Hay bridge are avoided in the Wien bridge, shown in Fig. 15-43.[1] The criterion for balance gives the following equation:

$$\frac{r_2 r_3}{r_4} = \left(r_1 - \frac{j}{\omega C_1}\right)(1 + jr_2\omega C_2) \tag{15-41}$$

Equating real terms and imaginary terms of Eq. (15-41) gives the two relations

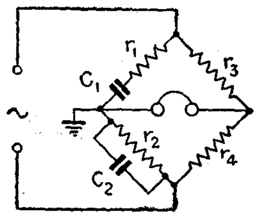


FIG. 15-43.—Wien bridge.

$$r_1 = r_2\left(\frac{r_3}{r_4} - \frac{C_2}{C_1}\right) \tag{15-42}$$

$$\omega^2 = \frac{1}{C_1 C_2 r_1 r_2} \tag{15-43}$$

If $r_3 = 2r_4$ and $C_1 = C_2$, Eq. (15-42) simplifies to

$$r_1 = r_2 \tag{15-44}$$

and Eqs. (15-43) and (15-44) give

$$f = \frac{1}{2\pi r_1 C_1} \tag{15-45}$$

Since the value of r_1 required for balance varies inversely with frequency, this bridge cannot be made direct reading. Because r_1 is equal to r_2 when the bridge is balanced, r_1 and r_2 may be adjusted by means of a single dial which is calibrated to read frequency.

15-17. Cathode-ray Oscillographs (**Oscilloscopes**).—An increasingly important electron tube device is the cathode-ray oscillograph, or *oscilloscope*.[2] Although designed originally for the study of voltage and

[1] FIELD, R. F., *Gen. Radio Expt.*, **6** (November, 1931).

[2] RIDER, J. F., "The Cathode-ray Tube at Work," J. F. Rider, New York, 1935; MALOFF, I. G., and EPSTEIN, D. W., "Electron Optics in Television," McGraw-Hill Book Company, Inc., New York, 1938; WATSON-WATT, R. A., HERD, J. F., and BAINBRIDGE-BELL, L. H., "The Cathode-ray Oscillograph in Radio Research," H. M. Stationery Office, London, 1933; ARDENNE, M. VON, "Die Kathodenstrahlrohren," Julius Springer, Berlin, 1934; MACGREGOR-MORRIS, J. T., and HENLEY, J. A., "Cathode Ray Oscillography," Chapman & Hall, Ltd., London, 1936; KNOLL, M., *Arch. tech. Mess.*, **1**, 76 (1931) (summary of literature to 1931); ZWORYKIN, V. K., *Electronics*, November, 1931, p. 188 (bibliography of 12 items); BATCHER, R. R., *Proc. I.R.E.*, **20**, 1878 (1932) (bibliography of 29 items); STINCHFIELD, J. M., *Elec. Eng.*, **53**, 1608 (1934); BATCHER, R. R., *Instruments*, **9**, 6, 38, 77, 112, 140, 166, 197, 231, 255, 286, 312, 341 (1936); PARR, G., "The Low Voltage Cathode Ray Tube," Chapman and Hall, Ltd., London, 1937 (extensive bibliography); STOCKER, A. C., *Proc. I.R.E.*, **25**, 1012 (1937); MAYER, H. F., *Electronics*, April, 1938, p. 14; PREISMAN, A., *RCA Rev.*, **3**, 473 (1939); OVERBECK, W. P., and LÖF, J. L. C., *Rev. Sci. Instruments*, **11**, 375 (1940); GEOHAGEN, W. A., *Electronics*, November, 1940, p. 36.

current wave form, the cathode-ray oscillograph finds many other useful applications. Its advantages over electromagnetic oscillographs are that it draws negligible current from a source of voltage under observation and that it responds at frequencies ranging up to 100 megacycles. Cathode-ray tubes are also an essential part of television equipment. Cathode-ray tubes for oscillographic use are made in a variety of sizes, ranging from 1-in. to 18-in. screen.

The cathode-ray oscillograph tube consists essentially of an electron gun for producing a beam of rapidly moving electrons called *cathode rays*, a fluorescent screen upon which a luminous spot is produced by the impact of the cathode rays, and means for displacing the spot from its quiescent position as the result of current or voltage applied to the deflecting mechanism. Although the electron beam may be focused by

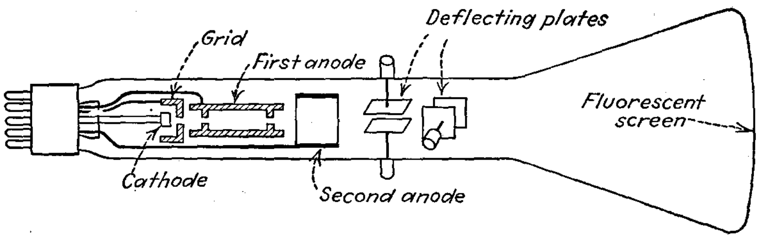


Fig. 15-44.—Electrode structure of typical cathode-ray tube.

means of magnetic fields, electrostatic focusing is usually used. Figure 15-44 shows the electrode structure of a typical cathode-ray tube having an electron gun with electrostatic focusing. The electron gun consists of a thermionic cathode, a grid for controlling the electron density of the beam and hence the brightness of the luminous spot, and two anodes. The final velocity with which the electrons leave the gun is determined by the potential of the second anode, which is normally maintained constant in the operation of the tube. The electrostatic field between the grid and the first anode and between the two anodes focuses the stream of electrons in a manner somewhat analogous to the focusing of light rays by lenses.[1] If the electrodes are properly shaped, the potential of the first anode, which is intermediate between that of the cathode and the second anode, can be adjusted so that the electron beam is brought to a sharp focus in the plane of the fluorescent screen.

The diaphragms in the grid and first anode serve to intercept electrons having velocity components normal to the axis of the beam and thus prevent loss of sharpness of the luminous spot.

[1] Maloff, I. G., and Epstein, D. W., *Proc. I.R.E.*, **22**, 1386 (1934); **23**, 263 (1935); Epstein, D. W., *Proc. I.R.E.*, **24**, 1095 (1936); Maloff, I. G., and Epstein, D. W., "Electron Optics in Television," McGraw-Hill Book Company, Inc., New York, 1938.

Magnetic focusing of the electron beam may be accomplished by means of either a uniform axial field extending the entire length of the tube or a nonuniform field, symmetrical about the tube axis, produced by one or more short concentrated coils, the axes of which coincide with that of the tube. Because of the difficulty of producing a uniform axial field, magnetic focusing is ordinarily accomplished by means of one or more short coils.

15-18. Deflection of Electron Beam.—The electron beam in a cathode-ray tube may be deflected either by means of an electric field produced by two electrodes between which deflection voltage is impressed, or by means of a magnetic field produced by electromagnets through which deflection current flows. The general arrangement of electrodes in a tube in which both vertical and horizontal deflection is produced by electric fields is shown in Fig. 15-44. The deflection is in a plane normal to the deflection plate surfaces and, as pointed out in Sec. 1-20, after deflection the electrons move as though they had originated at a point on the axis opposite the center of the pair of deflecting plates.

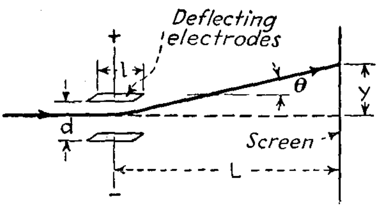


FIG. 15-45.—Electrostatic deflection of cathode-ray beam.

Inspection of Fig. 15-45 shows that the linear displacement of the luminous spot along the screen, under the assumption that the surface of the screen is plane, is

$$y = L \tan \theta \quad \text{cm} \tag{15-46}$$

in which L is the distance in centimeters from the center of the deflection plates to the screen. Substitution of Eq. (1-21) in Eq. (15-46) gives

$$y = \frac{E\mathcal{E}lL}{v_o^2 m_e d} \quad \text{cm} \tag{15-47}$$

in which E is the deflection voltage in statvolts, $\mathcal{E}$ is the charge of an electron in e.s.u., l is the length of the deflection plates in centimeters, v_o is the velocity in centimeters per second with which the electrons enter the deflecting field, m_e is the mass of an electron in grams, and d is the spacing of the deflection plates in centimeters. But

$$\tfrac{1}{2} m_e v_o^2 = E_a \mathcal{E} \tag{15-48}$$

in which E_a is the potential difference in statvolts between the cathode and the second anode. Substitution of Eq. (15-48) in Eq. (15-47) gives

$$y = \frac{ElL}{2E_a d} \tag{15-49}$$

in which E and E_a may be expressed in volts.

Equation (15-49) shows that the deflection of the luminous spot varies linearly with deflection voltage but is inversely proportional to the second anode voltage. Since the brightness of the spot increases with the energy with which the electrons strike the screen, which is in turn proportional to the second anode voltage, a high value of brightness is attained at the expense of deflection sensitivity. For this reason some tubes contain a third anode, which may take the form of a grid adjacent to the screen or of a graphite coating on the wall of the tube. The second anode voltage may thus be made low enough to give high deflection sensitivity and the high velocity required for spot brilliance be given to the electrons subsequent to deflection. Equation (15-49) shows that deflection sensitivity is also increased by lengthening the plates. In order to prevent the electrons from striking the deflection plates at large values of deflection, long plates are usually divergent over a portion of their length.

In most applications of cathode-ray tubes the deflection voltage varies. The deflection is then proportional to the integral of the deflection voltage throughout the time of transit of electrons between the plates. If the rate of change of voltage is sufficiently small so that the voltage is essentially constant during the time of transit, the deflection is proportional to the instantaneous voltage. This is not true, however, when the period of alternating deflection voltage is of the order of magnitude of the transit time between the deflection plates. The high-frequency limit of the deflection voltage that can be analyzed by means of a cathode-ray tube is therefore dependent upon the transit time. Because the transit time is decreased by increase of second anode voltage and by decrease of length of deflection plates, both of which reduce the deflection sensitivity, it follows that tubes designed for use at very high frequencies have low deflection sensitivity. The limiting frequency at which a cathode-ray tube can be used is reduced by capacitance between the deflection electrodes and their leads and by lead inductance. These are minimized by bringing the leads directly through the wall of the tube, as in Fig. 15-44, but for the sake of simplicity of manufacture and convenience of making connections, the leads are brought out through the base in most tubes designed for oscillographic use.

Two coil arrangements that may be used in producing electromagnetic deflection are shown in Figs. 15-46 and 15-47. Because of the large air gap and the low magnetizing force that is ordinarily used, an iron core does not greatly increase the sensitivity but is sometimes used to help in shaping the field. In the arrangement of Fig. 15-47 the coils may be bunch-wound and wrapped about the neck of the tube, as shown, or they may be wound in a slotted core similar to the stator of a two-phase induction motor. In order to prevent interaction between the horizontal and vertical pairs, the coils must be carefully designed and mounted.

The deflection produced magnetically is normal to the flux (see Sec. 1-21). It follows from Eq. (1-29) that the deflection over the surface of a spherical screen is given by the relation

$$y = \frac{1}{3 \times 10^{10}} \sqrt{\frac{\varepsilon}{2m_e}} \frac{sBL}{\sqrt{300E_a}} = 0.297 \frac{sBL}{\sqrt{E_a}} \qquad (15\text{-}49A)$$

in which s is the distance in centimeters throughout which the magnetic field acts, B is the flux density in gauss (assumed to be constant throughout the distance s), L is the distance in centimeters from the center of the field to the screen, and E_a is the second anode voltage in volts. At low frequencies the deflection sensitivity may be increased by increasing the number of turns in the deflection coils, but at high frequencies it is usually necessary to keep the coil inductance small.

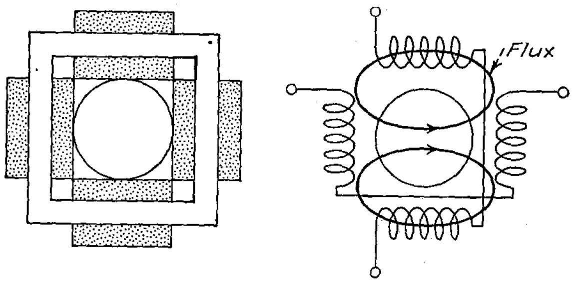


FIG. 15-46.—Structure and flux pattern of coils for electromagnetic deflection of cathode-ray beam.

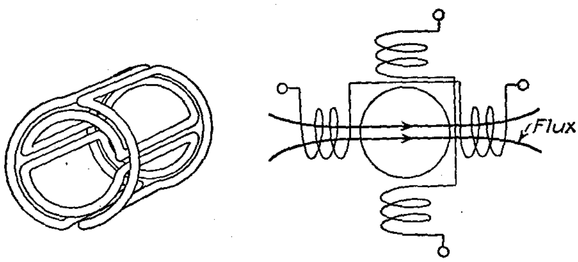


FIG. 15-47.—Structure and flux pattern of coils for electromagnetic deflection of cathode-ray beam.

Most cathode-ray tubes designed for oscillographic use employ electrostatic focusing and deflection. Sometimes deflection in one direction is produced electrostatically, and at right angles magnetically. The obvious advantage of electrostatic deflection is that little or no deflection current or power is required. The auxiliary circuits are therefore much simpler, and difficulties arising from coil inductance are avoided.

Large angular deflection requires that electrostatic deflection plates shall either be short or have large separation in order that they will not intercept the electrons. Since either small length or large separation results in low deflection sensitivity, tubes designed for electrostatic deflection usually use relatively small angular deflection and must consequently be longer than tubes designed for magnetic deflection. The greater structural simplicity of magnetic-deflection tubes makes them cheaper to manufacture and more rugged than electrostatic-deflection tubes. Magnetic deflection is not satisfactory in the study of voltage wave form, since inductance of the deflection coils causes the current wave form to differ from that of nonsinusoidal voltage impressed upon the coils or upon the amplifier that excites the coils (see Fig. 15-54). For this reason magnetic deflection is ordinarily used only in producing a time base.

15-19. Cathode-ray Tube Screens.[1]—The screen of a cathode-ray tube consists of a thin layer of a *phosphor*, which is a material that luminesces as the result of bombardment by rapidly moving electrons. The bombardment gives rise to both *fluorescence* or emission of light during bombardment and *phosphorescence* or emission of light after bombardment. The phosphor is applied to the inside of the end of the tube by spraying, dusting, or precipitation from a liquid. Factors governing the choice of the phosphor for a specific type of tube are the required luminous intensity of the spot per watt of the incident electron beam, the desired rate of decay of phosphorescence, and the desired color of the light, which depends in turn upon whether the tube is to be used visually or photographically. Slow decay of phosphorescence makes possible the visual observation of nonrepeating transients and prevents flicker in the visual observation of periodic voltages or currents of low frequency, but causes blurring whenever an image on the screen moves or changes form. A compromise must therefore be made between absence of blurring and freedom from flicker.

Continued bombardment of one point of the screen reduces the sensitivity of the phosphor and may eventually result in its destruction. Furthermore, at high voltages the impact of the electrons upon the screen may raise its temperature sufficiently to melt the glass. For these reasons the beam should be turned off when the spot is stationary. To avoid reduction of sensitivity of the screen it is also best not to operate the tube over an extended period with an unchanging pattern. Many oscilloscopes are provided with a switch that turns off the beam by

[1] NICHOLS, E. L., HOWES, H. L., and WILBER, D. T., "Cathode-luminescence and Luminescence of Incandescent Solids," *Carnegie Inst. Pub.*, Washington (1928); TOMASCHEK, R., *Die Physik*, **2**, 33 (1934); PERKINS, T. B., and KAUFMANN., H. W., *Proc. I.R.E.*, **23**, 1324 (1935); LEVY, L., and WEST, D. W., *J. Inst. Elec. Eng.* (*London*), **79**, 11 (1936); MALOFF, I. G., and EPSTEIN, D. W., *Electronics*, November, 1937, p. 31.

applying a high negative potential to the grid. Where such a switch is lacking, the beam may be turned off by means of the intensity control. The life of the tube is greatly extended if the operator forms the habits of turning off the beam when not making an observation and of using as low brilliance as possible.

15-20. Oscilloscope Time Bases.—Cathode-ray oscillograph tubes are provided with two sets of deflection plates or coils, mounted so that the deflection produced by one is in the horizontal or X direction, and by the other in the vertical or Y direction. Simultaneous application of direct voltage to both sets of plates or of current to both sets of coils produces a displacement that is the vector sum of those which result from the individual voltages or currents. Since the results obtained with magnetic and electrostatic deflection are essentially the same, much of the following discussion will mention only electrostatic deflection.

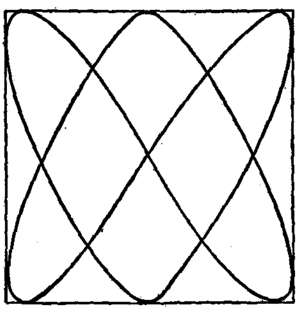


FIG. 15-48.—Lissajous figure for 3-to-2 frequency ratio, enclosed by rectangle.

When alternating voltage is applied to one pair of plates, the spot oscillates and, because of phosphorescence and persistence of vision, is seen on the screen as a line. Application of alternating voltage to both sets of plates causes the spot to trace a complicated path which does not in general close upon itself if the frequencies of the two voltages are different and which is, therefore, seen as a moving pattern. If the quotient of the two frequencies is rational, the path closes and the image is seen as a stationary pattern. A rational frequency ratio can be determined by enclosing the pattern by a rectangle the sides of which are parallel to the X and Y axes and tangent to the pattern. The ratio of the Y to the X frequency equals the number of points of tangency of the curve to a horizontal side of the rectangle divided by the number of points of tangency to a vertical side. Thus, in Fig. 15-48 the ratio of the Y frequency to the X frequency is 3:2. The form of the pattern is also affected by the harmonic content, amplitude, and phase relation of the two voltages. For sinusoidal voltages of the same frequency, that are of equal amplitude and in phase, the pattern is a straight line making a 45-degree angle with the horizontal. As the phase of one voltage is changed, the pattern becomes an ellipse which widens with increase of phase angle and becomes a circle when the phase angle reaches 90 degrees. Although these patterns, called *Lissajous figures*, may be analyzed quite readily if one component is sinusoidal, it is usually desirable to use an X voltage of such wave form that the unknown voltage is plotted on a time base.

In order that the image shall show the unknown voltage as a function of time, it is necessary that the spot shall periodically *sweep* across the screen horizontally with uniform velocity up to a certain point and then return instantaneously to its zero position. If the time taken for one timing *sweep* is equal to the period of the voltage applied to the Y plates, the pattern will consist of one cycle of the Y voltage. If the sweep frequency is equal to f_y/n, the image will show n waves of the Y voltage. The image will be stationary if n is a rational number. The required horizontal movement of the fluorescent spot can be produced by means of an X voltage that periodically increases uniformly with time and falls to zero instantaneously upon reaching a given value. The wave form of such a voltage is shown in Fig. 15-49. The earliest method of obtaining a saw-tooth voltage was by means of a rotating voltage divider. The usefulness of this method was limited by friction of sliding contacts and by contact resistance and other mechanical difficulties. The development of vacuum-tube oscillators for producing sweep voltages greatly increased the field of application of the cathode-ray oscillograph[1] and

FIG. 15-49.—Wave form of linear sweep voltage for cathode-ray oscillograph.

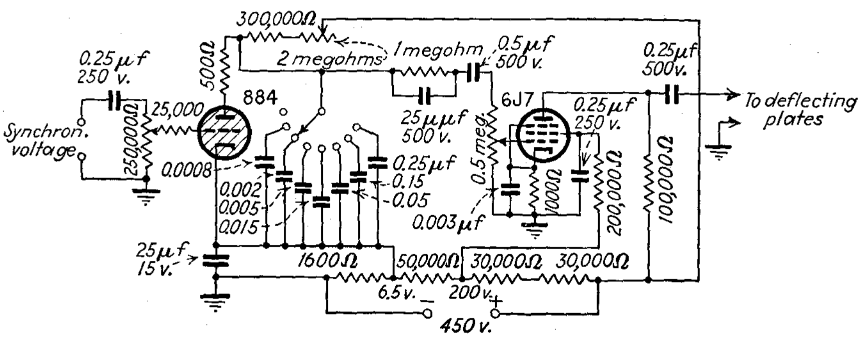


FIG. 15-50.—Sweep oscillator and amplifier.

made possible the synchronization of the sweep voltage to the unknown voltage (see Secs. 10-13, 10-15, and 12-33).

Methods of producing saw-tooth voltages were discussed in Secs. 10-13, 12-6, and 12-33. Although circuits using only high-vacuum tubes have a higher frequency limit of oscillation, the arc-tube circuits of Fig. 12-38 or 12-50 or modifications of these circuits are usually used as sweep oscillators in oscilloscopes.[2] Figure 15-50 shows a circuit that

[1] BEDELL, F., and REICH, H. J., *J. Am. Inst. Elec. Eng.*, **46**, 563 (1927).

[2] FRUEHAUF, G., *Arch. Elektrot.*, **21**, 471 (1929); SAMUEL, A. L., *Rev. Sci. Instruments*, **2**, 532 (1931) (with bibliography); FINCH, G. I., SUTTON, R. W., and TOOKE, A. E., *Proc. Phys. Soc. London*, **43**, 502 (1931); BREWER, F. T., *Electronics*, December,

incorporates an amplifier for varying the sweep amplitude.[1] Practical sweep oscillators do not furnish a voltage that completely satisfies the requirements for a perfect sweep voltage, since some time is required for the voltage to fall to zero at the end of the cycle. Because of this, a portion of the observed phenomenon occurs during the return sweep. The fraction of the cycle taken by the return sweep is ordinarily negligible at low frequency, but at very high frequencies the return time may be large enough so that a pattern of appreciable intensity is traced during the return sweep. The pattern traced during the return sweep may be eliminated by biasing the grid of the cathode-ray tube beyond cutoff during the return sweep. The required pulse of negative voltage may be obtained from a small resistor in series with the tube that discharges the sweep-oscillator condenser, such as the 500-ohm current-

FIG. 15-51.—Wave form of sweep voltage used when it is desired to observe only a portion of the cycle of voltage impressed upon the oscilloscope.

limiting resistance in series with the 884 tube in the circuit of Fig. 15-50. The voltage pulse may also be obtained by passing the saw-tooth voltage through the pulse-forming circuit of Fig. 10-10.

When it is desired to observe in detail a small portion of the entire cycle of a voltage wave, or a periodic voltage pulse of short duration repeated at a relatively low frequency, the luminous spot should sweep only during the portion of the cycle to be observed and remain at rest during the remainder of the cycle. This result can be attained by the use of a sweep voltage of the form of Fig. 15-51, which may be generated by the circuit of Fig. 15-52.[2] This circuit is a combination of the thyratron relaxation oscillator of Fig. 12-50, the pulse-forming circuit of Fig. 10-11, the trigger circuit of Fig. 10-8, and the circuit of Fig. 10-31.

1931, p. 222; SCOTT, H. H., *Gen. Radio Expt.*, **7**, No. 1 (1932); ULBRICHT, G., *Hochf. tech. u. Elektroakus.*, **39**, 130 (1932); FIELD, G. S., *Can. J. Research*, **7**, 131 (1932); FUCKS, W., and KROEMER, H., *Arch. Elektrotech.*, **27**, 125, 606 (1933); HALLER, C. E., *Rev. Sci. Instruments*, **4**, 385 (1933); ZWORYKIN, V. K., *Proc. I.R.E.*, **21**, 1655 (1933); *J. Inst. Elec. Eng.*, **71**, 82 (1932); NOTTINGHAM, W. B., *J. Franklin Inst.*, **211**, 751 (1931); RICHARDS, P. A., and MEIER, W. L., *Electronics*, April, 1934, p. 110; LENSIL, C. E., *Electronics*, May, 1935, p. 156; GOLDSMITH, T. T., JR., and RICHARDS, L. A., *Proc. I.R.E.*, **23**, 653 (1935); MCCARTHY, D., *Wireless World*, **37**, 367 (1935); *Electronics*, February, 1936, p. 42; LEEDS, L. M., *Proc. I.R.E.*, **24**, 872 (1936); POTTER, J. L., *Proc. I.R.E.*, **26**, 713 (1938). See also PARR, *op. cit.*, pp. 164–165.

[1] WALLER, L. C., *Radio Retailing*, January, 1937, p. 65. This article gives complete specifications for the construction of an oscilloscope using a type 913 cathode-ray tube.

[2] HAWORTH, L. J., *Rev. Sci. Instruments*, **12**, 478 (1941).

The functions of the various portions of the circuit are indicated by the wave forms of the voltages at various points.

A much longer time base can be obtained by the use of a circular or spiral sweep. A circular or eliptical sweep is obtained by applying to the two sets of deflection plates voltages that are 90 degrees out of phase, obtained from a series combination of resistance and capacitance or inductance, as shown in Fig. 15-53. The voltage under observation

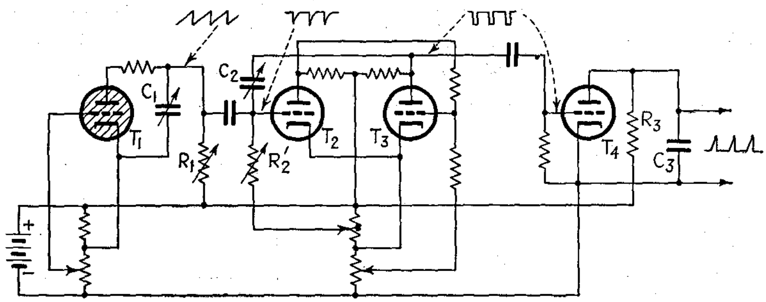


FIG. 15-52.—Expanded-sweep circuit for the observation of a small portion of the cycle of voltage impressed upon the oscilloscope.

is made to produce a radial displacement of the spot proportional to the instantaneous voltage by superimposing it upon the direct voltage of the second anode,[1] or by modulating the input voltage of the circuit of Fig. 15-53 in accordance with the voltage under observation, by means of a linear modulator. A spiral sweep is obtained by modulating the input voltage of the circuit of Fig. 15-53 in accordance with a saw-tooth voltage.

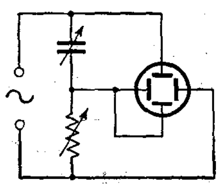


FIG. 15-53.—Circuit for producing a circular sweep.

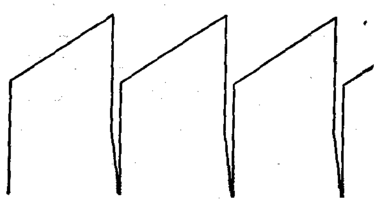


FIG. 15-54.—Form of exciting voltage required to produce linear electromagnetic sweep.

The voltage under observation is superimposed upon the saw-tooth voltage in modulating the input to the phase-splitting circuit.

In order to obtain a linear sweep by electromagnetic means, the current through the deflecting coils must be of saw-tooth form. The linear rise in current through the coils requires a linear rise in voltage across the coil resistance and a constant voltage across the coil inductance.

[1] Variation of second anode voltage varies the deflection sensitivity, and hence the radius of the circular sweep. See N. V. KIPPING, *Wireless World*, **13,** 705 (1924).

The periodic wave of voltage across the coil must therefore be the sum of a saw-tooth wave and a rectangular wave. The desired result may be obtained by impressing a wave of the form of Fig. 15-54 upon the input of the circuit of Fig. 15-55. The function of the diode, condenser, and resistance in the circuit of Fig. 15-55 is to damp out transient oscillations set up in the transformer and deflection coils by the discontinuities of current. Voltage of the form of Fig. 15-54 may be obtained from a saw-tooth-wave oscillator by the addition of resistance in series with the condenser, as shown in Fig. 15-56. The constant condenser charging current produces a constant voltage across R' and a linearly rising voltage across the condenser. R' is varied until the current wave has the correct wave form.

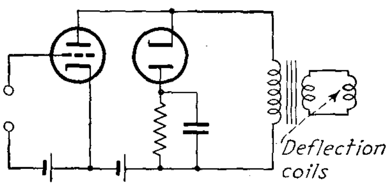


FIG. 15-55.—Amplifier for producing linear electromagnetic sweep.

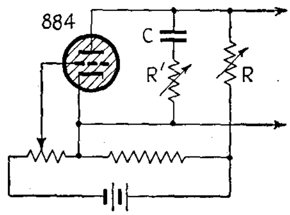


FIG. 15-56.—Oscillator for the generation of voltages of the form of Fig. 15-54.

15-21. Oscilloscope Amplifiers.—Oscilloscope amplifiers serve the dual purpose of providing sufficient voltage to give the desired deflection, and of making possible the variation of deflection voltages without drawing appreciable current from the source under observation or from the sweep oscillator. Sweep amplifiers must be capable of amplifying frequencies ranging from the fundamental frequency up to approximately the tenth harmonic. In order to cover the sweep frequency range of 50 to 50,000 cps, therefore, the sweep amplifier must have negligible frequency and phase distortion in the range from 50 to 500,000 cps. A similar frequency range is desirable in the Y-deflection amplifier in the study of wave form, particularly when the wave contains near-discontinuities. Compensated amplifiers embodying cathode-follower stages, similar to the amplifier of Fig. 6-24, are used in the best oscilloscopes.

Figure 15-57 shows the usual manner of coupling the deflection plates to the output of a single-sided amplifier. The voltage divider P changes the direct voltage between the plates and thus displaces the image on the screen. Since the second anode and one deflection plate of each pair may be connected within the tube, the use of this circuit makes possible a simplified tube structure. It has the disadvantage, however, that the variation of the potential of only one plate relative to the second anode and the cathode produces a corresponding variation of the effective accelerating voltage and thus destroys the linearity between deflection voltage and deflection.[1] This difficulty may be avoided by using a push-pull amplifier to excite the deflection plates. The potential of one plate relative to the second anode then falls by the same amount that the other rises and the average potential of the deflection plates relative to the cathode remains constant. The cathode-phase-inverter circuit of Fig. 5-16 may be conveniently used for this purpose. Figure 15-58 shows a form of this circuit that makes it possible to displace the image on the screen. Increasing the magnitude of the bias of T_1 above that of T_2 by means of the potentiometer P decreases the plate current of T_1 and makes the voltage drop across R_{b1} less than

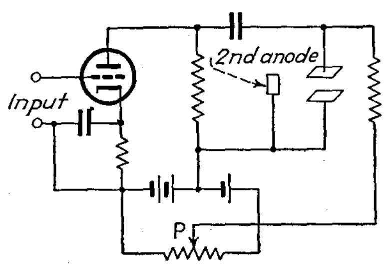


FIG. 15-57.—Single-sided oscilloscope amplifier with a control for displacing the pattern.

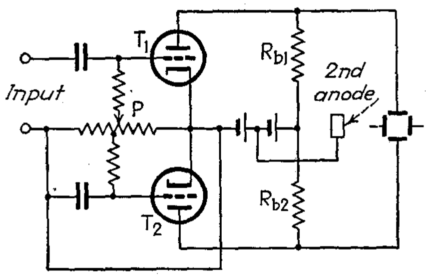


FIG. 15-58.—Push-pull oscilloscope amplifier.

that across R_{b2}. A direct difference of potential is thus established between the deflecting plates, causing a steady deflection of the image.

The complete circuit diagram of a typical oscilloscope is shown in Fig. 15-59 (see Prob. 15-1).

15-22. Electronic Switches for the Observation of Two Waves.—Most cathode-ray tubes that are now on the market contain only one set of elements and can therefore be used directly for the observation of only one voltage as a function of time. Two or more waves can be observed simultaneously by the use of a commutator,[2] but this method is not very convenient. A number of electronic circuits have been developed that make possible the observation of two waves.[3] In these circuits the two voltages to be observed are applied to two amplifiers, the outputs of which are impressed upon the Y deflecting plates. Some form

[1] DUMONT, A. B., *Electronics*, January, 1935, p. 16.

[2] BEDELL, and REICH, *loc. cit.*

[3] SEWIG, R., *Z. f. tech. Physik.*, **14**, 152 (1933); BROWN, C. B., *Electronics*, **6**, 170 (June, 1933); DAVIDSON, I. B., *J. Sci. Instruments*, **11**, 359 (1934); GARCEAU, L., *Rev. Sci. Instruments*, **6**, 171 (1935); DUMONT, A., *Electronics*, March, 1935, p. 101; WOODRUFF, L. F., *Elec. Eng.*, **54**, 1045 (1935); GEORGE, R. H., HEIM, H. J., and ROYS, C. S., *Elec. Eng.*, **54**, 1095 (1935); HUGHES, H. K., *Rev. Sci. Instruments*, **7**, 89 (1936); SHUMARD, C. C., *Elec. Eng.*, **57**, 209 (1938); REICH, H. J., *Rev. Sci. Instruments*, **12**, 191 (1941).

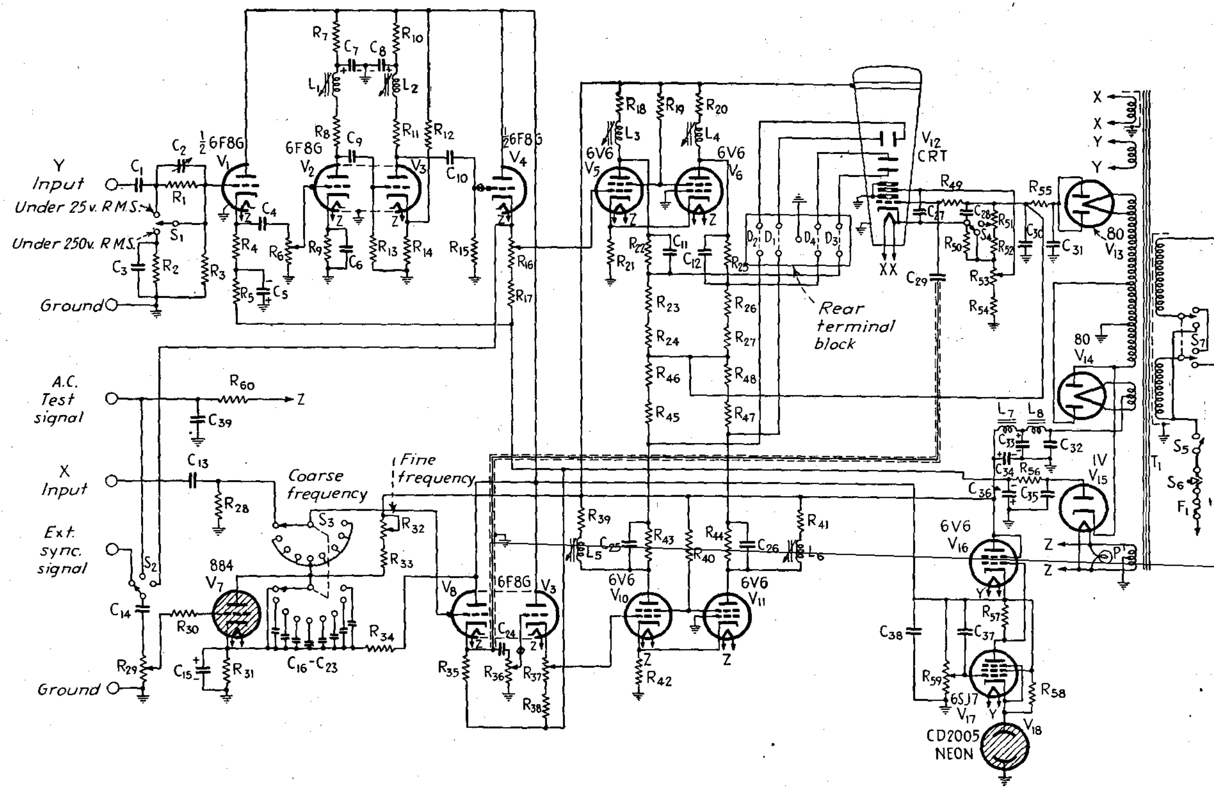


FIG. 15-59.—Complete circuit of a typical oscilloscope. (*Courtesy of Allen B. DuMont Laboratories, Inc.*)

of trigger switching circuit, such as those of Fig. 10-8, 10-9, or 12-34, synchronized to the sweep voltage, is used to control the control-grid, screen, or suppressor voltages of the amplifiers in such a manner that the amplifiers amplify during alternate sweeps. The images of the two waves are formed on the fluorescent screen during alternate sweeps. Because of phosphorescence and persistence of vision, they appear to be seen simultaneously.

A typical switching circuit is shown in Fig. 15-60.[1] This circuit uses pentode amplifier tubes controlled by a parallel thyratron switching circuit of the form of Fig. 12-34*a*. The screens of the amplifier tubes T_3 and T_4 are connected to the anodes of the thyratrons T_1 and T_2,

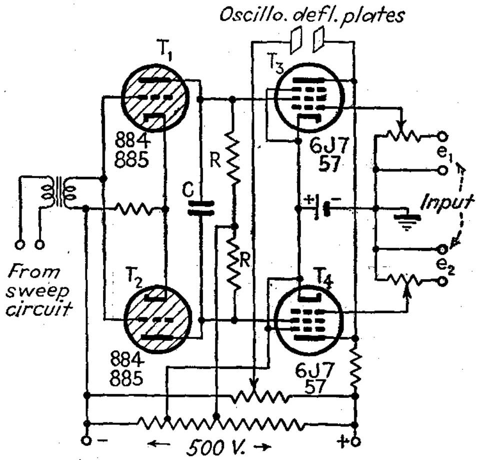


FIG. 15-60.—Electronic switch for the simultaneous observation of two waves with a single-element cathode-ray tube.

respectively. The amplifier cathodes are positive relative to the thyratron cathodes by a voltage that exceeds the thyratron tube drop. Hence, when either thyratron fires and the voltage of its anode relative to its cathode falls to a value equal to the tube drop, the screen voltage of the corresponding amplifier tube is made slightly negative relative to its cathode, so that its plate current is reduced to zero, and the tube ceases to amplify.

15-23. Electronic Transient Visualizers.—The periodic contactor circuits mentioned on pages 457 and 491 are of value in the study of transient voltages and currents by means of the cathode-ray oscillograph.[2] The sweep oscillator voltage, in addition to providing time deflection of the fluorescent spot, varies the grid voltage of an amplifier tube, the plate circuit of which contains a relay. The relay is used to open and

[1] GEORGE, HEIM, and ROYS, *ibid.*

[2] REICH, H. J., *Rev. Sci. Instruments*, **5**, 7 (1934).

close the circuit in which the transient is produced, the adjustment being such that the transient is initiated at the beginning of the timing sweep. Since the relay operates at the same instant in each sweep, the transient is observed on the screen as a stationary figure. Limitations imposed by relay inertia and chattering are overcome by the use of an 884 thyratron tube in place of the mechanical relay, as shown in Fig. 15-61.[1] Extinction of the 884 switching tube T_4 is accomplished by virtue of the ability of the 884 grid to interrupt anode current of 50 ma or less when the grid is made sufficiently negative. The saw-tooth voltage of the sweep oscillator is applied to the grid of T_4 with proper polarity so that the sudden change in voltage when the sweep oscillator tube T_2 fires causes the negative bias of T_4 to increase enough to cut

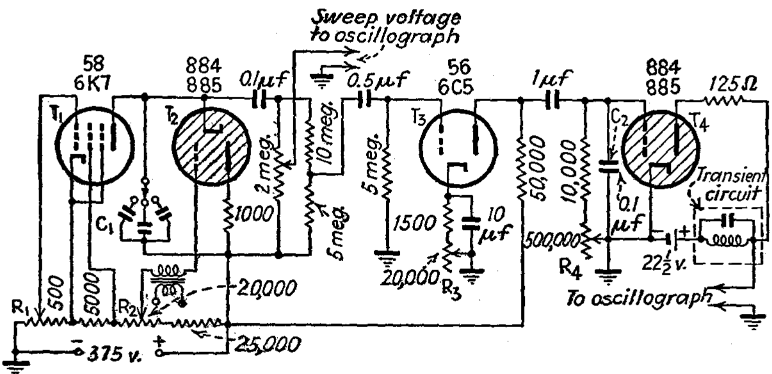


FIG. 15-61.—Electronic transient visualizer.

off the anode current. Since this change in voltage is very rapid, the extinction time of T_4 may be made as small as 10 μsec. As the condenser C_1 charges, the negative grid voltage of T_4 decreases and at some time in the cycle becomes so small that T_4 again fires. If the plate circuit of T_4 contains only resistance, the wave of anode current, and the voltage across this resistance, are observed to be square topped, as they should be if firing and extinction of T_4 are rapid. The condenser C_2 of Fig. 15-61 serves to delay the extinction of T_4, so that a transient initiated by the extinction of T_4 will not start during the return sweep of the fluorescent spot. The frequency of repetition is controlled by C_1, R_1, and R_2. R_3 and R_4 control the portion of the cycle during which T_4 conducts. Figure 15-62 shows how this device may be used in the oscillographic study of transient oscillations in coupled oscillatory circuits and in lines.

Another type of visualizer, in which the transient is initiated by a voltage impulse, is based upon the relaxation circuit of Fig. 12-43.[2] The

[1] BENNETT, J. A., master's thesis, Cornell University, June, 1935; REICH, H. J., *Elec. Eng.*, **55**, 1314 (1936), *Trans. Am. Inst. Elec. Eng.*, **56**, 873 (1937).

[2] REICH, *ibid.*

voltage that causes the transient is preferably taken from a resistance in series with C_1, either directly or through an amplifier, and the circuit is tripped by means of voltage from the sweep oscillator. A similar circuit of somewhat reduced flexibility is obtained by taking the transient exciting voltage pulse from a resistance in series with the sweep oscillator condenser, through a suitable amplifier.[1]

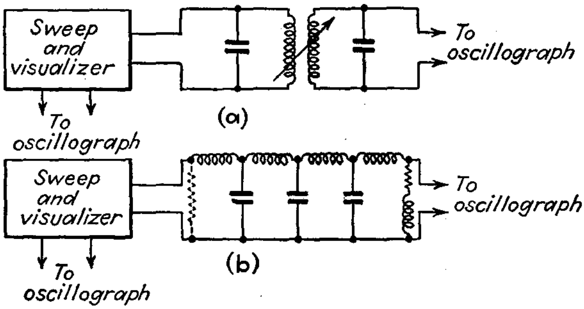


FIG. 15-62.—Circuits for using transient visualizer in the study of transient oscillations in (a) coupled circuits and (b) lines.

A third type of transient visualizer is illustrated in Fig. 15-63.[2] The action of this circuit is as follows: During one-half of the cycle of alternating supply voltage, the condenser C_1 is charged from the secondary of the transformer through the rectifier T_2. At the positive crest of the following half cycle the thyratron T_1 fires, and C_4 begins to charge,

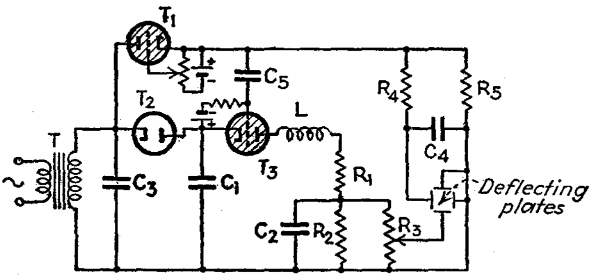


FIG. 15-63.—Transient visualizer that initiates the transient by the discharge of a condenser through a thyratron.

causing a horizontal deflection of the fluorescent spot. Firing of T_1 also applies a positive impulse to the grid of T_3, causing it to fire. C_1 discharges through T_3, producing a voltage pulse across R_3, the shape of which depends upon the parameters C_1, C_2, L, R_1, and R_2. This voltage, or a portion of it, may be applied to the circuit in which it is

[1] Unpublished paper by P. W. Ryburn, Univ. of Illinois, 1938.

[2] ROHATS, N., *Gen. Elec. Rev.*, **39**, 146 (1936); *Trans. Am. Inst. Elec. Eng.*, **56**, 873 (1937).

desired to initiate a transient. The transient is thus repeated in synchronism with the sweep at the frequency of the alternating supply voltage.

15-24. Oscillographic Comparison of Frequencies.—The cathode-ray oscillograph can be used to advantage in the calibration of oscillators by comparison with known frequencies. The simplest method is by the use of Lissajous figures. The output of the oscillator that is to be calibrated is applied to one pair of deflecting plates and the standard frequency to the other pair. Either frequency is adjusted to give a stationary pattern. The frequency ratio is equal to the ratio of the number of horizontal and vertical points of tangency of a rectangle that encloses the pattern.

A second method of frequency comparison was developed by N. V. Kipping.[1] In Kipping's circuit, shown in Fig. 15-64, the lower frequency is applied to a series combination of resistance and capacitance. The voltage across the resistance is applied to one set of deflecting plates, and the voltage across the condenser to the other. Because these voltages are 90 degrees out of phase, a circular pattern is obtained when the resistance is adjusted so as to make their amplitudes equal. The higher frequency is introduced into the anode circuit. Variation of anode voltage changes the electron velocity and, therefore, the deflection sensitivity. This causes the diameter of the circular pattern to change, the effect being equivalent to a radial displacement of the spot relative to the circle. If the ratio of the two frequencies is rational, a stationary pattern is obtained. The frequency ratio is equal to the number of whole waves superimposed on the circle, divided by the number of revolutions that the spot makes in completing the pattern. With modern tubes, which incorporate a control grid, the higher frequency may also be introduced into the grid circuit to produce a variation of intensity along the circle. The latter method is not satisfactory for other than integral frequency ratios, however, because of the difficulty of determining how many times the spot moves about the circle in completing the pattern.

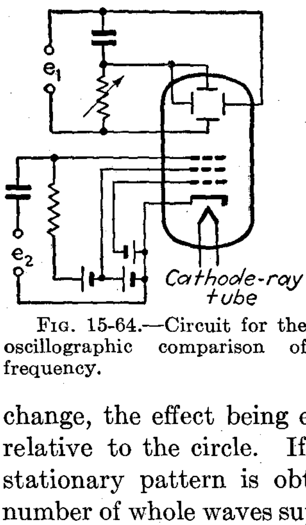


FIG. 15-64.—Circuit for the oscillographic comparison of frequency.

A third method[2] makes use of the circuit of Fig. 15-65. The two voltages e_1 and e_2, of frequency f_1 and f_2, are first applied individually, and the resistances and capacitances are adjusted so that each voltage

[1] KIPPING, N. V., *Wireless World*, **13**, 705 (1924).

[2] RANGACHARI, T. S., *Wireless Eng.*, **5**, 264 (1928), **6**, 184 (1929); KURRELMEYER, BERNHARD, *Rev. Sci. Instruments*, **7**, 200 (1936), **8**, 348 (1937).

produces a circular pattern. When the two voltages are applied simultaneously, the pattern is some form of roulette, of which the cycloids are examples.[1] If the amplitude of the higher frequency is somewhat smaller than that of the lower frequency, the pattern will in general have one or more cusps or loops. A stationary pattern is obtained if the ratio of the higher to the lower frequency f_2/f_1, reduced to its simplest form, is N_2/N_1, in which N_2 and N_1 are integers. For the circuit of Fig. 15-65, there are $N_2 + N_1$ loops or cusps, pointed outward. In generating the pattern, the fluorescent spot moves from one cusp or loop to the N_1th next one. Thus N_1 may be determined by adding one to the number of cusps skipped by the fluorescent spot. If N is the total number of cusps, $N = N_2 + N_1$. If the ratio of the frequencies cannot be expressed as the ratio of two integers, the pattern will have $N_2 + N_1$ cusps, where N_2/N_1 is the rational number that most nearly approximates the frequency ratio f_2/f_1, but will rotate. If f_2/f_1 is less than N_2/N_1, the pattern will rotate in the same direction as the spot moves when the lower frequency f_1 is applied alone, which may be termed the positive direction. The frequency ratio is

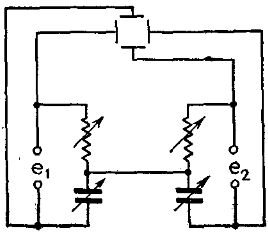


FIG. 15-65.—Circuit for the oscillographic comparison of frequency.

$$\frac{f_2}{f_1} = \frac{(1 - f_3/f_1)N_2}{N_1} \tag{15-50}$$

in which f_3 is the number of revolutions of the pattern per second. If f_2/f_1 is greater than N_2/N_1, on the other hand, the direction of rotation is negative, and the numerical value of f_3 should be taken as negative in Eq. (15-50). The positive direction of rotation can be readily determined by increasing f_1 or decreasing f_2 slightly. The rotation is positive if the speed of rotation increases.

If the position of one condenser and its associated resistor is interchanged in the circuit of Fig. 15-65, there are $N_2 - N_1$ cusps or loops, pointed inward. For frequency ratios for which N_1 exceeds $N_2 - N_1$, interpretation of the patterns is more difficult with this modified circuit since the fluorescent spot makes more than one complete trip around the center between two successively generated cusps, and the patterns are intricate. Figure 15-66 shows typical patterns obtained for several frequency ratios. All patterns are for the circuit of Fig. 15-65 with the exception of the lower right pattern, which was obtained with the transposed circuit.

[1] REYNOLDS, J. B., and WEIDA, F. M., "Analytic Geometry and Elements of Calculus," p. 246, Prentice-Hall, Inc., New York, 1930.

The circuit of Fig. 15-65 cannot be used with an oscillograph in which there is a common connection between one horizontal and one vertical deflecting plate. With this type of oscillograph the circuit of

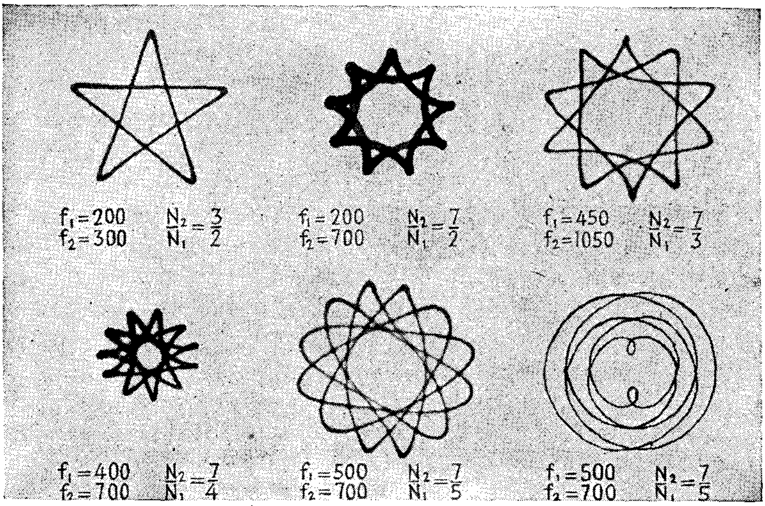


FIG. 15-66.—Oscillograms obtained with the circuit of Fig. 15-65 for various frequency ratios.

Fig. 15-67 may be used.[1] The impedance of the isolating transformer T should be high in comparison with the resistance R_1. For frequencies in the audible range an ordinary interstage coupling transformer and a

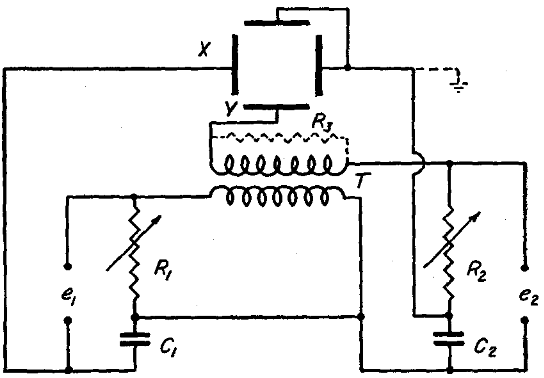


FIG. 15-67.—Modified form of the circuit of Fig. 15-65 that makes possible the use of a tube with a common connection between two deflecting plates.

5000-ohm variable resistance may be used. At frequencies above 5000 cps, the distributed capacitance of the transformer tends to prevent the formation of a circular pattern for the voltage e_1. For this reason e_1

[1] RANGACHARI, *loc. cit.;* REICH, H. J., *Rev. Sci. Instruments,* **8**, 348 (1937).

should be the lower frequency. Resistance or capacitance between the Y plates or leads produces a phase shift which makes it difficult to obtain a circular pattern for the voltage e_2 at very high frequencies. No difficulty is encountered, however, even above 10,000 cps, if the Y plates are not shunted by a leak and if the connection from the ungrounded Y plate to the transformer is short. A high resistance R_3 across the secondary is sometimes of help at high frequencies. If the primary and secondary of Fig. 15-67 are wound in the same direction, the cusps or loops of the pattern will point inward. The pattern may be transformed into one in which the cusps point outward by transposing the primary or secondary connections or by interchanging the position of either resistor and its associated condenser. Provision should be made for the variation of the amplitudes of e_1 and e_2 both in the circuit of Fig. 15-65 and in that of Fig. 15-67.

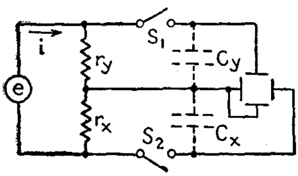


FIG. 15-68.—Circuit in which the capacitance of the deflecting plates may affect the form of the oscillogram.

15-25. Effect of Capacitance of Deflecting Plates.—Special precautions must be taken in the use of the cathode-ray oscillograph in the study of small voltages and currents. The capacitances of deflecting plates and their leads, although small, may under certain circumstances produce surprising results. In Fig. 15-68, r_x and r_y are two resistances through which a common alternating current flows. If r_x and r_y are high, it is found that when the voltages across them are applied to the cathode-ray oscillograph by closing the switches S_1 and S_2, the image may be an ellipse, instead of the straight line that would be expected with voltages that are in phase. If only one resistor were known to be

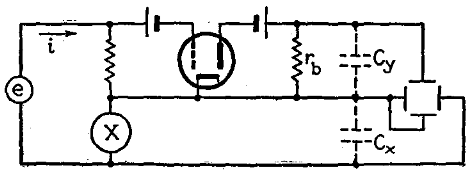


FIG. 15-69.—Circuit for the oscillographic determination of current-voltage characteristics of the element X. Errors may result from deflecting-plate capacitances C_x and C_y.

nonreactive, the incautious observer might conclude that the phase difference of the voltages results from reactance in the other resistor. Actually, the difference in phase results from inequality of the electrode capacitances C_x and C_y, or of the resistances r_x and r_y.

Although the simple circuit of Fig. 15-68 is not likely to be used, modified forms of it may be. Thus, the circuit of Fig. 15-69 may, for example, be used to determine the current-voltage characteristic of the element X. The load resistor r_b may be of the order of $\frac{1}{2}$ megohm, so that the capacitance of the deflecting plates and amplifier tube may produce appreciable shift in the phase of the voltage across r_b. Similarly, if the resistance of X is high, C_x may appreciably affect the phase of the voltage across it.

15-26. Electron-ray Tube.—Another type of cathode-ray tube is the *electron-ray tube* or *magic eye* which is used as a tuning indicator in radio

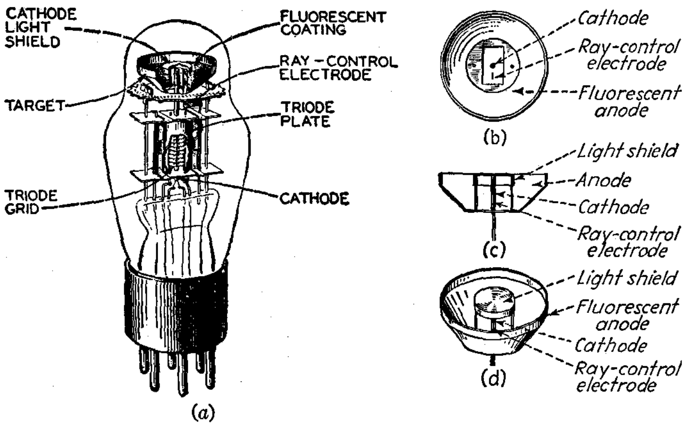


FIG. 15-70.—Structure of the type 6AB5 electron-ray tube.

receivers.[1] This tube, the construction of which is shown in Fig. 15-70, contains two sets of elements, one of which is a triode amplifier and the other a cathode-ray indicator. The latter consists of a cathode, a fluorescent anode (target), and a control electrode which controls the portion of the fluorescent anode upon which the electrons strike. The ray-control electrode is connected to the plate of the amplifier section, which is connected to the supply voltage through a high resistance, as shown in Fig. 15-71. Variation of the amplifier grid voltage changes the voltage of the ray-control electrode, and thus the portion of the target that fluoresces. The fluorescence is observed as an annular sector of varying angular width. The tube may be calibrated for use as a voltmeter where rough measurements suffice. When the electron-ray tube is used as a tuning indicator, the a-v-c biasing voltage, which increases with amplifier output, is used as the control voltage.

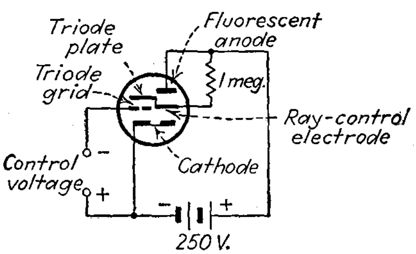


FIG. 15-71.—Electron-ray-tube circuit.

[1] WALLER, L. C., and RICHARDS, P. A., *Radio Retailing*, December, 1935, p. 47; THOMPSON, H. C., *Proc. I.R.E.*, **24**, 1276 (1936); WALLER, L. C., *QST*, October, 1936, p. 35, November, 1936, p. 23; WALLER, L. C., *RCA Rev.*, **1**, 111 (1937).

15-27. Determination of Static Tube Characteristics.—It would seem at first thought that the determination of static tube characteristics by means of direct voltages and d-c meters is such a simple procedure that it requires no discussion. Unless certain precautions are observed, however, erratic readings of plate current may be observed. These can usually be traced to high-frequency oscillations in the battery leads, the necessary inductance and capacitance being furnished by the meters, leads, and electrodes. A type 56 triode, for instance, will oscillate readily with a plate voltage of only 45 volts if the grid and plate leads happen to be of correct lengths to bring the grid and plate circuits into resonance. The difficulty may be prevented by the use of by-pass condensers from grid and plate to cathode or a resistance in series with the grid, if grid current does not flow.

15-28. Oscillographic Determination of Tube Characteristics.—The ordinary static method of obtaining static triode characteristics cannot be satisfactorily applied in certain ranges of current and voltage because of excessive plate and grid dissipation. The flow of high grid and plate currents for a sufficient time to take meter readings may result in temporary changes in characteristics as the result of change of cathode temperature or in permanent damage to the tube as the result of gas emission from the plate. A method for overcoming this difficulty by the use of oscillographic recording was devised by Kalin.[1] The oscillograph used by Kalin was electromagnetic and incorporated a rotating drum for photographic recording. The oscillograph element was placed in the plate circuit, so that the vertical deflection of the light spot was proportional to the plate current. The grid voltage was made to vary linearly with time by means of a rotary voltage divider. Since the recording drum rotated at constant speed, the grid voltage was proportional to the horizontal displacement of the spot on the film. The oscillogram consequently represented a plot of plate current against grid voltage. By this method the average plate dissipation was kept sufficiently low so that the complete family of transfer characteristics could be obtained. The substitution of a linear sweep oscillator in place of the rotating voltage divider improves Kalin's method. The sweep oscillator may be synchronized to the drum motor.

In the positive grid region the average plate and grid dissipations may be too great, even when the grid voltage is periodically varied as a linear function of time. Kozanowski and Mouromtseff[2] reduced the average dissipation to a very low value by using for the grid excitation a single pulse of voltage from the discharge of a large condenser through

[1] KALIN, ALBERT, *Univ. Wash. Eng. Expt. Sta. Bull.* 30, (1924); SCHNEIDER, W. A., *Proc. I.R.E.*, **16**, 674 (1928).

[2] KOZANOWSKI, H. N., and MOUROMTSEFF, I. E., *Proc. I.R.E.*, **21**, 1082 (1933).

a resistance. Although the instantaneous values of dissipation may be very high, the currents flow for such short intervals that the electrode temperatures do not rise so high as to damage the tube. Peak input powers of twenty to thirty times the nominal rating can be recorded by this method. The grid and plate characteristics are derived from oscillograms of grid and plate voltages and current, recorded simultaneously by a four-element moving-mirror oscillograph. The circuit is shown in Fig. 15-72. All resistances used in this circuit are non-inductive. The oscillograph elements must, of course, be accurately calibrated. The condenser is charged from the voltage supply preparatory to the taking of an oscillogram, and the discharge of the condenser is initiated automatically by the closing of the switch in synchronism with the opening of the oscillograph shutter.

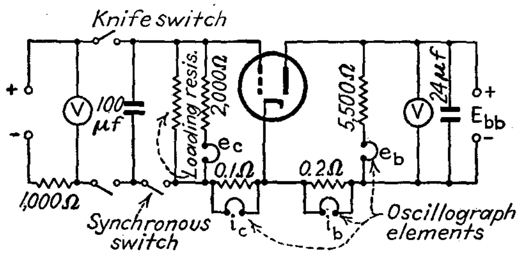


FIG. 15-72.—Circuit for the determination of tube characteristics by means of an electromagnetic oscillograph.

A cathode-ray oscillograph may also be used in the determination of static transfer characteristics, but its use involves difficulties. If electrostatic deflection is used, the plate current must be passed through a resistance in order to provide deflecting voltage. A resistance that is high enough to provide adequate voltage may appreciably alter the characteristic. If an amplifier is used in conjunction with a low resistance, the amplifier must have negligible nonlinear, frequency, and phase distortion over a wide frequency range in order to ensure distortionless amplification. The use of an inverse feedback amplifier is recommended for this purpose. Because of the impedance of the deflecting coils, difficulties are also encountered when the plate current is used to produce magnetic deflection.

Chaffee has described a cathode-ray circuit for the determination of static characteristics in which the grid of the tube under observation is excited by a voltage pulse produced by the discharge of a thyratron.[1]

15-29. Dynamic Measurement of Tube Factors.—Bridge circuits for the dynamic measurement of tube factors were first described by

[1] CHAFFEE, E. L., *Electronics*, July, 1937, p. 30; see also H. F. MAYER, *Electronics*, April, 1938, p. 14.

J. M. Miller.[1] Other circuits have since been developed. Before proceeding to a discussion of tube-factor bridges it is advisable to mention general limitations of these bridges and precautions that should be observed in their use. The equations that must be employed in the applications of these bridges are derived by means of equivalent tube circuits, which are strictly applicable only at small signal amplitudes. The accuracy of the experimentally determined values therefore increases as the signal amplitude is reduced. For this reason it is advisable to use small voltages in exciting tube-factor bridges and to employ an amplifier between the bridge and the telephone receivers if necessary. To prevent direct voltage drops that reduce operating voltages it is important to provide direct-plate-current paths of low resistance by shunting the phones with low-resistance chokes. This is particularly necessary in the circuits of Figs. 15-73, 15-74, 15-76, 15-78, 15-79, and 15-80. Transformers having low primary resistance may, of course, be used in place of chokes. In the circuits of Figs. 15-73 to 15-76, and 15-80, direct plate current also flows through the secondary of the exciting voltage input transformer, which should consequently have low d-c resistance. The circuits should be set up so that stray capacitances between leads and batteries are as small as possible. For the sake of simplicity, the tube is shown as a triode in the following circuits, multigrid tubes merely requiring the addition of other supply voltages.

15-30. Amplification-factor Bridge.—The basic bridge circuit generally used for the measurement of amplification factor is that of Fig. 15-73.[1] By constructing the equivalent plate circuit and assuming that $V_p = -(r_2/r_1)V_g$ and that grid current does not flow, the student may readily show that the alternating plate current is zero when

$$\mu = \frac{r_2}{r_1} \tag{15-51}$$

Actually the effect of tube capacitances prevents V_p and V_g from being exactly 180 degrees out of phase at high frequencies. They can be made so by the use of the balancing condenser C. The voltages V_p and V_g are opposite in phase when the ratio of the reactances of the condensers shunting r_1 and r_2 is equal to r_1/r_2, or

$$\frac{C + C_{gk}}{C_{pk}} = \frac{r_2}{r_1} \tag{15-52}$$

The voltages can also be made 180 degrees out of phase by means of a small mutual inductance M, as in Fig. 15-74. As measurements are

[1] MILLER, J. M., *Proc. I.R.E.*, **7**, 112 (1919).

usually made at 1000 cps or lower, the balancing condenser or mutual inductance is not ordinarily necessary.

r_1 should preferably be small enough so that the direct voltage drop in r_1 will not greatly alter the operating grid voltage. For accurate measurements this voltage drop should be taken into account in adjusting the tube voltages.

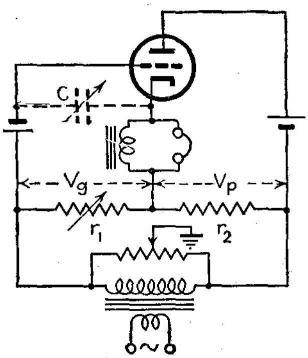


FIG. 15-73.—Bridge for the measurement of amplification factor.

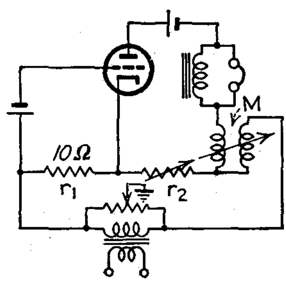


FIG. 15-74.—Bridge for the measurement of amplification factor.

15-31. Plate-resistance Bridges.—The plate resistance of triodes or other low–plate-resistance tubes may be readily measured by means of the simple four-arm bridge of Fig. 15-75.[1] When the bridge is balanced,

$$r_p = r_3 \frac{r_1}{r_2} \tag{15-53}$$

and

$$C = (C_{pk} + C_{gp}) \frac{r_1}{r_2} \tag{15-54}$$

The condenser C may be omitted if measurements are made at 1000 cps or lower. To avoid excessive direct-voltage drop in r_1, r_1 should preferably be small (10 or 100 ohms). The voltage drop should be taken into account in adjusting the operating plate voltage.

A second bridge for the measurement of plate resistance is formed by the addition of resistances r_3 and r_4 to the amplification-factor bridge of Fig. 15-74, as shown in Fig. 15-76.[2] The bridge is first balanced by means of r_2 and M, with S_1 closed and S_2 open. Then S_1 is opened and S_2 closed, and the bridge is again balanced. The student may show that the plate resistance is then given by the relation

$$r_p = 100r_4 \tag{15-55}$$

[1] BALLANTINE, S., *Proc. I.R.E.*, **7**, 134 (1919).
[2] MILLER, *loc. cit.*

It is difficult to measure the high plate resistance of tetrodes and pentodes accurately with the circuits of Figs. 15-75 and 15-76. High values of plate resistance can be conveniently measured with the circuit of Fig. 15-77.[1] If r_1 is a calibrated resistance, r_p is found by adjusting r_1 so that the voltage across r_2 is the same for the two positions of the

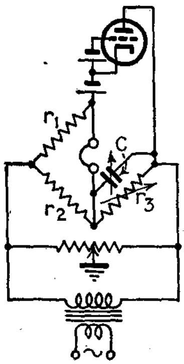


FIG. 15-75.—Bridge for the measurement of plate resistance.

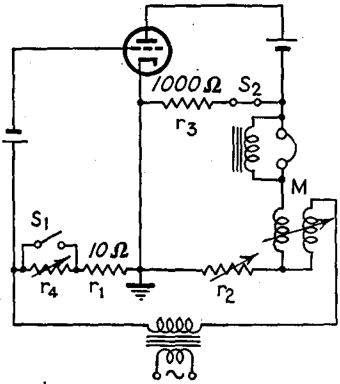


FIG. 15-76.—Bridge for the measurement of plate resistance.

switch S. If the applied voltage V is kept constant, the vacuum-tube voltmeter may be calibrated to read resistance directly, r_1 being used for calibration. A third method of measurement is to observe the voltage V necessary to give a definite deflection of the vacuum-tube voltmeter. For high accuracy r_2 should be small in comparison with r_p. This necessitates the use of a vacuum-tube voltmeter of high sensitivity.

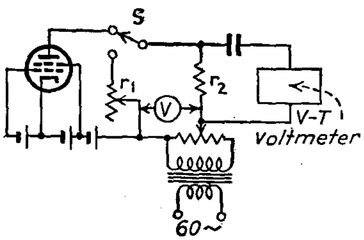


FIG. 15-77.—Bridge for the measurement of high plate resistances.

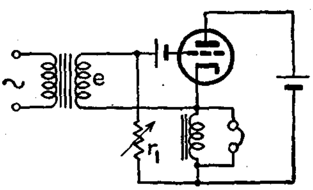


FIG. 15-78.—Bridge for the measurement of transconductance of tubes with low plate resistance.

The direct voltage drop in r_2 must be taken into consideration in adjusting the plate voltage.

15-32. Transconductance Bridges.—The transconductance of tubes with low plate resistance can be readily measured by the simple circuit of Fig. 15-78, which is a modified form of a circuit first suggested by

[1] Standards on Electronics, p. 27, Institute of Radio Eng., New York, 1938.

Appleton.[1] Examination of this bridge shows that the alternating current through the phones is made up of two components, one of which flows through r_1 and the source e, and the other through the plate of the tube, and that these two currents are opposite in phase if the electrode capacitances are small. When the circuit is balanced, these two currents are equal and, furthermore, there is no voltage drop through the phones. Therefore,

$$\frac{e}{r_1} = \frac{\mu e}{r_p} \tag{15-56}$$

or

$$g_m = \frac{1}{r_1} \tag{15-57}$$

For greater accuracy the unbalancing caused by the interelectrode

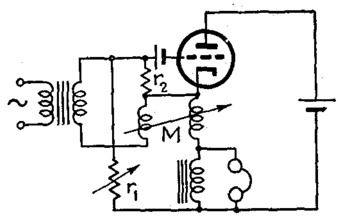


FIG. 15-79.—Transconductance bridge of Fig. 15-78, modified to incorporate reactance balance.

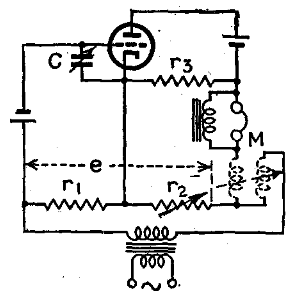


FIG. 15-80.—Bridge for the measurement of transconductance of tubes with high plate resistance. Either C or M is used for reactive balance.

capacitance at high frequency can be corrected by the addition of a mutual inductance, as in Fig. 15-79.

For tubes with high plate resistance the circuit of Fig. 15-80 is more satisfactory.[2] In this circuit the alternating voltage across r_2 is balanced against that across r_3. The bridge is balanced reactively by the condenser C, which may be replaced by a mutual inductance M, shown dotted in Fig. 15-80. For the circuit in which C is used, the voltages across r_3 and r_2 are opposite in phase and equal in magnitude when the bridge is balanced. Since no alternating current flows through the phones and the reactance of the interelectrode capacitances is high in comparison with r_3, the load impedance is practically equal to r_3. If grid current does not flow,

[1] APPLETON, E. V., *Wireless World*, **6**, 458 (1918).

[2] BALLANTINE, *loc. cit.*

$$\frac{er_2}{r_1 + r_2} = \frac{\mu e_g r_3}{r_p + r_3}$$
$$= \frac{\mu e r_3}{r_p + r_3} \cdot \frac{r_1}{r_1 + r_2} \tag{15-58}$$

or

$$g_m = \frac{r_2}{r_1 r_3} \cdot \frac{r_p + r_3}{r_p} \tag{15-59}$$

The plate resistance of tetrodes and pentodes is so high that the ratio of r_p to r_3 may be made large enough so that Eq. (15-59) reduces to

$$g_m = \frac{r_2}{r_1 r_3} \tag{15-60}$$

Equations (15-59) and (15-60) may also be shown to hold when M is used in place of C.

15-33. Bridge for the Measurement of μ, r_p, and g_m.—Inspection of Figs. 15-74, 15-76, and 15-80 shows that these three bridges may be very readily combined into a single instrument for the measurement of μ, r_p, and g_m.

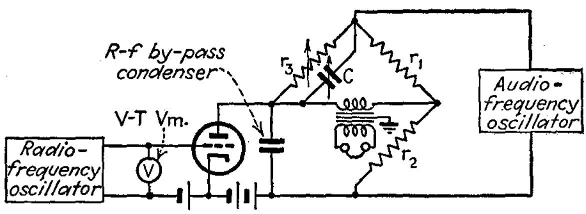


FIG. 15-81.—Bridge for the measurement of detection plate resistance.

Detection-plate-resistance Bridge.—Figure 15-81 shows a bridge for the measurement of detection plate resistance r_p'.[1] The action of this bridge is similar to that of the plate-resistance bridge of Fig. 15-75. When balance is obtained,

$$r_p' = \frac{r_2 r_3}{r_1} \tag{15-61}$$

In the measurement of detection plate resistance of diodes, the radio-frequency excitation is applied in series with the plate.

Measurement of Other Tube Factors.—The bridges discussed above may be adapted to the measurement of other tube factors of multigrid tubes.[2] Bridges have also been designed for the measurement of negative tube factors.[3] Circuits for the determination of inter-

[1] BALLANTINE, S., *Proc. I.R.E.*, **17**, 1164 (1929).

[2] Standards on Electronics, Institute of Radio Engineering, New York, 1938.

[3] CHAFFEE, E. LEON, "Theory of Thermionic Vacuum Tubes," Chap. IX, McGraw-Hill Book Company, Inc., New York, 1933; HICKMAN, R. W., and HUNT, F. C., *Rev. Sci. Instruments*, **6**, 268 (1935).

electrode capacitances, leakage conductances, tube characteristics, and other quantities are discussed in the 1938 report on electronics of the Standards Committee of the Institute of Radio Engineers.

15-34. Negative-resistance Bridges.—A bridge for the measurement of negative resistance is shown in Fig. 15-82.[1] The condition for balance is

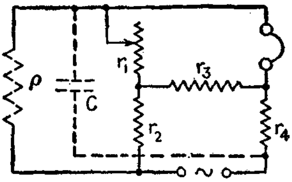


FIG. 15-82.—Bridge for the measurement of negative resistance.

$$\rho = \left(\frac{r_2}{r_3} + 1\right) r_1 + r_2 \qquad (15\text{-}62)$$

The resistances of Fig. 15-82 must be taken into account in determining the operating voltages of the negative-resistance element. Proper choice of resistances simplifies the criterion for balance. If $r_2 = 99$ ohms and $r_3 = 1$ ohm, Eq. (15-62) reduces to

$$\rho = 100r_1 + 99 \qquad (15\text{-}63)$$

If the negative-resistance element has appreciable capacitance, the balancing condenser C should be added.

Three modifications of Dingley's bridge that are capable of measuring a greater range of negative resistance were developed by Terman.[2] These circuits and the criteria for balance are shown in Fig. 15-83. r

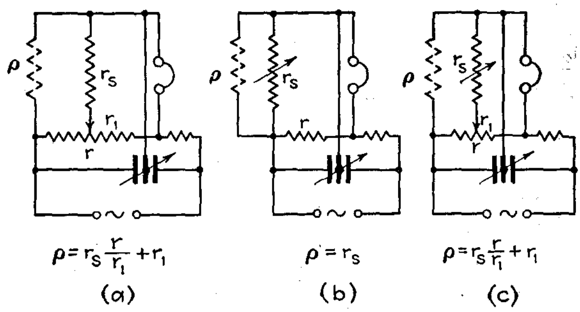


FIG. 15-83.—Three modifications of the negative-resistance bridge of Fig. 15-82.

should not exceed several hundred ohms, and r_s should be much larger than r_1. Ten thousand ohms is suggested for r_s in circuits b and c. The capacitance balance can be constructed from a two-gang variable condenser.

15-35. Tuttle Tube-factor Bridge.—The circuits that have been discussed employ a single source of excitation voltage. A different type of tube-factor bridge, in which currents caused by three separate

[1] DINGLEY, E. N., JR., *Proc. I.R.E.*, **19**, 1948 (1931).
[2] TERMAN, F. E., *Electronics*, December, 1933, p. 340.

voltages are balanced, has been developed by W. N. Tuttle.[1] The three forms of Tuttle's circuit used for the measurement of amplification factor, electrode resistance, and transconductance are shown in Figs. 15-84, 15-85, and 15-87. The transformers and attenuators in these

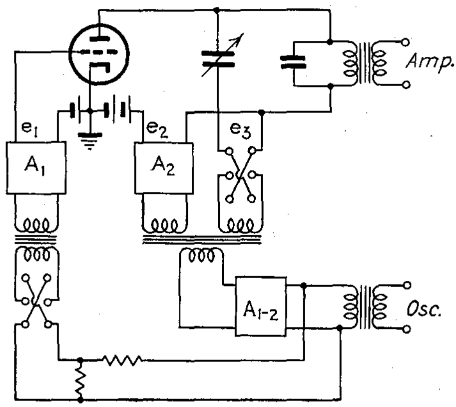


FIG. 15-84.—Bridge for the measurement of mu-factor.

circuits have been designed so that the three voltages e_1, e_2, and e_3 are exactly in phase (or 180 degrees out of phase). The function of e_3 is to furnish a reactive component of current to balance the reactive current caused by electrode and circuit capacitances.

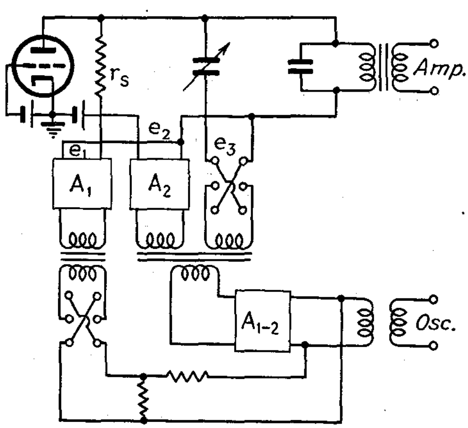


FIG. 15-85.—Bridge for the measurement of electrode resistance.

The principle of the circuit of Fig. 15-84, which measures amplification factor, is the same as that of the more common amplification-factor bridge of Fig. 15-73. The plate current resulting from e_2 in the plate circuit is balanced against the current resulting from e_1 in the grid circuit. The net

[1] TUTTLE, W. N., *Proc. I.R.E.*, **21**, 844 (1933).

current is zero when $e_2 = \mu e_1$. The voltage ratio is read directly on the attenuators.

The circuit of Fig. 15-85 measures plate resistance. An exact analysis of the operation of the circuit can be made by constructing the complete equivalent circuit and determining the conditions necessary to reduce the current in the output branch to zero. For the purpose of explaining the basic principle of operation, it is instructive to neglect the tube capacitances and to consider only the nonreactive components of current. Under the assumption that the capacitances are zero, e_3 is also zero, and the equivalent circuit is of the form shown in Fig. 15-86, in which z_0 is the impedance of the output transformer and condenser. The network equations are

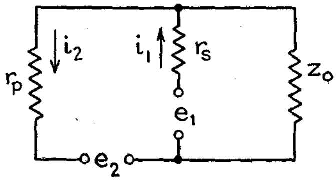


FIG. 15-86.—Equivalent circuit for the bridge of Fig. 15-85.

$$e_1 = i_1(r_s + z_o) - i_2 z_o \tag{15-64}$$
$$e_2 = -i_1 z_o + i_2(r_p + z_o) \tag{15-65}$$

Balance is obtained when $i_1 = i_2$, so that the current in the output branch is zero. Setting the currents equal and dividing Eq. (15-65) by Eq. (15-64) gives the condition for balance.

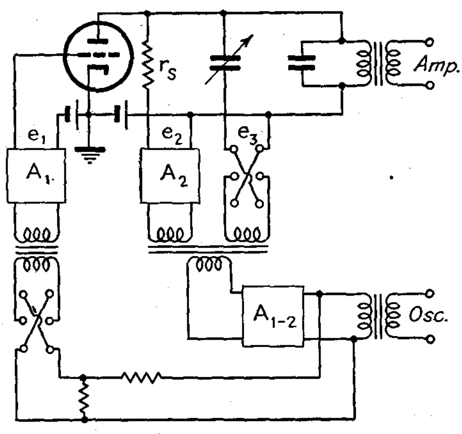


FIG. 15-87.—Bridge for the measurement of transconductance.

$$r_p = \frac{r_s e_2}{e_1} \tag{15-66}$$

The attenuators may be calibrated to read r_p directly.

The circuit of Fig. 15-87 is used to measure transconductance. The equivalent circuit, neglecting tube capacitances and grid current, is similar to that of Fig. 15-86; but since e_1 is applied in the grid circuit,

e_1 must be replaced by μe_2, and e_2 by e_1, in the equivalent circuit. The condition for balance is

$$\frac{\mu e_1}{r_p} = \frac{e_2}{r_s} \qquad \text{or} \qquad g_m = \frac{e_2}{e_1 r_s} \tag{15-67}$$

To measure negative factors, the phase of e_1 is reversed by means of the double-throw switch.

By using Tuttle's circuits, all tube factors may be read with ease with a single bridge. Another advantage results from the fact that all batteries may be kept at ground potential. The construction of a reliable portable instrument of this type necessitates great care in transformer design and in the location of component parts.

15-36. Harmonic Analyzers.—Harmonic analyzers for the direct measurement of harmonic amplitudes by electrical means are of four types: tuned circuit, heterodyne, dynamometer, and fundamental-suppression. Indirect measurements may also be made by taking oscillograms of the current or voltage and analyzing the oscillograms by mechanical or electrical analyzers or by graphical or selected-ordinate methods. The oscillographic method requires more time and is not so accurate as some of the direct methods.

15-37. Analyzers Using Tuned Circuits.—The use of tuned circuits for harmonic analysis was first suggested by Pupin in 1893 and was

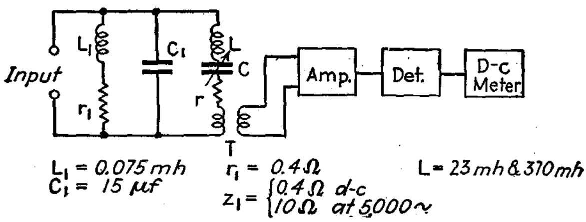


FIG. 15-88.—Tuned-filter type of harmonic analyzer.

subsequently used by him in the study of electrical apparatus.[1] In 1912, R. Beattie made an analysis to determine the best form of circuit for the tuned-circuit type of analyzer.[2] A portable analyzer based upon the tuned-filter principle was designed by Wegel and Moore in 1924.[3] The general features of their circuit are shown by Fig. 15-88. The voltage to be analyzed is applied across the series resonant circuit $LCrT$, and the voltage induced across the secondary of the transformer T is applied to the amplifier. The rectified output of the amplifier is read by

[1] PUPIN, M., *Am. J. Science*, **45**, 429 (1893); *Trans. Am. Inst. Elec. Eng.*, **11**, 523 (1894).

[2] BEATTIE, R., *Electrician*, **69**, 63 (1912).

[3] WEGEL, R. L., and MOORE, C. R., *Bell System Tech. J.*, **3**, 299 (1924); *Trans. Am. Inst. Elec. Eng.*, **43**, 457 (1924).

means of a d-c meter. The harmonic content of the applied voltage is determined by tuning the filter successively to the various harmonic frequencies and reading the output meter. The purpose of the parallel resonant circuit made up of L, r, and C_1 is to correct for the variation, with frequency, of a-c resistance of the series circuit and of the amplifier gain, so that the sensitivity remains essentially constant over the frequency range for which the instrument is designed.

Instead of taking the output of the filter from the transformer, as in Fig. 15-88, it is possible to use the voltage developed across L, C, or r. Beattie[1] and Morgan[2] have shown that it is usually best to take the voltage from the inductance. Since inductive reactance is directly proportional to frequency, the voltage across the inductance at a given current increases with frequency. This discrimination against the lower components of the input wave is desirable if the amplitudes of the lower frequency components of the input are greater than the higher frequency

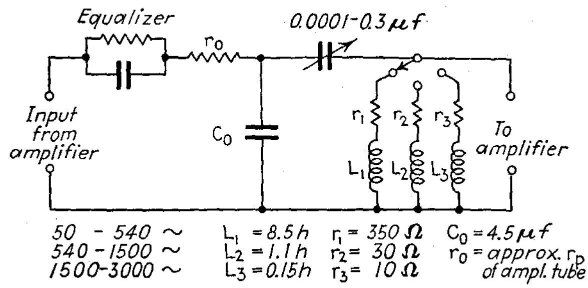


FIG. 15-89.—Amplifier-coupling filter for tuned-filter harmonic analyzer.

components, which is usually true. It helps to make possible the measurement of a small second harmonic in the presence of a large fundamental.

A singly tuned filter does not tune sharply enough to make possible measurement of the amplitude of a small-amplitude component whose frequency does not differ greatly from that of a large-amplitude component. To overcome this difficulty McCurdy and Blye[3] designed an analyzer which uses two tuned filters as coupling elements between amplifier tubes. The circuit of one filter stage is shown in Fig. 15-89. The first amplifier is preceded by a filter for suppressing the fundamental frequency. The output is read by means of a thermocouple meter.

15-38. Heterodyne Harmonic Analyzers.—In addition to the problem of obtaining sufficiently sharp tuning, the design of the tuned-filter type

[1] BEATTIE, *loc. cit.*

[2] MORGAN, F., *J. Inst. Elec. Eng.*, **71**, 819 (1932) (contains bibliography of 40 items).

[3] MCCURDY, R. G., and BLYE, P. W., *Trans. Am. Inst. Elec. Eng.*, **48**, 1167 (1929).

of analyzer is complicated by the variation of filter resistance with frequency. The audible frequency band cannot be readily covered with a single filter inductance. It is necessary to use a number of inductances and associated equalizers, each of which covers a portion of the band. These problems can be readily solved by using a fixed-frequency filter and heterodyning the output of a variable-frequency oscillator with the input voltage so as to produce a sum or difference frequency equal to the filter frequency. The fixed filter frequency makes it possible to use a highly selective filter.

A block diagram of a typical heterodyne harmonic analyzer is shown in Fig. 15-90. The frequency of the filter is higher than that of the highest input signal to be measured. The use of a balanced modulator to convert the impressed frequency to the frequency of the filter affords

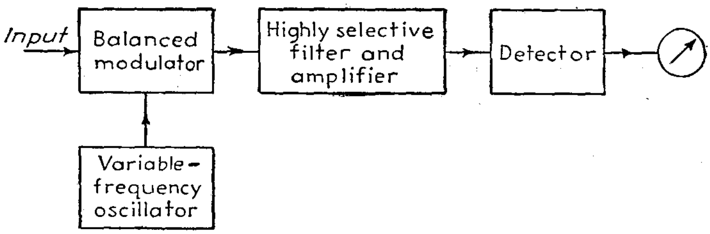


FIG. 15-90.—Block diagram of a heterodyne harmonic analyzer.

a simple means of eliminating the impressed frequency and at the same time ensures lower nonlinear distortion than could be obtained with a single-tube frequency converter. Some heterodyne analyzers may be calibrated to read directly; in others the voltage of the impressed signal is determined by measuring the voltage of a reference signal, the amplitude of which is made equal to that of the impressed signal. The various heterodyne analyzers that have been developed differ mainly as to the type of filter used. Early circuits made use of mechanical filters[1] and of filters tuned to a frequency lower than that of the input voltage.[2] In more recent analyzers, however, the filter consists of quartz crystals[3] or of two or more stages of inverse-feedback filters such as those of Fig. 6-45.[4] An important advantage of inverse-feedback filters is that they make possible the variation of selectivity at constant amplification by adjustment of feedback.

A heterodyne analyzer incorporating a two-section 50-kc quartz-crystal filter is shown in Fig. 15-91. The oscillator is tuned so that the

[1] MOORE, C. R., and CURTIS, A. S., *Bell System Tech. J.*, **6**, 217 (1927).

[2] LANDEEN, A. G., *Bell System Tech. J.*, **6**, 230 (1927).

[3] CASTNER, T. G., *Bell Laboratories Record*, **13**, 258 (1935).

[4] TERMAN, F. E., BUSS, R. R., HEWLETT, W. R., and CAHILL, F. C., *Proc. I.R.E.*, **27**, 649 (1939).

sum of the input frequency and the oscillator frequency is equal to 50 kc. The resistance R and the capacitance C correct for any slight unbalance of the balanced modulator. They are adjusted so as to eliminate the oscillator fundamental frequency from the output of the modulator. The output stage is a vacuum-tube voltmeter with a balancing circuit for eliminating the zero-signal current from the meter. The tuned-plate oscillator is designed to give constant amplitude throughout its range of 35 to 50 kc.

The instrument is calibrated by adjusting the amplifier gain to give a 2-volt reading of the output meter when 1 volt of direct voltage, obtained from the filament battery, is applied between the grids of the

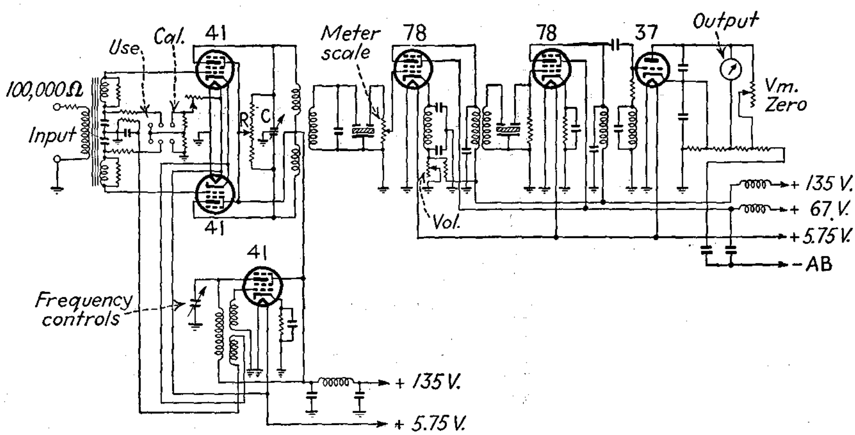


FIG. 15-91.—Heterodyne wave analyzer using a quartz-crystal filter. (*Courtesy of General Radio Co.*)

balanced modulator and the oscillator is tuned to the filter frequency. The justification for this method of calibration follows from the analysis of the balanced modulator. Equation (9-13) demonstrates that two components of the output voltage of a balanced modulator are $2Aa_2E_1E_2 \cos 2\pi(f_1 + f_2)t$ and $2Aa_2E_1E_2 \cos 2\pi(f_1 - f_2)t$ when the input voltage is made up of two components $E_1 \sin \omega_1 t$ and $E_2 \sin \omega_2 t$. As f_2 approaches zero, the sum and difference frequencies become more and more nearly equal and, in the limit, when f_2 is zero, are both equal to f_1. When one voltage is direct, therefore, the output voltage is $4Aa_2E_1E_2 \cos 2\pi f_1 t$, which has twice the amplitude of the sum- and difference-frequency components obtained when both input voltages are alternating. The frequency range of the General Radio wave analyzer is 20 to 15,000 cps, and the voltage range 200 μv to 200 volts.

In another type of heterodyne analyzer, devised by C. G. Suits, the unknown voltage and the voltage from the oscillator are applied simultaneously to the grid circuit of a square-law detector.[1] The plate current contains the impressed frequencies and their harmonics, steady components, and intermodulation components. A d-c milliammeter in the plate circuit responds to the steady components and to any intermodulation frequencies that do not greatly exceed the resonance frequency of the meter movement. By adjusting the oscillator frequency to differ only slightly from the frequency of a component of the unknown voltage, the needle can be caused to oscillate at the resulting difference frequency. Ordinarily only one difference frequency at a time will be low enough to affect the meter. According to principles set forth in Chaps. 3 and 9, the change in reading of the meter with application of the signal voltage will then be

$$\Delta I = a_2[\tfrac{1}{2}E_o^2 + \Sigma\tfrac{1}{2}E_n^2 + E_oE_x \cos 2\pi(f_x - f_o)t] \tag{15-68}$$

where f_o and E_o are the frequency and amplitude of the oscillator voltage, f_n and E_n are the frequency and amplitude of any component of the input voltage, and f_x and E_x are the frequency and amplitude of the particular component of the input voltage that is being measured. The milliammeter needle will oscillate through a current range that is twice the amplitude of the difference-frequency component of the plate current. This fact may be expressed by the equation

$$\Delta I_x = 2a_2E_oE_x \tag{15-69}$$

If the oscillator voltage does not vary with frequency, the amplitude of oscillation of the meter is proportional only to the amplitude of the unknown voltage, and the instrument may be calibrated by means of a single input voltage of known amplitude. In the determination of per cent harmonic relative to the fundamental, only relative values are required, and the instrument need not be calibrated if the oscillator voltage does not vary with frequency. If the oscillator is not designed to give constant output at all frequencies, the analyzer may be calibrated by applying inputs having known voltage and frequency.

Errors in reading result from two sources: oscillator harmonics and departure of the detector characteristic from a true square law. The first type of error is small if the harmonic content of the oscillator is below 1 or 2 per cent. The second results mainly from the presence of the third-order term in the series expansion for the detector plate current and may be minimized by keeping the oscillator and input voltages small. It may be practically eliminated by the use of a balance detector, similar to that of Fig. 15-2.[2]

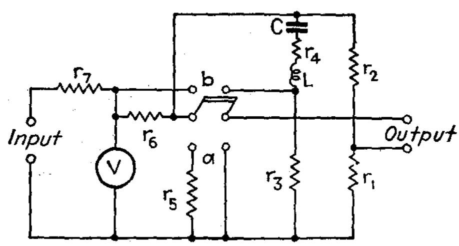
[1] SUITS, C. G., *Proc. I.R.E.*, **18**, 178 (1930). See also R. R. CHILTON, A Practical Wave Analyzer for Distortion Measurements, *Technical Bull., Inst. Rad. Eng. (Australia)*, 1941.

[2] GREENWOOD, W., *Wireless Eng.*, **9**, 310 (1932).

15-39. Dynamometer-type Harmonic Analyzers.—In the dynamometer type of analyzer the unknown voltage is applied, through an amplifier, to one coil of the dynamometer and the output of an oscillator to the other.[1] If the oscillator frequency is adjusted so that it differs only slightly from the frequency of the component of the unknown voltage to be measured, the meter needle will oscillate at the difference frequency. The amplitude of oscillation is proportional to the product of the amplitudes of the currents in the two coils. If the current from the oscillator is kept constant with the help of an additional meter, the amplitude of oscillation of the dynamometer needle is proportional only to the amplitude of the component of the unknown voltage. The instrument may, therefore, be calibrated to read the input amplitude directly.

It is also possible to construct dynamometer-type analyzers that give steady deflection.[2] The difficulty of obtaining an oscillator current of good wave form which remains exactly in phase with the voltage that is being measured makes it impractical to use this type of analyzer over the frequency range required for the testing of audio-frequency amplifiers.

15-40. Fundamental-suppression Harmonic Analyzers.—In the fundamental-suppression type of analyzer the fundamental component of the



$$\frac{r_1}{r_2} = \frac{r_3}{r_4} = k$$
$$r_1 + r_2 = r_5 + r_6$$
$$r_1 + r_2 < r_3 + r_4$$

FIG. 15-92.—Practical form of fundamental-suppression type of harmonic analyzer.

unknown voltage is removed by some form of bridge. The remaining r-m-s output from the bridge is proportional to the square root of the sum of the squares of all the harmonics. From the values of the total harmonic content and of the complete input, the fundamental component can be determined. The resonance bridge of Fig. 15-41*a* may be used for this purpose.[3] Figure 15-92 shows this bridge in convenient form for

[1] NICHOLSON, M. G., and PERKINS, W. M., *Proc. I.R.E.*, **20**, 734 (1932).

[2] COCKROFT, J. D., COE, R. T., TYACKE, J. A., and WALKER, M., *J. Inst. Elec. Eng.*, **63**, 69 (1925); *Proc. Phys. Soc. London*, **40**, 228 (1928).

[3] BELFILS, G., *Rev. gén. élec.*, **19**, 523 (1926).

harmonic measurement.[1] When the switch is in position a, the output is approximately

$$(\text{total input}) \times \frac{k}{k+1} \times \frac{r_5}{r_5 + r_6}$$

When the switch is in position b, the bridge output is equal to

$$(\text{vector sum of the harmonics}) \times \frac{k}{k+1}.$$

The bridge output is measured by means of a vacuum-tube voltmeter. The input voltage to the bridge may change with adjustment of the bridge and should be kept constant. Then if the voltmeter reads the same for the two switch positions, small distortion factors are equal to R_5/R_6. The parallel resonance bridge of Fig. 15-41b may be used in place of the series resonance bridge.[2] High- and low-pass filters may also be used for fundamental suppression.[3]

The fundamental-suppression type of bridge is particularly useful in determining the distortion factor δ (see Sec. 4-18). The fundamental-suppression bridge may often be used advantageously in combination with other methods of analysis. If the fundamental is removed by means of a bridge, for instance, the harmonic content may be analyzed with considerable accuracy by oscillographic and graphical methods.[4] Fundamental suppression is often advisable in connection with tuned-circuit and heterodyne analyzers because it greatly reduces the production of harmonics in the analyzer; it also reduces the selectivity requirements of the analyzer. Reduced selectivity is essential when the frequency of the unknown voltage drifts appreciably during the time required to make a measurement.

15-41. Measurement of Voltage Amplification.—Figure 15-93 shows the circuit that is commonly used for the measurement of voltage amplification of an amplifier. When the attenuation of the attenuator is equal to the gain of the amplifier, the deflection of the vacuum-tube voltmeter will not change when the position of the switch is changed. If the attenuator is calibrated in decibels, it gives a direct reading of the amplifier gain in decibels. The purpose of the resistor R is to provide the proper terminating impedance for the attenuator. If a calibrated attenuator is not available, it may be replaced by a voltage divider, as in Fig. 15-94. The voltage amplification A is equal to the ratio r_2/r_1

[1] WOLFF, IRVING, *J. Opt. Soc. Am. and Rev. Sci. Instruments*, **15,** 163 (1927).

[2] WAGNER, H. M., *Proc. I.R.E.*, **23,** 85 (1935); see also H. H. SCOTT, *Proc. I.R.E.*, **26,** 226 (1938).

[3] MCCURDY and BLYE, *loc. cit.*

[4] BROWN, S. L., *Phys. Rev.*, **31,** 302 (1928); PIDDINGTON, J. H., *Proc. I.R.E.*, **24,** 591 (1936).

when the same voltmeter reading is obtained with the switch in the two positions. Although affording a less accurate means of comparing the input and output voltages, a cathode-ray oscilloscope or an electron-ray tube may be used in place of the vacuum-tube voltmeter in the circuits of Figs. 15-93 and 15-94.

The response curve of an amplifier or other network may be observed directly by means of a cathode-ray oscilloscope.[1] This is accomplished by modulating the frequency of the amplifier excitation voltage in accordance with the sweep voltage and applying the amplifier output voltage to the vertical deflection plates of the oscilloscope. If the excitation amplitude is independent of frequency, the envelope of the pattern

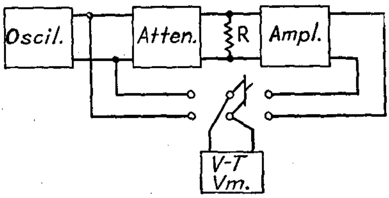


FIG. 15-93.—Circuit for the measurement of voltage amplification.

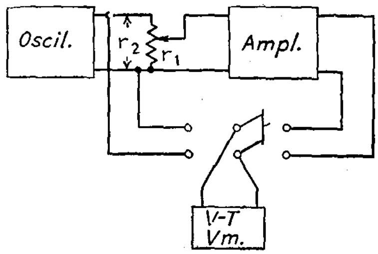


FIG. 15-94.—Circuit for the measurement of voltage amplification.

obtained on the screen is the response curve, plotted on a uniform frequency scale.

15-42. The Use of Triangular and Rectangular Waves in Amplifier Analysis.—In order to amplify a triangular or rectangular wave of voltage without observable distortion, an amplifier must have negligible frequency and phase distortion over a frequency range covering the fundamental frequency and approximately the first ten harmonics. This fact may be used to advantage in checking amplifiers oscillographically for frequency distortion. The distortion resulting from insufficient amplification of the lower-frequency components, and of the higher-frequency components, is shown by waves *a* and *b*, respectively, of Fig. 15-95. Transient oscillations resulting from resonance of transformer inductance and distributed capacitance result in an output wave of the form of wave *c*. The ratio of the resonance frequency to the frequency of the rectangular wave may be readily determined from such a wave. Rectangular waves may be used in the determination of amplifier response by applying the output of a variable-frequency rectangular-wave generator to the input of the amplifier and observing the output of the

[1] DIAMOND, H., and WEBB, J. S., *Proc. I.R.E.*, **15**, 767 (1927).

amplifier by means of an oscilloscope.[1] Negligible departure of the output wave from rectangular form over any range of frequency is an indication that the amplifier will give essentially undistorted amplification of complicated waves throughout this frequency range.

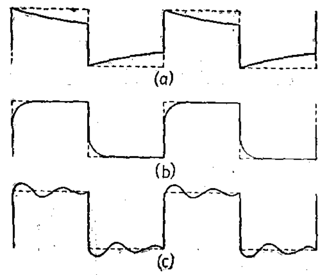


FIG. 15-95.—Distortion of rectangular wave caused by: (a) poor low-frequency response; (b) poor high-frequency response; (c) transient oscillations in coupling transformers.

15-43. Measurement of Power Output. The power output of an amplifier with known load resistance may be determined by measuring the r-m-s voltage across the load or the r-m-s current through the load. A thermal meter or a copper oxide rectifier meter is usually used for the purpose. Figure 15-96 shows the circuit diagram of the General Radio output meter, by means of which power may be quickly measured at various values of effective load resistance. The meter consists of a variable-ratio transformer compensated by means of resistances so that a constant percentage of the power is expended in the secondary load. The load is a constant-resistance network which also serves the function of a four-range multiplier for the voltmeter. The indicating instrument is a copper oxide rectifier

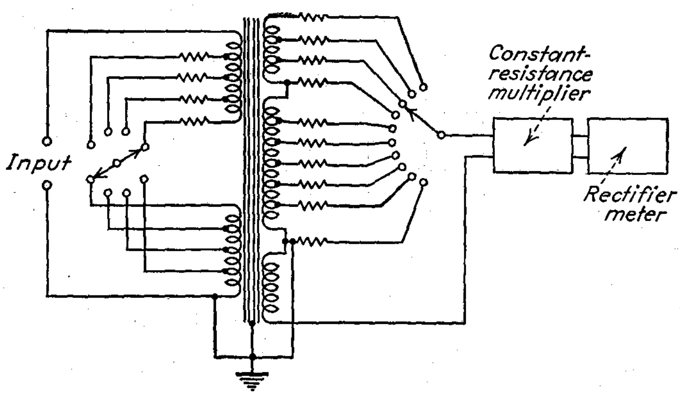


FIG. 15-96.—Circuit diagram of General Radio power-output meter.

voltmeter. The effective load may be varied in 40 steps from 2.5 to 20,000 ohms. The multiplier provides 5-, 50-, 500-, and 5000-mw ranges of power. The scale is also calibrated to read decibel power level relative to 1 mw zero level.

[1] REICH, H. J., *Proc. I.R.E.*, **19**, 401 (1931); STOCKER, A. C., *Proc. I.R.E.*, **25**, 1012 (1937); SWIFT, G., *Communications*, February, 1939, p. 22; BEDFORD, A. V., and FREDENDAHLE, G. L., *Proc. I.R.E.*, **27**, 277 (1939); ARGUIMBAU, L. B., *Gen. Rad. Expt.*, **14**, December, 1939, p. 1; WAIDELICH, D. L., *Proc. I.R.E.*, **32**, 339 (1944).

15-44. Determination of Optimum Power Output and Optimum Load.[1]—The determination of optimum power output and optimum load

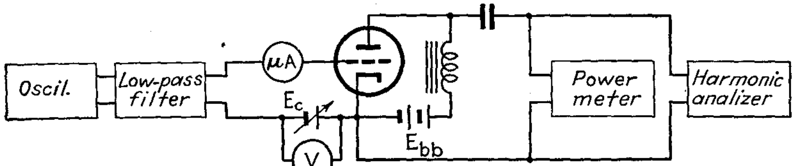


FIG. 15-97.—Circuit for the determination of optimum power output and optimum load resistance.

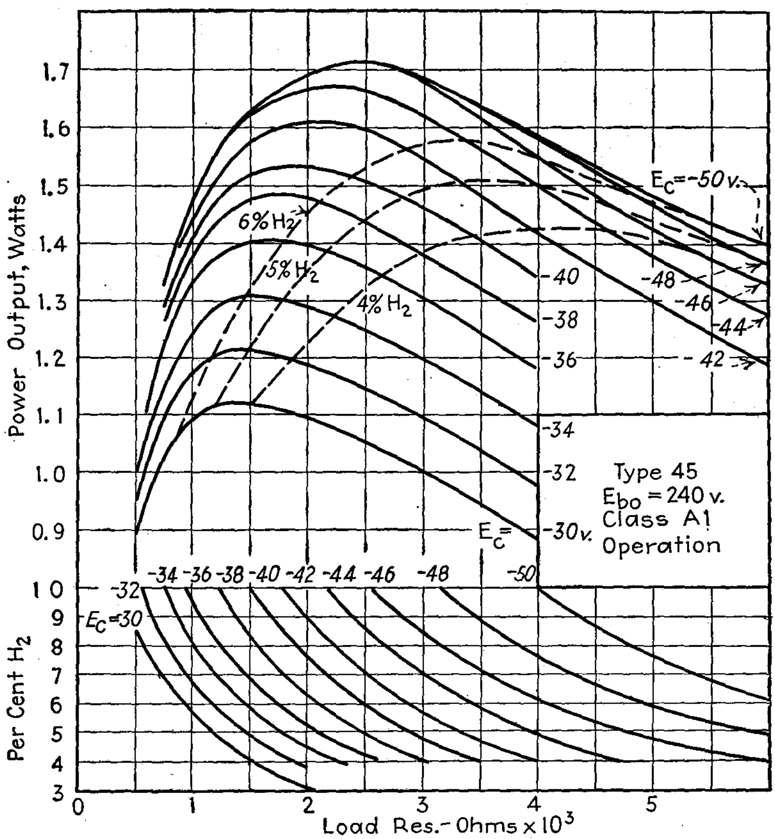


FIG. 15-98.—Experimentally determined curves of power output and harmonic content of a type 45 tube. The dashed curves show power output at constant percentage of second harmonic.

requires a series of measurements. A suitable circuit is shown in Fig. 15-97. For each load impedance there is a limiting value of bias and

[1] KELLOGG, E. W., *J. Am. Inst. Elec. Eng.*, **40,** 490 (1925); WARNER, J. C., and LOUGHREN, A. V., *Proc. I.R.E.*, **14,** 735 (1928); HANNA, C. R., SUTHERLIN, L., and UPP, C. B., *Proc. I.R.E.*, **16,** 462 (1928).

signal amplitude beyond which the harmonic content exceeds the allowable values. Curves of power output and harmonic content as a function of load impedance must be constructed for various values of bias. The simplest laboratory procedure is to set the bias at fixed values and to take readings of power output and harmonic content at five or six values of load impedance. If operation is to be restricted to the region in which no grid current flows, the signal must be adjusted for each value of bias and load so that grid current does not quite flow. The signal amplitude at which grid current commences can be determined by a galvanometer or headphones in series with the grid.

Figure 15-98 shows curves of power and harmonic content for a type 45 tube. The load impedance at which the harmonic content is just equal to the maximum allowable value at each value of bias may be determined from the curves of harmonic content. These values of impedance and bias then determine points on the power curves corresponding to the given harmonic content. The dashed curves of Fig. 15-98 show the power output at 4, 5, and 6 per cent second harmonic as a function of load resistance. From these the optimum load and power output at the given value of distortion may be read. Readings of third harmonic taken simultaneously with second harmonic showed the former to be negligible in comparison with the latter.

Problem

15-1. Carefully study the circuit of Fig. 15-59 and explain the functions of all tubes and circuit elements.

APPENDIX

A-1. Parallel Equivalent Circuits.—Figure A-2 shows the series equivalent plate circuit for the general form of triode circuit of Fig. A-1, in which the load z_b may include one or more impressed voltages. The voltage between A and B is $\mu E_g - I_p r_p$, which may be written in the form $r_p(g_m E_g - I_p)$. It is evident, therefore, that the

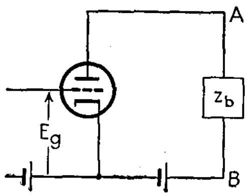


FIG. A-1.

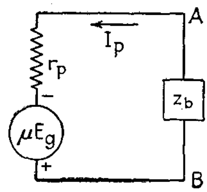


FIG. A-2.

voltage between A and B and, hence, all voltages and currents in the load, are unaltered if the equivalent voltage μE_g is removed and an additional current $g_m E_g$ is sent through r_p in the direction opposite to that in which μE_g tends to send current. The series equivalent circuit of Fig. A-2 may therefore be replaced by the parallel equivalent circuit of Fig. A-3. The current $g_m E_g$ may be assumed to be caused to flow by a constant-current generator, as shown in Fig. A-4.

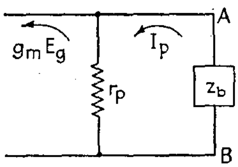


FIG. A-3.

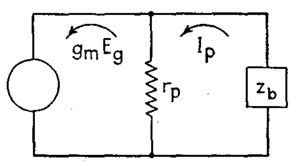


FIG. A-4.

In the series equivalent circuit, in which the equivalent generator is assumed to supply a constant voltage μE_g, the current must vary with load and plate impedance. In the parallel equivalent circuit, on the other hand, in which the generator is assumed to supply a constant current $g_m E_g$, the voltage across the load and generator must vary with load and plate impedance. It should be noted that the polarity of the voltage across the constant-current generator is determined solely by the voltage across the load. If the load contains an alternating e.m.f., the upper terminal may conceivably be positive, *i.e.*, the voltage across the generator may oppose the current through the generator. For this reason the polarity of the constant-current generator is not indicated.

Although the positive directions of most of the circuit currents and voltages may be chosen at random, the direction of the equivalent current must be properly chosen relative to the alternating grid voltage. If the instantaneous grid voltage is assumed to be positive when it makes the grid positive relative to the cathode, the equivalent current $g_m E_g$ should be indicated as flowing through the constant-current generator from the plate terminal to the cathode terminal as shown in Fig. A-4.

The procedure to be followed in the formation of the parallel equivalent circuit differs from that given in Sec. 4-2 for the series equivalent circuit only in the first step, which is as follows:

1. Connect the plate resistance and an equivalent constant-current generator in parallel between the plate and the cathode of the tube. The current supplied by the generator is g_mE_g and should be indicated as flowing through the generator from the plate terminal to the cathode terminal. If alternating voltage is impressed upon more than one grid, there is a similar component of current for each additional grid that is excited.

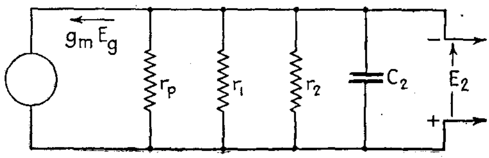


FIG. A-5.

The parallel equivalent circuit is particularly useful when the load consists of a number of parallel branches which do not contain an e.m.f., as, for example, in the resistance-capacitance–coupled amplifier at high frequencies, the parallel equivalent circuit of which is shown in Fig. A-5. It can be seen that the output voltage E_{g2} is minus the product of the current g_mE_{g1} by the resultant impedance of the four parallel branches. Dividing this product by E_{g1} at once gives the voltage amplification, Eq. (6-15). In this example the parallel equivalent circuit clearly affords a more direct method of determining the voltage amplification than does the series equivalent circuit, used in Sec. 6-5. The parallel equivalent circuit does not, however, ordinarily simplify the analysis when voltages are impressed in both the grid and the plate circuits and simultaneous network equations must be solved.

A-2. Power Relations in Vacuum-tube Plate Circuits.—The power supplied to the plate circuit is

$$P_i = \frac{1}{T}\int_0^T E_{bb}i_b\,dt = E_{bb}\frac{1}{T}\int_0^T i_b\,dt = E_{bb}I_{ba} \tag{A-1}$$

where T is the period of the fundamental component of plate current. But

$$i_b = I_{ba} + i_{pa} \tag{A-2}$$

where i_{pa} is the instantaneous alternating plate current, measured relative to the average value I_{ba}. Furthermore,

$$E_{bb} = E_{ba} + I_{ba}R_b + e_{pa} + e_{zb} \tag{A-3}$$

where e_{pa} is the instantaneous alternating plate voltage, measured relative to the average value E_{ba}, and e_{zb} is the instantaneous alternating voltage across the load, measured relative to the average value $I_{ba}R_b$.

Equation (A-1) may, therefore, be written in the form

$$P_i = \frac{1}{T}\int_0^T (I_{ba} + i_{pa})(E_{ba} + I_{ba}R_b + e_{pa} + e_{zb})\,dt \tag{A-4}$$

$$= I_{ba}E_{ba} + I_{ba}^2R_b + I_{ba}\frac{1}{T}\int_0^T e_{pa}\,dt + I_{ba}\frac{1}{T}\int_0^T e_{zb}\,dt +$$

$$E_{ba}\frac{1}{T}\int_0^T i_{pa}\,dt + I_{ba}R_b\frac{1}{T}\int_0^T i_{pa}\,dt + \frac{1}{T}\int_0^T i_{pa}e_{pa}\,dt + \frac{1}{T}\int_0^T i_{pa}e_{zb}\,dt \tag{A-5}$$

Since i_{pa}, e_{pa}, and e_{zb} are measured with respect to average values, they do not contain steady components. Therefore, all but the last two integrals of Eq. (A-5) are zero. The power input is

$$P_i = I_{ba}E_{ba} + I_{ba}^2R_b + \frac{1}{T}\int_0^T i_{pa}e_{pa}\,dt + \frac{1}{T}\int_0^T i_{pa}e_{zb}\,dt \tag{A-6}$$

The first term of Eq. (A-6) represents the d-c power expended in the tube, or d-c plate dissipation. The second term represents the d-c power developed in the load. The third term represents the a-c power expended in the tube, or a-c plate dissipation. The last term represents the a-c power developed in the load.

The total plate dissipation is

$$P_p = I_{ba}E_{ba} + \frac{1}{T}\int_0^T i_{pa}e_{pa}\,dt \tag{A-7}$$

The a-c power output is

$$P_o = \frac{1}{T}\int_0^T i_{pa}e_{zb}\,dt \tag{A-8}$$

Since the plate supply voltage is constant, rise in voltage across the load is accompanied by an equal reduction in plate voltage, and so the alternating plate and load voltages are equal in magnitude, but opposite in phase.

$$e_{pa} = -e_{zb} \tag{A-9}$$

Therefore

$$\frac{1}{T}\int_0^T i_{pa}e_{pa}\,dt = -P_o \tag{A-10}$$

and

$$P_p = I_{ba}E_{ba} - P_o \tag{A-11}$$

Since the resistive component of load impedance causes the instantaneous plate voltage to fall with increase of plate current, e_{pa} has a component that is in phase opposition to i_{pa} and the second term of Eq. (A-7) is negative. Excitation therefore causes the plate dissipation to decrease. Equation (A-11) shows that the reduction of plate dissipation is equal to the power output. This fact and the negative value of the second term of Eq. (A-7) can be interpreted as indicating that the tube acts as a source of power delivered to the load or, more correctly, that the tube converts d-c power furnished by the plate supply into a-c power in the load.

A-3. Linear Modulation.—Application of a sinusoidal carrier excitation voltage to a circuit containing an element that conducts in only one direction results in the production of pulses of current. By Fourier analysis the current may be analyzed into a steady component and components having frequencies equal to the applied frequency and its harmonics. The addition of a steady biasing voltage in series with the carrier excitation voltage changes the portion of the cycle during which current flows. It thereby changes the amplitude of the current pulses and, therefore, of the fundamental component of current. Expressions for the fundamental component of current I_k and for the fundamental component of output voltage E_k across a resistance load may be derived as follows. Let the carrier excitation voltage be $E_2 \cos \omega_k t$, and let the biasing voltage be E_b. Current flows only when the total instantaneous applied voltage $E_2 \cos \omega_k t + E_b$, exceeds zero. If θ_0 is the value of $\omega_k t$ at which the current is cut off, then $E_2 \cos \theta_0 + E_b = 0$, or $\cos \theta_0 = -E_b/E_2$. If the circuit contains only resistance and the characteristic curve of the rectifier is linear, the instantaneous current is proportional to the instantaneous applied voltage when this voltage is positive, and is zero during the remainder of the cycle. Thus the current is

$$i = KE_2\left(\cos \omega_k t + \frac{E_b}{E_2}\right) = I_{\max}(\cos \omega_k t - \cos \theta_0), \qquad -\theta_0 < \omega_k t < +\theta_0 \tag{A-12}$$

in which K is a constant of proportionality, and $I_{\max} = KE_2$ = the crest instantaneous current when $E_b = 0$. Expressed as a Fourier series the current is

$$i = A_0 + A_1 \sin \omega_k t + B_1 \cos \omega_k t + \cdots \tag{A-13}$$

$$A_1 = \frac{1}{\pi}\int_{-\pi}^{\pi} i \sin \omega_k t \, d(\omega_k t) = \frac{I_{\max}}{\pi}\int_{-\theta_0}^{+\theta_0} (\cos \omega_k t - \cos \theta_0) \sin \omega_k t \, d(\omega_k t) = 0 \tag{A-14}$$

$$B_1 = \frac{I_{\max}}{\pi}\int_{-\theta_0}^{+\theta_0} (\cos \omega_k t - \cos \theta_0) \cos \omega_k t \, d(\omega_k t) = \frac{I_{\max}}{\pi}(\theta_0 - \sin \theta_0 \cos \theta_0) \tag{A-15}$$

Therefore, the amplitude of the fundamental component of current is

$$I_k = \frac{I_{\max}}{\pi}(\theta_0 - \sin \theta_0 \cos \theta_0) \tag{A-16}$$

The amplitude of the fundamental component of output voltage across a resistance load is

$$E_k = \frac{I_{\max}R_e}{\pi}(\theta_0 - \sin \theta_0 \cos \theta_0) \tag{A-17}$$

in which R_e is the effective resistance of the load at carrier frequency. The solid curve of Fig. 9-10 was plotted by means of Eq. (A-17).

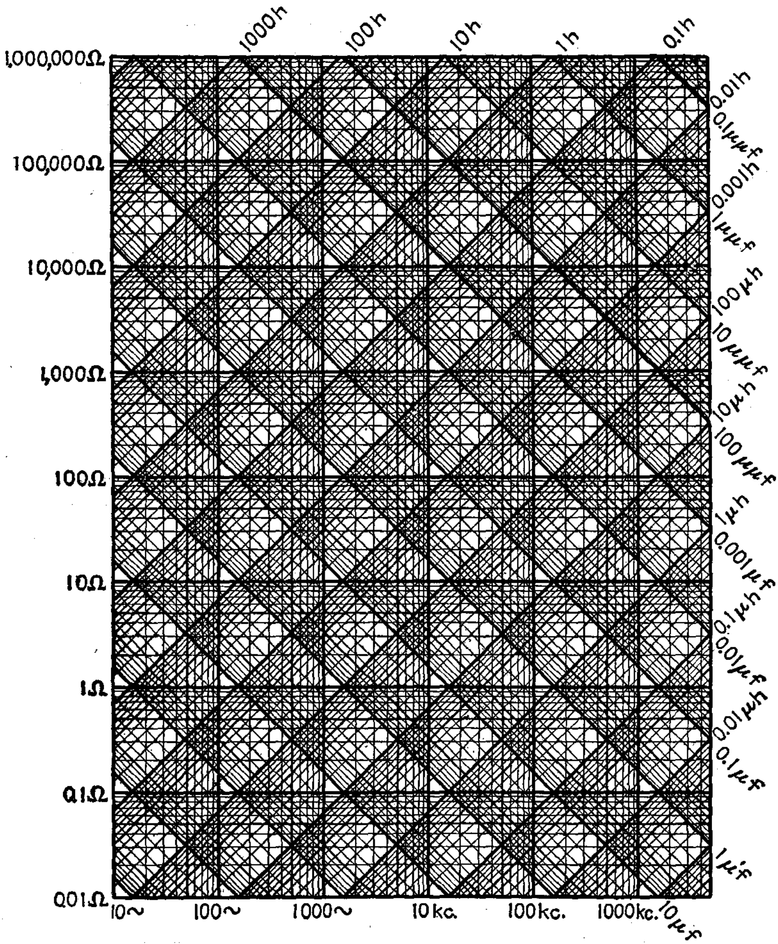


FIG. A-6.—Chart for the determination of inductive and capacitive reactance and of resonance frequency. The reactance at any frequency is determined by the intersection of the vertical line corresponding to the frequency with the diagonal line corresponding to the capacitance or inductance. The intersection of an inductance line with a capacitance line gives the frequency at which the inductance and capacitance are in resonance.

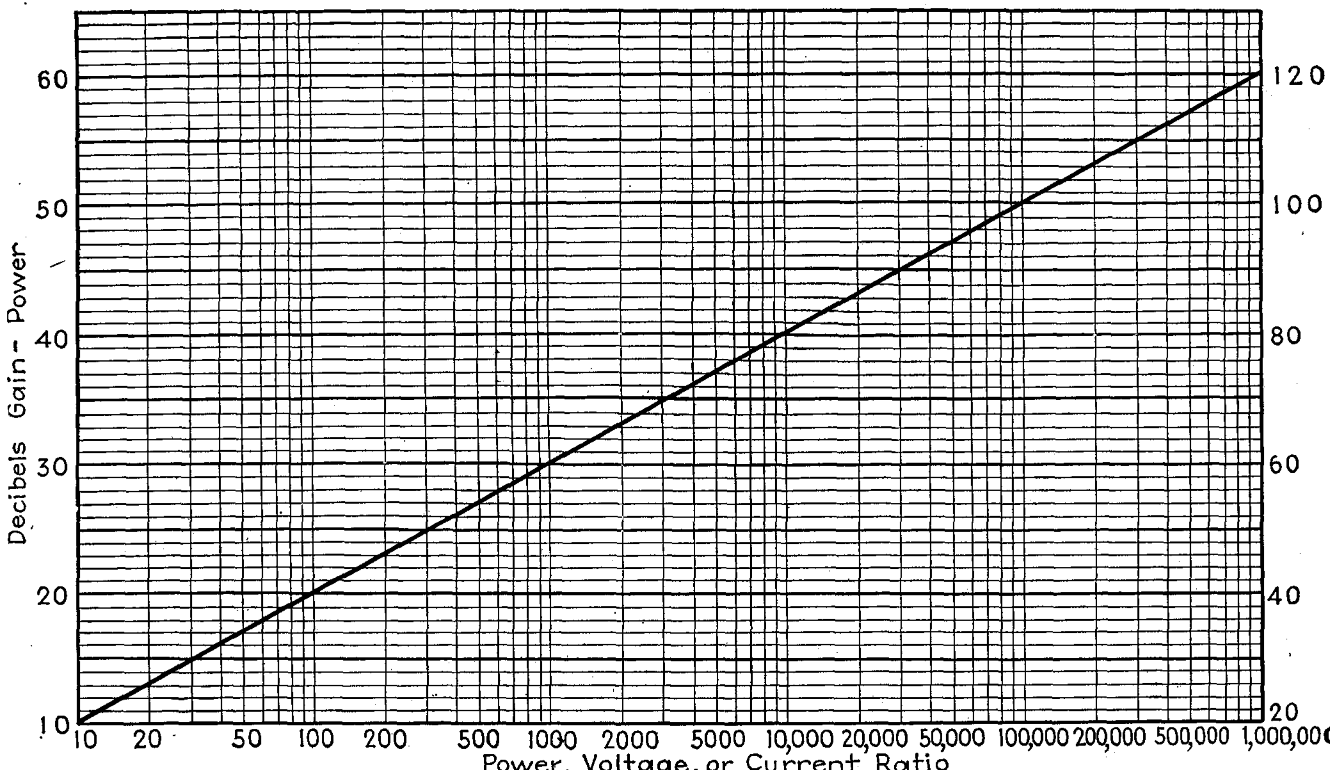


FIG. A-7.—Chart for the determination of decibel gain.

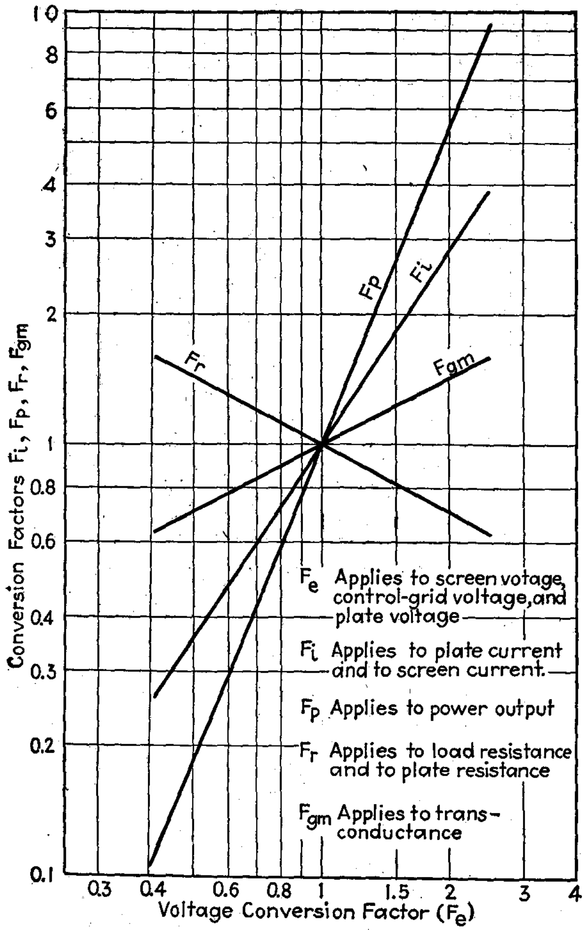


FIG. A-8.—Conversion factors for power amplifier triodes and pentodes.

In using these curves, the ratio of the new plate voltage to the published plate voltage nearest the desired new operating point is first determined. This ratio, the voltage conversion factor F_e, is then used to determine, from the curves, the factors F_i, F_p, F_r, and F_{gm}. F_e is also used to determine the new screen and control-grid voltages. *(Courtesy of RCA Manufacturing Company, Inc.)*

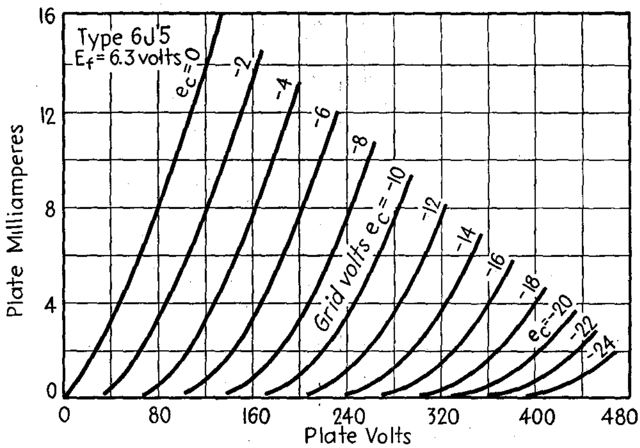


FIG. A-9.—Average plate characteristics for the type 6J5 triode.

AVERAGE PLATE CHARACTERISTICS

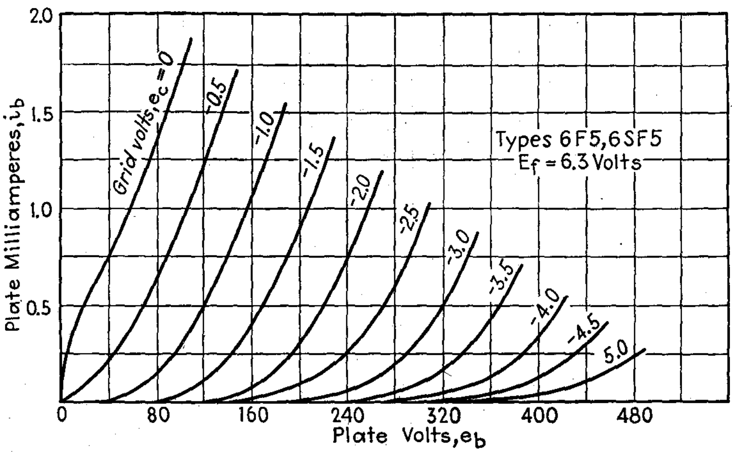


FIG. A-10.—Average plate characteristics for the types 6F5 and 6SF5 triodes.

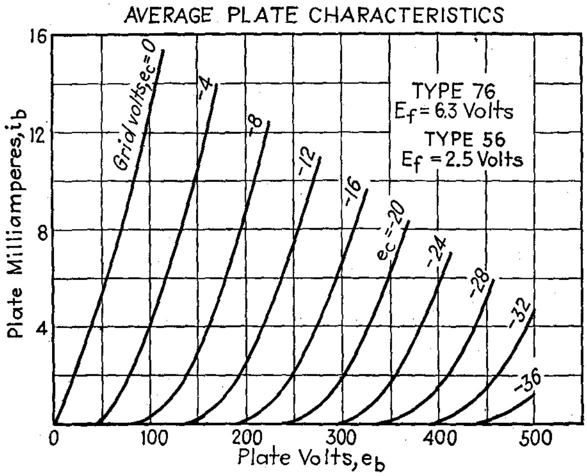


FIG. A-11.—Average plate characteristics for the types 76 and 56 triodes.

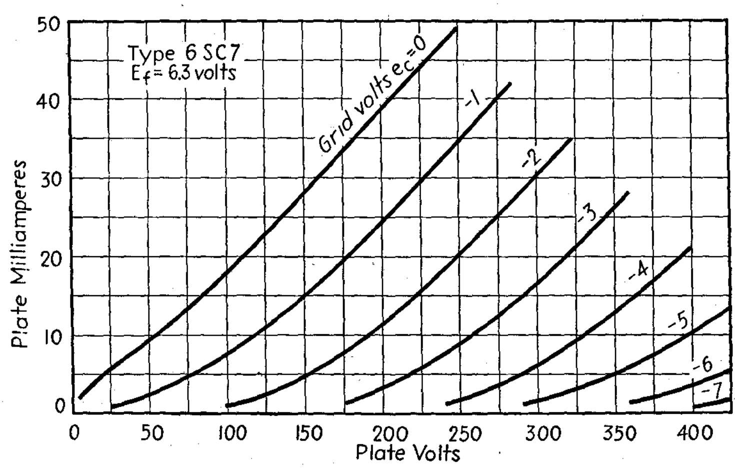


FIG. A-12.—Average plate characteristics for each unit of the type 6SC7 twin triode.

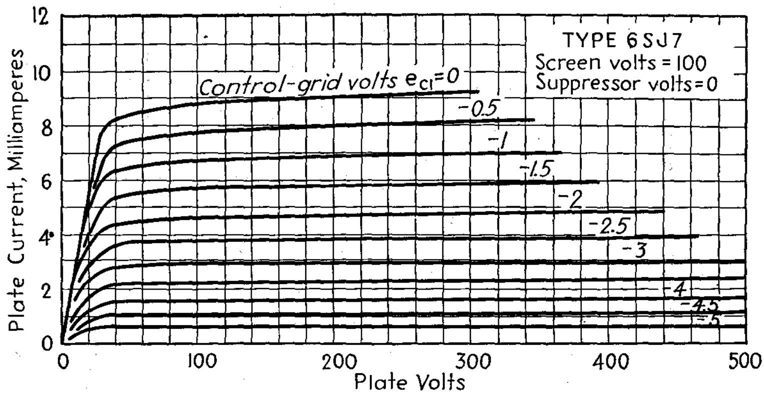


FIG. A-13.—Average plate characteristics for the type 6SJ7 pentode with 100-volt screen voltage.

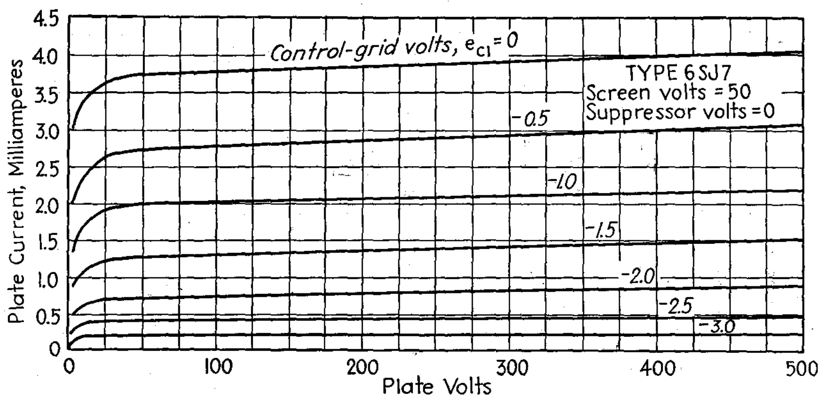


FIG. A-14.—Average plate characteristics for the type 6SJ7 pentode with 50-volt screen voltage.

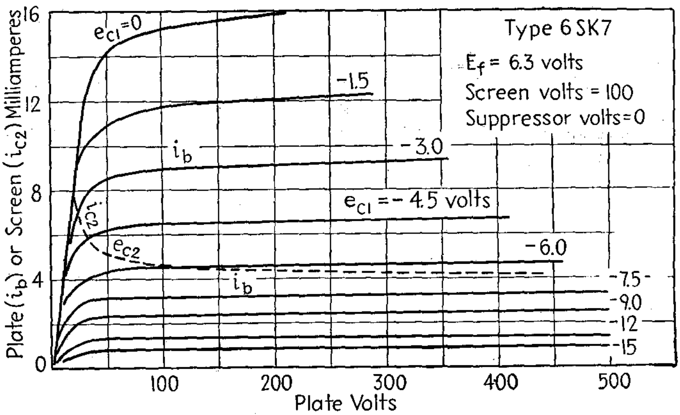


FIG. A-15.—Average plate characteristics for the type 6SK7 pentode.

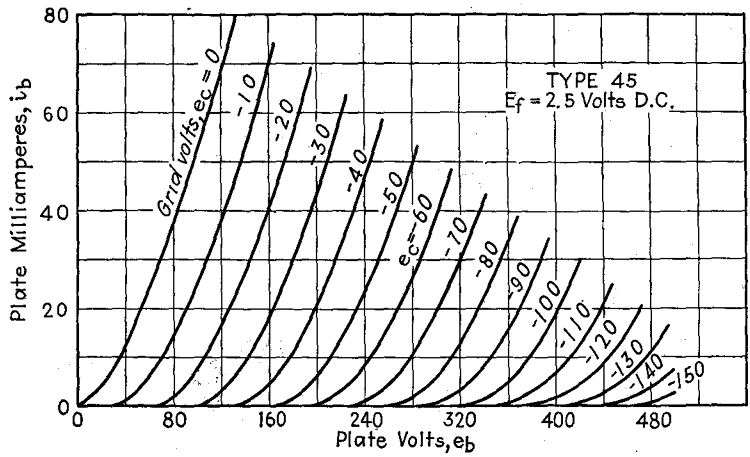


FIG. A-16.—Average plate characteristics for the type 45 triode.

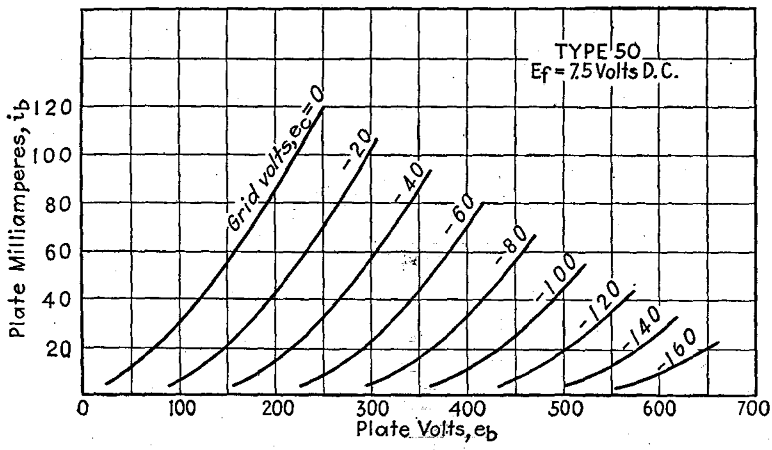


FIG. A-17.—Average plate characteristics for the type 50 triode.

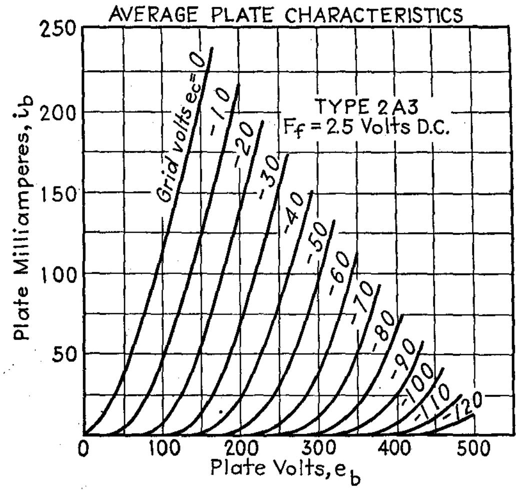


Fig. A-18.—Average plate characteristics for the type 2A3 triode.

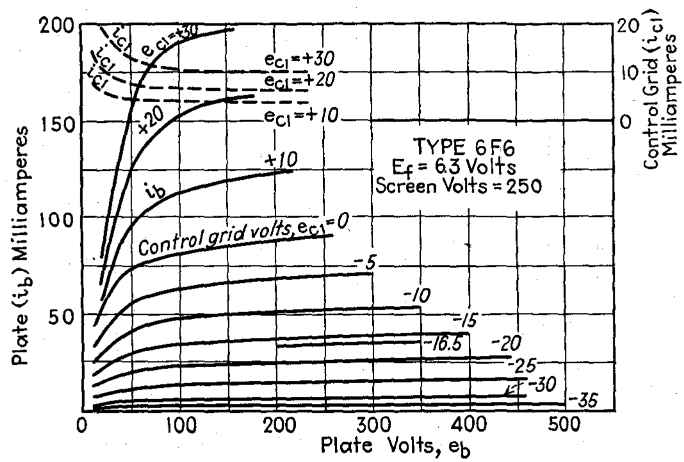


Fig. A-19.—Average plate characteristics for the type 6F6 triple-grid tube, connected as a pentode. See Fig. 3-4 for characteristics for triode connection.

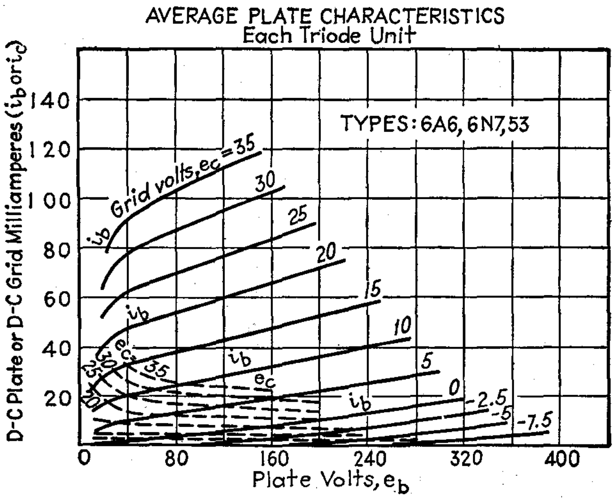


FIG. A-20.—Average plate characteristics for types 6A6, 6N7, and 53 twin triodes.

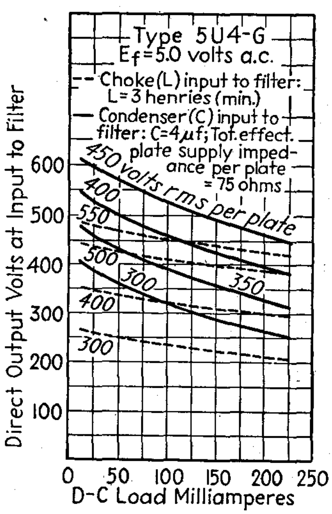


FIG. A-21.—Operation characteristics for the type 5U4-G rectifier.

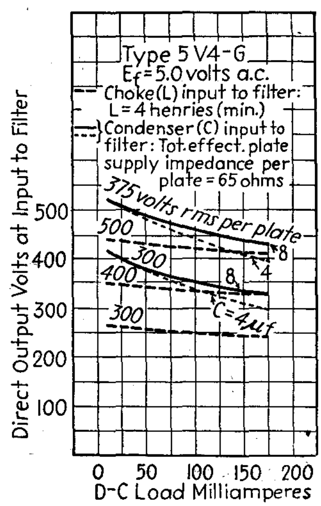


FIG. A-22.—Operation characteristics for the type 5V4-G rectifier.

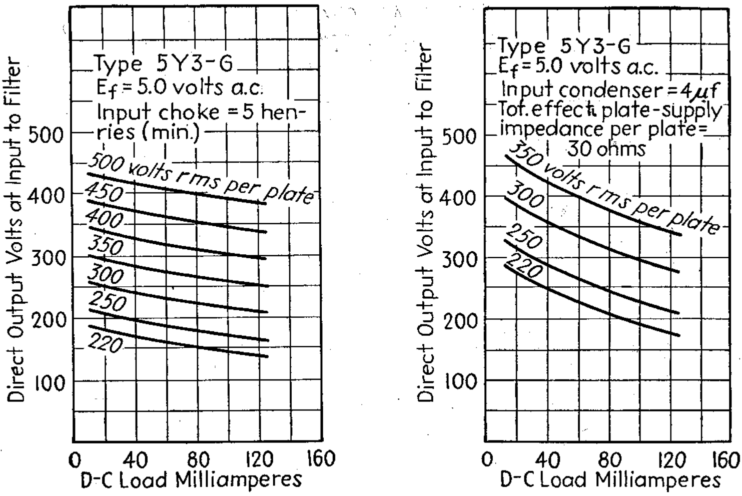


FIG. A-23.—Operation characteristics for the type 5Y3-G rectifier.

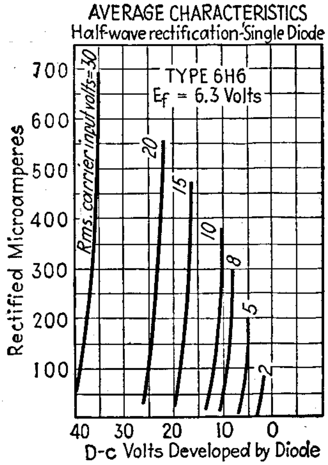


FIG. A-24.—Rectification characteristics for the type 6H6 twin diode.

Operating Data for Typical Amplifier Tubes (Receiving Type)

Type	Classification by construction	Type of cathode	Filament or heater voltage	Filament or heater current, amp	Maximum plate voltage	Maximum screen voltage	Allowable plate dissipation, watts	Allowable screen dissipation, watts	C_{gk},* $\mu\mu f$	C_{gp}, $\mu\mu f$	C_{pk}*, $\mu\mu f$	Connection and class of operation	Static operating voltages and currents: Grid voltage	Screen voltage	Screen current, ma	Plate voltage	Plate current, ma per tube	μ	r_p, ohms	g_m, μmhos	Load resistance, ohms†	P_0 watts‡	Total harmonic content, per cent
2A3	Power triode	Filament	2.5	2.5	250	....	....	....	9	13	4	Class A1, single tube	−45	...	...	250	60	4.2	800	5,250	2,500	3.5	5.0
												Class AB1, two tubes	−62	...	...	300	40	....	....	.....	3,000	15.0	2.5
2A5		Heater	2.5	1.75	375	285	11	3.75	...	......	....	Single Class A1 pentode	−20	285	7.0	285	38	200	78,000	2,560	7,000	4.8	9.0
	Triple-grid power amplifier				350	....	10	....	...	......	....	Single Class A1 triode	−20	...	...	250	31	6.8	2,600	2,600	4,000	0.85	6.5
6F6		Heater	6.3	0.7	375	285	11	3.75	...	......	....	Class AB2 pentodes	−26	250	2.5	375	17	....	....	.....	10,000	18.5	3.5
					350	....	10	....	...	......	....	Class AB2 triodes	−38	...	...	350	24	....	....	.....	6,000	13.0	2.0
6J5	Detector, voltage amplifier triode	Heater	6.3	0.3	300	....	2.5	....	3.4	3.4	3.6	Voltage amplifier	− 8	...	...	250	9	20	7,700	2,600			
6L6	Beam power pentode	Heater	6.3	0.9	360	270	19	2.5	...	......	....	Single, Class A1	−14	250	5.0	250	72	135	22,500	6,000	2,500	6.5	10.0
					360	270	19	2.5	...	......	....	Push-pull, Class A1	−16	250	5.0	250	60	....	....	.....	5,000	14.5	2.0
					360	270	19	2.5	...	......	....	Push-pull, Class AB1	−22.5	270	2.5	360	44	....	....	.....	6,600	26.5	2.0
6SC7	Twin high-mu triode	Heater	6.3	0.3	250	....	....	....	3	2.4	4	Voltage Amplifier	− 2	...	...	250	2	70	53,000	1,325			
6SF5	High-mu triode	Heater	6.3	0.3	250	....	....	....	4	2.4	3.6	Voltage Amplifier	− 2	...	...	250	0.9	100	66,000	1,500			
6SJ7	Triple-grid detector amplifier	Heater	6.3	0.3	300	125	2.5	0.3	6	0.005	7	Pentode voltage amplifier	− 3	100	0.8	250	3	Greater than 1,650	Greater than 1,000,000	1,650			
					250	....	2.5	....	3.4	2.8	11	Triode voltage amplifier	− 8.5	...	...	250	9.2	19	7,600	2,500			
6SK7	Triple-grid supercontrol (variable-mu) amplifier	Heater	6.3	0.3	300	125	4.0	0.4	6	0.005	7	Pentode voltage amplifier	− 3	100	2.4	250	9.2	Approximately 1,600	Approximately 800,000	2,000			
45	Power triode	Filament	2.5	1.5	275	....	....	....	4	7	3	Single Class A1	−50	...	...	250	34	3.5	1,610	2,175	3,900	1.6	5.0
												Push-pull Class AB2	−68	...	...	275	14	....	....	.....	3,200	18.0	5.0
50	Power triode	Filament	7.5	1.25	450	....	....	....	4.2	7.1	3.4	Single Class A1	−84	...	...	450	55	3.8	1,800	2,100	4,350	4.6	5.0
													−54	...	...	300	35	3.8	2,000	1,900	4,600	1.6	5.0

* C_{gk} and C_{pk} are total input and output capacitances (including capacitance to suppressor and screen) in the case of pentodes.
† Plate-to-plate load resistance in push-pull operation. ‡ Power output for two tubes in push-pull operation.

OPERATING DATA FOR TYPICAL POWER THYRATRONS

Type	No. of electrodes	Cathode				Tube drop, volts		Approx. starting characteristics			Deionization time, μsec	Maximum ratings				
		Type	Voltage	Current, amps	Heating time	Max.	Min.	Anode voltage	Shield-grid voltage	Control-grid voltage		Max. peak inverse voltage	Max. peak forward voltage	Amps max. inst. anode current at 25 cycles and above	Amps max. inst. anode current below 25 cycles	Average current, amps
FG-17	3	Fil.	2.5	5.0	5 sec	16		50	....	−0	1,000	5,000	2,500	2.0	1.0	0.5
								100	....	−2.25						
								1,000	....	−6.5						
FG-27-A	3	Fil.	5.0	4.5	1 min	16		70	....	0	1,000	1,000	1,000	10	5.0	2.5
								100	....	−2.25						
								1,000	....	−8.0						
FG-33	3	Heater	5.0	4.5	5 min	24	10	100	....	+10	1,000	1,000	1,000	15	5.0	2.5
								1,000	....	+10						
FG-41	3	Heater	5.0	20	5 min	24	10	250	....	0	100	10,000	10,000	75	25	12.5
								1,000	....	−1.5						
								10,000	....	−4.5						
FG-57	3	Heater	5.0	4.5	5 min	16		60	....	0	1,000	1,000	1,000	15	5.0	2.5
								100	....	−1.75						
								1,000	....	−6.5						
FG-67	3	Heater	5.0	4.5	5 min	24	10	100	....	+3.0	100	1,000	1,000	15	5.0	2.5
								1,000	....	−0.5						
FG-81-A	3	Fil.	2.5	5.0	5 sec	24	8	50	....	0	1,000	500	500	2.0	1.0	0.5
								100	....	−3.75						
								500	....	−5.25						
FG-95	4	Heater	5.0	4.5	5 min	26	10	100	0	+3.5	1,000	1,000	1,000	15	5.0	2.5
								1,000	0	−9.0						
FG-97	4	Fil.	2.5	5.0	5 sec	26	10	100	0	+1.0	1,000	1,000	1,000	2.0	1.0	0.5
								1,000	0	−10.0						
FG-98-A	4	Fil.	2.5	5.0	5 sec	26	8	60	0	0	1,000	500	500	2.0	1.0	0.5
								500	0	−10.0						

OPERATING DATA FOR TYPICAL POWER THYRATRONS.—(*Continued*)

Type	No. of electrodes	Cathode: Type	Cathode: Voltage	Cathode: Current, amps	Cathode: Heating time	Tube drop, volts: Max.	Tube drop, volts: Min.	Approx. starting characteristic: Anode voltage	Approx. starting characteristic: Shield-grid voltage	Approx. starting characteristic: Control-grid voltage	Deionization time, μsec	Maximum ratings: Max. peak inverse voltage	Maximum ratings: Max. peak forward voltage	Maximum ratings: Amps max. inst. anode current at 25 cycles and above	Maximum ratings: Amps max. inst. anode current below 25 cycles	Maximum ratings: Average current, amps
FG-105	4	Heater	5.0	10.0	5 min	20	10	100	0	+1.0	1,000	2,500	2,500	40	12.8	6.4
FG-154	4	Fil.	5.0	7.0	10 sec	30	10	60	0	0	1,000	500	500	10.0	5.0	2.5
								100	0	−3.5						
FG-172	4	Heater	5.0	10.0	5 min	24	6	100	0	+1.0	1,000	1,000	1,000	40	13.0	6.4
								1,000	0	−9.0						
FG-178-A	3	Fil.	2.5	2.25	6 sec	24	8	25	....	0	1,000	500	500	500	250	125
								100	....	−3.0						
								500	....	−7.0						
GL-414	4	Heater	5.0	20.0	10 min	24	10	100	0	−1.0	1,000	2,000	,2000	100	25	12.5
								2,000	0	−14.0						
KU-627	3	Fil.	2.5	6	10 sec	12		100	....	−2	1,000	2,500	1,250	2.5		0.64
								1,000	....	−8						
KU-628	3	Fil.	5	11.5	40 sec	12		100	....	−2	1,000	2,500	1,250	8		2
								1,000	....	−8						
WL-672	4	Heater	5.0	6.0	5 min	12		1,000	+10	−18	1,000	1,500	1,500	30		2.5
								100	+10	−3						
								1,000	0	−10						
								100	0	+3						
								1,000	−5	+12						
								100	−5	+21						
								1,000	−30	+170						
								100	−30	+175						
WL-677	3	Heater	5	9.5	5 min	12		1,000	....	−6	1,000	10,000	10,000	15		4
								500	....	−30						

Operating Data for Typical Ignitrons

	FG-235-A	FG-238-B	FG-258-A	FG-259-B	FG-271	GL-415	GL-427	WL-651/656	WL-652/657	WL-655/658	WL-681/686
Tube drop	8–20	12½–19	8–20	12½–20	8–20	8–20	10–15	8–20	8–20	8–20	8–20
Maximum demand, kva	1,200	2,400	2,400	1,200	600	105	.....	1,200 400	600 200	2,400 800	300 100
Corresponding average anode current, amps	75.6	135	192	75	30.2	3	5	75.6 140	30.2 56	192 355	12.1 22.4
Max. surge peak anode current at 250 v.	......	......	......	......	.....	.....	.....	13,500	6,750	27,000	3,360
Max. surge peak anode current at 500 v.	......	......	......	......	.....	.....	.....	6,750	3,375	13,500	1,680
Ignitor voltage, max. inst. pos. allowed	900	anode volt.	900	anode volt.	900	900	350	900	900	900	900
Ignitor voltage, max. inst. pos. required	200	150	200	150	200	200	150	200	200	200	200
Ignitor voltage, max. inst. neg. allowed	5	5	5	5	5	5	5	5	5	5	5
Ignitor current, max. inst. allowed, amps	100	100	100	100	100	100	100	100	100	100	100
Ignitor current, max. inst. required, amps	40	40	40	40	40	40	30	40	30	40	40
Ignitor current, average allowed	1	2.0	1	2.0	1	1	2	1	1	1	1
Ignition time, μsec	100	100	100	100	100	100	100	100	100	100	100

AUTHOR INDEX

O

P

Q

R

S

T

SUBJECT INDEX

Q

R

S

SYMBOL INDEX

The following index lists the pages on which each symbol, used with a particular connotation, is defined. It does not include most of the special symbols for electrode voltages and currents listed on pages 70 and 71, some symbols used only in the course of derivations or in a single figure, or well-known standard symbols such as π or the vector operator j.

ANSWERS TO PROBLEMS

1-1. (*a*) 4.24×10^{-10} sec. (*b*) 4×10^{-10} erg.

1-2. (*a*) Electron returns to first electrode. It moves to a maximum distance of 0.237 cm from the first electrode. At the instant of field reversal it is 0.118 cm from the first electrode. (*b*) The energy acquired before field reversal is returned to the source of applied potential. That acquired during the return to the first electrode is delivered to the cathode and is converted principally into heat.

1-3. (*a*) The energy lost by the electron in moving to the second electrode against the field is 6.4×10^{-10} erg. As this is less than the initial energy, the electron will reach the second electrode. (*b*) 6.4×10^{-10} erg is delivered to the source of applied potential. The remainder is delivered to the second electrode and is converted mainly into heat.

1-4. 192.7 volts.

1-5. 11.38 gausses.

3-1.[1] $r_p = 11{,}600\ \Omega$; $\mu = 19.5$; $g_m = 1680\ \mu$mhos.

3-2.[1] $g_m = 1950\ \mu$mhos (2000 μmhos from slope of transfer characteristic); $r_p = 350{,}000\ \Omega$; $\mu = 680$.

3-3.[1] (*a*) $\mu \cong 8$; $g_m \cong 1700\ \mu$mhos; $r_p = \mu/g_m \cong 4700\ \Omega$.
(*b*) $\mu \cong 15$; $r_p \cong 10{,}000\ \Omega$; $g_m = \mu/r_p \cong 1500\ \mu$mhos.

3-4. 60, 100, 900; 120, 200, 1800; 180, 300, 2700; 160, 960, 1000; 40, 800, 840; 740, 860, 940, 1060; 220, 260, 1020, 1100, 1860, 1900; 20, 140, 700, 780, 1700, 1740.

3-8.[1] (*a*) At $e_b = 250$ *v.* and $E_c = -8v.$, $\mu = 3.5$; $r_p = 1550\ \Omega$; $\partial r_p/\partial e_b = -16$ Ω/volt; $a_1 = 2.5$ ma/volt; $a_2 = 4.07 \times 10^{-2}$ ma/(volt)^2.

(*b*) $a_1 = 1.25$ ma/volt; $a_2 = 5.1 \times 10^{-3}$ ma/(volt)^2.

(*c*) $a_1 = 0.83$ ma/volt; $a_2 = 1.5 \times 10^{-3}$ ma/(volt)^2.

4-7.[1] (*a*) $\mu = 3.5$; $r_p = 1785\ \Omega$; $g_m = 1960\ \mu$mhos.

(*b*) $R_b = 425\ \Omega$.

(*d*) For a 50-volt grid swing, $H_1 = 24.8$ ma and $H_2 = 1.88$ ma. For a 40-volt grid swing, $H_1 = 20$ ma and $H_2 = 1.25$ ma.

(*e*) $P_o = 1.54$ watts and 1.0 watt.

5-2. Voltage gain = 40 db; current gain = 20 db; power gain = 30 db.

6-2. (*c*) $\mu = 19$ and $r_p = 12{,}800\ \Omega$ (determined from the plate diagram[1]); $A_1 =$ 6.42 at 60 cps, 9.15 at 100 cps, and 14.52 at 1000 cps; $A_2 = 15.1$; over-all gain = 39.7 db at 60 cps, 42.8 db at 100 cps, and 46.8 db at 1000 cps.

(*d*) $A_1 = 19.85$ at 60 cps, 27.7 at 100 cps, and 40.3 at 1000 cps; $A_2 = 48.4$; over-all gain = 59.6 db at 60 cps, 62.6 db at 100 cps, and 65.8 db at 1000 cps.

(*f*) $I_{bo} = 2$ ma, $R_{cc} = 3000\ \Omega$.

6-3. Assume that the operating plate voltage is 250 volts and the bias -2 volts (in practice the plate *supply* voltage would probably be 250 volts and the bias about -1.5 volts). Then $\mu = 100$ and $r_p = 66{,}000\ \Omega$.

Second stage: Use $r_o = 500{,}000\ \Omega$. Then $r_h = 58{,}300\ \Omega$. At 10,000 cps, $\omega r_h C_2 =$ 0.0183 and $A = A_m = 87.4$.

[1] For accurate results it is recommended that the student use characteristic curves contained in the RCA Receiving Tube Handbook, or a similar handbook in which the curves are plotted on closely ruled graph paper.

First stage: $C_2 = 222\ \mu\mu$f. If $r_1 = 50{,}000\ \Omega$ and $r_2 = 250{,}000\ \Omega$, $r_h = 25{,}600\ \Omega$; then $A_m = 38.4$, $\omega r_h C_2 = 0.356$ at 10,000 cps; and $A = 0.93A_m = 35.7$. $r_l = 278{,}000\ \Omega$. Make $\omega r_l C_c = 10$ at 100 cps to make $A = A_m$ at this frequency. Then $C_c = 0.057\ \mu$f. Use $C_c = 0.06\mu$f. If r_1 is decreased to 25,000 Ω and $r_2 = 250{,}000\ \Omega$, $r_h = 16{,}900\ \Omega$; then $A_m = 25.4$, $\omega r_h C_2 = 0.236$ at 10,000 cps, and $A = 0.97A_m = 24.6$. $r_l = 268{,}000\ \Omega$. If $C_c = 0.06\ \mu$f, $\omega r_l C_c = 10$ at 100 cps, and $A = A_m = 25.4$ at this frequency.

6-4. A of second stage is constant at 15.1 below 100,000 cps; A of first stage is 14.23 at 50,000 cps and 13.42 at 100,000 cps.

6-5.[1] Crest value of fundamental component of output voltage $= (210 - 40)/2 = 85$ volts; $A = 85/5.5 = 15.47$; per cent $H_2 = (250 - 264)/(2 \times 170) = 4.12$.

6-6.[1] *(a)* Crest value of fundamental component of output voltage $= (208 - 90)/2 = 59$ volts; $A = 59/0.75 = 78.6$; per cent $H_2 = 1.0$. *(b)* $r_b = r_1r_2/(r_1 + r_2) = 333{,}333\ \Omega$; $R_b = r_1 = 500{,}000\ \Omega$; crest fundamental output voltage $= 53$ volts; $A = 70.6$; per cent $H_2 = 1.7$.

6-7.[1] *(a)* Crest value of fundamental output voltage $= (188 - 38)/2 = 75$ volts; $A = 75$. *(b)* $r_p = 208{,}000\ \Omega$; $\mu = 96$; $A = 79.5$. *(c)* $I_{bc} = 0.125$ ma; $R_{cc} = 12{,}000\ \Omega$. *(d)* $A = 40.5$.

6-9.

Frequency	A	Frequency	A	Frequency	A
20	21	550	60	10,000	102
50	40	1000	60.1	12,600	109
100	52	2000	62.5	16,500	60
200	57.5	5000	65	20,000	36
400	59.7	7000	78		

6-10. $k = 0.034$.

7-3.[1] *(a)* Operating point is determined by allowable plate dissipation. $E_c = -63$ volts. *(b)* Load line intersects the voltage axis at 490 volts. *(c)* Opt. $P_o = 2.48$ watts. *(d)* $r_b = 5780\ \Omega$. *(e)* $r_p = 1500\ \Omega$. *(f)* $I_p = 20.8$ ma; $P_o = 2.5$ watts. *(g)* $P_o = 2.47$ watts. *(h)* $P_i = 9.99$ watts; $\eta_p = 24.8$ per cent. *(i)* Power sensitivity $= 3.13 \times 10^{-4}$ mho. *(j)* $n = 12{:}1$.

7-4.[1] The following values are found by the use of Eq. (7-46): $r_p = 2000\ \Omega$; $I_{bo} = 18$ ma; $E_{bo} = 278$ volts; $E_c \cong -65$ volts; Opt. $P_o = 1.34$ watts; per cent $H_2 = 4.5$; r_b (determined from the dynamic load line) $= 13{,}300\ \Omega$. (Note that the allowable plate dissipation was made low in order that the problem could be solved by the use of plate characteristics normally available.)

7-6.[1] $r_{bb} = 2360\ \Omega$; per cent $H_3 = 2.75$; per cent $H_5 = 0.4$; $P_o = 2.24$ watts; $\eta_p = 44.5$ per cent; $n = 6.89{:}1$.

9-8.[1] *(a)* Fundamental modulation-frequency output is 15.0 volts; second-harmonic output is negligible. *(b)* Fundamental modulation-frequency output is 19 volts; second-harmonic output is negligible.

9-9.[1] *(a)* Fundamental modulation-frequency output is 14.0 volts; second-harmonic output is 0.75 volt (5.35 per cent). *(b)* Fundamental modulation-frequency output is 16.4 volts; second-harmonic output is 1.6 volts (9.75 per cent).

10-3. *(a)* $e_{c3} = -10$ volts. *(b)* $\Delta e_{c2}/\Delta e_{c2} = \frac{1}{2}$. *(c)* $\rho_o = -6900\ \Omega$. *(d)* $C = 1.45\ \mu$f; $L = 1.75h$.

[1] See footnote, p. 713.

6-11.[1] A type 2A3 triode operated at $E_{bo} = 250$ volts and $E_c = -45$ volts will deliver the 50 ma crest current. Crest fundamental alternating current of $(120 - 15)/2 = 52.5$ ma with 5 per cent second harmonic is obtained with a load resistance of 2500 Ω and a grid bias of -43.5 volts (-45 volts relative to mid-point of filament). $E_{gm2} = 43$ volts for 50 ma crest current. Required voltage amplification of first stage $= \frac{43}{1} = 43$. For the first stage use 6SF6 tube with 500,000-Ω plate-coupling resistor. For $E_{bb} = 250$ volts and $E_c = -1.5$ volts, graphically determined $A_1 = 71$. Input voltage to amplifier for full deflection = 0.605 volt. Use direct coupling in first stage and 2500-Ω resistor in series with oscillograph element. The circuit is shown in Fig. 6-55. Circuit of the form of Fig. 5-5 could also be used, but degeneration caused by second-stage plate current in the voltage divider or cathode resistor would reduce the sensitivity.

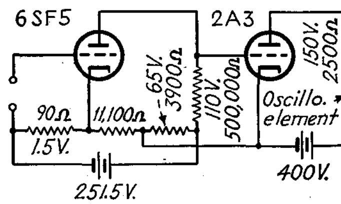


FIG. 6-55.—Solution for Prob. 6-11.

6-12.[1] (a) 2A3 triode or 6L6 pentode. (b) For 6L6 pentode operated on 200-volt supply with 2000-Ω load (field), plate current is 50 ma at $e_c = -16$ volts and 60 ma at -13.5 volts (determined from dynamic transfer characteristic). (c and d) Required voltage amplification for first stage $= 2.5/0.05 = 50$. Use the circuit of Fig. 6-37 without the exciter. Use a 6SF5 tube in the first stage; $E_{ob} = 100$ volts, $E_c = -1.0$ volt, and 200,000-Ω coupling resistor. A of first stage (determined graphically) is 50. With 500,000-Ω coupling resistor, A of first stage is 62.

7-2.[1] (a) $\mu = 3.6$ by graphical determination. Opt. $E_c = -0.7 \times 200/3.6 = -39$ volts. Use $E_c = -40$ volts. Zero-signal plate dissipation $= 200 \times 21 \times 10^{-3} = 4.2$ watts, which is within the allowable value. (b) Load line for 5 per cent second harmonic intersects the voltage axis at 328 volts, and the current axis at 54 ma. (c) $P_o = 0.94$ watt. (d) Opt. $r_b = 6100$ Ω. (e) $r_p = 1780$ Ω.

From the following graphically determined values it can be seen that some increase in power output can be obtained, at the expense of plate circuit efficiency, by changing the grid bias from -40 volts to -35 volts:

E_c volts	Zero-signal P_p, watts	P_o at 5 per cent H_2, watts	r_b ohms	r_p ohms	r_b/r_p	η_p per cent
−40	4.2	0.93	6100	1790	3.4	22.1
−37.5	5.2	0.97	4850	1600	3.03	18.6
−35	6.3	1.05	3540	1500	2.36	16.7
−32.5	7.5	1.04	2640	1400	1.89	13.9
−30	8.8	1.03	1710	1300	1.32	11.7

(f) $I_p = 0.707\mu E_c/(r_p + r_b) = 17.7$ ma for $E_c = -35$ volts; $P_o = I_p^2 r_b = 1.11$ watts. (g) $P_o = 1.17$ watts; (h) $P_i = 6.3$ watts; plate-circuit efficiency is listed in the foregoing table. (i) P. S. $= 4.3 \times 10^{-4}$ mho. (j) $n = 9.4$ for $r_b = 3540$ Ω.

13-1. Wrong type of amplifier tube; crest phototube voltage too high for gas phototube; coupling resistance too low; relay should be shunted by a condenser; circuit should include bias adjustment; polarity of phototube is incorrect (coupling resistor should be shunted by a condenser for best results)

[1] See footnote, p. 713.

13-2. (*a*) See Fig. 13-83. (*b*) See Fig. 13-83. (*c*) For 72-ft-candle illumination, phototube flux = 0.5 lumen. Phototube current = 3.5 μa. Voltage drop across coupling resistance = 28 volts. Bias caused by coupling-resistance drop = -28 volts. For E_{bb} = 100 volts and i_b = 7 ma, $e_c = -1$ volt (determined from transfer characteristic for 1000-Ω load). Required applied biasing voltage = $+27$ volts. (*d*) For E_{bb} = 100 volts and i_b = 8 ma, $e_c = -0.5$ volt. Δe_c = 0.5 volt. Δi = 0.0625 μa. ΔL = 0.00893 lumen. Change in illumination = 1.285 ft-candles.

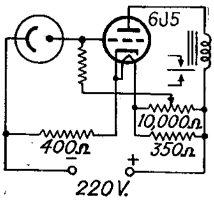


FIG. 13-83.—Solution for Prob. 13-2.

13-3. (*a*) Use the circuit of Fig. 13-11. (*b*) Relay closes at $e_c = -1$ volt; opens at $e_c = -2$ volts. (*d*) i = 9.4 μa when L = 0.2 lumen. $E_r = +9.4$ volts. $E_{cc} = -11.4$ volts. (*e*) To close relay, E_r = 10.4 volts, i = 10.4 μa, L = 0.22 lumen, ΔL = 0.02 lumen. Change in illumination required to close relay = 3.2 ft-candles. A common voltage supply may be used for anode and plate. C supply voltage causes the anode voltage to be one or two volts greater than 90 volts, but this will have negligible effect upon the computations.

10-5. 0.83 by planimeter; 0.85 by selected ordinates.

14-1. A 25Z6 tube may be used. $\hat{r}_p$ = 138 Ω; Assume R_s = 5 Ω; R = 14,000 Ω; $\hat{R}_s/R$ = 0.0102; ωRC = 105; $C \cong 2$ μf; $\bar{R}_s/R$ = 0.0116; E_{dc}/E_m = 1.7; E_{dc} = 288 volts; $\hat{I}_p$ = 160 ma; peak inverse voltage at no load = 338 volts.

14-2. Use full-wave, single-phase rectifier with 5Y3G tube. For two-stage L-section filter, required LC = 64.3. For 20,000-Ω bleeder, $R_{\text{min.}}$ = 2600 Ω, and L_1 should exceed 5.2 henrys at full load. L_1 should exceed 20 henrys at minimum load. Use 25-henry choke. If L_1 drops to 12.5 henrys at full load, required condenser capacitance per stage is 5.15 μf. Use 6 μf. Resonant frequency is 18.4 cps at full load. Voltage drop in chokes at full load is approximately 17 volts. Required secondary voltage is 425 volts each side of center. Effective output impedance of the filter at 50 cps is 560 Ω. Terminal voltage rises to about 365 volts at minimum load.

14-3. A 5Y3-GT tube may be used. r_p = 344 Ω; assume R_s = 125 Ω; Assume two chokes will be used having a resistance of 75 Ω each; R = 4150 Ω; Use C_1 = 4 μf; ωRC = 6.25; $\hat{r}_p$ = 320 Ω; $\hat{R}_s/R$ = 0.113; ρ_1 = 0.09; $\bar{R}_s/R$ = 0.124; E_{dc}/E_m = 0.7; E_{rms} = .707(400 + 15)/.7 = 420 volts each side of center. Remainder of solution is similar to that of Prob. 14-2.

14-4. The peak inverse anode voltage will be somewhat greater than 4700 volts. A type 866A/866 mercury-vapor rectifier tube may be used (see transmitting tube manual for tube data). Use a 50,000-Ω bleeder and two chokes, having an estimated resistance of 75 Ω each. α_1 = 472; $LC = 40 \times 10^{-6}$ henrys $\times$ farads; R_t = 5520 Ω at full load; L_1 should exceed 11 henrys at full load; R_t = 18,950 Ω at minimum load; L_1 should exceed 18.9 henrys at minimum load. Use chokes having an inductance of 25 henrys at no load. If the inductance falls to half this value at full load, the required value of C is 3.2 μf. Use 4-μf condensers. Resonance frequency = 22.5 cps at full load. For a tube drop of 15 volts, the required full-load secondary voltage per anode is 1730 volts.

[1] See footnote, p. 713.