McGraw-Hill Electrical and Electronic Engineering Series

FREDERICK EMMONS TERMAN, *Consulting Editor*

ENGINEERING ELECTRONICS

McGraw-Hill Electrical and Electronic Engineering Series

FREDERICK EMMONS TERMAN, *Consulting Editor*

BAILEY AND GAULT · Alternating-current Machinery
BERANEK · Acoustics
BRUNS AND SAUNDERS · Analysis of Feedback Control Systems
CAGE · Theory and Application of Industrial Electronics
CUCCIA · Harmonics, Sidebands, and Transients on Communication Engineering
EASTMAN · Fundamentals of Vacuum Tubes
EVANS · Control-system Dynamics
FITZGERALD AND HIGGINBOTHAM · Basic Electrical Engineering
FITZGERALD AND KINGSLEY · Electric Machinery
GEPPERT · Basic Electron Tubes
GLASFORD · Fundamentals of Television Engineering
HAPPELL AND HESSELBERTH · Engineering Electronics
HARMAN · Fundamentals of Electronic Motion
HESSLER AND CAREY · Fundamentals of Electrical Engineering
HILL · Electronics in Engineering
JOHNSON · Transmission Lines and Networks
KRAUS · Antennas
KRAUS · Electromagnetics
LEPAGE · Analysis of Alternating-current Circuits
LEPAGE AND SEELY · General Network Analysis
MILLMAN AND SEELY · Electronics
MILLMAN AND TAUB · Pulse and Digital Circuits
ROGERS · Introduction to Electric Fields
RÜDENBERG · Transient Performance of Electric Power Systems
SEELY · Electronic Engineering
SEELY · Electron-tube Circuits
SEELY · Radio Electronics
SISKIND · Direct-current Machinery
SKILLING · Electric Transmission Lines
SKILLING · Transient Electric Currents
SPANGENBERG · Fundamentals of Electron Devices
SPANGENBERG · Vacuum Tubes
STEVENSON · Elements of Power System Analysis
STORER · Passive Network Systems
TERMAN · Electronic and Radio Engineering
TERMAN AND PETTIT · Electronic Measurements
THALER · Elements of Servomechanism Theory
THALER AND BROWN · Servomechanism Analysis
THOMPSON · Alternating-current and Transient Circuit Analysis
TRUXAL · Automatic Feedback Control System Synthesis

ENGINEERING ELECTRONICS

GEORGE E. HAPPELL
Associate Professor of Electrical Engineering
Purdue University

WILFRED M. HESSELBERTH
Associate Professor of Electrical Engineering
Purdue University

NEW YORK TORONTO LONDON
McGRAW-HILL BOOK COMPANY, INC.
1953

ENGINEERING ELECTRONICS

Library of Congress Catalog Card Number: 53-5166

V

THE MAPLE PRESS COMPANY, YORK, PA.

PREFACE

The authors present herewith a textbook for use in a beginning course in electronics for electrical-engineering students. Most of the material used was first published in planograph form and has been used as a text at Purdue University for the past two years. It has been revised and brought up to date as use in the classroom and the advice of critics have indicated that improvements could be made.

The usual course in technical schools consists of two or three class periods and one laboratory period each week throughout the school year. Sufficient material is included for such a program. It has been the authors' experience that the average student enrolling in such a course has the following status: He is a junior and has already had courses in general physics, mathematics through calculus, and direct-current circuits. He is starting courses in alternating-current circuits, electrical measurements, and possibly differential equations, as well as electronics. The student plans to enter one of various fields—communications, electronic control, servomechanisms, power machinery, power transmission, business, graduate study. What he will actually do after graduation is often something else. In any event electronics will be useful knowledge for one engaged in nearly any branch of electrical engineering and in many allied fields. The material in this book presents the fundamental ideas of electronics in both a theoretical and a practical fashion to provide a good foundation for further study, as well as useful knowledge for a terminal course.

The first four chapters provide material for a brief study of the physics of vacuum tubes, not covered in the usual previous physics courses. They also serve to delay the study of circuits until the student has gained some knowledge of a-c circuits elsewhere. Chapter 5 presents a very elementary description of the circuits and actions of certain very common electronic devices. It also acquaints the student with some common electronic nomenclature. The authors have found it fills a very real need —to provide a background for those students who have not picked it up in their experience. Even with very rapid coverage it should be valuable.

Chapter 6 presents the usual methods employed in electronic-circuit analyses, analytic and graphical. Great stress is laid on the use of the linear-equivalent-circuit theorem. Also considerable attention is paid to graphical methods with nonlinear circuits. Only elementary aspects of this fascinating subject are presented because of the limitations of time and space.

Although in theory a student should have well in hand all the tools he has studied, as a practical matter the authors feel that a brief restatement or treatment of certain ideas often helps enough to pay for its inclusion in a volume designed principally as a textbook. Hence the short treatments of such subjects as network theorems and Fourier analysis are included. The practical use of this material begins at once in the following chapter, although for some of it the delay is great enough, as in the case of power-series expansion of plate current, so that the student will wish to refer back to the discussion again. At any rate he knows where to find the material.

The chapters following the sixth present a selection of the various aspects of electronics which can reasonably be included in a beginning course. No claim is made that all the interesting and useful developments in the field are discussed or even mentioned.

In the numerous cases in which a mathematical development is attempted, the authors have endeavored to provide, first, a facile word explanation for the behavior. Then follows the sensing of current and voltage symbols, the writing of circuit equations, the solution of these, the simplifying assumptions and rearrangement needed to place the solution in a usable form (which often involves the drawing of a simplified equivalent circuit), and the final interpretation of the results. Numerous worked-out examples are provided to help in understanding. The authors feel that much practice is needed in these matters for students who plan to continue in fields allied to electronics.

The authors have freely consulted periodicals and engineering texts by many writers. They wish to acknowledge the valuable criticism and encouragement given by their colleagues. Especially do they appreciate the assistance of their former colleague, Dr. K. J. Hammerle. In addition, thanks are due the unknown critics engaged by the publisher. They have made many valuable suggestions.

George E. Happell
Wilfred M. Hesselberth

CONTENTS

CHAPTER 1

ELECTRON BALLISTICS

1-1. Introduction. The science of electronics can be considered as beginning with the discovery of the "Edison effect." In 1883 Thomas A. Edison was endeavoring to determine why his incandescent lamps were burning out at the positive end of the filament. (Only direct current was used at the time, as alternating current did not come into common use until a much later date.) In the course of his experiments he sealed a second electrode into some of his experimental lamps and found that, if this second electrode was connected through a galvanometer to the positive end of the filament, current would flow. When the electrode was connected to the negative end of the filament, no current would flow. This effect was the first actual proof that current was the flow of negative charge. Edison, however, was unable to explain it, and since he had more urgent work at hand, he did not continue the investigation.

In 1897 Sir J. J. Thomson of the Cavendish Laboratories definitely established the existence of the electron and explained the Edison effect on this basis. He proved that the electron was a negatively charged particle and that it was the motion of this particle under the influence of electric fields that constituted current flow. He also established the fact that the ratio of charge to mass of the electron was constant for all electrons.

In 1910 Dr. Robert Millikan measured the electrostatic charge on the electron and found it to be constant for all electrons. This result, combined with Thomson's determination of the charge-to-mass ratio, allowed the mass of the electron to be calculated. It also proved that all electrons are identical.

Sir J. A. Fleming, another Englishman, was the first to put the Edison effect to work. In 1904 he patented the first electronic tube, which was called the Fleming valve. This was a simple diode used as a rectifier. However, since alternating current was not in common use at that time, the tube was of little practical value.

Probably the greatest early contribution to the electronic art was made in 1906 by Dr. Lee De Forest, who added a third element, which he called a grid, to the diode. By use of this grid he was able to control relatively large electron currents in the anode circuit by means of varying potentials applied to the grid and with the expenditure of very little power in the grid circuit. Since that time many improvements have been made, so that today we have not only diodes and triodes, but also tetrodes, pentodes, heptodes, and others, each type being named in accordance with the total number of electrodes it possesses. The electronic tube is no longer a laboratory plaything but has advanced to the stage where it is the basis of an entire industry which is rapidly increasing in magnitude.

New and very promising additions to the electronic art have recently become prominent with the development of such devices as improved crystal diodes, thermistors, transistors, and magnetic amplifiers. These devices may some day become as important to the engineer as the vacuum tube is today.

1-2. The Electron. In the late nineteenth century the atom was thought to be the smallest indivisible particle of matter. Today it is known that the atom is made up of a relatively large mass, called the nucleus, about which electrons are revolving in orbits. The electrons have negative charges, and the nucleus has a positive charge. The total positive charge on the nucleus is equal to the sum of the negative charges of all the electrons revolving about that nucleus.

Millikan's experiments proved that the electron is the smallest indivisible unit of electric charge. It exhibits the properties of a particle or corpuscle as well as those of a wave motion. In some instances it is necessary to use the wave properties to explain experimental results. At other times it acts as if it were a charged particle.

In considering the electron as a particle it is usually imagined to be spherical in shape. This concept may not be true, since as yet no experiment has been devised which will give us a picture of the electron. The charge of the electron has been measured as 1.60×10^{-19} coulomb. The charge-to-mass ratio for an electron at rest has been determined as 1.76×10^{11} coulombs per kg. From these two figures the rest mass can be calculated and is found to be 9.11×10^{-31} kg. When in motion the mass of the electron increases with increase in velocity. It has been shown that $m_e = m_r/\sqrt{1 - (v^2/c^2)}$, where m_r is the rest mass of the electron, v is the velocity of the electron, and c is the speed of light. However, for a velocity as great as one-fifth the speed of light, the error in using the rest mass is only 2 per cent. For most applications in this book, this error is negligible. The diameter of the electron has been computed as about 3.8×10^{-15} m.

1-3. Other Charged Particles.[1*] The nucleus of the atom is made up of one or more protons and in some cases one or more neutrons and positrons. The proton has a charge numerically equal to that of the electron but is positive in polarity. The mass of the proton is 1849 times the mass of the electron. An atom of hydrogen, which is the simplest atom found,

* Superior numbers are used to refer to supplementary reference material listed at the end of the chapter.

is made up of one proton for a nucleus and one electron revolving in an orbit about it. The neutron has a mass about 0.08 per cent larger than that of the proton but has no electrical charge. The positron, or "positive electron," has a positive charge numerically equal to the charge of the electron, and it also has the same mass as the electron. It occurs very rarely in the free state and then only as the result of atomic disintegration.

Particles having the mass of a molecule but a charge numerically equal to a multiple of that of the electron are known as ions. If an atom in a molecule has lost one or more electrons, the net charge is positive and the particle is known as a positive ion. If the molecule has an extra electron attached to it, it is known as a negative ion. In electronics these particles are usually encountered in gaseous conduction although they do occur in high-vacuum tubes since it is impossible to remove all molecules of gas from a tube. Ions are usually formed by collisions between electrons and molecules—a collision may remove an electron from its orbit about the nucleus to form a positive ion and a free electron.

1-4. Units.[2] The rationalized mks system has become fairly common in the engineering field, and since it has some distinct advantages, it will be used in this book. This, of course, means that all electrical quantities will be expressed in the common or practical units—amperes, volts, coulombs, watts, ohms, and so forth. In the case of the mechanical units the use of mass in kilograms and distance in meters results in a new unit of force. It has been named the newton and is equal to 10^5 dynes. It was necessary to invent this new unit in order to be able to use Newton's second law, $f = ma$, and have mass in kilograms and acceleration in meters per second per second. Energy is measured in joules, which again is the practical unit.

1-5. Properties of Charged Bodies. Electrons obey the same physical laws as do any other bodies. For instance, Newton's second law states that the rate of change of momentum of a body is proportional to the applied force and occurs in the same direction as the force acts. Mathematically this can be stated as

$$f = \frac{d(mv)}{dt} \tag{1-1}$$

but if the mass is constant, this becomes

$$f = m\frac{dv}{dt} = ma \tag{1-2}$$

Gravitational force acts on the electron just as it does on larger bodies. In this case the acceleration will be the same as it is for all bodies falling in free space close to the earth, 32 ft per sec², which is equal to 9.8 m per sec². This is not a very high rate of acceleration. The force acting on the electron can be calculated and is found to be

$$f_g = ma = 9.11 \times 10^{-31} \text{ kg} \times 9.8 \text{ m per sec}^2 = 89.3 \times 10^{-31} \text{ newton} \quad (1\text{-}3)$$

In addition to being acted upon by gravity the electron is affected by neighboring charged bodies. Like charges repel, and unlike charges attract. Coulomb's law states that the force between two charged bodies varies as the product of their charges and inversely as the square of the distance between them, *i.e.*,

$$f = \frac{Q_1 Q_2}{4\pi\epsilon_0 d^2} \quad (1\text{-}4)$$

where ϵ_0 is the permittivity of free space. In the cgs system $4\pi\epsilon_0$ is equal to unity, but in the mks system $4\pi\epsilon_0$ is equal to $1/(9 \times 10^9)$ farads per m.

If the force between two electrons spaced 1 cm apart in a vacuum is calculated from Coulomb's law, it is found to be

$$f_e = \frac{(1.6 \times 10^{-19})^2}{[1/(9 \times 10^9)] \times (0.01)^2} = 2.3 \times 10^{-24} \text{ newton} \quad (1\text{-}5)$$

This is a very small force when compared with forces with which we commonly work. But when compared with that due to gravity, it is seen that it is a very large force, the ratio being

$$\frac{f_e}{f_g} = \frac{2.3 \times 10^{-24}}{89.3 \times 10^{-31}} = 2.57 \times 10^5 \quad (1\text{-}6)$$

In other words, the force between two electrons 1 cm apart, which is an extremely large distance when working with electrons, is 257,000 times as great as the force due to gravity. Therefore the effects of gravity will be neglected in the remainder of this book.

1-6. The Electric Field.[3,4] In order to understand the action of electrons in electronic tubes it is necessary to have some concept of an electric field. The presence of an electric field in space can be detected by bringing a charged body into that region and noting whether there is any force exerted on the body. If a field is present, it is due to other charged bodies and hence a force is exerted on the test body. All definitions are set up on the idea of using a test charge which is positive in polarity. Therefore, if the positive test charge has a force exerted on it, it seems logical that the strength of the field be specified as having the same direction as the force. However, this method of measuring the field strength has one objection—the force exerted on the test charge is dependent upon the magnitude of the test charge. Hence field strength, or field intensity, is defined as a vector equal in magnitude to the force per *unit* test charge (positive) and having the same direction as the force. Unit charge in the mks system may be defined as that amount of charge which, if placed 1 m from a like charge in free space, repels it with a force of 9×10^9 newtons. This unit of charge is called a coulomb. $\boldsymbol{\varepsilon}$ is used as a symbol* for field intensity, and

$$\boldsymbol{\varepsilon} = \frac{\mathbf{f}}{Q} \tag{1-7}$$

A complete knowledge of an electric field involves the determination of the magnitude and direction of the field-intensity vector for all points in the space. If one supposes a small positive test charge to be placed at an infinite number of points in an electric field and a small arrow drawn indicating the direction of the force at each point, it may be noted that these arrows tend to line up and form continuous curves. This is the same process that one would use in drawing flux lines in a magnetic field using a small compass. Thus the direction of the field-intensity vector can be mapped, and it remains only to specify the magnitude of the field at all points. A special but important type of field is one in which the force, or field-intensity, lines are all parallel straight lines and in which the force is constant at all points. Such a field is called uniform and can exist in only a limited region of space. Generally the lines are curved, and the magnitude of the field intensity varies from point to point. If charge is measured in coulombs and force in newtons, the field intensity is in newtons per coulomb. Later it will be shown that field intensity can be expressed also as volts per meter.

1-7. Potential. Since an electric field is due to electric charges, the lines of force representing electric flux must begin on a charge of one polarity and terminate on a charge of the opposite polarity. If a test charge is moved around a closed loop in an electric field, the total work done in moving the charge is zero. When the charged body is moved against the force of the field, work is done on the body by the external force and it gains potential energy. When the body is moved in the same direction as the force due to the field, the body loses potential energy. The movement of a charged body perpendicular to the field direction involves no work. If a charged body moves from a position of high potential energy to one of lower potential energy, the work done, if computed in terms of the field force, will take a positive sign and hence is often alluded to as the work done by the field. Conversely, the work done in raising the potential energy of a charged body in the field, when computed from the force of the field, takes a negative sign. The work that must be done in moving a test charge from one point to a point of higher potential energy can all be regained by allowing the body to return to the first point.

* In this text the boldface letter has either of two meanings. As in the present case, it indicates that such a letter represents a vector quantity. Later, it will be used to designate that the letter represents a complex number.

The difference in electric potential between two points is defined as the negative of the work done by the field in moving a unit positive test charge from one point to the other. Therefore, if the unit positive test charge is moved between two points, against the force of the field, and unit work is done, there is a unit potential difference between the two points, with the second one being at a higher potential than the first. In the mks system this potential difference is called 1 volt.

1-8. Equipotential Surfaces. In order better to understand what happens to charged bodies in an electric field, the field between two infinitely long parallel cylinders will be studied. The two cylinders will be given charges by means of a battery as shown by Fig. 1-1, which also shows a cross section of the two cylinders taken at right angles to their axes.

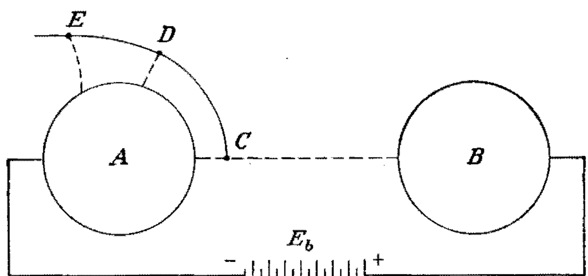


FIG. 1-1. Plotting equipotential lines between two parallel cylinders.

If a unit positive test charge is moved from the surface of cylinder A toward cylinder B, a point C can be found where unit work has been done on the test charge against the force of the field. But since the potential difference is defined as the negative of the work done by the field on a unit positive test charge, point C is at a higher potential than the surface of cylinder A. In the mks system point C has a potential of $+1$ volt with respect to A. If the test charge is moved along any other path, a point D can also be found which has a potential of $+1$ volt with respect to A. In fact, by following different paths an infinite number of points can be found which have a potential of $+1$ volt with respect to A. If all of these points are joined by a line, the result is an equipotential line, the cross section of an equipotential surface. The potential difference between any two points on an equipotential surface is zero; therefore, if a test charge is moved along an equipotential surface, no work is done. The same process can be followed to plot the 2-volt, the 3-volt, and other equipotentials for the two cylinders. The zero equipotential is the surface of cylinder A, and the 10-volt equipotential is the surface of cylinder B.

It has been found that if two electrodes are immersed in a slightly conducting solution, the equipotential lines due to the conduction of current through the solution are the same as those in free space where no current is flowing. Figure 1-2 shows the schematic diagram of the apparatus for making such a plot. A 1000-cps voltage is used instead of a direct voltage to avoid polarizing the solution and also to allow a pair of headphones to be used to indicate balance. The electrodes are made of metal and placed in a nonconducting tray filled with a thin layer of conducting solution. If the electrodes are made of copper, then the solution can be a dilute solution of copper sulfate; thus battery action will be avoided. The bottom of the tray has coordinate lines marked on it, or if it is made of glass, the tray can be placed on a piece of coordinate paper. The 1000-cps signal is applied to the two electrodes and to a potentiometer as shown. The slider of the potentiometer is moved until the voltmeter reads the potential of the equipotential line that is to be plotted. The probe is then immersed in the solution and moved around until a point is found where no 1000-cps tone is heard in the headphones. The potential of the probe is then the same as the potential of the slider on the potentiometer, which in turn has the potential (with respect to A) indicated by the voltmeter. The coordinates of the point are then recorded. This same procedure is repeated as many times as necessary to draw the equipotential line. The slider on the potentiometer is then changed to the potential of the next line and the procedure repeated until all equipotentials are plotted. Figure 1-3 shows the plot of the equipotential lines between two parallel cylinders, taken by means of an electrolytic tank.

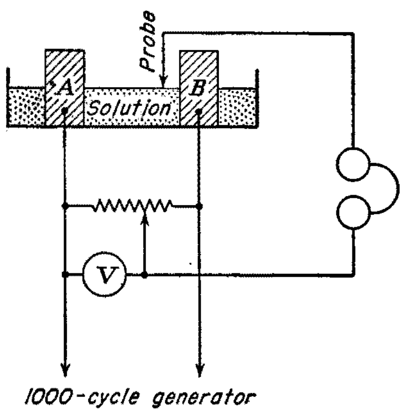


FIG. 1-2. Diagram of a potential-plotting tank.

1-9. Potential Gradient. When a unit positive charge is moved between two points in an electric field, the negative of the work done by that field is a measure of the potential difference between these points. Hence in Fig. 1-4 if the unit charge is moved a distance Δs, the work done by the field is given by

$$\Delta W = f_s \, \Delta s \tag{1-8}$$

where f_s is the average value of the force in the space interval Δs. Since we are using a unit charge, we may rewrite Eq. (1-7) as

$$f_s = \mathcal{E}_s \tag{1-9}$$

If we represent potential by the symbol E, then for a unit charge

$$\Delta E = -\Delta W \tag{1-10}$$

and we may combine Eqs. (1-8) to (1-10) to obtain

$$\Delta E = -\mathcal{E}_s \, \Delta s \tag{1-11}$$

or

$$\mathcal{E}_s = -\frac{\Delta E}{\Delta s} \tag{1-12}$$

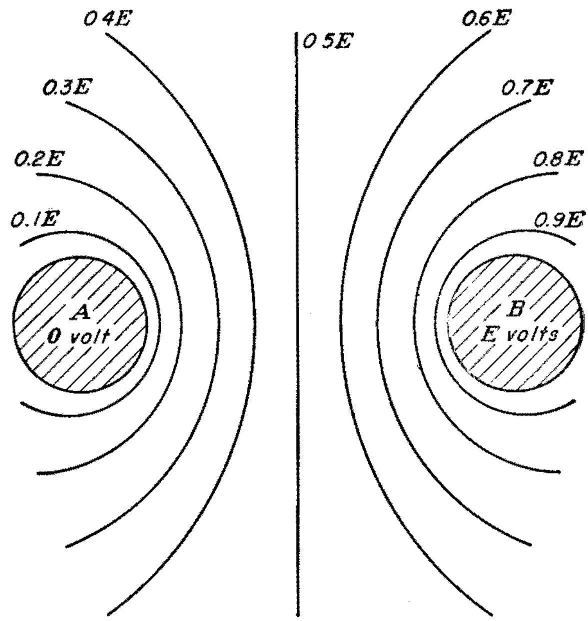


FIG. 1-3. The potential field between two parallel cylinders.

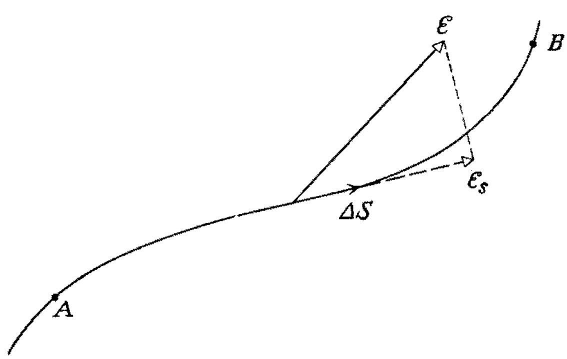


FIG. 1-4.

where $\mathcal{E}_s$ is the average field intensity along the increment of path Δs. In order to find the field intensity at a point, Δs should be made to approach zero. Equation (1-12) then yields

$$\mathcal{E}_s = \lim_{\Delta s \to 0} \left(- \frac{\Delta E}{\Delta s} \right) = - \frac{dE}{ds} \tag{1-13}$$

Thus from the definitions it has been shown that field intensity is equal to the negative of the rate of change of potential with position.

The space derivative in Eq. (1-13) is known as a directional derivative since its value depends upon the direction of ds. If ds is taken along an equipotential, the value of dE/ds equals zero since dE is zero. However, if ds is taken in any other direction, dE/ds is not zero. The direction along which it is a maximum is that of the shortest path which joins two equipotential lines. This is the value of dE/ds known as the gradient of E, abbreviated as grad E, and it is a vector quantity.

From the above discussion it can be seen that the field intensity at a point is a vector equal in magnitude to the potential gradient at that point and having a direction the same as the direction of the shortest path. This may be expressed as

$$\boldsymbol{\mathcal{E}} = - \textbf{grad } E \tag{1-14}$$

Since the field-intensity vector is always along the direction of the gradient and since the gradient is along the shortest path between two equipotential lines, the flux lines must be at right angles to the equipotential lines.

1-10. Experimental Determination of Flux Pattern.[5] A picture of the electric field between two cylinders can be obtained by using a rather deep glass jar, such as a large battery jar, and placing in it two metallic cylinders standing on end. Kerosene is then poured in to a depth of about 2 in. By means of a long-stemmed funnel carbon tetrachloride is poured into the bottom of the jar under the kerosene. If care is taken, the kerosene, having a lower specific gravity than the carbon tetrachloride, will float on top. There will be a sharp surface of separation between the two liquids. Enough carbon tetrachloride is poured into the jar so that this surface is at least 1 in. off the bottom of the jar. Next silk fibers are cut into about $\frac{1}{16}$-in. lengths and dropped into the jar. The silk fibers have a density such that they sink to the surface of separation and remain there, pointing in random directions. The two cylinders are then connected through a switch to a transformer furnishing several thousand volts. The setup is then left undisturbed for several minutes in order that the small fibers may come to rest. The switch is then closed. The fibers, which carry electric flux with greater ease than do the solutions, tend to line up along the flux lines. The flux pattern can then be noted or photographed if a record is desired. The pattern stays for only a short period of time since the fibers soon become charged and are pulled over to one electrode or the other. Figure 1-5 shows a

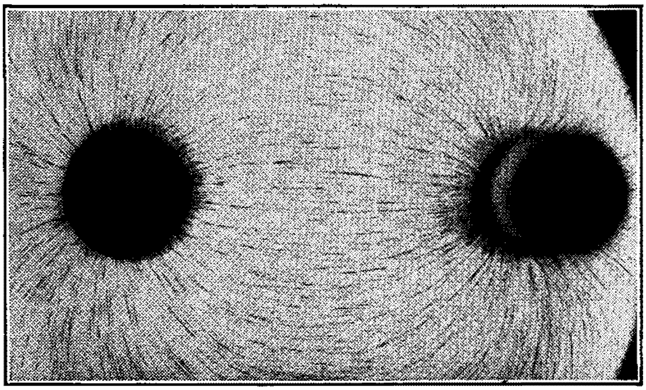


Fig. 1-5. Flux pattern between two parallel cylinders.

photograph of such a flux pattern, while Fig. 1-6 shows the same field with the equipotential lines of Fig. 1-3 drawn to the same scale and superimposed on the flux pattern. Here it may be noted that the two families of curves cross each other at right angles, except at the right edge where the field is somewhat distorted by the walls of the container.

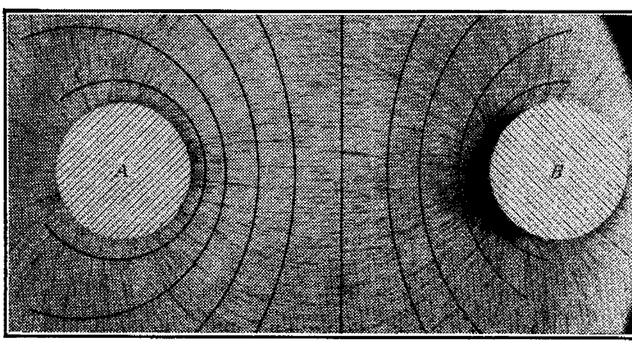


Fig. 1-6. Flux and potential field between two parallel cylinders.

1-11. The Behavior of an Electron in an Electric Field. With this elementary background a study of the motion of charged particles in an electric field may be begun. Newton's second law was expressed mathematically in Eq. (1-2) as $f = m\,dv/dt = ma$. If it is desired to study the motion of an electron in an electric field, m becomes equal to the mass of an electron, $m_e = 9.11 \times 10^{-31}$ kg, while f is the force due to the electric field and Q is the charge on an electron, $Q_e = -1.60 \times 10^{-19}$ coulomb (the charge on the electron is negative). Thus

$$\mathbf{a} = \frac{d^2\mathbf{s}}{dt^2} = \frac{\mathbf{f}}{m} = \frac{Q_e\mathcal{E}}{m_e} = \frac{-Q_e}{m_e}\,\mathbf{grad}\,E \tag{1-15}$$

If the electric field is due to two infinitely large parallel plates connected to a battery, the field between the plates is uniform. This means that the flux lines are straight parallel lines perpendicular to the plates and that the equipotential surfaces are planes parallel to the plates and equally spaced. Therefore the potential gradient is constant at all points in the field and is equal to E_b/d, where E_b is the emf of the battery and d is the distance between the plates. Figure 1-7 shows the circuit and a plot of the potential of all points as a function of the perpendicular distance between the two plates.

Equation (1-15) involves no coordinate system. If axes are chosen so that the origin is at the negative plate and the X axis is as shown, then the force and hence the motion are along the flux lines (unless the electron had an initial velocity due to some other force). The forces parallel to the Y and Z axes equal zero since they are parallel to the equipotential planes. Equation (1-15) may be rewritten as

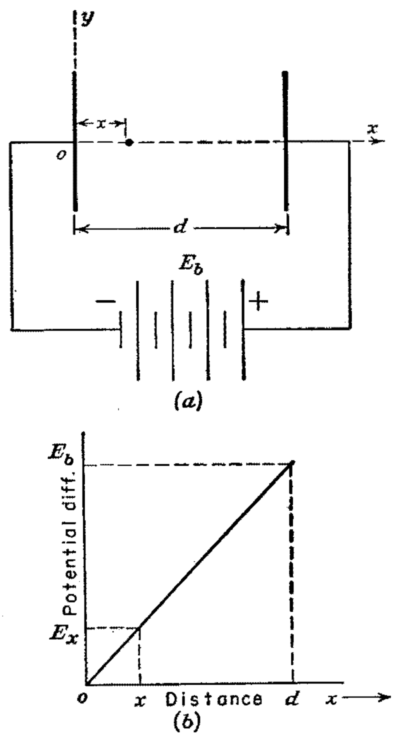


FIG. 1-7. Potential distribution between parallel plane plates.

$$\frac{d^2x}{dt^2} = -\frac{Q_e}{m_e}\frac{E_b}{d} \tag{1-16}$$

If an electron is released with zero velocity at the surface of the negative plate when $t = 0$, then its instantaneous velocity can be calculated by integrating Eq. (1-16),

$$\frac{dx}{dt} = -\int \frac{Q_e}{m_e}\frac{E_b}{d}\,dt = -\frac{Q_e}{m_e}\frac{E_b}{d}\,t + C_1 \tag{1-17}$$

but since the initial velocity is zero, $C_1 = 0$ and therefore

$$\frac{dx}{dt} = v_x = -\frac{Q_e}{m_e}\frac{E_b}{d}t \tag{1-18}$$

If this is integrated, it becomes

$$x = -\frac{1}{2}\frac{Q_e}{m_e}\frac{E_b}{d}t^2 + C_2 \tag{1-19}$$

The electron was released at the negative plate at time $t = 0$, and therefore $C_2 = 0$. Hence

$$x = -\frac{1}{2}\frac{Q_e}{m_e}\frac{E_b}{d}t^2 \tag{1-20}$$

If it is desired to calculate the velocity of an electron after it has traveled a distance x, this can be done by solving Eq. (1-18) for t and substituting this expression into Eq. (1-20). Then solving for v_x yields

$$v_x = \sqrt{-2\frac{Q_e}{m_e}\left(\frac{E_b}{d}x\right)} \tag{1-21}$$

If the curve of Fig. 1-7 is examined, it is seen that the potential of any point can be expressed in terms of the battery voltage and the distance x; thus

$$E_x = \frac{E_b}{d}x \tag{1-22}$$

which is the same as the part of Eq. (1-21) enclosed in parentheses. Thus the velocity at any point is seen to be proportional to the square root of the potential difference through which the electron falls.

If the transit time of the electron is desired, then x is set equal to d and Eq. (1-20) yields

$$t_d = \sqrt{-2\frac{m_e}{Q_e}\frac{d^2}{E_b}} \tag{1-23}$$

If x is set equal to d, Eq. (1-21) gives the velocity of the electron when it reaches the positive plate:

$$v_d = \sqrt{-2\frac{Q_e}{m_e}E_b} \tag{1-24}$$

It is shown in Prob. 1-4*b* that the result in (1-24) can be derived from energy relations which do not involve a knowledge of the character of the electric field. Hence, whether the field is uniform or not, the gain in velocity of an electron depends only on the total change in potential.

1-12. An Electron with an Initial Velocity in a Uniform Electric Field. In Fig. 1-8, if an electron is released at a distance x_0 from the negative plate, with a velocity v_0,

$$\frac{dx}{dt} = -\frac{Q_e}{m_e}\frac{E_b}{d}t + v_0 \tag{1-25}$$

and Eq. (1-19) takes the form

$$x = -\frac{1}{2}\frac{Q_e}{m_e}\frac{E_b}{d}t^2 + v_0t + x_0 \tag{1-26}$$

In all the previous discussions the voltage between the plates was supplied by a battery and hence was held constant. The question now arises as to what is the effect of a variable voltage on the motion of an

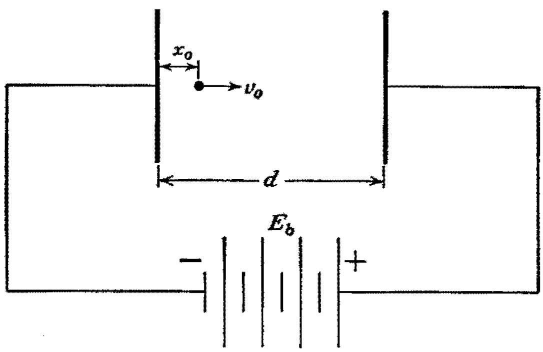


FIG. 1-8. Action of a uniform electric field on an electron with an initial velocity.

electron. Let us represent this voltage by e_b, using lower case to indicate that it varies with time. Since the instantaneous value of the potential gradient changes, the acceleration likewise changes with time. Expressed mathematically, this becomes

$$\frac{d^2x}{dt^2} = -\frac{Q_e}{m_e}\frac{e_b}{d} \tag{1-27}$$

Since e_b is dependent upon t, the integral of the right member of Eq. (1-27) is different from that found for constant E_b. As an example, assume that e_b is zero when $t = 0$ and rises uniformly from 0 to 100 volts in 1×10^{-6} sec. Then the equation for e_b becomes

$$e_b = 10^8t \tag{1-28}$$

Substituting e_b from Eq. (1-28) into Eq. (1-27) yields

$$\frac{d^2x}{dt^2} = -\frac{Q_e}{m_e}\frac{10^8t}{d} \tag{1-29}$$

and hence

$$\frac{dx}{dt} = -\frac{1}{2}\frac{Q_e}{m_e}\frac{10^8}{d}t^2 + C_1 \tag{1-30}$$

However, if $dx/dt = v_0$ when $t = 0$, then Eq. (1-30) becomes

$$\frac{dx}{dt} = -\frac{1}{2}\frac{Q_e}{m_e}\frac{10^8}{d}t^2 + v_0 \tag{1-31}$$

and

$$x = -\frac{1}{6}\frac{Q_e}{m_e}\frac{10^8}{d}t^3 + v_0t + C_2 \tag{1-32}$$

But if in addition $x = x_0$ when $t = 0$, then Eq. (1-32) becomes

$$x = -\frac{1}{6}\frac{Q_e}{m_e}\frac{10^8}{d}t^3 + v_0t + x_0 \tag{1-33}$$

We should always remember that Q_e is a negative quantity since the charge on the electron is negative; hence Eqs. (1-31) and (1-33) render positive quantities if the potential gradient is positive.

The equations derived in Arts. 1-11 and 1-12 for the behavior of an electron in an electric field can be extended to the case of any charged particle by appropriate substitutions of charge and mass values.

1-13. Moving Electrons in a Magnetic Field.[7] A current-carrying conductor in a magnetic field has a force exerted on it if the directions of current flow and of the magnetic flux are not parallel. This is because of the interaction between the external magnetic field and the magnetic field set up by the current.

If the conductor is at right angles to the magnetic-flux lines, then the force on the conductor can be written as

$$f = Bil \tag{1-34}$$

where l is the length, in meters, of the conductor in the magnetic field, B is the flux density in webers per square meter, i is the current, in amperes, flowing in the conductor, and f is the force in newtons. However, if the directions of current flow and of the flux lines make an angle θ with each other as shown in Fig. 1-9, the flux-density vector can be resolved into two components, B_x in the same direction as the current flow and B_y at right angles to the direction of current flow. There is no interaction between B_x and the flux due to the current flow. But B_y produces a force on the conductor. This force is at right angles to the plane of B and the conductor. It is up out of the plane of the paper, and its magnitude is equal to

$$f_z = B_y il \tag{1-35}$$

But since $B_y = B \sin \theta$, Eq. (1-35) can be written

$$f_z = Bil \sin \theta \tag{1-36}$$

The force per unit length of conductor is then

$$f'_z = Bi \sin \theta \tag{1-37}$$

Since current flow is actually the motion of electrons, the force on the conductor is really a force on the electrons in it, which are moving in a direction opposite to that conventionally assigned to current flow.

If the electrons in a conductor are moving with a velocity v, then in 1 sec all the electrons in a length of conductor equal to v m pass a certain reference point. If the conductor has n charged particles (or n electrons) per

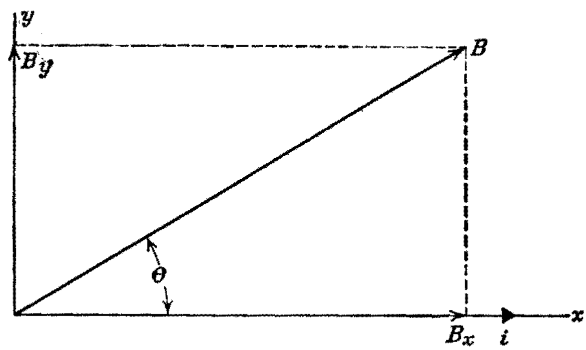


Fig. 1-9.

unit length of path, then in 1 sec nv electrons pass the reference point. Since the charge on the electron is Q_e coulombs and current is the rate of flow of charge, the current can be expressed as

$$i = Q_e nv \tag{1-38}$$

and therefore Eq. (1-37) may take a form

$$f'_z = BQ_e nv \sin \theta \tag{1-39}$$

and the force per electron is

$$f_e = BQ_e v \sin \theta \tag{1-40}$$

If an electron with constant velocity is projected into a magnetic field, with the direction of motion at right angles to the flux lines, the value of θ is 90° and Eq. (1-40) becomes

$$f_e = BQ_e v \tag{1-41}$$

This force is always at right angles to the direction of motion and hence will not change the speed of the electron. The path which results will be curved. The acceleration caused by the force is, from kinematics, equal to v^2/r, where r is the radius of curvature of the path. From Newton's second law, $f = ma$,

$$BQ_ev = \frac{m_ev^2}{r} \tag{1-42}$$

or

$$r = \frac{m_ev}{BQ_e} \tag{1-43}$$

From this it can be seen that the path of an electron moving in a uniform magnetic field which is at right angles to the direction of motion must be circular since the radius is constant.

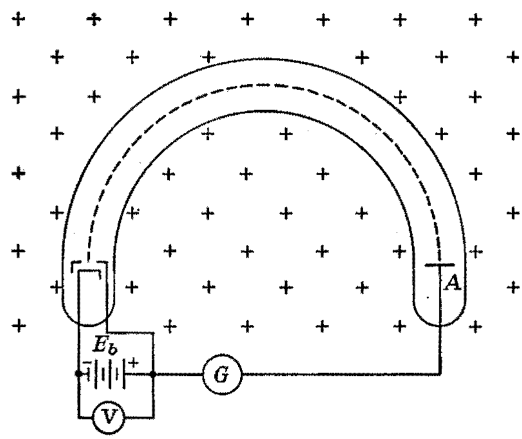


FIG. 1-10. Tube for measuring the charge-to-mass ratio for electrons.

The above discussion suggests one method of measuring the charge-to-mass ratio of an electron by means of a moving electron in a magnetic field. A beam of electrons is projected at a known velocity into a uniform magnetic field in a special semitoroidal tube as shown in Fig. 1-10. The velocity of the electrons can be calculated from the relationship

$$v = \sqrt{-2\frac{Q_e}{m_e}E_b} \tag{1-44}$$

which is the velocity of the electrons as they leave the electron gun. In the drawing their direction of motion would normally be toward the top of the page, but because of the action of the magnetic field, which is shown as flux lines into the plane of the paper, the path will be curved. By proper adjustment of the voltage E_b and of the strength of the magnetic field, which is produced by means of large electromagnets, the electrons can be made to follow a circular path such that they strike the electrode A. This will be indicated by a current in the galvanometer G. The distance between the center of the gun and the pickup electrode determines the radius of curvature. The flux density B can be determined by means of a fluxmeter or by calculations. From Eqs. (1-43) and (1-44)

$$\frac{Q_e}{m_e} = -\frac{2E_b}{r^2B^2} \tag{1-45}$$

1-14. Magnetic Focusing of an Electron Beam. If an electron with an initial velocity v_0 is projected into a magnetic field so that the direction of its motion is not at right angles to the flux lines, then the force on the electron is given by

$$f_e = BQ_ev_0 \sin \theta \tag{1-46}$$

where θ is the angle between the directions of motion and of the magnetic-flux lines. The velocity vector can then be resolved into two components, one of which is parallel to the flux and the other at right angles to the flux as shown in Fig. 1-11. The magnitude of the velocity vector parallel to the flux is given by

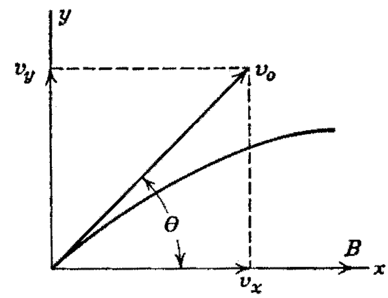


FIG. 1-11.

$$v_x = v_0 \cos \theta \tag{1-47}$$

and results in zero force on the electron. The magnitude of the velocity vector at right angles to the flux is given by

$$v_y = v_0 \sin \theta \tag{1-48}$$

and results in a force on the electron such that the electron travels in a circular path having a radius

$$r = \frac{m_ev_0 \sin \theta}{BQ_e} \tag{1-49}$$

But at the same time that the electron is moving in a circular path perpendicular to the flux, it is also moving parallel to the flux at a constant velocity given by Eq. (1-47). The result of these two motions is a helical path with constant pitch. The peripheral velocity for the circular motion is given by Eq. (1-48); therefore the angular velocity is

$$\omega = \frac{v_0 \sin \theta}{r} \tag{1-50}$$

Substituting r from Eq. (1-49) into Eq. (1-50) yields

$$\omega = \frac{Q_eB}{m_e} \tag{1-51}$$

The time T for one complete revolution of the circular motion is available from

$$Tv_0 \sin \theta = 2\pi r \tag{1-52}$$

or

$$T = \frac{2\pi r}{v_0 \sin \theta} \tag{1-53}$$

hence upon substituting r from Eq. (1-49) into Eq. (1-53) we obtain

$$T = 2\frac{\pi m_e}{BQ_e} \tag{1-54}$$

If L is the pitch of the helix,

$$L = Tv_x = Tv_0 \cos \theta = \frac{2\pi}{B}\frac{m_e}{Q_e} v_0 \cos \theta \tag{1-55}$$

A magnetic field is sometimes used to focus an electron beam in a cathode-ray tube. Such a field is known as a magnetostatic electron

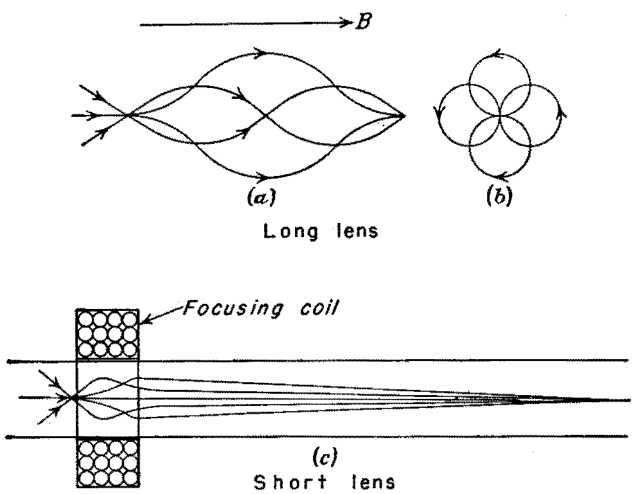


FIG. 1-12. Focusing action of magnetic fields.

lens and takes the form of either a long lens or a short lens. The magnetic field in the long lens extends over the whole length of the electron path. Figure 1-12a shows how a divergent beam of electrons, produced by an electron gun, is focused by a long lens. The electrons in the divergent beam follow helical paths because of the magnetic field along the axis of the tube. The projections of the paths of the electrons, on a plane perpendicular to the direction of motion, are circles which pass through the axis of the tube shown in Fig. 1-12b. The time of revolution for each electron is independent of the angle of divergence of the beam. This is shown by Eq. (1-54). If the angle of divergence is small, then $\cos \theta$ is essentially equal to unity and Eq. (1-47) indicates that all electrons will have approximately the same translational velocity. Therefore at the end of one revolution they will all be in the same relative positions as when they entered the magnetic field. By adjusting the magnitude of the magnetic field, which is produced by a long solenoid, the electrons can be made to go through one or more complete revolutions during their flight down the tube, thus focusing the aperture of the gun on the end of the tube or on another electrode, as shown in the figure.

In the short form of the magnetostatic lens the focusing field may be produced by a short coil (or a permanent magnet) surrounding the tube. The axis of the coil is concentric with the axis of the tube as shown in Fig. 1-12*c*. The magnetic field produced by the coil, which is short compared with the length of the electron path, causes the electrons to follow helical paths while passing through it. The strength of the magnetic field is adjusted so that the electrons leave the field with radial components of velocity of such magnitude that they are focused into a small spot at the terminus of the paths. This is the form of lens most commonly used to focus television-picture tubes.

1-15. The Cathode-ray Tube[6]—Electrostatic Focusing. One of the simplest applications of electron ballistics to electronic tubes is found in the cathode-ray tube. Here low-velocity electrons are accelerated by means of an electric field. While they are being accelerated, they are also being focused into a very narrow beam. They may then pass between electrodes called deflecting plates, where their direction of motion is changed by a voltage applied between the plates. The electrons then are allowed to travel the length of the tube. When they strike the end of the tube and are stopped suddenly, they give up their kinetic energy. The end of the tube is coated with a substance called a phosphor. This material has the property of being able to convert the energy of the electron into light energy. Where the electrons strike the end of the tube, a bright spot of light appears.

Thus it can be seen that the most important parts of the cathode-ray tube are those which produce, accelerate, and focus the electron beam, the deflection system, and the phosphor or screen. The elements which produce, accelerate, and focus the electrons constitute the electron gun. Figure 1-13 is a sketch of the cross section of a very simple cathode-ray tube. The electron gun has a cathode K which when heated produces or emits electrons. These electrons are emitted in a direction normal to the surface of the cathode, but because of the force of repulsion between electrons they have a tendency to spread, or fan out. The grid G is a cylinder with a cap over one end. In the center of the cap is a very small hole, or aperture. This aperture permits passage of only that portion of the beam which is moving approximately parallel to the axis of the tube. The potential of G can be made more or less negative, and thus it controls the rate of flow of electrons passing through the grid. While passing between the grid and the first anode A_1, the electrons are accelerated and also focused into a convergent beam. After passing through a crossover, or focal, point, the beam again becomes divergent but soon after enters the space between the first anode A_1 and the second anode A_2. Here again it experiences a force, an acceleration, and a focusing action. But by this time the velocity of the electrons has become so great that they do not go through another crossover point until they reach the screen. Hence the spot of light on the screen should be very small. Figure 1-13 shows the focusing action of the cylindrical electrodes used in this simple gun.

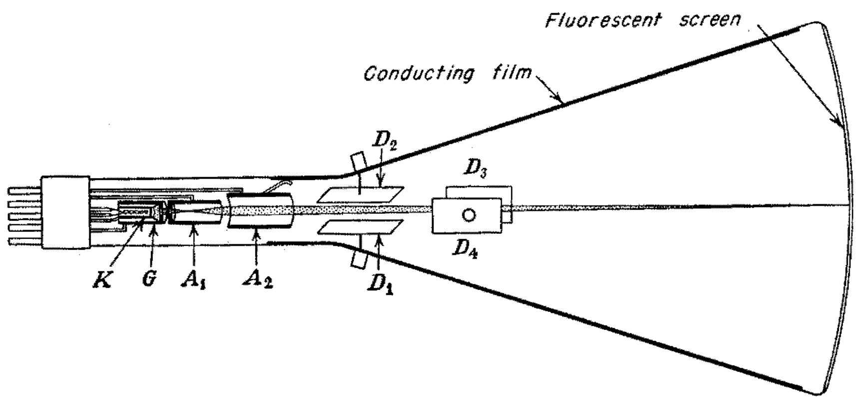


FIG. 1-13. Construction of a cathode-ray tube having electrostatic focusing and deflection.

The reason why an electric field can be used to focus an electron beam will be better understood if Fig. 1-14 is studied. This sketch shows three equipotential lines between the first and the second anodes. They are symmetrical about the x axis. If an electron traveling along the x axis arrives at point P with a velocity v_1, it has a force

$$\mathbf{f}_1 = -Q_e \,\mathbf{grad}\, E \tag{1-56}$$

exerted on it by the electric field. **grad** E in this case has a direction along the x axis, and its magnitude may be variable between P and Q. The result is that the electron is accelerated and arrives at point Q with a velocity which is greater in magnitude but still directed along the x axis.

However, if an electron arrives at the equipotential E_1 and is not traveling in a direction which is normal to the equipotential, then the direction of motion of the electron will be changed. If the electron reaches point R with a velocity v_2, it has a force exerted on it by the field equal to

$$\mathbf{f}_2 = -Q_e \,\mathbf{grad}\, E \tag{1-57}$$

Since **grad** E always has a direction normal to the equipotential, at R it has a direction shown in the figure. The resultant motion of the electron is due to the force f_2 in Eq. (1-57) and to the initial velocity v_2 of the electron. The path therefore is somewhere between the direction of f_2 and that of v_2 and is bent in toward the axis of the tube. The amount of

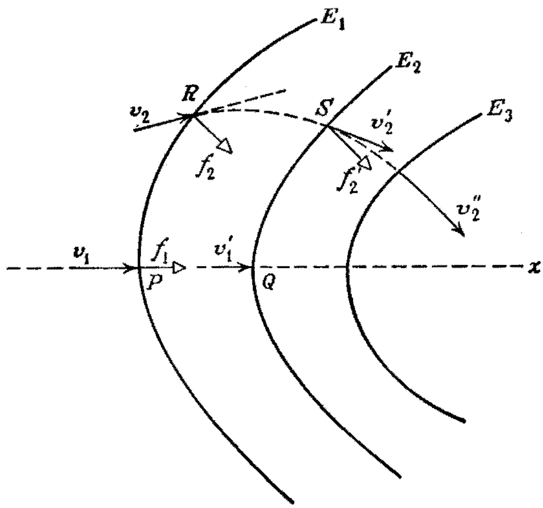


FIG. 1-14. Focusing action of the electric field in an electron lens.

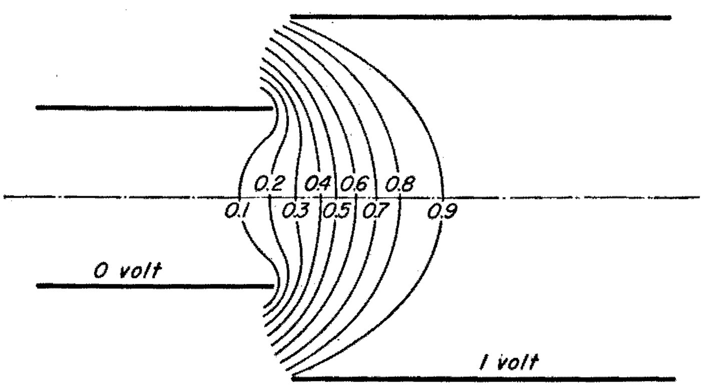


FIG. 1-15. Potential field of a two-cylinder electrostatic lens.

bending depends upon the magnitude and direction of both the field intensity at each point on the path and the velocity of the electron.

Making the proper choice of electrodes for an electron gun depends largely upon experience gained in laboratory experimentation. The field for any symmetrical electrode configuration can be determined by plotting the equipotential lines using an electrolytic tank in a manner similar to that explained in Art. 1-8. Figure 1-15 shows such a plot. Here it may be noted that the equipotentials are shaped to cause a convergent beam until the electrons reach the plane between the two cylinders. There they enter a portion of the field that tends to cause the beam to diverge. However, because of the larger radii of curvature of the equipotential surfaces together with the higher velocities of the electrons in their convergent paths, the divergent effect of the new portion of the field is relatively weak. With proper adjustment of electrode potentials the paths of the electrons can be made somewhat divergent, approximately parallel, or convergent, as desired.

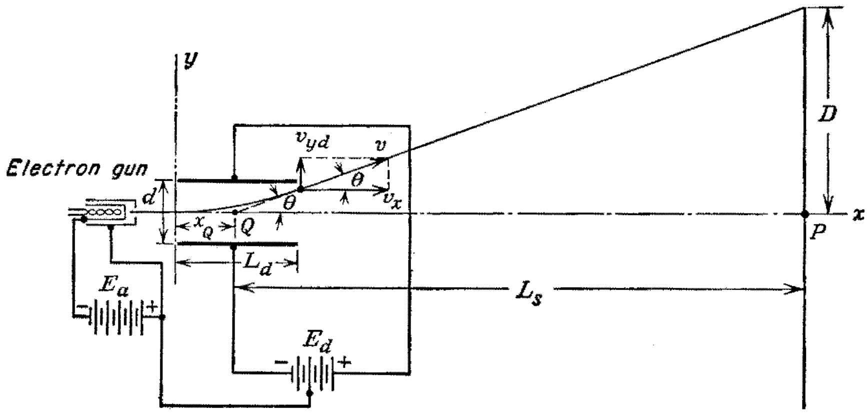


Fig. 1-16.

1-16. The Deflection of the Electron Beam in a Cathode-ray Tube. As an electron beam passes through the field of the electron gun, it is accelerated and leaves the gun with a velocity

$$v_x = \sqrt{-2\,\frac{Q_e E_a}{m_e}} \tag{1-58}$$

where E_a is the potential of the second anode relative to the cathode. The electrons then travel along the x axis with this constant velocity until they strike the end of the tube at point P in Fig. 1-16, unless they enter another field which causes further acceleration.

If the latter field is due to two flat plates parallel to the axis of the tube and with a voltage E_d between them, then the acceleration is at right angles to the x axis and the beam of electrons travels in a curved path as shown in Fig. 1-16. The acceleration of the beam is given by its components

$$a_y = \frac{d^2y}{dt^2} = -\frac{Q_e}{m_e}\frac{E_d}{d} \tag{1-59}$$

$$a_x = 0 \tag{1-60}$$

Equation (1-60) holds since there is no field parallel to the x axis in this region. Integrating Eq. (1-59) gives

$$\frac{dy}{dt} = v_y = -\frac{Q_e}{m_e}\frac{E_d}{d}t + v_0 \tag{1-61}$$

but $v_0 = 0$ since the electrons have no component of velocity parallel to the y axis when they enter the space between the plates. Omitting v_0 and integrating (1-61) yields

$$y = -\frac{Q_e}{2m_e}\frac{E_d}{d}t^2 + y_0 \tag{1-62}$$

But $y_0 = 0$ when the electron enters the field since it is traveling along the x axis.

Integrating Eq. (1-58), the equation for the motion in the x direction can be written as

$$x = \sqrt{-2\frac{Q_e E_a}{m_e}}\,t + x_0 \tag{1-63}$$

If the origin is set so that $x = 0$ when the electron just enters the field between the plates, then $x_0 = 0$ when $t = 0$. It follows that

$$t = \frac{x}{v_x} = \frac{x}{\sqrt{-2(Q_e E_a/m_e)}} \tag{1-64}$$

Combining Eqs. (1-62) and (1-63) to eliminate t yields

$$y = \frac{1}{4d}\frac{E_d}{E_a}x^2 \tag{1-65}$$

and since d, E_d, and E_a all are constant, (1-65) is the equation of a parabola, which holds only as long as the electrons are between the two plates and are being accelerated by the field between them. As the electrons leave the field between the plates, they continue in straight lines. The angle their direction makes with respect to the horizontal is given by

$$\tan\theta = \frac{v_{yd}}{v_x}, \tag{1-66}$$

where v_{yd} is the y velocity of the electrons at the instant they leave the field. In order to calculate this velocity, the time required for the electrons to travel the length of the plates must be known, and this can be calculated from Eq. (1-64) by letting $x = L_d$,

$$t_d = \frac{L_d}{\sqrt{-2Q_e E_a/m_e}} \tag{1-67}$$

Substituting t_d into Eq. (1-61) results in

$$v_{yd} = -\frac{Q_e}{m_e}\frac{E_d}{d}\frac{L_d}{\sqrt{-2Q_eE_a/m_e}} \tag{1-68}$$

Substituting from Eq. (1-58) and (1-68) into Eq. (1-66) gives

$$\tan\theta = \frac{1}{2}\frac{E_d}{E_a}\frac{L_d}{d} \tag{1-69}$$

If the path of the electrons after leaving the plates is extended back as a straight line until it intersects the x axis, it does so at the point Q where $x = L_d/2$. This can be proved as follows.

$$\tan\theta = \frac{y_d}{L_d - x_Q} \tag{1-70}$$

From Eq. (1-62)

$$y_d = -\frac{1}{2}\frac{Q_e}{m_e}\frac{E_d}{d}t_d^2 \tag{1-71}$$

Substituting t_d from Eq. (1-67) into Eq. (1-71) yields

$$y_d = \frac{E_d}{4E_a}\frac{L_d^2}{d} \tag{1-72}$$

Substituting from Eq. (1-72) into Eq. (1-70) and equating the result to $\tan\theta$ given by Eq. (1-69),

$$\tan\theta = \frac{E_d}{4E_a}\frac{L_d^2}{d}\frac{1}{L_d - x_Q} = \frac{1}{2}\frac{E_d}{E_a}\frac{L_d}{d} \tag{1-73}$$

gives, when solved for x_Q,

$$x_Q = \frac{L_d}{2} \tag{1-74}$$

Thus the projection of the straight-line portion of the path of the electrons always intersects the x axis at the center of the region between the plates. Since this is true, it is no longer necessary to use the expression for the parabolic path, Eq. (1-65), in order to calculate the deflection of the electrons. Instead, the path is taken to be along the axis of the tube until the electrons reach the point Q, where they make a sharp turn at an angle θ with the tube axis. They then travel along a straight line until they strike the screen.

The deflection sensitivity of a cathode-ray tube is the ratio of the deflection D of the beam at the screen to the voltage causing that deflection. From similar triangles we may write

$$\tan\theta = \frac{D}{L_s} = \frac{v_{yd}}{v_x} \tag{1-75}$$

and from this and Eq. (1-69),

$$\frac{D}{L_s} = \frac{E_d L_d}{2E_a d} \tag{1-76}$$

Solving for the deflection sensitivity,

$$\frac{D}{E_d} = \frac{L_s L_d}{2E_a d} \tag{1-77}$$

Thus we see that the deflection sensitivity is a function of the tube geometry and of the accelerating voltage. In order to have a high sensitivity the plates should be as long and as close together as possible. Since the beam should not strike the plates, the length and the spacing between them are limited by the maximum angle θ required for the beam to reach the edge of the screen. Also for high sensitivity the length of the tube from the center of the deflecting plates to the screen should be as long as practicable.

It should be noted that all the discussion up to this point has been for a constant voltage on the deflecting plates. However, if the voltage is approximately constant for the time that it takes an electron to travel the length of the plates, then these same relationships hold for each electron. Therefore, a varying voltage may be applied to the deflecting plates, and if the variation in voltage is small during the time the electron is between the plates, then that instantaneous value of voltage determines the amount of deflection at the screen. Equation (1-77) then gives

$$D = \frac{L_s L_d}{2E_a d} e_d \tag{1-78}$$

If $e_d = E_{dm} \sin \omega t$ and if the time of 1 cycle of the sinusoidal voltage is large compared with the time it takes an electron to pass between the plates, then the deflection at the screen is proportional to the instantaneous values of the deflecting voltage and hence is sinusoidal.

The important fact that the electric field extends, or fringes, beyond the edges of the deflecting plates has been neglected in the derivation. Actually the deflection sensitivity is greater than the above equations predict. Empirical relationships have been developed to account for this effect of fringing. For the case of the parallel plates it is found that the length of the plates should be considered as the actual length plus an amount equal to the distance between the plates, or

$$L_d' = L_d + d \tag{1-79}$$

Equation (1-78) then becomes

$$D = \frac{L_s L_d'}{2E_a d} e_d \tag{1-80}$$

Actually a cathode-ray tube with electrostatic deflection has two sets of deflecting plates, one in front of the other and with their planes at right angles as shown in Fig. 1-13. Thus one set of plates deflects the beam in one direction, while the other set deflects it in a perpendicular direction. If the tube is oriented so that one set of plates deflects along the horizontal, or x, axis, then the other set deflects along the y, or vertical, axis. Now if the first set, or horizontal plates, can be made to move the spot across the screen at a constant rate and the second set to move it vertically at a sinusoidal rate, the spot should trace a sine wave on the face or screen of the tube. This can be effected by causing the horizontal-plate potential to vary linearly with time by means of a saw-tooth wave such as shown in Fig. 1-17. At the same time the voltage applied to the vertical deflecting plates is sinusoidal. If the time for 1 cycle of the saw-tooth is the same as the time for 1 cycle of the sinusoidal wave, then 1 cycle of the sine wave will appear on the screen of the tube. If the time of 1 cycle of the saw-tooth is equal to the time of 2 cycles of the sine wave, 2 cycles will appear on the screen.

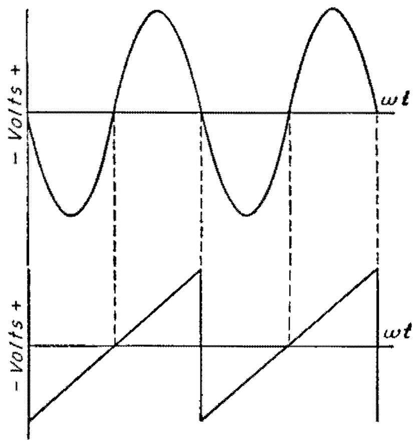


FIG. 1-17. Sinusoidal and saw-tooth waveforms.

The saw-tooth waveshape necessary for the linear-time axis can be generated by means of an electronic circuit known as a sweep, or saw-tooth, generator. The voltage output of this circuit is usually so small that the saw-tooth wave must be amplified before it is applied to the deflecting plates. Likewise the voltage which is applied to the vertical plates may also be amplified if it is not large enough to deflect sufficiently the beam in the cathode-ray tube.

Thus it can be seen that the cathode-ray tube with its necessary circuits, such as rectifiers, sweep generators, and amplifiers, can be used for studying or examining the operation of a-c circuits. Such a device is known as a cathode-ray oscilloscope.

REFERENCES

1. Millikan, R. A.: "Electrons, Protons, Photons, Neutrons, Mesotrons and Cosmic Rays," University of Chicago Press, Chicago, 1947.
2. Kennelly, A. E.: Recent Developments in Electrical Units, *Elec. Eng.*, **58**, 78 (1939).

3. Dow, W. G.: "Fundamentals of Engineering Electronics," John Wiley & Sons, Inc., New York, 1937.
4. Boast, W. B.: "Principles of Electric and Magnetic Fields," Harper & Brothers, New York, 1949.
5. George, R. H., K. A. Oplinger, and C. F. Harding: Improved Method of Visualizing and Photographing the Dielectric Field, *Purdue Univ. Eng. Expt. Sta. Bull.* 29 (November, 1927).
6. "The Cathode-ray Tube and Typical Applications," A. B. DuMont Laboratories, Clifton, N.J., 1948.
7. E. E. Staff, MIT: "Applied Electronics," John Wiley & Sons, Inc., New York, 1943.
8. Millman, J., and S. Seely: "Electronics," 2d ed., McGraw-Hill Book Company, Inc., New York, 1951.
9. Ryder, J. D.: "Electronic Fundamentals and Applications," Prentice-Hall, Inc., New York, 1950.
10. Spangenberg, K. R.: "Vacuum Tubes," McGraw-Hill Book Company, Inc., New York, 1948.

PROBLEMS AND QUESTIONS

1. A uniform tungsten wire is sealed in vacuo and a direct voltage applied as shown in Fig. 1-18. (*a*) Sufficient current is passed through the wire to warm it but not to cause electrons to be emitted. Is one end of the wire warmer than the other, and if so, why? (*b*) The battery voltage is raised until the tungsten wire emits electrons. Should the switch be in position 1 or 2 to cause a deflection of the milliammeter? (*c*) Should the positive terminal of the milliammeter be down or up in the diagram?

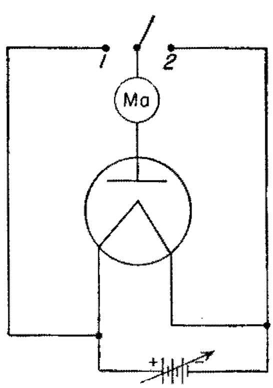


FIG. 1-18.

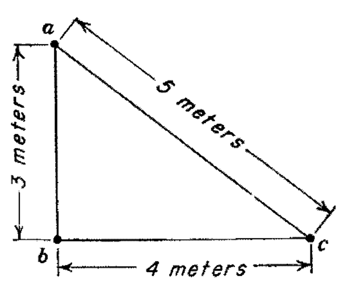


FIG. 1-19.

2. Three particles, each with a negative charge of 1 microcoulomb, are located at points *a*, *b*, and *c* (Fig. 1-19). (*a*) Determine the magnitude of the resultant force on the particle at (*c*). (*b*) Determine the direction of this force, using a horizontal reference.

3. A free electron is in a uniform electric field whose strength is 10,000 newtons per coulomb. If the electron is initially at rest, what is its velocity at the end of 0.002 μsec?

4. In Fig. 1-20 an electron traveling from left to right reaches hole A (very small) with a velocity $v_0 = 7 \times 10^6$ m per sec. Determine its velocity v when it passes through hole B (also very small) by two methods: (*a*) Begin with the relation $f = ma$, and write all steps to a final formula for v in terms of v_0, E_b, and the charge-to-mass ratio. (*b*) Use the first and third members of the equality $(\text{KE})_B - (\text{KE})_A = -(U_B - U_A) = -Q(E_B - E_A)$, where KE is kinetic energy and U is potential energy, to obtain an expression for v at position B.

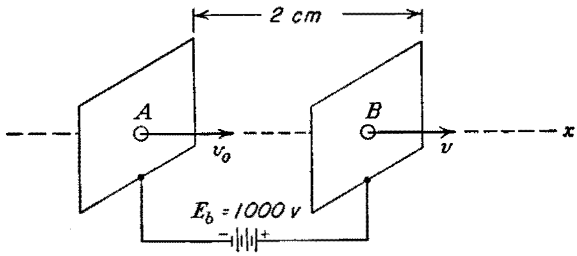


FIG. 1-20.

5. Two parallel plane plates, A and B, are 1 cm apart. An electron is at rest very close to plate A. At time $t = 0$, the voltage on plate B is made 1 volt positive relative to plate A by means of a square-wave generator. The 100-Mc-voltage waveform thereafter is shown in Fig. 1-21. (*a*) Determine the velocity and the distance traveled at the end of ½ cycle. (*b*) Repeat for 1 cycle. (*c*) Does the electron ever reach plate B?

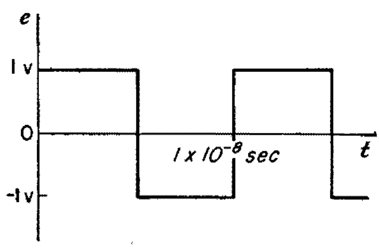


FIG. 1-21.

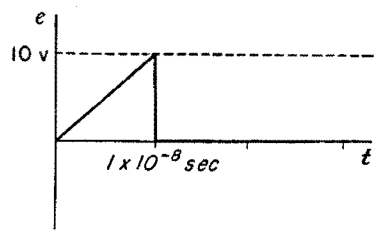


FIG. 1-22.

6. The square-wave generator in Prob. 5 is replaced by a pulse generator which supplies to plate B the waveform of voltage shown in Fig. 1-22. (*a*) How far has the electron traveled at the end of 1×10^{-8} sec? (*b*) What is its velocity? (*c*) How long will it take the electron to reach plate B?

7. The generator in Prob. 5 is replaced by a sine-wave generator for which $e_b = E_{bm} \sin \omega t$. Derive expressions for the velocity of and the distance traveled by the electron.

8. An electron initially at rest is first accelerated by an electric field caused by a potential difference of 1000 volts between two parallel plates 0.1 m apart. The second plate has a small hole through which the electron passes into a space free of electric field. This space does have, however, a uniform magnetic field of 0.5 weber per m² perpendicular to the direction of motion of the electron at the instant of entering this space. (*a*) Determine the speed of the electron while in the magnetic field. (*b*) Determine the radius of the circular path traveled by the electron. (*c*) Calculate the time required for the electron to complete a semicircle.

9. Why does a beam of electrons have a tendency to spread?

10. At a given time two electrons are very close together and have the same speed. One has a velocity parallel to a uniform magnetic field, while the other makes a small angle with the field. Both electrons will eventually hit a screen in the tube. (*a*) Can a screen distance be computed so the electrons will hit the same spot? (*b*) Will the electrons hit at the same time? Explain.

11. Refer to Fig. 1-16. If the distance from the center of the deflecting plates to the screen is 30 cm, the length of the deflecting plates is 2.5 cm, the distance between them is 0.7 cm, and the second-anode voltage is 2000 volts, calculate (*a*) the velocity of the electron along the axis of the tube, (*b*) the time required for the electron to pass between the deflecting plates, (*c*) the deflection sensitivity of the tube, (*d*) the magnitude of the deflection at the screen (the deflection voltage being constant and equal to 100 volts), and (*e*) the corrected deflection for this tube if fringing of the field about the deflecting plates is to be considered.

12. An electron with energy of 1000 electron volts enters the field between the two large parallel plates shown in Fig. 1-23. The motion is initially perpendicular to the field, and the point of entrance is halfway between the plates. Determine the distance A at which the electron strikes the upper plate (neglect field fringing).

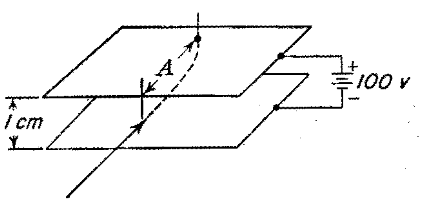


FIG. 1-23.

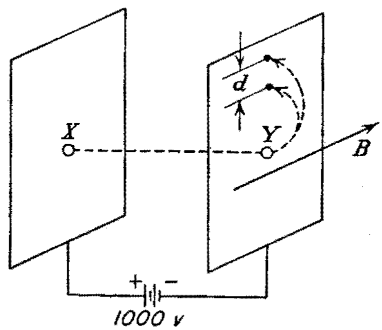


FIG. 1-24.

13. In Prob. 12 what direction and strength of magnetic field are required to prevent deflection of the electron while between the plates?

14. In order to distinguish between two isotopes of helium having atomic weights 3 and 4, respectively, ions of each are introduced through hole X (Fig. 1-24) and accelerated to pass through hole Y. At Y they enter a magnetic field which is parallel to the plates. $B = 0.1$ weber per m^2. Determine the distance d between the marks produced on a photographic film lying on the right-hand plate.

15. The anode of an X-ray tube is at 100,000 volts relative to the cathode. (*a*) What energy in joules will an electron have which starts from rest at the cathode and travels to the anode? (*b*) If the anode current is 5 ma, what power must be dissipated by the anode? Neglect the energy possessed by the X rays.

CHAPTER 2

EMISSION

2-1. Introduction. In the previous chapter the effects of electric and magnetic fields on the motions of electrons were discussed in some detail. These constitute a very important phase of electronics, but free electrons in space must be produced by one or more of the following processes:

1. Thermionic emission
2. High-field emission
3. Secondary emission
4. Photoelectric emission
5. Ionization of gases
6. Radioactivity

The first five of these will be studied in turn, but this chapter deals only with the first three in detail and photoelectric emission very briefly. Ionization of gases will be discussed in Chap. 13, Conduction through Gases. Radioactivity does not assume great importance in our study.

2-2. Structure of the Atom.[1] If a periodic chart of the atoms is examined, it will be noted that there are over 90 different elements appearing in nature. Each element has a different atomic weight and its own chemical properties. If the atoms of the various elements could be magnified until they become visible, it would be seen that all atoms of any element are almost identical. Each would resemble a miniature solar system consisting of a large central body called a nucleus and one or more electrons revolving in elliptical paths, or orbits, about it. Furthermore, it would be noted that the electrons travel in certain distinct orbits at definite but different energy levels. The more energy an electron has, the larger the radius of its orbit. Also an electron never goes from one energy level to another without either absorbing or giving up energy. In the lowest level, which is closest to the nucleus, there is room for 2 electrons; in the next higher level there is room for 8 electrons, for 18 in the next, etc. Each level is able to accommodate a certain maximum number and no more. In a normal atom each level is filled to capacity before any electrons appear in the next higher level. Thus in the normal atom all electrons are revolving in orbits as close to the nucleus as they can get, with the low-energy, or tightly bound, electrons close to the nucleus and the high-energy, or loosely bound, electrons farther from the nucleus. These loosely bound, or high-energy, electrons are the ones in which we are interested since they are the ones which will become the free electrons necessary for the flow of current in a conductor or for the emission of electrons in a vacuum tube.

2-3. The Nucleus. A few facts concerning the nucleus might well be reviewed at this time. The nucleus, as stated in Chap. 1, is a very heavy body when compared with an electron. The mass of the hydrogen nucleus is 1849 times that of an electron. The nucleus includes one or more protons and in some cases one or more neutrons and positrons. The net charge on the nucleus is positive and equal in magnitude to the total charge on the electrons associated with it. This results in forces of attraction between the nucleus and the electrons. These are part of the forces which hold the electrons in their orbits about the nucleus.

2-4. Free Electrons in a Metal. When the structure of any metal is studied, it is found to be an aggregate of crystals of various sizes. A crystal in turn is found to consist of an orderly or geometrical arrangement of atoms in three-dimensional space, these atoms being held in position by interatomic forces. Each atom is composed of a positive nucleus surrounded by orbital electrons, as explained in the previous article. Thus when the atoms are brought into some sort of regular pattern, as they are in a crystal, the outermost electrons in any atom will be acted upon not only by their parent nucleus but by neighboring nuclei as well. In fact some of the orbits may actually overlap, and the forces on the electron due to the two positive nuclei cancel each other. The electron is then no longer bound to a nucleus and hence is free to be acted upon by any other force or field which may be present. If now a voltage sets up an electric field in the metal, the electrons will move or drift toward the point of more positive potential. It is this motion or drift of electrons which constitutes current flow in a metal.

When an electron becomes free, it has a certain amount of energy associated with it. In fact, if the energies of all free electrons in a certain metal at zero degrees Kelvin ($0°K = -273°C$) could be measured, it would be found that they possess kinetic energies ranging from very small values up to a definite maximum value which would depend upon the properties of the metal. There would be no electrons possessing energies greater than this amount no matter how many samples of the metal were examined. If the metal were heated, it would be found that the energies of the free electrons would be increased. At ordinary room temperatures these free electrons could move about in the metal but very few would be able to escape from the body of the metal.

2-5. Electron Escape from a Metal. The reasons why an electron has difficulty in escaping from the surface of a metal can probably best be shown by investigating the forces between a positive nucleus and an electron as the electron is moved farther and farther away from the nucleus.

Consider a single electron in the field of an isolated nucleus. The nucleus has a positive charge and the electron a negative one; hence there is a force of attraction between them in accordance with Coulomb's law,*

$$f = \frac{(ZQ_e)Q_e}{4\pi\epsilon_0 r^2} \tag{2-1}$$

where f = force of attraction
Z = atomic number of the atom, which equals the number of electrons surrounding the nucleus
Q_e = charge on the electron
$-Q_eZ$ = positive charge on the nucleus
r = distance between the electron and the nucleus
ϵ_0 = permittivity of free space, $1/(36\pi \times 10^9)$ in the mks system

Figure 2-1 shows the force on the electron as a function of the distance r from the nucleus. From this figure and from Eq. (2-1) it can be seen that the force attracting the electron to the nucleus does not equal zero until the distance r is equal to infinity. Therefore, if an electron is to escape the attractive force of a positive nucleus, it must be given enough kinetic energy to carry it to infinity.

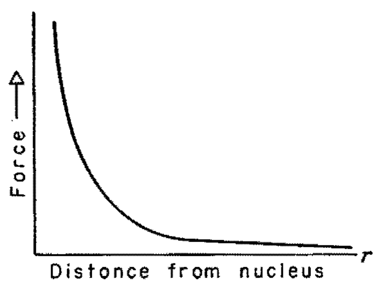


FIG. 2-1. Force on an electron in the vicinity of an isolated nucleus.

Now let us assume that, on account of the energy it already possesses, an electron arrives with zero velocity at a point located a distance r from a nucleus. It is desired to calculate the added energy required by this electron in order to free it. This can be done by calculating the amount of work done against the force, given by Eq. (2-1), while moving the electron to infinity.

$$W = \int_r^\infty f\,dr = \int_r^\infty \frac{ZQ_e^2}{4\pi\epsilon_0 r^2}\,dr = \frac{ZQ_e^2}{4\pi\epsilon_0 r} \tag{2-2}$$

This relationship is shown in graphical form in Fig. 2-2. It will be useful later.

When electrons escape from their orbits within a metal, they can do so because the attractive forces of adjacent nuclei cancel the forces holding

* Coulomb's law holds only for values of r which are many times greater than atomic radii. Hence the discussion following Eq. (2-1) is valid only for relatively large distances between nucleus and electron.

them in their orbits. However, in the case of electrons trying to escape from the surface of the metal, there are no adjacent nuclei on the outside of the metal to cancel the forces of attraction between the electrons and the nuclei of the metal. Therefore, if an electron arrives at the surface of a metal with a given amount of energy, it will try to leave the surface but will be pulled back in unless it has enough kinetic energy to carry it to infinity. This is analogous to throwing a ball up into the air. It will always return unless you can give it enough energy to carry it out past the gravitational field, which would mean out to infinity unless other heavenly bodies interfere. At room temperature few electrons possess sufficient energy to escape from the binding forces at the surface of a

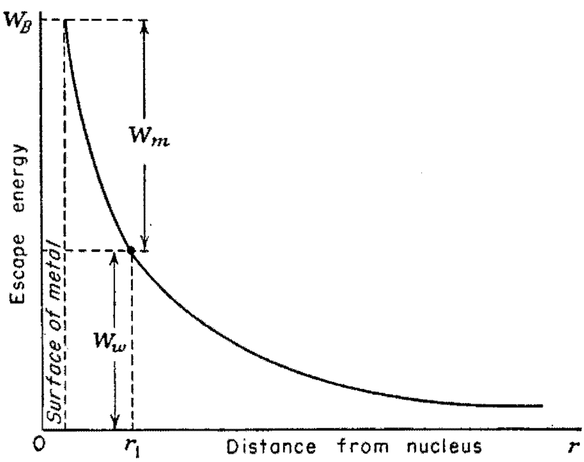


FIG. 2-2. Energy required by an electron to escape the force attracting it to an isolated nucleus.

metal. But as the metal is heated the energies of the free electrons are increased, until finally at a very high temperature many of the electrons will have enough energy to escape from the surface. We then have what is known as thermionic emission.

2-6. The Work Function. An analysis of the energies involved in the escape of electrons from the surface of a metal is considerably more complicated than that discussed in the preceding article for the escape of one electron from an isolated nucleus. For a simplified qualitative explanation, however, let us assume that the graph of Fig. 2-2 is applicable to the problem of electron escape from the surface of a metal.

Let us denote by W_B the energy required for the escape of an electron which arrives with zero velocity at the surface of a metal. Also let us indicate by W_m the maximum energy, at 0°K, possessed by any electron arriving at the surface of a metal and moving in a direction normal to it.

The added energy required for the escape of the latter is called the work function W_w of the metal. Thus

$$W_w = W_B - W_m \tag{2-3}$$

W_w is therefore the minimum energy that must be supplied to the electrons in a material at 0°K in order to give them enough energy to overcome the boundary forces trying to hold them within the surface of the material. Note that this definition specifies the minimum energy. This means the energy which must be supplied to the highest-energy electrons which exist in the material at 0°K.

2-7. The Electron-volt as a Unit of Energy. The ordinary unit of energy in the mks system is the joule, but the work function is usually expressed in electron-volts. If an electron falls through a potential difference of 1 volt, it acquires an energy given by

$$-Q_eE = 1.6 \times 10^{-19} \times 1 = 1.6 \times 10^{-19} \text{ joule} \tag{2-4}$$

The energy gained is proportional to the voltage through which an electron falls. Therefore energy can be expressed in terms of voltage. If an electron has fallen through a potential difference of 1 volt, we say that it has 1 electron-volt of energy instead of saying that it has 1.6×10^{-19} joule of energy.

If the minimum energy W_w required for the escape of an electron is equivalent to that acquired by the electron in falling through ϕ_e volts, then

$$W_w = -\phi_e Q_e \qquad \text{joules} \tag{2-5}$$

where ϕ_e is the work function expressed in electron-volts.

2-8. Contact Difference in Potential. When two dissimilar metals are put in good electrical contact, a potential difference will be found to exist between them. This potential difference is known as the contact potential difference and is numerically equal to the difference between the work functions of the two metals when expressed in electron-volts. It is due to the fact that electrons can pass from the material of low work function to that of the high work function more easily than in the reverse direction. The high-work-function material then becomes negatively charged, and electrons find it increasingly difficult to pass from the low- to the high-work-function material. The charge builds up until the net transfer of charge is zero, and at this point the high-work-function material has a negative potential relative to the other material. As a rule the contact potential difference is not large enough to cause serious trouble in electronic tubes, but occasionally the effect must be considered in explaining their action, particularly when the element is sensitive to small changes of voltage as is the grid in a triode.

2-9. Thermionic Emission. When any solid is heated, an increase in the random motion of free electrons is one of the results. These electrons are moving in all directions in the material, and therefore some of them arrive at the surface with components of velocity normal to it. As the temperature is increased, they arrive with higher and higher velocities until a temperature may finally be reached where they have enough energy to free themselves from the boundary forces. This phenomenon is known as thermionic emission.

An equation relating the temperature of a solid to the emission current has been derived both theoretically and experimentally. It is

$$J_s = AT^2\epsilon^{-b/T} \tag{2-6}$$

and is known as Richardson's equation.* J_s is the emission-current density and is determined by the rate of escape of the electrons. A is a constant that is determined by the physical characteristics of the material and is most easily determined by experimental means. On the other hand, b, which is also a constant, can be determined theoretically and is found to be

$$b = \frac{W_w}{k} = \frac{-\phi_e Q_e}{k} \tag{2-7}$$

where k is Boltzmann's constant and is equal to 1.38×10^{-23} joule per °K. When numerical values are substituted for the constant quantities in Eq. (2-7),

$$b = 11{,}605\phi_e \tag{2-8}$$

where ϕ_e is the work function expressed in electron volts. Thus we see that the emission current obtainable from a given material is partially determined by the work function. The best emitters have low work functions.

Of equal and often greater importance than a low work function is the ability of a material to be heated to a high temperature without damage. Temperature, expressed in degrees Kelvin, appears both as T^2 and in the exponent of ϵ in Eq. (2-6). Thus a high enough value of T will offset the disadvantage of a high value of ϕ_e. Unfortunately all metals have melting points, some being so low that melting occurs before an efficient emission point is reached. Even though a material has a high melting point, it cannot be operated at a temperature anywhere near that value or it will begin to evaporate. The melting point of tungsten is 3643°K; yet a tungsten filament is operated at only about 2600°K. At temperatures above this the rate of evaporation becomes so high that the life of the filament is very short.

* Also known as Dushman's equation. Richardson and Dushman independently derived this form of the equation.

Another important characteristic of an emitter is that it must be mechanically strong at the temperature necessary for efficient operation. Otherwise it will break when the tube is subjected to mechanical shock or sustained vibrations.

2-10. Measurement of Thermionic Emission. To determine the emission constants for any particular metal, a tube with special construction is necessary. Figure 2-3 shows such a tube in which F is a filament along the axis of the cylindrical plate, or anode A. All filaments are cooled at the ends because of the lead wires, which are of large diameter in order that they will remain cool. Since the emission constants must be determined for a filament of uniform temperature, the emission from the cooled ends of the filament should not be measured. Therefore two guard rings are connected as shown.

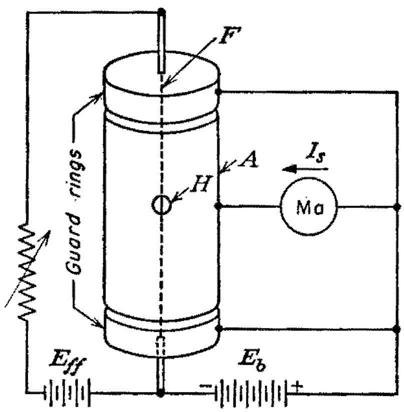


FIG. 2-3. Tube for determining the emission characteristics of a filament.

Emission from only that part of the filament with uniform temperature is measured, and any end effects due to variation in temperature or electric field are eliminated. The entire structure is mounted on a stem, sealed into a bulb, and evacuated. The temperature of the filament can be measured by means of an optical pyrometer sighted on it through the hole H. The emitted electrons are drawn to the anode by the field set up by the voltage E_b, which is made large enough to attract all those emitted. The emitted current I_s is measured by the milliammeter and is called the saturation current. From the measured values of I_s and the dimensions of the filament, the current density J_s may be found for any particular value of temperature T. In order to find the values of A and b from these data, it is most convenient first to write Richardson's equation in the following form:

$$\frac{J_s}{T^2} = A\epsilon^{-b/T} \tag{2-9}$$

Then write Eq. (2-9) in logarithmic form,

$$\log \frac{J_s}{T^2} = -0.434b\frac{1}{T} + \log A \tag{2-10}$$

the logarithms being to base 10. When the data are plotted on rectangular-coordinate paper using $\log (J_s/T^2)$ as the ordinate and $1/T$

as the abscissa, the result is a straight line with a slope of $-0.434b$ and an intercept on the axis of the ordinates equal to log A. Such a graph is shown in Fig. 2-4, where the constants for three very common types of emitters have been determined.

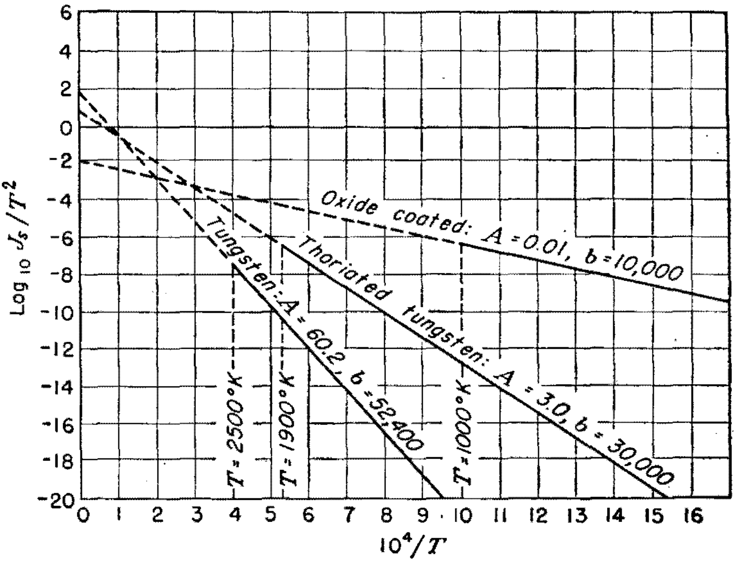


FIG. 2-4. Determination of constants in Richardson's equation.

Another method of depicting the emitting properties of thermionic cathodes is based on the Stefan-Boltzmann law of radiated power, which is

$$P = Ke_tT^4 \tag{2-11}$$

where P = power per unit area radiated by the body
T = temperature, °K
e_t = radiation emissivity, which depends upon the material and to some extent on the operating temperature
K = Stefan-Boltzmann constant, $(5.673 \pm 0.004) \times 10^{-8}$ watt per m² per °K⁴

If Richardson's equation and the Stefan-Boltzmann law of radiated power are combined, we can obtain an expression for the logarithm of current density in terms of the power radiated. This was done by Davisson, who devised a cross-section paper* in which he used a curvilinear axis in such a way that, when J_s is plotted as a function of P, the result is a straight line for any given emitter. Figure 2-5 shows such a plot for the same emitters used in Fig. 2-4.

* Manufactured and sold by the Keuffel and Esser Company.

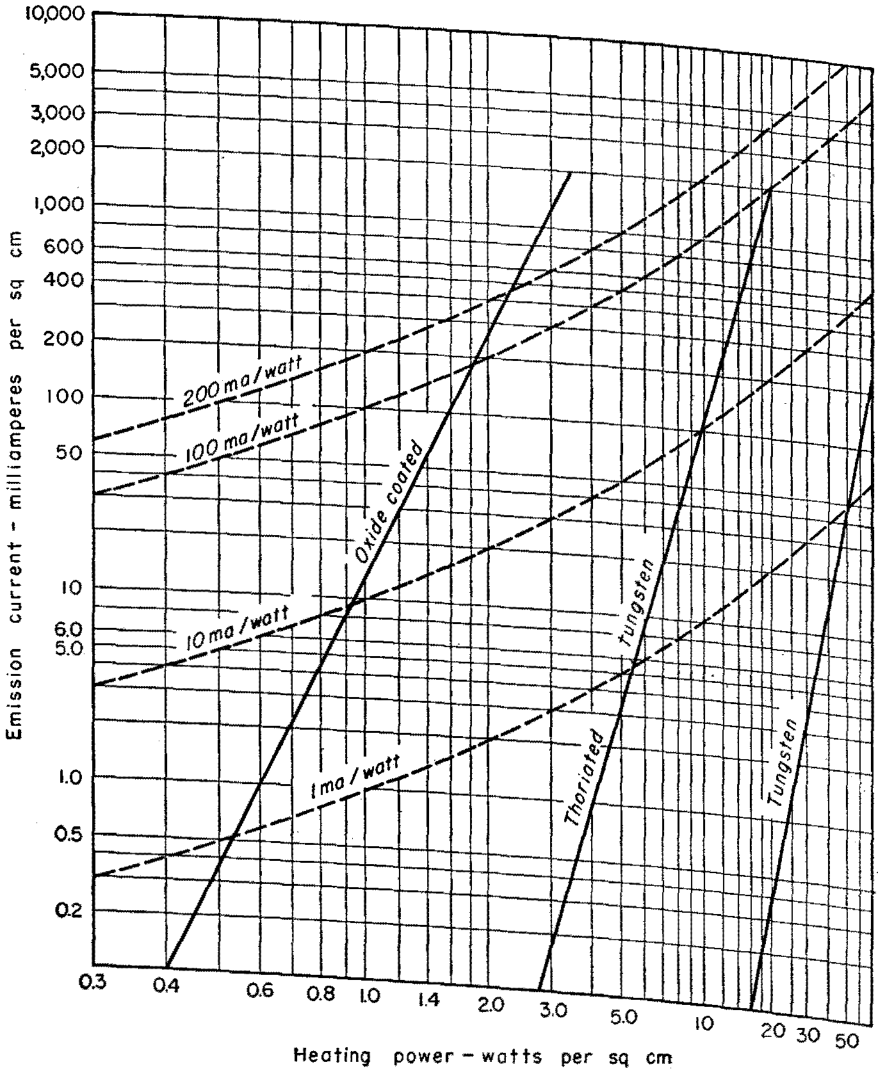


FIG. 2-5. Emission-current density as a function of the heating power for three types of emitters. The curves are plotted on "power-emission" paper.

This type of plot has the advantage that measurements of emission at a few low values of power may be made when the emission current from the cathode and the heating of the anode are small. The curves may then be extrapolated to give the emission at higher values of power where such prolonged operation might be injurious to the tube.

In telling how well an automobile performs we usually express the efficiency, not in per cent, but in miles per gallon of gasoline. Likewise in electronics we express the emission efficiency in terms of emission current per watt input to the filament. Thus to get emission efficiency from the power-emission chart, we have merely to divide the ordinate by the abscissa at any point.

2-11. Thermionic Emission from Tungsten.[2] Tungsten is about the only pure metal used as an emitter. Its melting point is given as 3643°K, and the emission current can be calculated for any temperature by means of Richardson's equation, where A is equal to 60.2 amp per cm^2 per °K^2 and b is equal to 52,400°K. The curves in Figs. 2-4 and 2-5 give the emission characteristics in different ways. In practice, when tungsten is used as a cathode, it takes the form of a wire bent in the shape of a hairpin or a spiral, and it is heated by passing a current through it. The practical life of a tungsten filament has been found to end when the cross-sectional area has decreased by 10 per cent, because burnouts occur soon thereafter. The higher the temperature of the filament, the greater the emission efficiency and the shorter the life. A study which includes economic considerations indicates that the temperature of the filament should be such that the cross-sectional area decreases 10 per cent in 1000 to 2000 hr.

Operating temperatures for tungsten filaments range from 2400 to 2600°K, with a corresponding range of emission efficiencies of 2 to 10 ma per watt. Compared with other commercial emitters tungsten has a very low emission efficiency. But it does have the advantage of being electrically rugged and will withstand positive-ion bombardment when extremely high voltages are applied to the tube. Mechanically it is fragile because it is very brittle. It is used chiefly in tubes where the plate voltages are to be above about 2500 volts, finding its greatest use in high-voltage rectifiers, power tubes, and X-ray tubes.

2-12. Thoriated-tungsten Filaments.[3] If during the manufacturing process about 1 per cent thorium oxide is added to the tungsten wire, then by proper heat-treatment we can form a thoriated-tungsten filament. This type of filament seems to be simply a tungsten wire on the surface of which has been formed a monatomic layer of pure thorium. However, the melting point of thorium is 2118°K, while that of tungsten is 3643°K. Since the layer of thorium must be maintained, this type of filament can be operated at only 1900°K compared with 2500°K for pure tungsten. Even so, because of its lower work function, the emission efficiency is very much higher than that for tungsten.

To form an activated thoriated-tungsten filament, the wire containing thorium oxide is formed or bent into the desired shape and mounted in the tube. Then during the evacuation process the filament is first heated to 2800°K for about 1 min. This reduces part of the thorium oxide to metallic thorium and oxygen, the latter being removed by means of the pumps exhausting the tube. The thorium thus formed is distributed through the entire cross section of the tungsten. Next the filament is operated at about 2100°K for approximately 15 min. At this temperature the thorium will diffuse from the interior of the filament toward the surface, where some of it will be evaporated and deposited on other elements or the walls of the tube. This evaporation is not desirable but is necessary in order to drive the thorium out to the surface. The evaporation occurs at a lower rate than the diffusion of the thorium to the surface, and as a result the filament gradually attains a thin and perhaps monatomic layer of thorium. The filament temperature is then reduced to about 1900°K, which is the operating temperature. It has been found that a layer of tungsten carbide on the surface of the filament lowers the rate of evaporation of thorium and makes possible operation at higher temperature with resultant increase in emission efficiency. Because of many variables the emission efficiency may range all the way from 50 to 100 ma per watt.

This type of emitter, like tungsten, is brittle. But unlike tungsten it cannot withstand much positive-ion bombardment. No matter how well a tube is exhausted, there is always some gas remaining. Hence there are some positive ions (molecules which have lost one or more electrons) present. These ions, being positive in polarity, are attracted to the filament. If the potential difference between plate or anode and filament is high enough, the positive ions will arrive at the filament with enough velocity to bombard and physically knock the thorium off the surface of the wire. Thoriated-tungsten filaments are used only in tubes where the plate voltage never exceeds 3000 to 4000 volts. Since they have emission efficiencies of 50 to 100 ma per watt, which is about 10 times that of pure tungsten, they are preferred in low-voltage transmitting tubes.

The life of the thoriated-tungsten filament depends on the amount of free thorium present in the filament. At the operating temperature there is a gradual evaporation of thorium from the surface and also a gradual diffusion of thorium from the interior to the surface. When all the free thorium has been driven out of the tungsten and evaporated off the surface, it will then act as a pure-tungsten-filament with its characteristic low emission efficiency, which at the operating temperature of a thoriated filament is extremely low.

2-13. Oxide-coated Emitters.[4] The Wehnelt, or oxide-coated, cathode was discovered in 1904 when Wehnelt found that if a platinum ribbon was coated with the oxides of certain rare earths a very efficient emitter was the result. However, no commercial use was made of this emitter until battery tubes came into common use in the early twenties.

The present-day oxide-coated cathode consists of a layer of rare-earth oxides coated on a core metal which may be nickel, konel, or a platinum alloy. The oxides are usually rather unstable in air; so the coatings are sprayed on in the form of finely ground barium and strontium carbonates in a thin collodion lacquer. After drying, the cathodes are mounted in the tube, which is then put on the vacuum system and evacuated. During the pumping process the cathode is usually heated to about 1500°K for a short time. At this high temperature the carbonates break down into the oxides plus carbon dioxide, which is removed by the pumps. Because of the very low work function these cathodes can be operated at about 1000°K, at which temperature they have very high emission efficiencies and long lives. The emission efficiency is on the order of 100 to 1000 ma per watt of heating power.

The greatest advantage of the oxide-coated cathode over other emitters is that it can be either a filament type or an indirectly heated cathode. In the filamentary type of cathode a ribbon or wire of the core metal is coated with the emitting material. Then a current is passed through it, which heats the oxide coating up to emitting temperature.

In the indirectly heated cathode the emitting material is usually sprayed onto a cylinder made of the core metal, usually nickel. Into this cylinder is inserted a heater, which may be made of tungsten wire coated with a porcelainlike cement insulating the heater from the cathode sleeve. We thus have an indirectly heated cathode in which the emitting surface is equipotential and which also isolates the emitter from the cyclic potential variations if alternating heating current is used. Figure 2-6 shows photographs of several types of directly and indirectly heated cathodes used in high-vacuum receiving tubes.

Principal disadvantages of the oxide-coated cathodes include the fact that the tubes cannot be evacuated to such a high degree as can tubes using other types of emitters. This is probably due to the incomplete breakdown of the carbonates into oxides. Therefore during use there is a further gradual breakdown of these materials. Also the oxide-coated cathodes are susceptible to positive-ion bombardment and hence are not ordinarily used in tubes in which the plate voltage exceeds 500 to 1000 volts. These two facts limit the uses of the oxide-coated cathodes chiefly to radio receiving tubes and, as we shall find later, gas-filled tubes. However, this is not a hard and fast rule. Cathode-ray tubes have oxide-coated indirectly heated cathodes, and here the anode voltage is of the order of 1500 to 2000 volts and in some cases may be as high as 27,000 volts. Oxide-coated cathodes may also be used in larger sizes of audio-frequency (a-f) power-amplifier tubes, but here the plate voltage does not usually exceed 1000 volts.

2-14. Schottky Effect. According to the discussion thus far, the total emission current from any emitter is given by Richardson's equation

$$J_s = AT^2\epsilon^{-b/T} \tag{2-6}$$

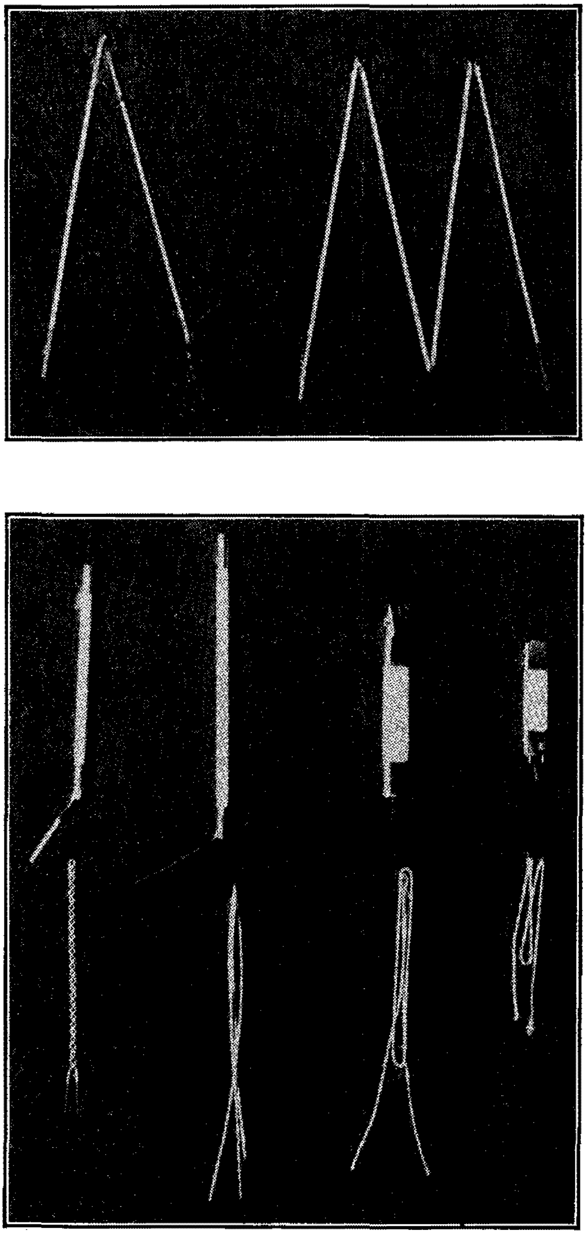


FIG. 2-6. Indirectly and directly heated cathodes used in high-vacuum receiving tubes.

which indicates that the emission current from any filament is independent of the anode voltage. In other words, no matter what voltage is applied to the anode, the total emission current should be constant at any given temperature. However, if tests are made on actual filaments, for

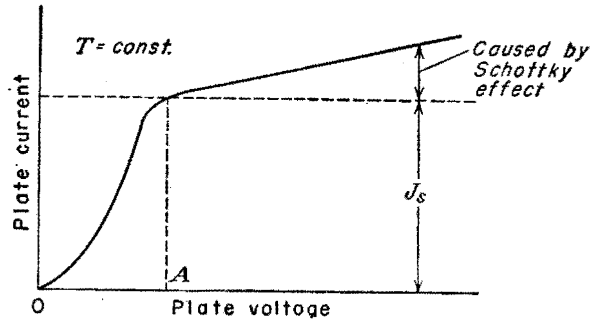


FIG. 2-7. A typical volt-ampere characteristic for a diode, showing the Schottky effect.

instance pure tungsten, a curve such as shown in Fig. 2-7 will be the result. The current for voltages between O and A is limited by space charge, which will be discussed in the next chapter. In the region beyond A the current should be a constant, given by Richardson's equation. Instead, we have a current which gradually increases with plate voltage. This phenomenon is known as the Schottky effect and is caused by additional energy which is supplied to the electrons from some source other than that used to heat the emitter. The potential difference between plate and cathode sets up an electric field which exerts a force on the cathode surface. This force is in such a direction as to help free electrons that otherwise would never escape from the boundary forces. Figure 2-8 shows the various forces acting on an electron as a function of the distance from the surface. The net force at point B is zero. To the left of point B it is in such a direction as to oppose the escape of an electron, and to the right it aids in the removal. Thus in order to be free an electron needs to have only enough energy to carry it beyond point B instead of to infinity. Since this force depends upon the magnitude of the plate voltage E_b, the

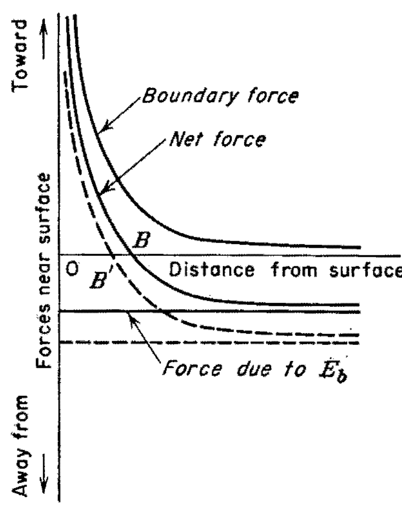


FIG. 2-8. Forces on an electron near an emitting surface which is in an electric field.

position of point B also depends upon E_b. As the value of E_b increases, point B approaches the surface; hence lower-energy electrons can escape, and J_s increases.

2-15. High-field Emission. If the potential difference between two electrodes in a high-vacuum tube is made increasingly larger, a point will finally be reached where electrons are emitted from the negative electrode even though it may be at 0°K. The field intensity required is of the order of magnitude of 10^{10} volts per m. This effect is known as high-field emission, and it differs from thermionic emission only in the way in which electrons are given the energy to escape the boundary forces. In thermionic emission they are given this energy by heating, while in high-field emission they are given the energy by the attractive force due to the external field.

High-field emission is found to be independent of the temperature of the emitter and follows the law

$$J_s = a\mathcal{E}^2\epsilon^{-B/\mathcal{E}} \tag{2-12}$$

where J_s = current density
a = a constant
$\mathcal{E}$ = field intensity at the surface of the negative electrode
B = a constant determined mainly by the work function of the material

Note that the current is independent of the temperature and depends only on the constants of the material and on the field intensity.

The field between a point and a plate or between a cylinder and a fine concentric wire may reach extremely high values with the application of a high voltage, particularly if the spacing is small. Thus in the design of high-voltage tubes, the shapes of the electrodes as well as the spacing must be considered in order to keep the potential gradient below a value at which field emission may result.

This type of emission is sometimes employed in certain kinds of cold-cathode devices and possibly may be the source of electrons in mercury arcs. Currents due to high-field emission may reach values of the order of thousands of amperes per square centimeter. These high currents undoubtedly would damage a tube unless it were designed to pass them.

2-16. Secondary Emission. It has been shown earlier in this chapter that when an electron acquires a certain amount of energy with velocity normal to and outward from the surface of the material, it will be emitted. Energy can be given to the electron by heating the material—this is thermionic emission. Energy given to the electron by an external field exerting a force on it may lead to high-field emission. Or energy may be given to the electron by the impact of another body, and this may cause secondary emission.

In traversing the interelectrode space in a vacuum tube, electrons may acquire a considerable amount of kinetic energy. Eventually they must strike an electrode or the walls of the tube and give up their energies. This may cause heating of the bombarded material; or if the impacts are with other electrons within the surface of the material, secondary emission may occur. Now, if there is an electrode within the tube which is at a higher potential than the bombarded material, the secondary electrons will be attracted to it and a current will flow.

Figure 2-9 shows a tube where the secondary-emission current can be measured. The electron gun, which is the source of the bombarding, or primary, electrons, gives them a velocity in such a direction that they strike the anode A_2. Upon impact, a primary electron may give up its energy to one or more electrons within the surface of A_2, causing them to be emitted. They are then attracted to a collector anode A_3, which is 30 to

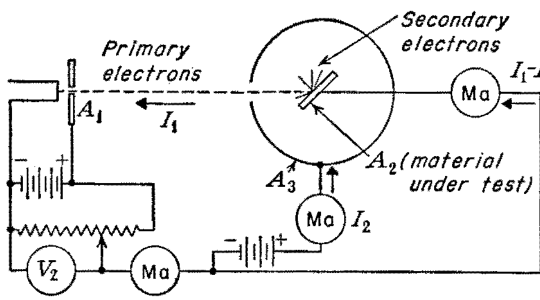


FIG. 2-9. Tube for measuring secondary-emission effects.

50 volts positive with respect to A_2. The meters shown in the circuit indicate the magnitudes and directions of the various currents which flow. The primary current can be obtained from a reading of the milliammeter in the cathode circuit. The secondary current can be read from the meter in the collector-anode (A_3) circuit. The current flowing in the second-anode circuit is the difference between the primary and the secondary currents. Since it is customary to signify as positive a current which flows from the external circuit to the electrode and thence across the vacuum, the primary current in the second-anode circuit is positive and the secondary current is negative. Thus the second-anode current may be positive, negative, or zero depending upon the relative number of primary and secondary electrons. If the number of primaries exceeds the number of secondaries, the current is positive, while if the secondaries exceed the primaries, it is negative. Note that the secondary-emission current is positive in the collector-anode (A_3) circuit.

If several such tubes are constructed which are identical in all respects except that each uses a different material for the surface of A_2, it is found that the ratio of secondary to primary current increases with decreasing work function of the material. It is also found that the ratio is dependent upon the potential of A_2. In fact this ratio may reach a value as high as 10 upon using very low work-function materials and the proper voltages.

Figure 2-10 shows how the secondary-emission ratio varies with the value of the voltage applied to A_2, which is also known as the accelerating voltage. From these curves we see that, when the accelerating voltage is low, the current to the second anode is positive since the ratio I_2/I_1 is less than unity. As the voltage is increased the secondary-to-primary ratio increases and may reach a value greater than unity. This means that there are more secondaries than primaries, and hence the current to

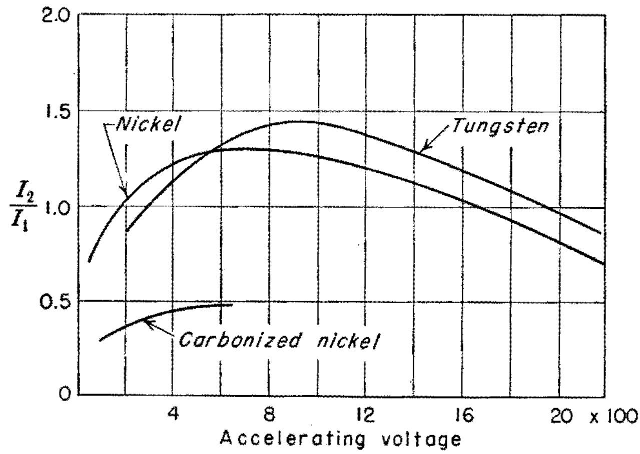


FIG. 2-10. Secondary-emission characteristics for three common materials used in vacuum tubes. [*From J. H. O. Harries, Secondary Electron Radiation, Electronics,* **17,** 100–108, 180 (*September,* 1944), *by permission.*]

the second anode reverses and becomes negative. At low accelerating voltages the bombarding electrons have very little energy and hence can knock out only a few of the higher-energy electrons in the surface. As the accelerating voltage is increased, the bombarding electrons arrive at the surface with more energy and hence remove more electrons. Finally a point is reached where the bombarding electrons arrive with such high velocity that they simply bury themselves deep in the material of the anode. Hence the secondary-emission ratio begins to fall off.

Secondary emission can occur, not only from conducting materials, but from insulators as well. It can take place if a stream of electrons strikes the wall of a glass tube with high enough velocity. It may even cause physical damage to the tube due to the heating of the bombarded surface. If the electrons arrive at the glass with sufficient energy to cause a secondary-to-primary ratio greater than unity, the wall is left with a net positive charge. The secondary electrons are attracted to the anode since it is the point of highest potential. The positive charge on the glass accelerates the primary electrons, which means that they arrive at the glass with still higher velocity. This in turn means that the secondary-to-primary ratio becomes larger, and hence more secondaries are emitted and the glass becomes even more positive. The effect is accumulative and may build up until the glass becomes so hot that it melts, allowing air to rush in and destroy the vacuum. In commercial tubes the electrodes are usually designed so that they shield the glass from the electron stream and thus prevent this type of damage.

2-17. Photoelectric Emission. The fourth process for obtaining free electrons from a solid body is known as photoelectric emission. As the name implies, light is used to give electrons enough added energy to free themselves from the boundary forces.

The law of conservation of energy states that energy cannot be destroyed. Therefore when light, which is a form of energy, falls on the surface of a solid, it is either absorbed by the surface or reflected by it. A portion of the energy absorbed will be supplied to some of the electrons which are free within the body but unable to escape without external assistance. If the work function of the surface is low enough, the electrons now have sufficient energy to free themselves from these boundary forces and photoelectric emission occurs. If an external electric field is present, a current flows which is proportional to the amount of light falling on the surface.

A more complete discussion of photoelectric emission will be found in a later chapter.

REFERENCES

1. Richtmyer, F. K., and E. H. Kennard: "Introduction to Modern Physics," McGraw-Hill Book Company, Inc., New York, 1947.
2. Jones, H. A., and I. Langmuir: Characteristics of Tungsten Filaments as Functions of Temperature, *Gen. Elec. Rev.*, **30**, 310, 354, 408 (1927).
3. Langmuir, I.: The Electron Emission for Thoriated-tungsten Filaments, *Phys. Rev.*, **22**, 357 (1923).
4. Blewett, J. P.: Properties of Oxide-coated Cathodes, *J. Applied Phys.*, **10**, 668 (1939).
5. E. E. Staff, MIT: "Applied Electronics," John Wiley & Sons, Inc., New York, 1943.
6. Millman, J., and S. Seely: "Electronics," 2d ed., McGraw-Hill Book Company, Inc., New York, 1951.
7. Ryder, J. D.: "Electronic Fundamentals and Applications," Prentice-Hall, Inc., New York, 1950.

PROBLEMS AND QUESTIONS

1. Of the processes which can produce free electrons in the space of an electronic tube, which supply these electrons directly from the emitter?

2. Are the higher-energy electrons in an atom close to or far away from the nucleus?

3. Explain why there are free electrons in a metal.

4. A pure-tungsten wire 5 cm long and 0.0127 cm in diameter is operated at 2500°K temperature. A is 60.2 amp per cm^2 per °K^2, and b is 52,400°K. Find the total saturation current for the wire.

5. (*a*) An electron moves from a point of zero potential to one having a potential of +100 volts. How much kinetic energy, expressed in electron-volts, does it gain? (*b*) A double-charged positive ion replaces the electron in part *a*. How many electron-volts of potential energy are now gained?

6. With Eqs. (2-6) and (2-11) known, find an expression for saturation-current density for a tube in terms of filament power per unit area and filament and other constants [the constants in Eqs. (2-6) and (2-11) are assumed known].

7. Can a large-diameter filament be operated at a higher temperature than one with a smaller diameter and yet have the same length of life? Explain.

8. State the approximate emission efficiencies and the chief applications of each of the three types of emitters.

9. Explain the difference between the Schottky effect and high-field emission.

10. In a diode what happens to the secondary electrons driven off the plate?

CHAPTER 3

THE HIGH-VACUUM DIODE

3-1. Space Charge. Chapter 1 dealt with the effects of uniform electric fields on the motion of single electrons. This method is applicable to situations in which there are relatively few electrons present, such as in a cathode-ray tube. But in the great majority of electronic tubes we seldom have electrons in such small numbers that we can ignore the effects of neighboring electrons. For instance, if a tube is passing a current of 1 ma, there are approximately 6.25×10^{15} electrons traveling from the cathode to the anode each second. This means that at any instant there must be billions of electrons present in the interelectrode space, each carrying a charge of 1.60×10^{-19} coulomb and each exerting a force on neighboring electrons. Thus we see that the field through which an electron must pass in its flight from cathode to anode is due not only to the potential difference between the electrodes but also to the charges carried by all the electrons in the interelectrode space. The field can no longer be considered uniform, and hence the forces on and the acceleration of an electron depend upon its position in the field.

The only reason why it is possible for the plate voltage to control the magnitude of the current flowing through a tube is this effect of the space charge. Without space charge the saturation current of the filament would always flow. This current depends only on the material and the temperature of the emitter and not upon the plate voltage of the tube. Thus the only way in which we could control the current would be by changing the temperature of the cathode.

3-2. The Diode. The thermionic diode is the simplest electronic tube which can be constructed, and its action is the easiest to understand. It consists of either a filamentary or an indirectly heated cathode, and an anode or plate, which usually surrounds the cathode. Figure 3-1 shows two radio-receiver-type diodes. On the right is a type 80 tube which is really two diodes in a single envelope. The two filaments are connected in series, and the two plate leads, or connections, are brought out separately. In this figure one of the plates has been removed and is lying beside the tube, in order to show the filament. On the left is a type 5V4, which is also a double diode. Here again the plate has been removed in order to show the indirectly heated cathode.

All of the smaller diodes, such as those used in radios, have oxide-coated cathodes. This is simply because of the desirable high emission efficiency of the oxide-coated cathode and because the positive-ion bombardment is

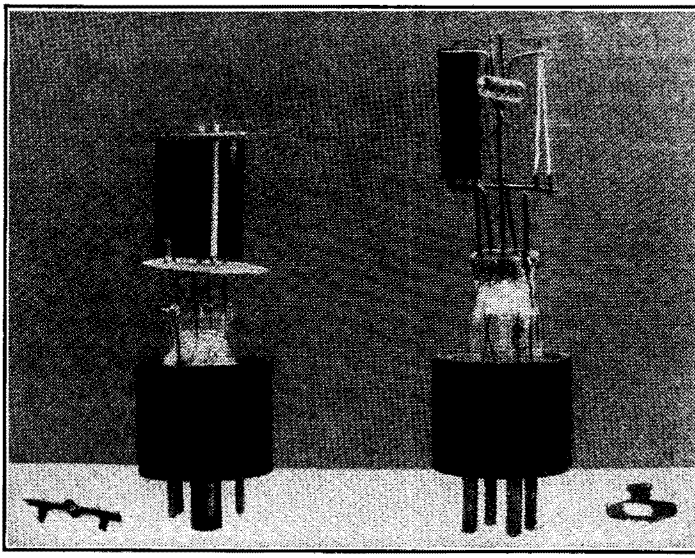


FIG. 3-1. Radio-receiver-type diodes.

not serious at the low applied voltages. Large rectifiers, such as are used in transmitters, may use thoriated-tungsten or pure-tungsten filaments, the choice depending on the voltages applied to the plates.

3-3. Experimental Determination of the Current in a Diode. If a diode with a pure-tungsten filament is connected as shown in Fig. 3-2 and data are taken for curves of plate current vs. plate voltage, for various filament temperatures, curves such as those shown in Fig. 3-3 will be the result. Here it can be seen that at low voltages the plate current is independent of the temperature of the filament but is dependent on the plate voltage. At higher voltages the plate current becomes independent of the plate voltage but does become dependent on the temperature of the emitter. Each curve may be divided into three regions. From A to B current flows even though the plate

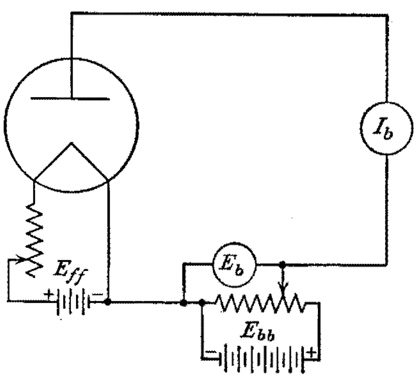


FIG. 3-2. Circuit for determining the characteristics of a diode.

is at a negative or zero potential. This is because some of the electrons are emitted with enough energy to carry them over to the plate even though the field is in such a direction as to oppose this motion. From B to C the plate current increases with plate voltage but is less than the total emission current for the filament as calculated from Richardson's equation. The region from A to C is known as the space-charge-limited region and is the portion which we shall study in this chapter. From C to D all the electrons emitted are pulled over to the plate as fast as they are emitted. This means that there is very little space charge present between the electrodes, and hence the current can be approximated from Richardson's equation.

Actually, when the voltage is increased above the values at C required for saturation, the current increases very slowly. This is the Schottky

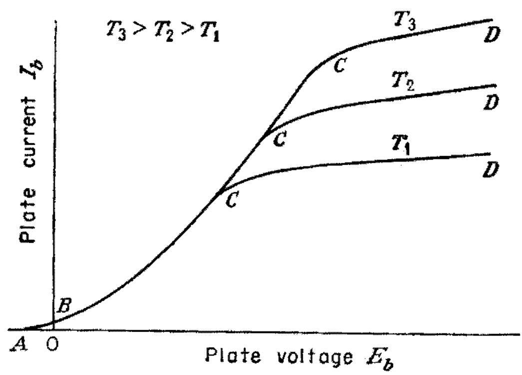


FIG. 3-3. Characteristics for a high-vacuum diode.

effect and is due to the external field helping lower-energy electrons to free themselves from the forces tending to hold them in the cathode. Pure tungsten shows this Schottky effect very nicely, but thoriated-tungsten and oxide-coated cathodes show too much increase in plate current, with increasing plate voltage, after saturation supposedly has been reached. In fact these two materials show very little tendency to saturate.

3-4. Potential Distribution in a Diode with Parallel Plane Electrodes. For the sake of simplicity all of our derivation and most of our discussion in this section will be on a diode with parallel, plane, infinitely large electrodes. This eliminates the effects of fringing at the edges of the electrodes and also allows us to work with an electric field which is relatively simple to describe mathematically.

If a tube were built with large parallel plane electrodes and with the negative electrode arranged so that it could be heated to such a temperature that it would emit large numbers of electrons, some very interesting data could be obtained. Suppose we pretend that such a tube has been constructed and theorize on its properties. A schematic diagram for the tube and circuit is shown in Fig. 3-4, where the means of heating the cathode has been omitted. Now if the plate potential is held constant at some value, say 100 volts, and the temperature of the cathode is increased, a point finally is reached where we have a measurable plate current flowing. If, with this small current flowing, we vary the value of E_{bb}, we find that it has very little effect on the plate current as long as the voltage is greater than a certain value. This simply means that there are so few electrons in the interelectrode space near the cathode that the plate current is determined solely by the emission of the cathode. If

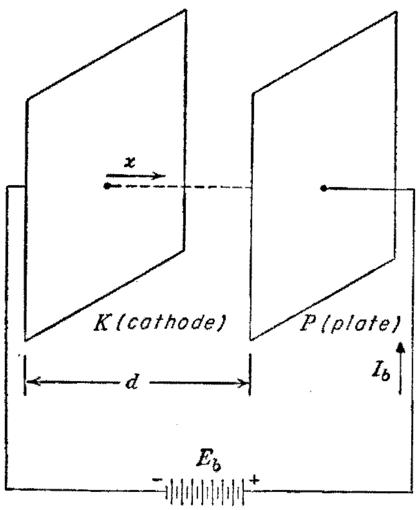


FIG. 3-4. Diode with parallel plane electrodes.

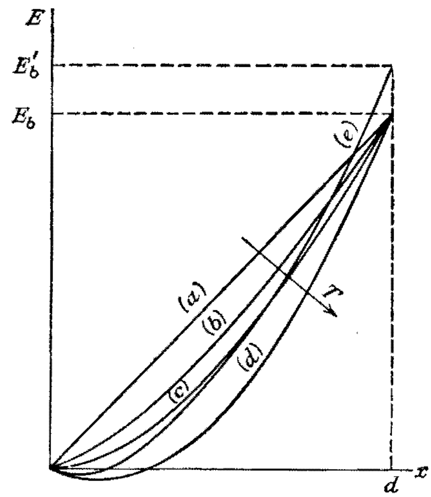


FIG. 3-5. Potential-distribution curves for a high-vacuum diode with parallel plane electrodes.

under these conditions the potentials of all points on a line normal to and between the two electrodes are plotted as a function of the distance of the point from the cathode, an approximately straight line, such as shown by curve a in Fig. 3-5, is the result. This shows that the potential gradient or slope is constant at all points and hence there is a uniform field.

If the temperature of the cathode is now increased to a value T_2, we may notice that the plate current increases but also that the plate voltage has some effect on the magnitude of the plate current. If once more the potentials of all points on the line normal to the electrodes are plotted on the same set of axes as before, curve b of Fig. 3-5 is the result. Here we see that the potential for any given point, except for the end points, is less than it was when the temperature of the cathode was lower. The reason for this is that at the lower temperature T_1 there are so few electrons in the interelectrode space that they have very little effect on the field configuration. Now at the higher temperature there are many more electrons in the space, and since the plates set up a field such as to attract the electrons toward the plate, while all the electrons carry a charge of such polarity as to repel other electrons as they are emitted, the net result is a lower potential at all intermediate points. At this time it might be well to recall that the force on an electron is equal to the charge on the electron multiplied by the negative of the potential gradient. If we once more examine curve b of Fig. 3-5, we see that the potential gradient near the cathode is less than for the condition of no space charge (curve a). However, close to the anode the gradient of curve b is greater than the gradient of curve a, which means that the force on an electron is greater than with no space charge present. This seems logical because there are more electrons between a given electron (close to the plate) and the cathode than there are between the electron and the plate. Therefore the net force of repulsion near the plate is in such a direction as to assist the field set up by the electrodes.

If the temperature of the cathode is raised to a value T_3, we find that the potential distribution is similar to that shown in curve c of Fig. 3-5. Here we note that the potential gradient is zero at the surface of the cathode. This means that the force of attraction to the plate on an electron at the cathode surface is just canceled by the forces of repulsion of the electrons in the interelectrode space. Therefore, an electron emitted with zero velocity arrives at the surface of the emitter with zero force exerted on it and probably returns to the cathode. When an electron in the interelectrode space arrives at the plate, the potential distribution shifts to a position somewhat like that of curve b. The resulting gradient near the cathode surface now allows another electron to escape to the interelectrode space. This now shifts the distribution temporarily back to curve c. We then have what is known as space-charge-limited conditions.

If the potential distribution in an actual diode could be plotted, it probably would be found to be more like curve d of Fig. 3-5. Here the potential actually becomes negative and goes through a minimum a short distance from the cathode. This is because of the initial velocity with which electrons are emitted. Referring to curve c, suppose an electron is emitted with a finite velocity from the surface of the cathode. Initially there would be no force on it, since the gradient is zero; but it would pass out into interelectrode space, where the force on it would be in a direction to pull it toward the plate. Another electron emitted just behind the first, however, would find a small force of repulsion trying to make it return to the cathode. This is because that extra electron in the interelectrode space makes the space charge large enough so that it more than cancels the field due to the electrodes. Therefore, the field at the surface of the cathode is slightly negative, and the potential minimum moves a short distance out. Since millions of electrons are actually emitted simultaneously, this effect is vastly magnified and the minimum-potential point takes a position as shown in curve *d*. Unless emitted electrons have enough velocity to carry them out past this minimum, they will be returned to the cathode. Thus only the higher-velocity electrons can get across the space. The current will be less than the saturation value since it is limited by space charge. As the potential of the plate is changed, the number of electrons in the interelectrode space will vary. Curve *e* shows the possible potential distribution if E_b is increased. The point of zero gradient moves to the left, and the negative gradient near the cathode is decreased. Hence more electrons escape into the interelectrode space. Thus different values of space-charge-limited currents result. It is this effect which we shall now study and for which we shall derive mathematical relationships. But first it might be well to review some of the fundamentals of electric fields.

3-5. Gauss' Theorem.[1] In the mks system we write Coulomb's law for two point charges in a vacuum as

$$f = \frac{q_1 q_2}{4\pi\epsilon_0 r^2} \tag{3-1}$$

where f is the force (in newtons) between the two charges q_1 and q_2 (in coulombs), r is the distance between them (in meters), and ϵ_0 is the permittivity of free space.

The region in the neighborhood of a electric charge is said to contain an electric field, and in order to measure the intensity of that field at a point, we measure the force per unit charge on a small positive test charge placed at the point. Hence from Eq. (3-1), the field intensity at a point a distance r from a charge q is given by

$$\mathcal{E} = \frac{q}{4\pi\epsilon_0 r^2} \tag{3-2}$$

When more than one charged body is present, the intensity can be found by determining the vector sum of the intensities due to each individual charge; thus

$$\mathcal{E}_{\text{total}} = \vec{\sum_n} \frac{q_n}{4\pi\epsilon_0 r_n^2} \tag{3-3}$$

Plotting an electric field was discussed in Art. 1-6. If we let one line of flux per unit area represent one unit of field intensity, a flux pattern will indicate the strength as well as the direction of the field at any point.

Referring to Eq. (3-2), we see that the field intensity varies as the reciprocal of the square of the distance from a charge q. If now we enclose this charge by any surface, all the flux emanating from the charge passes through the surface. Therefore, if we circumscribe a sphere of radius r about the charge, the total flux emanating from the charge can be calculated by multiplying the field intensity, which is numerically equal to the flux density, by the area of the sphere. Thus from Eq. (3-2) we get

$$\text{Total flux} = \mathcal{E}4\pi r^2 = \frac{q}{\epsilon_0} \tag{3-4}$$

This is a mathematical expression of Gauss' theorem, which states that the total flux emanating from an electric charge in a medium is equal to the magnitude of that charge divided by the permittivity of the medium.

3-6. The Space-charge Equation.[2] A mathematical expression which gives the relationship between space-charge-limited current and the plate voltage of a diode was derived independently by Child in 1911 and by Langmuir in 1913. This relationship is now known as the Child-Langmuir equation or the three-halves-power law.

In order to derive this relationship for the simplest possible electrode configuration, the following assumptions are made:

1. The electrodes are infinite, parallel, plane, equipotential surfaces.
2. The zero-potential electrode, or cathode, is heated to a high enough temperature so that the supply of emitted electrons is greater than the demand.
3. The electrons start from rest at the surface of the cathode.
4. The anode is maintained at a constant positive potential E_b.
5. Only electrons are present in the space between electrodes.

If the two electrodes are placed so that the x axis of the coordinate system is normal to the two surfaces as shown in Fig. 3-6, the field can be represented by a vector whose direction is from right to left as shown. If no charge is present in the interelectrode space, the field is uniform. However, if there is a cloud of electrons between the plates, the field will be due not only to the potential difference between the electrodes but also to the charges on all the electrons. The field is no longer uniform but is still from right to left. The y and z components of the field equal zero since the electrodes are assumed to be infinite in extent. Assume a small element of volume Δx, Δy, Δz. The field-intensity vector varies in magnitude at different points between the electrodes.

According to the definition of flux density, the total flux entering the small volume through the right face equals the field intensity multiplied by the area of the face. Hence

$$\text{Flux entering} = \mathcal{E}\,\Delta y\,\Delta z \tag{3-5}$$

The flux leaving the left face is given by

$$\text{Flux leaving} = \left(\mathcal{E} + \frac{d\mathcal{E}}{dx}\,\Delta x\right)\Delta y\,\Delta z \tag{3-6}$$

where $d\mathcal{E}/dx$ is the rate of change of field intensity with distance. The difference in flux entering and leaving the small volume is found by subtracting Eq. (3-5) from Eq. (3-6).

$$\text{Difference in flux} = \frac{d\mathcal{E}}{dx}\,\Delta x\,\Delta y\,\Delta z \tag{3-7}$$

This difference must be due to flux emanating from the charge enclosed by the surface of the small element of volume.

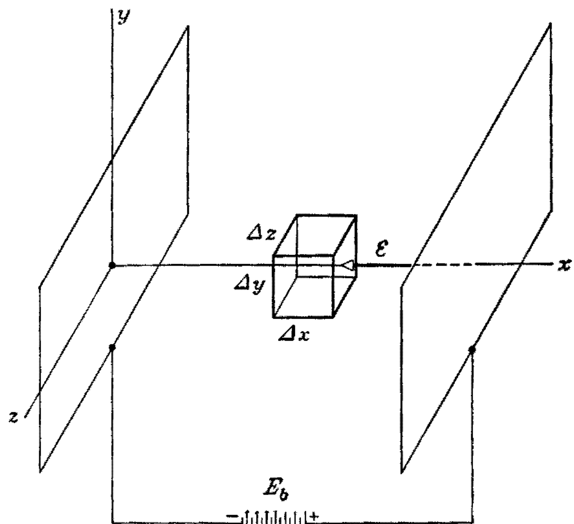


FIG. 3-6. Electrode configuration and the element of volume used in the development of Poisson's equation.

Since Gauss' theorem states that the total flux emanating from an enclosed charge is equal to the magnitude of the total charge divided by the permittivity of the medium, we may write

$$\text{Flux from enclosed charge} = \frac{\rho}{\epsilon_0}\,\Delta x\,\Delta y\,\Delta z \tag{3-8}$$

where $\rho\,\Delta x\,\Delta y\,\Delta z$ is the total enclosed charge, ρ being the charge density.

Now since the difference in flux given by Eq. (3-7) is due to the enclosed charge, we may equate the right members of Eqs. (3-7) and (3-8), obtaining the relation

$$\frac{d\mathcal{E}}{dx} = \frac{\rho}{\epsilon_0} \tag{3-9}$$

It was previously proven that the field intensity is equal to the negative of the potential gradient, or

$$\mathcal{E} = -\frac{dE}{dx} \tag{3-10}$$

and hence

$$\frac{d\mathcal{E}}{dx} = -\frac{d^2E}{dx^2} \tag{3-11}$$

Substituting this into Eq. (3-9) yields

$$\frac{d^2E}{dx^2} = -\frac{\rho}{\epsilon_0} \tag{3-12}$$

This is known as Poisson's equation in one dimension.

What we are trying to do is to find a relationship between anode voltage and the space current in the tube. Therefore, if we can find a relationship between charge density and current, Eq. (3-12) should lead to the desired result.

If the velocity of the electrons is designated by v, the density J_b of the current flowing between two electrodes can be written as

$$J_b = -\rho v \tag{3-13}$$

The negative sign is used because the direction of positive current flow is taken from anode to cathode and ρ is negative, while v is positive since it is from left to right in the direction of the positive x axis.

If we solve Eq. (3-13) for ρ and substitute the result in Eq. (3-12), we have

$$\frac{d^2E}{dx^2} = \frac{J_b}{\epsilon_0 v} \tag{3-14}$$

Now, since the velocity of an electron in the electric field is given by

$$v = \sqrt{-2\frac{Q_e}{m_e}E} \tag{3-15}$$

we may substitute into Eq. (3-14) and obtain the relation

$$\frac{d^2E}{dx^2} = \frac{J_b}{\epsilon_0}\sqrt{\frac{-m_e}{2Q_e}}\,E^{-1/2} \tag{3-16}$$

A first integration is achieved by multiplying both members of Eq. (3-16) by $2(dE/dx)\,dx$ and then integrating.

$$2\int \frac{dE}{dx}\frac{d^2E}{dx^2}\,dx = 2\frac{J_b}{\epsilon_0}\sqrt{\frac{-m_e}{2Q_e}}\int E^{-1/2}\frac{dE}{dx}\,dx \tag{3-17}$$

The current density J_b appears outside the integral because it is constant for all values of x. If it were variable, the rate of flow of charge would be different at different points in the tube; this would mean a building up of charge. Now if we note that $\frac{d}{dx}\left(\frac{dE}{dx}\right)^2 = 2\frac{dE}{dx}\frac{d^2E}{dx^2}$ and $\frac{dE}{dx}dx = dE$, then we can write Eq. (3-17) in the form

$$\int \frac{d}{dx}\left(\frac{dE}{dx}\right)^2 dx = 2\frac{J_b}{\epsilon_0}\sqrt{\frac{-m_e}{2Q_e}}\int E^{-\frac{1}{2}}\,dE \qquad (3\text{-}18)$$

or

$$\left(\frac{dE}{dx}\right)^2 = 4\frac{J_b}{\epsilon_0}\sqrt{\frac{-m_e}{2Q_e}}\,E^{\frac{1}{2}} + C_1 \qquad (3\text{-}19)$$

The boundary conditions are such that when x is zero the potential E is likewise zero, and since for space-charge-limited current the electrons at the surface of the cathode must have zero force exerted on them, the potential gradient at the surface is likewise zero. Thus C_1 is found to be zero. Hence

$$\frac{dE}{dx} = 2\sqrt{\frac{J_b}{\epsilon_0}}\left(\frac{-m_e}{2Q_e}\right)^{\frac{1}{4}} E^{\frac{1}{4}} \qquad (3\text{-}20)$$

In Eq. (3-20) the variables may be separated and the result integrated,

$$\int E^{-\frac{1}{4}}\,dE = 2\sqrt{\frac{J_b}{\epsilon_0}}\left(\frac{-m_e}{2Q_e}\right)^{\frac{1}{4}}\int dx \qquad (3\text{-}21)$$

whence

$$\frac{4}{3}E^{\frac{3}{4}} = 2\sqrt{\frac{J_b}{\epsilon_0}}\left(\frac{-m_e}{2Q_e}\right)^{\frac{1}{4}} x + C_2 \qquad (3\text{-}22)$$

But since E is zero when x equals zero, C_2 is also zero. Solving Eq. (3-22) for J_b yields

$$J_b = \frac{4\epsilon_0}{9\sqrt{-m_e/2Q_e}}\,\frac{E^{\frac{3}{2}}}{x^2} \qquad \text{amp per unit area} \qquad (3\text{-}23)$$

Substituting numerical values for the constants,

$$J_b = \frac{2.33 \times 10^{-6}E^{\frac{3}{2}}}{x^2} \qquad \text{amp per unit area} \qquad (3\text{-}24)$$

Since this equation holds for all values of x, it holds for x equal to d, the distance between cathode and plate, where E is equal to E_b. Therefore,

$$J_b = \frac{2.33 \times 10^{-6}}{d^2}\,E_b^{\frac{3}{2}} \qquad \text{amp per unit area} \qquad (3\text{-}25)$$

For a given area A of the plate we may write

$$I_b = \frac{2.33 \times 10^{-6}}{d^2} A E_b^{3/2} \qquad \text{amp} \tag{3-26}$$

In this equation all quantities are constant except the plate voltage and the plate current. In other words the plate current is proportional to the three-halves power of the plate voltage, or

$$I_b = kE_b^{3/2} \tag{3-27}$$

where $k = [(2.33 \times 10^{-6})/d^2]A$. This is the Child-Langmuir law for parallel plane electrodes.

A similar relationship may be derived for the case where the anode and cathode are concentric cylinders. The result of such a derivation* is

$$I_b = 14.68 \times 10^{-6} \frac{L}{r_a} \frac{E_b^{3/2}}{\beta^2} \qquad \text{amp} \tag{3-28}$$

where I_b is the plate current, E_b is the plate potential in volts, r_a is the radius of the anode, L is the length of the cathode or of the anode, whichever is the shorter, β^2 is a number depending on the ratio r_a/r_k, and r_k is the radius of the cathode. Table 3-1 gives the values of β^2 for various values of r_a/r_k. If r_a is large compared with r_k, it may be seen that β^2 may often be taken equal to unity without introducing appreciable error.

TABLE 3-1
VALUES* OF β^2 FOR VARIOUS VALUES OF r_a/r_k

r_a/r_k	β^2	r_a/r_k	β^2	r_a/r_k	β^2
1.0	0.000	3.0	0.517	7.0	0.887
1.5	0.119	4.0	0.667	8.0	0.925
2.0	0.279	5.0	0.767	9.0	0.955
2.5	0.412	6.0	0.836	10.0	0.978

* Data are from I. Langmuir, Currents Limited by Space Charge between Coaxial Cylinders, *Phys. Rev.*, **22,** 353 (1923), Table III, with permission.

Equation (3-28) shows that in this case, also, the plate current is proportional to the three-halves power of the plate voltage. In fact it can be shown that no matter what the shape of the electrodes, the three-halves power holds, the only difference being the constant of proportionality, which is determined by the geometry of the electrodes.

3-7. Experimental Proof of the Child-Langmuir Equation. In the above proof it was assumed that the cathode is an equipotential surface

* I. Langmuir and K. T. Compton: Electrical Discharges in Gases, Part II, Fundamental Phenomena in Electrical Discharges, *Rev. Mod. Phys.*, **3,** 245–248 (1931).

and that the electrons have zero velocity at the surface of the cathode. In actual practice these conditions are seldom fulfilled. The cathode cannot be an equipotential surface unless it is the indirectly heated type, while most rectifiers are of the filamentary type. Neither are the electrons emitted with zero velocity.

If the assumption is made that a diode follows a power law, but not necessarily with the three-halves exponent, we can write

$$I_b = kE_b^n \tag{3-29}$$

If this is written in logarithmic form, we have

$$\log I_b = \log k + n \log E_b \tag{3-30}$$

This is the equation for a straight line if $\log I_b$ is plotted as the ordinate and $\log E_b$ as the abscissa in a rectangular-coordinate system.

The constants k and n for an actual tube may be evaluated as follows: The circuit of Fig. 3-7 is used, and values of I_b as a function of E_b are obtained while E_{ff} is held constant. Note that the negative end of the plate battery is connected to a center tap on the filament battery. This is done in order to counteract or minimize the effects of the potential gradient along the filament. The data thus taken are plotted, as shown in Fig. 3-8, and a straight line is obtained. If the locus is extrapolated to the point where $\log E_b = 0$, this value of I_b is equal to k. The slope of the line equals the constant n, and its value will be in the neighborhood of 1.5. If, while taking the above-mentioned data, E_b is set equal to zero, it will be noted that an appreciable current still flows in the plate circuit. This current is due to the initial velocity of the electrons, contact potential, and the effects of the potential gradient along the filament. This, of course, introduces factors which we have not accounted for in our derivation. If an indirectly heated cathode is used, then these factors are reduced to two in number.

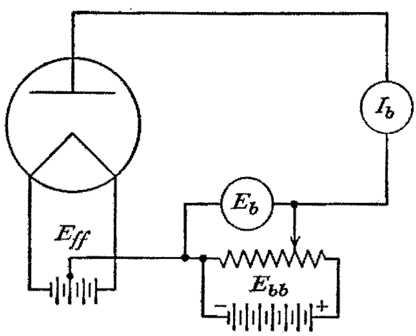


FIG. 3-7. Circuit used to evaluate the constants of the Child-Langmuir equation.

3-8. Power Loss in the Diode. Each electron arriving at the plate in a diode has an energy which, neglecting the initial or emission velocity, is given by

$$W = -Q_eE_b \qquad \text{joules} \tag{3-31}$$

When this electron is stopped by impact with the plate, its kinetic energy is converted into heat. If n electrons arrive at the plate in a time t, then the power delivered to the plate is given by

$$P = \frac{W}{t} = \frac{-Q_e n E_b}{t} \qquad \text{watts} \tag{3-32}$$

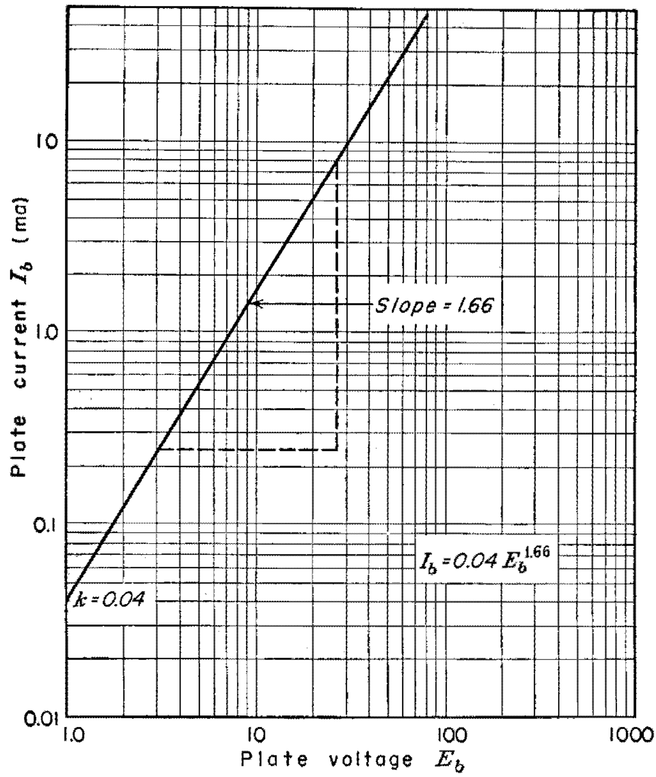


FIG. 3-8. Curve used to determine the coefficients in the Child-Langmuir equation.

The term $(-Q_e n)/t$ is the rate of flow of charge and therefore is the current I_b flowing; hence Eq. (3-32) may be written

$$P = I_b E_b \qquad \text{watts} \tag{3-33}$$

Thus in any tube part of the power which the plate must be able to dissipate can be calculated by Eq. (3-33). In addition the anode absorbs some of the heat radiated from the filament. All this power must be removed largely by radiation; hence the anode or plate material should be a good radiator. In a receiving tube the plate is usually made of nickel, and in order to improve the heat-radiating ability it is sometimes coated with carbon black. This plate must be held to temperatures below the point where it shows any color, otherwise it may release some occluded gas and thus spoil the vacuum. In low-power radiation-cooled transmitting tubes the plates may be made of tantalum, molybdenum, or graphite. Tantalum has the distinct advantage of being able to absorb gases when heated to a red heat; in other words, it acts as a getter. Although it is a relatively poor radiator, it does have a very high melting point and can be operated safely at a red heat; thus it acts as a getter without evaporation of the plate material. Molybdenum can be operated at a high temperature and is a better radiator than

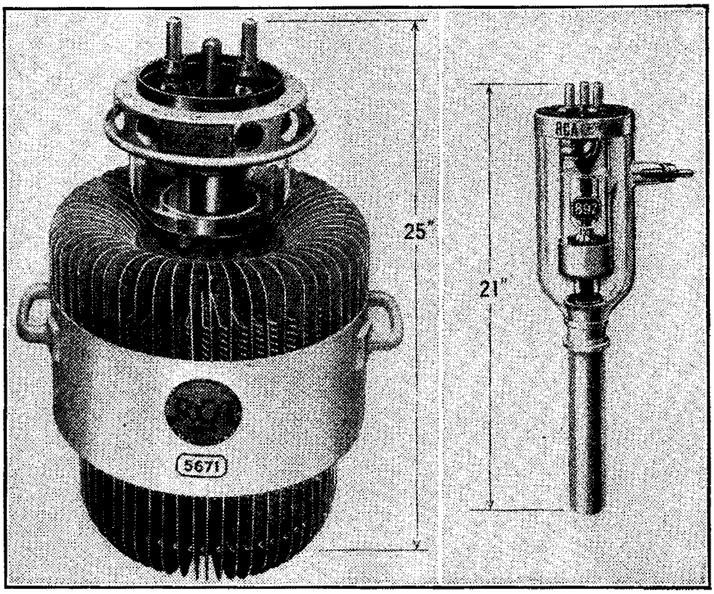


FIG. 3-9. Large transmitting tubes. (*Courtesy of RCA.*)

tantalum, but it does not absorb gas. Graphite is the best radiator and can likewise be operated at a high temperature, but it does have a tendency to have considerable amounts of occluded gas, which must be removed by high-temperature treatment while the tube is being exhausted.

High-power transmitting tubes are cooled either by forced air or by circulating water. The plate is made of copper, which also serves as part of the envelope of the tube. These tubes have ratings of 5 to 1000 kw. If air-cooled, the copper anode has cooling fins through which air is blown. If water-cooled, the anode is immersed in a water jacket through which the cooling water is circulated. Figure 3-9 shows a photograph of an air-cooled tube and a water-cooled tube with the water jacket removed.

3-9. Diode Ratings. As just explained, a diode is limited in its plate-current and voltage ratings by the ability of the plate to dissipate heat. This energy radiated from the plate must pass through the glass bulb, where a certain portion is absorbed, thus heating the glass. So our envelope also limits the ratings and must have a large enough area to be able to reradiate this heat without the temperature reaching the point where gases are driven off. Because of this, tubes with ratings above 1 kw are never enclosed in glass envelopes but are made with either forced-air or water cooling.

When used to rectify high voltages, the maximum plate voltage may be limited by flashover across the surface of the glass. For this reason the plate and filament leads are brought out at opposite ends of the envelope. These tubes can be made to withstand thousands of volts, while a receiving type of rectifier with all leads brought out the same end will withstand a maximum of only 500 to 1000 volts.

When used in rectifier circuits, the polarity of the plate voltage reverses each cycle and the tube must be able to withstand an inverse peak voltage (a voltage which makes the plate negative with respect to the filament) of one to two times the crest value of the alternating voltage it is rectifying. Therefore tubes are usually given a peak-inverse-voltage rating.

To summarize, the current and voltage ratings of diodes are determined by the following:

1. The plate current, limited by the emission ability of the filament with rated filament voltage applied.
2. The heat-dissipating ability of the plate.
3. The heat-dissipating properties of the envelope.
4. The peak inverse voltage the tube will withstand.
5. Flashover voltage between lead wires.

3-10. Characteristic Curves for a Diode. If the values of plate current vs. plate voltage are plotted for a diode with rated filament voltage, a curve similar to Fig. 3-10 is the result. This is known as a static plate characteristic. The resistance of the diode for direct current varies from point to point on the curve. It is the reciprocal of the slope of the straight line drawn from the origin to the point and is given by

$$R_p = \frac{E_b}{I_b} \tag{3-34}$$

The slope of the static characteristic itself is given by

$$\text{Slope} = \frac{dI_b}{dE_b} = g_p \tag{3-35}$$

which has the dimensions of a conductance and is known as the plate conductance. The reciprocal of g_p is known as the dynamic plate resistance and is denoted by the symbol r_p.

$$r_p = \frac{dE_b}{dI_b} \tag{3-36}$$

It should be noted that this factor also is different for different points on the curve. The dynamic plate resistance is useful for expressing approximate relationships between small voltage and current changes.

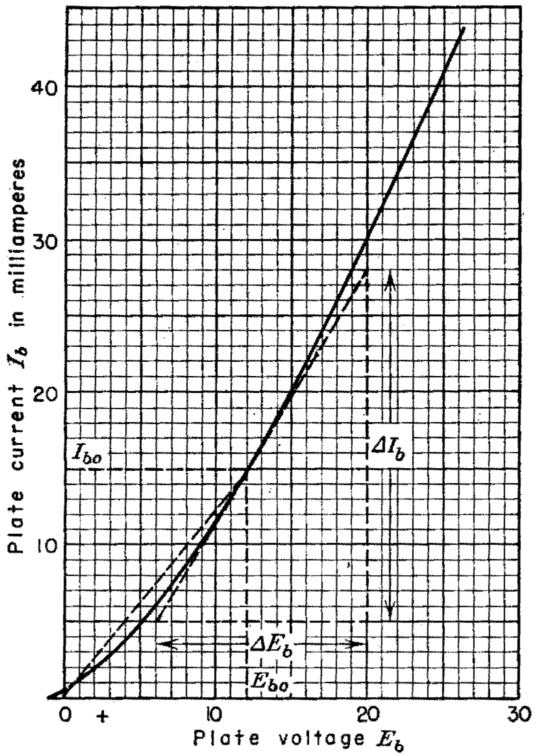


FIG. 3-10. Plate characteristic of a type 6H6 diode.

If a diode is connected in the series circuit of Fig. 3-11 with a load resistor R_L and an alternating-voltage source, d-c power can be obtained from the alternating source. The sinusoidal voltage v_p does not appear across the tube. Rather it is this voltage minus the IR drop in the load which appears across the tube.

To study the action in this circuit, it is desirable to draw another curve showing the relationship between the values of voltage applied to the circuit and the current through the circuit. This curve is known as a dynamic characteristic and is obtained by adding to the abscissa values

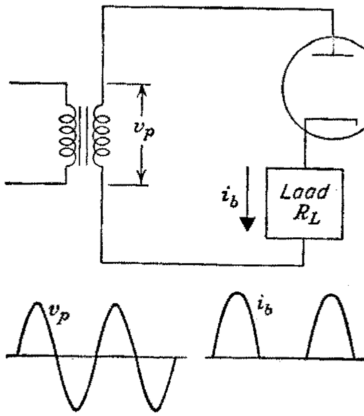


FIG. 3-11. A diode with a resistance load in an a-c circuit.

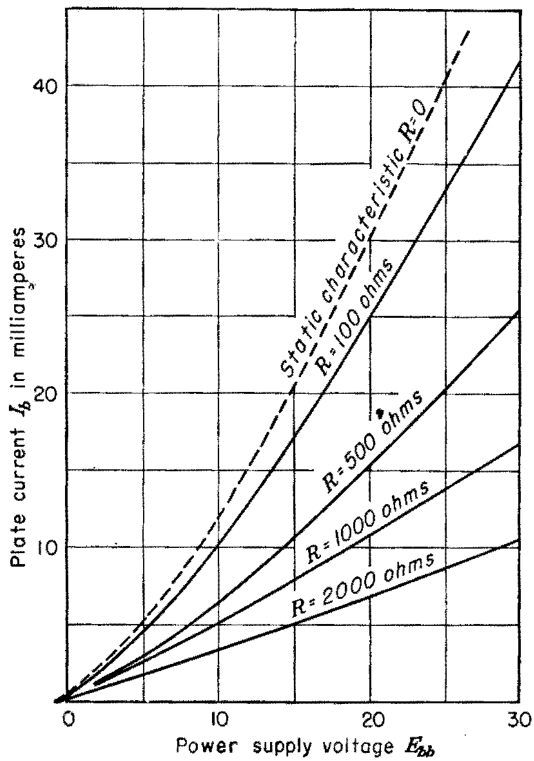


FIG. 3-12. Dynamic plate characteristics for a type 6H6 diode with various load resistors.

of the static curve an amount that is equal to the I_bR_L drop in the load at each value of I_b. A family of dynamic characteristics is drawn in Fig. 3-12 for various values of the load resistance R_L. The variable supply voltage is E_{bb}. Note that the slope of any dynamic curve is less than that of the static one and this difference is more pronounced with increasing load resistance. Likewise the curve becomes straighter with increased load resistance. This is because we are adding a linear element

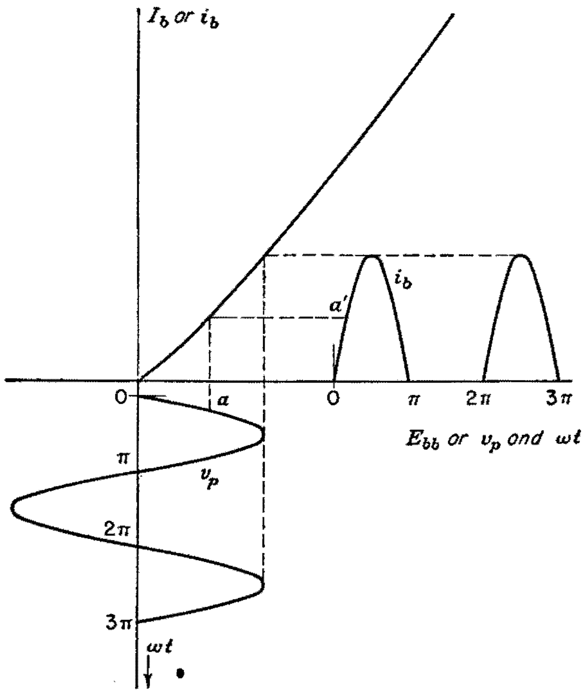


FIG. 3-13. Graphical method of determining the plate current in a diode when an alternating voltage is applied.

to the circuit containing a nonlinear one, and the larger R_L is, the more it influences the shape of the characteristic.

Assume that the dynamic characteristic of a tube connected as in Fig. 3-11 is given by the curve in Fig. 3-13. If an alternating voltage v_p is applied, the resulting waveform of the current through the diode and the load R_L can be found graphically from the dynamic characteristic. The alternating-voltage wave is plotted on the curve by using the same voltage scale but adding an axis for time, which is increasing in the negative-y direction. A time scale is also added to the horizontal axis. Thus the instantaneous values of plate current can be found for any value of ωt. Note that the plate current does not consist of exact half sinusoids since the dynamic characteristic is not a straight line. However, if it is almost straight, the plate current can then be assumed to have the shape of a half sinusoid without appreciable error. Further study of this rectifying operation will be made in later chapters.

REFERENCES

1. Frank, N. H.: "Introduction to Electricity and Optics," 2d ed., McGraw-Hill Book Company, Inc., New York, 1950.
2. Spangenberg, K. R.: "Vacuum Tubes," McGraw-Hill Book Company, Inc., New York, 1948.
3. Ryder, J. D.: "Electronic Fundamentals and Applications," Prentice-Hall, Inc., New York, 1950.
4. Millman, J., and S. Seely: "Electronics," 2d ed., McGraw-Hill Book Company, Inc., New York, 1951.
5. E. E. Staff, MIT: "Applied Electronics," John Wiley & Sons, Inc., New York, 1943.

PROBLEMS AND QUESTIONS

1. The anode of a diode with parallel plane electrodes has a potential of E_b volts above that of the cathode. Sketch the potential distribution for each of the following cases: (*a*) The plate current is zero. (*b*) There are few electrons in the interelectrode space. (*c*) The plate current is space-charge-limited, and electrons are emitted from the cathode with zero velocity. (*d*) The plate current is space-charge-limited, and the electrons are emitted from the cathode with appreciable velocity. In each of the foregoing cases does an electron just emitted from the cathode have a net force exerted on it if the potential distribution is as described?

2. Two charges Q_1 and Q_2 are enclosed by a glass sphere of 1 m diameter. Q_1 is $+3 \times 10^{-3}$ coulomb, and Q_2 is -2×10^{-3} coulomb. (*a*) Calculate the total flux emanating from the sphere. (*b*) Calculate the average flux density at the surface of the sphere.

3. One mathematical step in the derivation of the Child-Langmuir law is the solution of an equation of the form $d^2y/dx^2 = y^{-1/2}$. If both y and dy/dx are zero when x is zero, find x as a function of y.

4. A diode is operated at such a cathode temperature that the plate currents for the voltages considered in the problem are determined by the Child-Langmuir law. At a plate voltage of 50 volts the current is 20 ma. Find the plate current if the plate voltage is raised to 125 volts.

5. The plate current of a certain diode is known to obey the relation $I_b = KE_b^n$. It is found that when $E_b = 100$ volts, $I_b = 10$ ma. Also when $E_b = 10$ volts, $I_b = 0.3$ ma. Find the coefficient K and the exponent n.

6. Suppose the space between the plates of Fig. 3-6 contains no charges of any kind. If the voltage E_b is 1000 volts and the separation is 0.1 m, find the following: (*a*) the potential gradient, (*b*) the field strength, (*c*) d^2E/dx^2. Does the value in (*c*) satisfy Poisson's equation?

7. (*a*) Use Eqs. (3-24) and (3-25) for the case of the parallel plane electrodes, and determine the equation for the potential at any point in the interelectrode space. (*b*) For $d = 1$ cm and $E_b = 100$ volts, plot the potential distribution curve, and compute J_b. (*c*) Repeat (*b*) if E_b is made 150 volts.

8. A free electron, just off the surface of the cathode in a thermionic diode, has zero velocity. (*a*) Calculate the energy, in joules, of the electron when it strikes the plate, which has a potential of $+250$ volts with respect to the cathode. (*b*) If the plate current is 250 ma, calculate the number of electrons striking the plate each second. (*c*) Calculate the power supplied to the plate by the bombarding electrons.

9. State the merits and demerits of tantalum, of molybdenum, and of carbon as anode materials.

10. Name five factors which determine the current and voltage ratings of a diode.

11. Use Fig. 3-10 to determine the value of R_p for a 6H6 type tube when $E_b = 12$ volts and $I_b = 15$ ma.

12. Calculate the dynamic plate resistance of a 6H6 tube at the operating point of Prob. 11.

13. State the meanings of the terms "static characteristic" and "dynamic characteristic."

14. A 200-ohm resistor is inserted in series with a plate lead of a 6H6 tube. If the cathode is normally heated and $E_{bb} = 20$ volts, determine I_b.

15. Draw the dynamic characteristic for a 6H6 diode which has a 400-ohm resistor in the plate lead.

16. A certain diode has a pure-tungsten filament in a straight piece 5 cm long and 0.0127 cm in diameter. The operating temperature is 2500°K. The anode is a cylinder 1 cm in diameter, 6 cm long, and concentric with the filament. What plate current flows if the anode is 100 volts positive relative to the filament?

CHAPTER 4

THE VACUUM TRIODE AND OTHER MULTIELEMENT TUBES

4-1. De Forest's Triode. The triode was first developed by Lee De Forest in 1906. His first tube, which was quite different from present-day tubes, consisted of two plane parallel plates between which was located an incandescent filament. By making one of the plane electrodes (which we shall call the plate, or anode) positive with respect to the filament, he had what amounted to a space-charge-limited diode. However, when he made the other plane electrode (which we shall call a grid) negative with respect to the filament, the emitted electrons were repelled toward, as well as attracted to, the anode. By varying the magnitude of the potential on the grid he could increase or decrease the amount of repulsion and thus control the anode current. Since the grid had negative polarity, no current flowed in its circuit and control of anode current was obtained with the expenditure of no power in the grid circuit. This control, however, was very limited, and in order to improve it, he soon tried the control electrode in the form of a ladderlike arrangement, or grid, between the filament and the plate. By changing the negative potential applied to this grid he found that he had much better control over the plate current, with the expenditure of little or no power in the grid circuit. This is the tube that has developed into our present-day triode.

4-2. Triode Construction. Today's triode usually consists of a filamentary or an indirectly heated cathode surrounded by a helical grid wound with fine nickel wire. This grid in turn is surrounded by a plate, which may be cylindrical in shape. In some cases the plate, instead of being circular in cross section, may be elliptical or rectangular.

Figure 4-1 shows photographs of the parts going into the construction of two present-day triodes. Note that while both filamentary and indirectly heated cathodes are shown, most tubes now are the indirectly heated type. The envelope may be either glass or metal. Some types of tubes are made in both the glass and metal versions, while others are made only in glass or in metal. The electrical properties of a glass tube are approximately the same as those for a metal tube of the same type. There are some differences in the capacitances between electrodes and in the shielding from external fields.

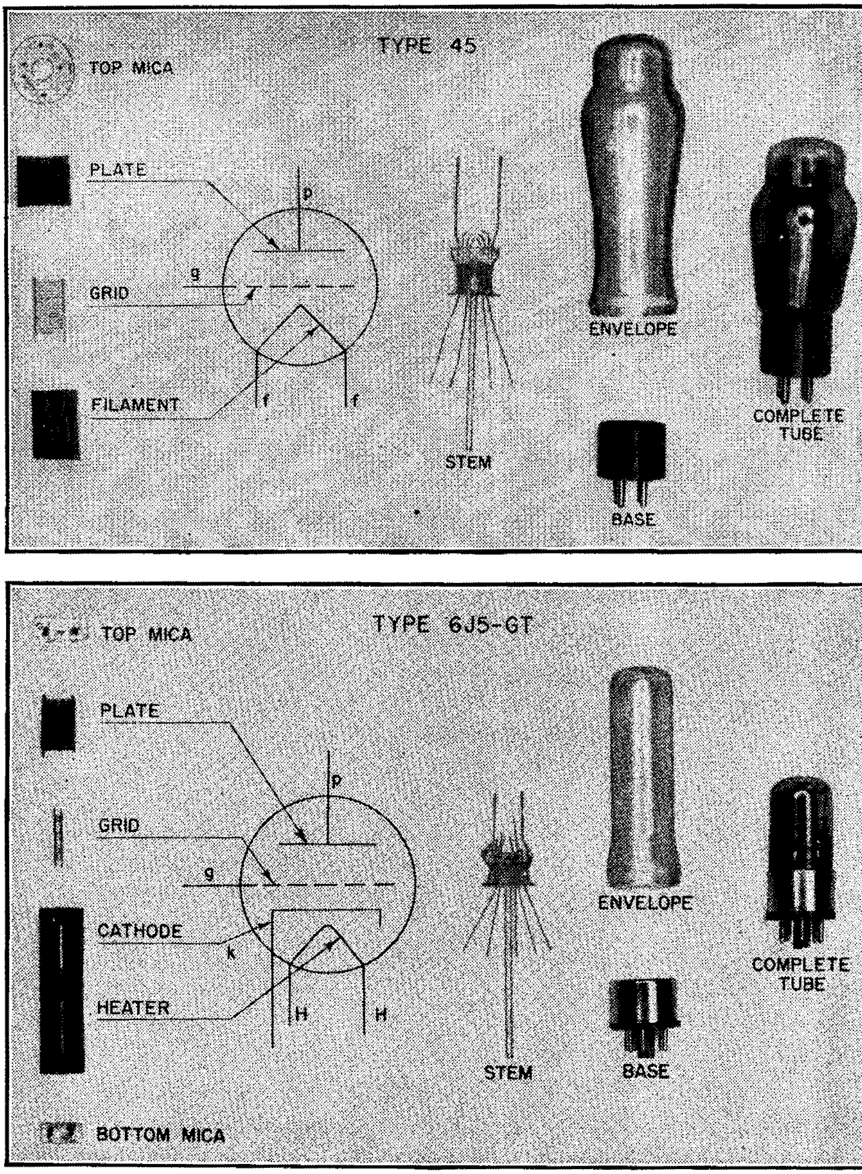


FIG. 4-1. Photographs showing the construction of two types of triodes.

4-3. The Function of the Grid. In the preceding chapter it was shown that the plate current in a diode is dependent on the potential distribution, which in turn is determined by the amount and distribution of the space charge in the interelectrode space. If the initial velocity of the emitted electrons is assumed to be zero, the spacing between parallel plane electrodes to be d, and the plate voltage to be E'_b, the potential distribution will be as shown by curve (1) in Fig. 4-2. The value of cathode current (equal to the plate current in this case) may be called I'_k. If the plate voltage is increased to E''_b, the potential distribution will be as shown by curve (2) and the cathode current will increase to a new value I''_k.

Now suppose it to be possible to insert a third parallel plane electrode at a distance d' from the cathode. If its potential is made E'_c, the potential distribution between it and the cathode is the same as in curve (1) and Eq. (3-24) shows that the cathode current will again be I'_k. In order to increase the cathode current to I''_k, the third-electrode voltage need be raised to only E''_c as shown. Note that less change is needed for third-electrode voltage than for plate voltage to obtain the same change in cathode current.

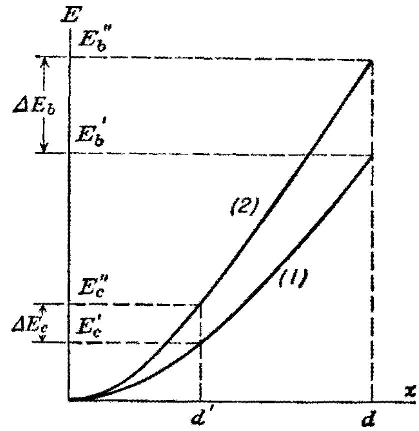


FIG. 4-2. Potential distribution in a diode for two values of plate voltage. Electrons have zero initial velocity.

If the third electrode is perforated, we call it a grid. Now, many of the electrons will pass through the holes and eventually arrive at the plate. This follows because the potential gradient between the grid and the plate is everywhere positive. Although the potential distribution now depends upon both grid and plate potentials, a change in value of cathode current is still more dependent upon changes in grid potential than upon changes in plate potential.

To avoid grid current, it is usually advisable to operate the grid at a negative potential, in which case the plate current is the same as the cathode current. The potential distribution in a triode is shown in Fig. 4-3. Because of the distortion of the field caused by the irregular geometry of the grid, it is impossible to show the potential distribution by means of a single curve. Let us assume the grid structure to consist of parallel bars. Then the upper curve represents the distribution along a line midway between two grid wires, while the lower curve represents the distribution along a line through the center of a grid wire. These are the two extremes of potential distribution, and all others must lie between them. Note that the field close to the cathode is quite similar to the field close to the cathode in a diode. If the plate voltage of the triode is decreased, the potential-distribution curve will be lowered and the plate current will decrease. The dotted curves in Fig. 4-3a show this condition. The solid curves of Fig. 4-3b are for the same conditions as are the solid curves of Fig. 4-3a. The dotted curves of Fig. 4-3b show the effect of a more negative grid on the potential distribution. The grid potential has been so adjusted that the potential distribution in the vicinity of the cathode is the same as for the dotted curves of Fig. 4-3a. Hence approximately the same plate current will flow. Note that in Fig. 4-3a a change of ΔE_b volts was necessary to cause a given plate-current change. In Fig. 4-3b a change of only ΔE_c volts was needed to cause the same change in plate current. Hence the grid is more effective in controlling plate current, and at the same time, since it is negative with respect to all other electrodes, it draws no current.

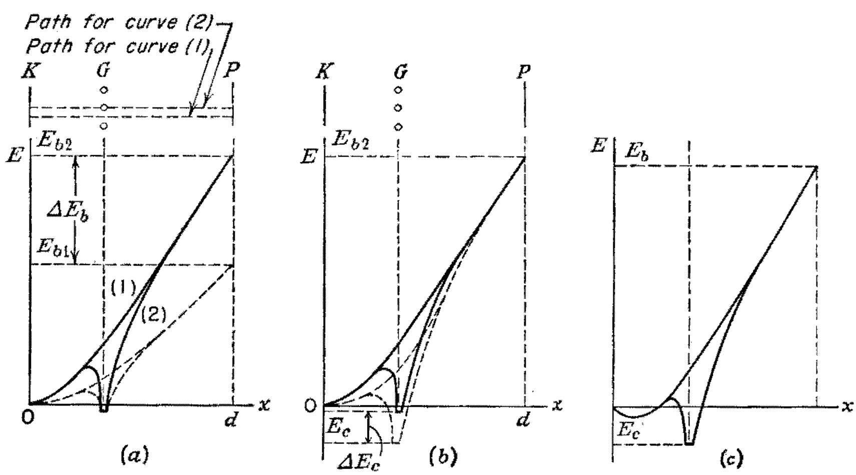


FIG. 4-3. Potential distributions in a triode with a negative grid.

In an actual tube the electrons are not emitted with zero velocity, and hence a potential minimum is established as it is in the diode. The potential distribution will then be more as shown in Fig. 4-3c. Only those electrons with sufficient energy to carry them out past the point of minimum potential will make up the plate current. Hence our theory would have to be slightly modified to account for this effect.

Let us call the relative effectiveness of changes in the grid and plate potentials on the plate current the amplification factor and designate it by the symbol μ. If a change of ΔE_c volts on the grid causes a certain change in the plate current, then a change $\Delta E_b = \mu \, \Delta E_c$ volts on the plate would be necessary to cause the same change in current. The plate current depends on both voltages and may be expressed by the functional relation

$$I_b = f(E_b, E_c) \tag{4-1}$$

Since for the purposes of this discussion the grid is assumed to be negative, one might expect to find no current flowing in the grid circuit. However, the grid does have an appreciable physical size, and some electrons will have high enough velocities so that they will penetrate the negative field of the grid and cause a small current to flow. For most applications this current is so small that it can be neglected. There may be positive ions present in the actual tube, which will also cause grid-current flow. Some tubes are so designed that they operate with positive

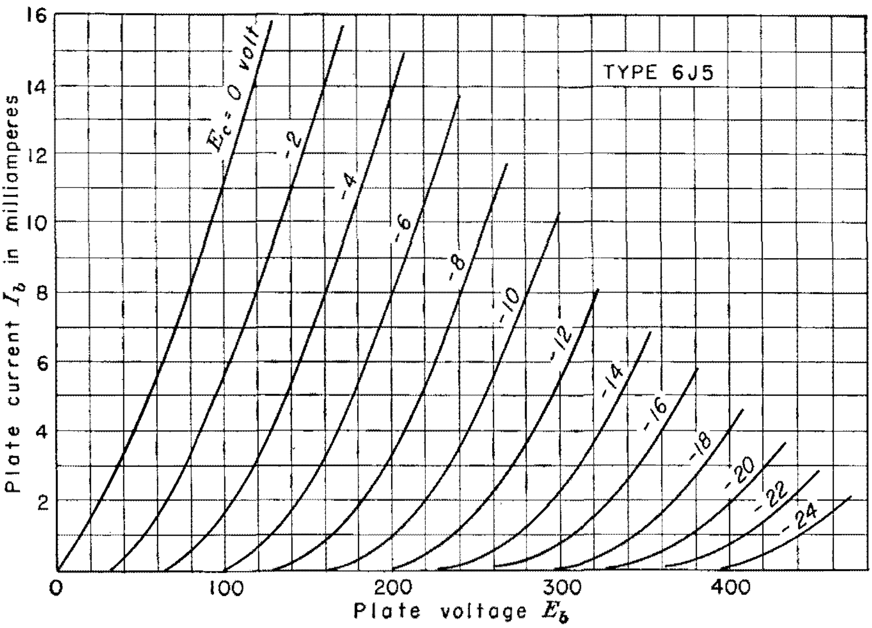


FIG. 4-4. Plate characteristics of a 6J5 triode.

grids and hence draw considerable grid current. However, at the present time our interests lie with the tubes with negative grids and having no appreciable grid current.

4-4. Triode Characteristics. Since in general a vacuum tube is a nonlinear device, the most convenient method of expressing its operating characteristics is by means of a number of graphs. The plate current I_b for a vacuum triode is a function of two independent variables, the plate voltage E_b and the grid voltage E_c. Therefore the function cannot be plotted as a single curve but must be represented either by a space model or by families of curves called static characteristics.

One of these families is shown in Fig. 4-4 and is known as the plate-current plate-voltage characteristics, or simply the plate characteristics. Here the grid voltage is a parameter. Note that the curve shapes are all similar, but each curve is spaced from its neighbor by approximately equal horizontal distances. This is the family which is usually drawn in the tube manuals, and from it the plate current can be determined for any combination of electrode voltages.

A second family of curves is known as the transfer, or mutual, characteristics, and a typical set is shown in Fig. 4-5. It is named transfer, or mutual, characteristics because it gives the relationship between plate current and grid voltage for constant plate voltage. This family may appear at first glance to be more useful than the plate family in predicting the operation of a particular tube. This is not necessarily the case since it does not lend itself so readily to the study of the tube in an actual circuit. Again all the curves are of about the same shape but spaced from their neighbors by an approximately constant horizontal distance. They are not straight lines over their entire length but are nearly so over a limited region. If, in the use of this tube, we so adjust our voltages that we never go beyond this straight portion, we can treat the tube as a linear device and make a simple analytical analysis of the operation of the circuit. For this type of analysis we need to know the tube coefficients, which will be discussed later. If in our operation of the tube we do go beyond the linear region, the plate current is no longer related to the grid voltage by a constant of proportionality and the device usually is no longer treated analytically but is analyzed graphically.

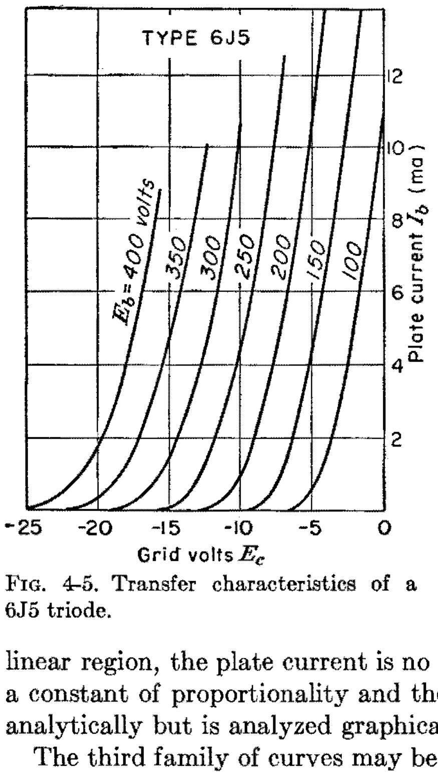


FIG. 4-5. Transfer characteristics of a 6J5 triode.

The third family of curves may be obtained if the plate current is held constant by the simultaneous variations of the grid and plate voltages. Such a family is plotted in Fig. 4-6 and is known as the constant-current characteristics. This family does not give us any additional information over that shown by the other two families. It may, however, occasionally be used for the analysis of transmitting-tube operation.

4-5. Triode Coefficients. In order thoroughly to understand the actions of thermionic triodes, we must consider two conditions of operation. The first is the static condition and involves only constant voltages and currents. The second, and probably the more important, is the dynamic condition. Here the currents and voltages are varying about a certain reference point which we shall call the operating point. The chief use of the static condition is to locate this operating point and thus allow calculations for the dynamic condition.

Up to this point the properties of vacuum tubes have been considered under conditions of constant applied voltages and currents, *i.e.*, the first condition named above. Following customary procedure, upper-case letters have been employed as symbols for these quantities, both in graphs and in equations. Now it is time to begin the study of tube action under the second-named condition, which allows some of these

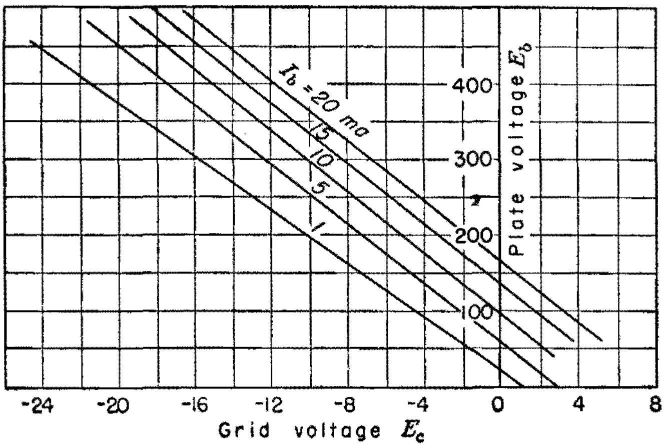


FIG. 4-6. Constant-current curves for a 6J5 triode.

quantities to be variable with time. Following the same standards of practice, lower-case letters should be used. However, the graphs and equations obtained under the former condition are useful in this new study and hence some confusion in usage of symbols is likely to prevail.

Equation (4-1) stated a relationship between plate current, plate voltage, and grid voltage. Since it is not essential that the quantities involved be constants, we may rewrite the relation in the form

$$i_b = f(e_b, e_c) \tag{4-2}$$

This equation has a very limited usefulness, when in this form, and hence it would be desirable to find some expression which would give the actual relationship among these three variables. This can be done by means of a rather difficult mathematical derivation, which is beyond the scope of this book, the result being

$$i_b = k(e_b + \mu e_c)^n \tag{4-3}$$

where k is a constant determined by the geometry of the tube, μ is the amplification factor, and n has a value of approximately 1.5. Note the similarity between this and the statement of the Child-Langmuir law for a diode [Eq. (3-27)].

Figure 4-4 is the plate family of characteristics for a 6J5 tube. Since the data for each of these curves were taken for a constant value of e_c and since μ remains approximately constant for all curves (see Fig. 4-8), the shape of each curve is determined only by k and n. An examination of

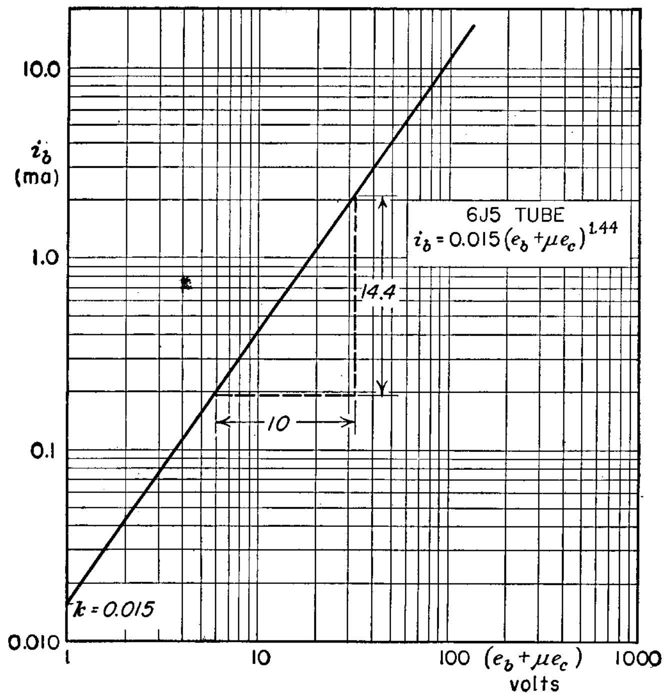


FIG. 4-7. Determination of the constants in the Child-Langmuir equation, which has been modified for the triode.

the family shows that all have almost identical shapes, which indicates that k and n must be constant for that particular tube.

Experimental evidence to prove that Eq. (4-3) gives the relationship between the three variables is not difficult to obtain. If we first write Eq. (4-3) in logarithmic form, we have

$$\log i_b = \log k + n \log (e_b + \mu e_c) \qquad (4\text{-}4)$$

which is a straight-line graph when plotted with $\log i_b$ as the ordinate and $\log (e_b + \mu e_c)$ as the abscissa. If we now take values of e_b, e_c, and i_b from the curves of Fig. 4-4 and plot $\log i_b$ vs. $\log (e_b + \mu e_c)$, a single straight line is obtained as shown in Fig. 4-7. In order to do this, it was necessary to know the value of μ, which for this particular tube is approximately 20. The calculation of this amplification factor μ will be taken up in a later section. We can now determine the values of n and k for the triode, just as we did in the case of the diode. For this particular tube they are found to be $n = 1.44$ and $k = 0.015$. Having evaluated these two factors, we may write

$$i_b = 0.015(e_b + 20e_c)^{1.44} \tag{4-5}$$

which is the formula that gives the approximate relationship between the three variables in a 6J5 triode. Of course this relation also applies for the case of constant voltages and current.

Equation (4-3) gives us the general expression for the plate current in any triode. But even this rather simple mathematical expression becomes too complicated to use in predicting the operation of a particular tube under a given set of operating conditions. It is therefore desirable to express the properties of triodes in the form of three coefficients which give the rate of change of one variable with respect to a second variable while the third is held constant. These are known as the tube coefficients and are closely related to the slopes of the three families of characteristics.

The amplification factor has already been mentioned and may be defined mathematically as

$$\mu = -\frac{\partial e_b}{\partial e_c} = -\left.\frac{de_b}{de_c}\right|_{i_b \text{ constant}} \tag{4-6}$$

the partial derivative being used to imply that the plate current i_b is held constant. In this definition it would also be proper to use upper-case symbols. The coefficient may be interpreted as being the negative of the slope of the constant-current characteristics (Fig. 4-6). Since the slope of any of these curves is negative, a negative sign is used before the derivative to give a more desirable positive value to the coefficient.

The dynamic plate resistance $r_p (= 1/g_p)$ has already been discussed for the diode, and everything said about it in that discussion holds for the triode. It is the reciprocal of the slope of the plate characteristic at a given operating point and can be expressed mathematically as

$$r_p = \frac{\partial e_b}{\partial i_b} = \left.\frac{de_b}{di_b}\right|_{e_c \text{ constant}} \tag{4-7}$$

From an examination of the plate-family characteristics shown in Fig. 4-4, it can be seen that the slope, and hence the plate resistance, will vary over a considerable range of values depending on where the operating point is chosen.

The third tube coefficient is the mutual conductance, or the transconductance, either name being acceptable. As the name implies, it is a coefficient which is mutual to, or determined by, two electrodes in the tube. There are other mutual conductances which could be defined, but to us, unless other electrodes are specified, the term will always mean the plate-current grid-voltage mutual conductance. Its symbol is g_m, and it is defined by

$$g_m = \frac{\partial i_b}{\partial e_c} = \left.\frac{di_b}{de_c}\right|_{e_b \text{ constant}} \tag{4-8}$$

It is equal to the slope of the transfer characteristic at the operating point, and its value is usually given in micromhos.

The relationship between these three coefficients can be shown by taking the total differential of i_b [Eq. (4-2)]; thus

$$di_b = \frac{\partial i_b}{\partial e_b} de_b + \frac{\partial i_b}{\partial e_c} de_c \tag{4-9}$$

An examination of this equation shows that Eqs. (4-7) and (4-8) furnish substitutions for the partial derivatives; thus

$$di_b = \frac{1}{r_p} de_b + g_m \, de_c \tag{4-10}$$

Now if we take the special case where the plate current is held constant, di_b becomes equal to zero and Eq. (4-10) yields

$$-\left.\frac{de_b}{de_c}\right|_{i_b \text{ constant}} = g_m r_p \tag{4-11}$$

Here the left term is, from Eq. (4-6), the amplification factor μ. Hence Eq. (4-11) becomes

$$\mu = g_m r_p \tag{4-12}$$

The three tube coefficients, amplification factor μ, dynamic plate resistance r_p, and mutual conductance g_m determine the tube's operation at a given operating point. Since the slopes of the plate characteristics and of the transfer characteristics vary considerably, the coefficients r_p and g_m likewise vary over quite a range of values. On the other hand the slope of the constant-current curves is almost the same for the whole family, and therefore μ is practically constant for all operating points. Figure 4-8 shows the variations in the three tube coefficients for various operating points for a 6J5 triode.

4-6. Calculation of Tube Coefficients. The three tube coefficients may be easily calculated from any one of the three families of characteristics. The results will be only approximate, but accurate enough for most calculations in which we wish to use them. Since the plate family is the one most commonly available, we shall use this in our discussion. Figure 4-9 gives the plate family for a type 6J5 medium-μ triode.

Amplification Factor. If we were using the constant-current characteristics, the definition of μ could be used directly to determine its value. However, in this case it is not possible to apply the definition exactly.

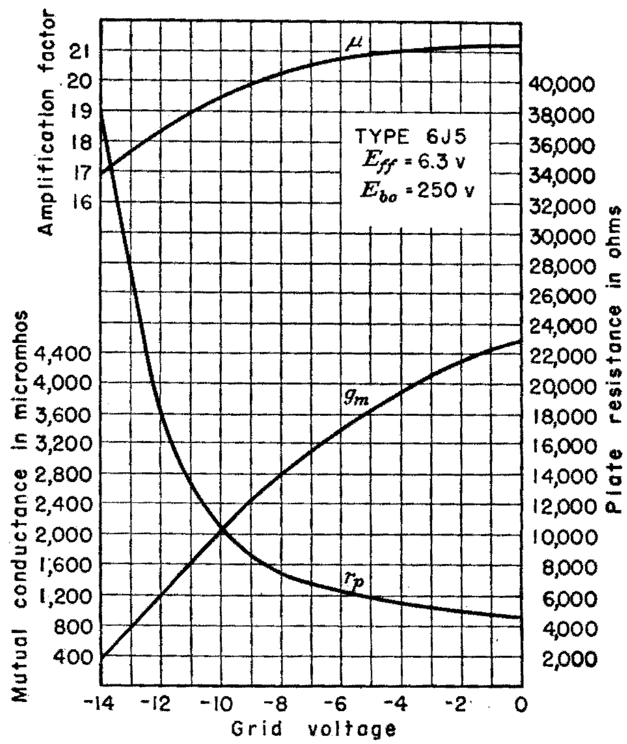


FIG. 4-8. Curves showing the dependence of μ, g_m, and r_p on the operating point of a 6J5 triode.

An approximate determination can be made by using the relation

$$\mu = -\frac{\partial e_b}{\partial e_c} = \lim_{\substack{\Delta e_b \to 0 \\ \Delta e_c \to 0}} -\frac{\Delta e_b}{\Delta e_c}\bigg|_{i_b \text{ constant}} \approx -\frac{\Delta e_b}{\Delta e_c}\bigg|_{\substack{i_b \text{ constant} \\ \Delta e_b,\ \Delta e_c \text{ small}}} \tag{4-13}$$

Let us now choose an operating point. If the values of any two of the three variables are specified, then the value of the third can be obtained from the plate family. For instance, if we choose the point $E_{bo} = 250$ volts and $E_{co} = -8$ volts, then from the curves we find that $I_{bo} = 9$ ma. The additional subscript *o* has been appended to indicate an operating-point value (also called a bias value). Now if the plate voltage is kept constant and the grid voltage changed to -10 volts, the plate current will drop to 4.2 ma. In order to bring the plate current back to 9 ma, the plate voltage must be increased to 290 volts. Likewise, if the grid voltage is reduced to -6 volts, the plate voltage must be decreased to 210 volts to maintain the plate current constant. Thus when the grid voltage is changed from -6 volts to -10 volts, or $\Delta e_c = -4$ volts, the plate voltage must change from 210 volts to 290 volts, or $\Delta e_b = 80$ volts, if no change in plate current is allowed. From Eq. (4-13) we may write $\mu = -(290 - 210)/(-4) = 20$. The advantage of straddling the operating

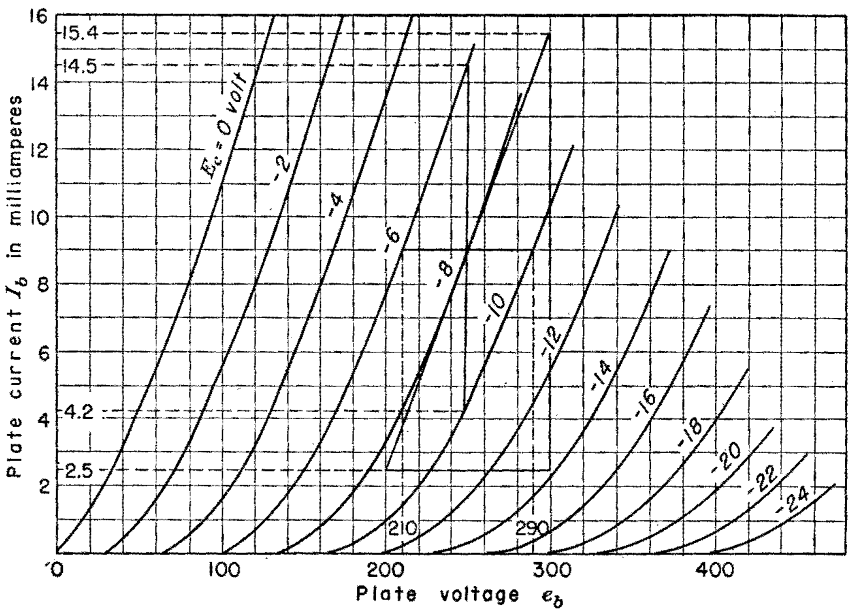


FIG. 4-9. Curves showing the graphical determination of the tube coefficients at a given operating point for a 6J5 tube.

point in this determination is that a better average value is obtained. The value obtained by this method may be checked by referring to that tabulated in a tube manual for the same operating point.

Dynamic Plate Resistance. The dynamic plate resistance can be determined directly from the definition as the reciprocal of the slope of the plate characteristic at the operating point. Drawing a tangent line in Fig. 4-9, we obtain data which yield

$$r_p = \frac{\partial e_b}{\partial i_b} = \frac{\Delta e_b}{\Delta i_b}\bigg|_{\text{tangent-line values}} = \frac{300 - 200}{0.0154 - 0.0025} = 7750 \text{ ohms}$$

The tube-manual value for this coefficient is 7700 ohms, which is a satisfactory check.

The Mutual Conductance. Here again we may refer back to the definition in order to write an approximate expression for the mutual conductance.

$$g_m = \frac{\partial i_b}{\partial e_c} = \lim_{\substack{\Delta i_b \to 0 \\ \Delta e_c \to 0}} \left. \frac{\Delta i_b}{\Delta e_c} \right|_{e_b \text{ constant}} \approx \left. \frac{\Delta i_b}{\Delta e_c} \right|_{\substack{e_b \text{ constant} \\ \Delta i_b,\ \Delta e_c \text{ small}}} \tag{4-14}$$

From the plate family we see that, with the plate voltage held constant at 250 volts, a change in the grid voltage from -10 volts to -6 volts results in a plate-current increase from 4.2 ma to 14.5 ma. Hence $g_m = (0.0145 - 0.0042)/4 = 2575$ micromhos. This is an average value obtained in the region of the operating point. In order to check the values of the coefficients, let us use the relation $\mu = g_m r_p$; $2575 \times 10^{-6} \times 7750 = 20$, which checks the value of μ which was obtained before.

4-7. Dynamic Transfer Characteristics. Up to now we have discussed transfer characteristics from the viewpoint of variable grid voltage and constant plate voltage. This means that, unless E_{bb} is varied, there cannot be a load impedance in the plate circuit of the tube. In order to use a tube as an amplifier, it is necessary to have a voltage output, which can be obtained by placing a load resistor in the plate circuit. The voltage drop across this resistor is the voltage output. As a result the plate voltage will not be constant, since it is equal to the battery voltage E_{bb} minus the drop across the load resistor. It is then seen that it might be desirable to have a transfer characteristic for the tube with a load R_L in the plate circuit and with a fixed value of E_{bb}.

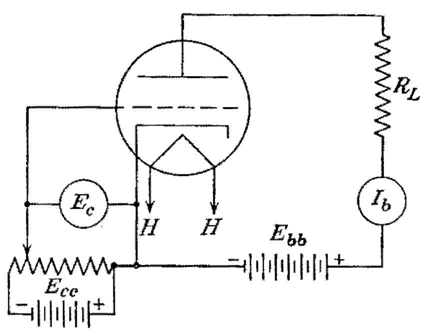


FIG. 4-10. Circuit used to determine the dynamic transfer characteristics of a triode.

If a load resistor is put in the plate circuit of the tube as shown in Fig. 4-10 and data are taken to plot a transfer characteristic with E_{bb} held constant, the resulting curve differs considerably from the static curve. The new curve is known as the dynamic transfer characteristic. It actually is the characteristic of the tube and its load resistor together, while the static curve is the characteristic of the tube alone.

Figure 4-11 shows a static transfer characteristic ($R_L = 0$) for a 6J5 tube. It also shows dynamic characteristics for E_{bb} equal to 250 volts and load resistors of 10,000, 50,000, and 100,000 ohms. Note that as the magnitude of the load resistor increases the characteristic tends to become more linear. This seems reasonable because we have effectively a linear and a nonlinear element in the same series circuit, and the larger the linear element becomes, the more the circuit should have over-all linearity.

Use of the dynamic transfer characteristic allows one to make a graphical determination of the waveform of the plate current in a triode for a given waveform of the grid-signal voltage v_g. From the circuit diagram

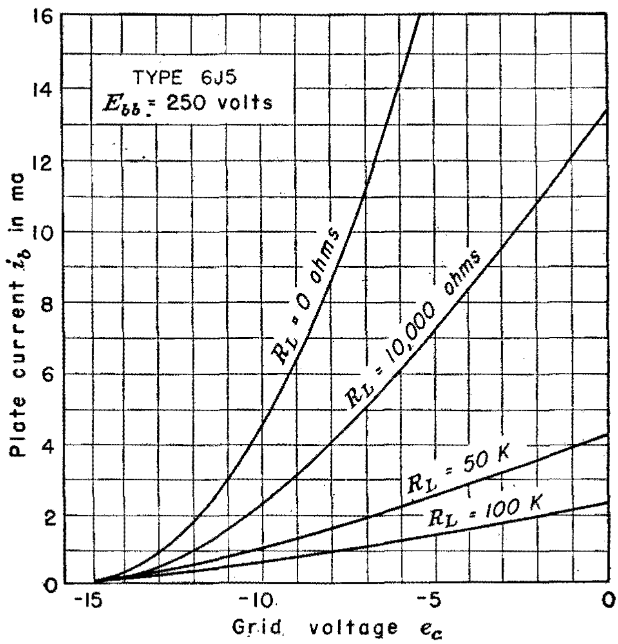


FIG. 4-11. Dynamic transfer characteristics for a 6J5 tube.

of Fig. 4-12*a* we can write

$$e_c = E_{cc} + v_g \tag{4-15}$$

If v_g is sinusoidal and V_{gm} is used to designate the maximum value of the signal voltage, we can also write

$$v_g = V_{gm} \sin \omega t \tag{4-16}$$

and

$$e_c = E_{cc} + V_{gm} \sin \omega t \tag{4-17}$$

Referring to Fig. 4-12*b*, we see that v_g can be plotted on the e_c axis with the time axis running vertically at $e_c = E_{cc}$. By projection on a dynamic characteristic drawn for particular E_{bb} and R_L values, the waveform of the plate current can be plotted along a time axis running horizontally at $i_b = I_{bo}$. (The dynamic curve in the figure is for a 6J5 tube with E_{bb} 250 volts and R_L 50,000 ohms.) This waveform may or may not be similar to the waveform of the signal voltage v_g, depending on whether the operating portion of the dynamic curve is linear or not. If E_{cc} in the example had been chosen more negative, it is possible that the lower portion of the current waveform would have been a poor reproduction of the signal-voltage waveform.

The dynamic curve predicts the performance of a tube for particular values of E_{bb} and R_L. If either of the latter is changed, a new dynamic characteristic must be plotted. This is a rather laborious process if many such curves have to be drawn. As a result, although the dynamic characteristic can be used in an explanation or as a starting point in an analysis of a vacuum-tube problem, it is seldom used for the design and calculation of practical amplifier circuits.

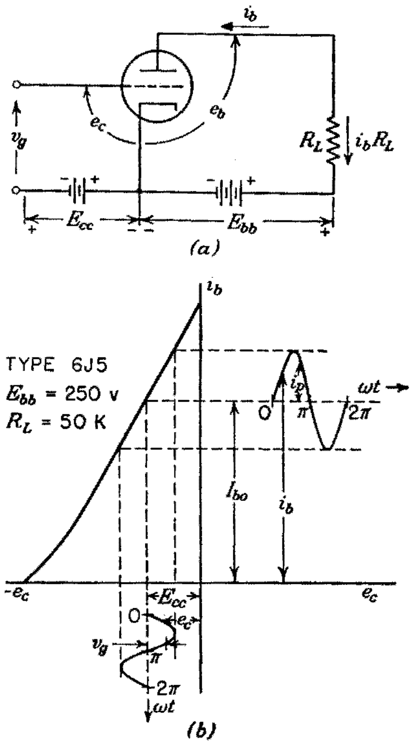


FIG. 4-12.

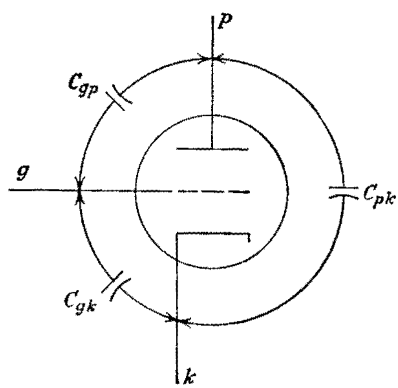


FIG. 4-13. Interelectrode capacitances in a triode.

4-8. Shortcomings of the Triode. In the evolution of the triode from its first crude beginning to its present-day form, it soon became apparent that this type of tube was handicapped by its relatively low amplification factor. Triodes are commonly used in voltage amplifiers, and, as will be shown later, the maximum theoretical gain of an amplifier is given by the amplification factor of the tube. Thus it can be readily understood why tubes with high value of μ might be desirable. It has usually been found impractical to use triodes with μ greater than 100. They can be built, but their operation in the usual amplifier circuits is far from satisfactory.

If we examine the schematic diagram of a triode as pictured in Fig. 4-13, we see that three interelectrode capacitances are shown. These capacitances are actually within the tube itself and are determined by the physical sizes and spacings of the electrodes, lead wires, and base pins. While they may all be important at times, the only one of interest to us at present is the grid-to-plate capacitance C_{gp}, which acts as a link or coupling between the grid and the plate circuits. If the tube is used as an alternating-voltage amplifier, the varying voltages in the grid and plate circuits cause an alternating current to flow through C_{gp} into the grid circuit. The magnitude of this current is proportional to the size of C_{gp}. This feeding of energy from the plate circuit of the tube to the grid circuit is called feedback. Under some circumstances it can cause the circuit to oscillate. The frequency of the oscillations is largely determined by the inductance and capacitance of the circuit. Obviously we should not allow a circuit to act as a generator of an alternating voltage while trying to use it as an amplifier.

Triodes with a high value of μ show much more tendency to oscillate than do those with low values of μ. This can be readily understood. If an amplifier has a high gain, a small alternating voltage in the grid circuit causes a large alternating voltage in the plate circuit. This in turn causes a relatively large current to be fed back to the grid. On the other hand an amplifier with low gain has a smaller alternating voltage in the plate circuit, and hence a smaller current flows back to the grid.

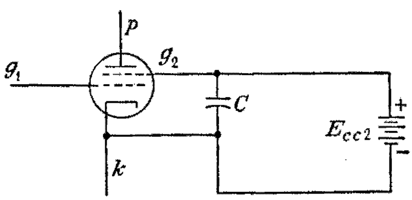


Fig. 4-14.

Thus we see that in order to prevent oscillations in amplifiers, we have our choice of several alternatives. We can use tubes with low values of μ, or we can find some means of reducing the interelectrode capacitance C_{gp}. A third method would be purposely to feed current back to the grid circuit in such a manner as to be 180° out of phase with that fed back through C_{gp} and sufficient to cancel its effect. This is known as neutralization, and it is often resorted to when triodes are used at radio frequencies.

4-9. The Tetrode. About twenty years after De Forest invented the triode, Dr. A. W. Hull decided to try to reduce the undesirable interelectrode capacitance C_{gp} by introducing a fourth element, which he called a screen grid, between the control grid and the plate. He found that, if this grid was held at a positive fixed potential with respect to the cathode but at cathode potential for alternating voltages, the effective capacitance between the control grid and plate was greatly reduced and the tube characteristics changed. In order to have the screen grid with the correct potentials, this electrode can be connected as shown in Fig. 4-14, where C is a large bypass capacitor connecting the screen grid to the cathode.

This capacitor is large enough so that it has negligible reactance at the frequency of any voltage that is to be applied to the control grid of the tube. Since the screen grid is at zero alternating potential with respect to the cathode, it acts somewhat as a grounded plane between the plate and the control grid, thus reducing their mutual capacitance. If the screen grid had been a solid sheet of metal completely surrounding the plate, the capacitance between the control grid and plate would have been zero. Obviously this could not be used since it would stop the flow of electrons to the plate. Actually the screen grid is in the form of a helix surrounding the control grid, and it reduces C_{g1p} to a very small value. Figure 4-15 shows some of the interelectrode capacitances in a tetrode, and while the screen grid causes more interelectrode capacitances to exist, it has reduced the value of C_{g1p}.

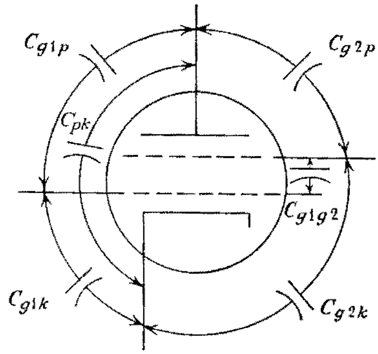


FIG. 4-15. Interelectrode capacitances in a tetrode.

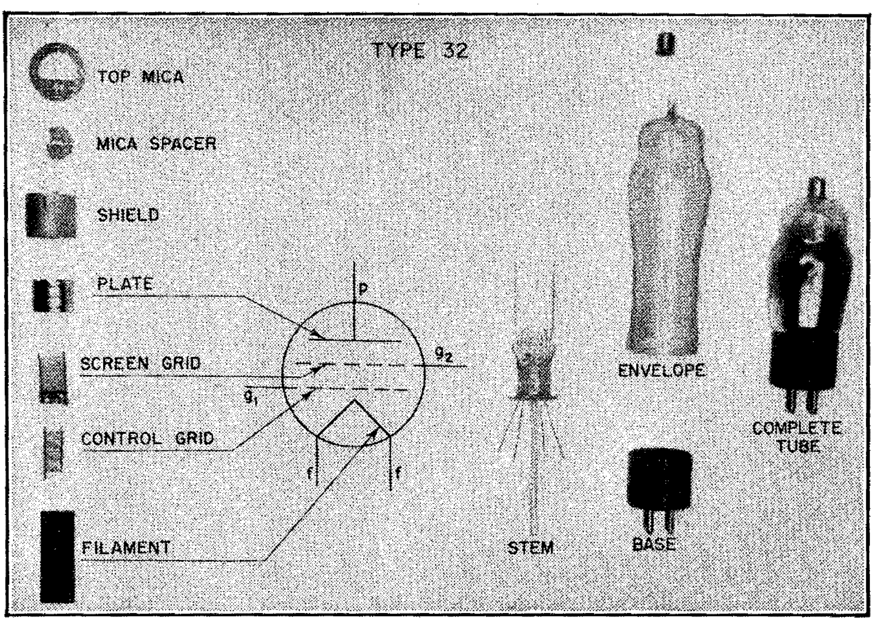


FIG. 4-16. Photograph showing the construction of a type 32 tetrode.

In addition to a second grid some tetrodes have a shield which surrounds the plate and is connected to the screen grid. The purpose of this shield is to reduce the capacitance between the anode and other parts of the circuit external to the tube. Figure 4-16 shows the elements of a type 32 tube. Note that the control-grid lead is brought out the top of the envelope. This is done to reduce the capacitance between the grid lead and the other leads and elements in the tube. Receiving-type

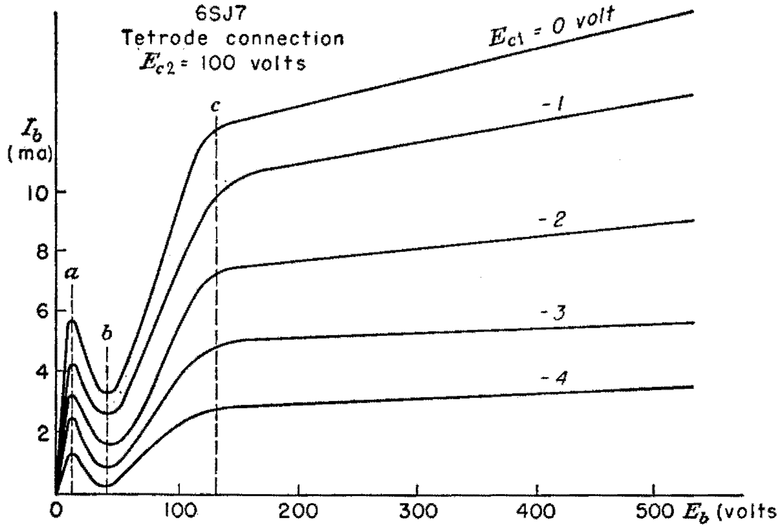


FIG. 4-17. Plate characteristics of a type 6SJ7, tetrode connected.

tetrodes are seldom used in present-day circuits, and hence the 32 is a discontinued type.

4-10. Tetrode Characteristics. Since the plate current in a tetrode is dependent on three electrode voltages E_{c1}, E_{c2}, and E_b, a four-dimensional model would be necessary completely to represent its characteristics. Or, as an alternative, several two-dimensional families of curves may be plotted, each family having one of the variables held constant and another used as a parameter. The most common diagram of this type is the plate family with E_{c2} held constant and E_{c1} the parameter. Such a family is shown in Fig. 4-17.

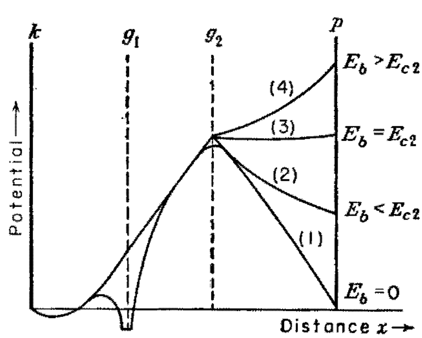


FIG. 4-18. Potential distribution in a tetrode with various plate voltages.

In order to explain the shapes of these curves, let us first examine the potential distribution in a tetrode as pictured in Fig. 4-18. Here it is seen that, in the region between the cathode and the screen grid, the potential distribution is almost identical to that in a triode and that the magnitude of the plate voltage has practically no effect on the shape of the electric field in this region. This is a logical conclusion since, in shielding the control grid from the plate by means of the screen grid, the cathode region is also shielded from the action of the plate. If the screen grid is maintained at a constant positive potential, the relationship between E_{c1} and the cathode current ($I_k = I_{c2} + I_b$) will be similar to that for a triode. This can be shown by plotting a transfer characteristic of cathode current vs. control-grid voltage for a tetrode. This has been done in Fig. 4-19 for various values of E_{c2}. When the electrons making up the cathode or total space current arrive at the plane of the screen grid, a portion of them strike this grid, resulting in I_{c2}. However, the structure of the screen grid is such that the wires are of small diameter and the spacing between them rather large, which results in most of the electrons passing through into the space between the screen grid and the plate. What happens to them there is determined by the relative potentials of the plate and the screen.

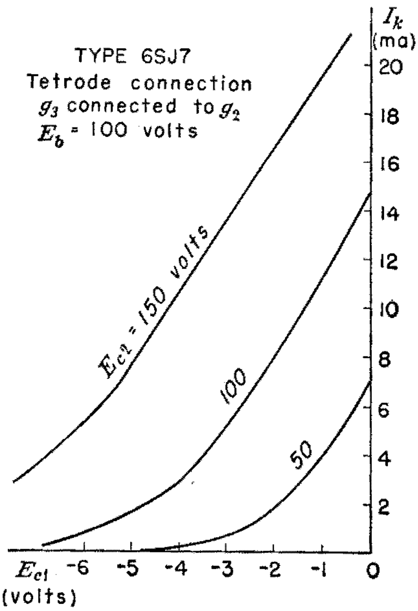


FIG. 4-19. Transfer characteristics for a tetrode—cathode current vs. grid voltage.

Once more referring to Fig. 4-17, the family of plate characteristics for a tetrode, we shall now discuss the shapes of the various portions of the curves. Taking any one of the family of curves shown, say for $E_{c1} = 0$, we note that when $E_b = 0$ the plate current is very small. Any current which flows under these conditions is caused by initial velocity of the electrons and by contact potential difference. If the electrons had started out with zero velocity at the cathode, they would have been accelerated by the positive potential on g_2. When they arrived at the plane of g_2, they would have reached a maximum velocity depending on the magnitude of E_{c2}. If they missed the grid wires and passed into the space between g_2 and p, they would have been decelerated by the field, which has a negative-potential gradient, represented by curve 1 in Fig. 4-18. These electrons would arrive at the surface of the plate with zero velocity and would then reverse direction and be accelerated back toward the screen grid. They might once more miss the wires of the screen grid, pass through, come to a stop, and then again reverse direction of motion. In fact, they might oscillate about this grid a few times before striking it and causing screen-grid current. Electrons, however, do not all leave the cathode with zero initial velocity. Therefore some of them will arrive at the plate with enough energy so that they will penetrate the surface, causing plate current to flow. As the plate voltage is increased, the deceleration of the electrons in the region between the second grid and the plate becomes less. Hence more of them will strike the plate, causing an increased plate current. A value of plate voltage finally will be reached where electrons arrive at the plate with sufficient energy to cause secondary emission.

According to convention, the plate current due to the secondary electrons is a negative current. The net plate current is the algebraic sum of the primary and secondary currents. As the potential of the plate is increased, the primary electrons bombard the plate with higher velocities and hence cause more secondary current, while the primary current remains approximately constant. As a result the plate current decreases with increasing plate voltage as shown in the region *ab* of Fig. 4-17. The slope of this portion of the curve is negative, which means that the dynamic plate resistance is likewise negative. Tetrodes usually are not operated on this part of the curve, which is known as the dynatron region. There are, however, a few special circuits where this negative-resistance characteristic is desirable. One such circuit is the dynatron oscillator.

As the plate voltage is further increased, until it approaches E_{c2}, fewer electrons are attracted to the screen grid. This results in an increased plate current (see region *bc* of Fig. 4-17). When the plate voltage becomes greater than that of the screen grid, as shown by curve 4 of Fig. 4-18, all the secondary electrons return to the plate. Since, as explained earlier, the plate voltage has very little control over the plate current and since, at these higher plate voltages, the secondary current neither adds to nor subtracts from the plate current, the plate characteristic should be flat or slowly rising with increasing E_b. This proves to be the case, as is shown by the portion of the curve to the right of *c* in Fig. 4-17.

When examining the plate characteristic of a tetrode, the student often comes to the conclusion that the flat portion of the characteristic is due to the limitation of emission current from the cathode. This is not the case. The cathode in a tetrode is capable of emitting much more current than the plate and screen grid are able to take care of without overheating. This can be proved by increasing the screen-grid voltage, which will result in an increased space current. The emission current is independent of the electrode voltages and is determined only by the temperature of the cathode [Richardson's equation (2-6)]. In this discussion the temperature of the cathode is held constant; yet the space current increases with screen-grid voltage. This is shown by the transfer characteristics in Fig. 4-19. Therefore, we cannot have reached saturation current for the emitter, and the flat part of the plate characteristic must represent the effect of control of plate current by the plate voltage. This is the usual useful operating region for the tube.

4-11. Tube Coefficients for the Tetrode. From an examination of the plate family we can reach the conclusion that, since the slope of the characteristics is very small in the operating region, the dynamic plate resistance is very high. Measurements of this parameter for small tetrodes show that it has values of a few hundred thousand ohms to more than a megohm. This is much higher than values of r_p for a triode, which in most cases is less than 100,000 ohms. Likewise, we can conclude that, since the control-grid voltage has considerable effect on the plate current while the plate voltage has very little effect, the μ of the tube is quite high. Measurements show that it may be as high as 1000. Since the action of the control grid is similar to the action of the control grid of a triode, the mutual conductances of both types should be of the same order of magnitude. Measurements on tetrodes yield a value on the order of 2000 micromhos.

4-12. Shortcomings of the Tetrode. The tetrode has one distinct shortcoming, and that is that the operating region is limited at the low-plate-voltage end by secondary-emission effect. The first attempt to eliminate this difficulty was to use carbonized plates (a graphite coating on nickel). Carbon, having a higher work function than nickel, had the effect of reducing the secondary emission. Although some improvement resulted from this treatment, other methods gave better results. The most successful of these involved the addition of a third grid, called a suppressor grid. This development was made about 1930.

4-13. The Pentode. Because of their more desirable characteristics, pentodes and beam power tubes have almost completely superseded tetrodes in present-day low-power circuits. This is because of the elimination of the secondary-emission effects which occur when the plate voltage is less than the screen-grid voltage.

The pentode has three grids. The first, or control, grid, which is closest to the cathode, serves the same purpose as the control grid in a triode or tetrode. The second, or screen, grid is the same as the screen grid in a tetrode. The third, or suppressor, grid is usually connected to the cathode and thus sets up a field between itself and the plate such that secondary electrons return to the plate. This action can probably best be seen by referring to Fig. 4-20, which shows the potential distributions in a pentode when the three grid voltages are held constant and the plate voltage is varied. The suppressor grid is connected to the cathode and is therefore at zero potential. Note that, in the region between the suppressor grid and the plate, the potential gradient is always positive, even for small values of E_b. This means that all electrons in this region will be acted upon by forces which move them toward the plate.

A question which is often raised is how and why do the electrons get into the space between the plate and suppressor grid if the latter is a cathode potential. This is easily ansered by assuming that electrons are emitted from the cathode with zero or a small initial velocity and then examining the effects of the various parts of the field between the cathode and the plate.

The portion of the field between the cathode and the screen grid is almost the same as the field between the cathode and the plate in a triode.

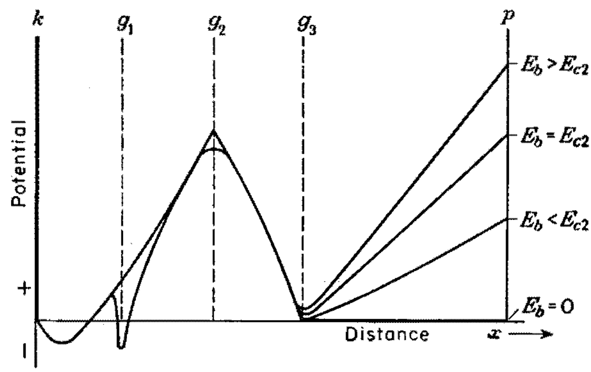


FIG. 4-20. Potential distribution in a pentode.

Therefore the control-grid voltage is able to influence the space current just as it does in a triode. On arrival at the plane of g_2 most of the electrons pass between the grid wires and enter the field between g_2 and g_3. The potential gradient of this field is negative, which means that the electrons are slowed down. They arrive at the point of zero-potential gradient with a velocity at least as great as the velocity with which they left the cathode. This velocity carries them into the space between the suppressor grid and the plate. They are then accelerated toward the plate if E_b has a value greater than zero. If the plate voltage is high enough, secondary electrons will be emitted from the plate but the potential gradient of the field is such as to force them back from whence they came. Thus the secondary-emission effect is eliminated.

Figure 4-21 shows the plate characteristic of a 6SJ7 tube (normally a pentode) operated as a tetrode by connecting the screen grid and the suppressor grid together. Note that this results in a typical tetrode characteristic. When the tube is connected normally (as a pentode), the dip is eliminated and the flat portion extends down to lower values of E_b.

4-14. The Static Characteristics of a Pentode. A family of static plate characteristics can be plotted for a pentode by using a procedure somewhat similar to that used for a triode. The only difference is that the screen grid and the suppressor grid must be maintained at constant potentials with respect to the cathode. Such a family for a 6SJ7 tube is

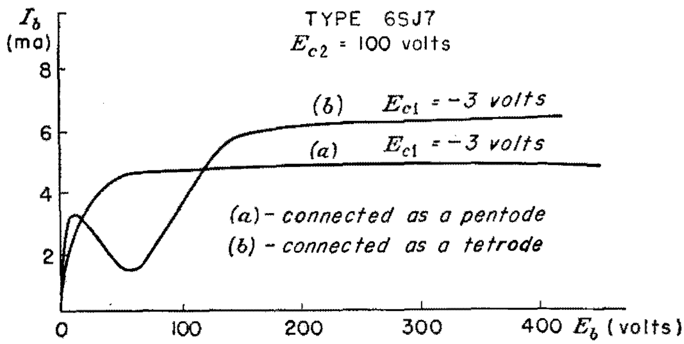


FIG. 4-21. Plate characteristics of a 6SJ7, connected (a) as a pentode, (b) as a tetrode.

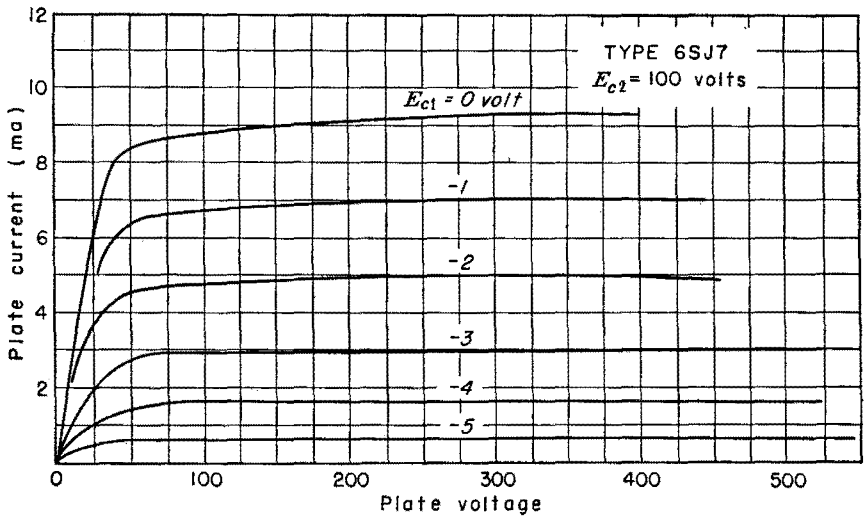


FIG. 4-22. Family of plate characteristics for a 6SJ7 pentode.

shown in Fig. 4-22. The screen voltage in this instance was held constant at a value of 100 volts, and the suppressor grid was connected to the cathode. Note that, as in the tetrode, the plate voltage has very little effect on the plate current in the operating region of the tube. In fact, the plate voltage has even less effect than in a tetrode. This is because of the additional shielding by the suppressor grid. If we plot a family of static transfer characteristics for a pentode, using values of E_b in the operating range as a parameter and with E_{c2} held constant, it will be seen that the family consists practically of a single curve. Such a family is shown in Fig. 4-23.

4-15. Dynamic Transfer Characteristics for a Pentode. If a given load is inserted in the plate circuit of a pentode, as shown in Fig. 4-24, an infinite number of transfer characteristics can be plotted. This is because of the two parameters, screen-grid voltage and suppressor-grid voltage. However, since the suppressor grid is usually connected to the cathode in most circuits, we shall not consider it as a parameter during the remainder of this discussion.

Let us first examine the effects of varying the load resistor R_L, while holding E_{c2} and E_{bb} constant. Such a family is pictured in Fig. 4-25.

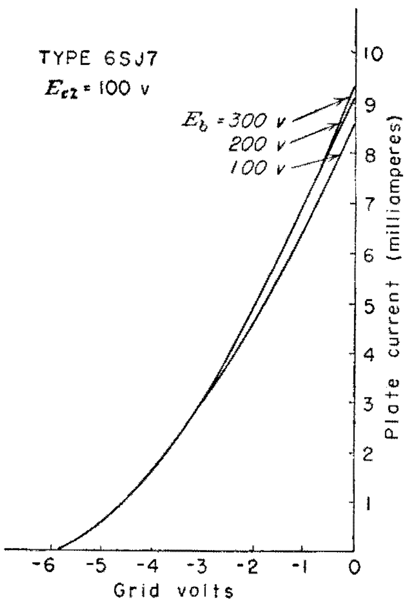


FIG. 4-23. Transfer characteristics for a 6SJ7 pentode.

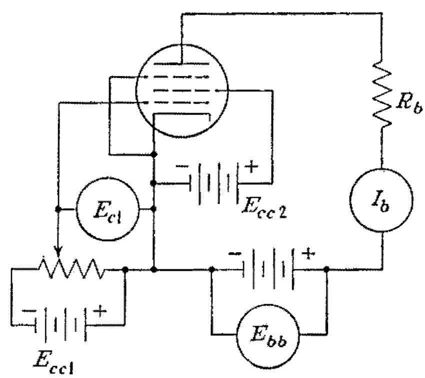


FIG. 4-24. Circuit for obtaining dynamic transfer characteristics.

Note how the straight portion of the curves decreases in length with increasing values of R_L. This seems to indicate that as the load resistance in a pentode amplifier is increased, the magnitude of the grid swing (peak-to-peak value of grid signal) must be decreased if we wish to keep the distortion of the plate-current waveshape down to a reasonable value. For example, the swing for the b curve might be from -2.5 to -4.5 volts, while for the c curve the limits might be from -3.5 to -4.5 volts.

A second set of dynamic transfer characteristics for the 6SJ7 pentode appears in Fig. 4-26. In this instance the load resistance is held constant, and the screen-grid voltage is the parameter. Two load-resistance values are used. For the lower resistance value ($R_L = 25{,}000$ ohms) the curves are somewhat similar to the static characteristics in general appearance. Note that in the negative control-grid region a much higher range of plate-current values can be obtained if the screen voltage is made high. This shows the importance of high screen voltage for power-amplifier operation. For the higher load-resistance value ($R_L = 100{,}000$ ohms) the character of the curves is quite different. The curves all have similar shapes, and they flatten out at the same value of plate current. These latter curves are useful for voltage-amplifier operation. Curves for lower screen voltages have somewhat longer straight portions, but the magnitude of screen voltage used is not very critical.

4-16. Tube Coefficients for a Pentode. Referring to the family of plate characteristics for a 6SJ7 tube, shown in Fig. 4-22, we see that the slope of the curves in the operating region is very small. Therefore the dynamic plate resistance must be very high. It is even higher than for a tetrode. For most pentodes it is of the order of 1 to 2 megohms.

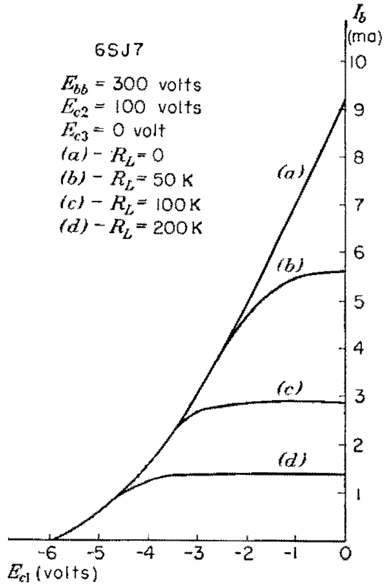


FIG. 4-25. Dynamic transfer characteristics for a 6SJ7 pentode.

Since the control grid in a pentode has about the same effect on space current that it has in a triode, we would expect the mutual conductances to be of about the same order of magnitude. By referring to a tube manual we can easily check that this is true. The value of g_m is usually found to be in the range of 1500 to 3000 micromhos.

The amplification factor for a pentode is not given in a tube manual. However, by means of a vacuum-tube bridge this tube coefficient can be measured and may be found to be as high as 2000 or more. One should not be misled by this large value of μ. It was previously stated that the gain of an amplifier could approach the value of μ as a limit. Theoretically this is true, but practically it is impossible to attain gains much greater than 250, even though the tube may have a μ of 2000. One reason for this is that as soon as we use larger load impedances in order to increase the gain, the power-supply voltage becomes so large as to be impractical. We must therefore be satisfied with less gain than the tube is capable of giving.

It might be well to mention in passing that there are other tube coefficients, besides those mentioned above, which are sometimes useful. For instance, the screen grid and plate have a mutual conductance, which is defined as $g_{g2p} = \partial i_b/\partial e_{c2}$. Likewise the screen grid has an amplification factor, which may be defined as $\mu_{g2p} = -\partial e_b/\partial e_{c2}$. There are other coefficients, but since they are seldom used, no more will be said about them here.

4-17. Remote-cutoff Tubes. So far the only pentode discussed has been the type similar to the 6SJ7, which is a sharp-cutoff tube. It has been named this because there is a rather definite low value of grid voltage which makes the tube cut off. This characteristic is shown in Fig. 4-27, where curve *a* is a static transfer characteristic for a 6SJ7. Curve *b* is for a 6SK7, a remote-cutoff type. Note that for the latter the curve

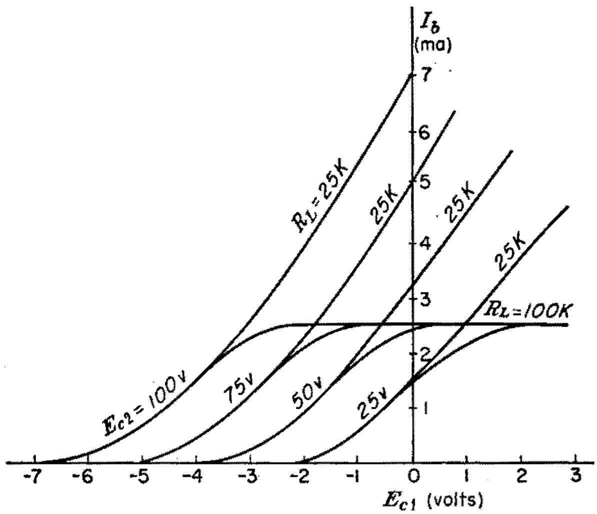


FIG. 4-26. Dynamic transfer characteristics of a 6SJ7 pentode for various screen voltages.

becomes almost asymptotic to the zero-current axis and that the cutoff voltage is at some indefinite remote point. This type is also called variable-μ because of the variation of the amplification factor at various operating points on the characteristic curve.

The remote-cutoff, or variable-μ, characteristic is obtained by designing the tube with a variable-pitch grid winding. In the 6SK7 the grid wires are fairly close together at the top and bottom of the grid but are spaced farther apart near the center. Such a construction allows the grid to stop the flow of current through the upper and lower portions at a certain low negative voltage while current still flows through the center portion. At some higher negative grid voltage the center portion will cut off. The result is a transfer characteristic like that shown in the figure.

4-18. Tube Classification by Use. Tubes may be classified according to the type of service they are to perform. One such method of grouping them is either as voltage amplifiers or as power amplifiers.

Voltage amplifiers are intended for use with high impedance loads in the plate circuit and hence are designed for small values of plate current. They are therefore constructed with small cathodes, grids, and plates. They may be in either glass or metal envelopes and can sometimes be obtained in both versions. The metal tube has the advantage of better shielding from external fields and of slightly lower interelectrode capac-

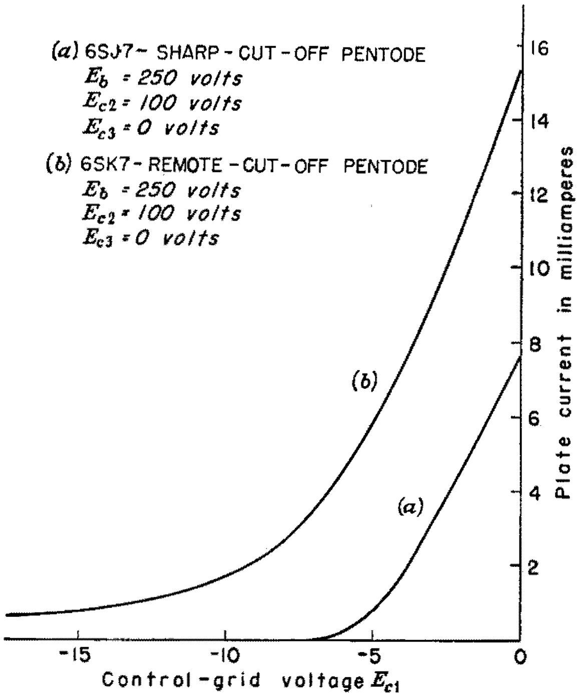


FIG. 4-27. Transfer characteristics for a sharp-cutoff pentode and a remote-cutoff pentode.

itances. Voltage-amplifier tetrodes and pentodes have been made in both sharp- and remote-cutoff types.

Power tubes, on the other hand, are necessarily high-current tubes. They are usually used after one or more stages of voltage amplification and supply relatively large amounts of power to low-impedance loads. They must have large cathodes capable of supplying the currents demanded of them. The currents to the plate and to the screen grid are therefore large, compared with the currents in a voltage-amplifier tube, and hence these elements must be designed with large physical size to dissipate considerable power.

Figure 4-28 is a photograph of the elements from various types of tubes. The control grids from the 6SJ7 and 6SK7 have been enlarged in order to show the difference in the winding of the wires. The heavier construction of the types 6F6 and 6K6, which are power tubes, may be noted.

4-19. Beam Power Tubes. The beam power tube is sometimes called a tetrode although it has the characteristics of a pentode. The difference between it and a true pentode is in the manner in which suppression of secondary electrons is accomplished.

The principal limitation of the power output of the pentode tube is due to the curvature of the lower end of the I_b vs. E_b curves at low plate volt-

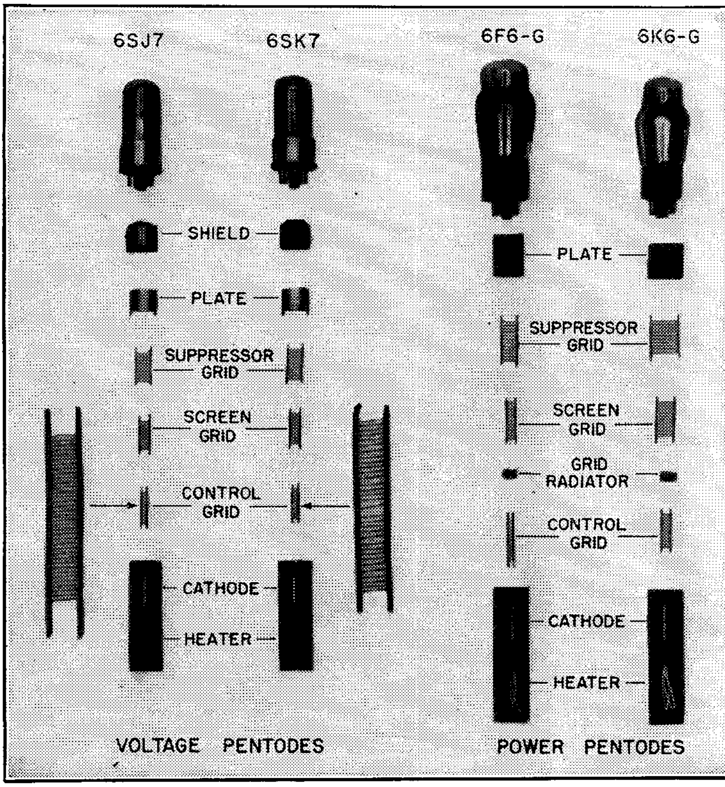


FIG. 4-28. Elements of voltage and of power pentodes.

ages. This curvature causes distortion of the output signal. In order to keep these curves straight over a greater length, the beam power tube was developed. This type of tube accomplished the suppression of secondary electrons by causing the formation of a dense cloud of electrons between the screen grid and the plate, this space charge establishing a field such that it counteracts the action of the screen grid on the secondary electrons.

In order to produce this effect, the cathode must be physically large so that it is capable of emitting copious numbers of electrons. It is also flattened so that it is somewhat oval in shape instead of circular. This causes the electrons to be emitted in a somewhat broad beam. In addition, beam-forming plates, which are shown in Fig. 4-29, are outside of the screen grid and connected to the cathode. These further focus the beam into one of high charge density. This cloud of electrons is so dense that a powerful space charge is set up in the g_2p interelectrode region and acts in such a manner as to repel secondary electrons back to the plate.

Since the space current passed by a beam power tube must necessarily be high in order to accomplish suppressor action, the screen would inter-

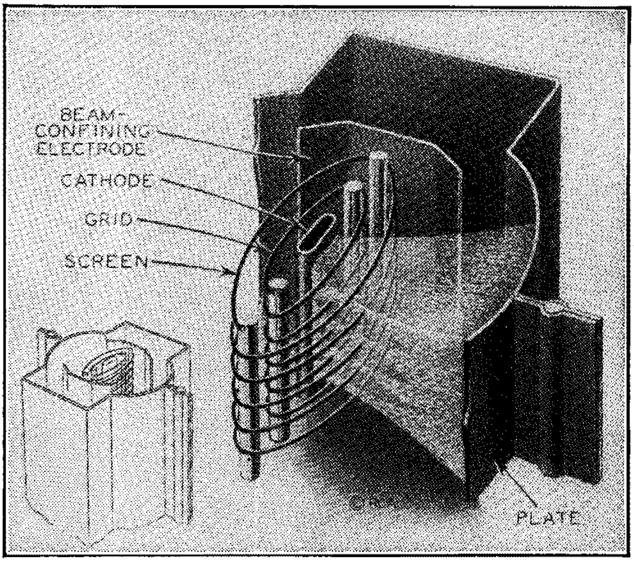


Fig. 4-29. Internal structure of a beam power tube. (*Courtesy RCA.*)

cept large numbers of electrons if precautions were not taken to prevent this by winding the screen grid with the same pitch as the control grid. When the tube is assembled, care is taken to see that the screen-grid wires are in line with the control-grid wires. Thus when plate current is flowing, the screen-grid wires lie in the shadow of the control-grid wires, and they therefore intercept few electrons, and so the screen-grid current is low.

The plate family of characteristics for a type 6L6 beam power tube is shown in Fig. 4-30. Here it can be seen that the curves are fairly straight until they come to the knee, where they fall off very rapidly.

4-20. Miscellaneous Types of High-vacuum Tubes. There are many other types of high-vacuum tubes, some of which are somewhat similar to those already described and others quite different.

A combination of diodes, triodes, and pentodes may be enclosed by a single envelope, and while it may be considered as a single tube, it acts as several tubes. The different plates may all use a common cathode, or each may have a separate cathode. An example of a duodiode triode is the 6SQ7. This tube has a single cathode with two small anodes for the diodes mounted below the triode section. Naturally this tube can be used only in circuits where it is possible to have the cathodes for the three sections common. There are also duodiode pentodes, twin triodes, twin pentodes, diode-triode pentodes, and many other combinations. Some of the triode sections are low-μ and some are high-μ. An examination of a tube manual will soon acquaint the student with these various types.

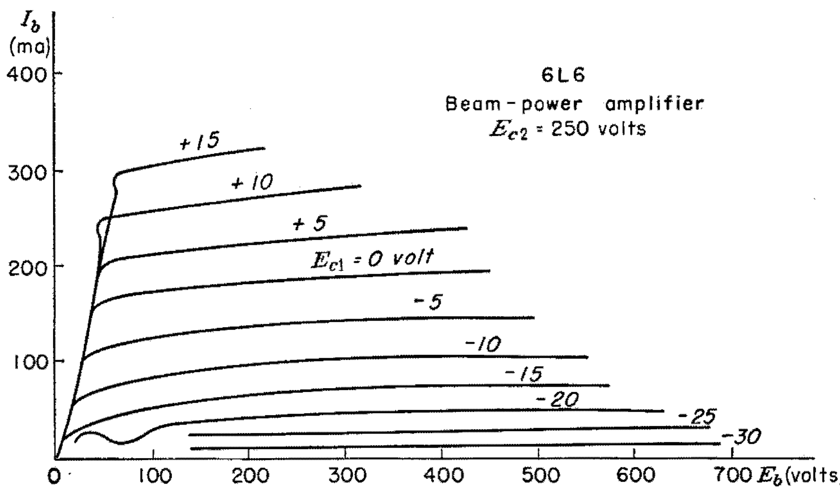


FIG. 4-30. Plate characteristics of a beam power tube.

There are many other types of tubes such as acorn tubes, lighthouse tubes, magnetrons, klystrons, etc., which are used at very high frequencies, but which we shall not describe or discuss, since they are beyond the scope of this book. There are also miniature and subminiature types, which are principally small versions of the ones we have already studied.

REFERENCES

1. E. E. Staff, MIT: "Applied Electronics," John Wiley & Sons, Inc., New York, 1943.
2. Ryder, J. D.: "Electronic Fundamentals and Applications," Prentice-Hall, Inc., New York, 1950.
3. Geppert, D. V.: "Basic Electron Tubes," McGraw-Hill Book Company, Inc., New York, 1951.

PROBLEMS AND QUESTIONS

Characteristic curves for several types of tubes are contained in Appendix A.

1. A hypothetical triode has parallel plane electrodes. The distance from cathode to grid is 1 cm, from grid to anode 2 cm. Relative to the cathode the grid voltage is

-10 volts, the anode voltage $+100$ volts. Sketch approximate graphs of the potential distributions in the interelectrode space (*a*) with normally hot cathode and along a line intersecting a grid wire, (*b*) with normally hot cathode and along a line passing midway between the grid wires. (*c*) Repeat (*b*) but with the plate potential raised to 150 volts. (*d*) Repeat (*a*) but with the grid voltage changed to a value which makes the potential distribution near the cathode approximately the same as in (*c*). Compare the plate currents in parts (*c*) and (*d*).

2. Name three families of tube-characteristic curves which might be useful in the study of electronic circuits.

3. Name and give the mathematical symbols and definitions for the three tube coefficients for a triode. Also write the mathematical relationship between these three coefficients.

4. The plate current in a certain triode is 20 ma when $E_b = 200$ volts and $E_c = -20$ volts. If μ is approximately constant and equal to 9, what should be the approximate plate current if $E_b = 300$ volts and $E_c = -30$ volts?

5. Determine the approximate values of k and n in Eq. (4-3) for a triode-connected 6SJ7 so as to make this equation useful for performance calculations in the region near $E_b = 275$ volts, $E_c = -4$ volts. Use I_b in milliamperes.

6. Use the plate characteristics for a 6J5 to draw a static transfer characteristic for $E_b = 125$ volts.

7. From the plate characteristics for a 6J5 tube graphically determine the approximate numerical values of μ, r_p, and g_m, independently of each other, for $E_c = -6$ volts, $E_b = 160$ volts. Use the relation $\mu = r_p g_m$ to check these answers.

8. (*a*) Use Fig. 4-5 to determine the approximate value of μ for $E_b = 150$ volts, $I_b = 4$ ma. (*b*) Repeat for $E_b = 350$ volts, $I_b = 4$ ma. (*c*) Repeat for $E_b = 250$ volts, $I_b = 9$ ma.

9. Use Fig. 4-5 to determine the approximate value of g_m for $E_b = 250$ volts, $I_b = 9$ ma.

10. Use Fig. 4-5 to find the approximate values of μ, g_m, and r_p for $E_b = 200$ volts, $I_b = 5$ ma.

11. Use Fig. 4-6 to determine the approximate values of μ, g_m, and r_p for $E_b = 200$ volts, $I_b = 5$ ma.

12. Draw the dynamic transfer characteristic for a 6J5 tube for $E_{bb} = 200$ volts and $R_L = 10{,}000$ ohms. The following procedure is suggested: First assume a plate-current value. From this the voltage drop in R_L can be calculated, and hence the voltage from cathode to plate can be found. The plate voltage and current determine a point on the plate characteristics, and hence E_c is determined. This procedure repeated several times gives the dynamic transfer characteristic.

13. What are the two principal shortcomings of a triode?

14. A certain tetrode has $E_{cc1} = -10$ volts, $E_{cc2} = 150$ volts. Sketch the approximate potential-distribution curves for a straight-line path which intersects the wires of the control and of the screen grids, for the following conditions. Label each with identifying letters. (*a*) $E_{bb} = 300$ volts, (*b*) $E_{bb} = 150$ volts, (*c*) $E_{bb} = 0$ volt. Is the potential distribution near the cathode greatly different in the three cases?

15. Data for a tetrode type 24A can be found in a tube manual. (*a*) If the plate and screen voltages are held constant at 250 and 90 volts, respectively, by approximately how much will the plate current increase if the control-grid voltage is increased from -3.0 to -2.9 volts? (*b*) If the control-grid and screen-grid voltages are maintained constant at -3.0 and 90 volts, respectively, by approximately how much will the plate current increase if the plate voltage is increased from 250 to 300 volts?

16. For a type 24A tube $C_{g1p} = 0.007\ \mu\mu f$, and $C_{in} = C_{g1k} + C_{g1g2} = 5.3\ \mu\mu f$. $\mathbf{E}_g = 2/\underline{0°}$ volts at 796 kc. Since the tube amplifies 100 times, $\mathbf{E}_o = 200/\underline{180°}$ volts, where $\mathbf{E}_o$ is the alternating output voltage. Calculate the current $\mathbf{I}_g$ (in microamperes) which flows through the tube capacitances. See Fig. 4-31.

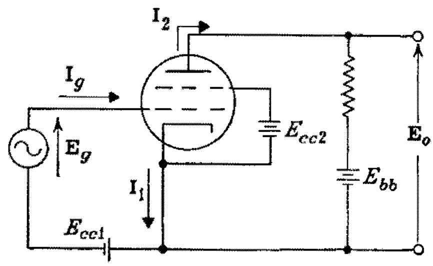


FIG. 4-31.

17. In a certain pentode with plane electrodes the values of the bias voltages are $E_b = 200$ volts, $E_{c3} = 0$ volt, $E_{c2} = 100$ volts, $E_{c1} = -2$ volts. The potential distribution along a certain path is shown in Fig. 4-32. Many electrons leave the cathode with zero velocity. Determine the speed of one of these when it reaches (*a*) the plane of g_1, (*b*) the plane of g_2, (*c*) the plane of g_3, (*d*) the plate.

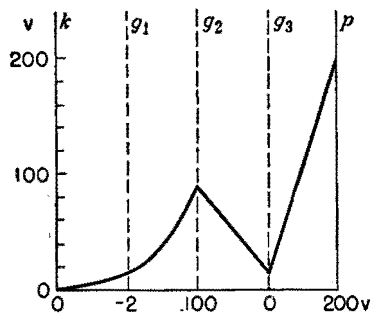


FIG. 4-32.

18. What three methods might be used to avoid oscillations in an amplifier?

19. Use the plate characteristics for a 6SJ7 tube. $E_{c2} = 100$ volts, $E_{c3} = 0$ volt. Draw the static transfer characteristics for $E_b = 50$ volts and for $E_b = 300$ volts.

20. Use Fig. 4-22 to determine the approximate value of g_m for $E_b = 200$ volts, $E_{c1} = -3$ volts. Can these curves be used to determine the values of r_p and μ?

21. Draw the dynamic transfer characteristic for a 6SJ7 tube with $E_{bb} = 200$ volts, $E_{cc2} = 100$ volts, $E_{cc3} = 0$ volt, $R_L = 25{,}000$ ohms. Use the procedure suggested in Prob. 12. Can you suggest a better method?

22. In the manufacture of a type 6SJ7 tube some of the wires of the control grid are bent apart because of rough handling. What happens to the tube characteristics?

CHAPTER 5

SOME APPLICATIONS OF VACUUM TUBES

5-1. Radio Communication. Now that we have some notion of how an electronic tube works, let us begin to find out how to make it useful. Its usefulness depends upon its characteristics, and it must be incorporated into an electrical circuit which will properly exploit these properties.

Let us begin with an application with which most readers have some acquaintance, communication by radio, for example. Since sound waves in air lose their energy rapidly as the distance from the source increases, and since noise is likely to be at an interfering energy level, direct vocal communication can be held over only short distances. If a microphone is used to convert sound energy at the source into electrical energy, wires may be employed to convey some of this energy to a distant receiver which reconverts to sound energy again. This transfer of electrical energy is attained by the electric and magnetic fields surrounding the telephone wires. It might be attempted to dispense with these wires and to create a widespread electromagnetic field at the transmitter, varying at the sound frequency. This could be done by using an antenna. A similar antenna at the receiver could be used to extract some energy from the varying field. However, it is found that this scheme works for very short distances only. Investigation shows that for the successful radiation of energy, the antenna should have certain dimensions approaching a quarter wavelength of the electromagnetic radiation. For a reasonable antenna size this requires the frequency of the signal to be quite high, say 50,000 cps or more. Since the human voice has far too low a frequency content to meet this requirement, some frequency-changing device is needed. Also another frequency changer (detector) is necessary at the receiver to get the audible voice frequency again.

If the sinusoidal waveform for a current is examined, it is conceivable that two changes can be made in it. First, the amplitude can be varied, and, second, the spacing between waves can be changed. Figure 5-1 shows a radio-frequency-current waveform before (a to b) and after (b to c) it has been amplitude-modulated at an audio frequency, while Fig. 5-2 shows the same radio-frequency current when it is angle-modulated. It can be proved mathematically that either of these types of variation or modulation results in the formation of new frequencies which differ from both the original radio frequency (carrier) and the audio frequency and lie in a band which includes the carrier.

For the case of Fig. 5-1 notice that the audio-frequency (a-f) waveform (assumed sinusoidal here) may be plainly discerned in the envelope of the upper extremities of the modulated radio-frequency (r-f) waveform. If the amplitude of the audio signal is increased, the distance between peak and trough of the envelope increases further. Thus amplitude modulation can convey in a high-frequency (h-f) waveform both pitch and intensity of a signal to be transmitted.

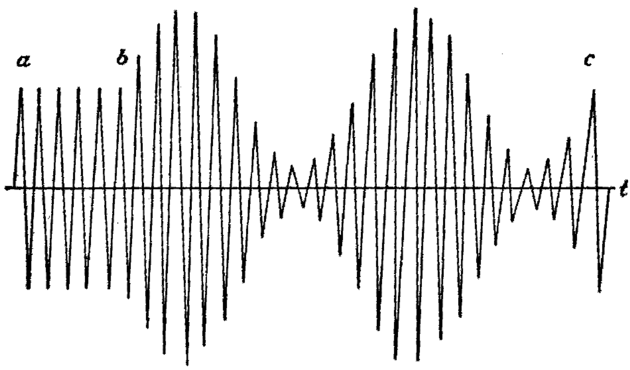


Fig. 5-1. Amplitude-modulated wave.

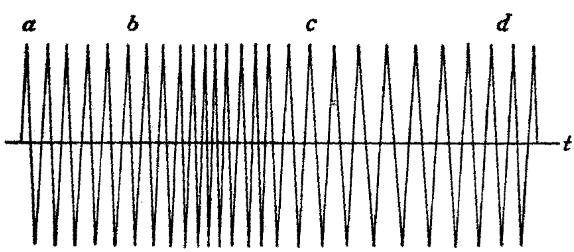


Fig. 5-2. Angle-modulated wave.

In Fig. 5-2 the unmodulated r-f-carrier waveform is shown from *a* to *b*. From *b* to *c* the frequency is varying but is higher than that of the carrier, while from *c* to *d* the frequency is lower. The frequency thus varies higher and lower than the carrier at the rate of the audio signal. The audio amplitude determines the extent of the frequency variation. Thus this system too is capable of conveying both pitch and intensity of an audio signal.

5-2. A-M Transmitter System. Figure 5-3 shows a so-called block diagram of one amplitude-modulation (a-m) transmitter system. The microphone produces electrical energy of frequency and amplitude corresponding to those of the sound-pressure waves. Many types of microphones have been devised. A carbon microphone is simple, gives more a-c electrical energy than the energy contained in the intercepted portion of the sound wave, and is used in ordinary telephone work. Its reproduction lacks the faithfulness desirable in radio practice. Other types include condenser, moving coil, ribbon, and crystal. They do not employ electron tubes and need not be discussed here except to say that they produce an output voltage more or less closely corresponding to the sound-pressure wave. The output voltage is never more than a fraction of a volt and always needs to be amplified by a voltage amplifier, often to a value as high as hundreds of volts. This latter voltage is applied to an amplifier designed to deliver a heavy current at a high voltage. This

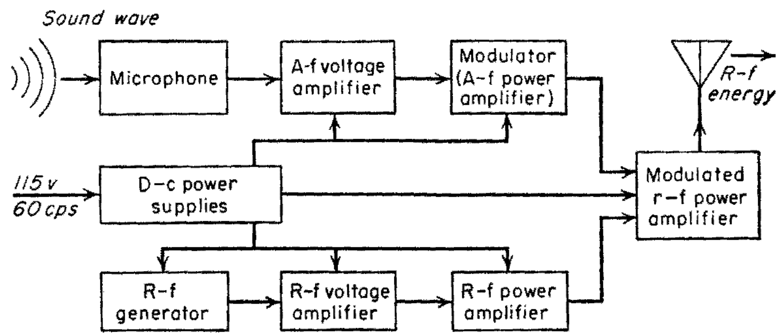


FIG. 5-3. An a-m transmitting system.

amplifier is called a power amplifier. As the frequencies are still those of the audio signal, this power is fed into a frequency changer.

The r-f generator usually is a vacuum-tube type of oscillator, a device which changes d-c energy into a-c energy at almost any frequency desired. The output voltage being rather small, an r-f voltage amplifier raises its value and at the same time serves as a buffer to prevent the action of the r-f power amplifier which follows from affecting the action of the oscillator, which might otherwise vary the frequency of its output. The r-f power amplifier is needed because the frequency-changing device used here requires r-f power input as well as a-f power input. The two frequencies are mixed in the modulated r-f power amplifier, which must be nonlinear so that different frequencies will be produced. A large h-f current having a waveform represented very simply in Fig. 5-1*b* to *c* now flows into the antenna, where a magnetic field and an accompanying electric field with corresponding variations are formed. This electromagnetic energy is radiated in directions depending upon the antenna configuration, and we shall suppose that some of it is intercepted by the antenna system of the receiver.

5-3. An A-M Receiver System. The voltage induced in the receiving antenna by the desired electromagnetic wave is often extremely small. If it considerably exceeds that induced from noise sources (atmospheric static, neon signs, automobile ignition, etc.) and from other transmitters in the same frequency range, it can be made to deliver satisfactorily the signal contained in its modulation.

The receiver system shown in block-diagram form in Fig. 5-4 has the induced antenna voltage applied to the input of an r-f filter, which can be adjusted to reject to some extent all signals having their energies outside the desired frequency band. A tunable *L-C* circuit is commonly used for this. Next, an r-f amplifier raises the voltage to a higher level, more desirable for operating the mixer (frequency changer), and at the same

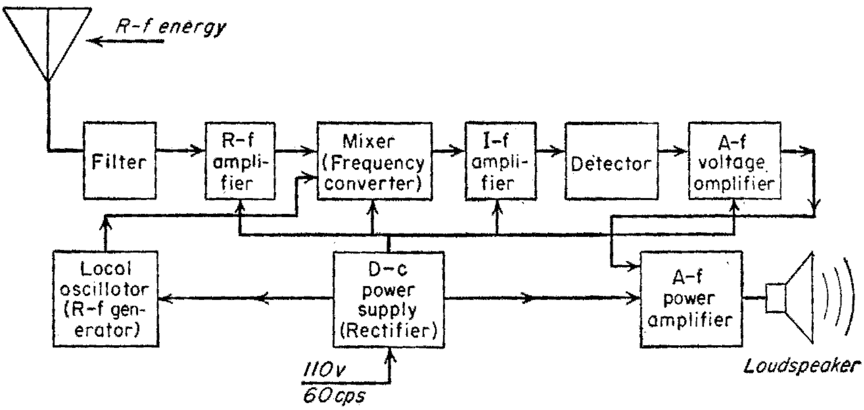


FIG. 5-4. Block diagram of a radio receiver.

time further filters out the undesired signals. The type of frequency changer used here mixes the modulated r-f signal with an r-f signal from a local oscillator to give a new modulated r-f (called intermediate-frequency) signal lying in a different band, where it will be further amplified and filtered by the intermediate-frequency (i-f) amplifier. The advantage of this system lies in the fact that all desired signals will be changed from their original r-f bands to the same i-f band, and it is possible to design the amplifier handling this fixed band of frequencies to do the amplifying and filtering much better than if it were to be an adjustable one.

The modulated signal in the i-f band is now perhaps as great as 10 volts rms and under favorable conditions is almost wholly free of any interference. Its waveform envelope may be imagined to be that of Fig. 5-1. Now we want to get a voltage having the waveform of the upper envelope, without the r-f variations. The demodulator, or detector, is used for this purpose. Then an a-f voltage amplifier raises the voltage to a level sufficient to operate the power amplifier, which in turn drives a loudspeaker.

In both the transmitter system and in the receiver, batteries or power supplies are necessary for operating the tubes. Sometimes rotating machines are used to convert alternating current to direct current, but more often rectifiers and filters are employed.

5-4. Background Material. Let us investigate a few devices, including some introduced in connection with a-m systems. These devices have a great many applications, even outside the communications field. The study of vacuum-tube circuits is somewhat complicated because they contain d-c supplies and, in addition, generators which supply either periodic or nonperiodic variable voltages. Although the circuits external to the tube may be approximately linear, the tube itself has a nonlinear characteristic in general and the total currents which flow usually do not have the same waveforms as do the applied voltages. However, problems of design and analysis we are leaving for later chapters. Here we wish to get acquainted in a general way with some vacuum-tube circuits. Thus we hope to gain a desirable background for the discussions which follow.

In all types of vacuum tubes the tube characteristics determine the applications to which the tube can be put. For example, the most useful property of a diode is its unilateral conductivity. Anode current flows when this element is positive relative to the cathode; no current flows when the voltage is reversed. For a triode two properties are most important. First, if the operation is conducted in what is called a linear region, where the plate-current vs. grid-voltage characteristic curve is approximately straight, then alternating voltages applied between cathode and grid produce an anode current with the variable component of approximately the same waveform as the alternating grid voltage. This property is useful in amplifiers. Second, if the operation is conducted in a nonlinear region, where the characteristic of plate current vs. grid voltage is sharply curved, the tube becomes a distorter. Distortion can prove quite useful: it makes possible detection, modulation, frequency changing, limiting, and many other processes. In the circuits which follow, the appearance of the circuit does not always indicate what it will do, as the region of tube operation is also quite important.

No attempt is made in this chapter to give elaborate details of how a circuit works. In many cases waveforms are drawn to indicate what the circuit does. Arrangements for heating the cathodes are omitted from the diagrams in all cases.

5-5. The Diode Rectifier. The diode is most often used as a rectifier or as a detector. Figure 5-5 shows a simple half-wave rectifier circuit. The anode- or plate-current waveform shown is representative if the load is a resistor. The usual load, however, would be a storage battery under charge or some other electrochemical device where a current having a positive average value is needed and a pulsating waveform is not objectionable. For use as a supply of direct current from an a-c source, a filter is needed, and Fig. 5-6 shows a rectifier with a filter. The inductor L permits direct current to pass freely, but it offers high impedance to the varying "ripple" components. The large capacitors C, on the other hand, offer low impedance to alternating current and tend to eliminate it from the load current. The load current thus becomes direct current equal to the average value of the tube anode current. The load in this case is often the tubes of a radio, which is thus enabled to operate from an a-c power line without the use of batteries.

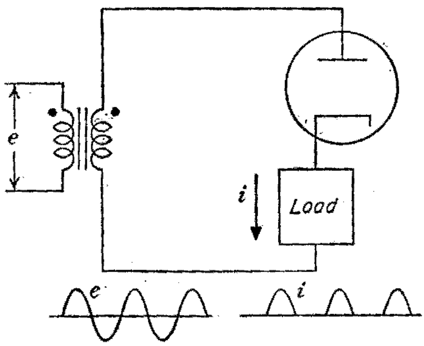


FIG. 5-5. A half-wave rectifier.

5-6. The Diode Detector. The diode is much used as a detector of a-m radio signals. Figure 5-7 shows a circuit diagram for such a detector. The a-m r-f waveform is shown for e_{in}. The envelope of the peaks of the r-f waves represents the a-f intelligence carried by the wave. Note that the circuit is very similar to that of the rectifier of Fig. 5-6. However,

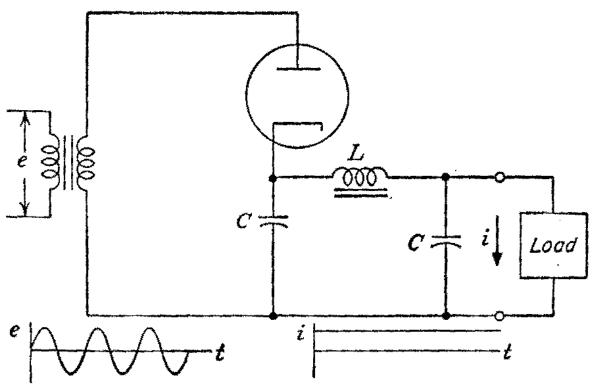


FIG. 5-6. A half-wave rectifier with a filter.

in this case the transformer coil is part of a tuned circuit, which causes a comparatively high voltage to be built up between ground and the anode of the tube for signals corresponding to the tuned frequency and a low voltage for signals differing in frequency from the tuned frequency by more than a small amount. Thus it acts as a filter for high-frequency voltages, allowing only the modulated signal of the desired frequency band to be acted upon by the detector circuit. When e_{in} makes the anode positive relative to the cathode, the tube conducts with a comparatively low voltage drop and the capacitor C rapidly charges to nearly the peak value of the secondary voltage. When the voltage e_{in} decreases from its

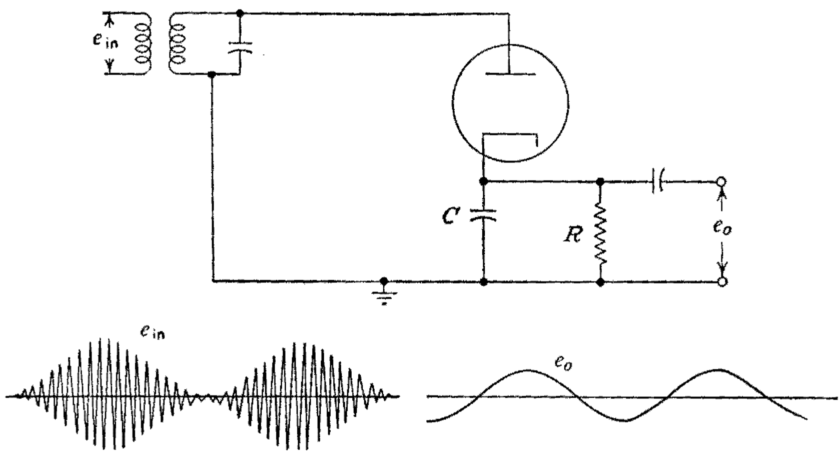


FIG. 5-7. A diode detector.

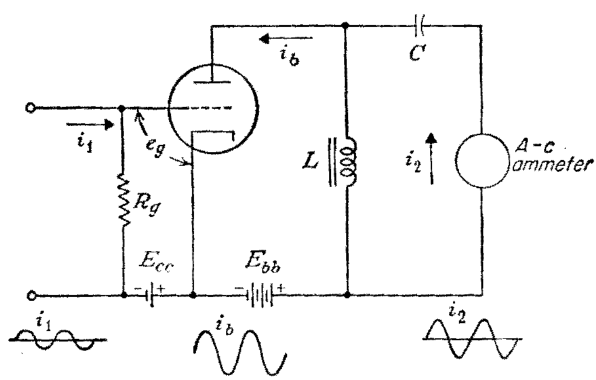


FIG. 5-8. A current amplifier.

peak value, the charge appears to be trapped on C, since the tube is now nonconducting, but C can discharge through R, which it does slowly, R being large in ohmic value. If R and C are properly chosen, the voltage across C rises and falls as the upper envelope of e_{in} rises and falls; but R must not discharge C fast enough to allow the output voltage e_o to follow the rapid r-f variations in the e_{in} waveform. Thus e_o is a voltage of the same waveshape as that of the original audio signal used to obtain the modulated wave e_{in}.

5-7. The Triode Current Amplifier. If a current is too feeble to actuate a certain indicating instrument, it is often expedient to amplify the current rather than to obtain a more sensitive instrument. Figure 5-8 shows a form of current amplifier. The batteries furnish bias to the grid and to the plate in order to place operation of the tube in a desired linear region. The small current i_1, which is to be measured, flows through R_g, producing a variable voltage which alternately increases and decreases the cathode-to-grid potential. The anode current therefore also rises and falls, but the variations are much larger than those of i_1. If it is desirable that only the variable component of i_2 pass through the load, a filter may be used. Its action should be clear from the explanation, given in Art. 5-5, of the functions of the inductor and of the capacitor.

5-8. The Triode Voltage Amplifier. The direct voltages E_{cc} and E_{bb} in Fig. 5-9 place the tube operation in the linear portion of the tube characteristics. Signal voltage v_g varies the cathode-to-grid potential about the value set by E_{cc}, thus raising and lowering the plate current from its no-signal, or quiescent, value. This varying current causes a variable voltage drop in R_b; as the grid potential rises, the plate potential falls. The direct component of the voltage between cathode and plate is prevented from reaching the output by means of a blocking capacitor C. The scales for v_g and e_o in Fig. 5-9 are not the same, since the magnitude of e_o may be up to 50 or more times as great as that of v_g. This ratio is called the voltage gain, or the amplification. The gain usually increases as the value of R_b is made higher, but the practical size of R_b is limited by the voltage E_{bb} available.

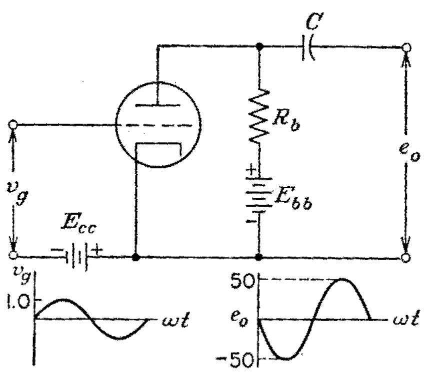


FIG. 5-9. A triode voltage amplifier with fixed grid bias.

5-9. Two-stage Voltage Amplifier with Resistance-Capacitance Coupling between Stages. The voltage gain of the resistance-capacitance-coupled (R-C-coupled) amplifier diagramed in Fig. 5-10 is the product of the gains of the individual stages and therefore may be as high as 2500 or more. The capacitor C_1 is called a blocking, or coupling, capacitor. It prevents the high direct plate potential of tube 1 from improperly biasing the grid of tube 2, while permitting the variable plate voltage to be impressed on the grid. In order properly to bias the latter grid, a d-c path between cathode and grid must exist, and hence R_g is inserted to provide this path. The ohmic value of R_g must be quite high since C_1 and R_g form a voltage divider, and as large a portion of the output voltage of tube 1 as practicable should be applied to the grid of tube 2. Note that a common E_{bb} supply for the two tubes is used, this being the general practice.

5-10. Two-Stage A-F Voltage Amplifier with Transformer Coupling and Cathode Bias. In the circuit of Fig. 5-11 the interstage iron-core

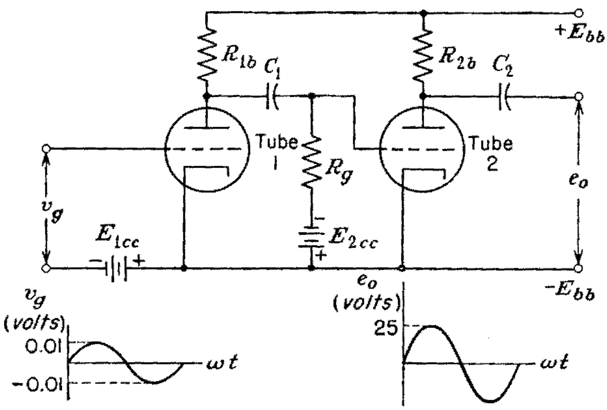


FIG. 5-10. A two-stage R-C-coupled amplifier.

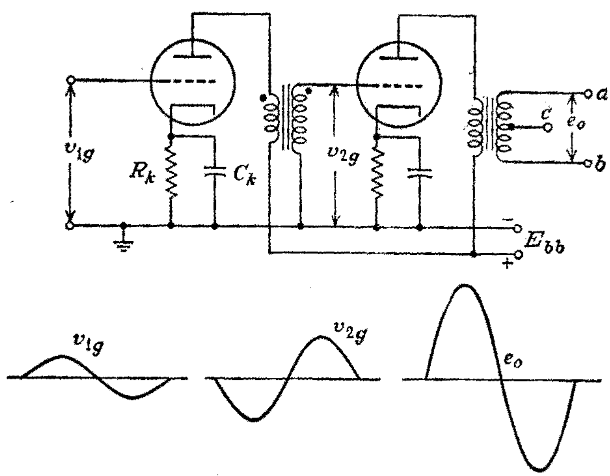


FIG. 5-11. A two-stage transformer-coupled amplifier with cathode bias.

transformer apparently has three advantages over an R-C type of interstage coupler. First, it provides a very high impedance load in the plate circuit of tube 1, which, as we shall see later, results in this tube having a gain almost equal to its μ. Second, it causes very little loss in direct voltage, and therefore a high value of E_{bb} is not needed. Third, the transformer may have a step-up turns ratio, which results in some additional voltage gain. However, these advantages over R-C coupling are offset to some extent by the following facts: The transformer is susceptible to stray magnetic fields which may cause hum in the output. Also the transformer costs more, it is heavier, its frequency-gain characteristic is usually not so good, and it usually must be used with a low- or medium-μ tube, while the R-C-coupled stage may use a high-μ tube. In practice the R-C-coupled amplifier is much more common.

In the circuit of Fig. 5-11, voltage for the grid bias is provided by the anode power supply. If the negative side of E_{bb} is grounded, as is usually the case, the fall in potential across R_k, due to direct plate-current flow, maintains the cathode at some desired positive voltage (bias) above ground. Thus the grid, which is at the same direct potential as is ground, is negative relative to the cathode when the signal is zero. The capacitor C_k serves to bypass the alternating components of the plate current and prevents R_k from becoming an additional a-c load in the plate circuit of the tube. Cathode bias is a very common arrangement, which can be used instead of fixed bias in any of the circuits of Figs. 5-8 to 5-10.

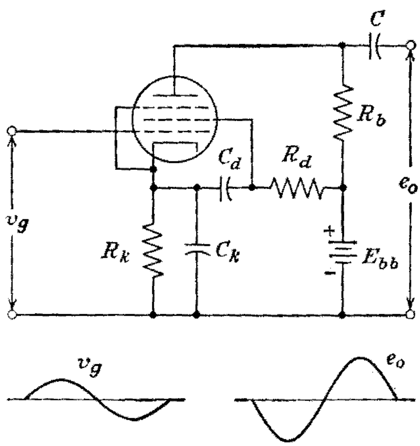


FIG. 5-12. A pentode voltage amplifier.

5-11. The Pentode Voltage Amplifier with Resistance Load. The pentode amplifier of Fig. 5-12 operates in a manner similar to that of the triode. The No. 3 grid usually operates with zero bias, *i.e.*, it is tied to the cathode. The No. 2, or screen, grid usually has a positive bias of about 100 volts when the plate-supply voltage E_{bb} is as high as 300 volts. R_d is a voltage-dropping resistor used to allow E_{bb} to furnish the screen-grid bias. To prevent R_d from becoming a screen-grid a-c load, it is bypassed to the cathode by a capacitor C_d. It will be recalled from Chap. 4 that, when the screen grid is maintained at a suitable fixed positive potential relative to the cathode, μ and r_p are high and also g_m is reasonably high. Thus the gain is high, values up to 250 being readily obtainable. The amplifier in this form is useful mostly for operation over a moderately wide band of frequencies, for example, the range of audible frequencies.

5-12. The Pentode Voltage Amplifier with Tuned Load. For amplifying r-f voltages the gain of circuits like that of Fig. 5-12 is too low. This is because unavoidable capacitance between circuit elements connected to the top of R_b and ground tends to short R_b at high frequencies. This effect can be compensated for to some extent. This is done when the wide band extends into the r-f range. However, it is often desirable to amplify only a narrow band of radio frequencies. The arrangement of Fig. 5-13 then works very well. The input and output tuned circuits act to give high grid and output voltages at the tuned frequency and to discriminate against signals of other frequencies. By properly designing the tuned circuits, very high voltage gain is obtainable. If a triode is sub-

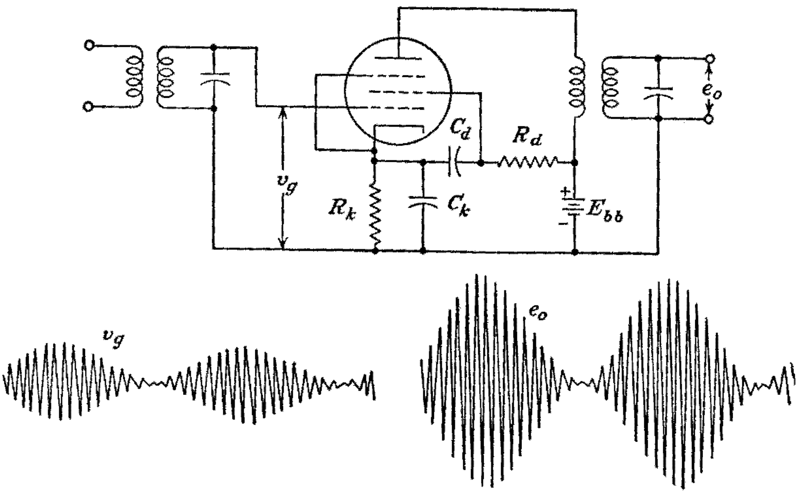


FIG. 5-13. An r-f amplifier.

stituted for the pentode, the circuit will probably oscillate because of the energy feedback from the plate to the grid. This may be avoided if neutralization of some kind is used. However, pentodes are nearly always employed.

5-13. The Power Amplifier. The circuit shown in Fig. 5-14 resembles that of the first stage of the triode voltage amplifier of Fig. 5-11. However, the interests of the designer differ. In Fig. 5-11 high voltage output is wanted, along with small distortion, and both are obtained by using a high-impedance plate load. In this case, however, the interest is in power output, and the load arrangement of the voltage amplifier gives high voltage, little current, and little power. A properly chosen load of moderate size would give less voltage, more current, and more power. Since R is seldom the correct value of load for the tube, a transformer is used, making the load $(N_1/N_2)^2R$, which becomes the correct value if the transformer turns ratio is correct. For large power output the grid must be driven with a high voltage and a tube chosen which has a large cathode

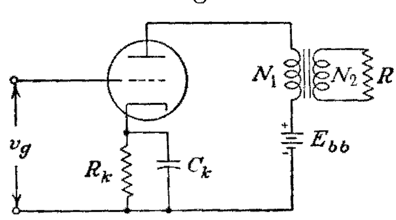


FIG. 5-14. A triode power amplifier.

for ample emission and a large anode to dissipate safely considerable power. Distortion is a problem to be dealt with and has a bearing on the choice of load.

Instead of the triode, a pentode or beam power tube may be used if desired. The load must then be very carefully chosen if high distortion is to be avoided. Also a screen-grid voltage supply must be provided. This is usually the same as the anode supply, since more power output can be obtained with this arrangement. Hence usually no voltage-dropping resistor is used. Although the pentode power amplifier operates with less alternating grid voltage than does a triode and is more efficient, it possesses some drawbacks, which will be discussed later in more detail.

The iron-core transformer, shown in Fig. 5-14, for a-f use, may be replaced by a tuned air-core transformer (see Fig. 5-13) if radio frequencies are to be handled. As might be expected, the design and adjustment of this transformer are not the same for the power amplifier as for the voltage amplifier.

5-14. Classes of Vacuum-tube Amplifier Operation. The plate-circuit efficiency of an amplifier is defined as the ratio of the a-c power delivered to the load to the d-c power furnished by the plate power supply. Any change in circuit adjustment or mode of operation which increases the a-c power or decreases the d-c power will tend to increase the efficiency. Increasing the alternating voltage applied to the grid has the effect of increasing the a-c power output, whereas it may increase the d-c power input only slightly. It seems desirable then to apply voltages of high amplitude to the grid of a power tube. In fact the alternating voltage applied to the grid may be made so great in amplitude that the grid becomes positive relative to the cathode during part of the cycle and possibly swings beyond plate-current cutoff during another part of the cycle. This sort of operation can result in very high efficiency, but the output-current or -voltage waveform is often considerably misshapen compared with that of the grid voltage under smaller-drive conditions. For if the grid becomes positive, grid current flows and any impedance in the grid circuit causes a voltage drop; this results in a decrease in the voltage reaching the grid of the tube. As this occurs only on the positive peaks, there is a tendency to flatten off this portion of the waveform of the grid voltage and hence of the corresponding portion of the output-current and -voltage waveforms. When the grid voltage swings negatively beyond plate-current cutoff, the lower portion of the plate-current waveform is of course flattened. One sees then that high-efficiency operation often brings with it considerable distortion, and one must make a choice as to which is more important or take measures to eliminate the distortion.

Since this matter often comes up for discussion, it is usual to classify the various modes of operation. If plate current flows during all of the cycle, the operation is called class A. If it flows considerably more than half but less than all of the cycle, it is called class AB. Class B means that current flows for approximately a half cycle and class C for less than a half cycle. Under ordinary circumstances the efficiency increases with the classification in the order named, but so does the distortion.

In addition to these classifications, subscripts 1 and 2 are often used. Class 1 means that during the cycle grid current never flows sufficiently to affect the waveform of the grid voltage. This limits the grid-voltage swing in the positive direction to a degree depending upon the amount of impedance in the grid circuit. With power amplifiers it is usually considered that zero is the limit of the grid-voltage swing. In the case of the

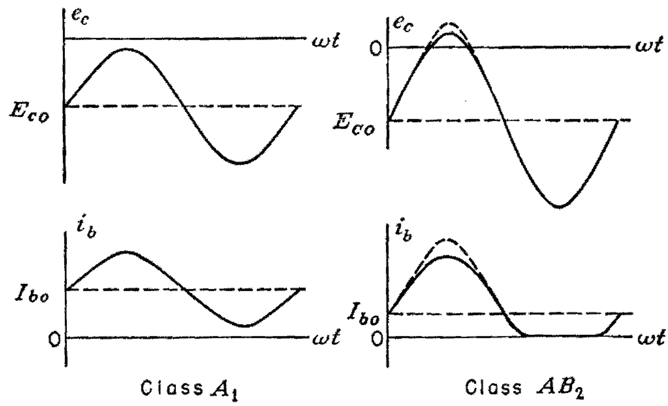


FIG. 5-15. Graphs of grid voltage and of plate current for two classes of operation. The broken-line curves are portions of sine waves drawn for comparison.

first stage of a voltage amplifier a safe limit is about 1 volt less than that, *i.e.*, a grid swing to -1 volt. Operation in which the grid voltage exceeds this limit is referred to by numeral 2. For example, class AB_2 operation means that the grid draws current during part of the cycle and that the plate current flows less than all but more than half the cycle. Figure 5-15 shows typical graphs of grid voltage and of plate current for class A1 and for class AB_2 operation.

Most voltage amplifiers are intended to deliver a voltage much higher than the input voltage and of the same waveform. Efficiency is of no importance, and therefore class A_1 operation is always employed. Audio-frequency power amplifiers are class A_1 if operated single-sided, *i.e.*, with only one tube in each stage. However, when the operation is push-pull (see Art. 5-16) class A_1, AB_1, AB_2, B_1, or B_2 may be employed since the double-tube arrangement patches up the waveform of the output current and voltage. For r-f operation over a narrow band, filters may be used to eliminate some of the distortion-frequency components, and the classes giving higher efficiencies are sometimes usable.

5-15. A Phase Inverter. Using a single-voltage source a phase inverter is employed to furnish an output of two numerically equal voltages 180° out of phase (relative to ground). The second stage in the circuit of Fig. 5-11 can be changed to a phase inverter if a ground connection to the transformer secondary is made to mid-tap c.

One R-C type of phase inverter is diagramed in Fig. 5-16. Tube 1 and its circuit constitute an ordinary amplifier. Suppose an alternating voltage of 1 volt peak is fed to its grid and that the gain is 15. Then e_{1o} is 15 volts peak value and 180° out of phase with v_{1g}, as shown in the diagram. The tap at x is situated at one-fifteenth of the resistance value from the ground end of the resistor, and hence the voltage at x is 1 volt below

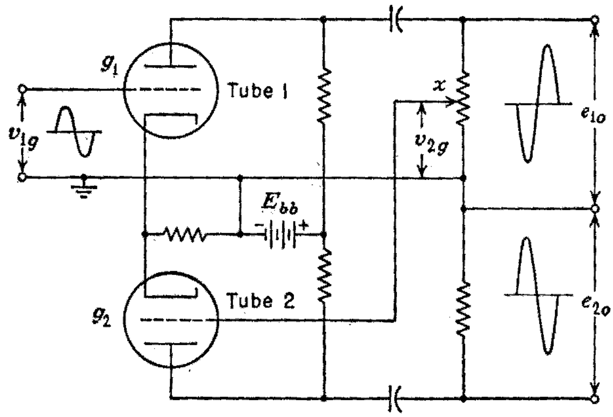


FIG. 5-16. A phase inverter.

ground when g_1 is 1 volt above ground. This voltage at x is applied to the grid of tube 2, which likewise amplifies to produce e_{2o} with 15 volts peak value and with 180° phase shift. Thus e_{1o} and e_{2o} are both 15 volts and 180° out of phase with each other. Tube 2, which reversed the phase of e_{2o} relative to e_{1o}, is called an inverter tube. Sometimes the whole circuit of two tubes is called a phase splitter.

5-16. A Push-Pull Power Amplifier. A push-pull power amplifier such as that of Fig. 5-17a is a type of balanced amplifier which has two main features. (1) It needs to be preceded by a phase inverter or another balanced amplifier. (2) To obtain maximum benefit from the arrangement, it always uses a transformer in the plate circuits. The principal advantages of the push-pull amplifier are twofold. (1) Each tube may be operated under conditions which result in high efficiency but which would produce great distortion with a single-tube circuit. (2) The plate-circuit transformer combines the two tube currents in such a way that the secondary load receives a current or voltage with very little distortion (no even harmonics).

The sketches of Fig. 5-17*b* show first the waveforms of the voltages applied to the two grids. Next are the current waveforms, evincing very great second-harmonic distortion. However, the current i_2 flowing in the secondary has a waveform bearing an excellent resemblance to that of v_{1g}.

5-17. An Amplifier with a Cathode Load. If an impedance is placed between cathode and ground, it is in the cathode-to-grid circuit as well as the cathode-to-plate circuit. This placement is described as being in

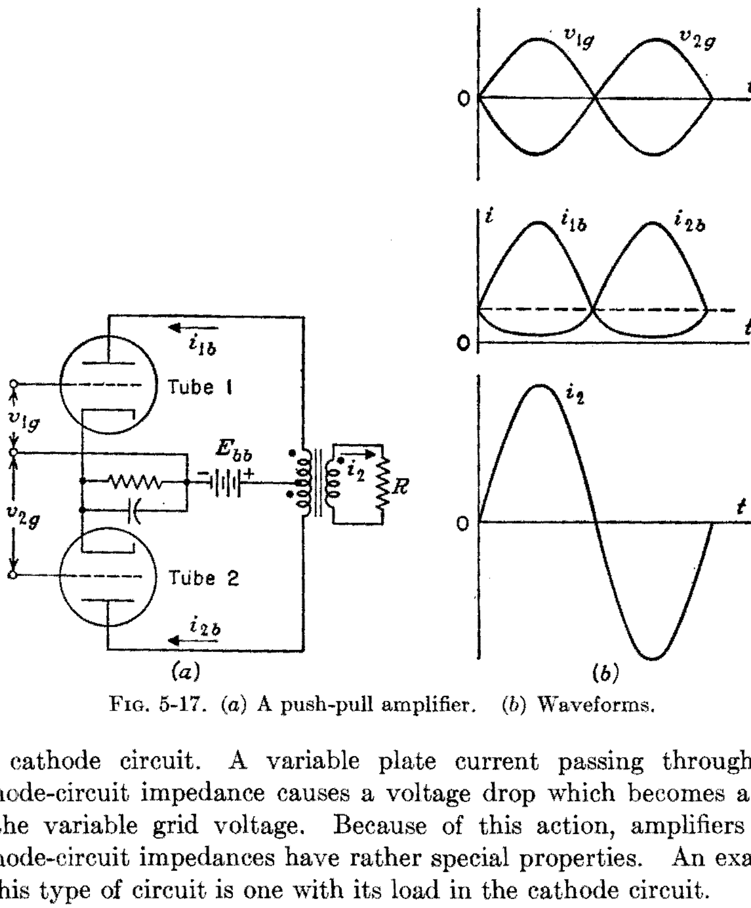


FIG. 5-17. (*a*) A push-pull amplifier. (*b*) Waveforms.

the cathode circuit. A variable plate current passing through the cathode-circuit impedance causes a voltage drop which becomes a part of the variable grid voltage. Because of this action, amplifiers with cathode-circuit impedances have rather special properties. An example of this type of circuit is one with its load in the cathode circuit.

A commonly used type of amplifier with a cathode load is shown in Fig. 5-18. By using a power tube and a transformer-coupled load this could be used also as a power amplifier. The principal features of this type of amplifier are as follows: (1) The input impedance of the grid circuit is very high. (2) The output impedance (looking back from the output terminals) is very low. (3) The distortion, even for large input signals, is very low. (4) The gain is less than unity. (5) There is no phase shift between input and output. The circuit is known as a cathode follower because of the fact that the potential of the cathode follows that of the grid.

5-18. A Feedback Oscillator. An oscillator is an electronic-tube circuit used for changing a direct voltage into an alternating one. Its mechanical analogue is the clock with its weights, pendulum, escapement, and gear train. A pendulum, once set in motion by being given a small

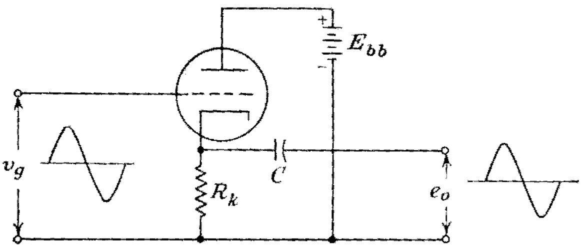


FIG. 5-18. A cathode-follower voltage amplifier.

displacement, swings in an approximately simple-harmonic manner, but the amplitude of swing decreases gradually because of friction and windage. The escapement applies a force to the pendulum in synchronized pulses, and thus the swing continues with constant amplitude. Figure 5-19 shows an L-C parallel circuit which receives a pulse of unidirectional current every time the switch is momentarily closed and opened again. Following each pulse, the current i_t, which circulates through the coil and capacitor combination (tank), is alternating with a frequency depending upon L, C, and the circuit resistance, and this current diminishes in amplitude each cycle because of loss of energy. However, if the switch is closed in proper timing with the oscillations in the tank, an approximately sinusoidal current of constant amplitude can be maintained.

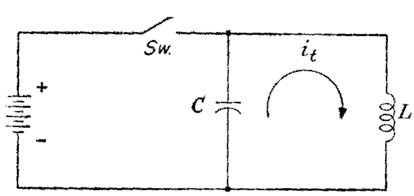


FIG. 5-19. A parallel-resonant circuit.

Figure 5-20 shows one form of feedback oscillator. L_1 and C_1 constitute the tank circuit. Any random impulse in the tube causes a change in the plate current. This causes a small sinusoidal current to circulate in the tank, and a voltage is induced in L_2. The two coils are so wound and placed that when the upper end of L_1 is positive, the lower end of L_2 is also positive (observe the two dots). Hence, the cathode-to-grid voltage is positive when the cathode-to-plate voltage is negative. This grid voltage results in a plate voltage which is amplified and reversed in phase. Hence, the original impulse, now strengthened in amplitude, excites a larger tank current, which in turn induces a larger grid voltage, and so on, until a condition of equilibrium is finally reached. At this point, the tank current is sinusoidal and large, as is also the grid voltage.

5-19. A Class C Amplifier. The idea expressed in the explanation of the feedback oscillator can be employed to make a very efficient controlled power converter. Instead of deriving a grid signal from the out-

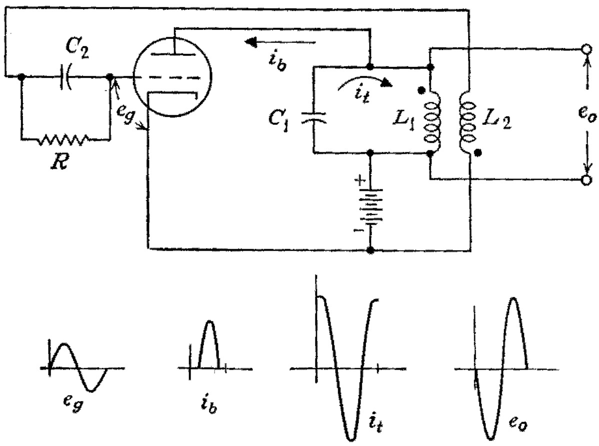


FIG. 5-20. A feedback oscillator.

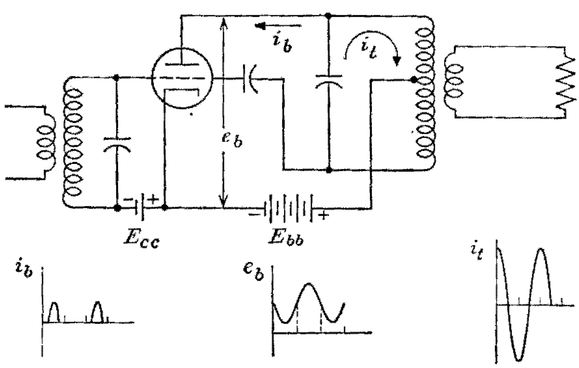


FIG. 5-21. A class C r-f power amplifier.

put tank, it may be obtained as the output of another amplifier. Thus, in effect, the opening and closing of the switch in Fig. 5-19 is externally controlled. The energy fed to the tank comes from the plate power supply. If the closing and opening of the switch occurs at a time in the cycle when the voltage across the tank is near its maximum, the voltage across the switch is a minimum and less energy is lost in arcing at the contacts. Turning from the analogue back to the tube, the efficiency of power conversion is high if the plate current flows in short, heavy spurts, timed when the plate voltage is low. To achieve this operation the plate-supply voltage is made high and the grid-bias voltage set far beyond cutoff. A very high grid drive is employed, which causes a large plate current to flow for perhaps one-third of a cycle. If the plate tank is tuned to agree with the frequency of the grid signal, the tank responds to the excitation with a very large circulating current of approximately sinusoidal waveform. A practical class C amplifier circuit and waveforms of voltages and currents important in its operation are shown in Fig. 5-21. The connection from the lower end of the output transformer back to the grid is provided to give neutralization and thus to prevent the amplifier from acting as an oscillator.

Since the class C amplifier needs a tuned-plate load in order to give a sinusoidal output from a sinusoidal input voltage, it is useless for audio frequencies. It is employed to amplify only unmodulated, key-interrupted, or f-m r-f voltages. It can also be used in the production of a-m signals.

REFERENCES

1. "The Radio Amateur's Handbook," American Radio Relay League, Inc., West Hartford, Conn., annual.
2. Henney, K.: "Electron Tubes in Industry," 1st ed., McGraw-Hill Book Company, Inc., New York, 1934.
3. Jordan, E. C., P. H. Nelson, W. C. Osterbrock, F. H. Pumphrey, and L. C. Smeby: "Fundamentals of Radio," Prentice-Hall, Inc., New York, 1942.

PROBLEMS AND QUESTIONS

1. (*a*) Sketch an a-m wave where the audio signal is small and of constant amplitude and frequency. (*b*) Repeat (*a*) for the same conditions except that the frequency of the audio signal is doubled. (*c*) Repeat (*a*) for the same condition except that the amplitude of the audio signal is doubled.

2. Which determines the extent of the frequency variation in an angle-modulated wave, the audio signal's amplitude or its frequency?

3. Why is a system of filters or tunable *L-C* circuits needed in a radio receiver?

4. Most broadcast (a-m) receivers use diodes to perform two functions. Name them.

5. A diode detector circuit and its input-voltage waveform are shown in Fig. 5-7. (*a*) C and R are proper sizes for correct detector action. Sketch one cycle of the voltage across R. (*b*) C and R are made much too small for proper detector action. Sketch the waveform of the voltage across R. (*c*) C and R are made much too large for proper detector action. Sketch the waveform of voltage across R.

6. A current amplifier uses the circuit of Fig. 5-8. $R_g = 200{,}000$ ohms, $g_m = 2000$ micromhos. The resistance of the a-c milliammeter is negligible so that the plate voltage of the tube remains constant. If $I_1 = 5\mu$a rms, determine the reading of the a-c milliammeter. *Hint:* See Art. 4-6.

7. Instead of having no load resistor as in the case of Prob. 6, a plate load of 100,000 ohms is used in series with the a-c milliammeter. (*a*) Why does the method of that problem not give the meter reading in this case? (*b*) If the voltage amplification is 15, what is the reading of the meter?

8. Which of the following tubes are quite suitable for use preceding an iron-core interstage transformer: 6J5, 6SF5, 6SJ7?

9. (*a*) In Fig. 5-12 state the purpose of each of the following circuit elements: R_k, C_k, C_d, R_d, R_b. (*b*) Why is the suppressor grid connected to the cathode instead of the screen grid? (*c*) What is the purpose of the capacitor C?

10. Which of the blocks in Fig. 5-4 could utilize the device diagramed in Fig. 5-13?

11. For Fig. 5-3, pick circuits from this chapter which could perform the function named in each box. The microphone and the modulated r-f amplifier may be omitted.

12. In Fig. 5-14 the load on the secondary of the output transformer (ratio $N_1/N_2 = 10$) is 10 ohms. What a-c load does the tube see in its plate circuit?

13. The dynamic characteristic for a certain triode power tube is shown in Fig. 5-22. The grid-current characteristic is also shown. The grid bias E_{co} is -30 volts. The load is $R_L = 2500$ ohms. (*a*) The grid signal is sinusoidal and 20 volts peak value. Draw the grid-voltage and plate-current waveforms. (*b*) Determine the average plate current I_{ba}. (*c*) Determine the effective value I_p of the sinusoidal alternating plate current as 0.707 times its peak value. (*d*) Compute the power P_{bb} furnished by the plate power supply as $E_{bb}I_{ba}$. (*e*) Compute the a-c power P_{ac} supplied to the load as $I_p{}^2R_L$. (*f*) Compute the plate-circuit efficiency η_p as P_{ac}/P_{bb}. (*g*) Repeat the above computations for a grid signal of 25 volts peak value. (*h*) Does a larger value of grid signal result in higher efficiency in this example? (*i*) What class is the operation in each of the above cases? (*j*) The grid signal is changed to 40 volts peak value. $E_{co} = -30$ volts as before. What is now the class of operation? (*k*) The bias is changed to $E_{co} = -80$ volts, and a signal of 100 volts peak value is used. What is the class of operation? Why do you think the efficiency might be higher than in any of the above cases?

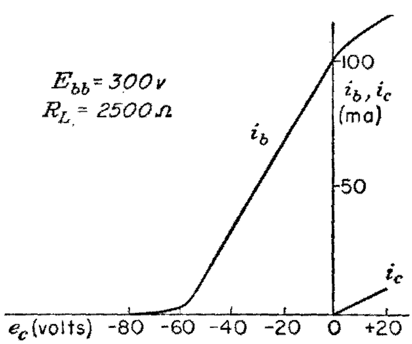


FIG. 5-22.

14. In the cathode follower of Fig. 5-18, if the grid is made 1 volt more positive than normal relative to the cathode, a test reveals that e_o is 15 volts higher than before. What value of v_g is required to cause this to happen and what is the amplification e_o/v_g?

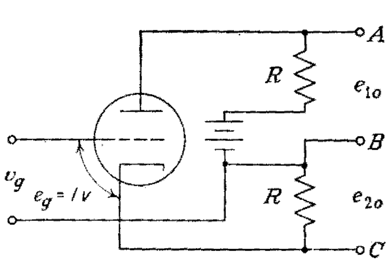


FIG. 5-23.

15. In Fig. 5-23 the operation is class 1. (*a*) When $e_g = 1$ volt, a test reveals that the voltage from C to A is 16 volts. What value of v_g is being applied? (*b*) If v_g is made 4.5 volts, what are the values of e_{1o} and e_{2o}? (*c*) In case (*b*) what is the phase angle between e_{1o} and e_{2o}? What is meant by this phase angle?

16. (*a*) For the circuit of Fig. 5-17 sketch the waveform of the current through the plate power supply. What is the frequency of this current in terms of that of v_{1g}? (*b*) Graphically subtract i_{2b} from i_{1b}. What waveform does the result resemble?

CHAPTER 6

CONCEPTS USEFUL IN VACUUM-TUBE-CIRCUIT ANALYSIS

6-1. Introduction. Electronic apparatus already constructed may be tested in the laboratory for performance. The information obtained is useful in designing new equipment, but often insufficient data are taken to determine what the performance would be if some circuit details were changed. Hence, it is desirable that one be able to make on paper an analysis of the operation of a circuit, as this often most readily determines the approximate influence of any circuit component. Furthermore, a simple mathematical or graphical analysis of a circuit is often cheaper and more quickly made than a laboratory test.

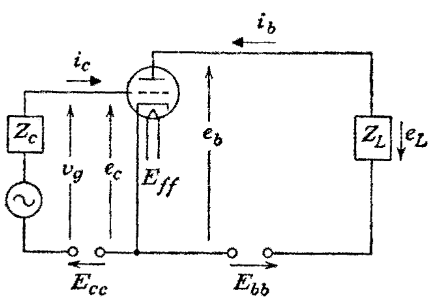


FIG. 6-1. A triode with a plate load.

Any mathematical method of analysis requires symbols for the quantities involved. Many sets of symbols are in current use, and some confusion results from this multiplicity. The symbols* of this book are based upon standards of the Institute of Radio Engineers (IRE). To facilitate understanding, not all of them will be tabulated at once, but some will be introduced gradually as the study continues.

6-2. The Triode with a Plate Load. A rather general representation for the circuit of this type is shown in Fig. 6-1. Symbols for quantities of principal immediate interest are shown. The head end of the arrow, used in connection with an algebraic symbol for voltage, indicates its positive sense. As an example, if $v_g = 5 \sin \omega t$, v_g is positive for values of ωt between 0 and π. During this interval the voltage v_g makes the grid more positive than the bias value, E_{cc}. v_g is negative for values of ωt between π and 2π, and the situation is now reversed from that indicated by the sense arrow. Without the sense arrow one would not know whether a positive value of v_g means the grid potential is above or below the bias value. Closed-head arrows are used to indicate positive

* A table of symbols will be found in Appendix C.

sense of current. If at some moment i_c is negative, then the conventional current is flowing in a direction opposite to that of the arrow. Note that most of the voltage symbols mean rise in potential in a direction from the cathode to some other electrode. The symbol e_L is an exception, as it should agree with the convention of Ohm's law, $e_L = i_b R_L$, which states that the fall in potential in the direction of conventional current flow through a resistor is in phase with the current as well as being proportional to its instantaneous magnitude.

Definitions:

E_{ff} is the effective value of heater or filament supply voltage.
E_{bb} is the direct voltage supplied to the plate circuit.
E_{cc} is the direct voltage supplied to the grid circuit.

E_{bb} and E_{cc} are commonly supplied by either batteries or electronic devices. In either case they may be represented by battery symbols. Usually the internal impedances are so small that they can be neglected.

v_g is the instantaneous value of signal voltage introduced into the grid circuit.

In general v with a suitable subscript is used to represent an applied signal voltage.

e_c is the instantaneous total value of cathode-to-grid voltage rise.
e_b is the instantaneous total value of cathode-to-plate voltage rise.
e_L is the instantaneous total value of voltage drop across Z_L in the direction of positive sense of i_b.
i_c is the instantaneous total grid current.
i_b is the instantaneous total plate current.
Z_L represents the series equivalent of all R, L, and C elements in the plate circuit external to the tube.
Z_c represents the series equivalent of all R, L, and C elements in the external grid circuit.

From the circuit of Fig. 6-1 it is apparent that

$$e_c = E_{cc} + v_g \tag{6-1}$$
$$e_b = E_{bb} - e_L \tag{6-2}$$

6-3. Quiescent Operation of the Triode Circuit. When E_{ff}, E_{bb}, and E_{cc} are active in the circuit of Fig. 6-1 but v_g is zero for a considerable length of time, the operation is said to be quiescent.

Definitions:

E_{co} is the quiescent value of e_c.
E_{bo} is the quiescent value of e_b.
E_{Lo} is the quiescent value of e_L.
I_{bo} is the quiescent value of i_b.
I_{co} is the quiescent value of i_c.

The i_b vs. e_b curves for a typical triode are shown in Fig. 6-2. With no signal applied in the grid circuit, v_g is zero, and E_{cc} and E_{bb} together with

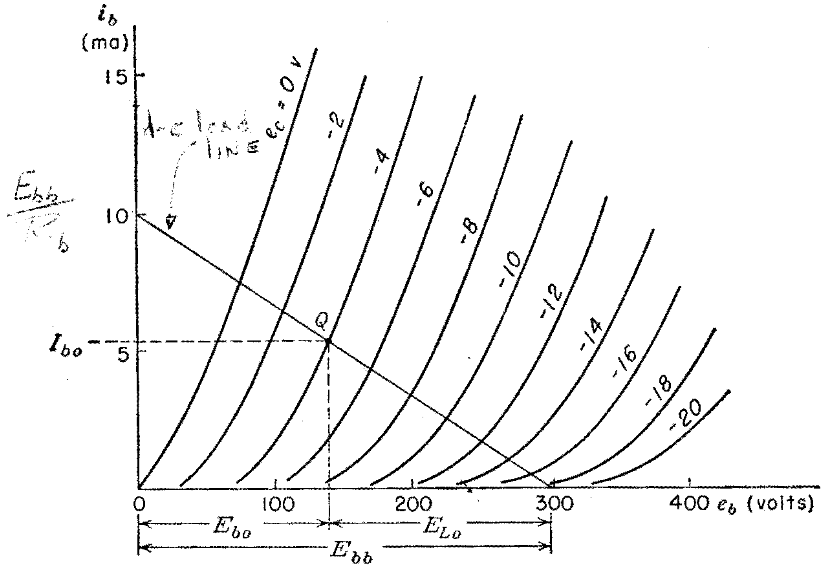


FIG. 6-2. Plate diagram for a triode.

Z_c and Z_L determine the tube operation. The quiescent grid voltage may be obtained from the equation

$$E_{co} = E_{cc} - I_{co}R_c \tag{6-3}$$

In this case R_c is the d-c resistance of Z_c. For the usual case, E_{cc} is negative so that I_{co} is zero, and the quiescent grid voltage or bias E_{co} equals E_{cc}. The quiescent value of plate voltage is given by the equation

$$E_{bo} = E_{bb} - E_{Lo} = E_{bb} - I_{bo}R_{dc} \tag{6-4}$$

where R_{dc} is the d-c resistance in the plate circuit of the tube. The locus of E_{bo}, I_{bo} is obtained by plotting Eq. (6-4), using the axes of e_b and i_b in Fig. 6-2. It is readily determined that this straight-line locus always crosses the e_b axis at E_{bb}, and this point together with any other point may be used to fix the line. This locus is called the static- or d-c-load line.

For any value of direct current through the tube and plate circuit, it shows the voltage furnished by E_{bb}, the part of this voltage used by the load R_{dc}, and the remaining part left as the drop across the tube. The quiescent operating point Q lies on the line, and if E_{co} is known, Q is fixed.

6-4. Signal Voltages. In many industrial electronic circuits the wave-shape of the voltage applied to the grid is sinusoidal. In some others it is unidirectional and slowly changing. In communications circuits a signal of steady character, such as a sinusoid or uniformly and regularly interrupted constant voltage, transmits no intelligence; therefore the waveforms used are irregular and usually do not repeat in cycles. In radar and in television circuits the signal voltages are often in the form of pulses.

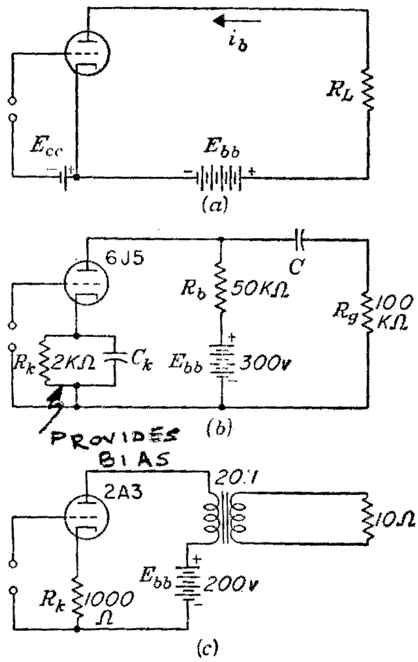


FIG. 6-3. Some common plate-load arrangements.

These various waveforms require different techniques for making a study of circuit behavior. The commonest method, which works well for cyclic signals, is that of using a Fourier analysis (see Art. 6-16) of the signal waveform and applying each sinusoidal frequency component to the circuit. This method will be used throughout most of this book. Usually a single-frequency signal is assumed, but it must be remembered that this is often but one of the many that are applied simultaneously.

Signals are sometimes applied to the plate circuit or, in multielement tubes, to other grids besides the first. In such cases it is important to learn how to handle this aspect of the analysis. In most cases, however, it is assumed that the only signal applied is that to the first grid.

In actual cases, rapidly varying signals usually have zero average value. Even if a d-c component is present, it is often blocked from reaching the grid by a series capacitor or by a transformer. In cases where it is not blocked from the grid circuit, the d-c component is often considered to be part of the grid bias for the tube, and only the remaining alternating part is considered as the signal.

6-5. The Grid-bias Line and the Dynamic-load Line for Resistive Loads. Let us now consider some details about Z_L. Figure 6-3 shows some examples of common load arrangements. In Fig. 6-3a the d-c plate load is R_L and the d-c-load line may be drawn in the manner described in Art. 6-2 and shown in Fig. 6-2. The load for alternating current is R_L, the same as for direct current. As the grid voltage changes, if the locus of points of instantaneous operation is plotted from the relation $e_b = E_{bb} - i_bR_L$, it is easily seen that the a-c-load line is identical with the d-c-load line [see Eq. (6-4)].

In the circuit of Fig. 6-3b numerical values have been assigned in order to make the explanation clearer. Let us assume C and C_k to be so large

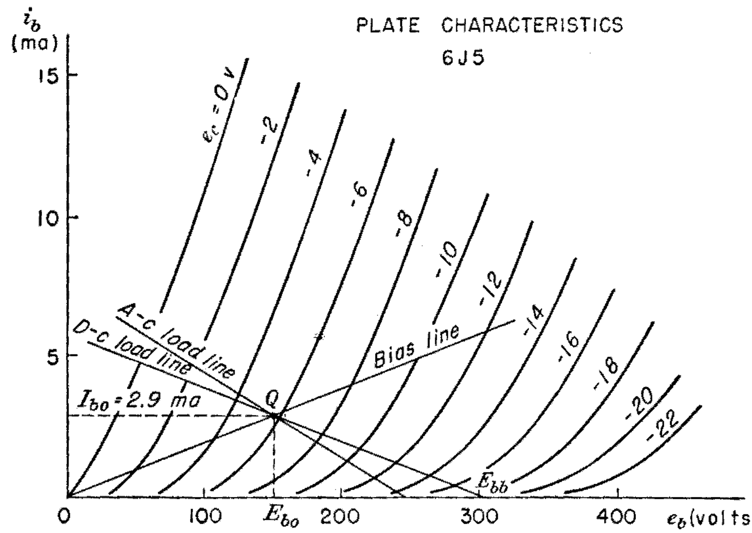


FIG. 6-4. Plate diagram for the circuit of Fig. 6-3b.

that their ohmic values to the sinusoidal current are negligibly small. The d-c plate load is therefore $R_{dc} = R_b + R_k = 50{,}000 + 2000 = 52{,}000$ ohms. The static-load line is drawn in Fig. 6-4. To do this, two points are located, one at $e_b = 300$ volts on the $i_b = 0$ axis and another point, say for $e_b = 100$ volts. This means that the voltage drop in R_{dc}, of $300 - 100 = 200$ volts, is caused by $i_b = 200/52{,}000 = 3.85$ ma. The quiescent operation point Q is on this line, but since the cathode bias caused by plate-current flow through R_k is unknown, Q cannot be located immediately.

Next, we may draw a grid-bias line showing how the bias varies with plate current. For zero grid bias $i_b = 0$ is required. For $E_{co} = -2$ volts, $I_{bo} = 2/2000 = 1$ ma. For $E_{co} = -4$ volts, $I_{bo} = 2$ ma, and so on. These points determine a locus (not usually straight) upon which Q, determined by E_{co} and I_{bo}, must be situated. Q is therefore at the intersection as shown.

The varying part of the plate current follows a path for which R_b and R_g in parallel constitute virtually all the impedance, if C and C_k are very large. The a-c load is therefore 33,300 ohms. Little error is occasioned by assuming that the a-c-load line passes through Q, and a second point may be determined for drawing it. When $i_b = I_{bo} = 2.9$ ma, a voltage drop in the a-c load of $0.0029 \times 33{,}300 = 96.5$ volts results, which means that the line intersects the e_b axis at $e_b = E_{bo} + 96.5 = 150 + 96.5 = 246.5$ volts.

In the circuit of Fig. 6-3c, $R_{dc} = R_k$ if the output transformer has negligible primary d-c resistance. The d-c-load line and the bias line are

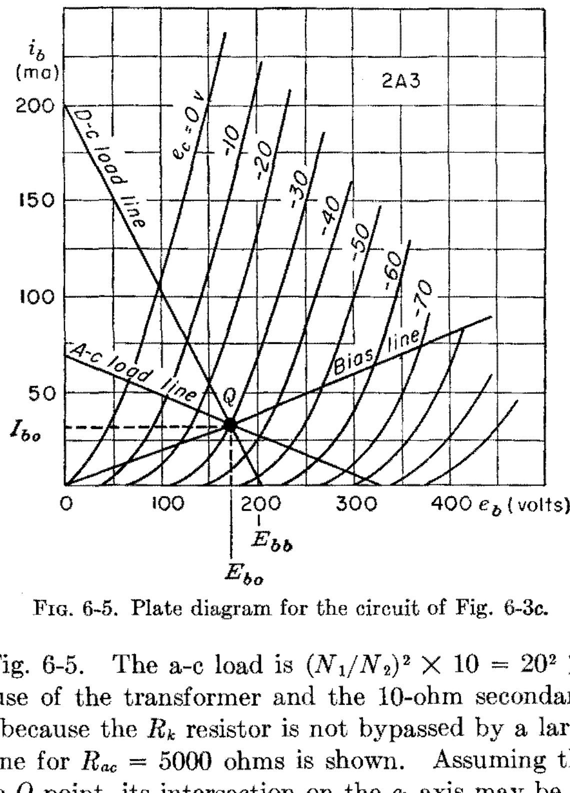


FIG. 6-5. Plate diagram for the circuit of Fig. 6-3c.

shown in Fig. 6-5. The a-c load is $(N_1/N_2)^2 \times 10 = 20^2 \times 10 = 4000$ ohms because of the transformer and the 10-ohm secondary load, plus 1000 ohms because the R_k resistor is not bypassed by a large capacitor. The load line for $R_{ac} = 5000$ ohms is shown. Assuming that it passes through the Q point, its intersection on the e_b axis may be computed as being $0.030 \times 5000 = 150$ volts greater than E_{bo}.

It should be noted that, whenever the a-c load exceeds the d-c load, the plate diagram will resemble Fig. 6-5; an a-c load smaller than the d-c load results in a diagram similar to that of Fig. 6-4. The grid-bias line is useful only when R_k is known and the bias is not. The d-c-load line is useful when the Q point is not known. It usually serves no useful purpose after Q has been determined as the dynamic operation does not follow it.

6-6. The Case of a Reactive Load. The load in the plate circuit of a vacuum tube need not be a pure resistance and often is not. As an example, consider the triode of Fig. 6-6 with an inductive load. With no signal applied, the plate current I_{bo} is steady, and the inductive reactance has no effect. The d-c load external to the tube itself is therefore 5000 ohms. A load line for this value is constructed and labeled "d-c-load line" in Fig. 6-7. The quiescent operating point is now located at Q.

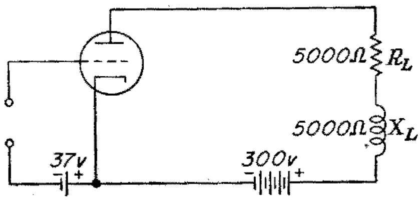


FIG. 6-6. A triode with an inductive load.

With a signal applied to the grid circuit the plate current varies, and Kirchhoff's equation for the plate circuit becomes

$$e_b = E_{bb} - e_L = E_{bb} - i_bR_L - L\frac{di_b}{dt} \tag{6-5}$$

This relation between i_b and e_b, constrained by the load, when plotted will give the path of operation of the tube. The equation has a term involving the time rate of change of i_b, and therefore this term is dependent upon

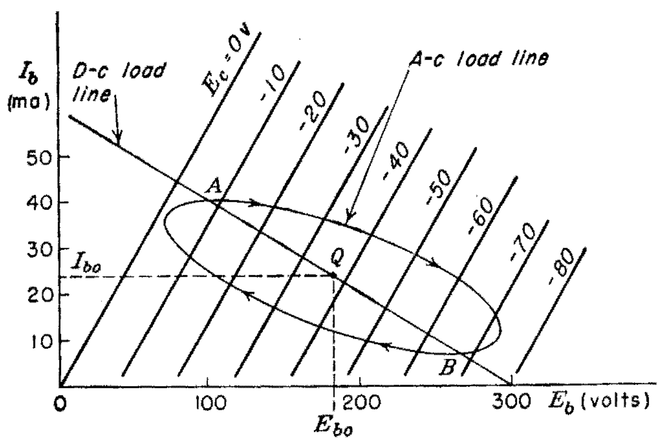


FIG. 6-7. Plate diagram for the circuit of Fig. 6-6.

the rate of variation of the grid voltage. Thus the load line, or path of operation, has a shape dependent upon the grid signal and the shape and spacing of the plate characteristics of the tube and cannot readily be plotted. If, for the sake of completing the discussion, one is willing to approximate by considering the tube characteristics as parallel, straight, and equidistant lines, so that no distortion appears in the plate current, then a sinusoidal grid drive results in a sinusoidal current,

$$i_p = I_{pm} \cos \omega t \tag{6-6}$$

where i_p is the variable part of i_b (see definition 4 under Art. 6-7), and

$$i_b = I_{bo} + I_{pm} \cos \omega t \tag{6-7}$$

Therefore, upon substitution in Eq. (6-5),

$$e_b = E_{bb} - I_{bo}R_L - R_L I_{pm} \cos \omega t + \omega L I_{pm} \sin \omega t$$

or

$$e_b = E_{bo} - I_{pm} \sqrt{R_L{}^2 + \omega^2 L^2} \cos (\omega t + \theta) \tag{6-8}$$

where

$$\theta = \arctan \frac{\omega L}{R_L} \tag{6-9}$$

Equations (6-7) and (6-8) constitute the parametric equations of an ellipse, ωt being the parameter. The locus may be drawn by assuming values of ωt and plotting the points obtained. For this example let us assume the drive to be sufficient to make $I_{pm} = 16$ ma and that $I_{bo} = 24$ ma and $E_{bo} = 180$ volts, for the tube and circuit values used. Then $\omega t = 0$ yields $e_b = E_{bo} - I_{pm}R_L$ and $i_b = I_{bo} + I_{pm}$, or $e_b = 180 - 0.016 \times 5000 = 100$ volts and $i_b = 24 + 16 = 40$ ma. In like manner, $\omega t = 90°$ yields $e_b = 260$ volts, $i_b = 24$ ma; $\omega t = 180°$ gives $e_b = 260$ volts, and $i_b = 8$ ma; $\omega t = 270°$ corresponds to $e_b = 100$ volts, $i_b = 24$ ma. The elliptical locus is drawn in Fig. 6-7. It can be shown that the points on the ellipse corresponding to maximum and minimum values of current (A and B) are the intersections with the ellipse of a load line corresponding to the resistive part of the load. This is noticeably true in this case and is also generally true.

There should now be a realization of the shape of the path of operation and of the direction of instantaneous-operating-point travel along the path for an inductive load, but little can be gained by using this method. Since no distortion was assumed, the circuit can be much more easily solved by application of the linear-equivalent-plate-circuit theorem to be developed later. If actual tube characteristics are used, the path is no longer elliptical nor is its shape readily determined. Hereafter, if the graphical method using load lines is attempted, the load must be at least approximately a pure resistance.

6-7. Variations of Plate Current and Plate Voltage Resulting from a Grid Signal.

Definitions:

e_g or Δe_c is the variable part of e_c.
e_p or Δe_b is the variable part of e_b.
e_z or Δe_L is the variable part of e_L.
i_p or Δi_b is the variable part of i_b.

The positive sense of each of these quantities is assumed the same as that of the total quantity of which it is a part. Thus e_g is positive when it increases the potential of the grid above that of the cathode.

When the grid is made more positive than the bias value, the plate current increases, resulting in greater voltage drop in the load. This causes the difference in potential between the plate and the cathode to decrease. The action is shown graphically in Fig. 6-8, using the plate diagram for a triode with a resistance load. In drawing this diagram it has been assumed that the i_b vs. e_b curves for different grid voltages are equally spaced in the region cut by the load line. If this is true, a sinusoidal grid voltage e_g will result in a sinusoidal waveshape for both i_p and e_p.

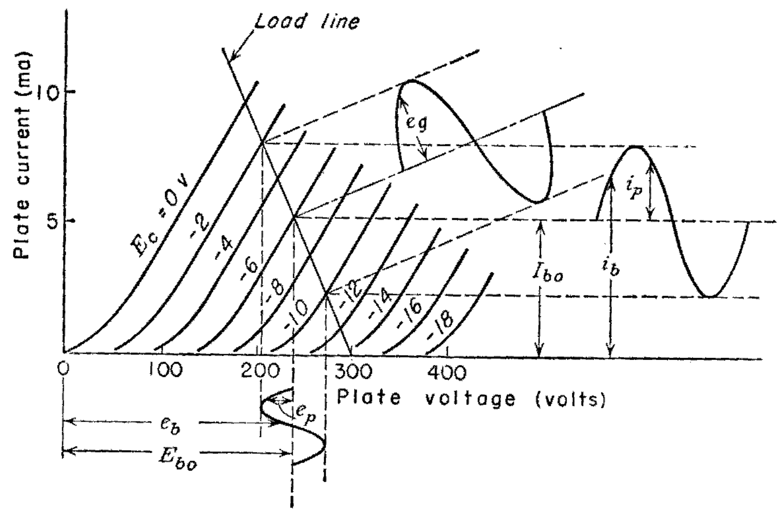


FIG. 6-8. Plate diagram for a triode amplifier with current and voltage waveforms.

It is often desirable to refer to sinusoidal voltages and currents by their effective (rms) values, which are expressed by the symbols V_g, E_g, E_p, I_p, and E_z, each corresponding to the same lower-case letter. In expressing mathematical relationships between sinusoidal quantities the notation of complex numbers is convenient. For example, the complex algebraic symbol for V_g is $\mathbf{V}_g$. This symbol carries its own magnitude and phase angle. Thus the polar expression for a sinusoidal v_g is $\mathbf{V}_g = V_g\underline{/\theta}$, and similarly for the other quantities.

Under the idealized conditions of Fig. 6-8, where no grid current flows and no distortion of waveform occurs, the following relations may be seen to hold:

$$e_c = E_{co} + e_g \tag{6-10}$$
$$e_b = E_{bo} + e_p \tag{6-11}$$
$$e_L = E_{Lo} + e_z \tag{6-12}$$
$$i_b = I_{bo} + i_p \tag{6-13}$$

6-8. The Dynamic Characteristic. Under most conditions of operation, application of a signal to the grid produces a plate current of distorted waveshape and a new set of symbols is needed.
Definitions:

> E_{ca} (grid bias) is the average value of e_c with signal applied.
> E_{ba} is the average value of e_b with signal applied.
> I_{ba} is the average value of i_b with signal applied.

Average values may be read on a d'Arsonval type of instrument.

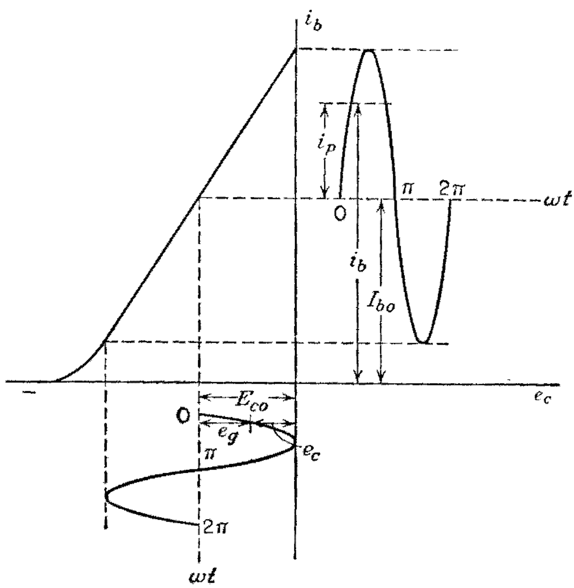


FIG. 6-9. Dynamic transfer characteristics of a triode amplifier with waveforms of current and voltage.

If corresponding values of i_b and e_c, occurring on the a-c-load line for a triode, are plotted on rectangular i_b and e_c axes, a so-called dynamic transfer characteristic results (see also Art. 4-7). For example, the dynamic transfer characteristic corresponding to the plate diagram of Fig. 6-8 is shown in Fig. 6-9. If the points of intersection of the characteristic curves with the load line in Fig. 6-8 are equally spaced, the dynamic characteristic is a straight line. Figure 6-9 also pictures clearly the steady and variable components of grid voltage and of plate current.

If the spacing of the points on the load line mentioned in the preceding paragraph is not uniform, the plate diagram has the appearance as in Fig. 6-10 and the dynamic transfer characteristic resembles the example shown in Fig. 6-11. If a sinusoidal grid signal is applied, the waveform of i_b is distorted and, because the load is resistive, that of e_b suffers a

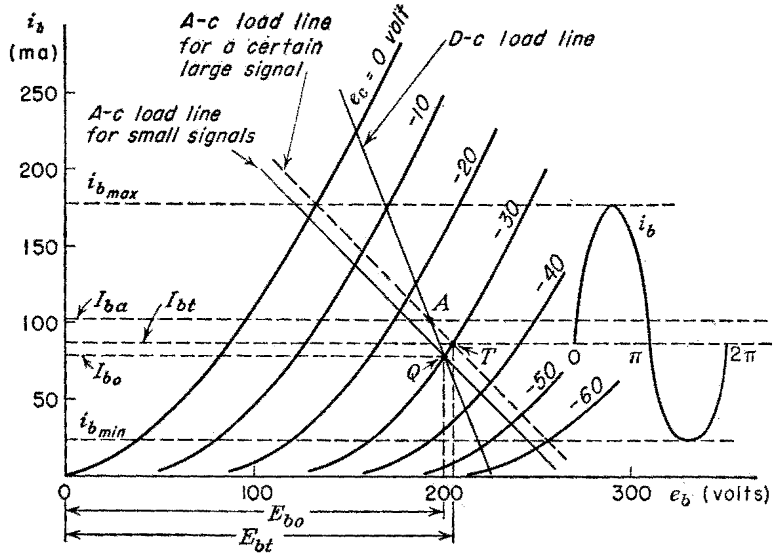


FIG. 6-10. Plate diagram of a triode with fixed grid bias and nonlinear operation (exaggerated).

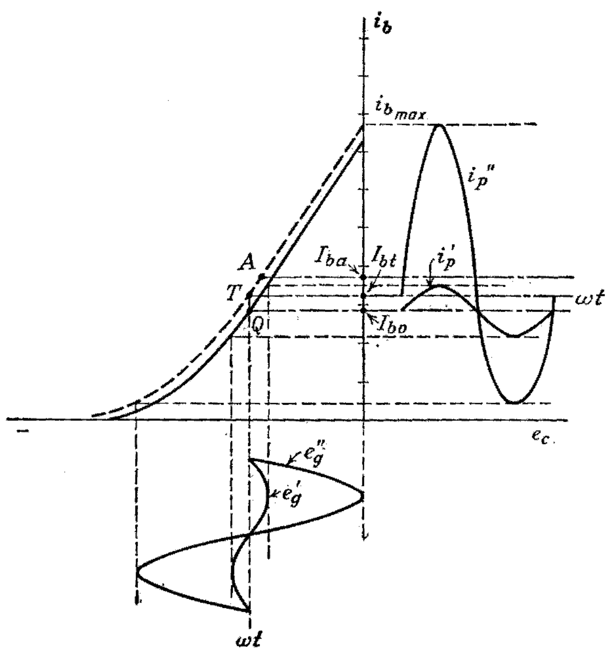


FIG. 6-11. A dynamic transfer characteristic for a triode with fixed grid bias, for linear and for nonlinear operation.

similar distortion. As a consequence, the positions of the load line and of the dynamic characteristic shift. To show this, we need to study both figures. As a preliminary step, let it be assumed that a small sinusoidal grid voltage e_g' is applied (shown in Fig. 6-11 only). A small alternating plate current i_p' flows, sinusoidal in form, so that the total plate current averages I_{bo}. The grid-voltage swings are still centered at Q.

Now allow the grid swing to increase in size (e_g''). The resulting plate-current waveform (i_p'') is drawn, together with the average value I_{ba}, which in this case exceeds I_{bo}. Since the d-c-load line is the locus of average values of e_b and i_b, the point A representing (E_{ba},I_{ba}) lies on the d-c-load line (see Fig. 6-10). Also, since the load is resistive, i_b and e_b have the same waveshapes (inverted), and whenever i_b passes through its average value I_{ba}, the plate voltage e_b passes through its average value E_{ba} at the same instant. Hence (E_{ba},I_{ba}) is a point on the dynamic-load, or a-c-load, line also. In other words, as the grid drive is increased, the a-c-load line moves upward (in this case) so that it always passes through A. The new line is drawn broken in Fig. 6-10, as is the corresponding dynamic transfer characteristic in Fig. 6-11. If the bias is fixed, the grid voltage still swings about $E_{co} = -30$ volts as an axis (called the time axis) and the variational values of i_b (or e_b) should be measured from the corresponding time-axis values obtaining at T, whenever it is desired to relate values of e_b, i_b, and e_c at the same instant.

Definitions:

E_{ct} is the time-axis value of e_c.
E_{bt} is the time-axis value of e_b.
I_{bt} is the time-axis value of i_b.

Strictly speaking then,

e_p is the variable part of e_b measured from E_{bt}.
i_p is the variable part of i_b measured from I_{bt}.

These values are related as follows:

$$e_b = E_{bt} + e_p \tag{6-14}$$
$$i_b = I_{bt} + i_p \tag{6-15}$$

T is the point determined by E_{ct}, E_{bt}, I_{bt}. It is on the actual a-c-load line and the dynamic transfer characteristic.

The dynamic operating point T does not usually coincide with the static operating point Q because of the nonsymmetrical waveforms of plate current and voltage produced by nonlinear distortion. This is sometimes referred to as the effect of plate-circuit rectification.

For triodes the nonlinear distortion is not very severe under single-tube class A operation, and the shift of operating point is usually small. It is therefore common procedure to assume T to be Q, which makes $i_b = I_{bo} + i_p$. Similar relations hold for e_p and e_z.

For other classes of triode operation and for power pentodes in any class, the distortion is often rather severe, and for most accurate results the procedure for determining T outlined in Art. 6-18 or some other equivalent one should be followed.

6-9. Circuit Theorems.[1,2] In addition to the laboratory method for determining circuit behavior, numerous theorems have been developed which are helpful in determining the performance of a circuit or in designing one. The student already has some familiarity with some of them; others are new. The topics listed here will be discussed in some detail in the articles which follow. These are (1) Ohm's law, (2) Thévenin's theorem and Norton's theorem, (3) the equivalent-linear-plate-circuit theorem, (4) the harmonic analysis of a periodic function, (5) graphical harmonic analysis of plate current, and (6) the series expansion for plate current.

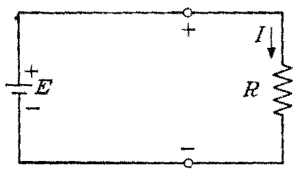


Fig. 6-12.

6-10. Ohm's Law. It is assumed that the student already has considerable familiarity with Ohm's law. However, a brief restatement of certain facts may be desirable.

For a d-c circuit, $E = IR$. E is the voltage drop across the resistor in the direction of conventional current flow (see Fig. 6-12). If the value of R is independent of the magnitudes of E and I, the resistor is said to be linear. (For linearity it is assumed that the effect is the same for either polarity of the applied voltage.)

If the voltage and current vary with time, we may write $e = iR$. Since inductance or capacitance in the circuit also causes a difference in potential, it is desirable to incorporate the effects of $L\,di/dt$ and $(1/C)\int i\,dt$ into Ohm's law. This can be done if we assume L and C also to be linear elements (the value is independent of e or i) and the waveform of e to be sinusoidal. Recalling that $\mathbf{E}$ and $\mathbf{I}$ are complex numbers, in algebraic, polar, trigonometric, or exponential form, which represent the effective values and the phases of the voltage and the current, respectively, we can write $\mathbf{E} = j\mathbf{I}X_L$ and $\mathbf{E} = -j\mathbf{I}X_C$ as satisfactory solutions to the problem of obtaining the more desirable forms. The student should remember that $X_L = 2\pi fL$ and $X_C = 1/2\pi fC$.

If an alternating voltage is applied to a complex network of R, L, and C linear elements, the ratio $\mathbf{E}/\mathbf{I}$ can be expressed as a complex number which is denoted by $\mathbf{Z}$. The value of $\mathbf{Z}$ can be computed by using the R, jX_L, and $-jX_C$ representations for the components and combining them in the same manner as for a similar network composed entirely of resistors. The angle of $\mathbf{Z}$ will be found never to exceed 90° in magnitude. Thus it can be stated that the instantaneous potential drop across a linear network containing no generators is in the direction of the instantaneous current flow at least half the time. One result is that although the instantaneous value of the power, fed from the source of applied voltage to a linear network containing some resistance, may be sometimes positive and sometimes negative, yet its average value is positive.

Although the chief applications for Ohm's law are in connection with networks containing linear elements R, L, and C only, it offers possibilities for investigating the character of any network between two terminals (see Fig. 6-13). A voltage **E** is assumed to be applied to the box terminals with arbitrary positive sense. Positive sense for **I** is taken to agree with that for **E**, as though the box were a load. By writing Kirchhoff's circuit equations around various loops of the network and using other available theorems, suppose we obtain **Z** from the relation $\mathbf{Z} = \mathbf{E}/\mathbf{I}$. For example, suppose $\mathbf{Z} = 3 + j4$ ohms. Then the network is the equivalent of 3 ohms resistance in series with 4 ohms inductive reactance at the frequency of the voltage used. However, suppose **Z** is determined as $-3 + j4 = 5\underline{/127^\circ}$ ohms. Thus, in its original a-c meaning, Ohm's law has failed since $\theta > 90°$. However, a new interpretation may be gained by reference to Fig. 6-14 if one solves for the power delivered to the box. $P = EI \cos \theta = EI \cos 127° = -0.6EI$. In words, power has been delivered from the box rather than to it. Thus the box must contain the equivalent of a generator.

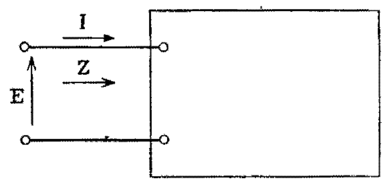


FIG. 6-13.

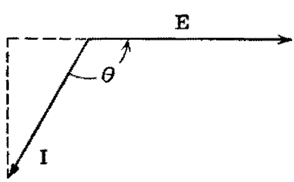


FIG. 6-14.

With **E** and **I** properly sensed, if Ohm's law yields **Z** with an angle numerically less than 90° or with R positive, the network absorbs average power. If, on the other hand, the angle of **Z** numerically exceeds 90°, or if R is negative, the network contains the equivalent of a generator.

6-11. Thévenin's Theorem and Norton's Theorem. Thévenin's theorem is useful for converting any linear network with two terminals into an equivalent series circuit. Thévenin's theorem may be used as follows:

1. Referring to Fig. 6-15*a*, remove the load, either actually in the laboratory or on paper.
2. Determine the open-circuit voltage $\mathbf{E}_{oc}$ between the terminals.
3. Remove all generators from the network, replacing them by their internal impedances. Measure the complex impedance between the network terminals. Call this $\mathbf{Z}_{oc}$.

4. Draw Thévenin's equivalent series circuit as in Fig. 6-15*b*. The generator voltage is $\mathbf{E}_{oc}$, and the impedance in series with it is $\mathbf{Z}_{oc}$.

5. Reconnect the load impedance.

The current through the load and the power delivered to the load are the same as though the original circuit were used.

The proof of Thévenin's theorem is easily obtained from Fig. 6-16, which is the same as Fig. 6-15*a* except that an additional series generator with voltage $\mathbf{E}_{oc}$ has been inserted with sense as shown. The impedance elements in the circuit, being linear, are unaffected by the magnitude of the currents, and hence each generator produces a current independent of the action of any other generator. (This independence of action is known as the *superposition theorem.*)

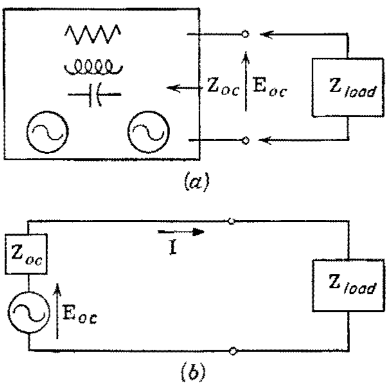


FIG. 6-15. (*a*) Linear bilateral network. (*b*) Thévenin's equivalent circuit.

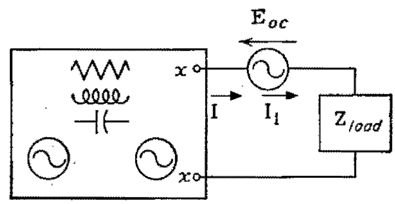


FIG. 6-16. Diagram for proving Thévenin's theorem.

The current $\mathbf{I}_1$, produced by the newly introduced generator alone, is

$$\mathbf{I}_1 = -\frac{\mathbf{E}_{oc}}{\mathbf{Z}_{oc} + \mathbf{Z}_{\text{load}}} \tag{6-16}$$

If we tentatively assume that the total current $\mathbf{I}_2$ equals zero, then the voltage between network terminals *xx* is E_{oc}, with the result that the total voltage drop around the loop through $\mathbf{Z}_{\text{load}}$ is zero. This justifies the assumption that $\mathbf{I}_2 = 0$, since there can be only one value of total current. The current $\mathbf{I}$, produced by the generator of the original network, is $\mathbf{I}_2 - \mathbf{I}_1$, and hence

$$\mathbf{I} = +\frac{\mathbf{E}_{oc}}{\mathbf{Z}_{oc} + \mathbf{Z}_{\text{load}}} \tag{6-17}$$

This is the same as the current $\mathbf{I}$ determined by the circuit of Fig. 6-15*b*. Hence we may use Fig. 6-15*b* to replace Fig. 6-15*a*.

Norton's theorem is easily obtained from that of Thévenin. In Fig. 6-15*b* consider a short circuit applied across the load impedance.

$$\mathbf{E}_{oc} = \mathbf{I}_{sc}\mathbf{Z}_{oc} \tag{6-18}$$

With the short circuit removed

$$\mathbf{I} = \frac{\mathbf{E}_{oc}}{\mathbf{Z}_{oc} + \mathbf{Z}_{\text{load}}} \tag{6-19}$$

Substituting to eliminate $\mathbf{E}_{oc}$ yields

$$\mathbf{I} = \frac{\mathbf{Z}_{oc}}{\mathbf{Z}_{oc} + \mathbf{Z}_{\text{load}}} \mathbf{I}_{sc} \tag{6-20}$$

Equation (6-20) is indistinguishable from that obtained from Fig. 6-17, which features a constant-current generator delivering a current $\mathbf{I}_{sc}$ to $\mathbf{Z}_{oc}$ in parallel with $\mathbf{Z}_{\text{load}}$. If $\mathbf{Z}_{\text{load}}$ is varied, $\mathbf{I}_{sc}$ remains constant but $\mathbf{I}$ varies.

The concept of a generator which delivers the same current regardless of the magnitude of the load may worry us a little. Let us use Ohm's law to determine its emf and internal impedance Z_{int}. Making Z_{load} in Fig. 6-17 equal to zero yields $I_{sc} = \text{emf}/Z_{\text{int}}$. With Z_{load} infinite, $I_{sc} = \text{emf}/(Z_{\text{int}} + Z_{oc})$. Hence $\text{emf}/Z_{\text{int}} = \text{emf}/(Z_{\text{int}} + Z_{oc})$. Since $Z_{oc} \neq 0$, the only manner of satisfying these relations is to claim that $Z_{\text{int}} = \infty$ and $\text{emf} = \infty$. Therefore in drawing Norton's equivalent circuit, it is well to omit the drawing of the generator and to show only the arrow indicating the positive sense of $\mathbf{I}_{sc}$.

Norton's equivalent circuit for any linear network may be obtained as follows:

1. Short the load, and determine the magnitude and phase of the short-circuit current. (*Caution:* This is safe procedure on paper but may be dangerous in the laboratory with low-impedance circuits.) Label a current arrow with this magnitude.

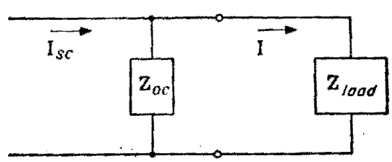


Fig. 6-17. Norton's equivalent circuit.

2. Remove the short and the load, and determine the open-circuit impedance in the same manner as for Thévenin's theorem. Use this impedance as a shunt as shown in Fig. 6-17.

3. Reconnect the load.

Both Thévenin's and Norton's theorems give the same results when their application is legitimate. Many vacuum-tube circuits do not behave in a linear manner, and discretion must be exercised.

6-12. The Equivalent-plate-circuit Theorem. In Chap. 4 it was shown how μ, g_m, and r_p for a triode may be determined from the i_b vs. e_b family of curves. If these curves in some region are approximately straight and parallel, then r_p is almost constant. If they are also about equally spaced, then g_m and μ are almost constant. These conditions can usually be assumed to exist over at least a small region, and if the tube is constrained to operate in this region, the operation is said to be linear. In other words, if $i_b = f(e_b,e_c)$ were sketched in three dimensions, an approximate plane surface in the operating region would result and the operation of the tube follows this plane relationship.

In many cases of vacuum-tube-circuit analysis the variations in plate current and voltage caused by a grid signal are of greater importance than the steady or quiescent values. If the operation of the circuit is linear, these variational quantities can be readily computed by the use of the equivalent-plate-circuit theorem. This theorem states the effect in the plate circuit of a signal voltage applied in the grid circuit. We shall now develop this theorem.

Allow e_b and e_c to vary independently, and from the total differential of the calculus obtain

$$\Delta i_b \approx di_b = \frac{\partial i_b}{\partial e_b}\Delta e_b + \frac{\partial i_b}{\partial e_c}\Delta e_c \tag{6-21}$$

or in the special symbols of this book (and dropping the approximation symbol)

$$i_p = \frac{1}{r_p}e_p + g_m e_g \tag{6-22}$$

or

$$i_p r_p = e_p + \mu e_g \tag{6-23}$$

Recalling that

$$e_b = E_{bb} - e_L \tag{6-2}$$

and

$$e_b = E_{bo} + e_p \tag{6-11}$$

$$e_L = E_{Lo} + e_z \tag{6-12}$$

we obtain

$$E_{bo} + e_p = E_{bb} - E_{Lo} - e_z \tag{6-24}$$

Since

$$E_{bo} = E_{bb} - E_{Lo} \tag{6-4}$$

it follows that

$$e_p = -e_z \tag{6-25}$$

which is more or less obvious. Hence Eq. (6-23) may be changed to the form

$$\mu e_g - e_z - i_p r_p = 0 \tag{6-26}$$

Equation (6-26) represents the action in a vacuum-tube plate circuit, such as that shown in Fig. 6-18, because of the action of grid-circuit signal v_g. We note the following points:

1. Only the part of v_g which reaches the grid, namely, e_g, is effective. e_g is μ times as effective, per volt, in the plate circuit as is e_z.

2. Only varying components of current and voltage occur in the equation. Therefore the equation is useless for determining direct currents and voltages.

3. Any kind of passive load may be used, and e_z is the voltage drop across it.

4. The ordinary nonelectronic circuit of Fig. 6-19 has the same loop equation.

It follows that the circuit of Fig. 6-19 is useful as an approximate substitute for the actual tube circuit, and in so far as μ and r_p remain constant, it is a good substitute for computing variational quantities.

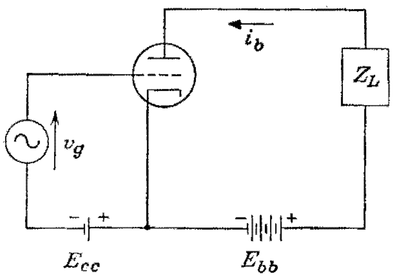


FIG. 6-18. A triode circuit.

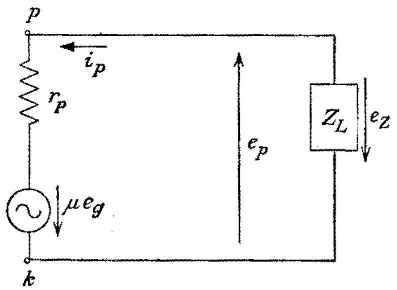


FIG. 6-19. Linear equivalent circuit for the network of Fig. 6-18.

If the load consists of linear elements equivalent to R, L, and C in series, then

$$e_z = i_pR + L\frac{di_p}{dt} + \frac{1}{C}\int i_p\,dt \tag{6-27}$$

If this be substituted into Eq. (6-26), then the value of i_p can be determined if the other quantities are known. If e_g is sinusoidal, a great saving in labor results by the use of effective values of currents and voltages as shown in Fig. 6-20. Since

$$\mathbf{E}_z = \mathbf{I}_p\mathbf{Z}_L \tag{6-28}$$

Eq. (6-26) changes to the form

$$\mu\mathbf{E}_g - \mathbf{I}_p\mathbf{Z}_L - \mathbf{I}_p r_p = 0 \tag{6-29}$$

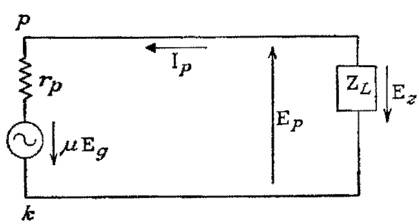


FIG. 6-20. Thévenin's equivalent circuit for sinusoidal voltages.

The complex-number symbols indicate that each carries its own magnitude and phase angle. Since $e_p = -e_z$ from Eq. (6-25), the symbol for e_p may be shown in Fig. 6-19 and its effective-value counterpart in Fig. 6-20.

For many purposes it is convenient to think of certain points of these circuits as corresponding to actual elements in Fig. 6-18. Thus symbols k and p may be added to represent cathode and plate, respectively. The grid circuit may be added, if desired, as shown in Fig. 6-21. This series circuit is often called Thévenin's or the constant-voltage equivalent plate circuit.

If the load in Fig. 6-21 is shorted,

$$\mathbf{I}_{sc} = \frac{\mu \mathbf{E}_g}{r_p} = g_m \mathbf{E}_g \tag{6-30}$$

Since the open-circuit impedance is r_p, it follows that Norton's equivalent circuit has the form shown in Fig. 6-22. In some applications it is more

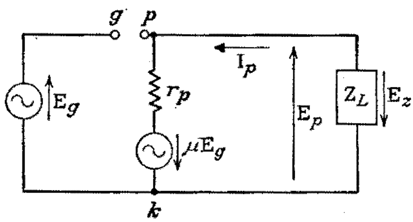


FIG. 6-21. Equivalent plate and grid circuits.

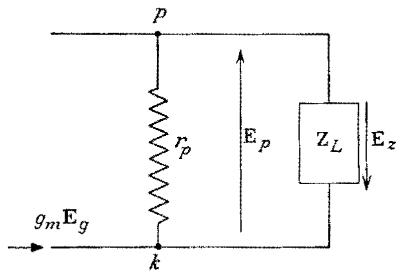


FIG. 6-22. Norton's equivalent circuit.

convenient than the constant-voltage-generator or Thévenin's circuit. Both circuits give the same answers in solving a problem.

An amended form of the plate-circuit theorem, in which signals are applied in both the grid and the plate circuits, is given in Prob. 9 at the end of this chapter.

6-13. The Equivalent Plate Circuit for the Pentode and Other Multigrid Tubes. Figure 6-23 shows a pentode connected in a simple circuit

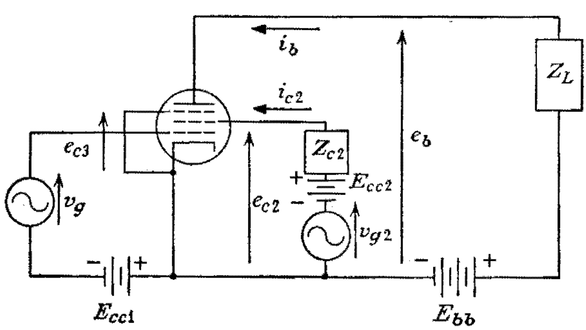


FIG. 6-23. A pentode circuit.

with a plate load. Most of the symbols adopted for triodes are useful here, with minor changes. Voltages applied to grid 1 may take a subscript 1 if necessary to avoid confusion. Those applied to other grids always take subscripts according to their numbers. As examples:

E_{cc2} is the direct voltage applied between cathode and grid 2.

e_{c3} is the total instantaneous voltage rise from cathode to grid 3 (in this case e_{c3} equals zero).

v_{g2} is the instantaneous voltage introduced into the grid 2 circuit from some other circuit.

Other symbols which are formed in the same fashion should be readily recognized.

For many applications of the pentode, v_{g2} is zero, E_{cc2} is positive, and e_{c3} is zero. The impedance Z_{c2} can be made practically zero. A signal v_{g1} causes variation of currents i_{c2} and i_b, but there being no impedance in the second-grid circuit, the value of e_{c2} does not vary. Thus, although the bias voltages E_{c1}, E_{c2}, E_{c3}, and E_b determine the operating point, the only elements with varying voltages relative to the cathode are the No. 1 grid and the plate, the same as for a triode. Under these circumstances and when the operation is linear, the equivalent plate circuit for a pentode is identical with that of the triode. If a signal is applied to another grid also, say the screen (and again only if the operation is linear), it is necessary only to insert in the series equivalent circuit another generator with voltage label $\mu_{g2p}e_{g2}$, where $\mu_{g2p} = -\partial e_b/\partial e_{c2}$.

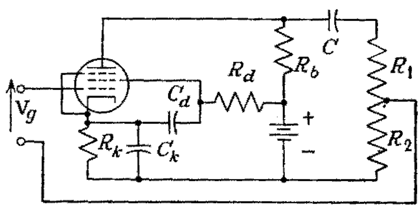


FIG. 6-24. A pentode circuit.

6-14. An Example of the Construction of an Equivalent Plate Circuit. A somewhat complex pentode circuit is shown in Fig. 6-24. Assume in this circuit that C, C_d, and C_k are all very large, so that at the signal frequency their reactances are negligibly small. The quiescent operating point is determined by the characteristics of the tube, the voltage E_{bb}, and the values of the resistors R_k, R_d, and R_b. Assume that the values of μ and r_p have been determined at this operating point, which lies in a linear region.

1. The part of the equivalent circuit which represents the tube between p and k can be made by drawing a resistor for r_p and a generator for each alternating voltage applied to a grid. Each generator is labeled with its emf $\mu\mathbf{E}_g$, with the positive-sense sign toward k. The sense signs are chosen in this way so that $\mathbf{I}_p$ may be assumed with positive sense in the same direction as was i_b in the actual tube circuit. The μ and $\mathbf{E}_g$ subscripts are determined by the grid being used.

2. If a signal voltage $\mathbf{V}_p$ has been introduced into the plate circuit (none has in this example), insert in the series plate circuit a generator labeled and representing $\mathbf{V}_p$, with positive sense in the same direction in the circuit as that of $\mu\mathbf{E}_g$.

3. Omit all direct supply voltages and currents, and delete resistors used to regulate the d-c values if they are not part of an a-c path.

4. Complete the series plate circuit by inserting the load impedance.

5. It is generally convenient to choose a point g to which the components comprising the grid circuit can be connected. This includes $\mathbf{V}_g$, which should have positive sense toward the grid g. Likewise draw connections for other grids if signals are applied to them.

6. Draw arrows, as convenient, to represent positive sense of currents. These directions are arbitrary except in the case of $\mathbf{I}_p$, which has been standardized in preceding articles.

7. Rearrange the circuit in any convenient fashion.

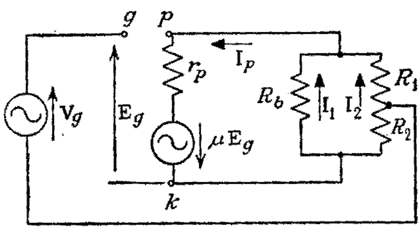


FIG. 6-25. Thévenin's equivalent for the circuit of Fig. 6-24.

8. $\mathbf{E}_g$ is the voltage rise from k to g. By any convenient path determine $\mathbf{E}_g$ in terms of other voltages, currents, and impedances.

9. Write circuit-loop equations and node equations using Kirchhoff's laws. Sufficient equations should be obtained to solve for any quantities desired.

Following these steps for the circuit of Fig. 6-24 results in the equivalent circuit shown in Fig. 6-25 and these equations:

From step 8:

$$\mathbf{E}_g = -\mathbf{I}_2R_2 + \mathbf{V}_g$$

From step 9:

$$\begin{aligned} \mu\mathbf{E}_g - \mathbf{I}_1R_b - \mathbf{I}_pr_p &= 0 \\ \mathbf{I}_1R_b - \mathbf{I}_2(R_1 + R_2) &= 0 \\ \mathbf{I}_1 + \mathbf{I}_2 &= \mathbf{I}_p \end{aligned}$$

If $\mathbf{V}_g$ is known, then there are four equations in four unknowns $\mathbf{I}_1$, $\mathbf{I}_2$, $\mathbf{I}_p$, and $\mathbf{E}_g$. Any of these may now be determined.

6-15. Limitations on the Linear-equivalent-circuit Method of Analysis. The foregoing method of setting up an equivalent circuit for studying the performance of a vacuum-tube circuit is certainly one way of solving a problem which at first appeared to be quite involved, *viz.*, the performance of a nonlinear device like a vacuum tube combined with linear-impedance elements. But of course the truth is that the vacuum tube has been treated as though it were linear, which it often is not. However, if the tube's grid voltage varies only slightly, say a volt or so from the quiescent value, the equivalent-circuit device works very well. For somewhat greater grid drive the device still is valuable in making estimates. For large variations in grid potential a better method of analysis is needed.

It will be found later that voltage amplifiers usually have low enough grid drive so that the equivalent-circuit method gives results which are quite near those obtained in the laboratory, provided all effects are taken into account and the values of μ and r_p are known. This latter requirement is somewhat of a catch, for, except for a few operating points, the values of these tube parameters are rarely found in a tube manual. Of course they may be measured on a vacuum-tube bridge if one is available, determined by some other laboratory device, or calculated from a set of characteristic curves. For power amplifiers the grid drive is usually too great for the linear-equivalent-circuit method to be used for any purpose other than an estimate of results.

6-16. Fourier Analysis of a Periodic Function.[3] Voltages, currents, and other quantities for which the waveform repeats can be separated into sinusoidal components by the Fourier method. For example, if $f(t)$ represents such a periodic quantity, we can write

$$f(t) = B_0 + B_1 \cos \omega t + B_2 \cos 2\omega t + \cdots + B_n \cos n\omega t + \cdots \\ + A_1 \sin \omega t + A_2 \sin 2\omega t + \cdots + A_n \sin n\omega t + \cdots \qquad (6\text{-}31)$$

This is an infinite series. If the B and A coefficients progressively decrease in value to a sufficient degree, the series may converge and represent $f(t)$. This occurs in practical cases of voltages and currents. If the convergence is rapid enough, the first few cosine and sine terms may afford a satisfactory approximation to the function $f(t)$.

If the function is known, the values of the B's and A's can be determined. For example, we can find B_0 by multiplying both members of Eq. (6-31) by $d\omega t$ and integrating from 0 to 2π. Since the algebraic area under a sine or cosine wave over any number of complete cycles is zero, we obtain

$$B_0 = \frac{1}{2\pi} \int_0^{2\pi} f(t)\, d\omega t \qquad (6\text{-}32)$$

In other words, B_0 is the average value of $f(t)$ over one cycle.

Any coefficient B_n can be obtained by multiplying both members of Eq. (6-31) by $\cos n\omega t\, d\omega t$ and integrating between the limits 0 and 2π. All integrals on the right equal zero except the one involving B_n, and thus we determine the value

$$B_n = \frac{2}{2\pi} \int_0^{2\pi} f(t) \cos n\omega t\, d\omega t \qquad (6\text{-}33)$$

By multiplying both members of Eq. (6-31) by $\sin n\omega t\, d\omega t$ and proceeding as above, we obtain

$$A_n = \frac{2}{2\pi} \int_0^{2\pi} f(t) \sin n\omega t\, d\omega t \qquad (6\text{-}34)$$

If the waveform of $f(t)$ has symmetry about the $t = 0$ axis, all A_n values will be zero. Hence in choosing the position of the origin, it often pays to try to arrange this symmetry. A waveform with symmetry about the origin itself has all B_n values equal to zero.

6-17. Graphical Harmonic Analysis of Plate Current.[4] Figures 6-10 and 6-11 show that when a sinusoidal waveform of voltage is applied to the grid, the waveform of the plate current or of the plate voltage may be nonsinusoidal. It is often desirable to make a graphical analysis of a nonsinusoidal waveform to determine its components. The analysis depends upon being able to pick numerical values from a drawing of the waveform or from a path of operation (load line) on the tube characteristics. Since the path of operation is easily drawn only in the case of resistive loads, it will be assumed in these analyses that the loads are of this character. Figure 6-26 shows a current waveform similar to that of Fig. 6-11, along with the graph of grid voltage $e_g = E_{gm} \cos \omega t$, which will serve as a timing wave. An arbitrary scale of current has been adopted. The position of the i_b axis is arbitrary, and in this case it is chosen to give symmetry about this axis. A Fourier analysis can then be made in terms of cosines only and will take the form of

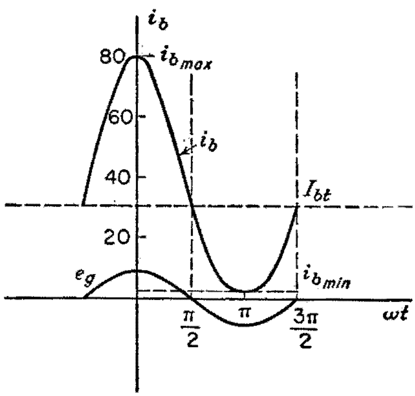


FIG. 6-26. Waveform of plate current resulting in nonlinear distortion in a triode amplifier.

$$i_b = B_0 + B_1 \cos \omega t + B_2 \cos 2\omega t + B_3 \cos 3\omega t + \cdots \tag{6-35}$$

where the B's are undetermined constants with current labels. Inspection of the waveform of Fig. 6-11 shows B_0, which is the average-current value, to be somewhat more than the T value. Since the waveform of i_b more or less resembles that of e_g, the value of B_1 relative to B_2, B_3, etc., must be considerable. But because of the fact that the waveform is higher than its fundamental value on both the upper and lower lobes, the second-harmonic term is important. In general it will be found that the third-harmonic term is small for triode amplifiers operated class A_1. As a convenient approximation therefore, this analysis will be made by assuming a simplified form:

$$i_b = I_{ba} + I_{p_1m} \cos \omega t + I_{p_2m} \cos 2\omega t \tag{6-36}$$

This identity contains three undetermined current values, which can be evaluated by properly choosing three points on the path of operation or on the i_b waveform, the former being in general more convenient of application. For greatest accuracy these points should be taken at the peak values of the component waves, and it seems fairly obvious that two of these occur at $\omega t = 0$ and π. As it is also of some advantage to choose points at equal grid-voltage intervals, $\omega t = \pi/2$ is taken as the third value. The following pairs of values are thus obtained: $\omega t = 0$, $i_b = i_{b_{\max}}$; $\omega t = \pi/2$, $i_b = I_{bt}$; $\omega t = \pi$, $i_b = i_{b_{\min}}$. Substituting these in order into the identity (6-36) yields

$$i_{b_{\max}} = I_{ba} + I_{p_1m} + I_{p_2m} \tag{6-37}$$
$$I_{bt} = I_{ba} - I_{p_2m} \tag{6-38}$$
$$i_{b_{\min}} = I_{ba} - I_{p_1m} + I_{p_2m} \tag{6-39}$$

These equations when solved give

$$I_{ba} = \frac{i_{b\max} + i_{b\min} + 2I_{bt}}{4} \tag{6-40}$$

$$I_{p_1m} = \frac{i_{b\max} - i_{b\min}}{2} \tag{6-41}$$

$$I_{p_2m} = \frac{i_{b\max} + i_{b\min} - 2I_{bt}}{4} = \frac{(i_{b\max} - I_{bt}) - (I_{bt} - i_{b\min})}{4} \tag{6-42}$$

The second form of Eq. (6-42) provides an easy graphical test for second-harmonic distortion. $i_{b_{\max}} - I_{bt}$ and $I_{bt} - i_{b_{\min}}$ are vertically measured distances along an i_b axis, but their ratio is the same as for corresponding distances measured along the load line. Thus, considering the two parts of the load line (Fig. 6-10) above and below the T point and extending to the limits of operation, it is the difference in their lengths which measures distortion. If they are equal, the second-harmonic distortion is zero.

Subtracting I_{bt} from both members of Eq. (6-40) gives

$$I_{ba} - I_{bt} = \frac{i_{b\max} + i_{b\min} - 2I_{bt}}{4} = I_{p_2m} \tag{6-43}$$

Since $I_{ba} - I_{bt}$ is the d-c or zero-frequency component of the varying part of the plate current i_p, it is often denoted as I_{p_0}. We may then state a convenient relation

$$I_{p_0} = I_{p_2m} \tag{6-44}$$

Equation (6-44) offers a laboratory method for approximately determining the second-harmonic component of the plate current for a triode. Since for a triode with fixed bias $I_{bt} = I_{bo}$ approximately, the difference between the no-signal value and the signal value of direct plate current is I_{p_0}. This is the peak value of the second-harmonic component of the plate current.

Although the assumption of Eq. (6-36) led to formulas for current, it is possible to begin with a similar assumption for plate voltage and obtain thereby formulas similar to (6-40) to (6-42) and (6-44) with current values changed to the corresponding voltage ones.

As an example of the use of these formulas, values may be obtained from the drawing of Fig. 6-26. Here $i_{b_{max}} = 80$ ma, $I_{bt} = 30$ ma, $i_{b_{min}} = 2$ ma. Then $I_{p_1m} = (80 - 2)/2 = 39$ ma, the peak value of i_{p_1}, and $I_{p_0} = I_{p_2m} = (80 + 2 - 2 \times 30)/4 = 5.5$ ma, the peak value of i_{p_2} and also the average value of i_p. The percentage of second-harmonic distortion is $100I_{p_2m}/I_{p_1m} = 100 \times 5.5/39 = 14.1$ per cent. In Fig. 6-27 the components of i_b have been sketched, and their sum yields approximately the same waveform as that of Fig. 6-26. If more harmonics had been calculated, the resemblance would have been closer, but the simple formulas of (6-40) to (6-42) are easily applied and are adequate for most cases of single-sided triode amplifiers operated class A_1.

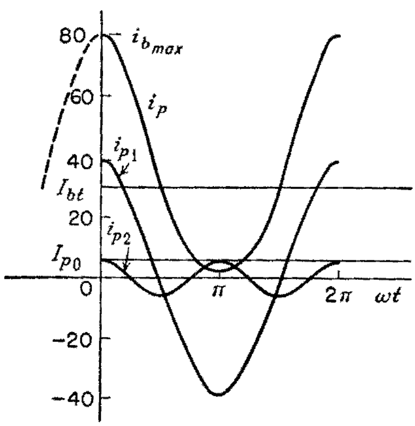


Fig. 6-27. Composition of plate current from its components.

When triodes are used in amplifiers in which, for the sake of higher efficiency, the operation is of a class other than A_1, or is push-pull, or in other circuits where nonlinear operation is deliberately chosen in order to cause some useful effect, the formulas developed above are generally inadequate. This is also true for pentode and beam-tube power amplifiers, even with class A_1 operation. Formulas have been developed in a similar manner to that above and using a considerable number of points. Properly chosen, these give an adequate determination of the components of any cyclic waveform occurring in a vacuum-tube circuit.

The general development of these formulas is outside the scope of this book, but as there is need for a set of them for use with certain applications, one additional group will be developed. As the distortion in the cases of pentode and beam-tube class A_1 amplifiers is small beyond the fourth harmonic, a useful set of formulas can be obtained by assuming

$$i_b = I_{ba} + I_{p_1m} \cos \omega t + I_{p_2m} \cos 2\omega t + I_{p_3m} \cos 3\omega t + I_{p_4m} \cos 4\omega t \tag{6-45}$$

and a timing wave of grid voltage, $e_g = E_{gm} \cos \omega t$, as before. Since the points where the peaks of the harmonics occur are not generally known, it is simpler to choose points at equal grid-voltage separation, as these can often be made to lie on characteristic curves. For the required five points let us use

$$\omega t = 0,\ i_b = i_{b_{\max}},\ \text{corresponding to } e_g = +E_{gm} \text{ or } e_c = E_{ct} + E_{gm}$$

$$\omega t = \frac{\pi}{3},\ i_b = i_{\frac{1}{2}},\ \text{corresponding to}$$

$$e_g = +\frac{1}{2}E_{gm} \text{ or } e_c = E_{ct} + \frac{1}{2}E_{gm}$$

$$\omega t = \frac{\pi}{2},\ i_b = I_{bt},\ \text{corresponding to } e_g = 0 \text{ or } e_c = E_{ct} \tag{6-46}$$

$$\omega t = \frac{2\pi}{3},\ i_b = i_{-\frac{1}{2}},\ \text{corresponding to } e_g = -\frac{1}{2}E_{gm} \text{ or}$$

$$e_c = E_{ct} - \frac{1}{2}E_{gm}$$

$$\omega t = \pi,\ i_b = i_{b_{\min}},\ \text{corresponding to } e_g = -E_{gm} \text{ or } e_c = E_{ct} - E_{gm}$$

Substituting into Eq. (6-45) gives, respectively,

$$i_{b_{\max}} = I_{ba} + I_{p_1m} + I_{p_2m} + I_{p_3m} + I_{p_4m} \tag{6-47}$$

$$i_{\frac{1}{2}} = I_{ba} + 0.5I_{p_1m} - 0.5I_{p_2m} - I_{p_3m} - 0.5I_{p_4m} \tag{6-48}$$

$$I_{bt} = I_{ba} - I_{p_2m} + I_{p_4m} \tag{6-49}$$

$$i_{-\frac{1}{2}} = I_{ba} - 0.5I_{p_1m} - 0.5I_{p_2m} + I_{p_3m} - 0.5I_{p_4m} \tag{6-50}$$

$$i_{b_{\min}} = I_{ba} - I_{p_1m} + I_{p_2m} - I_{p_3m} + I_{p_4m} \tag{6-51}$$

These equations when solved yield

$$I_{ba} = \frac{i_{b_{\max}} + 2i_{\frac{1}{2}} + 2i_{-\frac{1}{2}} + i_{b_{\min}}}{6} \tag{6-52}$$

$$I_{p_1m} = \frac{i_{b_{\max}} + i_{\frac{1}{2}} - i_{-\frac{1}{2}} - i_{b_{\min}}}{3} \tag{6-53}$$

$$I_{p_2m} = \frac{i_{b_{\max}} + i_{b_{\min}} - 2I_{bt}}{4} \tag{6-54}$$

$$I_{p_3m} = \frac{i_{b_{\max}} - 2i_{\frac{1}{2}} + 2i_{-\frac{1}{2}} - i_{b_{\min}}}{6} \tag{6-55}$$

$$I_{p_4m} = \frac{i_{b_{\max}} - 4i_{\frac{1}{2}} + 6I_{bt} - 4i_{-\frac{1}{2}} + i_{b_{\min}}}{12} \tag{6-56}$$

The meaning of the symbols in these equations is given in (6-45) and (6-46). As in the case of the simpler formulas (6-40) to (6-42), each of these formulas above has a voltage paraphrase. The method of harmonic analysis by using a finite number of terms of a Fourier's series is limited in its application to specific problems, since it is essentially numerical in its nature. For investigation of the general behavior of a circuit it is quite useless.

6-18. Determination of the T Point.[5] While rules have been devised for graphically determining the T point, it is sufficient to offer the method of repeated approximations. This method is outlined by the following procedure:

1. Use the plate characteristics for the tube. After drawing the d-c-load line and the bias line, if needed, the Q point is determined. Using this as the first approximate T point, draw the a-c-load line. Using the known grid drive, calculate the average plate current I_{ba} [see formula (6-40) or (6-52)]. Plot this approximate A point on the d-c load line.

2. If the tube operates with fixed bias, the new T point is on the $e_c = E_{co}$ curve. If, however, cathode bias is used, draw a horizontal line through A to intersect the grid-bias line, thus determining the new bias value E_{ct}. Sketch in the curve of i_b vs. e_b for $e_c = E_{ct}$. The new T point is on this curve.

3. Through A draw a line parallel to the original a-c-load line. This is the new a-c-load line, and its intersection with the characteristic curve for $e_c = E_{co}$ (or E_{ct} for cathode bias) is the second approximation to the T point.

4. The value obtained for I_{ba} by the procedure of (1) is probably incorrect, as it was assumed that Q determined the time axis for this calculation. Therefore, using the new T point obtained in step 3, one should recalculate I_{ba} and determine improved values for A', E_{ct}', and T'. This procedure may be repeated to obtain A'', E_{ct}'', and T'', and so on, until two successive sets of values are practically alike.

5. The process can often be shortened by guessing the A value, graphically finding T, and determining A'. If A and A' agree, the guess is correct.

6-19. Series Expansion of Plate Current.[8] For the determination of the general behavior of a circuit a method involving symbols for the circuit parameters is needed. The linear-equivalent-circuit theorem is one of these, but its application is limited to cases of linear operation, as was discussed in Art. 6-15. The student of calculus, familiar with infinite series, can probably readily perceive that the total differential of Eq. (6-21) used in the derivation of this theorem represents the results of using only the terms involving first derivatives in the Taylor-series expansion of the plate current. Under conditions which make the linear-equivalent-circuit theorem give very poor approximations it is possible to obtain much better results if a few more terms of the series expansion are used. Since the algebra involved in the derivation of a usable form of the series is rather complicated, it seems best to omit it from this book and to give only the beginning and the results of the process.

We may begin by assuming $i_b = f(e_b,e_c)$ to be the equation of a family of characteristic curves for a vacuum tube operated with negative grid in a circuit with a plate load. Also we may assume that the variable plate current can be expanded into an infinite series of the form

$$i_p = a_1 e + a_2 e^2 + a_3 e^3 + \cdots \tag{6-57}$$

where

$$e = e_g + \frac{v_p}{\mu} \tag{6-58}$$

and the a's are unknown coefficients involving the tube and circuit parameters. v_p represents a signal voltage introduced into the plate circuit. For this simplified case it must furthermore be stated that the load is a pure resistance R_L and that μ is a constant; otherwise the a's depend upon frequency, and e has a more complicated form. After considerable manipulation the following values of the a's are determined:

$$a_1 = \frac{\mu}{r_p + R_L} \tag{6-59}$$

$$a_2 = \frac{-\mu^2 r_p}{2(r_p + R_L)^3} \frac{\partial r_p}{\partial e_b} \tag{6-60}$$

$$a_3 = \frac{\mu^3 r_p}{6(r_p + R_L)^5} \left[(2r_p - R_L) \left(\frac{\partial r_p}{\partial e_b} \right)^2 - r_p(r_p + R_L) \frac{\partial^2 r_p}{\partial e_b^2} \right] \tag{6-61}$$

In these formulas the values of μ, r_p, and the derivatives are to be determined at the operating point T. The value of $\partial r_p/\partial e_b$ may be found by first plotting vs. e_b the values of r_p obtained along the tube's plate-characteristic curve through T, with e_c held constant at the E_{ct} value. Then $\partial r_p/\partial e_b$ is the slope of this curve at the T point. Similarly the second derivative can be determined by plotting vs. e_b values of $\partial r_p/\partial e_b$ obtained at several points on the curve plotted above. The second derivative desired is the slope of this curve at the point corresponding to T. The accuracy obtainable by this process, of course, rapidly diminishes as the order of the derivative rises. The usefulness of the series expansion appears in problems of general circuit analysis rather than in numerical ones.

A more general form of the series expansion in which the value of μ is not assumed to be constant, and in which the load may be reactive, has been made, but the discussion or inclusion of this form is outside the scope of this book. More complications resulting from grid-current flow through a grid-circuit impedance or from the application of varying voltages to more elements than the control grid and the plate of a multigrid tube can be conceived, and some of these have been treated elsewhere.[5,6] In the case of the pentode, under ordinary conditions of operation, the screen and the suppressor are at constant direct potential, and therefore the series derived for the triode suffices.

6-20. Distortion in Vacuum-tube Circuits. It is usually, though not always, desirable that the waveform of output current or voltage from a vacuum-tube circuit be a faithful reproduction of that of the input voltage. In previous articles it is shown that the nonlinear behavior of a vacuum tube results in the formation of harmonics in the plate current which were not present in the grid voltage. This is called harmonic or nonlinear distortion and is a form of amplitude distortion because its amount depends upon the amplitude of the grid voltage. It occurs whenever the parameters of a circuit are not constant with changing electrode voltages. This effect occurs not only in vacuum tubes but also in a similar way in devices having magnetic cores, because of nonlinearity of the *B-H* characteristic, and also in arcs, in electric lamps, varistors, and many other devices. Often the effect is useful; sometimes it is to be avoided.

Other important types of distortion occur in circuits where inductors or capacitors are used, either deliberately or inadvertently as in the wiring or between elements of the tube. The first of these is called frequency distortion and is the variation in amplitude of the output current or voltage as the frequency only of the input voltage is varied. Since most signals applied to the grid of a tube contain many frequency components, there will be a change in waveform between input and output because of this type of distortion. The effect can be minimized by proper choice of circuit elements and almost eliminated by using only resistors as circuit components. At high frequencies the ever-present tube and residual wiring capacitances still affect the circuit, of course, and at ultrahigh frequencies transit time for electrons in the tube causes frequency distortion. Also at high frequencies even a short length of wire has considerable inductive reactance.

Whenever the phase of the output voltage or current is shifted relative to the phase of the input voltage, there is said to be a time delay in the circuit. In order to keep the time delay independent of frequency, it is necessary that the phase shift be proportional to frequency (see Art. 7-18). If this is not so, then phase distortion results. This type of distortion, while not of great importance in ordinary a-f apparatus, is very undesirable in television circuits and in amplifiers of electronic switches and cathode-ray oscilloscopes. It usually occurs simultaneously with the amplitude-frequency distortion mentioned above.

A fourth form of distortion is called intermodulation distortion. It is related to the harmonic-production process and results from the same cause, *viz.*, the nonlinear character of the circuit or tube. It is the production of frequencies in the output which are the sum and differences of the signal frequencies and their harmonics. This effect occurs whenever two or more frequencies are present in the input voltage and the tube obeys a nonlinear law.

An example will suffice to explain this action. Suppose a tube obeys a quadratic law so that i_p may be expressed as the first two terms of the series expansion of Eq. (6-57): $i_p = a_1e + a_2e^2$. The tube is used in a simple amplifier circuit, and the voltage applied to the grid is $e_g = 4 \sin 100t + 6 \sin 250t$. Since $v_p = 0$, e has the value e_g and therefore the variable plate current is $i_p = a_1(4 \sin 100t + 6 \sin 250t) + a_2(4 \sin 100t + 6 \sin 250t)^2$, where a_1 and a_2 are constants which depend upon tube and circuit parameters. We may expand to obtain

$$\begin{aligned} i_p &= a_1(4 \sin 100t + 6 \sin 250t) + a_2(16 \sin^2 100t + 48 \sin 100t \sin 250t \\ &\quad + 36 \sin^2 250t) = a_1(4 \sin 100t + 6 \sin 250t) + a_2(8 - 8 \cos 200t \\ &\quad - 24 \cos 350t + 24 \cos 150t + 18 - 18 \cos 500t) = 26a_2 \\ &\quad + 4a_1 \sin 100t + 6a_1 \sin 250t - 8a_2 \cos 200t - 18a_2 \cos 500t \\ &\qquad - 24a_2 \cos 350t + 24a_2 \cos 150t \end{aligned}$$

The term $26a_2$ represents I_{p_o}, the approximate increase in direct current caused by the application of the signal to the grid. $4a_1 \sin 100t$ and $6a_1 \sin 250t$ are the fundamental components, having the same frequencies as the input voltage. $8a_2 \cos 200t$ and $18a_2 \cos 500t$ are the second-harmonic-distortion terms in the output. $24a_2 \cos 350t$ and $24a_2 \cos 150t$ are intermodulation products. It will be observed that their angular frequencies are the sum and the difference, respectively, of the input frequencies. Both the harmonic and the intermodulation products contain the coefficient a_2, which is given a meaning by Eq. (6-60):

$$a_2 = \frac{-\mu^2 r_p}{2(r_p + R_L)^3} \frac{\partial r_p}{\partial e_b}$$

The coefficient a_2 and the distortion products can be made small if $\partial r_p/\partial e_b$ is small, *i.e.*, if the characteristic curve for $e_c = E_{ct}$ is nearly straight near T. This can be controlled by choice of the operating point. Also it is seen that distortion is reduced if R_L is made large compared with r_p.

In the amplification of sound, harmonic distortion, while detracting from the fidelity of reproduction, does not result in the production of very objectionable sounds, whereas the intermodulation products may detract considerably from the pleasure of hearing. Since both types of distortion occur together, any measure taken to lessen one will also decrease the other. It should also be noted that since this intermodulation process can produce frequencies nonharmonically related to those originally present in the signal, it will prove to be a very valuable one in some applications.

REFERENCES

1. Cruft Laboratory, War Training Staff: "Electronic Circuits and Tubes," McGraw-Hill Book Company, Inc., New York, 1947.

2. Everitt, W. L.: "Communication Engineering," 2d ed., McGraw-Hill Book Company, Inc., New York, 1937.
3. Franklin, P.: "Differential Equations for Electrical Engineers," John Wiley & Sons, Inc., New York, 1933.
4. Espley, D. C.: The Calculation of Harmonic Production in Thermionic Valves with Resistive Loads, *Proc. IRE*, **21**, 1439 (1933).
5. Brainerd, J. G.: Mathematical Theory of the Four Electrode Tube, *Proc. IRE*, **17**, 1006 (1929).
6. Espley, D. C.: Harmonic Production and Cross Modulation in Thermionic Valves with Resistive Loads, *Proc. IRE*, **22**, 781 (1934).
7. Preisman, Albert: "Graphical Constructions for Vacuum Tube Circuits," McGraw-Hill Book Company, Inc., New York, 1943.
8. Reich, H. J.: "Theory and Applications of Electron Tubes," 2d ed., McGraw-Hill Book Company, Inc., New York, 1944.

PROBLEMS AND QUESTIONS

Characteristic curves for several types of tubes are contained in Appendix A.

1. See Fig. 6-28. (*a*) Draw the d-c-load line. (*b*) Determine E_{bo}, I_{bo}, E_{co}. (*c*) Draw the a-c-load line.

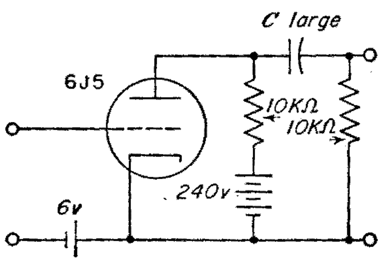


FIG. 6-28.

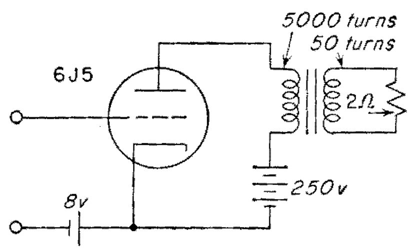


FIG. 6-29.

2. See Fig. 6-29. Assume the transformer winding resistances are zero. (*a*) Draw the d-c-load line, and locate Q. (*b*) Draw the a-c-load line.

3. See Fig. 6-30. Draw the d-c-load line, the bias line, and the a-c-load line.

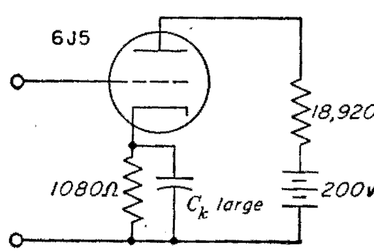


FIG. 6-30.

4. In Prob. 2 assume $e_g = 8 \sin \omega t$ and the T point is Q. Tabulate the values of i_b and of e_b for values of ωt at 30° intervals, 0 to 360°. Plot the waveforms of e_g, i_b, and e_b on the same set of coordinate axes.

5. Two boxes containing storage batteries and series resistors are connected as shown in Fig. 6-31. The d-c ammeter and voltmeter have their polarities indicated. Both meters read upscale. Which box supplies power, and which receives power?

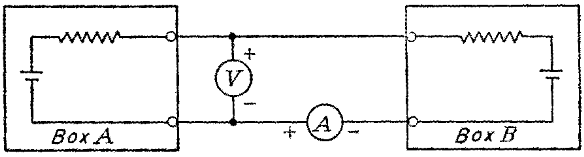


FIG. 6-31.

6. One of the two boxes in Fig. 6-32a is an a-c generator. The other is a load. R is a very small resistance (compared with the load) used to obtain a voltage in phase with the current. The electronic switch is used to show simultaneously on the cathode-ray oscilloscope the waveforms and phase relations for the two voltages applied between the G and A and the G and B terminals of the switch. The picture on the cathode-ray oscilloscope screen is shown in Fig. 6-32b. Which of the boxes is the load?

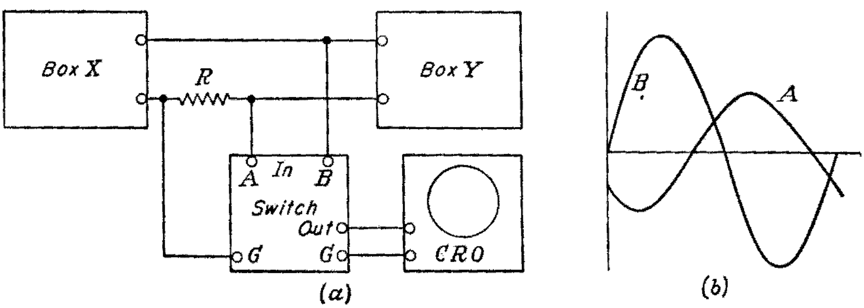


FIG. 6-32.

7. See Fig. 6-33. (a) Solve for the current I, using Thévenin's theorem. (b) Use Norton's theorem to find I.

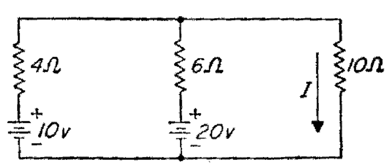


FIG. 6-33.

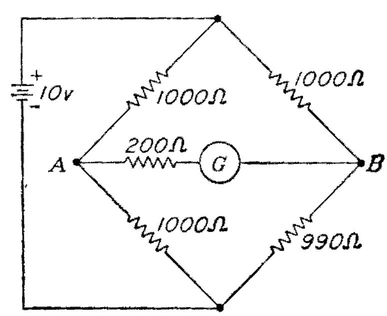


FIG. 6-34.

8. Compute the galvanometer current in the Wheatstone-bridge circuit of Fig. 6-34. *Hint:* Consider A and B as terminals and the 200-ohm galvanometer as the load. Use Thévenin's theorem.

9. A certain vacuum-tube circuit is similar to that of Fig. 6-18 except that the plate circuit contains in addition a generator of voltage v_p with positive sense toward the plate. Determine (*a*) the revised form of Eq. (6-29), (*b*) the revised form of Eq. (6-30).

10. For the tube of Fig. 6-35 $\mu = 20$, $r_p = 10{,}000$ ohms. The grid signal is $\mathbf{E}_g = 0.1 + j0$ volts at a frequency of 1000 cps. (*a*) Draw Thévenin's equivalent plate circuit labeled with $\mathbf{Z}_L$, r_p, $\mathbf{E}_g$, $\mu\mathbf{E}_g$, $\mathbf{I}_p$, $\mathbf{E}_z$, $\mathbf{E}_p$ and applicable sense markings. (*b*) Compute the values of $\mathbf{I}_p$, $\mathbf{E}_z$, and $\mathbf{E}_p$.

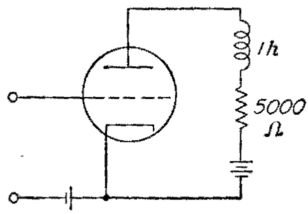


Fig. 6-35.

11. Draw Norton's equivalent plate circuit for the amplifier of Prob. 10. Label with $\mathbf{Z}_L$, r_p, $g_m\mathbf{E}_g$, $\mathbf{E}_z$, $\mathbf{E}_p$, $\mathbf{I}_p$, and applicable sense markings. Solve for $\mathbf{E}_p$ if $\mathbf{E}_g = 0.1 + j0$ volts.

12. In Fig. 6-36, $\mu = 20$, $r_p = 10{,}000$ ohms, $R_b = 5000$ ohms, $R_k = 1000$ ohms. (*a*) Draw Thévenin's equivalent plate circuit. (*b*) Solve for $\mathbf{I}_p$ if $\mathbf{V}_g = 0.5 + j0$ volts.

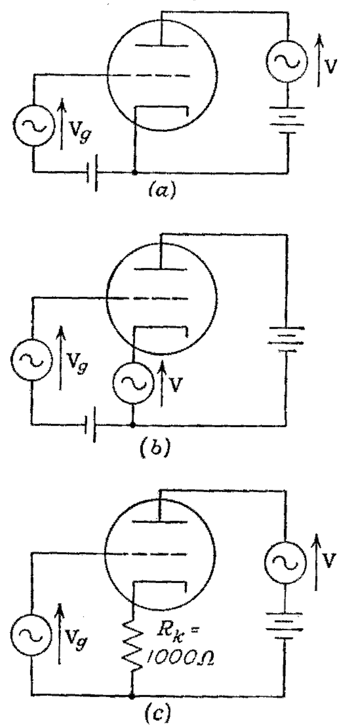


Fig. 6-37.

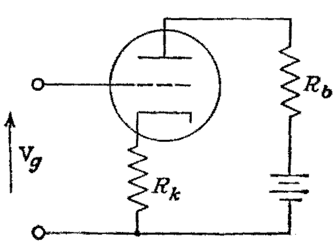


Fig. 6-36.

13. For each circuit in Fig. 6-37, $\mu = 20$, $r_p = 10{,}000$ ohms, $\mathbf{V}_g = 2\underline{/0^\circ}$ volts, $\mathbf{V} = 1\underline{/0^\circ}$ volt, and operation is linear. For each circuit (*a*) draw and label Thévenin's equivalent plate circuit, (*b*) determine $\mathbf{I}_p$, (*c*) determine the impedance $\mathbf{Z}$ seen by the generator $\mathbf{V}$ (consider the generator $\mathbf{V}_g$ shorted when computing this impedance).

14. The plate current for a certain diode consists of a succession of half sinusoids, as shown in Fig. 6-38. Determine the first four terms of the Fourier series for this waveform (*a*) if the origin is at *A*, (*b*) if the origin is at *B*. *Hint:* Replace ωt by $\omega t' - 90^\circ$ in the series for (*a*).

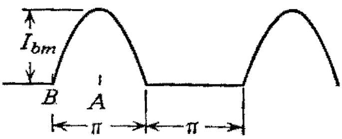


FIG. 6-38.

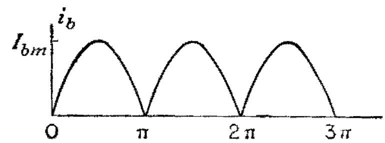


FIG. 6-39.

15. Figure 6-39 shows the waveform for the current from a full-wave rectifier as a succession of half sinusoids. Determine the first four terms of its Fourier-series expansion.

16. Determine the numerical values of I_{ba}, I_{p1}, I_{p2}, and I_{p0}, and the per cent of second-harmonic distortion for the 6J5 tube and circuit used in Probs. 2 and 4.

17. Determine the per cent of third-harmonic distortion for the circuit and operation of Prob. 4. Compared with the second-harmonic distortion determined in Prob. 16, is the amount of third-harmonic distortion appreciable?

18. Figure 6-40 shows the waveform for a certain plate current. Find the numerical values of I_{bt}, I_{ba}, I_{p0}, I_{p1}, I_{p2}, and I_{p3}.

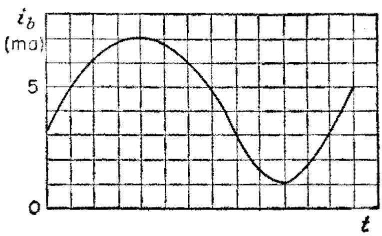


FIG. 6-40.

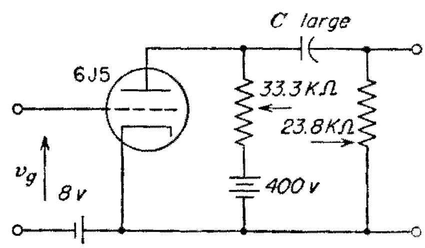


FIG. 6-41.

19. In Fig. 6-41, $v_g = 8 \sin \omega t$. Find the dynamic operating point T. Does the T point shift greatly from the Q point in this case?

20. In the operating region for a certain triode the plate characteristics are straight, and μ is constant. Signals are applied to both grid and plate circuits. The plate load is a pure resistance. (*a*) Write the Taylor-series expansion for the plate current i_p. (*b*) Draw an equivalent circuit using the results of (*a*). Label all elements. (*c*) $\mu = 20$, $r_p = 10{,}000$ ohms, and $R_L = 90{,}000$ ohms. $e_g = 2 \cos 2\pi \times 10^6 t$, and $v_p = 80 \cos 2\pi \times 1000t$ (volts). Find the value of $e_z = i_p R_L$, and determine the frequencies present in e_z.

21. In the operating region for a certain triode the plate characteristics are practically parabolic so that $i_b = 2 \times 10^{-6}(e_b + 20e_c)^2$ amp (e_b and e_c are in volts), $E_{bt} = 200$ volts, $E_{ct} = -8$ volts, $R_L = 93{,}750$ ohms. (*a*) Find μ, r_p, and $\partial r_p/\partial e_b$ at T. (*b*) Write the Taylor-series expansion (e and e^2 terms) for i_p. (*c*) If $e_g = 2 \cos 2\pi \times 10^6 t$ and $v_p = 80 \cos 2\pi \times 1000t$ (volts), what is the expression for e_z? What frequencies are present in e_z?

22. Two triodes are connected in cascade. Each tube operates in somewhat nonlinear fashion and produces distortion in its output. Suppose this action causes the grid voltage for the second tube to be $e_g = 6 \cos \omega t - 0.2 \cos 2\omega t$ volts. For the second tube assume $R_L = 12{,}000$ ohms, $E_{bt} = 156$ volts, $E_{ct} = -6$ volts and the equation of the plate family to be $i_b = 1.39 \times 10^{-5}(e_b + 20e_c)^{1.5}$ amp (e_b and e_c in volts). Determine (*a*) the equation for i_p in terms of e_g, (*b*) the equation for i_p in terms of ωt. (*c*) Isolate the terms representing generated second harmonic from those representing amplified second harmonic, and show that these tend to cancel. (*d*) Note the intermodulation-distortion terms.

23. The circuit of Fig. 6-42 can be used for measuring μ for a vacuum tube. R_1 is adjusted until the 1000-cps tone cannot be heard in the phones. (*a*) Draw the linear equivalent circuit. (*b*) Find μ in terms of R_1 and R_2.

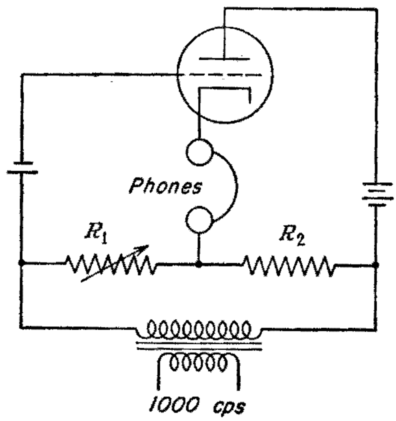


FIG. 6-42.

CHAPTER 7

VOLTAGE AMPLIFIERS

7-1. Classification of Voltage Amplifiers. The voltage produced by a pickup, an antenna, or one of many other devices is often so feeble that considerable amplification is necessary before it becomes great enough to drive the grid of a power tube or to operate some terminating device. Hence voltage amplifiers are of great importance. The tube used may be quite small, either a triode or a pentode, and usually its most important qualification should be a high value of μ. Since the purpose of a voltage amplifier is to reproduce the input waveform in magnified size and with no more distortion than necessary, class A_1 operation is always used. This results in low harmonic distortion, and usually the linear equivalent circuit is quite satisfactory for determining the performance, although the method of Art. 6-16 may be employed if the amount of harmonic distortion must be found. Amplitude-frequency and phase-frequency distortion both become quite troublesome in many voltage-amplifier circuits. This distortion occurs because elements in the tube and circuit have impedance values which vary with frequency.

Other further classification of voltage amplifiers is often desirable. A chemical-reaction-control process may produce a feeble direct voltage which varies very slowly with time. To amplify this voltage a direct-coupled amplifier is useful. This amplifier often will perform also on low- as well as higher-frequency alternating signals. Audio-frequency amplifiers must perform for frequencies over a quite wide range, the extreme case being, say, from 20 to 20,000 cps. Such an amplifier, for which the bandwidth is about twice the mean frequency value, is called a broad-band amplifier. An amplifier may operate in this same range but over only a narrow band, say around 1000 cps; its operation would be classified as narrow band. In cathode-ray oscilloscopes amplifiers are needed which have a range of a few cycles per second to 100,000 cps or higher. These are called wide-band or video-frequency amplifiers, the latter name sometimes being used because the same general type is employed in television, where a range from about 30 cps upward to beyond 4 Mc is needed. For the amplification of radio frequencies, it is usually desirable to have selective action to cause high amplification of wanted signals and rejection of unwanted ones. This calls for some variety of filtering device, tuned loads generally being employed. These r-f amplifiers have high amplification over only a very narrow range and hence are narrow-band amplifiers. For example, the r-f amplifier in a broadcast radio receiver may be adjusted to receive a station with a carrier at 890 kc and sidebands extending from 885 to 895 kc. The ratio of bandwidth to mean frequency is only $^{10}\!/_{890}$, or 0.0112. Some amplifiers of this kind operate on a fixed band and are called intermediate-frequency (i-f) amplifiers. The center of this band may be placed as low as 75 kc or as high as several megacycles. For a-m broadcast use, the ordinary value is 456 kc. For f-m broadcast receivers, 4.3 and 10.7 Mc are in common use at the present time.

7-2. A Simple Single-stage Amplifier with a Plate Load. A simple amplifier with the signal applied between the grid and ground and with a load in the plate-to-ground circuit appears in diagram form in Fig. 7-1. The tube, of course, may be a pentode instead of the triode shown. The desirable output may be voltage, current, or power. In amplifiers with a-c signals applied, usually only the alternating component of the output is of interest. In this case the output voltage $\mathbf{E}_o$ is considered to be $\mathbf{E}_p$ rather than E_b. The positive sense of $\mathbf{E}_p$ is preferable to that of $\mathbf{E}_z$ because $\mathbf{E}_o$ so often serves as $\mathbf{V}_g$ for a following stage, and the senses then agree without any change in sign.

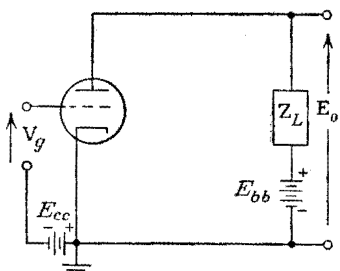


FIG. 7-1. A grounded-cathode grid-driven plate-loaded amplifier.

Before continuing it might be well to mention that the signal voltage can be applied to other circuits instead of between the No. 1 grid and ground, and also the useful output may be taken from the circuit of an element other than the plate. Some of these other possible arrangements will be discussed later, as they have some interesting and desirable features.

Returning to the discussion of the circuit of Fig. 7-1, an approximate analysis may be made, using the equivalent linear circuit. This method gives quite reliable results provided the quiescent operating point has been chosen in a linear region of the tube characteristics, so that if the grid-driving voltage is small, the tube parameters do not vary appreciably from their quiescent values. Figure 7-2 shows two forms of the equivalent circuit for low frequencies.

The voltage amplification **A** of an amplifier is defined as the ratio of the output voltage to the input voltage. For the circuit of Fig. 7-2a

$$\mathbf{I}_p = \frac{\mu \mathbf{V}_g}{r_p + \mathbf{Z}_L} \tag{7-1}$$

and

$$\mathbf{E}_o = \mathbf{E}_p = -\frac{\mu \mathbf{V}_g \mathbf{Z}_L}{r_p + \mathbf{Z}_L} \tag{7-2}$$

Hence

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{V}_g} = -\frac{\mu \mathbf{Z}_L}{r_p + \mathbf{Z}_L} \tag{7-3}$$

This can be written also as

$$\mathbf{A} = -\frac{\mu}{r_p}\frac{r_p \mathbf{Z}_L}{r_p + \mathbf{Z}_L} = -\frac{g_m}{\mathbf{Y}_{sh}} = -g_m \mathbf{Z}_{sh} \tag{7-4}$$

where $\mathbf{Y}_{sh}$ and $\mathbf{Z}_{sh}$ in this case combine r_p in parallel with $\mathbf{Z}_L$. Formula (7-4), which also follows directly from the circuit of Fig. 7-2*b*, is an important one which is useful for many forms of amplifiers which are classified under the title of Fig. 7-1.

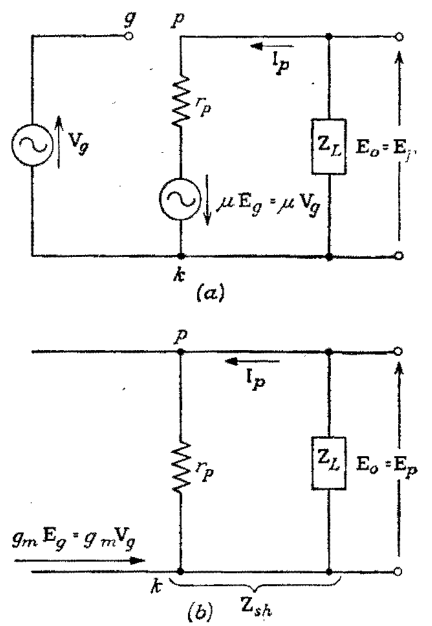


FIG. 7-2. Low-frequency equivalent circuits.

7-3. The Single-stage Amplifier at Higher Frequencies. As the frequency of the signal $\mathbf{V}_g$ is increased, the amplification decreases from the value for low frequencies. This is because of the effect in the circuit of various capacitances, whose reactances decrease with frequency. This effect will now be studied.

Figure 7-3*a* shows the equivalent circuit for this amplifier at higher frequencies, not including those which are very high where lead inductances and the transit time of the electron in the tube become important. C_{gk}, C_{gp}, and C_{pk} are interelectrode capacitances for a triode. For a given tube their values may be ascertained from a tube manual. For a pentode, C_{gk} should be replaced by C_{in}, which is the lumped value of $C_{g1k} + C_{g1g2} + C_{g1g3}$, since the screen and suppressor are at the same alternating potential as the cathode. Likewise C_{pk} must be replaced by C_{out}, representing $C_{pk} + C_{g2p} + C_{g3p}$. Otherwise the circuits are alike for both triodes and pentodes. C_w represents the capacitance to ground of the wiring and circuit elements which are at the alternating potential of the plate. For class A_1 operation there is no appreciable grid conduction current through the tube, and therefore $\mathbf{I}_g$ is mostly displacement current through C_{gk} and C_{gp}. Since $\mathbf{V}_g$ represents the voltage delivered by the source to the position shown, usually it is not necessary to concern ourselves with the effect of $\mathbf{I}_g$ and $\mathbf{Z}_c$, the impedance of the source.* This effect is mostly the concern of the device preceding our amplifier, which must furnish the voltage $\mathbf{V}_g$.

The circuit of Fig. 7-3*a* appears fairly complicated in this form, and so it may be changed to Norton's form. A short placed across xx gives a

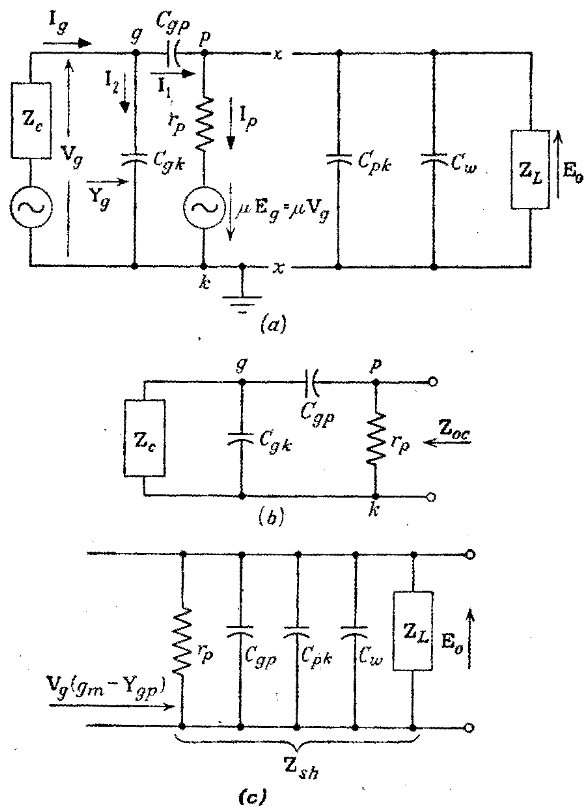


FIG. 7-3. (*a*) High-frequency equivalent circuit. (*b*) Open-circuit impedance to the left of xx. (*c*) Norton's form of (*a*).

short-circuit current of $\mu\mathbf{V}_g/r_p - \mathbf{Y}_{gp}\mathbf{V}_g$, flowing with positive sense upward in the shorting wire. $\mathbf{Y}_{gp}$ is the admittance of C_{gp}. With short and load removed, the open-circuit impedance looking toward the input from xx is shown in Fig. 7-3*b*. If Z_c has a high value (the output impedance of a pentode is an example), the open-circuit impedance is r_p in par-

* In some circuits self-generated oscillations may occur, dependent upon the characters of $\mathbf{Z}_c$ and of $\mathbf{Z}_L$. For an example refer to the tuned-plate tuned-grid oscillator described in Art. 11-4.

allel with an impedance which is approximately that of a capacitor somewhat smaller than C_{gp}. If Z_c has a low value (as in the case of the output impedance of a medium-μ triode), r_p is practically shunted by C_{gp}. To estimate the maximum effect, C_{gp} may be used as a shunt on r_p as shown in Fig. 7-3c. In applications, some judgment should be employed as to whether this maximum or a lesser quantity is the correct value.

Let $\mathbf{Z}_{sh}(= 1/\mathbf{Y}_{sh})$ represent the entire impedance shunted across the output. Then

$$\mathbf{E}_o = \mathbf{E}_p = -\mathbf{V}_g(g_m - \mathbf{Y}_{gp})\mathbf{Z}_{sh} \tag{7-5}$$

and

$$\mathbf{A}_{\text{high}} = -(g_m - \mathbf{Y}_{gp})\mathbf{Z}_{sh} \tag{7-6}$$

Some idea of the importance of the term $\mathbf{Y}_{gp}$ in the factor $g_m - \mathbf{Y}_{gp}$ may be gained by computing values for a 6J5 tube at 10 Mc. Here, approximate values are $g_m = 2600$ micromhos, $C_{gp} = 3.4$ $\mu\mu$f. Hence $g_m - \mathbf{Y}_{gp} = 2600 \times 10^{-6} - j2\pi \times 10^7 \times 3.4 \times 10^{-12} = (2600 - j214) \times 10^{-6}$ mho $= 2610 \times 10^{-6}\underline{/-5°}$ mho. For a pentode 6SJ7 under the usual conditions, approximate values are $g_m = 1650$ micromhos, $C_{gp} = 0.005$ $\mu\mu$f; hence $g_m - \mathbf{Y}_{gp} = 1650 \times 10^{-6} - j2\pi \times 10^7 \times 0.005 \times 10^{-12} = (1650 - j0.31)10^{-6}$ mho $\approx 1650 \times 10^{-6}\underline{/0°}$ mho. From this discussion it is seen that even for a triode the term $\mathbf{Y}_{gp}$ may often be omitted except perhaps for frequencies considerably exceeding 10Mc. Thus formula (7-6) simplifies to

$$\mathbf{A}_{\text{high}} = -g_m\mathbf{Z}_{sh} = -\frac{g_m}{\mathbf{Y}_{sh}} \tag{7-7}$$

which is the same form as for low frequencies. However, here $\mathbf{Z}_{sh}$ for a triode includes $\mathbf{Z}_L$, r_p, C_{pk}, C_w, and C_{gp} in parallel. In the case of pentodes $\mathbf{Z}_{sh}$ becomes $\mathbf{Z}_L$, r_p, C_{out}, and C_w in parallel, since C_{gp} is very small. As the frequency of the signal increases, the reactances of these capacitances decrease, $\mathbf{Z}_{sh}$ changes in value, and hence **A** decreases in magnitude and shifts its phase angle.

7-4. The Output Impedance of a Simple Amplifier. For various reasons it is sometimes important to determine the output impedance of an amplifier. The output impedance of a voltage-producing device is defined here as the rate of change of output voltage with respect to output current. These changes are produced by a change in load. For the simple Thévenin's circuit of Fig. 7-4a we can write $\mathbf{E} + \mathbf{E}_o - \mathbf{I}\mathbf{Z}_{oc} = 0$, and if $\mathbf{E}$ and $\mathbf{Z}_{oc}$ are constants, $d\mathbf{E}_o/d\mathbf{I} = \mathbf{Z}_{oc}$. Thus the output impedance in this case is equal to the open-circuit impedance used with Thévenin's equivalent circuit. There are other ways by which it can be determined, but in many cases this method is the easiest.

For the simple amplifier of Fig. 7-1, a value of the output impedance can be readily obtained by finding Thévenin's equivalent of Fig. 7-3*c*. This is shown in Fig. 7-4*b*. It follows that $\mathbf{Z}_{out}$ is equal to $\mathbf{Z}_{sh}$. It should be noted that this value is not the only one which can be named. For example, if the terminals for viewing $\mathbf{Z}_{out}$ are moved to the left of $\mathbf{Z}_L$, then $\mathbf{Z}_{out}$ will not include $\mathbf{Z}_L$. Hence a knowledge of the terminal position is necessary when one makes a computation of $\mathbf{Z}_{out}$.

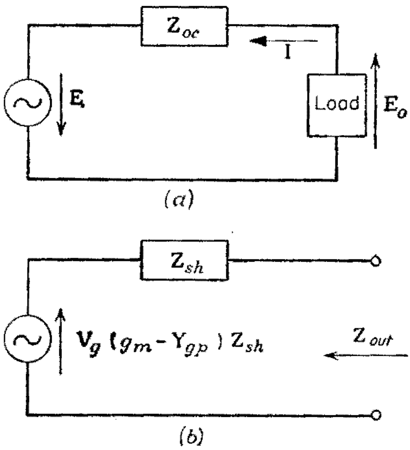


FIG. 7-4. Circuits for the determination of the output impedance.

7-5. The Input Impedance to a Simple Amplifier. The impedance looking into the input of an amplifier is not infinite. If the operation is class 2, there is conduction current between grid and cathode during part of the cycle of grid voltage, and even for a class 1 amplifier there is some displacement current through the interelectrode capacitances, particularly at high frequencies. The resulting input impedance will be an additional shunt on any amplifier or other device used to furnish the input signal. Because it is a shunt load, it is usual to determine the input admittance $\mathbf{Y}_g$, defined as $\mathbf{I}_g/\mathbf{V}_g$ (see Fig. 7-3*a*) rather than a value of $\mathbf{Z}_g$.

The input admittance $\mathbf{Y}_g$ of the simple amplifier can be determined by referring to Fig. 7-3*a*.

$$\mathbf{Y}_g = \frac{\mathbf{I}_g}{\mathbf{V}_g} = \frac{\mathbf{I}_1 + \mathbf{I}_2}{\mathbf{V}_g} \tag{7-8}$$

$$\mathbf{I}_1 = (\mathbf{V}_g - \mathbf{E}_o)j\omega C_{gp} = (\mathbf{V}_g - \mathbf{A}\mathbf{V}_g)j\omega C_{gp} \tag{7-9}$$

and

$$\mathbf{I}_2 = \mathbf{V}_g j\omega C_{gk} \tag{7-10}$$

Therefore

$$\mathbf{I}_g = \mathbf{V}_g j\omega[C_{gk} + (1 - \mathbf{A})C_{gp}] \tag{7-11}$$

and

$$\mathbf{Y}_g = j\omega[C_{gk} + (1 - \mathbf{A})C_{gp}] \tag{7-12}$$

For a pentode C_{gk} is replaced by C_{in}. For most pentodes C_{gp} is so small in value that $(1 - \mathbf{A})C_{gp}$ is negligible compared with C_{in}; hence for a pentode

$$\mathbf{Y}_g \approx j\omega C_{in} \tag{7-13}$$

In Eq. (7-12), $\mathbf{A}$ is in general a complex number so that $\mathbf{Y}_g$ will have both real and imaginary components. Thus

$$\mathbf{Y}_g = G_g - jB_g{}^* \tag{7-14}$$

Two examples may be helpful.

1. Suppose $\mathbf{A} = A\underline{/180°}$, as for an amplifier with $\mathbf{Z}_{sh}$ a pure resistance. Then $\mathbf{Y}_g = j\omega[C_{gk} + (1 + A)C_{gp}]$ and $G_g = 0, B_g = -\omega[C_{gk} + (1 + A)C_{gp}]$. The quantity $C_{gk} + (1 + A)C_{gp}$ is the input capacitance C_g. For a triode the contribution of C_{gp} is much more than that of C_{gk}.

2. Suppose $\mathbf{A} = A\underline{/225°} = A(-0.707 - j0.707)$, as for an amplifier with inductive $\mathbf{Z}_{sh}$. Then

$$\mathbf{Y}_g = j\omega[C_{gk} + (1 + 0.707A + j0.707A)C_{gp}] = -\omega 0.707AC_{gp} + j\omega[C_{gk} + (1 + 0.707A)C_{gp}]$$

Hence $G_g = -\omega 0.707AC_{gp}$ and $C_g = C_{gk} + (1 + 0.707A)C_{gp}$.

It is seen that, for purely resistive $\mathbf{Z}_{sh}$, G_g is zero and, for inductive $\mathbf{Z}_{sh}$, G_g may be negative. It may be determined by trial that if $\mathbf{Z}_{sh}$ is capacitive, G_g is positive. From the explanation of Art. 6-10, a negative value of G_g means that energy is being fed into the grid circuit from the plate circuit. At high frequencies, G_g, which increases with frequency, may become quite large. In parallel with other resistors in the plate circuit of the preceding tube the total G may become negative. In that case the second tube will supply some of its own grid excitation and oscillations will occur for some values of $\mathbf{Z}_c$.

7-6. Amplification, Decibels, Gain. The voltage amplification $\mathbf{A}$ of an amplifier has been defined as the ratio of the output voltage to the input voltage. The amplification produced by a single stage is often not enough to raise the output voltage to the required level. In such a case the output of the first stage is fed into a second stage, and hence if necessary to a third stage, and so on, until the output level is satisfactory. The result is a cascade, or multistage, amplifier. If the input to the first stage is called $\mathbf{V}_{1g}$, the output of this stage becomes the input $\mathbf{V}_{2g}$ for the second. Note that the numerical subscript preceding the letter subscript refers to the number of the stage and not to the number of the grid. Then for the first stage the definition for amplification becomes

$$\mathbf{A}_1 = \frac{\mathbf{V}_{2g}}{\mathbf{V}_{1g}} \tag{7-15}$$

$\mathbf{V}_{1g}$ and $\mathbf{V}_{2g}$ are both complex quantities, and it follows that $\mathbf{A}_1$ is also a complex quantity. Its size A_1 is, of course, the magnitude of the amplification; its angle θ_{1A} is the angle by which $\mathbf{V}_{2g}$ leads $\mathbf{V}_{1g}$. This phase shift is

* This definition of admittance is from ASA Definitions of Electrical Terms.

due partly to the action of elements in the circuit, but mostly to the choice of the sense of $\mathbf{E}_p$ rather than that of $\mathbf{E}_z$ as the sense for output voltage $\mathbf{E}_o$.

For two stages $\mathbf{E}_o = \mathbf{V}_{2g}\mathbf{A}_2$, and $\mathbf{V}_{2g} = \mathbf{V}_{1g}\mathbf{A}_1$; therefore $\mathbf{E}_o = \mathbf{V}_{1g}\mathbf{A}_1\mathbf{A}_2$, and $\mathbf{A} = \mathbf{E}_o/\mathbf{V}_{1g} = \mathbf{A}_1\mathbf{A}_2$. Thus the amplification of a cascade amplifier is the product of the amplifications of the individual stages.

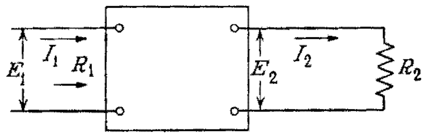


FIG. 7-5. A four-terminal network.

When several stages are used, it is often convenient to use a method based upon that employed with communication networks for computing the loss or gain in power (see Fig. 7-5). For this purpose the decibel (db) is defined:

$$\text{No. db gain} = 10 \log_{10} \frac{P_2}{P_1} \tag{7-16}$$

the factor 10 being employed since the bel is too large a unit for convenient use. If the input and load impedances are R_1 and R_2, respectively,

$$\text{No. db gain} = 10 \log \frac{I_2^2 R_2}{I_1^2 R_1} = 10 \log \frac{E_2^2/R_2}{E_1^2/R_1} \tag{7-17}$$

or if $R_1 = R_2$, as is sometimes the case,

$$\text{No. db gain} = 20 \log \frac{I_2}{I_1} = 20 \log \frac{E_2}{E_1} \tag{7-18}$$

These formulas (7-16) to (7-18) all measure *power* gain or loss.

In voltage amplifiers power gain is seldom discussed as the grid usually takes negligible power. Hence the definitions above sometimes are not strictly applied. Instead, the requirement that $R_1 = R_2$ is dropped, and

$$\text{No. db voltage gain} = 20 \log \frac{E_2}{E_1} = 20 \log A \tag{7-19}$$

is used as a definition, the word "voltage" being carefully inserted to avoid confusion. For a cascade amplifier this becomes

$$\text{No. db voltage gain} = 20 \log \cdot A_1 A_2 A_3 \cdots = 20 \log A_1 + 20 \log A_2 + 20 \log A_3 + \cdots \tag{7-20}$$

Hence the decibel voltage gain of a multistage amplifier is the sum of the decibel voltage gains of the separate stages. It should be noted that in the above formulas only scalar values are used, as the inclusion of the angles without further definition would be meaningless.

As the amplification A of an amplifier varies because of changing frequency, it is convenient to compare the amplification at one frequency with its value at another as a standard. This may be done by computing the ratio A_{f_1}/A_{f_0}, where f_1 refers to the frequency under consideration and f_0 to the standard frequency. The comparison may also be made using decibels. Since the A's represent ratios of output voltage to input voltage, the definition involving voltages may be adapted to use. Thus

$$\text{No. db drop in voltage gain} = 20 \log \frac{A_{f_0}}{A_{f_1}} \tag{7-21}$$

The word "drop" rather than "rise" is used since the ratio of A's has been inverted from its first expression above in order to avoid the necessity of determining the logarithm of a number less than unity. If A_{f_1} exceeds A_{f_0}, the ratio in expression (7-21) may be inverted and the result called decibel rise in voltage gain.

As an example, if the value of A at 100 cps is 70.7 per cent of its value at 1000 cps, the drop in voltage gain is 20 log (1/0.707) = 20 log 1.414 = 20 $\times$ 0.1505 $\approx$ 3 db.

The voltage output of pickups and microphones is often expressed in decibels. Thus a manufacturer's catalogue may state that a certain microphone has an output of -54 db, where 0 db = 1 volt, the load is 5 megohms, and the sound pressure is 1 dyne per cm^2. The actual voltage output under these conditions can be calculated as follows: 54 = 20 log $(1/E)$ or log $(1/E) = 2.7$, $1/E = 500$, and $E = 0.002$ volt.

The zero level is not always 1 volt for voltage decibels, nor is there a uniform standard for power use, 6 mw into a 500-ohm load being common but not universal in telephone practice. Unfortunately, many catalogues state the output of a device in decibels without any statement as to the standards used, and hence the information given is meaningless. If the output is high enough perhaps to overdrive the grid of an amplifier, it is best to state the output directly in volts.

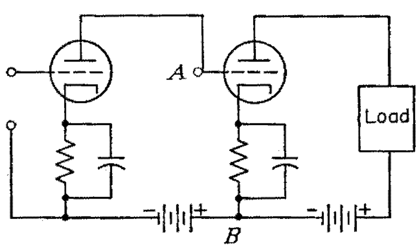


FIG. 7-6. A circuit like this will not work.

7-7. Methods of Coupling Multistage Amplifiers. A current i_b flows from plate to cathode in the first tube of an amplifier, while no current flows from cathode to grid in the second tube. Hence, one cannot connect the grid of the second tube directly to the plate of the first tube, in the manner shown in Fig. 7-6, to obtain satisfactory operation. Insertion of a suitable coupling network between the two tubes is necessary. One type sometimes employed is so-called direct coupling, produced by inserting a resistor between A and B in the circuit of Fig. 7-6. Other types are called resistance-capacitance coupling, untuned-transformer coupling, impedance coupling, and tuned-transformer coupling. The type of coupling used depends upon the requirements of the amplifier. A discussion of each of these follows.

7-8. Direct-coupled Amplifiers.[1,2] The purpose of direct coupling is to eliminate the capacitors, inductors, and transformers used in other types. These cause a loss in amplification at low frequencies and fail to give any response at all to direct voltages. Many circuits employing direct coupling have been devised. Some use several batteries, but this practice is generally undesirable.

The type of circuit shown in Fig. 7-7 was devised by Loftin and White. A rectifier type of power supply can be used instead of the battery, if

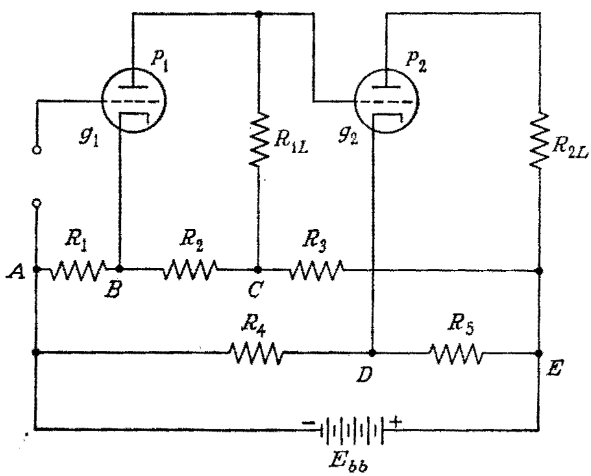


FIG. 7-7. A Loftin-White direct-coupled amplifier.

desired. The manner in which the resistor values are chosen, as well as the general action of the circuit, can be seen from the following example.

In Fig. 7-7 suppose that the tubes are type 6J5 and that R_{1L} is 100,000 ohms and R_{2L} is 10,000 ohms. Assume that the values of $E_{bo} = 90$ volts, $E_{co} = -3$ volts determine a suitable operating point for each tube. Measurements taken on a vacuum-tube bridge with this operating point yield $I_{bo} = 1$ ma, $r_p = 16{,}000$ ohms, and $\mu = 19.3$. Point A may be labeled 0 volt, and under quiescent conditions point B will be at +3 volts potential and p_1 will be at a potential of +93 volts. The voltage drop in R_{1L} is $0.001 \times 100{,}000 = 100$ volts, making point C at a potential $100 + 93 = 193$ volts above that at point A. The point g_2 is at a potential of +93 volts; therefore D should be at $93 + 3 = +96$ volts. p_2 has a potential of $96 + 90 = 186$ volts, and the drop in R_{2L} being $0.001 \times 10{,}000 = 10$ volts, the potential at E should be $186 + 10 = 196$ volts, the voltage required for E_{bb}.

For this problem the load will be considered to be a relay with 10,000 ohms resistance. The power-circuit contacts of the relay will be considered normally open and will close with a minimum current flow of 1.2 ma through the relay coil. Thus under quiescent conditions the relay contacts are open.

It is generally desirable to make the resistance of the bleeders R_1, R_2, etc., low enough so that changes in current through the tubes and these resistors due to the applied signal will change the potentials at points A, B, C, etc., by only a small percentage. For this reason it is assumed that these resistors have no effect on the operation of this circuit.

One may next draw Fig. 7-8, the linear equivalent circuit for the two-stage amplifier. Recalling that $i_p = \Delta i_b$, $e_g = \Delta e_c$, and $e_p = \Delta e_b$ as definitions, one may replace the usual quantities and label the circuit with

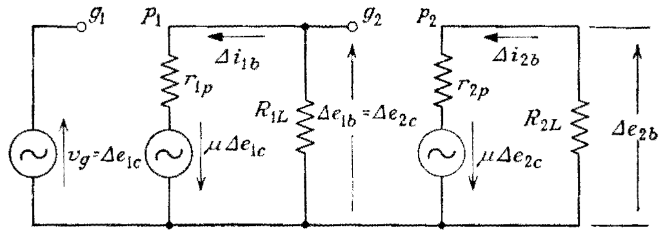


FIG. 7-8. An equivalent circuit useful for a two-stage amplifier handling slow grid-voltage changes.

the incremental ones, giving an equivalent circuit useful for small changes in voltage and current. It follows that

$$\Delta e_{1b} = -\Delta i_{1b} R_{1L} = -\frac{\mu R_{1L}\,\Delta e_{1c}}{r_p + R_{1L}} \qquad \text{and} \qquad \mathbf{A}_1 = \frac{\Delta e_{1b}}{\Delta e_{1c}} = -\frac{\mu R_{1L}}{r_p + R_{1L}}$$

Hence the value of $\mathbf{A}_1$ is $-19.3 \times 100{,}000/(16{,}000 + 100{,}000) = -16.6$. Likewise $\mathbf{A}_2 = -\mu R_{2L}/(r_p + R_{2L}) = -19.3 \times 10{,}000/(16{,}000 + 10{,}000) = -7.4$. Hence, $\mathbf{A} = \mathbf{A}_1\mathbf{A}_2 = -16.6 \times -7.4 = +123$.

The least value of Δi_{2b} required to cause the relay contacts to close is $+0.2$ ma, which requires a value of Δe_{2b} of $-0.2 \times 10^{-3} \times 10{,}000 = -2$ volts. The necessary value of input voltage Δe_{1c} is $-2/123 = -0.016$ volt or -16 mv. If this voltage is not available, a third stage may be added or perhaps tubes having a higher μ might be used. The negative value of Δe_{1c} obtained means that the input device must be connected to give a polarity opposite to that marked for the sense of Δe_{1c} in Fig. 7-8 above.

When direct-coupled amplifiers are used with l-f alternating-voltage signals, their operation should be quite satisfactory. But if the input is a low direct voltage, many factors contribute to make them troublesome.

1. The values of E_{bb} and E_{cc} must be maintained constant, either with batteries or with voltage-regulated power supplies. Otherwise the operating voltages and currents for the tubes will change. It is particularly

important to maintain the grid bias for the first tube constant, as any variation there is as potent as a signal of the same amount.

2. The initial velocity of electrons leaving the cathode has the same effect as the insertion of a small battery in the cathode lead, with the initial velocity assumed to be zero. If the heater voltage varies even slightly, the equivalent battery voltage changes the grid-to-cathode bias and hence the plate current. Thus the heater voltage must be maintained constant. In some amplifiers filament-type tubes are employed, and the current through the bleeder of the regulated plate-voltage supply also heats the filaments.

3. Grid current flowing through the impedance of the external grid circuit of the first tube causes a voltage drop which masks the desired input signal. This grid current is caused by one or more of several factors. High-initial-velocity electrons may strike even a negatively biased grid. A hot grid may emit because of cathode-emitting material sputtered on it during manufacture. Since even the most thoroughly evacuated tube contains some residual gas, positive ions may be present, which are attracted to the grid and may even knock secondary electrons from that electrode. There are more or less effective preventatives for grid current. Operation with low heater voltage reduces the initial velocities of the electrons, and the use of low plate voltages reduces the danger of ionization of residual gas. Careful manufacture to avoid sputtering of the grid and to reduce residual gas also helps.

4. The emission of coated and thoriated cathodes is not uniform over the cathode surface and varies slowly with time at any one spot. Because the grid wires produce a nonuniform field at the cathode surface, there is a resultant small effect on the space current. For amplifiers with input voltage of very low level this is probably the most serious difficulty in direct-coupled amplifier design.

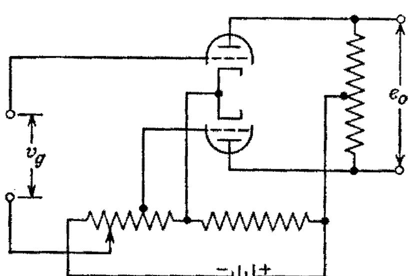


FIG. 7-9. A balanced voltage amplifier.

Balanced circuits are used with some success to cancel the effects of these random fluctuations. Figure 7-9 shows one stage of such an amplifier circuit. If an initial adjustment is made so that no voltage output is obtained with zero signal, then variations of cathode temperature or electrode-supply voltages have no effect providing the tubes behave identically. Practically, it is impossible perfectly to match the tubes so that their coefficients are identical, and hence complete cancellation of errors does not occur. Also the effect under item 4 is not helped by the method.

If possible it is probably better to avoid having to amplify a small direct voltage. Sometimes a vibrator-type switch may be used to break up a direct voltage, and the pulsating voltage obtained may be amplified with an ordinary a-c amplifier, which is relatively trouble free. Following this, the output may be rectified if desired.

7-9. Resistance-capacitance-coupled Amplifiers. The circuit diagram for two stages of an R-C-coupled amplifier using triodes is shown in Fig. 7-10, while one employing pentodes is shown in Fig. 7-11. Except for the voltage-dropping resistors and the bypass capacitors in the screen-grid circuits, the arrangements are identical.

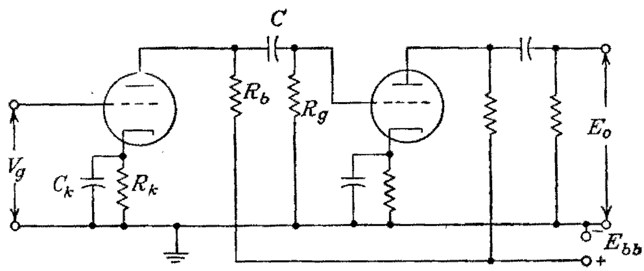


FIG. 7-10. An R-C-coupled triode amplifier.

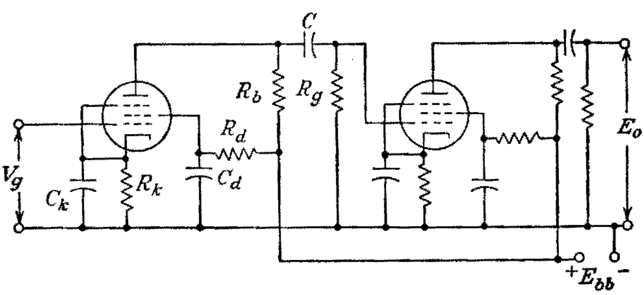


FIG. 7-11. An R-C-coupled pentode amplifier.

The capacitor C serves to prevent the high direct potential on the plate from being applied to the following grid and helps to block slow variations of voltage across R_b from passing on through the amplifier; yet it allows the higher-frequency signal to be amplified by the next stage.

R_g serves two or more purposes. It allows the grid to be at ground potential for direct currents (if the grid conduction current is negligible), while the cathode has a positive bias to ground because of current flow through R_k. Second, it permits C to charge as fast as slow variations in voltage across R_b change the potential of its left plate. This causes the total voltage drop across R_b, except the higher-frequency components due to the signal, to be established across C and none of it across R_g. This is an advantage for purposes of stability, a disadvantage if the signal itself is of low frequency. Third, R_k may be reduced in size or even omitted and R_g used to make a grid bias due to grid-current flow. Usually there is no advantage in doing this for an R-C-coupled amplifier, but the method is occasionally used. Under some circumstances it may be actually unsafe, since if grid current flows because of grid emission, it may give a positive bias instead of the negative one desired.

In the pentode circuit the bias for the screen grid in this case is obtained from the plate-voltage supply. For voltage-amplifier pentodes the screen voltage desired is often about 100 volts, while the plate-supply voltage is perhaps 250 volts. Although a screen-power supply could be employed, this is generally uneconomical. Or a voltage divider across E_{bb} could be tapped at 100 volts. The series dropping resistor used is the ordinary method, however. Since the plate current is approximately proportional

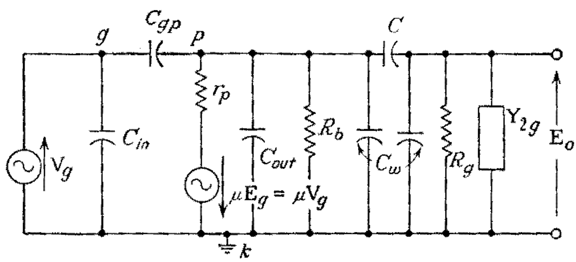


FIG. 7-12. Equivalent circuit for the entire frequency band.

to the screen current, but larger, if the direct screen current can be maintained constant, the direct plate current will remain approximately constant too, thus tending to keep the operating point of the tube from shifting. Any slow rise in screen current causes a corresponding drop in screen potential, which tends to prevent the current rise. Thus a series dropping resistor is a stabilizing element. It must always be bypassed by a capacitor of sufficient size to prevent appreciable signal-frequency variation in the screen potential; otherwise the tube will perform somewhat like a triode. It is assumed in the amplifier circuit under discussion that R_d and C_d have been properly chosen, and hence the screen grid is essentially at alternating ground potential.

The bias for the control grid is usually obtained by using R_k as shown. Besides being simple and cheap, the device is desirable from the stability standpoint, its action being similar to that of the screen-grid resistor for the pentode. Likewise C_k needs to be of adequate size, or the potential of the cathode will vary relative to ground. While this variation is sometimes desired, for the present it will be assumed that the cathode is at alternating ground potential.

7-10. The Linear Equivalent Circuit for the *R*-*C*-coupled Amplifier. In order to study the performance of this amplifier, the linear equivalent circuit is first drawn in Fig. 7-12. Unlike the direct-coupled amplifier of Art. 7-8, which is designed particularly for low- or zero-frequency use, this one fails at very low frequencies. Above these it operates over a wide band, and the capacitances between tube electrodes and between wiring and ground become important in the higher regions of this band. Tube capacitances are represented by C_{in}, C_{gp}, and C_{out}. C_{in} includes all capacitances between the control grid and tube elements which are at cathode potential. For a triode this is C_{gk}. For a pentode $C_{\text{in}} = C_{g1k} + C_{g1g2} + C_{g1g3}$, if the circuit is arranged in the manner of Fig. 7-11. C_{out} is C_{pk} for a triode and $C_{pk} + C_{g2p} + C_{g3p}$ for a pentode. Sinusoidal waveshape of applied voltage is assumed. Note that in this circuit $\mathbf{E}_g$ is equal to $\mathbf{V}_g$.

The box labeled $\mathbf{Y}_{2g}$ in the equivalent circuit represents the input shunt admittance to the next stage, if any (see Art. 7-5.). It is equivalent to a capacitor and a resistor in parallel. With the frequencies ordinarily used with R-C-coupled amplifiers of this type, the resistance is very high

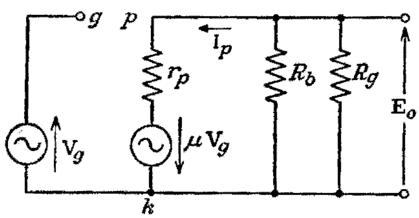


FIG. 7-13. Thévenin's circuit for the mid-frequency range.

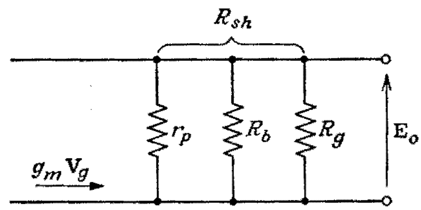


FIG. 7-14. Norton's equivalent circuit for the mid-frequency range.

compared with the grid leak R_g, and it may be either neglected or considered to be combined with R_g.

While an analysis of this circuit can be made which is valid for all frequencies for which the circuit is useful, the labor is considerable and the results will be in a form not particularly simple. Rather, it is better to separate the frequency band formally into three ranges. The middle frequencies are those high enough so that there is no appreciable loss in voltage in C, but low enough so that the numerous shunt capacitances collectively have no appreciable effect. Such a range may not exist for a particular amplifier circuit, but it is quite possible to design amplifiers which have a wide mid-frequency range. The low range includes those frequencies for which the coupling capacitor C noticeably affects the circuit performance. In the high range the shunt capacitors have low enough impedance so that the gain is no longer the same as in the mid-frequency range.

7-11. Behavior of the Amplifier in the Mid-frequency Range. Omission of the capacitors which are unimportant in this range yields the simplified drawing of Fig. 7-13. The grid circuit may be omitted if desired because the grid voltage equals $\mathbf{V}_{1g}$, unaffected by any circuit components in this amplifier stage. Norton's form is shown in Fig. 7-14.

Norton's form will be used in this analysis, it being the simpler of the two. Ohm's law yields

$$\mathbf{E}_o = -g_m \mathbf{V}_{1g} R_{sh} \tag{7-22}$$

or

$$\mathbf{A}_{\text{mid}} = \frac{\mathbf{E}_o}{\mathbf{V}_{1g}} = -g_m R_{sh} \tag{7-23}$$

where R_{sh} is the equivalent of r_p, R_b and R_g in parallel. Note that formula (7-4) applies here with $\mathbf{Z}_{sh} = R_{sh}$. The frequency range over which this formula reliably represents the amplification is yet to be determined.

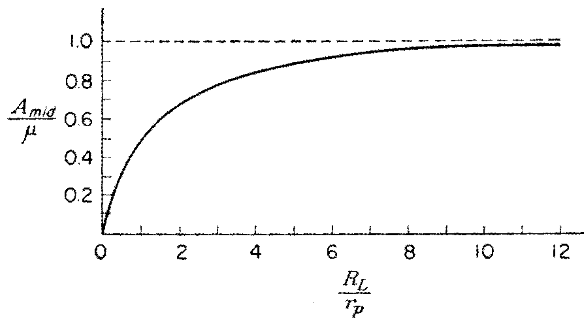


FIG. 7-15. Theoretical curves showing how amplification varies in the mid-frequency range with the a-c load.

It is well to investigate the implications of the simple formula (7-23). It may be written in other forms,

$$\mathbf{A}_{\text{mid}} = \frac{-\mu}{r_p\left(\dfrac{1}{r_p} + \dfrac{1}{R_b} + \dfrac{1}{R_g}\right)} = \frac{-\mu}{r_p\left(\dfrac{1}{r_p} + \dfrac{1}{R_L}\right)} = \frac{-\mu}{1 + \dfrac{r_p}{R_L}} \tag{7-24}$$

where R_L is the total plate-circuit a-c load. For a given quiescent operating point, μ and r_p are determined, and it is assumed that they remain constant as the grid voltage moves the instantaneous point of operation up and down the a-c-load line. This is essentially true as long as the grid signal voltage is small. The value of R_L may be varied, and of course A_{mid} varies with it. Figure 7-15 shows graphically the interdependence. This graph is said to be normalized because R_L is plotted in units of r_p and A_{mid} in units of μ; it applies, therefore, to any amplifier with a resistance plate load. As R_L increases, A_{mid} approaches μ; therefore μ is the upper limit to the amplification obtainable for this kind of circuit.

Since R_L is the parallel equivalent of R_b and R_g, it is always less than the smaller of these. R_g usually ranges between 250,000 ohms and 1 megohm, and R_b is smaller, for reasons to be given later. Hence, to increase R_L requires that R_b be increased. R_b is the d-c plate load if R_k is neglected. In Fig. 7-16, lines a, b, and c represent d-c loads for a medium-μ triode, with resistance values increasing in that order, and passing through the same Q point. Note that as the resistance R_b is increased, the value of E_{bb} required is increased. Therefore, for a fixed Q point it is not practicable to increase R_b greatly in an effort to obtain greater amplification. Lines a, d, and e represent increasing d-c load resistances, but the value of E_{bb} is held constant. Their intersections with the bias line f shifts, thus moving the operating point downward and to the left as R_b is made greater. Although formula (7-24) seems to show that A increases with R_L, this is not necessarily true, as μ may decrease and r_p may increase,

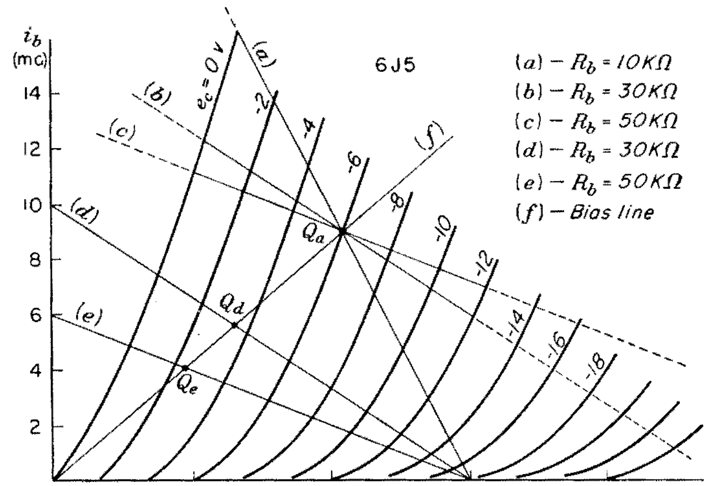


FIG. 7-16. Plate diagram of a triode with various loads.

which is exactly what happens in this case. Also the operation is somewhat more nonlinear for the very low Q points; hence the practical limit to R_b for medium-μ triodes is about $20r_p$, and for reasons to appear later it is often better to use a smaller value. For high-μ triodes the upper limit to R_b is about $10r_p$, and for pentodes its value is usually not higher than $0.3r_p$. In the latter case it might seem that little advantage is being taken of the very high value of μ, but even with this limit on the load an amplification up to around 200 can be obtained, which far exceeds that possible with a triode.

Nonlinear distortion for a triode is mostly second harmonic. Formula (6-71) is sufficiently accurate for its determination, and the method explained in connection with this formula may be used. If R_g is assumed to be very high, the lines of Fig. 7-16 become approximately the a-c-load lines. If drive to 0 volts is also assumed, although this is more than one is likely to use, then the lengths of segments of the operating region of the load lines above and below the Q_a point (shown solid) can be compared. It is fairly obvious by inspection that, for the tube represented in Fig. 7-16, a higher load value gives less distortion. That this distortion is very small for very high loads is also attested by the quite constant value of μ over a wide region, since the graphical determination of μ involves horizontal measurements between the curves of the plate family. If a similar study is made for high-μ triodes, one arrives at the same conclusion, *viz.*, that higher loads give less nonlinear distortion (see also Art. 6-19).

The grid bias to be used for triode voltage amplifiers is quite readily determined. The dynamic transfer characteristic of Fig. 6-11, which is the operating curve on I_b, E_c axes, is quite typical of triodes. It is straighter near the upper end, and hence one should locate the quiescent operating point there for minimum nonlinear distortion. For a medium-μ triode, grid current becomes appreciable for grid voltages more positive than about -1 volt, and operation should be kept below this region. Thus, if a signal of 2 volts peak value is to be amplified, the bias should be about -3 volts.

For pentode voltage amplifiers the upper end of the dynamic characteristic bends downward, making it undesirable to locate the operating point so near zero grid voltage. As an example, Fig. 4-26 shows the dynamic characteristics of a 6SJ7 pentode for $E_{bb} = 250$ volts, $R_b = 100{,}000$ ohms, and various screen voltages. The curves for low screen voltages have somewhat longer straight portions, and hence it is usually advisable to use low screen-voltage operation. However, since control-grid current flows for voltages more positive than about -1 volt, the curve for $E_{c2} =$ 25 volts should ordinarily be avoided with this value of R_b. The curve for $E_{c2} = +50$ volts has an inflection point near $I_b = 1.7$ ma, and operation at this point (requiring $E_{c1} = -1.25$ volts) will give higher gain with small signals than will operation at any neighboring point on the same curve. However, for large signals, it might be well to avoid the predominant third-harmonic distortion, which occurs with operation at an inflection point, and choose instead a somewhat lower value of I_{bo}, where the distortion is more second harmonic and less third. If the value of R_b were changed, a new set of curves would be obtained. Since in actual operation R_g is in parallel with R_b as an a-c load, the curves of Fig. 4-26, which are taken with direct applied voltages, cannot be used to determine exactly the allowable drive and the harmonic distortion. In general, the load for a pentode voltage tube is not critical and for low drive neither is the choice of operating point, as long as it lies in the straighter portion of the characteristic curve. For higher values of grid drive more care is needed correctly to place the operating point. The screen bias is usually obtained from the plate supply through a well-bypassed voltage-dropping resistor.

In practice the curves of Fig. 4-26 are not given in tube manuals, and hence, unless laboratory facilities are available, the method just explained is not particularly adaptable. If the value of I_{bo} were stated in the manual and either E_{c1} or E_{c2} were adjusted to obtain this value, successful operation for signals of moderate size could be obtained. Instead, however, most manuals present ready-made designs for amplifiers, and usually it is best to use them. The same can be done for triodes if desired.

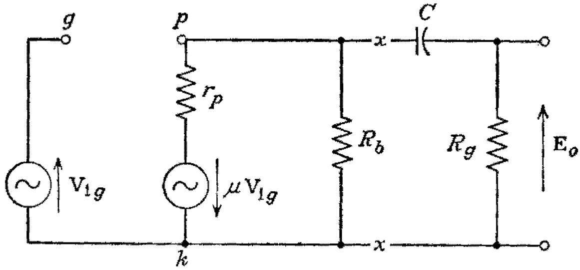


FIG. 7-17. Equivalent circuit for the l-f range.

7-12. Performance of an R-C-coupled Amplifier in the L-F Range. For low frequencies the equivalent circuit of Fig. 7-12 simplifies to that of Fig. 7-17. Because of the presence of C this circuit is somewhat more complicated to analyze than that for the middle frequencies. Although this is a Thévenin's form from the viewpoint of the whole load, if we consider the portion to the right of xx as the load, we obtain another Thévenin's form as shown in Fig. 7-18.

To simplify the analysis, let us define

$$R_{\text{low}} = R_g + \frac{r_p R_b}{r_p + R_b}. \tag{7-25}$$

Then

$$\mathbf{E}_o = \frac{-R_g}{R_{\text{low}} - jX_c} \frac{R_b}{r_p + R_b} \mu \mathbf{V}_{1g} \tag{7-26}$$

and

$$\mathbf{A}_{\text{low}} = -\frac{R_g}{R_{\text{low}} - jX_c} \frac{\mu R_b}{r_p + R_b} \tag{7-27}$$

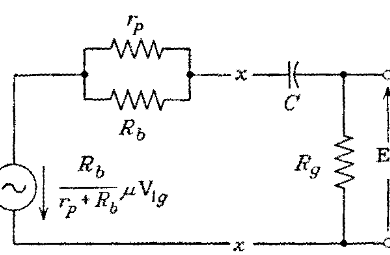


FIG. 7-18. Another Thévenin's form of the equivalent circuit of Fig. 7-17.

Note that if C were omitted from this circuit, one would again have a middle-frequency equivalent circuit. Hence, for middle frequencies

$$\mathbf{E}_o = -\frac{R_g}{R_{\text{low}}} \frac{R_b}{r_p + R_b} \mu \mathbf{V}_{1g} \tag{7-28}$$

and

$$\mathbf{A}_{\text{mid}} = -\frac{R_g}{R_{\text{low}}}\frac{\mu R_b}{r_p + R_b} \tag{7-29}$$

Dividing Eq. (7-27) by Eq. (7-29) and rearranging, we obtain

$$\frac{\mathbf{A}_{\text{low}}}{\mathbf{A}_{\text{mid}}} = \frac{R_{\text{low}}}{R_{\text{low}} - jX_c} = \frac{1}{1 - j(X_c/R_{\text{low}})} \tag{7-30}$$

Rather than use the already relatively simple formula (7-30), it is more convenient to employ a marker frequency and normalize the formula in terms of this and $\mathbf{A}_{\text{mid}}$. For this purpose any frequency may be used at which known results occur. From a mathematical standpoint the frequency at which X_c/R_{low} becomes unity is convenient. This will be used and designated as f_1; that is, f_1 is the frequency at which the number of ohms in C becomes equal to the number of ohms in the resistance part of the equivalent circuit of Fig. 7-18. Thus at the frequency f_1

$$\frac{1}{2\pi f_1 C} = R_{\text{low}} \tag{7-31}$$

which mathematically defines f_1. Solving,

$$f_1 = \frac{1}{2\pi R_{\text{low}} C} \tag{7-32}$$

Continuing the development of the formula for amplification,

$$\frac{\mathbf{A}_{\text{low}}}{\mathbf{A}_{\text{mid}}} = \frac{1}{1 - j(X_c/R_{\text{low}})} = \frac{1}{1 - j(1/2\pi fC) \times 2\pi f_1 C} = \frac{1}{1 - j(f_1/f)} \tag{7-33}$$

Formula (7-33) gives both the magnitude and the phase of the amplification at any low frequency, relative to the value at mid-frequency. Since the phase angle at mid-frequency is 180°, as indicated by the negative sign in formula (7-23), all angles for $\mathbf{A}_{\text{low}}$ lie in the third quadrant, assuming $\mathbf{V}_{1g}$ as the reference.

At any low-range frequency f, the drop in voltage gain below the mid-frequency value may be expressed as follows:

$$\text{db drop in voltage gain} = -20 \log \frac{A_{\text{low}}}{A_{\text{mid}}} = -20 \log \frac{1}{\sqrt{1 + (f_1/f)^2}}$$
$$= 10 \log \left[1 + \left(\frac{f_1}{f}\right)^2\right] \tag{7-34}$$

For N similar stages

$$db \text{ drop in voltage gain} = 10N \log \left[1 + \left(\frac{f_1}{f}\right)^2\right] \tag{7-35}$$

One may plot a graph of $\mathbf{A}_{\text{low}}$ in two advantageous ways: A_{low} and θ_A as functions of frequency, or A_{low} as a function of θ_A using polar coordinates. Figure 7-19 shows A_{low} as a function of f in seminormalized form. The scale of f is logarithmic, as this spreads out the curve for lower frequencies. When $f = f_1$, $\mathbf{A}_{\text{low}}/\mathbf{A}_{\text{mid}} = 1/(1 - j1) = 0.707\underline{/45°}$. At this frequency the amplification is down to about 71 per cent of the mid-frequency value, and the angle is 225°. The gain is down 3 db, and if power were important, the power would be reduced to $(0.707)^2 = 0.5$ of its mid-frequency value. Thus f_1 is sometimes called the lower half-power frequency, but

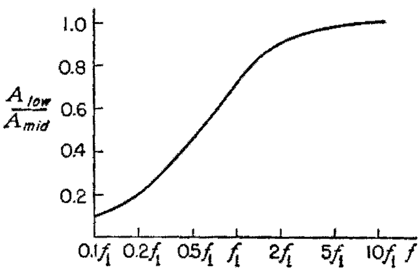


FIG. 7-19. Gain vs. frequency in the l-f range.

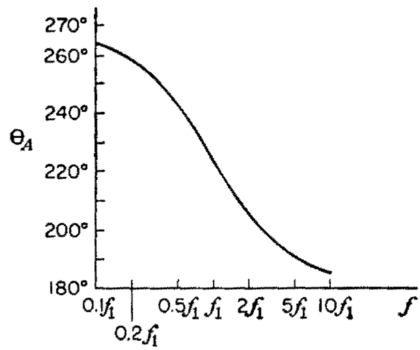


FIG. 7-20. θ_A vs. frequency in the l-f range.

the other relations are more important for a voltage amplifier. At $f = 5f_1$, $\mathbf{A}_{\text{low}}/\mathbf{A}_{\text{mid}} = 0.98\underline{/11.3°}$, or the amplification is down 2 per cent. At $f = 10f_1$, $\mathbf{A}_{\text{low}}/\mathbf{A}_{\text{mid}} = 0.995\underline{/5.7°}$, and the amplification is down one-half of 1 per cent. At $f = 0.1f_1$, $\mathbf{A}_{\text{low}}/\mathbf{A}_{\text{mid}} = 0.1\underline{/84.3°}$. Figure 7-20 shows how θ_A varies with frequency in this range. At $10f_1$ the angle is about 6° away from 180°, but the magnitude of the amplification is practically at its greatest value. Often $10f_1$ is considered the upper end of the l-f range and will be so taken in this book.

Examination of Eq. (7-33) shows that $\theta_A = 180° - \arctan(-f_1/f)$, where $\arctan(-f_1/f)$ is a fourth-quadrant angle. This equation can be solved to obtain

$$\frac{f_1}{f} = \tan \theta_A \tag{7-36}$$

The size of $\mathbf{A}_{\text{low}}/\mathbf{A}_{\text{mid}}$ obtained from Eq. (7-33) is

$$\frac{A_{\text{low}}}{A_{\text{mid}}} = \frac{1}{\sqrt{1 + (f_1/f)^2}} \tag{7-37}$$

Eliminating f_1/f between Eqs. (7-36) and (7-37) gives

$$\frac{A_{\text{low}}}{A_{\text{mid}}} = |\cos \theta_A| \tag{7-38}$$

This equation is plotted in polar coordinates with $A_{\text{low}}/A_{\text{mid}}$ as the radius vector and θ_A as the other variable, yielding the locus of Fig. 7-21. The student of analytic geometry should readily recognize that the locus is circular. Because of the restrictions that θ_A lies in only the third quadrant, only half of the mathematical locus is useful in this case.

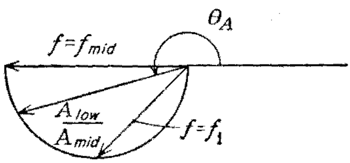


FIG. 7-21. Polar plot of $A_{\text{low}}/A_{\text{mid}}$ vs. θ_A.

Frequency was not used in plotting this locus, and so it is not obvious just what frequency places the relative-amplification radius vector in a given position. The position for f_1 is shown, it being determined from the known angle of 225°.

7-13. H-F Performance of R-C-coupled Amplifiers. The h-f equivalent circuit shown in Fig. 7-22 is obtained from that of Fig. 7-12 by omitting the coupling capacitor C. With proper interpretation, this is identical with the equivalent circuit of Fig. 7-3a for a single-stage ampli-

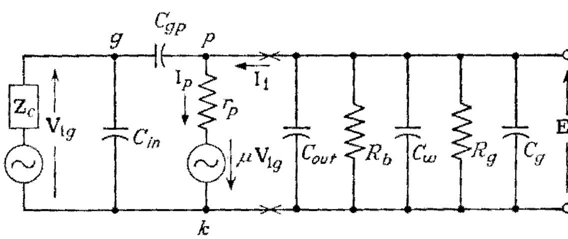


FIG. 7-22. Equivalent circuit for high frequencies.

fier, and hence Norton's form of Fig. 7-3c may be employed and is drawn in simplified form in Fig. 7-23. The resistor arrangement in Norton's form is the same as that in the equivalent circuit for mid-frequencies. Therefore we let

$$\frac{1}{R_{sh}} = \frac{1}{r_p} + \frac{1}{R_b} + \frac{1}{R_g} \tag{7-39}$$

Here

$$C_{sh} = C_{\text{out}} + C_w + C_{gp} + C_g \tag{7-40}$$

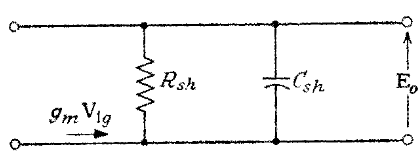


FIG. 7-23. Norton's equivalent circuit for high frequencies.

C_g being the input shunt capacitance of the next stage (see Arts. 7-3 and 7-5). It follows from the circuit of Fig. 7-23 that

$$\mathbf{Y}_{sh} = \frac{1}{R_{sh}} + \frac{j}{X_{c_{sh}}} \tag{7-41}$$

where $X_{c_{sh}}$ represents the reactance of C_{sh}. By formula (7-7)

$$\mathbf{A}_{\text{high}} = \frac{\mathbf{E}_o}{\mathbf{V}_{1g}} = -\frac{g_m}{\mathbf{Y}_{sh}} = -\frac{g_m}{(1/R_{sh}) + (j/X_{c_{sh}})} = \frac{-g_m R_{sh}}{1 + j(R_{sh}/X_{c_{sh}})}$$
$$= \frac{A_{\text{mid}}}{1 + j(R_{sh}/X_{c_{sh}})} \quad (7\text{-}42)$$

As a marker frequency f_2, one may conveniently use a value which makes $X_{C_{sh}} = R_{sh}$. Thus f_2 is defined by

$$R_{sh} = \frac{1}{2\pi f_2 C_{sh}} \quad (7\text{-}43)$$

Hence

$$f_2 = \frac{1}{2\pi R_{sh} C_{sh}} \quad (7\text{-}44)$$

Then

$$\frac{\mathbf{A}_{\text{high}}}{\mathbf{A}_{\text{mid}}} = \frac{1}{1 + j(2\pi f C_{sh}/2\pi f_2 C_{sh})} = \frac{1}{1 + j(f/f_2)} \quad (7\text{-}45)$$

This formula giving $\mathbf{A}_{\text{high}}$ is similar to that expressing $\mathbf{A}_{\text{low}}$, except that in this case any value of f renders θ_A in the second quadrant. Note that $f = nf_2$ yields the same value of A as does $f = f_1/n$, where n is any real number. Thus $A_{\text{high}}/A_{\text{mid}}$ for $f = 0.1f_2$, is the same as $A_{\text{low}}/A_{\text{mid}}$ for $f = 10f_1$. Likewise $f = f_2$ gives $\mathbf{A}_{f_2}/\mathbf{A}_{\text{mid}} = 0.707\underline{/135°}$, and so on.

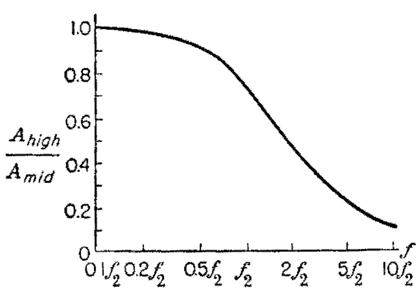


Fig. 7-24. Gain vs. frequency in the h-f range.

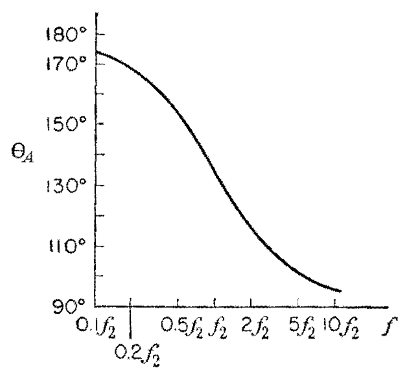


Fig. 7-25. θ_A vs. frequency in the h-f range.

Graphs similar to those for low frequencies are drawn in Figs. 7-24 to 7-26, the titles on each being sufficient identification.

Since $f = 0.1f_2$ gives $A = A_{\text{mid}}$ within one-half of 1 per cent, it is often considered the beginning of the h-f range. Thus the mid-frequency range extends from $10f_1$ to $0.1f_2$ provided $10f_1$ is less than $0.1f_2$, and in this range $\mathbf{A} = \mathbf{A}_{\text{mid}}$ with magnitude within one-half of 1 per cent (0.04 db) and angle within 6°.

7-14. The Analysis of an R-C-coupled Amplifier Circuit. The study just made of an R-C-coupled amplifier is sufficient to form a basis for the analysis of this device. The complete performance is shown in Fig. 7-27 as a composite of Figs. 7-19 and 7-24. It presumes that the performance of the amplifier can be broken up into three quite well defined ranges for which $10f_1$ and $0.1f_2$ have been arbitrarily chosen as dividing points. If $0.1f_2$ does not exceed $10f_1$, this presumption should be abandoned and another analysis made. The results of such an analysis can be written in a form which combines those we have obtained for high and low frequencies. This form is

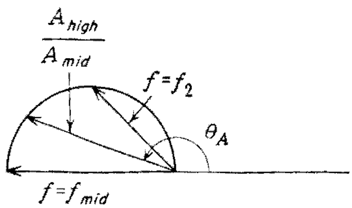


FIG. 7-26. Polar plot of $A_{\text{high}}/A_{\text{mid}}$ vs. θ_A for the h-f range.

$$\frac{\mathbf{A}}{\mathbf{A}_{\text{mid}}} = \frac{1}{1 - j(f_1/f)}\,\frac{1}{1 + j(f/f_2)} \tag{7-46}$$

In practice $0.1f_2$ usually exceeds $10f_1$, and hence the variation of gain with frequency indicated in Fig. 7-27 is a practical one. A composite drawing of Figs. 7-20 and 7-25 likewise would show the variation of θ_A to be from near 270° for very low frequencies to 180° somewhere in the

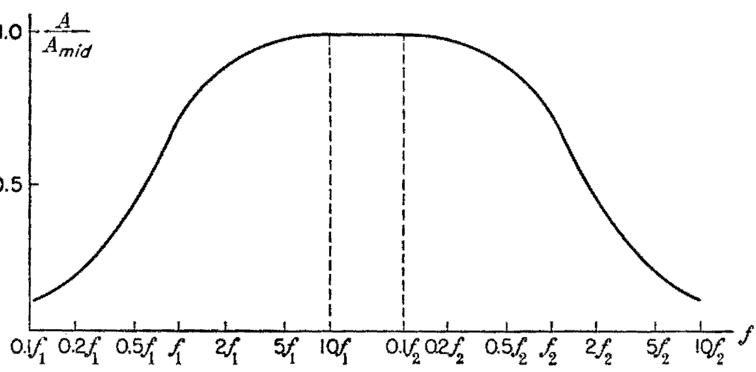


FIG. 7-27. Gain vs. frequency for an R-C-coupled amplifier.

middle-frequency range and approaching 90° as the frequency increases. Both magnitude and angle of $\mathbf{A}$ are shown in a circular locus which may be made for Figs. 7-21 and 7-26 combined.

As an example of the analysis of an R-C-coupled amplifier, assume the following: A 6J5 stage feeds another 6J5. $E_{bb} = 300$ volts, $R_k = 2440$ ohms, C_k is very large, $R_b = 100{,}000$ ohms, $R_g = 250{,}000$ ohms, $C = 0.01$ μf. From a tube manual $C_{gp} = 3.4$ $\mu\mu$f, $C_{pk} = 3.6$ $\mu\mu$f, and $C_{gk} = 3.4$ $\mu\mu$f. Suppose the mid-frequency gain of the second stage has already been determined as $15\underline{/180°}$. The measured value of C_w is 30 $\mu\mu$f.

1. On a sheet of i_b vs. e_b characteristics for a 6J5 draw a load line for $R_{dc} = 102{,}440$ ohms through $e_b = 300$ volts on the $i_b = 0$ axis. Also draw the bias line for $R_k = 2440$ ohms (see Art. 6-4). The intersection of these two lines gives Q: $E_{bo} = 112$ volts, $E_{co} = -1.8 \times 10^{-3} \times 2440 = -4.4$ volts. For voltage triodes the T point practically coincides with the Q point.

2. The values of μ and r_p can be determined by graphical construction on the plate characteristics, by using a vacuum-tube bridge, or by reference to a tube handbook where approximate values are given corresponding to various values of E_{bo} and of I_{bo} or of E_{co}. In this case the bridge method is used, yielding $\mu = 18$, $r_p = 16{,}000$ ohms.

3. Compute R_{sh}.

$$\frac{1}{r_p} + \frac{1}{R_b} + \frac{1}{R_g} = \frac{1}{0.016} + \frac{1}{0.1} + \frac{1}{0.25} = 62.5 + 10 + 4$$
$$= 76.5 \text{ micromhos}$$

$$R_{sh} = \frac{1}{76.5} = 0.0131 \text{ megohm} = 13{,}100 \text{ ohms}$$

4. Compute

$$R_{\text{low}} = R_g + \frac{1}{(1/R_b) + (1/r_p)} = 0.25 + \frac{1}{10 + 62.5}$$
$$= 0.25 + 0.0138 = 0.264 \text{ megohm}$$

5. Compute

$$A_{\text{mid}} = g_m R_{sh} = \frac{18}{16{,}000} \times 13{,}100 = 14.7$$

6. Compute

$$C_{sh} = C_{pk} + C_{gp} + C_w + C_g$$

where C_g for the next stage equals $C_{gk} + (1 - \mathbf{A})C_{gp} = 3.4 + (1 + 15)3.4 = 58\ \mu\mu\text{f}$ approximately. C_g varies in the h-f range, and 58 $\mu\mu$f represents its maximum value. $C_{sh} = 3.6 + 3.4 + 30 + 58 = 95\ \mu\mu\text{f}$.

7. Compute f_1.

$$f_1 = \frac{1}{2\pi R_{\text{low}} C} = \frac{1}{2\pi \times 264{,}000 \times 0.01 \times 10^{-6}} = 60 \text{ cps}$$

8. Compute f_2.

$$f_2 = \frac{1}{2\pi R_{sh} C_{sh}} = \frac{1}{2\pi \times 13{,}100 \times 95 \times 10^{-12}} = 128{,}000 \text{ cps}$$

9. Insert numerical values for f_1, f_2, and A_{mid} on the axes of Fig. 7-27 to obtain the proper response curve. *Note:* By assuming C_k to be very large the problem of feedback is avoided at this time. With a small value of C_k there would be an alternating-voltage drop across the R_k,C_k combina-

tion which would be applied to the grid in series with $\mathbf{V}_{1g}$. The result would be a lower value of amplification, particularly at low frequencies.

7-15. Design Considerations for an *R*-*C*-coupled Amplifier. The l-f performance is determined by f_1. Since $f_1 = 1/2\pi R_{\text{low}}C$, to make f_1 low in value, R_{low} must be made large. Therefore, a large value of C is desirable together with a large value of R_{low}. $R_{\text{low}} = R_g + [r_pR_b/(r_p + R_b)]$. For a triode, r_p is usually smaller than R_b. In the case of a medium-μ triode, r_p is less than 20,000 ohms; for a high-μ triode, r_p is usually less than 100,000 ohms. For a pentode, r_p may exceed 1 megohm, and hence R_{low} is approximately the value of R_g in series with R_b. R_g can be made much higher than R_b, but because of difficulties should there be grid emission from the following tube, it is generally limited to 2 megohms or less. Hence R_{low} is mostly dependent upon R_g, and its ordinary practical upper limit is about 2 megohms. C is limited in size by the leakage inherent to physically large capacitors, which affects the next tube's bias, and by the large capacitance to ground, which tends to spoil the h-f response. Furthermore, too large a product R_gC may result in blocking of the grid of the next tube. If a random pulse of voltage on the grid drives the grid positive, a flood of electrons charges C. Meanwhile, the causative pulse having receded, the grid is left highly negative, beyond cutoff, and the tube is inoperative until some of the charge on C has leaked off. If C's capacitance is high and R_g is high also, it takes considerable time for C to discharge. Various experimenters have formulated practical limits for the product R_gC, their values ranging from 0.1 to 0.01. If 0.01 is tentatively tried with $R_g = 2$ megohms, the maximum value of C will be 0.005 μf and $f_1 = 16$ cps is the lowest conservative value obtainable for a broad-band amplifier with simple R-C coupling.

In the h-f range the performance is determined by $f_2 = 1/2\pi R_{sh}C_{sh}$. R_{sh}, being the parallel combination of r_p, R_b, and R_g, is hence less than the smallest of these. For a medium-μ triode, r_p limits R_{sh}; for a high-μ triode, R_b is also an important factor. For a pentode, R_b usually limits R_{sh}. Since $C_{sh} = C_{\text{out}} + C_w + C_{gp} + C_g$, it is helpful, in attempting to raise the value of f_2, to keep these values small. By careful wiring and by the use of short connections and small components held away from the chassis, C_w can be made small. C_g depends upon the gain of the next stage and is quite large for triodes. For pentodes, C_g is approximately C_{in}; it is small and is almost independent of the stage gain. Hence, it is usually advantageous to feed into pentodes as a second stage if a very high value of f_2 is desired. Ordinarily nothing can be done to reduce $C_{\text{out}} + C_{gp}$ appreciably beyond carefully choosing a tube, unless one changes the type of amplifier from that being studied at present. In the case of our circuit the lower limit on C_{sh} is $C_{\text{out}} + C_{gp}$. To make f_2 high a small value of $R_{sh}C_{sh}$ is required. Since reducing R_{sh} also reduces the mid-frequency gain $g_m R_{sh}$, it is seen that very great bandwidth comes at the expense of lower stage gain.

Consultation with a tube manual shows that, although the values of μ and r_p vary widely with various voltage-amplifier types of triodes and pentodes, the values of g_m do not vary so much, usually being between 1500 and 3000 micromhos, except for television types. Thus the mid-frequency gain is mostly determined by the size of R_{sh}. Low-r_p tubes will have low gain, and the higher-r_p tubes will have higher gain unless R_b is made small. μ, which equals $g_m r_p$, is therefore an important factor when considering high-gain voltage amplifiers.

7-16. The Design of an *R-C*-coupled Amplifier. Although the process employed in analyzing an amplifier as in Art. 7-14 can be reversed in a fashion to design an amplifier, it is usually not desirable to do so. Instead, a tube manual is often employed. In the manual a special *R-C*-coupled amplifier chart has been prepared, which tabulates many circuit combinations and their performance figures. Minor changes in characteristics can be effected by slight changes in the circuit components. The procedure followed depends upon the tube manual used.

As an example, suppose an amplifier is desired which will have an output of 62 volts peak with an input of 10 mv peak. The output is to be constant within 3 db from 100 to 100,000 cps. The value of E_{bb} available is 250 volts. We shall use the charts contained in Appendix B.

1. The mid-frequency gain is $62/0.01 = 6200$, and the required output voltage is 62 volts peak, or $62/\sqrt{2} = 44$ volts rms.

2. If we wish to use a type 6J5 medium-μ triode, the available designs show that it is possible to get as high as 53 volts rms output with $E_{bb} = 250$ volts and with the proper circuit values. The distortion in the output stage alone is 4.6 per cent with a grid drive of 3.3 volts rms. It is of interest to note in passing that the distortion would be only 1.3 per cent if the drive were reduced to 1.0 volt. This indicates that, for a triode with a given load, the percentage distortion is roughly proportional to the amount of grid drive, as may be verified by reference to Eq. (6-57). The gain is given as 16.1. With this amplification, four stages would be needed to achieve an over-all gain as high as 6200.

3. If a high-μ triode such as a 6SF5 is tried, there is no design listed which will deliver 44 volts rms at the output with $E_{bb} = 250$ volts. With a gain of around 70 per stage, three stages would be needed.

4. A typical pentode 6SJ7 may be used to deliver as high as 54 volts with 5 per cent nonlinear distortion, using $E_{bb} = 250$ volts. The stage gain listed is 108, which makes two stages necessary. The value of R_b is 0.1 megohm, which is the lowest listed, and hence gives the best h-f response. This tube and circuit arrangement will be tentatively chosen for this example.

5. The lower marker frequency f_1 may be determined by using Eq. (7-35): $10 \times 2 \log [1 + (f_1/100)^2] = 3$, or $f_1 = 64$ cps.

6. Likewise f_2 may be determined with Eq. (7-45) as a guide: $10 \times 2 \log [1 + (100{,}000/f_2)^2] = 3$, or $f_2 = 156{,}000$ cps.

7. From the R-C-coupled amplifier data for the 6SJ7 tube in Appendix B the following values are obtained: $R_b = 0.1$ megohm, $R_g = 0.47$ megohm, $R_d = 0.39$ megohm, $R_k = 560$ ohms, $A = 108$ for heavy drive, 115 for light drive (0.1 volt rms).

8. The coupling capacitance C may be calculated by the use of formula (7-32), $f_1 = 1/2\pi R_{\text{low}}C$, where $R_{\text{low}} \approx R_g + R_b = 0.57$ megohm. Hence $C = 1/(2\pi \times 64 \times 0.57 \times 10^6) = 0.0044\ \mu$f.

9. The theory of the determination of C_k is discussed later under the subject of feedback (Art. 10-11). The value of C_k needed to limit the degenerative loss caused by R_k to 1 per cent at 64 cps is 46 μf. This result is obtained by using formula (10-41) of Art. 10-11 (see also Prob. 12, Chap. 10). A rule of thumb commonly used makes $X_{C_k} = 0.1R_k$ at the l-f end of the mid-frequency band. This rule makes C_k in this case only about one-tenth the size stated above, and hence the degenerative effect of R_k would be greater.

10. Since C_d does not perfectly bypass R_d, there is an alternating voltage applied to the screen grid, which causes some loss in gain. The theory of this action is beyond the scope of this book. One ordinary rule for choice of C_d is to make $X_{C_d} = 0.1R_d$. This rule is without rigorous foundation, but in ordinary amplifier circuits the results of its use are often satisfactory.

11. The h-f response of the amplifier is not stated in the chart. This is because some of the capacitances involved cannot be known until the amplifier is assembled and tested. If $f_2 = 156{,}000$ cps is the required value, then from formula (7-44) we may solve for the necessary value of C_{sh}. Assume r_p is approximately 1 megohm. $1/R_{sh} = 1/r_p + 1/R_b + 1/R_g = 1/1 + 1/0.1 + 1/0.47 \approx 1 + 10 + 2 = 13$ micromhos, and hence $R_{sh} = 77{,}000$ ohms. Therefore $C_{sh} = 1/2\pi f_2 R_{sh} = 1/(2\pi 156{,}000 \times 77{,}000) = 13.2\ \mu\mu$f. Since for a 6SJ7, $C_{\text{out}} = 7\ \mu\mu$f and $C_{\text{in}} = 6\ \mu\mu$f, there remains no allowance for wiring capacitance. Hence this design fails to meet the requirements of h-f response. Allowing 10 $\mu\mu$f as an approximate actual minimum wiring capacitance, $C_{sh} = 23\ \mu\mu$f, and f_2 for one stage has a value of 90,000 cps. Since the over-all gain is $115 \times 108 = 12{,}400$, which exceeds the gain requirement, it is possible to lower R_b in each stage and thus improve the h-f response. Other charts in tube manuals may be consulted and perhaps designs found which satisfy the frequency requirement. Or the 6SJ7 tube may be abandoned and the 6J5 or another tube tried.

7-17. Video-frequency Amplifiers.[3,4] We have seen that there is no great difficulty in designing R-C-coupled amplifiers to give excellent performance over the a-f range of 40 cps to 20 kc and even higher. The amplitude-frequency response can be made essentially flat, and for sound reproduction the small phase shift which occurs is undetected in our hearing process.

Where amplifiers are to be used in the production of visual images on the screen of a cathode-ray oscilloscope, far better performance is often required. This can be seen if we study the ordinary way in which a television image is formed on a kinescope screen. Figure 7-28 shows the path of the cathode-ray beam on the face during the production of one frame, or complete picture. Driven by the horizontal-deflection system using sawtooth voltages of high frequency and by the vertical-deflection system using sawtooth voltages of a much lower frequency, it describes slightly slanting lines as shown. By interlacing lines (as shown starting with the line numbered 8½) it is possible to operate with lower frame frequency and obtain a satisfactory picture. The actual number of lines in a frame greatly exceeds the 15 shown, the standard of the present time being 525 with a frame frequency of 30 per second. This makes the time for one line $1/(525 \times 30) = 63.5\ \mu$sec. A standard picture pattern has a width-to-height ratio of $4/3$, and hence there are $1.33 \times 525 = 700$ intervals on any line, with interline spacing. The sweep time for one of these intervals is $63.5/700 = 0.09\ \mu$sec. Hence, for excellent picture sharpness the amplifiers controlling the beam intensity must be able to change the spot from full to zero brilliance in approximately $0.1\ \mu$sec. Since synchronizing pulses for the vertical sweep occur 30 times each second, the amplifier must be able to handle this frequency. Practical television receivers contain video-frequency amplifiers which perform over a frequency range of 30 cps to about 4 Mc, the picture detail being somewhat inferior to that discussed above but nevertheless quite satisfactory.

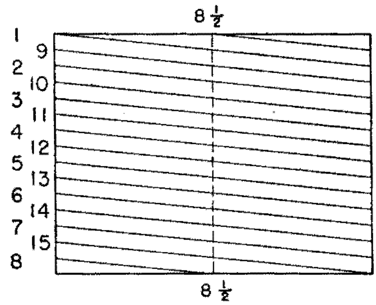


FIG. 7-28. Path of the beam on the face of a kinescope when scanning one frame of a television picture.

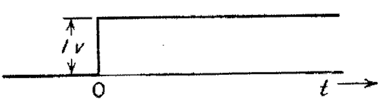


FIG. 7-29. A unit-step function.

Let us consider the effect of applying to an amplifier a signal voltage which possesses the property of fast rise and long duration. One such waveform is shown in Fig. 7-29. This is called a unit step voltage. We shall apply this voltage to the grid of an R-C-coupled amplifier and determine the character of the output voltage. Since the waveform is not repetitive, one cannot make a Fourier-series analysis of it, and hence the analyses previously made using sinusoidal waveforms are not directly applicable.

Figure 7-30a shows the schematic diagram including the switch, which is closed at time $t = 0$. Following this time an increased plate current flows through the tube. We can see better by looking at Fig. 7-30b what then occurs. At $t = 0$ the capacitors C_{out} and C_w have their quiescent charges, and some time will elapse before their voltages can fall to the new value. C meanwhile has only its quiescent charge, which can change only slowly because of R_g, and hence e_o decreases from zero practically at the same rate that the voltage across C_{out} and C_w falls. Thus for the events immediately following $t = 0$, the coupling capacitor is inactive, and the parallel combination C_{sh} of C_{out}, C_w, and C_g delays the fall of e_o. See now Fig. 7-30c, which shows Norton's equivalent circuit in simplified form. We can write for the time following $t = 0$,

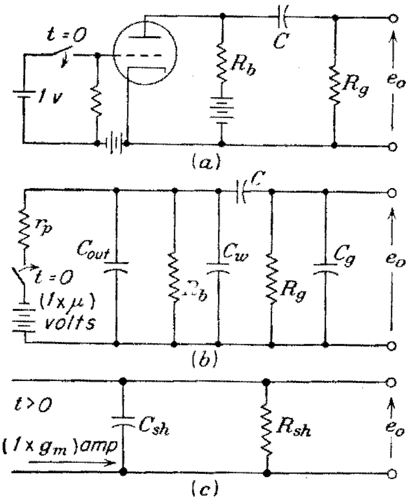


FIG. 7-30.

$$-1g_m = \frac{e_o}{R_{sh}} + C_{sh}\frac{de_o}{dt} \tag{7-47}$$

Separating variables,

$$dt = \frac{-R_{sh}C_{sh}de_o}{g_mR_{sh} + e_o} \tag{7-48}$$

Integrating,

$$t = -R_{sh}C_{sh} \ln (g_mR_{sh} + e_o) + C_1 \tag{7-49}$$

When $t = 0$, $e_o = 0$ and hence $C_1 = +R_{sh}C_{sh} \ln g_mR_{sh}$. Therefore

$$t = -R_{sh}C_{sh} \ln \left(1 + \frac{e_o}{g_mR_{sh}}\right) \tag{7-50}$$

Solving for e_o, we obtain

$$e_o = -g_mR_{sh}(1 - \epsilon^{-\frac{t}{R_{sh}C_{sh}}}) \tag{7-51}$$

A normalized graph of e_o as a function of t, given by this equation, is shown in Fig. 7-31. The time for e_o to attain 63 per cent of its final value is $t = R_{sh}C_{sh}$ sec. Thus $1/R_{sh}C_{sh}$ is a measure of the ability of the circuit to deliver a steep wave front of voltage. Note that this measure is the same as ω_2, obtained by the assumption of a sinusoidal grid voltage. Thus a high value of ω_2 for sinusoidal voltages implies a short rise time for step voltages.

After a short period of time the voltage e_o practically attains its final value $-g_mR_{sh}$ provided the coupling capacitor C did not change its charge. Since the latter occurrence usually is much slower than the events just

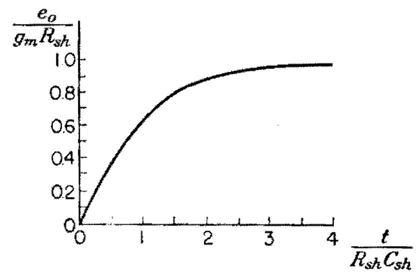


FIG. 7-31. The rise in output voltage with time in the circuit of Fig. 7-30.

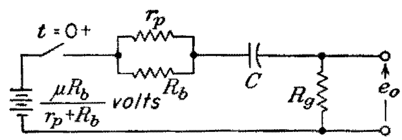


FIG. 7-32.

discussed, we can neglect C_{sh} and omit it from the circuit of Fig. 7-32 while we study the persistence of the value $-g_mR_{sh}$. This circuit should have a familiar appearance. It is not difficult to show that e_o in this case is given by

$$e_o = -g_mR_{sh}\epsilon^{-t/R_{\text{low}}C} \tag{7-52}$$

where R_{low} is defined in Eq. (7-25) (see Prob. 38, this chapter).

A normalized graph of e_o is shown in Fig. 7-33. From it we see that when $t = R_{\text{low}}C$ the voltage has reduced to 37 per cent of its greatest value. Thus $1/R_{\text{low}}C$ is a measure of the ability of the circuit to maintain the voltage at the value $-g_mR_{sh}$. Note that this measure is the l-f marker ω_1 obtained in an earlier study.

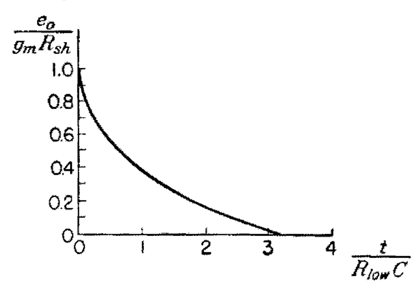


FIG. 7-33. The fall in output voltage with time in the circuit of Fig. 7-30.

If $1/R_{sh}C_{sh}$ can be made high and $1/R_{\text{low}}C$ can be made small, the output voltage can rise rapidly and fall only slowly from its top value. We could just as well state that the amplifier should have a bandwidth extending from very low frequencies to very high frequencies. Figure 7-34 shows a graph of e_o in which both rapid rise and long persistence are attained.

The circuit element which limits the h-f response is the shunt capacitance in the plate circuit. By wise choice of tubes and by careful wiring it is possible to reduce C_{sh} considerably below the value in the usual a-f amplifier. A criterion for this tube choice may be developed as follows: Over the range of constant amplification

$$A = g_m R_{sh} \tag{7-23}$$

If the upper limit of usefulness is taken as f_2, having particularly in mind the possibility of improving by some means the performance between $0.1f_2$ and f_2, then the maximum usable bandwidth is practically f_2, as given in Eq. (7-44). The product of gain and bandwidth is

$$Af_2 = \frac{g_m}{2\pi C_{sh}} \quad \text{cps} \tag{7-53}$$

and is called the gain area, it representing approximately the area under the usable portion of the response curve. It shows that a good tube for a video-frequency amplifier has a high transconductance and low input and output capacitances.

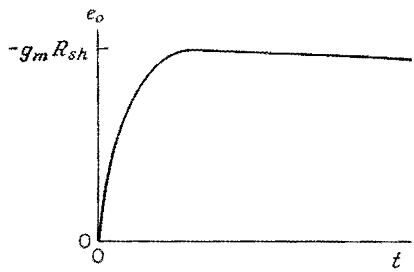


FIG. 7-34. Rise and fall of output voltage with time.

The gain-area criteria for several tubes are tabulated in Table 7-1, it being assumed that two like tubes are used in cascade, the criterion applying to the first tube. For a triode, C_g depends upon the gain, and this has been assumed to be $10\underline{/180°}$.

TABLE 7-1
GAIN-AREA CRITERIA

Tube	g_m, μmho	C_{out}, μμf	C_g, μμf	Gain area, Mc
6AC7	9000	5	11	90
6AK5	5100	2.8	4.2	116
6J5	2600	3.6	41	9
6SG7	4000	7	8.5	41
6SJ7	1650	7	6	20

The best tube for wide-band voltage-amplifier use, as listed in Table 7-1, is the 6AK5. This is due to its relatively high transconductance and low interelectrode capacitances. If a bandwidth of uniform response of 4 Mc is desired with an ordinary R-C circuit, the value of f_2 should be made 40 Mc and a gain of $^{116}\!/_{40} = 2.9$ per stage is indicated. The results will be somewhat inferior to this as the stray capacitances of wiring will increase the C_{sh} value above that due to the tabulated tube values.

It would be very desirable if the useful band could be made to extend farther than $0.1f_2$, and then f_2 could be lowered and more gain per stage obtained.

7-18. H-F Compensation. Various plans have been tried for raising the gain and controlling the phase shift of an *R-C*-coupled amplifier in the h-f range. Two of the commonest circuits are indicated in Fig. 7-35, where only the tube loads are shown. It is not difficult to show that either plan gives some promise of remedy.

In Fig. 7-35*a*, as the frequency rises, the action of L is to increase the current passing through the parallel combination of R_b, R_g, and C_g above the value obtaining without L. This of course will increase the output voltage, and if the value of L is correctly chosen, it is probable that this action can compensate for the drop in voltage due to C_g and also C_{out} and the stray capacitances. A careful mathematical analysis is needed to determine the details.

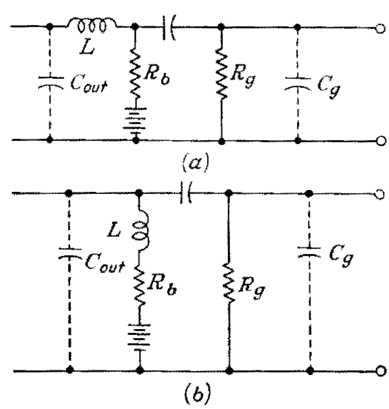


FIG. 7-35. High-frequency compensation methods; (*a*) using a series inductor, (*b*) using a shunt inductor.

In the circuit of Fig. 7-35*b*, the load on the tube increases as the frequency rises and hence the gain increases—we hope sufficiently to offset the effect of C_{sh}. Let us concentrate on the details of this method. Since to achieve great bandwidth the gain per stage will be small, a low value of R_b must be used. It will be much lower than either r_p or R_g. Hence

$$R_{sh} \approx R_b \tag{7-54}$$

Norton's equivalent circuit is drawn in Fig. 7-36. Then

$$\mathbf{E}_o = -\frac{(R_b + j\omega L)(1/j\omega C_{sh})}{R_b + j[\omega L - (1/\omega C_{sh})]} g_m \mathbf{V}_g \tag{7-55}$$

and

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{V}_g} = -g_m \frac{(L/C_{sh}) - j(R_b/\omega C_{sh})}{R_b + j[\omega L - (1/\omega C_{sh})]} \tag{7-56}$$

Making the denominator real, we obtain

$$\mathbf{A} = -g_m \frac{\dfrac{R_b}{\omega^2 C_{sh}^2} - j\left[\dfrac{\omega L^2}{C_{sh}} - \dfrac{L}{\omega C_{sh}^2} + \dfrac{R_b^2}{\omega C_{sh}}\right]}{R_b^2 + [\omega L - (1/\omega C_{sh})]^2} \tag{7-57}$$

It is desirable to graph A vs. f and θ_A vs. f, and some substitutions will help make this easier to do. Let f_2 be the upper half-power frequency of the uncompensated amplifier, defined by Eq. (7-43), which in this case becomes

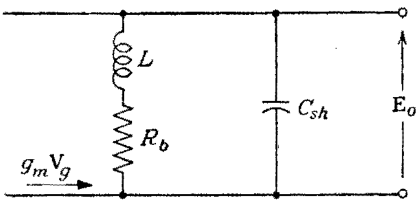


FIG. 7-36. Equivalent circuit for high frequencies.

$$R_b = \frac{1}{\omega_2 C_{sh}} \tag{7-58}$$

and hence

$$f_2 = \frac{1}{2\pi R_b C_{sh}} \tag{7-59}$$

Also, at this frequency f_2, let Q_2 be defined as

$$Q_2 = \frac{\omega_2 L}{R_b} \tag{7-60}$$

We can rewrite Eq. (7-57) in the form

$$\mathbf{A} = -g_m \frac{\dfrac{\omega_2{}^2 R_b}{\omega^2\omega_2{}^2 C_{sh}{}^2} - jR_b{}^3\left(\dfrac{\omega\omega_2{}^2L^2}{R_b{}^3\omega_2{}^2C_{sh}} - \dfrac{\omega_2{}^2L}{R_b{}^3\omega\omega_2{}^2C_{sh}{}^2} + \dfrac{\omega_2 R_b{}^2}{R_b{}^3\omega\omega_2 C_{sh}}\right)}{R_b{}^2 + \left(\dfrac{\omega\omega_2 L}{\omega_2} - \dfrac{\omega_2}{\omega}\dfrac{1}{\omega_2 C_{sh}}\right)^2} \tag{7-61}$$

or by Eqs. (7-58) and (7-60),

$$\mathbf{A} = -g_m \frac{\left(\dfrac{\omega_2}{\omega}\right)^2 R_b{}^3 - jR_b{}^3\left(\dfrac{\omega}{\omega_2}Q_2{}^2 - \dfrac{\omega_2}{\omega}Q_2 + \dfrac{\omega_2}{\omega}\right)}{R_b{}^2\left[1 + \left(\dfrac{\omega}{\omega_2}Q_2 - \dfrac{\omega_2}{\omega}\right)^2\right]} \tag{7-62}$$

Multiply numerator and denominator by $(\omega/\omega_2)^2$, and simplify.

$$\mathbf{A} = -g_m R_b \frac{1 - j\left(\dfrac{f}{f_2}\right)^2\left(\dfrac{f}{f_2}Q_2{}^2 - \dfrac{f_2}{f}Q_2 + \dfrac{f_2}{f}\right)}{\left(\dfrac{f}{f_2}\right)^2 + \left[\left(\dfrac{f}{f_2}\right)^2 Q_2 - 1\right]^2} \tag{7-63}$$

Therefore the simplified formula for the magnitude of $\mathbf{A}$ is

$$A = g_m R_b \frac{\sqrt{1 + \left(\dfrac{f}{f_2}\right)^2\left[\left(\dfrac{f}{f_2}\right)^2 Q_2{}^2 + 1 - Q_2\right]^2}}{\left(\dfrac{f}{f_2}\right)^2 + \left[\left(\dfrac{f}{f_2}\right)^2 Q_2 - 1\right]^2} \tag{7-64}$$

and that for θ_A is

$$\theta_A = 180° - \arctan \frac{f}{f_2}\left[\left(\frac{f}{f_2}\right)^2 Q_2{}^2 + 1 - Q_2\right] \tag{7-65}$$

With Q_2 as a parameter, several normalized curves of A/g_mR_b vs. f/f_2 have been drawn in Fig. 7-37.

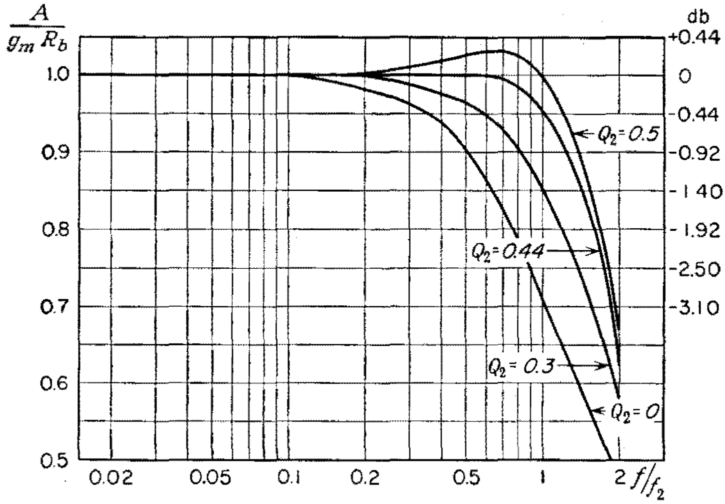


FIG. 7-37. Normalized curves of amplification vs. frequency.

Figure 7-38 illustrates in a simple manner how the waveshape is preserved if the phase shift is proportional to frequency. An amplifier should give either a phase shift of zero or a multiple of π for all frequencies in its useful band, or it should produce a shift angle proportional to the frequency, that is, $\phi = k\omega$. The time taken for a signal voltage to complete θ_A radians of its cycle is called the time delay between input and output. Thus the time delay in this case is

$$\frac{\phi}{2\pi}T = \frac{\phi}{2\pi f} = \frac{\phi}{\omega} = k \qquad (7\text{-}66)$$

where T is the period of one cycle. The statement that the time delay is a constant independent of frequency is thus equivalent to saying that the phase shift is proportional to frequency.

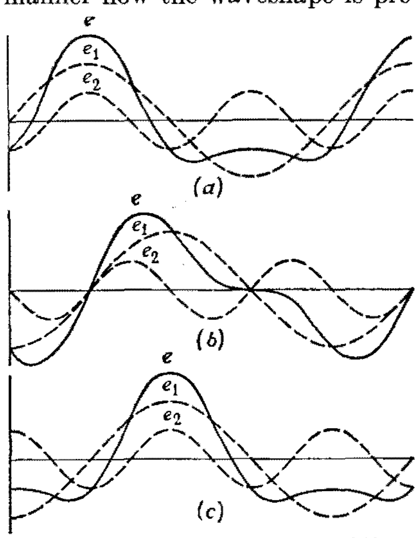


FIG. 7-38. The effect of phase shift on waveform. (a) Input voltage and its components. (b) Output voltage and its components. Each component shifted a constant angle of 90°. (c) Output voltage and its components. e_1 shifted 90°; e_2 shifted 2 × 90°.

In Fig. 7-39 several curves of phase shift vs. frequency have been drawn as $f_2\phi/f$ vs. f/f_2. The phase shift ϕ is f/f_2 times the ordinate. The time delay is $1/\omega_2$ times the ordinate.

We are now ready to judge the value of Q_2 desired. Examination of the normalized gain curves of Fig. 7-37 shows that if $Q_2 = 0.5$, the gain at $f = f_2$ is equal to that in the middle-frequency range. At about $f = 0.7f_2$ the gain rises to a maximum above this value by only about 0.25 db, which

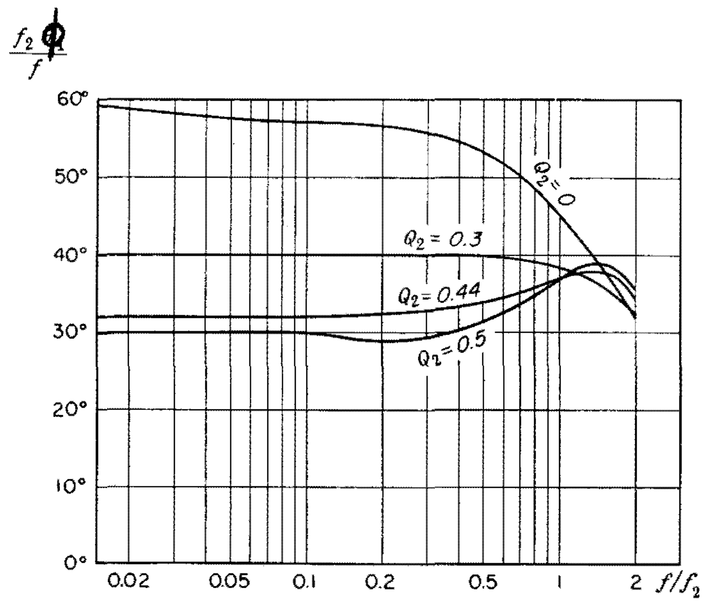


FIG. 7-39. Curves of phase shift vs. frequency.

is not usually objectionable. For $Q_2 = 0.44$ the response curve is flat over a greater range but is down 0.5 db at f_2. The value $Q_2 = 0.5$ is usually chosen when only amplitude of gain is important. The curves of Fig. 7-39 show that $Q_2 = 0.5$ has constant time delay only up to $0.1f_2$. On the other hand $Q_2 = 0.3$ gives a very good phase-shift characteristic up to $0.5f_2$ or higher. A value of Q_2 near 0.3 is ordinarily used when phase shift is the item of principal interest. If both amplitude and phase shift are important, $Q_2 = 0.44$ is a compromise value. Then the value of f_2 must be raised, and the subsequent loss in gain per stage accepted. More stages may be needed, of course, to offset this loss.

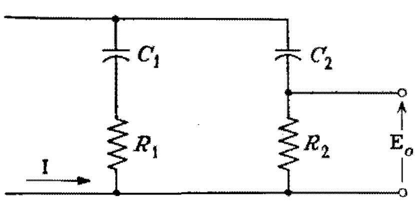


FIG. 7-40. $\mathbf{E}_o$ is independent of the frequency of $\mathbf{I}$ if $R_1C_1 = R_2C_2$.

7-19. L-F Compensation. In the l-f range R-C-coupled amplifiers suffer loss in gain, and an accompanying phase shift occurs, because of the impedance of the coupling capacitor and also because the bypass capacitors C_k and C_d (see Figs. 7-10 and 7-11) do not function as well as for higher frequencies. For example, as the frequency falls, C_k's mounting impedance allows more voltage drop across C_k and R_k. This voltage is in such phase that it cancels part of the signal in the grid circuit and makes the resultant gain of the stage less. Circuits may be devised to compensate for any of these effects. The one we shall study will be that involving the coupling capacitor.

Examination of the circuit of Fig. 7-40 shows that

$$\frac{\mathbf{E}_o}{\mathbf{I}} = -\frac{\left(R_1 + \dfrac{1}{j\omega C_1}\right) R_2}{R_1 + \dfrac{1}{j\omega C_1} + R_2 + \dfrac{1}{j\omega C_2}} = -\frac{\dfrac{j\omega R_1 C_1 + 1}{j\omega C_1} R_2}{\dfrac{j\omega R_1 C_1 + 1}{j\omega C_1} + \dfrac{j\omega R_2 C_2 + 1}{j\omega C_2}} \tag{7-67}$$

If

$$R_1 C_1 = R_2 C_2 \tag{7-68}$$

then

$$\frac{\mathbf{E}_o}{\mathbf{I}} = -\frac{j\omega R_2 C_2}{j\omega(C_1 + C_2)} = -\frac{R_2 C_2}{C_1 + C_2} = -\frac{R_1 R_2}{R_1 + R_2} \tag{7-69}$$

which is independent of frequency.

To adapt this principle to the problem at hand, one may propose the circuit of Fig. 7-41. The tube may be either a triode or a pentode, and a

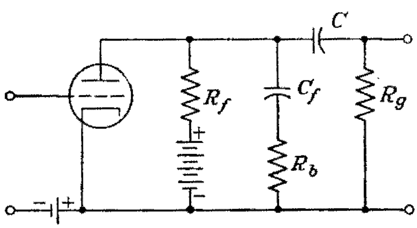


FIG. 7-41. One practical adaptation of the idea of Fig. 7-40, using shunt feed.

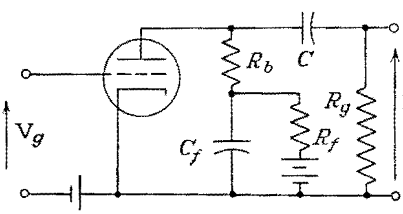


FIG. 7-42. A circuit using series feed.

fixed bias is shown because we are compensating for the effect of C only, although cathode bias is more practical to use. R_f is added to the circuit to provide a d-c path. Its presence complicates matters somewhat, and it is impossible to obtain **A** entirely independent of frequency. The most important objection to the circuit is the fact that R_f shunts R_L (R_b in parallel with R_g) and the gain in the middle-frequency range is reduced. Since there is only a small gain at best, this is a serious objection.

Figure 7-42 shows a series-feed adaptation of the same idea. In the middle-frequency range the load on the tube is R_L, and hence the gain is higher than in the previous case. R_f and C_f really form a decoupling unit, often used when several stages of an amplifier are fed from a common E_{bb} supply. A decoupler filters out the variable components of voltage caused by other stages and prevents feedback.

An analysis of the action of the series form of circuit of Fig. 7-42 is quite easy, and the results are interesting. First we may draw Norton's equivalent circuit in Fig. 7-43. Since r_p is much larger than R_b and R_f in series, it has been omitted. Considering that R_g is also much larger than R_b and R_f in series, most of the current $g_m V_g$ will pass up the left branch. Without serious error, therefore, we may write

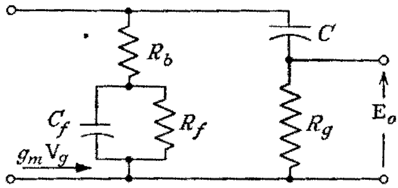


FIG. 7-43. Norton's equivalent of the circuit of Fig. 7-42.

$$\mathbf{E}_o = -g_m\mathbf{V}_g\left[R_b + \frac{R_f/j\omega C_f}{R_f + (1/j\omega C_f)}\right]\frac{R_g}{R_g + (1/j\omega C)}$$

$$= -g_m\mathbf{V}_g R_b\left[1 + \frac{R_f/j\omega C_f R_b}{R_f + (1/j\omega C_f)}\right]\frac{1}{1 + (1/j\omega C R_g)} \tag{7-70}$$

Hence

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{V}_g} = -g_m R_b \frac{R_f + \dfrac{1}{j\omega C_f} + \dfrac{R_f}{j\omega C_f R_b}}{R_f + (1/j\omega C_f)}\frac{1}{1 + (1/j\omega C R_g)} \tag{7-71}$$

or

$$\mathbf{A} = \mathbf{A}_{\text{mid}}\frac{1 + [(R_f + R_b)/j\omega C_f R_f R_b]}{[1 + (1/j\omega R_f C_f)][1 + (1/j\omega C R_g)]}$$

$$= \mathbf{A}_{\text{mid}}\frac{1 + \dfrac{1}{j\omega C_f[R_b R_f/(R_b + R_f)]}}{[1 + (1/j\omega R_f C_f)][1 + (1/j\omega C R_g)]} \tag{7-72}$$

We may let f_1 be the frequency at which $1/\omega_1 C = R_g$, or

$$\omega_1 = \frac{1}{R_g C} \tag{7-73}$$

which is almost identical with Eq. (7-32) if we note that R_{low} is practically R_g in this case. f_1 is the lower half-power frequency of the amplifier without the $R_f C_f$ compensation. It is interesting to observe that substitutions of this sort are often helpful whenever a term of the form ωRC occurs. For example, we may also let f_3 and f_4, respectively, be the frequencies at which

$$\omega_3 = \frac{1}{R_f C_f} \tag{7-74}$$

and

$$\omega_4 = \frac{1}{[R_b R_f/(R_b + R_f)]C_f} \tag{7-75}$$

Making these substitutions in Eq. (7-72) and simplifying,

$$\mathbf{A} = \mathbf{A}_{\text{mid}} \frac{1 - j(f_4/f)}{[1 - j(f_3/f)][1 - j(f_1/f)]} \tag{7-76}$$

Now f_1 is determined if the values of C and R_g have already been chosen, (see the example of Art. 7-20). R_b too has been determined if the upper half-power frequency has been specified. f_3 and f_4 hence depend upon R_f and C_f, both of which are arbitrary. Since f_4 is arbitrary, one can choose for its value that of f_1. This is desirable because it makes the result so simple. The gain now reduces to

$$\mathbf{A} = \mathbf{A}_{\text{mid}} \frac{1}{1 - j(f_3/f)} \tag{7-77}$$

which shows that the lower half-power frequency has been moved from f_1 to f_3 by this compensation method. As f_3 is arbitrary, this makes it possible to make the gain constant and the phase shift negligible down to any desired frequency value, $10f_3$.

The design equations are as follows: Since $f_4 = f_1$,

$$\frac{R_b R_f}{R_b + R_f} C_f = R_g C \tag{7-78}$$

Since f_3 is arbitrary,

$$f_3 = \frac{1}{2\pi R_f C_f} \tag{7-79}$$

it may be any desired value. If R_g, C, R_b, and f_3 are known, these equations are sufficient to determine R_f and C_f.

By making $\omega_4 > \omega_1$ it is possible to have a rising curve of gain as the frequency becomes lower. This idea is sometimes used to give an amplifier so-called bass boost.

7-20. An Example of Video-frequency Amplifier Compensation. An amplifier uses 6AK5 tubes with quiescent operation at $E_{bo} = 180$ volts, $E_{c2o} = 120$ volts, $E_{c1o} = -2$ volts, $I_{bo} = 7.7$ ma, $I_{c2o} = 2.4$ ma. Under these conditions $g_m = 5100$ micromhos, and $r_p = 0.7$ megohm. Suppose both uniform gain and phase shift proportional to frequency are important over a bandwidth from 20 cps to 5 Mc. Limit the h-f attenuation per stage to 0.33 db. Assume the stray capacitance can be reduced to 7 $\mu\mu$f.

Since both gain and phase shift are important, it is well to use $Q_2 = 0.44$. If the upper end of the useful band is taken at $f/f_2 = 0.6$, then $f_2 = 5/0.6 = 8.3$ Mc. Then the curve of Fig. 7-37 shows no deviation in uniformity of gain, and that of Fig. 7-39 shows a maximum angular deviation given by $(f_2/f)\theta_A = 34.5°$, or $\theta_A/0.6 = 34.5°$, $\theta_A = 20.7°$

per stage. At one-tenth this frequency, *i.e.*, at $f/f_2 = 0.06$, $\theta_A/0.06 = 32°$, or $\theta_A = 1.92°$ per stage. Since $10 \times 1.92° = 19.2°$ and not 20.7°, we see that the phase shift is not exactly proportional to frequency. These angles are of course in addition to the usual 180° phase shift per stage. We shall assume this departure from the ideal to be acceptable. If it were not for increased phase shift, it would be allowable to work to a higher value of f/f_2. If $Q_2 = 0.3$ is used, the phase-shift characteristic would be improved but we would be forced to use $f/f_2 = 0.5$ as the upper limit on account of the limit 0.33 db per stage attenuation of gain. Hence $Q_2 = 0.44$ seems to be the better choice.

From Table 7-1, $C_{out} = 2.8\ \mu\mu f$, $C_g = 4.2\ \mu\mu f$. Hence $C_{sh} = 2.8 + 4.2 + 7 = 14\ \mu\mu f$. $\omega_2 = 1/R_bC_{sh}$, and therefore $R_b = 1/\omega_2C_{sh} = 1/(2\pi 8.3 \times 10^6 \times 14 \times 10^{-12}) = 1360$ ohms. $Q_2 = \omega_2L/R_b$, and hence $L = 0.44 \times 1360/(2\pi \times 8.3 \times 10^6) = 11.5\ \mu h$. This completes the h-f compensation.

We are now ready to choose C and R_g. A conservative value of R_g for most tubes is 0.5 megohm, and the time constant R_gC usually should not greatly exceed 0.05 if grid blocking following an unusually high grid-signal pulse is to be avoided. In this problem let us use $R_gC = 0.01$. Hence $C = 0.01/0.5 \times 10^6 = 0.02\ \mu f$ is a conservative value.

The value of ω_1 by Eq. (7-73) is the reciprocal of the time constant R_gC, and hence $\omega_1 = 1/0.01 = 100$, making $f_1 = 100/2\pi = 16$ cps. We wish to lower the half-power frequency to 2 cps, and hence $\omega_3 = 2\pi \times 2 = 12.6$.

Equation (7-74) gives $R_fC_f = 1/\omega_3 = 1/12.6 = 0.079$. By Eq. (7-78), $[R_bR_f/(R_b + R_f)]C_f = R_gC = 0.01$. Dividing the second of these equations by the first yields $R_b/(R_b + R_f) = 0.126$, or $R_f = 6.93R_b = 6.93 \times 1360 = 9420$ ohms. $C_f = 0.079/9420 = 8.4\ \mu f$. The gain $A = g_mR_b = 5100 \times 10^{-6} \times 1360 \approx 7$. If C_f is an electrolytic capacitor, it should be paralleled by a small paper capacitor since many electrolytic capacitors are inductive at high frequencies.

If the shunt-feed circuit of Fig. 7-41 is used, the values of R_f and C_f do not greatly differ from those obtained for this case but the gain is reduced to 5.4.

In practice, the value of stray capacitance is usually determined experimentally. The amplifier chassis is wired and a load of a few thousand ohms inserted for R_b. By determining the frequency at which the gain drops to 71 per cent of its mid-frequency value, f_2 is determined. From this C_{sh} is computed. The load is then changed to the proper value which gives the desired f_2.

7-21. Balanced Voltage Amplifiers. Under some circumstances the two alternating voltages on the conductors bearing a signal are balanced to ground, *i.e.*, the potential of one wire is always as much above ground as the potential of the other wire is below ground. This is true in some measurement circuits, on ordinary telephone lines, in the input to a push-pull amplifier, and in the feed to the deflecting plates of a well-designed oscilloscope. Obviously an amplifier handling such a signal cannot be of the grounded-cathode single-ended variety. The two circuits of Fig. 7-44 represent typical balanced amplifiers. The coupling device in the output may be of the *R-C*, the transformer, or even the direct variety.

If an amplifier is used to step up the voltage, it can probably be arranged with choice of tubes, quiescent operation, and load so that the performance is approximately linear. The treatment may be that of two

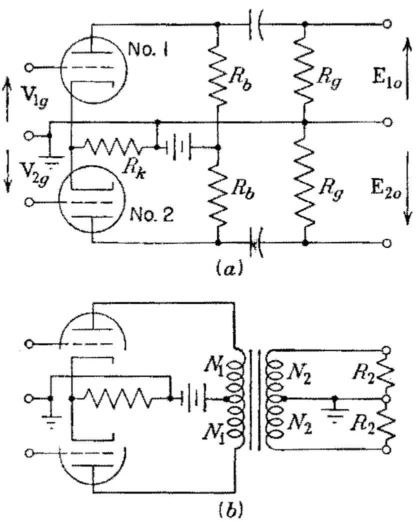


FIG. 7-44. Balanced amplifiers.

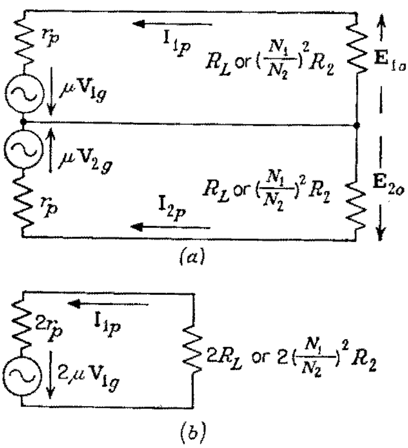


FIG. 7-45. Equivalent circuits of Fig. 7-44.

separate equivalent circuits as in Fig. 7-45*a*, or they may be combined as follows,

$$\mathbf{V}_{1g} = -\mathbf{V}_{2g} \tag{7-80}$$

and since the operation is linear,

$$\mathbf{I}_{1p} = -\mathbf{I}_{2p} \tag{7-81}$$

Hence $\mu\mathbf{V}_{2g}$ can be replaced by $\mu\mathbf{V}_{1g}$ with polarity signs reversed and $\mathbf{I}_{2p}$ by $\mathbf{I}_{1p}$ with arrow reversed. It should be clear that no alternating current flows through the center connecting wire, and so it may be omitted. We now have the equivalent circuit shown in Fig. 7-45*b*. Thus the a-c operation is the same as that of a circuit with the same total load, fed by a tube with twice the amplification factor and twice the plate resistance.

There is considerable advantage in the omission of C_k where balance is important. For example, if the No. 1 tube amplifies somewhat more than does the No. 2 tube because of minor differences in tubes or circuit values, the alternating-voltage drop in R_k caused by the sum of the two plate currents (which are 180° out of phase) will be a small amount, in phase with the plate current of the No. 1 tube. This will cancel some of the signal to the No. 1 tube and increase slightly the signal to the No. 2 tube, thus making the output voltages more nearly equal. An approximate formula for the ratio of output voltages with this self-balancing feature is given by

$$\frac{\mathbf{E}_{1o}}{\mathbf{E}_{2o}} = -\frac{\mathbf{A}_1}{\mathbf{A}_2}\frac{1 - 2(R_k/\mathbf{Z}_L)\mathbf{A}_2}{1 - 2(R_k/\mathbf{Z}_L)\mathbf{A}_1} \tag{7-82}$$

where the **A** values are those with R_k well bypassed.

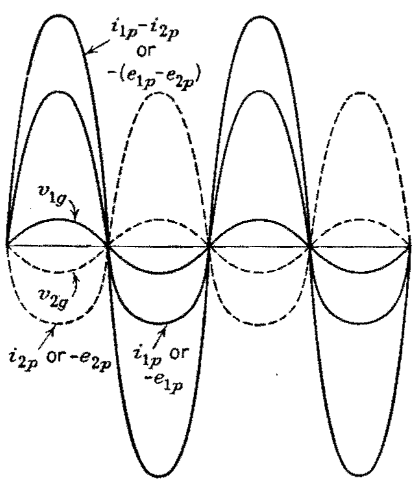


FIG. 7-46. Waveforms for the balanced amplifiers of Fig. 7-44.

If balanced amplifiers are operated with large grid signals, distortion of the plate currents occurs. This distortion may or may not appear in the output. In Fig. 7-44*a* a distorted plate current in tube 1 will cause a similar distorted waveform in its output voltage, and likewise for the No. 2 tube. However, if the output is taken between the top and bottom wires, even-ordered harmonics in the waveforms will cancel and the output will be essentially undistorted. Figure 7-46 shows waveforms of v_{1g}, v_{2g}, i_{1p}, i_{2p}, e_{1p}, e_{2p}, $i_{1p} - i_{2p}$, and $e_{1p} - e_{2p}$, the latter being the voltage between top and bottom wires. Note its good waveform.

The circuit of Fig. 7-44*b* acts in a somewhat different way, but the explanation can be made also by Fig. 7-46. As before, the plate currents are distorted, although in a slightly different manner. The flux in the transformer depends upon the mmf, which is proportional to the difference in the plate currents. Since $i_{1p} - i_{2p}$ has a good waveform, the flux and the voltages induced by it in all coils will also have a good waveform. In this case, not only the voltage between outside wires but voltages from either transformer secondary coil will be low in distortion. This action of the transformer is often used in power amplifiers even when balance is not needed. We shall study it further under push-pull power amplifiers.

7-22. Phase Inverters.[5] Devices which deliver two output voltages balanced to ground, but with only one input voltage, are called phase inverters (or splitters). A transformer with a center-tapped secondary is often used for the purpose, especially in cases where a low output impedance is necessary, as in the driver of a push-pull amplifier which is operated class 2, or in other cases where some voltage step-up is desirable. The phase-splitting transformer in the latter case comes under the interstage classification, and if it is of good quality, its response characteristic will usually meet moderate bandwidth requirements. For many purposes it is desirable to avoid the use of a transformer because of its restricted bandwidth and because it sometimes produces hum due to stray magnetic fields. Adaptations of the R-C-coupled amplifier supply other devices.

In Fig. 7-47 is shown the circuit of one type of inverter. The No. 1 tube and its load constitute an ordinary amplifier. Likewise the No. 2 tube and its load make a similar amplifier. The grid voltage for the latter should be 180° out of phase with that of the No. 1 tube, and hence the output voltage of the No. 1 tube is a good source for it. By arranging tap b so that R_{ab} is $1/A$ of R_g, the proper grid voltage will be available for the No. 2 tube. A is the numerical amplification of the No. 2 amplifier. The output $\mathbf{E}_{2o}$ is changed 180° in phase relative to the input to tube 2, and hence $\mathbf{E}_{1o}$ and $\mathbf{E}_{2o}$ have the proper phase relation. The No. 2 tube in this circuit is often called the inverter tube. Unfortunately, in the l-f and h-f ranges θ_A is not the 180° stated above, and the output voltages are not well balanced to ground. In ordinary a-f amplifiers this is not a troublesome matter, but for wider-band operation this is not a very good type of inverter.

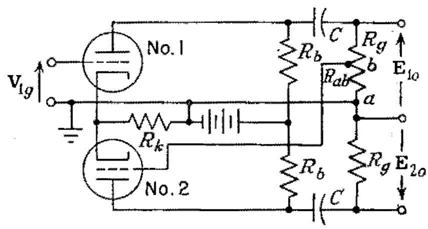


FIG. 7-47. A two-tube type of phase inverter.

The signal to the No. 2 grid really consists of two parts—that furnished by R_{ab} and that furnished by R_k. If the position of b is exactly correct and the a-c plate currents of the two tubes are equal, the signal voltage furnished by R_k will be zero. If some unbalance exists, R_k serves as a self-balancer. If position b is determined by experiment, R_k should be bypassed, during the adjustment, by a large capacitor to prevent the self-balancing action. After position b is found, removal of C_k will ensure an approximately balanced inverter even when old tubes are replaced by new ones.

The idea of R_k supplying a signal to the No. 2 tube is used in one type of inverter which has a circuit identical with that of Fig. 7-47 except that the No. 2 grid is grounded and the connection to tap b is removed. It may be shown that if

$$R_k \gg \frac{r_p + R_L}{\mu + 1} \tag{7-83}$$

satisfactory balance is achieved. Since such a value of R_k is likely to be too high to give satisfactory bias, the circuit is not particularly successful in its simplest form.

Another inverter circuit is shown in Fig. 7-48. It resembles an ordinary single-tube amplifier with its load split into halves. E_{bb} is placed in the middle to achieve a balance to ground. Since the output voltage E_{2o} is in the grid-to-cathode circuit, there is considerable feedback and a loss in gain results.

A simple explanation of the action of this inverter can be had if we design one of them. Suppose we want 21 volts rms on each side of the

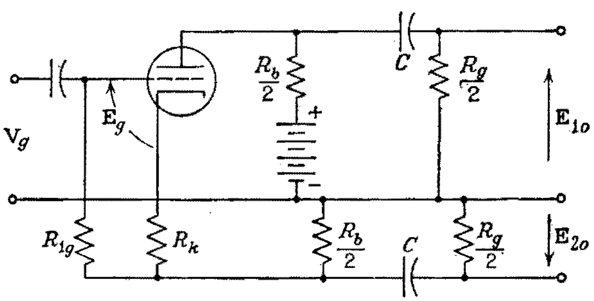


FIG. 7-48. A split-load type of phase inverter.

output, $E_{bb} = 225$ volts, and $f_1 = 50$ cps. If we choose a 6J5-type tube and consult the R-C-coupled amplifier chart in a tube manual, we can get along quite well. The procedure to be followed depends upon the tube manual employed; we shall use the chart in Appendix B for this example.

Under the type 6J5 we find several designs headed $E_{bb} = 250$ volts. Any of these designs is satisfactory for use with a plate-supply voltage within about ± 50 per cent of the tabular value. Our value of E_{bb} is 225 volts, and the obtainable value of E_o will be only $^{225}\!/_{250}$ as much as the published value for maximum grid drive under class A_1 conditions. This follows because the grid bias caused by the cathode resistor and hence the maximum allowable grid drive both change in approximately the same proportion as does E_{bb} if all circuit resistors remain unchanged. To allow for this lowered output voltage, we choose a design which lists E_o as $^{250}\!/_{225} \times 2 \times 21 = 47$ volts, or greater. One column which meets this condition lists the following circuit values: $E_{bb} = 250$ volts, $R_b = 0.047$ megohm, $R_g = 0.27$ megohm, $R_k = 2200$ ohms, $E_g = 3.5$ volts, $E_o = 52.5$ volts, $A = 15$, distortion $= 4.9$ per cent. If we use the resistors as listed, the data will be the same except that $E_{bb} = 225$ volts, $E_g = {}^{225}\!/_{250} \times 3.5 = 3.15$ volts, and $E_o = {}^{225}\!/_{250} \times 52.5 = 47.2$ volts. Since the desired output voltage is only 42 volts, the drive will be reduced to $(42/47.2) \times 3.15 = 2.8$ volts and the distortion to $(2.8/3.15) \times 4.9 = 4.4$ per cent. Because of feedback the actual distortion in the output voltage is much less than this figure.

In Fig. 7-49 the values of R_b and R_g are shown split in half. To find the values of the coupling capacitor, we may use formula (7-31). Thus $C = 1/2\pi R_{\text{low}} f_1 \approx 1/2\pi R_g f_1 = 1/(2\pi \times 0.27 \times 10^6 \times 50) = 0.012\ \mu\text{f}$. It should be noted that $10f_1$ is the lower working limit chosen for the amplifier. Since the tube sees two coupling capacitors in series, each one should be made twice the size calculated above. Thus in Fig. 7-49 each capacitor is labeled 0.025 μf.

The total output voltage from both sides being 42 volts, and the gain from the chart being 15, the grid voltage for the tube is $^{42}\!/_{15} = 2.8$ volts,

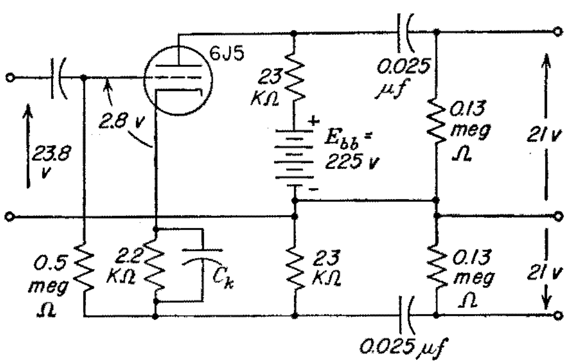


FIG. 7-49. A tube-manual design for a phase inverter.

as stated above. The voltages are shown in the figure with assumed polarity signs. Writing Kirchhoff's voltage equation for the grid-circuit loop yields $V_g = 21 + 2.8 = 23.8$ volts. The gain from the input of the amplifier to the output of one side is $21/23.8 = 0.88$.

R_k has been bypassed by a large capacitor C_k, but since there is considerable feedback anyway, the omission of C_k would not greatly change the voltage gain.

V_g drives a current into the grid lead, most of which flows down through the 0.5-megohm resistor. In fact, the voltage drop across this resistor is about 2.8 volts, and hence the current is $2.8/0.5 \times 10^6 = 5.6\ \mu\text{a}$. The impedance seen by V_g is therefore $23.8/5.6 = 4.3$ megohms, or 8.6 times the R_g (input) value. This high input impedance is characteristic of circuits which have negative feedback. In this case the R_{low} for the preceding stage is practically 4 megohms, and if f_1 for it is to be 50 cps also, $C_1 = 1/2\pi R_{\text{low}} f_1 = 1/(2\pi \times 4 \times 10^6 \times 50) \approx 0.0008\ \mu\text{f}$. Hence $C_1 = 0.001\ \mu\text{f}$ will be adequate.

This inverter gives an excellently balanced output for high frequencies up to very high values. Most wiring capacitances normally are balanced on the two sides. However, the heater is ordinarily at ground potential, and, the capacitance between cathode and heater being high, some unbalance exists on this account unless additional capacitance from plate to ground is provided. This is not difficult to do. Because of a high internal impedance to E_{1o} and a low internal impedance to E_{2o} (this is too advanced for us to see at the present time) the balance is somewhat upset in the very-high-frequency range. However, when circuit values like the ones in this example are used, the output voltages are still equal and the angle between them only a few degrees away from 180° at 1 Mc.

The principal fault of this inverter is the fact that the cathode is at a high a-c potential from ground and the heater is grounded. Because of capacitance between heater and cathode and some emission current between the two, a power-frequency current flows through R_k and R_L for the lower-half load, causing hum in the output. Hence this inverter should not be used where the voltage is very low but should preferably be installed immediately before the device which uses its output. By careful design it will produce enough voltage to drive push-pull amplifiers using most receiver types of power tubes in class AB_1.

7-23. Current Amplifiers. There are a few applications in which current output is a matter of interest. For example, it may be desired to actuate a moving-coil oscilloscope from a low-voltage high-impedance source. The coil usually has only a few ohms impedance, and around 100 ma may be needed for full-scale deflection. If alternating quantities are to be involved, one solution is an output transformer matching the coil to an amplifier tube in such a way that sufficient current flows with low distortion. This is essentially a power-amplifier design.

On the other hand suppose the current from a phototube is to operate a 100-ohm relay. The relay closes with 10 ma current flow and opens when the current falls to 8 ma. With peak illumination the tube current is 5 μa with no load in series with the tube. With no illumination the tube current is zero. The 5-μa current from the phototube is obviously quite insufficient to operate a directly connected relay. What is needed is current amplification of, say, $3 \times 10^{-3}/5 \times 10^{-6} = 600$ times to ensure reliable operation.

Continuing, the 100-ohm load on the output tube is so small it may be neglected. The quiescent operating point of the tube can be adjusted to make $I_{bo} = 7.5$ ma, so that normally the relay is open. The voltage gain A of the tube has no application in this case, but the transconductance g_m is quite useful, since $g_m = \partial i_b/\partial e_c$. Thus if $g_m = 1500$ micromhos, the required change in grid voltage for the output tube is $3 \times 10^{-3}/1500 \times 10^{-6} = 2$ volts. This input voltage can be obtained by using an R_g resistor in series with the phototube and in the grid circuit of the output tube. Its value should be $2/(5 \times 10^{-6}) = 400{,}000$ ohms. This resistor would change the current from the phototube only slightly because of the latter's very high plate resistance.

If a current is to be delivered to a small impedance load at the output of an amplifier and controlled by a voltage at the amplifier input several stages ahead, each stage except the last should be a voltage amplifier, the last one being a current amplifier. The ratio of I_{out}/E_{in} may be called g_c, the over-all-circuit transconductance. It may be expressed in formula form as

$$g_c = A_1 A_2 \cdots A_{n-1} g_{mn} \tag{7-84}$$

g_{mn} is the transconductance of the tube in the nth, or output, stage.

Current amplifiers operating with zero-impedance load give quite high distortion, and it is often advisable to use some load to reduce this distortion. In this case the computations can be made in the same manner as for a multistage voltage amplifier.

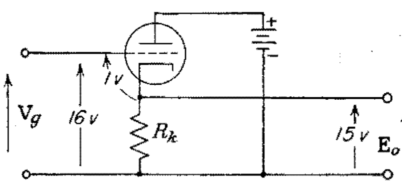


FIG. 7-50. A simple cathode follower.

7-24. The Cathode Follower.[6] It is not necessary that the input signal always be applied between cathode and grid or that the load be in the plate circuit. An interesting and useful example of other arrangements is the cathode follower, shown in its simplest form in Fig. 7-50. A rough picture of its performance can be gained as follows: Suppose $\mu = 18$, $r_p = 10{,}000$ ohms, and $R_k = 50{,}000$ ohms. Figured from a grid-cathode input to the output, the amplifier is conventional, and the gain is $\mu R_k/(r_p + R_k) = 18 \times 50{,}000/60{,}000 = 15$. Let us assume 1 volt rms applied between grid and cathode and that the operation is linear. Then E_o is 15 volts, and by Kirchhoff's voltage equation V_g is 16 volts. These voltages are shown on the circuit diagram. The over-all gain from input to output is only $^{15}\!/_{16} = 0.94\underline{/0°}$, the angle being zero because of the choice of positive sense for $\mathbf{E}_o$, it being generally convenient to regard ground as being at zero potential.

An "amplifier" which produces a loss in voltage might occasionally be useful for that reason alone, but the cathode follower has other very valuable properties besides, which often make it an object of use in spite of the lack of gain. To investigate these, let us assume that the operating point has been adjusted to a region of linear operation by taking the bias voltage from a proper point on R_k or by some other means and that the grid swing is small enough to keep operation in this region. If the output voltage is fed to another amplifier, a coupling capacitor C and a grid leak R_g may be needed. Practical circuits are shown in Fig. 7-51, using a triode and a pentode.

For low and middle frequencies the equivalent circuit of Fig. 7-52 applies to either triode or pentode if it is assumed that C_d is very large. The analysis for frequencies high enough so that the voltage drop in C is negligible is as follows:

$$\mathbf{E}_o = \frac{R_L}{r_p + R_L} \mu \mathbf{E}_g \qquad (7\text{-}85)$$

where R_L is the value of the parallel combination of R_k and R_g. Also

$$\mathbf{E}_g = \mathbf{V}_g - \mathbf{E}_o \qquad (7\text{-}86)$$

Hence

$$\mathbf{E}_o = \frac{R_L}{r_p + R_L} \mu \mathbf{V}_g - \frac{R_L}{r_p + R_L} \mu \mathbf{E}_o \qquad (7\text{-}87)$$

or

$$\mathbf{E}_o = \frac{\mu R_L \mathbf{V}_g/(r_p + R_L)}{1 + [\mu R_L/(r_p + R_L)]} = \frac{\mu R_L \mathbf{V}_g}{r_p + (1 + \mu)R_L} = \frac{[\mu/(1 + \mu)]R_L \mathbf{V}_g}{[r_p/(1 + \mu)] + R_L} \qquad (7\text{-}88)$$

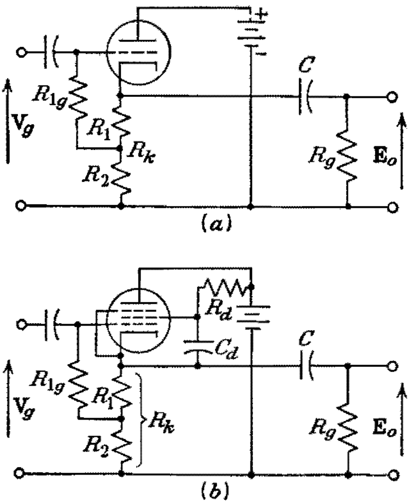


FIG. 7-51. Practical cathode-follower circuits.

The gain is

$$\mathbf{A}_{\text{mid}} = \frac{\mathbf{E}_o}{\mathbf{V}_g} = \frac{[\mu/(1 + \mu)]R_L}{[r_p/(1 + \mu)] + R_L} \qquad (7\text{-}89)$$

It is easy to show that A_{mid} can never exceed $\mu/(1 + \mu)$. It may also be shown that the position of the R_{1g} tap on R_k has little effect on the amplification with ordinary circuit values.

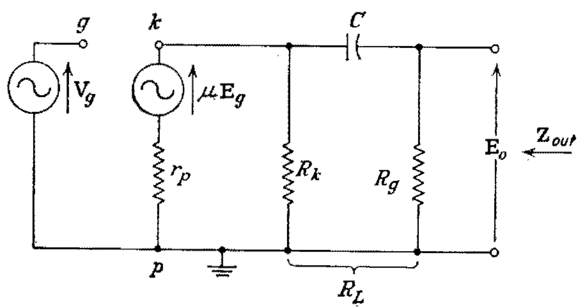


FIG. 7-52. The equivalent circuit for low frequencies.

Equation (7-88) suggests another equivalent circuit, valid in the mid-frequency range. It is shown in Fig. 7-53 (if we omit C), and it is easily verified that, as far as gain is concerned, it gives the result (7-89). In words, the cathode follower gives the same gain as would another tube, with plate resistance $r_p/(1+\mu)$ and amplification factor $\mu/(1+\mu)$, placed in the same circuit, but with the signal voltage applied directly between grid and cathode, that is, $\mathbf{E}_g = \mathbf{V}_g$.

To test whether the equivalence holds further than for gain, one may find the internal impedance from the output end by applying **E** volts to the output of each of these circuits and determining the current **I** which

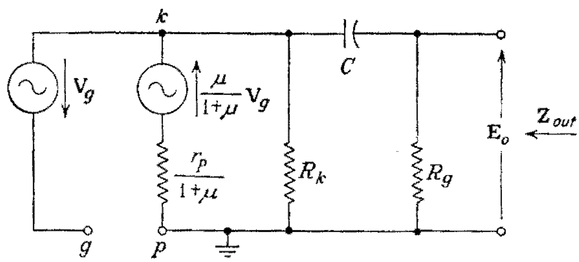


FIG. 7-53. A more useful equivalent circuit for a cathode follower.

flows. Since the current must be due to the application of **E** alone, let us make $\mathbf{V}_g$ equal to zero. In Fig. 7-53 the current **I** is

$$\mathbf{I} = \frac{\mathbf{E}}{R_L} + \frac{\mathbf{E}}{r_p/(1+\mu)} \tag{7-90}$$

and

$$\mathbf{Y}_{\text{out}} = \frac{\mathbf{I}}{\mathbf{E}} = \frac{1}{R_L} + \frac{1}{r_p/(1+\mu)} \tag{7-91}$$

In the circuit of Fig. 7-52, if we let $\mathbf{V}_g$ equal zero, $\mathbf{E}_g$ is not zero,

$$\mathbf{E}_g = -\mathbf{E} \tag{7-92}$$

and hence

$$\mathbf{I} = \frac{\mathbf{E}}{R_L} + \frac{\mathbf{E} - \mu(-\mathbf{E})}{r_p} = \mathbf{E}\left(\frac{1}{R_L} + \frac{1+\mu}{r_p}\right) \tag{7-93}$$

and hence

$$\mathbf{Y}_{\text{out}} = \frac{1}{R_L} + \frac{1}{r_p/(1+\mu)} \tag{7-94}$$

as before. $\mathbf{Z}_{\text{out}}$, not including R_L, is

$$\mathbf{Z}_{\text{int}} = \frac{r_p}{1+\mu} \approx \frac{1}{g_m} \qquad \text{if } \mu \gg 1 \tag{7-95}$$

In Fig. 7-52 this fact cannot be determined readily by inspection because application of a voltage at the output produces a grid voltage, which does not occur in the simpler equivalent circuit.

The analysis for low frequencies is quite lengthy although the general procedure is the same as that above, except that the impedance of C must be included. The result obtained using the circuit of Fig. 7-52 is

$$\mathbf{A}_{\text{low}} = \frac{\mathbf{A}_{\text{mid}}}{1 + (1/j\omega C R'_{\text{low}})} \tag{7-96}$$

where

$$R'_{\text{low}} = R_g + \frac{[r_p/(1+\mu)]R_k}{[r_p/(1+\mu)] + R_k} \tag{7-97}$$

If f_1 is the lower half-power frequency, given in formula form

$$f_1 = \frac{1}{2\pi R'_{\text{low}} C} \tag{7-98}$$

we may write

$$\mathbf{A}_{\text{low}} = \frac{\mathbf{A}_{\text{mid}}}{1 - j(f_1/f)} \tag{7-99}$$

as in the case of an R-C-coupled grid-driven plate-loaded amplifier.

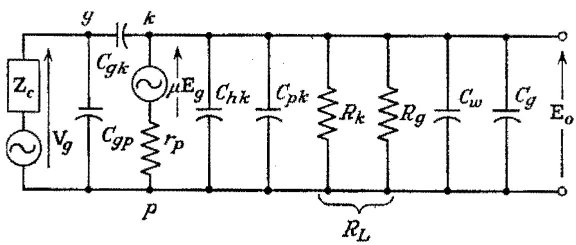


FIG. 7-54. The h-f equivalent circuit for a cathode follower.

Examination of the circuit of Fig. 7-53 shows that R'_{low} can be obtained directly from it by a simple conversion to a Thévenin's form. It follows that the equivalent circuit of Fig. 7-53 is valid for the l-f range.

For high frequencies the tube and circuit capacitances cause amplification loss and phase shift. The complete equivalent circuit for this frequency range is shown in Fig. 7-54. The capacitances involved may be lumped into C_{sh}, where

$$C_{sh} = C_{gk} + C_{pk} + C_{hk} + C_w + C_g \tag{7-100}$$

in which the first three belong to the cathode-follower tube. The degree of effectiveness of C_{gk} depends upon the value of Z_c. C_{hk} is the heater-cathode capacitance, and it is important because the cathode is at considerable alternating voltage from the grounded heater. C_w is the wiring capacitance and C_g is the input capacitance to the following tube. The development is somewhat involved, and only the results are given here.

$$\mathbf{A}_{\text{high}} = \frac{(g_m + \mathbf{Y}_{gk})\mathbf{Z}_{sh}}{1 + g_m \mathbf{Z}_{sh}} \tag{7-101}$$

where $\mathbf{Y}_{gk}$ is the admittance of C_{gk} and $\mathbf{Z}_{sh}$ is the impedance of the parallel combination of C_{sh}, r_p, R_k, and R_g. For circuit values commonly used and for frequencies in the useful range (less than f_2) the quantity $\mathbf{Y}_{gk}$ can be neglected with small error. After considerable manipulation the expression for $\mathbf{A}_{\text{high}}$ simplifies to

$$\mathbf{A}_{\text{high}} = \frac{\mathbf{A}_{\text{mid}}}{1 + j(f/f_2)} \tag{7-102}$$

as in the case of an ordinary amplifier. The value of $\mathbf{A}_{\text{mid}}$ is stated in Eqs. (7-89) and (7-106). f_2 is the upper half-power frequency, defined as

$$f_2 = \frac{1}{2\pi R'_{sh} C_{sh}} \tag{7-103}$$

where

$$\frac{1}{R'_{sh}} = \frac{1}{R_L} + \frac{1}{r_p/(1 + \mu)} \tag{7-104}$$

Figure 7-55 shows a Norton's equivalent circuit useful for high frequencies. It is interchangeable with that of Fig. 7-53 if the proper quantities are present for the frequency being used. From Fig. 7-55 one may determine R'_{sh} as with an ordinary amplifier circuit. For middle frequencies C_{sh} may be neglected so that

$$\mathbf{E}_o = g_m \mathbf{V}_g R'_{sh} \tag{7-105}$$

or

$$\mathbf{A}_{\text{mid}} = g_m R'_{sh} \tag{7-106}$$

It can be shown that this is identical to formula (7-89).

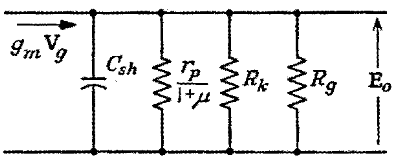


FIG. 7-55. Another h-f equivalent circuit for a cathode follower.

The input admittance to a cathode follower may be obtained by determining the current flowing into the input because of the applied voltage $\mathbf{V}_g$. Its value depends upon the circuit parameters, the signal frequency, and the position of the bias tap on R_k. For the case of the simple cathode-follower circuit of Fig. 7-50, it can be readily proved that

$$\mathbf{Y}_g = j\omega[C_{gp} + (1 - \mathbf{A})C_{gk}] \tag{7-107}$$

which resembles the formula applying to an ordinary amplifier. Note, however, the differences.

7-25. Graphical Treatment of Cathode Followers.[7] The treatment of cathode followers by means of an equivalent circuit leaves some matters in the realm of the unknown. For example, what operating point is best, what is the best load to use, how much drive may be applied for class A_1 operation, and what will be the distortion?

A graphical method serves very well in these matters and will be explained for a triode case. Figure 7-56 repeats the circuit, and Fig. 7-57 presents the plate characteristics of the 6J5 which is used.

The d-c load is R_k, and assuming $E_{bb} = 300$ volts, we draw sample d-c-load lines, say for 10,000 ohms and for 100,000 ohms. In most cases R_g is much larger than R_k, and these lines are the first approximation to the a-c-load lines also. The first item for investigation is the proper operating point. If the grid signal is small, almost any operating point not too

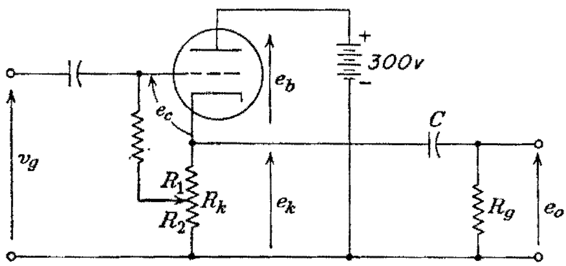


FIG. 7-56. A cathode follower.

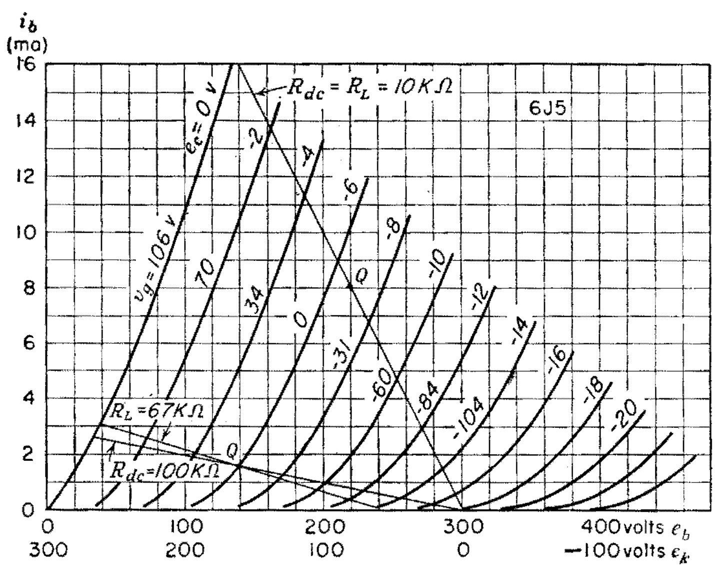


FIG. 7-57. Graphical analysis of a cathode follower.

near cutoff and which does not cause excessive plate dissipation will serve fairly well, but if it is desired to get the maximum output voltage from the amplifier, more care is required. Let us proceed graphically as follows:

1. Label e_k units on the e_b axis, using the relation

$$e_k = E_{bb} - e_b \tag{7-108}$$

where e_k is the total voltage of the cathode above ground.

2. Determine the relation between e_o, v_g, and e_g. This is best done by writing the alternating-voltage drop around the grid circuit, thus avoid-

ing the determination of the direct-voltage drop across the input-coupling capacitor. If we neglect the alternating-voltage drop in C,

$$v_g = e_o + e_g \tag{7-109}$$

3. Relate e_o and e_g to their total and quiescent values, and substitute these in Eq. (7-109).

$$e_o = e_k - E_{ko} \tag{7-110}$$
$$e_g = e_c - E_{co} \tag{7-111}$$

v_g has zero quiescent value.

$$v_g = e_k - E_{ko} + e_c - E_{co} \tag{7-112}$$

4. The proper operating point for greatest class A_1 output makes $e_k = e_{k_{\max}}$ on the peak positive swing of v_g ($v_{g_{+\max}}$) and $e_k = e_{k_{\min}}$ on the peak negative swing of v_g ($v_{g_{-\max}}$). For $e_{k_{\max}}$ we shall use a value that makes $e_c = 0$ and for $e_{k_{\min}}$, the value corresponding to $e_{c_{\min}}$, any arbitrary value, usually at or near the cutoff value of e_c. We may substitute into Eq. (7-112) to obtain two equations

$$v_{g_{+\max}} = e_{k_{\max}} - E_{ko} + 0 - E_{co} \tag{7-113}$$
$$v_{g_{-\max}} = e_{k_{\min}} - E_{ko} + e_{c_{\min}} - E_{co} \tag{7-114}$$

Since

$$v_{g_{+\max}} = -v_{g_{-\max}} \tag{7-115}$$

we get, upon adding (7-113) and (7-114) and solving for E_{ko},

$$E_{ko} = \frac{e_{k_{\max}} + e_{k_{\min}}}{2} - \left(E_{co} - \frac{e_{c\min}}{2}\right) \tag{7-116}$$

The quantity in parentheses is rather small compared with the first term, and hence

$$E_{ko} \approx \frac{e_{k_{\max}} + e_{k_{\min}}}{2} \tag{7-117}$$

5. For the 10,000-ohm-load line $e_{k_{\max}} = 163$ volts; $e_{k_{\min}} = 0$, if we drive to cutoff; $e_{c_{\min}} = -18$ volts. Hence $E_{ko} = {}^{163}\!/_{2} = 81.5$ volts, and from Fig. 7-57, $E_{co} = -6.8$ volts. Note that E_{co} is not one-half of $e_{c_{\min}}$. For the 100,000-ohm-load line $e_{k_{\max}} = 266$ volts, $e_{k_{\min}} = 22$ volts, if we limit the drive to $e_c = -16$ volts. Hence $E_{ko} = (266 + 22)/2 = 144$ volts, and $E_{co} = -7$ volts.

6. For the 10,000-ohm-load line $R_1 = (E_{co}/E_{ko})R_k = (6.8/81.5) \times 10{,}000 = 835$ ohms. For the 100,000-ohm-load line $R_1 = {}^{7}\!/_{144} \times 100{,}000 = 4850$ ohms.

This is one method for finding the best operating point. In case the a-c load differs greatly from the d-c load, a cut-and-try adjustment may be made after finding the tentative Q point, in order to operate between the proper limits.

We may now continue with a more specific example. Suppose R_k = 100,000 ohms and R_g = 200,000 ohms. Then R_L = 67,000 ohms, which differs considerably from the d-c load of 100,000 ohms. If the a-c-load line were drawn (it is not shown in the figure) through the tentative Q point obtained in procedure 5, the value of $e_{k_{max}}$ = 260 volts and $e_{k_{min}}$ = 50 volts would make $E_{ko} = (260 + 50)/2 = 155$ volts, which is to the left of the tentative Q point. Hence, whenever the a-c load is less than the d-c load, we shift to the left, in this case to, say, E_{ko} = 162 volts as a trial value. The a-c-load line through this operating point is shown in the figure. For it, $e_{k_{max}}$ = 262 volts, and we can make $e_{k_{min}}$ = 62 volts, or somewhat more if desired. For the 62-volt value, E_{ko} becomes 162.0 volts, which is a sufficient check on the trial value. Therefore E_{ko} = 162 volts, E_{co} = -6 volts, and R_1 = 3700 ohms are satisfactory circuit values.

Equation (7-112) may now be used to relabel the curves of Fig. 7-57 with input-voltage values corresponding to the various grid voltages.

For $e_c = 0$, $v_g = 262 - 162 + 0 + 6 = 106$ volts.

For $e_c = -2$ volts, $v_g = 228 - 162 - 2 + 6 = 70$ volts.

For $e_c = -4, -6, -8, -10, -12, -14$, respectively, the values of v_g are 34, 0, -31, -60, -84, and -104 volts.

If a sinusoidal input signal of 104 volts peak value is used, the fundamental peak-voltage output will be $(260 - 66)/2 = 97$ volts [see Eq. (6-41)]. The second-harmonic peak-voltage output is $(260 + 66 - 2 \times 162)/4 \approx 0.5$ volt, and the harmonic distortion is approximately 0.5 per cent of the fundamental.

If the same resistors had been used as a plate load in the amplifier with an input signal of 6 volts peak value applied directly to the grid, the fundamental peak output voltage would have been $(260 - 83)/2 = 88$ volts. The second-harmonic peak voltage would have been $(260 + 83 - 2 \times 162)/4 = 5$ volts. The distortion would then have been approximately 5.7 per cent of the fundamental.

Had the a-c load of 10,000 ohms been used in the cathode follower the distortion for maximum allowable drive would have been approximately 2 per cent. Hence, as in the case of ordinary amplifiers, moderately high load values are advantageous from the point of view of freedom from distortion as well as of allowable output voltage. But in any case the distortion for the cathode follower has been found to be low.

At the operating point for the 67,000-ohm load the value of g_m for the tube is 1000 micromhos, which makes the output impedance approximately 1000 ohms, while, for the 10,000-ohm load, g_m is 2600 micromhos and the output impedance is about 380 ohms. Thus for low output impedance lower loads are required.

The simple cathode-follower circuit of Fig. 7-50 is useful for only very restricted loads. If the load is too low (less than 1200 ohms for a 6J5 with $E_{bb} = 300$ volts), the allowable plate dissipation is likely to be exceeded. For loads which are too high the Q point is very near the cut-off point, which restricts the input voltage. However, for small input voltages this simple circuit will serve very nicely.

7-26. The Grounded-grid Amplifier.[8] Another circuit variation from the ordinary amplifier and one which has considerable usefulness is the grounded-grid or cathode-driven amplifier. A simple circuit arrangement is shown in Fig. 7-58*a*, in which a d-c path from cathode to grid is assumed to exist. If the operation is assumed to be linear, Fig. 7-58*b* is the equivalent circuit.

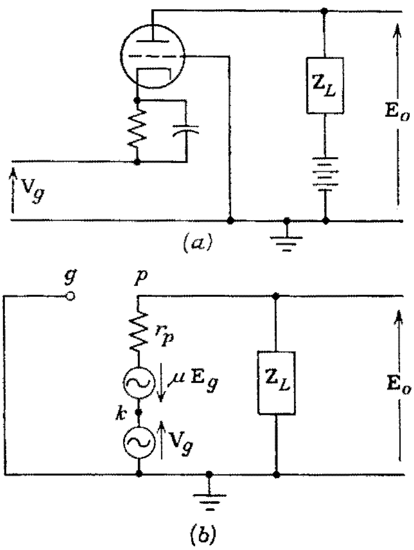


FIG. 7-58. A grounded-grid amplifier.

It is readily shown that the mid-frequency gain between input and output is

$$\mathbf{A}_{\text{mid}} = +\frac{(1+\mu)\mathbf{Z}_L}{r_p + \mathbf{Z}_L} \qquad (7\text{-}118)$$

For low and middle frequencies the input impedance for this amplifier is

$$\mathbf{Z}_g = \frac{\mathbf{V}_g}{-\mathbf{I}_p} = \frac{r_p + \mathbf{Z}_L}{1+\mu} \qquad (7\text{-}119)$$

which is very low. This low input impedance often makes the gain of the preceding stage lower than ordinary unless that stage itself has a very low internal impedance. The output impedance of the grounded-grid amplifier on the other hand is

$$\mathbf{Z}_{\text{out}} = r_p \qquad (7\text{-}120)$$

which is relatively high. Note that these characteristics are the opposite of those for a cathode follower, which operates with a grounded plate.

In this circuit the grounded grid acts as a shield between the input and output circuits, and most of the current through the interelectrode capacitances, caused by E_o, flows to ground via the grid rather than to ground via the cathode and input circuit. This eliminates much of the danger of oscillation which exists in most triode r-f circuits.

The commonest application of this circuit is to r-f use, where the effect just discussed makes it unnecessary (or at least easier) to neutralize the effect of feedback from the load circuit to the input circuit.

7-27. Transformers with Iron Cores.[9] Transformers are useful in many ways in electronic circuits. As power transformers they change the power-line voltage to other values suitable for plate-, grid-, and heater-supply systems. An output transformer makes the load or line impedance appear as a suitable value for the plate circuit of the power tube. An input transformer is useful with a low-impedance voltage source to raise its voltage to a higher level to apply to the grid of the first amplifier tube. This results in greater output voltage and also reduces the relative noise output caused by the first tube. Interstage transformers were once popular because they permitted quite high gain with low plate-supply voltage and a low-μ tube. As mentioned in Art. 5-15, a transformer with a center-tapped secondary can be used as a phase splitter. This is particularly useful with push-pull power amplifiers which draw grid current, as the apparent impedance of the source of grid signal may be kept low, thereby reducing the voltage loss on grid-current peaks.

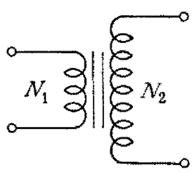


FIG. 7-59. A symbol for a transformer.

A two-winding iron-core transformer is a four-terminal device, which may be represented by the symbol shown in Fig. 7-59. The winding connected to the energy supply is called the primary, the other winding the secondary. Either coil may be used as the primary, but let us suppose in this case that the coil with N_1 turns is the primary.

If a well-designed power transformer of known turns ratio is tested in the laboratory, several important observations may be made. (1) With a direct voltage applied to the primary, the resistance of the primary winding may be determined. (2) With the secondary open-circuited and with the primary applied voltage E_1 of rated frequency and of magnitude no greater than rated, measurements with ammeter and wattmeter indicate that the primary current is very small and lags the applied voltage almost 90°. This indicates the effect of very high inductance and some resistance, calculations showing the latter to be greater than the winding resistance determined in test 1. A secondary terminal voltage E_2 may be measured, and it will be found that the relation $E_1/E_2 = N_1/N_2$ approximately holds. (3) If a variable load impedance $\mathbf{Z}_2$ is connected across the secondary terminals, a secondary current $\mathbf{I}_2$ flows, its magnitude being very closely proportional to the magnitude of the load admittance. Also this current lags $\mathbf{E}_2$ by approximately the power-factor angle of the impedance. The primary current will be found to be larger now, and it may be determined that the size of the increase $\mathbf{I}_1$ in primary current is related to the secondary current by the approximate relation $I_1/I_2 = N_2/N_1$. This additional primary current, due to the load, lags the applied primary voltage by approximately the phase angle of the load.

Now if we neglect the small no-load current determined in test 2 and declare the two approximate relations obtained in the laboratory to be exact, we have the performance of an ideal transformer. Its working rules may be recapitulated as the voltage rule

$$\frac{E_1}{E_2} = \frac{N_1}{N_2} \tag{7-121}$$

the E's being terminal voltages, and the current rule

$$\frac{I_1}{I_2} = \frac{N_2}{N_1} \tag{7-122}$$

the I's being the primary and secondary currents. Dividing Eq. (7-121) by Eq. (7-122) yields a form

$$\frac{Z_1}{Z_2} = \left(\frac{N_1}{N_2}\right)^2 \tag{7-123}$$

where Z_1 is the impedance seen by the generator and Z_2 is the load impedance.

It should be noted that an ideal transformer draws no primary current when the secondary is open. This means that the primary self-inductance is infinite, and because of the definite ratio of turns, the secondary self-inductance is infinite also. There can be no core losses, which would require no-load primary current. The currents $\mathbf{I}_1$ and $\mathbf{I}_2$ depend only on the applied primary voltage, the ratio of turns, and the load impedance. Hence there are no winding resistances or any capacitance or inductance effects due to the transformer itself, except one, and that is determined by the mutual inductance between windings, which causes the action expressed by the above equations. It is sometimes thought better to represent an ideal transformer by the symbol shown in Fig. 7-60 rather than by the more generally used symbol of Fig. 7-59. The letter a represents the ratio N_2/N_1. This alternate symbol emphasizes that the ideal transformer is another kind of circuit element in addition to resistors, inductors, and capacitors, this one being a ratio changer for voltages, currents, and impedances.

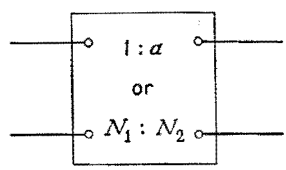


FIG. 7-60. An alternate symbol for an ideal transformer.

An actual transformer never possesses the perfect qualities of the ideal. The no-load current, required to supply core losses and to produce an mmf to force flux through the core, often is an important part of the total primary current, especially when the load impedance is high. The flux which produces the counter emf in the primary does not all link with the secondary. The nonlinking flux causes inductance effects which are not mutual, and these, together with the winding resistances, cause voltage losses.

Figure 7-61 shows one equivalent circuit for a transformer, neglecting capacitances and core losses. L_1 and L_2 are the self-inductances and R_p and R_s the winding resistances of the primary and secondary, respectively. The two polarity dots indicate that these two coil ends are simultaneously of like polarity. The condition with both dots at the same coil ends is sometimes called subtractive polarity. That with the dots at opposite coil ends is called additive polarity. To study a transformer it is often desirable to obtain a different form, which may be evolved in the following manner:

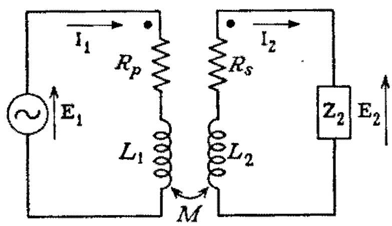


FIG. 7-61. An equivalent circuit for a transformer.

If we assume sinusoidal applied voltage and currents, we may write loop equations

$$\mathbf{E}_1 = \mathbf{I}_1(R_p + j\omega L_1) - \mathbf{I}_2 j\omega M \tag{7-124}$$
$$\mathbf{E}_2 = \mathbf{I}_1 j\omega M - \mathbf{I}_2(R_s + j\omega L_2) \tag{7-125}$$

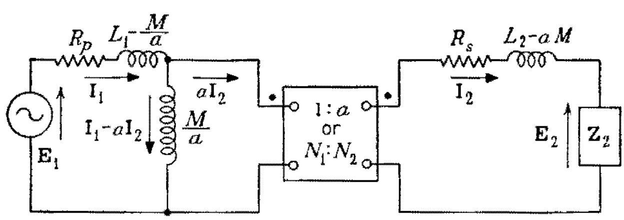


FIG. 7-62. Another equivalent circuit for a transformer.

These may be rewritten as

$$\mathbf{E}_1 = \mathbf{I}_1\left(R_p + j\omega L_1 - j\omega\frac{M}{a}\right) - \mathbf{I}_2 j\omega M + \frac{\mathbf{I}_1}{a}j\omega M \tag{7-126}$$
$$\mathbf{E}_2 = \mathbf{I}_1 j\omega M - \mathbf{I}_2 j\omega Ma - \mathbf{I}_2(R_s + j\omega L_2 - j\omega Ma) \tag{7-127}$$

or as

$$\mathbf{E}_1 = \mathbf{I}_1 R_p + \mathbf{I}_1 j\omega\left(L_1 - \frac{M}{a}\right) + (\mathbf{I}_1 - a\mathbf{I}_2)j\omega\frac{M}{a} \tag{7-128}$$
$$\mathbf{E}_2 = (\mathbf{I}_1 - a\mathbf{I}_2)j\omega\frac{M}{a}a - \mathbf{I}_2[R_s + j\omega(L_2 - aM)] \tag{7-129}$$

Figure 7-62 shows a circuit which yields the same loop equations (7-128) and (7-129) and hence will serve as another equivalent circuit for the transformer. The inductance L_1 is sometimes known as the magnetizing inductance, while $L_1 - (M/a)$ and $L_2 - aM$ are named the leakage inductances L_p and L_s of the primary and secondary, respectively.

The core losses in a transformer at any fixed frequency are roughly proportional to the second power of the maximum flux density in the iron and hence to the square of the voltage induced by the mutual flux. A resistor R_c of the proper size connected as a shunt in the position shown in Fig. 7-64 will be subjected to this primary induced voltage and hence will dissipate power equal to the core losses. The proper size for R_c may be determined very closely by an open-circuit test as

$$R_c = \frac{E_{1r}^2}{P_{oc}} \tag{7-130}$$

where E_{1r} is the rated primary voltage (approximately equal to the mutual induced voltage) and P_{oc} is the power consumed by the primary with the secondary circuit open. The value of R_c so obtained is reliable only for the frequency used in obtaining it.

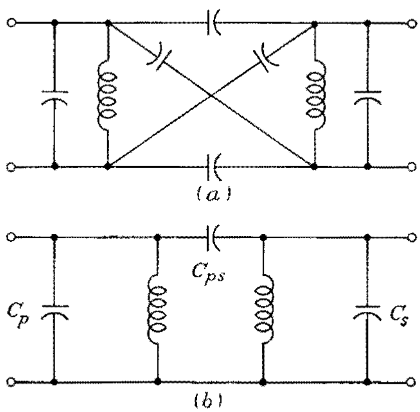


FIG. 7-63. Distributed capacitances in a transformer.

The effect of interturn or distributed capacitances on the transformer performance can be approximated. As measured externally the capacitances seem to be between terminals, and as shown in Fig. 7-63a, six

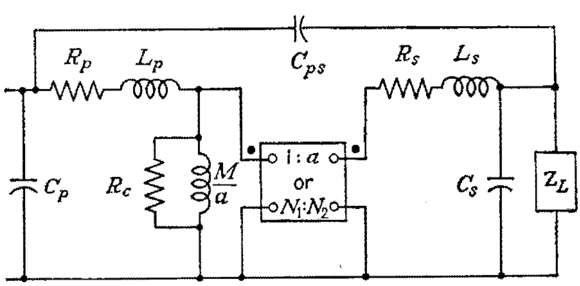


FIG. 7-64. The effects of distributed capacitances and core loss shown in the equivalent circuit.

quantities may be determined. Usually one end of the primary and one end of the secondary are at the same a-c potential, because of grounding requirements, and the six values can be lumped into three, which are designated in Fig. 7-63b as C_p, C_{ps}, and C_s. These capacitances are shown in the equivalent circuit of Fig. 7-64. At power frequencies the effect of the capacitances is generally slight, and they are often neglected in power-transformer analyses. For supplied voltages of higher audio frequencies, however, the ratio of output voltage to input voltage is considerably dependent upon these capacitances, and their effect should be considered in any study of performance at higher frequencies.

7-28. The Transformer with Load. Transformers which couple a power tube to its load, drive class B amplifiers, or serve for ordinary power purposes involve secondary load currents which dwarf any current

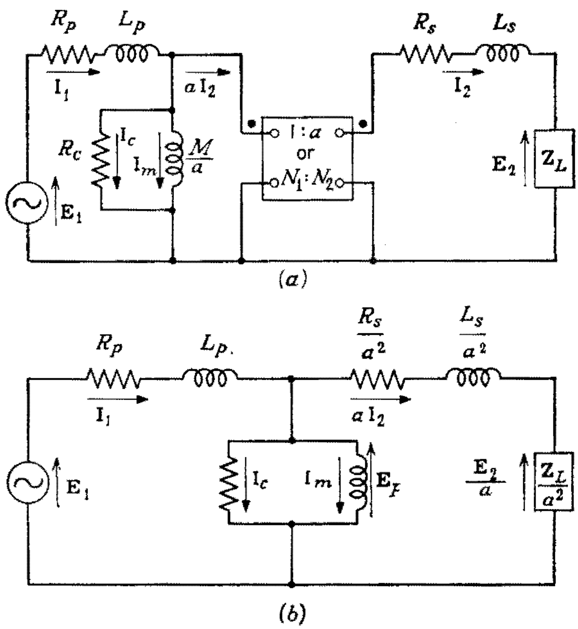


FIG. 7-65. Equivalent circuits for a power transformer.

through the transformer capacitances. Hence the capacitances of Fig. 7-64 are omitted from this study of power transformers.

In Fig. 7-65a suppose a generator supplies a sinusoidal voltage $\mathbf{E}_1$ to the transformer's primary terminals. Let us consider that $\mathbf{E}_1$ has three components, one to supply the drop in R_p, a second to supply the reactance drop in L_p, and a third, denoted by $\mathbf{E}_p$, the voltage across M/a or R_c.

Assume a load $\mathbf{Z}_L = R_L + jX_L$ to be connected to the secondary terminals, the load voltage being $\mathbf{E}_2$. The ratio of N_2 to N_1 is designated by a, and the polarity, as shown by the position of the two dots, is in this case such that both top terminals are positive at the same moment. A procedure, which is often convenient in coupled circuits such as this, is to change the transformer ratio to unity. This can be done without any effect upon the primary circuit if the secondary current is increased a

times, brought about by changing the secondary impedance by a factor $1/a^2$, $1/a$ times because the voltage has dropped $1/a$ times, and again $1/a$ times because the current must be raised to $a\mathbf{I}_2$. Since the primary current of the ideal transformer exactly equals $a\mathbf{I}_2$, and the polarities are correct, there is no need to show the ideal transformer in the circuit. Figure 7-65*b* shows the new form.

$\mathbf{I}_1$ has three components: $a\mathbf{I}_2$, which supplies the load, a core-loss current $\mathbf{I}_c$ in phase with $\mathbf{E}_p$, and a nonsinusoidal magnetizing current I_m, the fundamental component of which lags $\mathbf{E}_p$ by somewhat less than 90°. The latter two currents combined are called the exciting current, and with no load on the secondary they are the only currents flowing if we neglect capacitance currents. The small exciting current causes very little voltage drop in R_p and L_p, which are usually quite small themselves, and hence $\mathbf{E}_p$ is practically the same as $\mathbf{E}_1$ at no load. As the load impedance is decreased from a very high value, an increasing load current flows, causing voltage drops in the primary and secondary impedances and making E_2/a lower (if the load is resistive or inductive) than E_1. Hence the voltage regulation of a transformer from no load to full load is caused by the primary and secondary resistances and leakage reactances, and for good regulation these should be kept as small as practicable.

Because the contributions of I_c and I_m to I_1 are small, it is generally justifiable under full-load conditions to omit R_c and M/a from the equivalent circuit. Since, for coils wound on the same iron core, the inductances are proportional to the square of the number of turns, $L_s/a^2 = L_p$. In the case particularly of a large transformer, it is desirable for high efficiency to have $R_s/a^2 = R_p$, although this is not usually done in electronic applications. But if these conditions do hold, the circuit of Fig. 7-65*b* allows an approximate relation to be written as

$$\frac{\mathbf{E}_2}{a} = \mathbf{E}_1 - \mathbf{I}_1 \times 2(R_p + j\omega L_p) \tag{7-131}$$

and if the values of R_p and L_p are small enough, it may be justifiable to state that $\mathbf{E}_2/a = \mathbf{E}_1$, or

$$\frac{\mathbf{E}_2}{\mathbf{E}_1} = a = \frac{N_2}{N_1} \tag{7-121}$$

Hence Eqs. (7-121) to (7-123) can be applied to a power transformer, but with good approximation only when the winding resistances and leakage reactances are small and when Z_L/a^2 is very small compared with the value of $\omega M/a$; otherwise the effect of the magnetizing current cannot be neglected.

An output transformer is a power type of transformer which has an efficiency usually ranging between 70 and 90 per cent; under these circumstances the relation of (7-121) is not very reliable. A somewhat better expression can be derived from power relations. Let us assume the transformer to be of good design, with the copper losses (half the total) distributed equally between the primary and the secondary, and the core losses the other half of the total. Neglect the voltage drops due to leakage reactances. Then it can be shown that for the usual range of values of efficiency,

$$\frac{E_2}{E_1} \approx \frac{N_2}{N_1}\sqrt{\eta} \tag{7-132}$$

The proof of this statement is left as an exercise for the student (see Prob. 37 at the end of this chapter).

In order to avoid excessive magnetizing current, as well as very large odd-harmonic components in it, the maximum flux density in the iron core reached during a cycle should be limited to the lower region of the knee of the curve of B vs. H, approximately as shown in Fig. 7-66. The allowable maximum flux-density value depends upon the grade of iron employed. A typical path of operation for an applied primary alternating voltage is shown by the broken-line loop a. In some applications the primary carries a direct current as well as an alternating one, and care must be exercised in the design to keep the path of operation in a position not greatly higher than that shown by loop b. To do this, it may be necessary to have an air gap in the core.

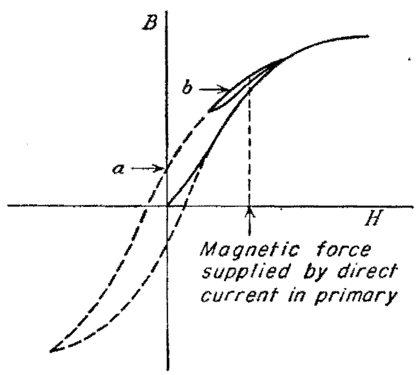


FIG. 7-66. Magnetic operation of a transformer.

7-29. The Transformer-coupled Amplifier.[12] If a transformer is used for coupling two stages of an amplifier, it effectively isolates the grid of the second tube from the direct plate voltage of the first and may also cause some voltage amplification because of a step-up turns ratio. If the second stage is operated class A_1, its grid draws no conduction current and the only load on the secondary is that due to currents flowing through the various capacitances. The grid current of the next tube may be said to flow through an admittance $\mathbf{Y}_g$, which may be calculated using either formula (7-12) or formula (7-13), and in any case $\mathbf{Y}_g$ is practically all capacitive for frequencies in the transformer's useful range. The effect is represented by C_g in Fig. 7-67.

Since this circuit is quite complicated, it is difficult to make an exact analysis. However, one may simplify the circuit to yield readily some approximate results. First, with subtractive polarity as marked, the secondary current will be almost unchanged and the primary current but slightly changed if C_{ps} is replaced by a capacitance $C_{ps}[(a-1)/a]$ in parallel with C_s. (If the transformer has additive polarity, the replacement value is $C_{ps}[(a+1)/a]$, which is larger, and hence additive polarity is undesirable.) This follows because the voltage drop across C_{ps} is $\mathbf{E}_2 - \mathbf{E}_1$, which is approximately $(a-1)\mathbf{E}_1$, whereas that across C_s is $\mathbf{E}_2$, or approximately $a\mathbf{E}_1$. In the same manner as in the previous article the

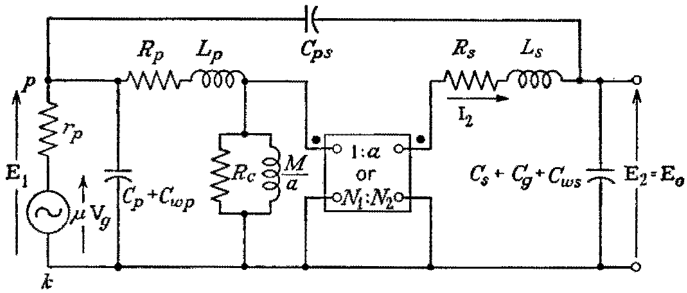


FIG. 7-67. An equivalent circuit for a transformer-coupled amplifier.

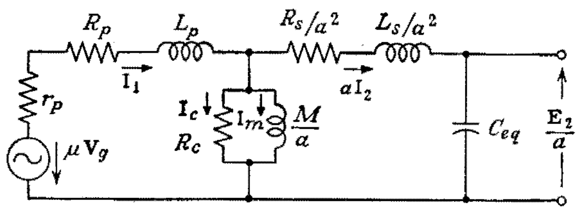


FIG. 7-68. Another equivalent circuit for a transformer-coupled amplifier.

transformer turns ratio may be changed to 1:1, yielding the simplified equivalent circuit of Fig. 7-68. In this circuit

$$C_{eq} = a^2\left(C_s + C_g + C_{ws} + C_{ps}\frac{a-1}{a}\right) \qquad (7\text{-}133)$$

where C_{ws} is the stray and wiring capacitance to ground in the grid lead.

At this point it is best to subdivide the analysis as in the case of R-C-coupled amplifiers. The l-f range is that in which the shunting effect of M/a is enough to affect appreciably the output voltage. The h-f range covers those frequencies for which the capacitances and the leakage inductances have important effects. For well-designed transformers there is a middle range of frequencies for which the amplification remains essentially constant, because of negligible shunting by either M/a or the capacitances alone or because of an antiresonant condition of their combination.

In the middle range there is no current flow beyond a usually negligible core-loss current, and therefore no voltage is lost. Hence, approximately,

$$\frac{\mathbf{E}_2}{a} = -\mu \mathbf{V}_g \tag{7-134}$$

or

$$\mathbf{A}_{\text{mid}} = \frac{\mathbf{E}_2}{V_g} = -\mu a \tag{7-135}$$

The result given by (7-135) was to be expected since the tube is operating with infinite a-c load impedance, giving an amplification of μ, and the transformer multiplies this by the turns ratio.

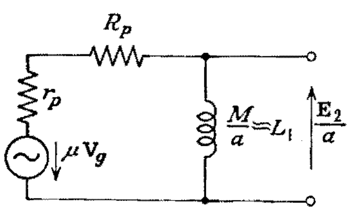


FIG. 7-69. Low-frequencye quivalent circuit for a transformer-coupled amplifier.

In the low range also, it is well to omit quantities having only secondary importance, such as L_p, which is small compared with M/a. We can also omit R_c, which is very high for well-designed transformers having adequate cores of good-quality steel. It is convenient also to replace M/a by the almost equal value L_1.* It follows from Fig. 7-69 that

$$\frac{\mathbf{E}_2}{a} = \frac{-j\omega L_1}{r_p + R_p + j\omega L_1} \mu \mathbf{V}_g \tag{7-136}$$

or that

$$\mathbf{A}_{\text{low}} = \frac{\mathbf{E}_2}{\mathbf{V}_g} = \frac{-\mu a}{1 - j[(r_p + R_p)/\omega L_1]} \tag{7-137}$$

An l-f marker f_1, which makes $\omega_1 L_1 = r_p + R_p$, can be used in helping to simplify this formula to

$$\frac{\mathbf{A}_{\text{low}}}{\mathbf{A}_{\text{mid}}} = \frac{1}{1 - j(f_1/f)} \tag{7-138}$$

where

$$f_1 = \frac{r_p + R_p}{2\pi L_1} \tag{7-139}$$

* Proof that $M/a \approx L_1$: $M = k\sqrt{L_1 L_2}$, where k, the coefficient of coupling, can be defined for our purpose as the fraction of the primary flux which cuts the secondary winding. $L_2 = a^2 L_1$, since the inductances of coils on an iron core are proportional to the square of the number of turns. Hence $a = \sqrt{L_2/L_1}$ and $M/a = kL_1$. For k near unity, as in this case, $M/a \approx L_1$.

Equation (7-138) is identical in form to Eq. (7-33) for the R-C-coupled amplifier.

The simplified circuit for high frequencies appears in Fig. 7-70, where C_p and the primary wiring capacitance have been neglected to make the analysis somewhat simpler without altering the validity significantly. New symbols R_{eq} and L_{eq} are defined in the diagram to aid in the simplification. Since L_{eq} and C_{eq} are in series, a resonant point will be reached at some frequency f_0 and the resulting current flow will tend to be high, being limited by R_{eq} only. The output voltage, being the drop across C_{eq}, will likewise tend to be somewhat high although not so much so as for frequencies somewhat below resonance where the reactance of C_{eq} is greater. Thus if R_{eq} is very low, we may expect the curve of gain to show a high peak at a frequency somewhat below the resonant point. If R_{eq} is greater, this peak should be less pronounced and for a large R_{eq} value the effect may be merely to keep the curve from falling so rapidly.

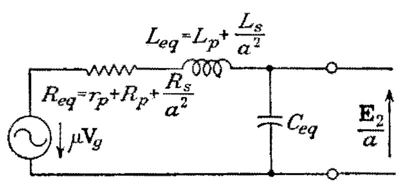


FIG. 7-70. High-frequency equivalent circuit for a transformer-coupled amplifier.

A more exact analysis may now be made. The circuit of Fig. 7-70 yields the relation

$$\frac{\mathbf{E}_2}{a} = \frac{-1/j\omega C_{eq}}{R_{eq} + j[\omega L_{eq} - (1/\omega C_{eq})]} \mu \mathbf{V}_g \tag{7-140}$$

or

$$\mathbf{A}_{\text{high}} = \frac{\mathbf{E}_2}{\mathbf{V}_g} = \frac{-\mu a}{\omega C_{eq}[(1/\omega C_{eq}) - \omega L_{eq}] + j\omega C_{eq} R_{eq}} \tag{7-141}$$

This complicated expression is considerably simplified by using a marker frequency f_0 defined by $\omega_0 L_{eq} = 1/\omega_0 C_{eq}$, or

$$f_0 = \frac{1}{2\pi \sqrt{L_{eq} C_{eq}}} \tag{7-142}$$

and a parameter Q_0 defined as

$$Q_0 = \frac{\omega_0 L_{eq}}{R_{eq}} \tag{7-143}$$

Then

$$\begin{aligned} \mathbf{A}_{\text{high}} &= \frac{\mathbf{A}_{\text{mid}}}{\omega C_{eq}[(\omega_0/\omega_0 \omega C_{eq}) - (\omega_0 \omega L_{eq}/\omega_0)] + j(\omega_0 \omega C_{eq} R_{eq}/\omega_0)} \\ &= \frac{\mathbf{A}_{\text{mid}}}{(f/f_0)[(f_0/f) - (f/f_0)] + j(f/f_0)(1/Q_0)} \end{aligned} \tag{7-144}$$

We may now calculate the response of the amplifier at several frequencies:

At $f = f_0$:

$$\mathbf{A} = Q_0\mathbf{A}_{\text{mid}}/\underline{90°}$$

At $f = 0.1f_0$:

$$\mathbf{A} = \frac{\mathbf{A}_{\text{mid}}}{0.1(10 - 0.1) + j(0.1/Q_0)} \approx \frac{\mathbf{A}_{\text{mid}}}{1 + j(1/10Q_0)}$$

At $f = 0.5f_0$:

$$\mathbf{A} \approx \frac{2\mathbf{A}_{\text{mid}}}{1.5 + (j/Q_0)}$$

At $f = 0.7f_0$:

$$\mathbf{A} \approx \frac{\mathbf{A}_{\text{mid}}}{0.51 + j(0.7/Q_0)}$$

At $f = 10f_0$:

$$\mathbf{A} \approx \frac{0.1\mathbf{A}_{\text{mid}}}{-10 + j(1/Q_0)}$$

The numerical value of gain in a normalized form is plotted against frequency in Fig. 7-71. Also the angle θ_A (for $Q_0 = 1$) is plotted as a function of frequency in Fig. 7-72.

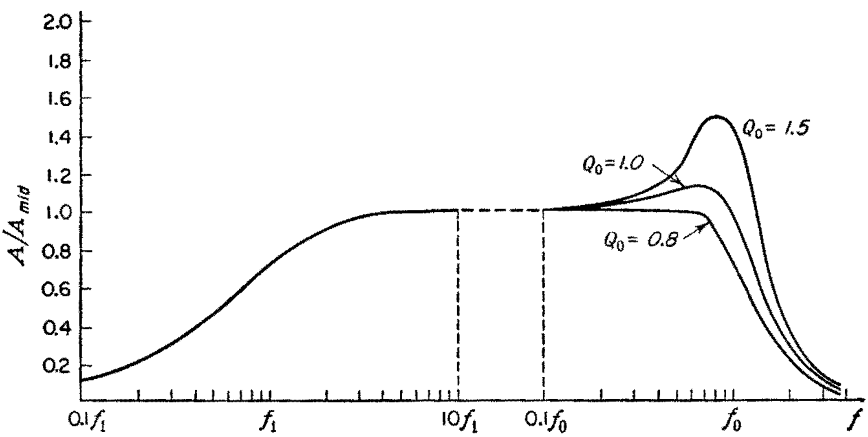


FIG. 7-71. Gain vs. frequency for a transformer-coupled amplifier.

Examination of the curve of gain in the h-f region shows that the general performance is not very uniform. It is usually satisfactory to assume that the range of constant gain extends from $10f_1$ to $0.1f_0$, except in cases where Q_0 is near 0.8, when $f = 0.7f_0$ may be used as the approximate upper limit. For $Q_0 = 0.8$ and $f = 0.7f_0$ one may readily show that $A = 0.98A_{\text{mid}}$.

The magnitudes and corresponding phase angles associated with the gain for various frequencies may be plotted in a polar system as was done in Arts. 7-12 and 7-13 in the case of R-C-coupled amplifiers. The result

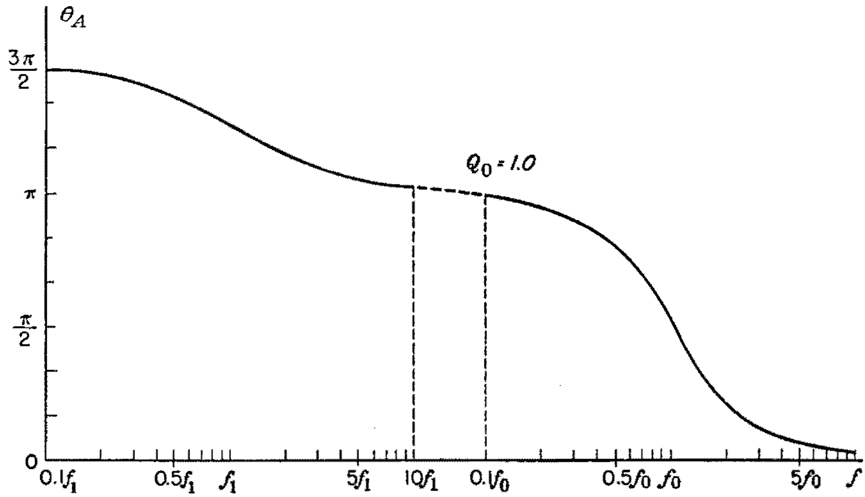


FIG. 7-72. Phase shift vs. frequency for a transformer-coupled amplifier.

in the case of $Q_0 = 1$ is drawn in Fig. 7-73. Such polar diagrams are useful in studies of feedback.

7-30. Parameters in Interstage-transformer Design. While the design of an interstage transformer is beyond the scope of this book, we can readily pick out what is needed for good performance. For high gain in the mid-frequency range, the formula $A = \mu a$ makes it appear that μ for the tube should be high and that the step-up turns ratio be high. But there are other important considerations. The mid-frequency bandwidth should generally be as wide as practicable, making it desirable to have f_1 quite low in value and f_0 high in value. Since $f_1 = (r_p + R_p)/2\pi L_1$, r_p and R_p should both be low in value. Examination of the curves of Fig. 7-71 shows that, for uniform gain over an extended range, a value of Q_0 near 0.8 is desirable. Now $Q_0 = \omega_0 L_{eq}/R_{eq} = \omega_0 L_{eq}/[r_p + R_p + (R_s/a^2)]$. Unless the designer uses especial care, Q_0 is likely to exceed 0.8, and hence it must be kept in mind to try to make L_{eq} small and the denominator large. As r_p and R_p should be low in value in the interests of l-f response, R_s should be made large.

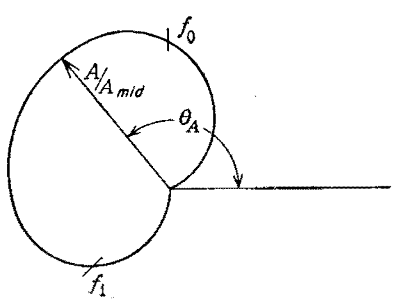


FIG. 7-73. Polar diagram for a transformer-coupled amplifier. $Q_0 = 1$.

Returning to μ and a, it now appears that only a medium value of μ can be used since r_p must be low. Let us see about the turns ratio a. A high value of L_1 usually means a large number of primary turns and, if a is large, many more secondary turns. Hence C_{eq} tends to be large, which is undesirable. Compromise is necessary, and a value between 3 and 4 for a is the general practice. The secondary is wound with small-gauge copper wire, which makes R_s high.

An important element in C_{eq} is C_{ps}. This capacitance can be reduced by shielding the primary from the secondary. To do this, a grounded copper or aluminum shield is placed between the primary and the secondary windings, taking care that the sheet is insulated in such a way that a shorted turn of winding is not formed by it.

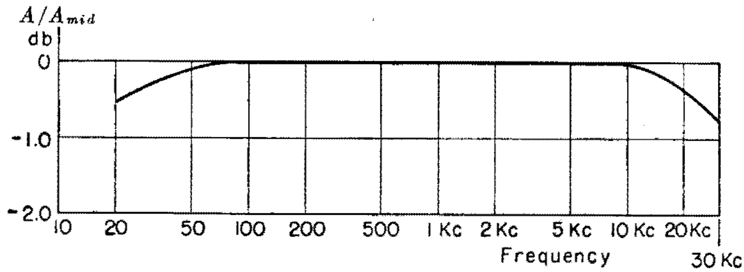


FIG. 7-74. Gain vs. frequency for a good-quality audio transformer.

The performance of a typical good-quality transformer is shown by the curve of A vs. f given in Fig. 7-74. With care in choice of core size and quality, coil turns and turns ratio, coil resistances and winding scheme, transformers may be made to perform with very uniform gain over a wide a-f band, when used with the proper tube. However, the cost of such a device is high compared with an R-C coupler, which can give equal or better results, especially when used with a pentode. Therefore, except for special purposes, interstage transformers are not often used.

7-31. Narrow-band Amplifiers.[15] The electromagnetic energy radiated by a radio transmitting antenna lies in a relatively narrow band of frequencies. For an a-m broadcast transmitter with a carrier of, say, around 1000 kc, the bandwidth is approximately 15 kc, and for an f-m transmitter with a carrier of, say, around 100 Mc, the bandwidth is approximately 150 kc. At the receiver there must be a separation of the desired from the undesired signals, and therefore it is an advantage to have an amplifier which amplifies over only a relatively narrow frequency band.

The ideal response curve desired is shown in Fig. 7-75. No amplifier previously studied in detail performs anything like this. In fact, the ordinary R-C-coupled amplifier has almost no output at frequencies greater than a few hundred kilocycles per second, because of the shunting effect of the tube and the wiring capacitances.

7-32. The Single-tuned-circuit R-F Amplifier. In considering an R-C-coupled amplifier the idea naturally occurs that it is possible to cancel the effect of the shunting capacitance by adding an inductor in parallel. Figure 7-76 shows a circuit based on this idea, in which R_b has been replaced by L. This gives the advantage of a lower d-c load as well as of canceling the effect of C at some certain high frequency. C may be increased above the stray circuit and tube values to help in the tuning, or, for very high frequencies, the inductor may be varied and only the stray and tube capacitances used. The amplification will be high at this antiresonant frequency f_0 of L and C and will decrease for either higher or lower frequencies. The coupling capacitor serves only to isolate the following grid from the high direct plate voltage and has no effect on the a-c action of the circuit. Norton's equivalent circuit for this amplifier is shown in Fig. 7-77. r_p, R_g, and R_{par} (the coil's a-c resistance) have been combined into R_{sh}. L and C represent pure reactance elements.

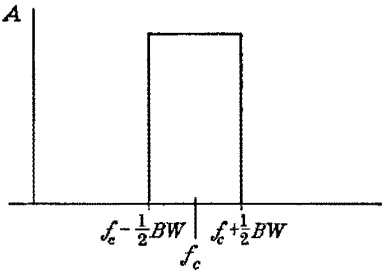


FIG. 7-75. The response of an r-f amplifier should be like this.

Equation (7-7) ($\mathbf{A} = -g_m\mathbf{Z}_{sh}$) may be used to compute the gain of this amplifier, and hence it is necessary only to determine the variation of $\mathbf{Z}_{sh}$ with frequency to find the amplifier's response. The admittance of the

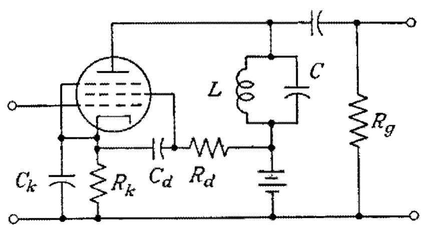


FIG. 7-76. An amplifier with a single-tuned load.

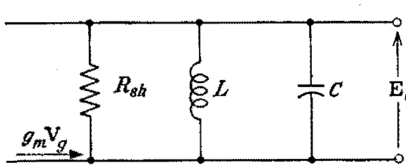


FIG. 7-77. Norton's equivalent circuit for the single-tuned amplifier.

load on the constant-current generator at the frequency f_0 is $1/R_{sh}$, and at any other frequency it may be stated as

$$\mathbf{Y}_{sh} = \frac{1}{R_{sh}} + j\left(\omega C - \frac{1}{\omega L}\right) \tag{7-145}$$

This equation may be put into a more useful form by factoring out $1/R_{sh}$, introducing ω_0/ω_0, and using the relation that $1/\omega_0 L = \omega_0 C$. Hence

$$\begin{aligned}\mathbf{Y}_{sh} &= \frac{1}{R_{sh}}\left[1 + j\left(\frac{R_{sh}\omega\omega_0 C}{\omega_0} - \frac{\omega_0 R_{sh}}{\omega\omega_0 L}\right)\right] \\ &= \frac{1}{R_{sh}}\left[1 + jR_{sh}\omega_0 C\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)\right] = \frac{1}{R_{sh}}\left[1 + jQ_0'\left(\frac{f}{f_0} - \frac{f_0}{f}\right)\right] \qquad (7\text{-}146)\end{aligned}$$

where

$$\omega_0 = \frac{1}{\sqrt{LC}} \qquad (7\text{-}147)$$

and

$$Q_0' = \frac{R_{sh}}{\omega_0 L} = R_{sh}\omega_0 C \qquad (7\text{-}148)$$

is a convenient parameter. (An explanation of Q_0 appears later in this article.) If δ is a real number, we may express any frequency in terms of the resonant frequency as

$$f = f_0 + \delta f_0 \qquad (7\text{-}149)$$

from which it follows that $\delta = (f - f_0)/f_0$, $f/f_0 = 1 + \delta$, and $(f/f_0) - (f_0/f) = (2 + \delta)/(1 + \delta)$. Substituting into Eq. (7-146), we obtain

$$\mathbf{Y}_{sh} = \frac{1}{R_{sh}}\left[1 + jQ_0'\delta\,\frac{2 + \delta}{1 + \delta}\right] \qquad (7\text{-}150)$$

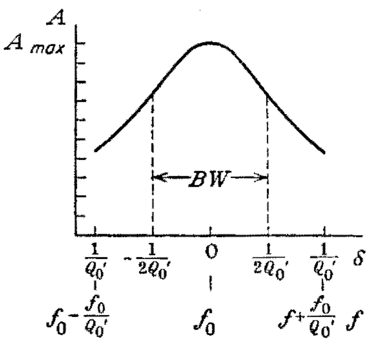


FIG. 7-78. Frequency response of a single-tuned amplifier.

or, since $(2 + \delta)/(1 + \delta) = 2 - \delta + \delta^2 - \delta^3 + \cdots \approx 2$, with an error less than δ if $\delta < 1$,

$$\mathbf{Y}_{sh} = \frac{1}{R_{sh}}(1 + j2Q_0'\delta) \qquad (7\text{-}151)$$

and

$$\mathbf{Z}_{sh} = \frac{1}{\mathbf{Y}_{sh}} = \frac{R_{sh}}{1 + j2Q_0'\delta} = \frac{R_{sh}}{\sqrt{1 + 4Q_0'^2\delta^2}}\,\underline{/-\arctan 2Q_0'\delta} \qquad (7\text{-}152)$$

It now follows that

$$\mathbf{A} = \frac{-g_m R_{sh}}{1 + j2Q_0'\delta} = \frac{\mathbf{A}_{\max}}{1 + j2Q_0'\delta} \qquad (7\text{-}153)$$

where $\mathbf{A}_{\max} = -g_m R_{sh}$ is the value of $\mathbf{A}$ for $\delta = 0$, that is, at $f = f_0$.

We may now plot the frequency-response curve for the single-tuned amplifier. This is done in Fig. 7-78. As in the case of the R-C-coupled amplifier, the half-power frequencies may be called f_1 and f_2. Since obviously there is no extended region of uniform response, it is generally convenient to consider the bandwidth in this case to extend from f_1 to f_2.

$$\text{BW} = 2 \times \frac{1}{2Q_0'} f_0 = \frac{1}{Q_0'} f_0 = \frac{1}{R_{sh}\omega_0 C} f_0 = \frac{1}{2\pi R_{sh} C} \qquad (7\text{-}154)$$

It is not possible to obtain an inductance by using a coil without resistance, and it is hence necessary to consider this resistance to be included in R_{sh}. The resistance accounting for copper loss in a coil is actually a series quantity, but for high frequencies much of the resistance is due to losses caused by induced voltages in shields, dielectrics, and cores. This may be considered as a parallel resistor. As a measurement usually determines the impedance of a coil on a series basis, we may transform as follows, using symbols shown in Fig. 7-79: At any frequency the admittances of the two circuits are equal.

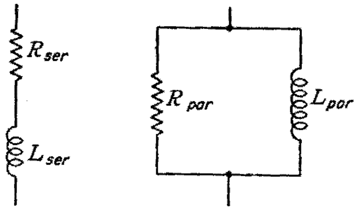


FIG. 7-79. Series and parallel equivalents.

$$\mathbf{Y} = \frac{1}{R_{ser} + j\omega L_{ser}} = \frac{R_{ser} - j\omega L_{ser}}{R_{ser}^2 + \omega^2 L_{ser}^2} = \frac{1}{R_{par}} - \frac{j}{\omega L_{par}} \qquad (7\text{-}155)$$

Hence

$$R_{par} = \frac{R_{ser}^2 + \omega^2 L_{ser}^2}{R_{ser}} \qquad (7\text{-}156)$$

$$\omega L_{par} = \frac{R_{ser}^2 + \omega^2 L_{ser}^2}{\omega L_{ser}} \qquad (7\text{-}157)$$

If an alternating voltage is applied to either of the circuits of Fig. 7-79, the total current which flows differs in phase from the voltage by an angle θ. The cosine of this angle, called the power factor, is a useful parameter. The tangent of θ is also often useful and is symbolized by Q. In the parallel-tuned circuit, an applied voltage E causes a current E/R_{par}, in phase with the voltage, in the R branch and a current $E/\omega L_{par}$, lagging the voltage by 90°, in the L branch. The value of $Q = \tan\theta$ would be $(E/\omega L_{par})/(E/R_{par}) = R_{par}/\omega L_{par}$. The same voltage applied to the series circuit would make $Q = \tan\theta = \omega L_{ser}/R_{ser}$. By substituting Eqs. (7-156) and (7-157) into the definition of the Q of a parallel circuit, we may verify the relation

$$Q = \frac{R_{par}}{\omega L_{par}} = \frac{\omega L_{ser}}{R_{ser}} \qquad (7\text{-}158)$$

Thus the value of Q obtained from either representation is the same.

If $R_{ser} \ll \omega L_{ser}$, that is, if Q has a large value,

$$R_{par} \approx \frac{(\omega L_{ser})^2}{R_{ser}} = Q^2 R_{ser} = Q\omega L_{ser} \tag{7-159}$$

and

$$L_{par} \approx L_{ser} \tag{7-160}$$

The percentage error in saying $L_{par} = L_{ser}$ is only $100/Q^2$ for ordinary values of Q. Of the above relations probably the most useful one is

$$R_{par} = Q_0\omega_0 L = Q_0\sqrt{\frac{L}{C}} \tag{7-161}$$

for this gives the impedance of a parallel-tuned circuit at the resonant frequency.

Returning to the discussion of the single-tuned amplifier, the value of Q_0' may be expressed in a different form by using the resonant condition and equating admittances,

$$\frac{1}{R_{sh}} = \frac{1}{r_p} + \frac{1}{R_g} + \frac{1}{Q_0\omega_0 L} \tag{7-162}$$

or, from Eq. (7-148),

$$\frac{1}{Q_0'\omega_0 L} = \frac{1}{r_p} + \frac{1}{R_g} + \frac{1}{Q_0\omega_0 L} \tag{7-163}$$

It should be noted that Q_0 belongs to the coil alone, while Q_0' belongs to the whole circuit. We may solve for Q_0'.

$$Q_0' = \frac{1}{\omega_0 L\left(\dfrac{1}{r_p} + \dfrac{1}{R_g} + \dfrac{1}{Q_0\omega_0 L}\right)} = \frac{1}{\dfrac{\omega_0 L}{r_p} + \dfrac{\omega_0 L}{R_g} + \dfrac{1}{Q_0}} \tag{7-164}$$

This formula is not particularly needed to calculate the value of Q_0' (see Example 1 of Art. 7-33), but it is useful in the following consideration: If, with a fixed f_0 value, either L or Q_0 of the coil is raised, R_{par} is thereby increased as shown by Eq. (7-161). Hence R_{sh} is increased, and the amplification A_{max} is made greater. If this is done by raising the value of L and keeping Q_0 fixed, formula (7-164) shows that Q_0' decreases and hence the bandwidth increases. But if L is fixed in value and Q_0 is increased, (7-164) shows that Q_0' increases and the tuning is sharpened by the decreased bandwidth. Hence in designing an amplifier it is seen that bandwidth as well as gain can be planned for.

7-33. Examples of Single-tuned Amplifiers

1. Let us use a 6SK7 tube in the circuit of Fig. 7-76. Using values recommended by a tube manual, assume $E_{bb} = 250$ volts, $E_{c1o} = -3$

volts, $E_{c2o} = 100$ volts, $E_{c3o} = 0$ volt. Use $R_g = 0.5$ megohm. The coupling capacitor can be 0.001 μf. For this academic problem assume f_0 to be 796 kc, since ω_0 then equals 5×10^6 radians per sec, which makes the arithmetic easy. A coil having $L = 200$ μh and $Q_0 = 150$ is available. The required capacitance C, as given by $1/\omega_0{}^2L$, is 200 $\mu\mu$f, which is easily obtained by adjusting a 360-$\mu\mu$f (maximum) variable capacitor. A tube manual yields the information that $r_p = 0.8$ megohm, $g_m = 2000$ micromhos, $I_{bo} = 9.2$ ma, $I_{c2o} = 2.6$ ma. The required R_k resistor is $3/11.8 \times 10^{-3} = 254$ ohms. A value of $R_d = (250 - 100)/(2.6 \times 10^{-3}) \approx 58{,}000$ ohms is required. The bypass capacitors C_k and C_d should have small impedances compared with R_k and R_d, respectively. For C_k a value of 0.01 μf having $X_c = 20$ ohms is adequate, while, for C_d, 0.001 μf with 200 ohms is ample. The Q_0 for the coil being 150, the equivalent shunting resistor for the idealized coil may be computed as $R_{\text{par}} = Q_0\omega_0L = 150 \times 5 \times 10^6 \times 200 \times 10^{-6} = 150 \times 1000 = 150{,}000$ ohms, and this is the impedance of the parallel combination of coil and capacitor at the antiresonant frequency. R_{sh} of Fig. 7-77 being R_{par} in parallel with r_p and R_g, it may be computed as $(1/0.150) + (1/0.8) + (1/0.5) = 6.67 + 1.25 + 2.0 = 9.92$ micromhos, or $R_{sh} = 1/9.92 = 0.101$ megohm $=$ 101,000 ohms. Hence $A_{\max} = g_mR_{sh} = 2000 \times 10^{-6} \times 101{,}000 = 202$. The value of Q_0' for the circuit determines the bandwidth (the prime has been appended to avoid confusion with the Q_0 of the coil alone), and it may be determined as $Q_0' = R_{sh}/\omega_0L = 101{,}000/1000 = 101$. Note that this value is lower than that of the coil alone because of the energy-absorbing qualities of r_p and R_g. The bandwidth BW $= (1/Q_0')f_0 = \frac{1}{101} \times 796 = 7.88$ kc. The response curve is that of Fig. 7-78 with proper numerical values inserted.

2. For a 6SK7 tube with the operating voltages, resistors, and frequency of Example 1, what value of L should be used to give a bandwidth of 10 kc and what gain can be obtained? Coils with $Q_0 = 150$ are available.

Solution. At 796 kc,

$$\frac{1}{r_p} + \frac{1}{R_g} + \frac{1}{Q_0\omega_0L} = \frac{1}{R_{sh}} \qquad \text{and} \qquad Q_0' = \frac{R_{sh}}{\omega_0L}$$

Therefore

$$\frac{1}{r_p} + \frac{1}{R_g} + \frac{Q_0'}{Q_0R_{sh}} = \frac{1}{R_{sh}} \qquad \text{or} \qquad R_{sh} = \frac{1 - (Q_0'/Q_0)}{(1/r_p) + (1/R_g)}$$

Since BW $= f_0/Q_0'$, $Q_0' = (796 \times 10^3)/10{,}000 = 79.6$. Therefore

$$R_{sh} = \frac{1 - (79.6/150)}{(1/0.8) + (1/0.5)} = \frac{1 - 0.531}{1.25 + 2.0} = 0.144 \text{ megohm}$$

The gain obtainable is $A = g_m R_{sh} = 2000 \times 10^{-6} \times 0.144 \times 10^6 = 288$. $\omega_0 L = R_{sh}/Q_0' = 0.144 \times 10^6/79.6 = 1810$ ohms, and $L = 1810/(5 \times 10^6) = 362\ \mu$h.

7-34. Other Types of R-F Voltage Amplifiers. The type of amplifier studied in Arts. 7-32 and 7-33 has the disadvantage in practice that, if C is made variable to adjust the tuning, both sets of plates must be isolated from ground and they are at high d-c potential above ground, making it dangerous to use in exposed positions. While the latter disadvantage may be circumvented, other types of circuits are usually favored, it being easier to maintain high effective Q_0 values with them.

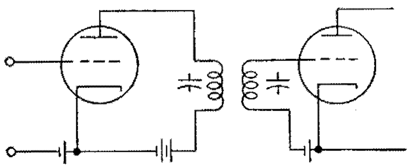


FIG. 7-80. Transformer coupling may be used.

We may use the idea of transformer coupling, and Fig. 7-80 shows an elementary circuit of this type. Because of core losses it is often impractical to use an iron core for the transformer, and air is ordinarily used for the greater part of the flux path, although powdered-iron slugs are sometimes employed to give high inductance with fewer turns. Because of the small inductance of the coils, the load on the first tube is too low to give much voltage output if no tuning is employed. Either the primary or the secondary or both may be tuned. The tuned-secondary type is widely used. If both coils are tuned, the tuning process is so complicated and critical that the capacitor adjustments are usually semipermanent and the amplifier is used for a fixed r-f band only.

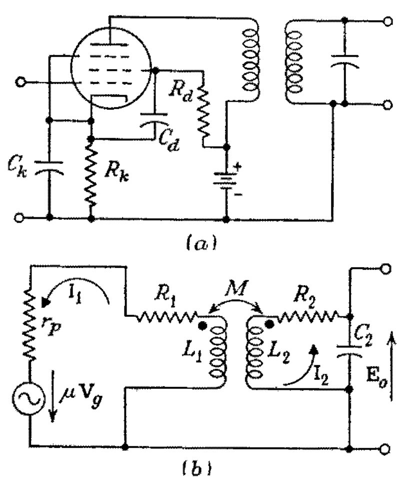


FIG. 7-81. (a) Tuned-secondary r-f amplifier. (b) Equivalent circuit of (a).

7-35. Tuned-secondary Type of R-F Amplifier. The circuit is shown in both schematic and equivalent forms in Fig. 7-81. In order to simplify the analysis, the stray capacitance in the primary and also the capacitance between primary and secondary have been neglected. At very high frequencies this would result in serious error. The transformer polarity is not important, but in order to make our analysis definite, it is assumed to be as shown by the dots. We can compute the secondary current, and it is obvious that the voltage drop in C_2 will yield the output voltage. Hence we shall first compute $\mathbf{I}_2$.

As a preliminary step let us consider the simple coupled circuit represented in Fig. 7-82. The ratio $\mathbf{E}_1/\mathbf{I}_1$, called the driving-point impedance,

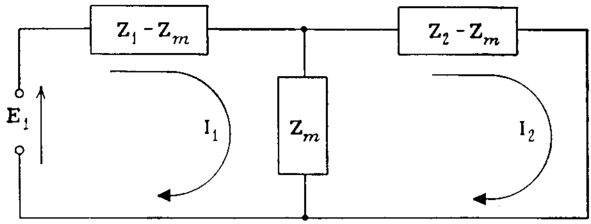


Fig. 7-82. Equivalent circuit of a simple coupled circuit.

can be shown to be

$$\mathbf{Z}_1' = \frac{\mathbf{Z}_1\mathbf{Z}_2 - \mathbf{Z}_m^2}{\mathbf{Z}_2} = \mathbf{Z}_1 - \frac{\mathbf{Z}_m^2}{\mathbf{Z}_2} \tag{7-165}$$

The ratio $\mathbf{E}_1/\mathbf{I}_2$, called the transfer impedance, is

$$\mathbf{Z}_{12} = \frac{\mathbf{Z}_1\mathbf{Z}_2 - \mathbf{Z}_m^2}{\mathbf{Z}_m} \tag{7-166}$$

From the equivalent circuit of Fig. 7-81

$$\mathbf{Z}_1 = r_p + R_1 + j\omega L_1 \tag{7-167}$$

$$\mathbf{Z}_2 = R_2 + j\left(\omega L_2 - \frac{1}{\omega C_2}\right) \tag{7-168}$$

$$\mathbf{Z}_m = j\omega M \tag{7-169}$$

Also

$$\mathbf{I}_2 = \frac{\mu \mathbf{V}_g}{\mathbf{Z}_{12}} \tag{7-170}$$

and

$$\mathbf{E}_o = \frac{-\mathbf{I}_2}{j\omega C_2} \tag{7-171}$$

Hence

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{V}_g} = \frac{-\mu}{\mathbf{Z}_{12} j\omega C_2} \tag{7-172}$$

where

$$\mathbf{Z}_{12} = \frac{(r_p + R_1 + j\omega L_1)[R_2 + j(\omega L_2 - 1/\omega C_2)] + \omega^2 M^2}{j\omega M} \tag{7-173}$$

Substituting,

$$\mathbf{A} = \frac{-\mu M/C_2}{(r_p + R_1 + j\omega L_1)[R_2 + j(\omega L_2 - 1/\omega C_2)] + \omega^2 M^2} \tag{7-174}$$

We would like to have a graph of A vs. f for f near f_0. Equation (7-174) looks formidable. It would be desirable to think of some simplifying approximations, since great accuracy is not a consideration in an affair like this. Also a few simplifying substitutions often help in understanding what is going on.

Let us list our assumptions as we progress:

1. $r_p \gg |R_1 + j\omega L_1|$. Hence

$$\mathbf{A} = \frac{-g_m M/C_2}{R_2 + (\omega^2 M^2/r_p) + j[\omega L_2 - (1/\omega C_2)]} \tag{7-175}$$

2. Let $R_2' = R_2 + (\omega^2 M^2/r_p)$ as a convenient substitution.

3. Also introduce ω_0, defined by $\omega_0 L_2 = 1/\omega_0 C_2$. Then

$$\mathbf{A} = \frac{-g_m\omega_0 M/R_2'\omega_0 C_2}{1 + j(1/R_2')[(\omega\omega_0 L_2/\omega_0) - (\omega_0/\omega\omega_0 C_2)]} \tag{7-176}$$

or

$$\mathbf{A} = \frac{-g_m\omega_0 M/R_2'\omega_0 C_2}{1 + j(\omega_0 L_2/R_2')[(\omega/\omega_0) - (\omega_0/\omega)]} \tag{7-177}$$

4. Let $Q_2' = \omega_0 L_2/R_2' = 1/R_2'\omega_0 C_2$.

$$\mathbf{A} = \frac{-g_m\omega_0 M Q_2'}{1 + jQ_2'[(f/f_0) - (f_0/f)]} \tag{7-178}$$

5. Let $f = f_0 + \delta f_0$, as was done in Eq. (7-149). Then $(f/f_0) - (f_0/f)$ reduces to $\delta[(2 + \delta)/(1 + \delta)]$, which is approximately 2δ if $\delta \ll 1$. Hence

$$\mathbf{A} = \frac{-g_m\omega_0 M Q_2'}{1 + jQ_2' 2\delta} \tag{7-179}$$

6. When $f = f_0$, $\delta = 0$. For this frequency

$$\mathbf{A} = \mathbf{A}_{\max} = -g_m\omega_0 M Q_2' \tag{7-180}$$

Thus we have for the amplification for any frequency near the resonant value

$$\mathbf{A} = \frac{\mathbf{A}_{\max}}{1 + j2Q_2'\delta} \tag{7-181}$$

This equation is identical in form to Eq. (7-153) for the previous case.

The bandwidth is given by

$$\mathrm{BW} = \frac{f_0}{Q_2'} \tag{7-182}$$

Note also the close similarity between the formulas for $A_{\max}$ and for BW and those of the former case of Art. 7-32. The curve of A vs. f may now be drawn, it being identical with that of Fig. 7-78 if proper numerical values are attached.

The value of A_{max} is dependent upon M, which is a function of the coefficient of coupling. We shall now investigate this fact to discover what coupling is best to use to obtain high gain. Since the value of Q_2' depends upon M, we may write

$$A_{max} = g_m Q_2' \omega_0 M = \frac{g_m \omega_0 M \omega_0 L_2}{R_2 + (\omega_0^2 M^2 / r_p)} = \frac{KM}{R_2 + (\omega_0^2 M^2 / r_p)} \tag{7-183}$$

where K is independent of M. If we set $dA_{max}/dM = 0$, we may be able to find the value of M which will make A_{max} the greatest. Doing this, it follows that

$$\left[R_2 + \frac{\omega_0^2 M^2}{r_p}\right] K - KM \frac{2\omega_0^2 M}{r_p} = 0 \tag{7-184}$$

or the optimum value of mutual inductance is

$$M_{opt} = \frac{\sqrt{R_2 r_p}}{\omega_0} \tag{7-185}$$

The coefficient of coupling k, which equals $M/\sqrt{L_1 L_2}$, has an optimum value

$$k_{opt} = \frac{\sqrt{r_p R_2}}{\omega_0 \sqrt{L_1 L_2}} = \sqrt{\frac{r_p}{\omega_0 L_1} \frac{R_2}{\omega_0 L_2}} = \frac{1}{\sqrt{Q_1 Q_2}} \tag{7-186}$$

where Q_1 is the Q_0 of the primary circuit and Q_2 is that of the isolated secondary circuit. Under optimum coupling conditions the load seen by the tube is, by Eq. (7-165), $\omega^2 M^2 / R_2 = \omega_0^2 R_2 r_p / R_2 \omega_0^2 = r_p$, which means that the conditions for maximum power transfer are satisfied.

For pentodes Q_1 ($\approx \omega_0 L_1 / r_p$) is usually so small that it is impossible to obtain the optimum value of coupling. This is not a matter of great importance, however, since the gain is quite satisfactorily high anyway, and too high gain per stage makes the problem of stray feedback troublesome. For triodes, which were formerly used in the ante-pentode days, and still are in high-power amplifiers, Q_1 is higher, and the meager gain is enhanced by careful adjustment of k. The bandwidth is somewhat dependent upon k, but for pentodes to only a slight extent. These points are perhaps made clearer by the following example.

7-36. An Example of a Tuned-secondary R-F Amplifier. For the circuit of Fig. 7-81a let the tube be a type 6SK7 with the same operating conditions as in Example 1 of Art. 7-33. For L_1 assume 50 μh with a Q_0 of 150 and for the secondary coil $L_2 = 200$ μh, $Q_0 = 150$, tuned by a capacitance C_2 to $f_0 = 796$ kc. Recall that $r_p = 0.8$ megohm, $g_m = 2000$ micromhos.

$Q_1 = \omega_0 L_1/r_p = 5 \times 10^6 \times 50 \times 10^{-6}/0.8 \times 10^6 = 312 \times 10^{-6}$; $Q_2 = 150$. Hence $k_{\text{opt}} = 1/\sqrt{Q_1 Q_2} = 4.6$. Since it is impossible to obtain this value and difficult even to approach unity for air-cored coils, let us assume that $k = 0.6$ is attainable. Then $M = k\sqrt{L_1 L_2} = 0.6\sqrt{50 \times 200} = 60\ \mu$h.

$$R_2' = R_2 + \frac{\omega_0{}^2 M^2}{r_p} = \frac{\omega_0 L_2}{Q_2} + \frac{\omega_0{}^2 M^2}{r_p} = \frac{5 \times 10^6 \times 200 \times 10^{-6}}{150} + \frac{(5 \times 10^6 \times 60 \times 10^{-6})^2}{0.8 \times 10^6} = 6.67 + 0.11 = 6.78 \text{ ohms}$$

which is practically the same as R_2. Hence $Q_2' = \omega_0 L_2/R_2' = 150$ approximately. $A_{\max} = g_m Q_2' \omega_0 M = 2000 \times 10^{-6} \times 150 \times 5 \times 10^6 \times 60 \times 10^{-6} = 90$. BW $= f_0/Q_2' = {}^{796}\!/_{150} = 5.3$ kc. As this bandwidth is likely to be too narrow, a secondary coil having a lower Q_0 should be chosen. As a result the gain will be lower. Changing k is futile, for even if k could be made unity, the bandwidth would be widened very little, as may be seen by recomputing Q_2'. Making the L_2/C_2 ratio higher does not have the effect of making the bandwidth greater as it did in the case of the amplifier of Art. 7-32.

7-37. Double-tuned Transformer-coupled R-F Amplifiers. In Fig. 7-80 if both the primary and the secondary are tuned, the frequency response of the amplifier may be quite different from that of the two circuits previously studied. It is not difficult to see that this is probably the case. Suppose the primary and secondary to be initially far apart and each tuned to a frequency f_c. The coils are then moved closer together so that a small degree of coupling exists between them. With the circuit in operation there will appear a voltage at the output. As the frequency of the signal is varied through the value f_c, the primary current, the secondary induced voltage, the secondary current, and the output voltage all pass through a peak at f_c. With the coils moved much closer together, impedance from the secondary considerably alters the equivalent primary impedance. For some frequency higher than f_c, the secondary circuit is inductive and hence the impedance coupled into the primary is capacitive [see formula (7-165)]. This alters the primary impedance to make it less inductive, thus raising the frequency for maximum response. The same explanation can be made for the action at some frequency lower than f_c. Thus two peaks in the frequency-response curve appear, there being a dip between them. Figure 7-85 shows three performance curves. For $a = 0.5$ the coupling is loose. For $a = 2$ the coupling is close, and two humps appear.

Thorough theoretical and design treatments of this type of amplifier are available in the literature, but an elementary study is not particularly difficult and will be made here. Referring to Fig. 7-80 we shall assume that the only coupling between primary and secondary is inductive (*i.e.*, there is no capacitive coupling between the coils) and that the two coils are identical. It is also assumed that no current is drawn by the following stage.

As a preliminary step let us find the equivalent Π for the two coupled coils (see Fig. 7-83). Most students find it easier to find the T first. As these are fairly familiar operations only the results are given here. Norton's equivalent circuit for this version of the amplifier of Fig.7-80 is shown in Fig. 7-84*a*, and its new form after replacing the simple coupled circuit by a Π is shown in Fig. 7-84*b*. This again is shown in a simplified form in Fig. 7-84*c*. A pentode tube is being used, and its r_p is assumed to be infinite. This gives the circuit an advantage of symmetry and makes $R = R_{\text{par}} = Q_0\omega_0 L$.

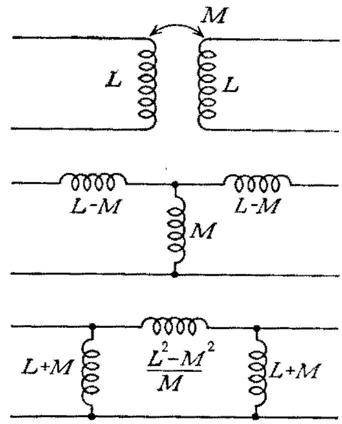


FIG. 7-83. Two coupled coils and their T and Π equivalents.

The impedance of the parallel combination of R, C, and $L + M$ is labeled $\mathbf{Z}_1$. That of the arch value $(L^2 - M^2)/M$ is called $\mathbf{Z}_2$. It should be easy to see that

$$\mathbf{I}_2 = g_m\mathbf{V}_g \frac{\mathbf{Z}_1}{2\mathbf{Z}_1 + \mathbf{Z}_2} \quad (7\text{-}187)$$

and that

$$\mathbf{E}_o = -\mathbf{I}_2\mathbf{Z}_1 = -g_m\mathbf{V}_g \frac{\mathbf{Z}_1{}^2}{2\mathbf{Z}_1 + \mathbf{Z}_2} \quad (7\text{-}188)$$

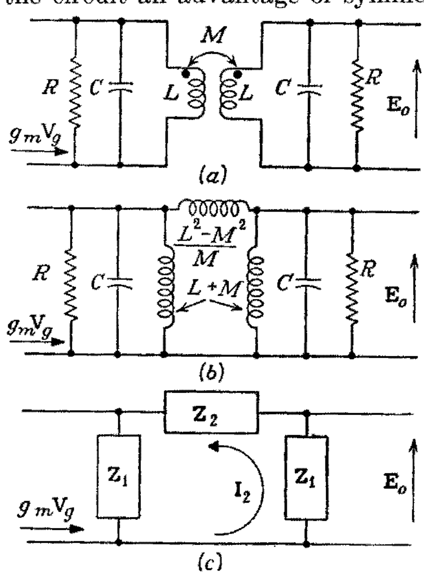


FIG. 7-84. Double-tuned r-f amplifiers. (*a*) Norton's equivalent circuit, (*b*) the circuit in Π form, (*c*) a simplified-version of the Π form.

Here $\mathbf{Z}_1$ is the impedance of R, C, and $L + M$ in parallel. At the antiresonant frequency f_0 of C and $L + M$, its value is R, while for nearby frequencies the value is given by adapting Eq. (7-152); thus

$$\mathbf{Z}_1 = \frac{R}{1 + j2Q_0\delta} \quad (7\text{-}189)$$

where δ is the fractional deviation in frequency from f_0. It should be noted that f_0 is not the resonant frequency of the isolated primary, as the mutual coupling increases the effective self-inductance by an amount M.

We may express $\mathbf{Z}_2$ as

$$\mathbf{Z}_2 = j\omega \frac{L^2 - M^2}{M} = jR\omega \frac{L^2 - M^2}{MR} = \frac{jR}{a} \tag{7-190}$$

where a is a useful parameter,

$$a = \frac{MR}{\omega(L^2 - M^2)} = \frac{kLR}{\omega L^2(1 - k^2)} = Q \frac{k}{1 - k^2} \approx Q_0 \frac{k}{1 - k^2} \tag{7-191}$$

For frequencies near f_0, Q is approximately constant at Q_0, and hence $\mathbf{Z}_2$ is also approximately constant.

Substituting these formulas for $\mathbf{Z}_1$ and $\mathbf{Z}_2$ into Eq. (7-188) yields

$$\mathbf{E}_o = -g_m\mathbf{V}_g \frac{\left(\dfrac{R}{1 + j2Q_0\delta}\right)^2}{\dfrac{2R}{1 + j2Q_0\delta} + j\dfrac{R}{a}} = -g_m\mathbf{V}_g \frac{Ra}{(1 + j2Q_0\delta)(2a + j - 2Q_0\delta)} \tag{7-192}$$

and

$$\mathbf{A} = \frac{\mathbf{E}_o}{\mathbf{V}_g} = \frac{g_m jRa}{(1 + j2Q_0\delta)[1 + j(2Q_0\delta - 2a)]} \tag{7-193}$$

The locus of this equation has symmetry about an axis whose location depends upon the value of the parameter a. In order to change this equation to a better form for plotting, obvious symmetry is desirable. The symmetry is made more apparent if one lets

$$2Q_0\delta = y + a \tag{7-194}$$

Then

$$\mathbf{A} = \frac{g_m jRa}{[1 + j(y + a)][1 + j(y - a)]} \tag{7-195}$$

and

$$A = \frac{g_m Ra}{\sqrt{[1 + (y + a)^2][1 + (y - a)^2]}} = \frac{g_m Ra}{\sqrt{y^4 + (2 - 2a^2)y^2 + a^4 + 2a^2 + 1}} \tag{7-196}$$

Note that y occurs in even powers only, and hence the graph of A from (7-196) has symmetry about the value where $y = 0$, that is, where $\delta = a/2Q_0$. Because of this symmetry A must have a maximum or minimum value at $y = 0$, and there may be other zero-slope values besides. To investigate this, let $dD^2/dy = 0$, where D is the denominator in Eq. (7-196). Then

$$4y^3 + 2(2 - 2a^2)y = 0 \tag{7-197}$$

or

$$y = 0, \pm \sqrt{a^2 - 1} \tag{7-198}$$

If $a = 1$, the zero-slope value is given by a triple root $y = 0$, which shows the curve of A vs. f to be very flat near this point. The coefficient of coupling k may be evaluated by solving Eq. (7-191) for k,

$$k = \frac{-Q_0 + \sqrt{Q_0^2 + 4a^2}}{2a} \tag{7-199}$$

which becomes, for $a = 1$,

$$k = \frac{-Q_0 + \sqrt{Q_0^2 + 4}}{2} \approx \frac{1}{Q_0} \tag{7-200}$$

This is very loose coupling for the usual Q_0 value. The value of A for $y = 0$ is given by substitution into Eq. (7-196) as

$$A = A_{\max} = \frac{g_m R}{2} \tag{7-201}$$

$a = 1$ and $k = 1/Q_0$ are called transitional values since $a < 1$ gives one maximum value, while $a > 1$ gives two maximum values, as will be seen. They are also called critical values as they are the lowest values which make the amplification the greatest obtainable.

For the case when $a = 1$, the bandwidth may be obtained by setting the expression for A in (7-196) equal to $\frac{g_m R}{2} \frac{1}{\sqrt{2}}$ and solving for y, yielding the value $y = \pm \sqrt{2}$. Then for $y = +\sqrt{2}$, $\delta = (1 + \sqrt{2})/2Q_0$, while for $y = -\sqrt{2}$, $\delta = (1 - \sqrt{2})/2Q_0$. Hence $f_1 = f_0 + [(1 - \sqrt{2})/2Q_0]f_0$, and $f_2 = f_0 + [(1 + \sqrt{2})/2Q_0]f_0$. We may now solve for the bandwidth as $f_2 - f_1$.

$$\text{BW} = \frac{\sqrt{2}\, f_0}{Q_0} \tag{7-202}$$

Since the bandwidth for the isolated Z_1 circuit is f_0/Q_0, it follows that for these conditions,

$$\text{BW} = \sqrt{2}\, \text{BW}_{z_1} \tag{7-203}$$

For $a < 1$ the only real zero-slope value is $y = 0$, and Eq. (7-196) then yields

$$A_{\max} = \frac{g_m R a}{a^2 + 1} \tag{7-204}$$

For $a > 1$, $y = \pm \sqrt{a^2 - 1}$ are also real values, and for these points

$$A = A_{\max} = \frac{g_m R}{2} \tag{7-205}$$

which is independent of a. At $y = 0$ formula (7-204) is again obtained, but here it is a minimum value;

$$A_{\min} = \frac{g_m Ra}{a^2 + 1} \qquad (7\text{-}206)$$

The half-power points occur at $y = \pm\sqrt{a^2 + 2a - 1}$, and hence the bandwidth is

$$\mathrm{BW} = \sqrt{a^2 + 2a - 1}\,\mathrm{BW}_{z_1} \qquad (7\text{-}207)$$

Typical graphs of A vs. y are shown in Fig. 7-85. The amplification is normalized relative to the value $g_m R/2$.

The center frequency f_c occurs where $y = 0$ and hence, by Eq. (7-194), where $\delta = a/2Q_0$. Hence

$$f_c = f_0(1 + \delta) = f_0\left(1 + \frac{a}{2Q_0}\right) \qquad (7\text{-}208)$$

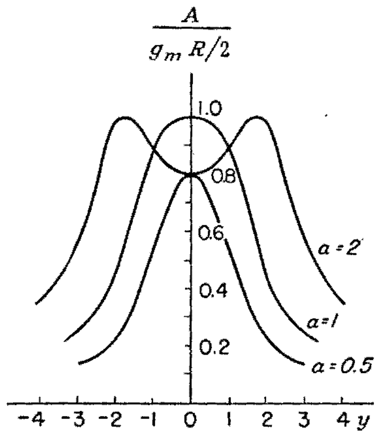


FIG. 7-85. Frequency response of a double-tuned amplifier.

This expression may be simplified. If we use Eq. (7-191) and assume $k \ll 1$,

$$\begin{aligned} f_c &= \frac{1}{2\pi\sqrt{(L+M)C}}\left[1 + \frac{k}{2(1-k^2)}\right] \\ &= \frac{1}{2\pi\sqrt{LC}\sqrt{1+(M/L)}}\left[1 + \frac{k}{2(1-k^2)}\right] \\ &= \frac{1}{2\pi\sqrt{LC}\sqrt{1+k}}\left[1 + \frac{k}{2(1-k^2)}\right] \approx \frac{1}{2\pi\sqrt{LC}\,[1 + (k/2)]}\left(1 + \frac{k}{2}\right) \\ &= \frac{1}{2\pi\sqrt{LC}} \end{aligned} \qquad (7\text{-}209)$$

Thus f_c equals the resonant frequency of the isolated primary or secondary. This approximation is very good for coupling near the critical value and with ordinary coils.

Amplifiers of the double-tuned type are almost universally used in the i-f stages of superheterodyne receivers, where by careful adjustment a response curve of suitable height and width and with steep sides may be obtained. The value of a is usually chosen slightly greater than unity, and the curve is then a fair approximation to the ideal of Fig. 7-75. For most i-f transformers the manufacturer permanently adjusts the two coils to this condition. The tuning of the coils, performed by adjusting the capacitors or the inductors, allows the technician to obtain the correct center frequency. With less than critical coupling it may be shown that stagger tuning, meaning the tuning of the primary to a frequency somewhat below center and the secondary to a frequency slightly above center, yields a characteristic somewhat resembling that of the more than critically coupled case, and hence the tuning adjustments allow a certain amount of bandwidth adjustment.

7-38. An Example of a Double-tuned Amplifier. A type 6SK7 tube is used in a double-tuned r-f amplifier. At the selected operating point $g_m = 2000$ micromhos, $r_p = 0.8$ megohm. The transformer has identical primary and secondary coils with $L = 200\ \mu$h, $Q_0 = 100$. The mutual inductance M is 4 μh. The tuning capacitors are equal, and $C = 200\ \mu\mu$f. The secondary load impedance may be considered infinite. (Usually the output i-f transformer in a receiver feeds into a diode which draws some current. The foregoing analysis is then somewhat in error.)

Since r_p is very high, its effect will be ignored. We shall find the bandwidth and the gain. The computations are

$$k = \frac{M}{\sqrt{LL}} = \frac{4}{200} = 0.02$$

$$a = \frac{Q_0 k}{1 - k^2} = \frac{100 \times 0.02}{1 - 0.0004} \approx 2$$

$$f_c = \frac{1}{2\pi \sqrt{LC}} = \frac{1}{2\pi \sqrt{200 \times 10^{-6} \times 200 \times 10^{-12}}} = 795 \text{ kc}$$

$$f_0 = \frac{1}{2\pi \sqrt{(L + M)C}} = \frac{1}{2\pi \sqrt{204 \times 10^{-6} \times 200 \times 10^{-12}}} = 787 \text{ kc}$$

The half-power points occur where $y = \pm \sqrt{a^2 + 2a - 1} = \pm \sqrt{7} = \pm 2.65$ or where $\delta = (a + y)/2Q_0 = (2 \pm 2.65)/200 = 0.0232$, or -0.0032. Thus

$$f_1 = 787(1 - 0.0032) = 784 \text{ kc} \qquad \text{and} \qquad f_2 = 787(1 + 0.0232) = 805 \text{ kc.}$$

$$\text{BW} = f_2 - f_1 = 805 - 784 = 21 \text{ kc}$$

$$A_{\max} = \frac{g_m R_{sh}}{2} = \frac{g_m Q_0 \omega_0 L}{2}$$

$$= \frac{2000 \times 10^{-6} \times 100 \times 2\pi 787 \times 10^3 \times 200 \times 10^{-6}}{2} = 99$$

$$A_{f_c} = \frac{g_m R_{sh} a}{a^2 + 1} = \frac{A_{\max} 2a}{a^2 + 1} = \frac{99 \times 4}{5} = 79$$

REFERENCES

1. Loftin, D. H., and S. Y. White: Cascaded Direct-coupled Tube Systems Operated from Alternating Current, *Proc. IRE*, **18**, 669 (1930).
2. Artzt, M.: Survey of D-C Amplifiers, *Electronics*, **18**, 112 (August, 1945).
3. Arguimbau, L. B.: "Vacuum Tube Circuits," John Wiley & Sons, Inc., New York, 1948.

4. Bedford, A. V., and G. L. Fredendall: Transient Response of Multistage Video-frequency Amplifiers, *Proc. IRE*, **27**, 277 (1939).
5. McProud, C. G., and R. T. Wildermuth: Phase-inverter Circuits, *Electronics*, **13**, 47 (December, 1940).
6. Richter, W.: Cathode-follower Circuits, *Electronics*, **16**, 112 (November, 1943).
7. Krauss, H. L.: Graphical Solutions for Cathode Followers, *Electronics*, **20**, 116 (January, 1947).
8. Jones, M. C.: Grounded-grid Radio-frequency Voltage amplifiers, *Proc. IRE*, **32**, 423 (1944).
9. Mueller, George: "Alternating-current Machines," McGraw-Hill Book Company, Inc., New York, 1952.
10. Cruft Laboratory, War Training Staff: "Electronic Circuit and Tubes," McGraw-Hill Book Company, Inc., New York, 1947.
11. Eastman, A. V.: "Fundamentals of Vacuum Tubes," 3d ed., McGraw-Hill Book Company, Inc., New York, 1949.
12. E. E. Staff, MIT: "Applied Electronics," John Wiley & Sons, Inc., New York, 1943.
13. Reich, H. J.: "Theory and Applications of Electron Tubes," 2d ed., McGraw-Hill Book Company, Inc., New York, 1944.
14. Seely, S.: "Electron-tube Circuits," McGraw-Hill Book Company, Inc., New York, 1950.
15. Terman, F. E.: "Radio Engineering," 3d ed., McGraw-Hill Book Company, Inc., New York, 1947.

PROBLEMS AND QUESTIONS

1. The triode section of a 6AT6 tube is used in the circuit of Fig. 7-86. $R_L = 100{,}000$ ohms, $E_{bo} = 250$ volts, and $E_{co} = -3$ volts. (*a*) Use a tube manual to determine

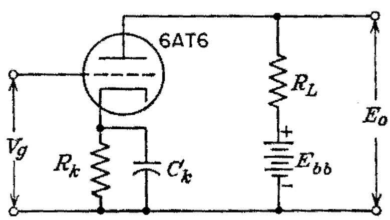


FIG. 7-86.

values of I_{bo}, E_{bb}, R_k, C_{gp}, C_{gk}, C_{pk}, μ, r_p, and g_m. (*b*) Draw Thévenin's equivalent circuit for low frequencies. Assume C_k is very large. (*c*) Find $\mathbf{A}_{\text{low}}$. (*d*) Draw Norton's equivalent circuit of high frequencies. Assume $C_w = 15\ \mu\mu f$. Determine the total shunt capacitance from plate to cathode, assuming a low-impedance source for V_g. (*e*) Determine the gain at 100 kc. (*f*) What modification should be made in Norton's circuit if the signal frequency is 100 Mc?

2. The gain of an amplifier using a 6AT6 triode is $40.7\underline{/157°}$ at 100 kc. The input voltage $\mathbf{V}_g$ is $2\underline{/0°}$. Find the grid current $\mathbf{I}_g$ if the operation is class 1.

3. A type 6AU6 pentode is used in a voltage-amplifier circuit. The gain is found to be $\mathbf{A} = 150\underline{/120°}$ at 100 kc. If $\mathbf{V}_g = 0.2\underline{/0°}$ volts, find $\mathbf{I}_g$ at this frequency.

4. The gain of a certain amplifier is $150\underline{/182°}$ at 1000 cps and $89\underline{/230°}$ at 60 cps. Calculate the decrease in voltage gain in decibels.

5. A certain crystal microphone under test conditions gives an open-circuit output voltage 50 db below 1 volt. How much amplification in decibels is needed to furnish 12.5 volts drive for the grid of a 6V6 power tube?

6. Figure 7-87 shows a T-type attenuator. Its insertion between the generator and the 500-ohm load decreases the power delivered to the load but should not change the

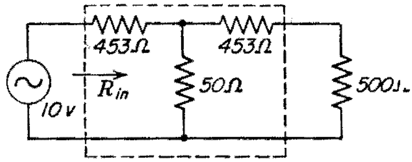


Fig. 7-87.

impedance presented to the generator. Check the design by proving $R_{in} \approx 500$ ohms, and compute the decibels loss caused by the insertion.

7. A two-stage direct-coupled amplifier of the Loftin-White variety is to be designed (see Fig. 7-88). For the first stage $E_{bo} = 52$ volts, $E_{co} = -2$ volts, $I_{bo} = 1$ ma. For the second stage $E_{bo} = 202$ volts, $E_{co} = -10$ volts, $I_{bo} = 1$ ma. (*a*) Connect the various leads to the voltage divider, and label the potential (relative to the tube 1

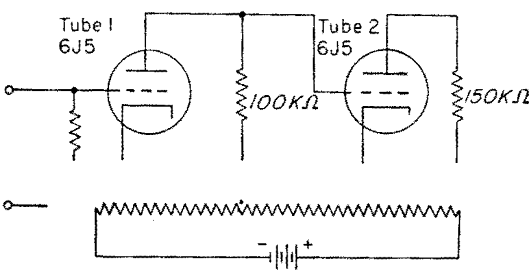


Fig. 7-88.

grid) at each connection. (*b*) Draw Thévenin's equivalent circuit for each stage. (*c*) If the grid of the first tube is made 0.5 volt more positive, what change in voltage appears across the 100,000-ohm resistor; across the 150,000-ohm resistor? Neglect the effect of the voltage divider. For tube 1, $\mu = 19$, $r_p = 14{,}300$ ohms. For tube 2, $\mu = 18$, $r_p = 22{,}800$ ohms.

8. Review problem. (*a*) Draw Thévenin's equivalent circuit for the amplifier of Fig. 7-89. The portion of the cathode resistor which furnishes bias is well bypassed, but R_1 is not. (*b*) Write the four equations needed for solving for $\mathbf{E}_o$.

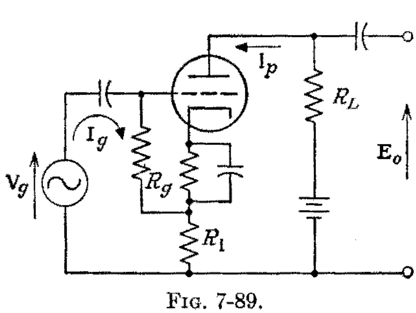


Fig. 7-89.

9. See Fig. 7-90. (*a*) Find **A** for middle frequencies. (*b*) The gain may be increased by increasing R_b. If this is done and the Q point is kept fixed, what practical difficulty arises? (*c*) If an attempt is made to increase the gain by increasing R_b and keeping E_{bb} fixed, what practical difficulty may arise? (*d*) In case (*b*) if R_g is large, what

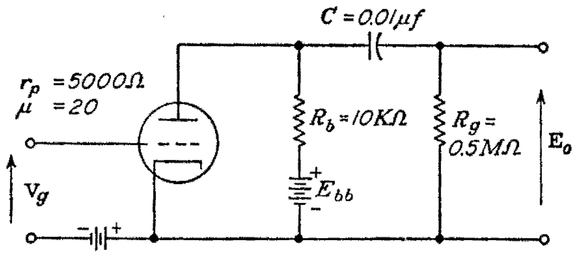


FIG. 7-90.

numerical value does the gain approach as R_b is made very large? (*e*) Find the lower half-power frequency. (*f*) If $\mathbf{V}_g = 1/\underline{0°}$ volts, find $\mathbf{E}_o$ at frequency f_1. (*g*) Although the load on the tube increases as the frequency decreases, the gain falls off. State in your own words why the gain of an R-C-coupled amplifier falls off at low frequencies. (*h*) To improve the l-f response, should C be increased or decreased; should R_g be increased or decreased?

10. Two of the three identical stages of an R-C-coupled amplifier are shown in Fig. 7-91. $R_g = 0.5$ megohm, $R_b = 0.25$ megohm, $C = 0.02$ μf, $C_w = 15$ $\mu\mu$f, C_d and C_k are very large. $E_{bb} = 250$ volts, $R_k = 1200$ ohms, $R_d = 1$ megohm. For this Q

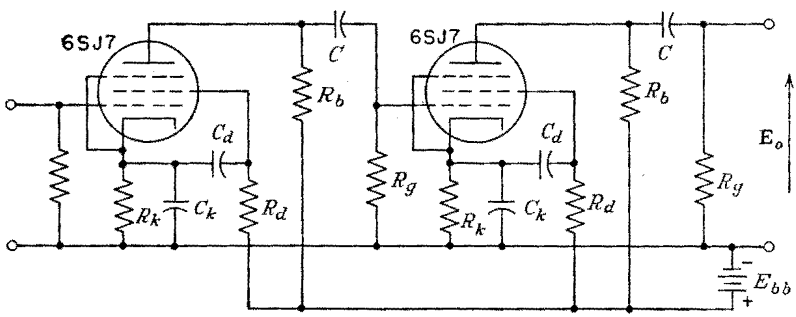


FIG. 7-91.

point $r_p = 1.5$ megohm, $g_m = 700$ micromhos. (*a*) For the first stage only determine R_{sh}, R_{low}, C_{sh}, f_1, f_2, A_{mid}. (*b*) Sketch the graph of A vs. f for the two identical stages. (*c*) At what frequencies is the gain of the two stages down 3 db from the mid-frequency value?

11. Determine the input admittance to the first tube of Fig. 7-91 at the frequency of 100 kc.

12. An R-C-coupled amplifier consists of three identical stages using 6J5 tubes. For each stage $E_{bb} = 250$ volts, $R_b = 100$ kilohms, $R_g = 500$ kilohms; $R_k = 3900$ ohms and is well bypassed. The coupling capacitor is 0.01 μf. $C_w = 12$ $\mu\mu$f. (*a*) Sketch the graph of gain vs. frequency for one stage. (*b*) Determine the input admittance at the upper half-power frequency.

13. A certain amplifier has $f_1 = 60$ cps and $f_2 = 500{,}000$ cps. A step voltage is applied to the grid. (*a*) What time is required for the output voltage to rise from

10 to 90 per cent of its greatest value? This time is often called the "rise time." (b) What time is required for the output to fall 10 per cent from its greatest value?

14. A type 6SH7 tube has the following characteristics: $C_g = 8.5\ \mu\mu f$, $C_{out} = 7\ \mu\mu f$, $r_p = 0.9$ megohm, $g_m = 4900$ micromhos. Assume that two of these tubes are used in cascade with a coupling capacitor of 0.2 μf and that C_w is 14.5 $\mu\mu f$. (a) What is the gain-area criterion for the first tube? (b) It is desired that the above tube be used in an amplifier which has a nearly uniform amplification from very low frequencies up to 1.5 Mc. This may require a different value of coupling capacitor but does not appreciably change C_w. Approximately what value of amplification per stage may be expected?

15. Compare 7G7, 7AG7, and 6BH6 type tubes as to suitability for use in a video-frequency voltage amplifier.

16. A compensated broad-band amplifier has been constructed with a plate load as shown in Fig. 7-35b. The quantity Q_2 as defined in Eq. (7-60) has been selected as 0.44. The upper half-power frequency of this amplifier *before* compensation is 2 Mc. For the tube used, $g_m = 5000$ micromhos, and $r_p = 0.7$ megohm. (a) Determine the value of R_b if the maximum gain per stage *after* compensation is 10. (b) If this value of R_b is used, at what frequency is the upper end of the mid-frequency band for this compensated amplifier? (c) What is the upper half-power frequency of the amplifier after compensation?

17. A wide-band amplifier is to have a flat gain characteristic and a phase shift proportional to frequency up to 4 Mc. What value of f_2 should be used for the design of the uncompensated amplifier?

18. (a) By inspection you should be able to express $\mathbf{E}_o$ in terms of $\mathbf{I}$ and the circuit constants for the circuit of Fig. 7-40 when $R_1 = R_2 = R$ and $C_1 = C_2 = C$. (This is a special case.) (b) Now by an extension of the reasoning of (a) you should be able to express, by inspection, $\mathbf{E}_o$ in terms of $\mathbf{I}$ and the circuit constants when $R_1/X_{C_1} = R_2/X_{C_2}$. (This is a more general case.)

19. Figure 7-92 is the same as Fig. 7-42 except that the C-R_g coupling is deleted. Find the gain at zero frequency and at a high frequency. (The results of this problem should help you to understand how l-f compensation comes about in the circuit of Fig. 7-42.)

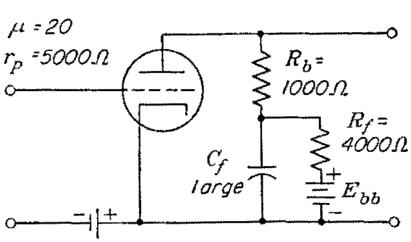


FIG. 7-92.

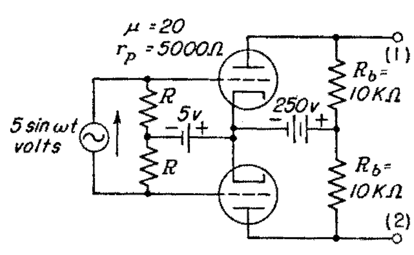


FIG. 7-93.

20. (a) In the balanced amplifier of Fig. 7-93, what is the potential of point 1 relative to point 2, with no applied signal? (b) A signal 5 sin ωt volts with polarity as shown is applied between the grids of the amplifier. Find the voltage at point 1 relative to point 2 as a function of time.

21. Refer to the circuit of Fig. 7-47. $R_b = 50$ kilohms, $R_g = 100$ kilohms, $R_k = 1000$ ohms. E_{bb} is the proper value to make $\mu = 20$, $g_m = 1500$ micromhos. $C_1 = C_2 = 0.01\ \mu f$. (a) Determine the mid-frequency gain of each tube and the value of R_{ab}. (b) What is the ratio of $\mathbf{E}_{2o}/\mathbf{E}_{1o}$ at the lower half-power frequency if R_k is well bypassed? (c) Repeat (b) with R_k not bypassed.

22. Determine the circuit values for a two-tube type of phase inverter employing a 6SN7 double triode. Each output voltage is to be 50 volts rms for middle frequencies. The lower half-power frequency is 20 cps. Assume R_k is well bypassed to avoid a determination of the feedback effect. What value of input voltage is needed?

23. Draw the circuit and determine the circuit values for a split-load type of phase inverter using a 6J5 tube. Each output is to be 25 volts rms. The lower half-power frequency is 20 cps. Assume this to be unaffected by feedback. Will the outputs be balanced at 20 cps? What value of input voltage is needed?

24. A current amplifier consists of three 6J5 stages. For the first and second stages $R_b = 100{,}000$ ohms, $R_g = 300{,}000$ ohms, $\mu = 20$, $r_p = 15{,}000$ ohms. The load on the third stage is practically 0 ohms and the tube operates with $\mu = 20$, $g_m = 1200$ micromhos. If the alternating grid signal for the first tube is 0.01 volt, what alternating current flows in the plate circuit of the third tube?

25. A current amplifier is to be designed to deliver 1 amp of current to operate a heavy-duty relay, with an amplifier input of 0.01 volt. The relay resistance is negligible. The output tube is a large power triode with $g_m = 10{,}000$ micromhos. How many stages of voltage gain will be required before the output tube if each stage has a gain of 10?

26. A 6C5 tube is used as a cathode follower. It is fed by another 6C5 used as a conventional amplifier with $Z_{out} = 10{,}000$ ohms. The follower feeds, in turn, another conventional amplifier (6C5) having a gain of 14. We are interested in the cathode follower. $\mu = 20$, $r_p = 10{,}500$ ohms, $R_k = 20{,}000$ ohms, $R_g = 500{,}000$ ohms, $C = 0.01$ μf, $C_{gk} = 3.0$ μμf, $C_{gp} = 2.0$ μμf, $C_{pk} = 11$ μμf, $C_{hk} = 12$ μμf, $C_w = 14$ μμf, $E_{bb} = 200$ volts. Determine A_{mid}, f_1, f_2, C_g at f_{mid}, Z_{out} at f_{mid}. See Fig. 7-94.

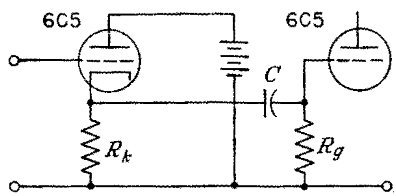


FIG. 7-94.

27. A 6J5 is to be used in a cathode-follower circuit similar to that shown in Fig. 7-56. R_k is to be 30,000 ohms. Assume $R_g \gg R_k$. Assume the limits on the grid drive to be 0 and -16 volts. $E_{bb} = 300$ volts. (*a*) Label e_k units on the e_b axis, and draw the load line on a sheet of 6J5 plate characteristics. (*b*) Determine values of E_{ko}, E_{co}, and R_1. (*c*) Relabel the curves with input-voltage values corresponding to the various grid voltages. (*d*) An input signal sufficient to drive the grid to 0 volts is used. Determine E_{k_1m}, E_{k_2m} and the percentage of second-harmonic distortion. What value of grid drive is needed?

28. In Fig. 7-95 the desired operating point for tube 2 is $E_{co} = -6$ volts, $I_{bo} = 1$ ma. (*a*) Determine the operating point for tube 1 and the value of R_k. (*b*) Use the

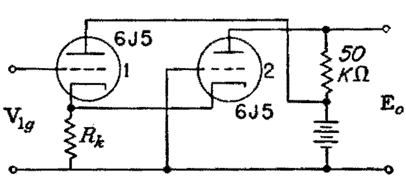


FIG. 7-95.

plate characteristics, and determine μ and r_p for each tube. (c) Assume linear operation, and draw the equivalent circuit. Determine the value of E_o if $V_{1g} = 1$ volt.

29. For an iron-core transformer prove that $L_p = (1 - k)L_1$ and $L_s = a^2L_p$.

30. A transformer-coupled amplifier has the following circuit values: Tube 6J5: $r_p = 7600$ ohms, $\mu = 20$, $C_{gp} = 3$ $\mu\mu$f, $C_{gk} = 3$ $\mu\mu$f, $C_{pk} = 4$ $\mu\mu$f. Transformer: $R_p = 400$ ohms, $R_s = 4500$ ohms, $L_1 = 15$ henrys, $C_{ps} = 300$ $\mu\mu$f, $C_p = 10$ $\mu\mu$f, $C_s = 50$ $\mu\mu$f, $k = 99.8$ per cent, $N_2/N_1 = 3$. Both polarity dots are at the top. The following stage is a 6J5 with a resistance load and a gain of $15\underline{/180°}$. The stray and wiring capacitance on the secondary side of the transformer is 49 $\mu\mu$f. Compute all values necessary to determine A vs. f, and plot the curve.

31. Why should an a-f voltage amplifier not use transformer coupling following a pentode?

32. A 200-μh inductor has an equivalent-series r-f resistance of 10 ohms at 796 kc. Determine (a) Q_0 on a series basis, (b) the equivalent shunt resistance at $f = f_0$, (c) the equivalent shunt inductance at $f = f_0$, (d) Q_0 on a shunt basis.

33. For a single-tuned-circuit r-f amplifier, $r_p = 1$ megohm, $g_m = 1800$ micromhos, $L = 300$ μh, $R_g = 0.75$ megohm, $f_0 = 796$ kc. If $Q_0 = 125$, determine R_{sh}, A_{max}, bandwidth, and the gain at 800 kc.

34. An amplifier of the type of Fig. 7-76 has $r_p = 1$ megohm, $g_m = 1500$ micromhos, $R_g = 1$ megohm, $f_0 = 796$ kc. What values of L and of Q_0 should be used if BW $= 20$ kc and $A_{max} = 150$?

35. An air-core r-f transformer has $L_1 = 200$ μh, $L_2 = 3200$ μh, $R_1 = 50$ ohms, $R_2 = 100$ ohms. The secondary is capacitor-tuned to a frequency of 796 kc. The primary is in the plate circuit of a triode for which $\mu = 20$ and $r_p = 8000$ ohms. (a) Determine the optimum coefficient of coupling. (b) What is the gain at 796 kc for this coupling? (c) What is the bandwidth with this coupling? (d) Repeat (b) and (c) if $k = 0.707$ is used instead.

36. A double-tuned transformer-coupled r-f amplifier is to be designed to have a frequency-response curve which is centered at $f = 1$ Mc. The bandwidth of the amplifier is to be 14.14 kc. The coupling is to be adjusted to the transitional value. (a) What should be the resonant frequency of the two tuned circuits? (*Hint*: Consider the center frequency to be the arithmetic mean of f_1 and f_2.) (b) Assume the resonant frequency of the tuned circuit to be very near 1000 kc. Find the approximate Q_0 of the tuned circuit if the bandwidth is to be as stated. (c) If the coils used have inductances of 200 μh and Q_0's of 150, how much resistance must be added in parallel with each to bring their Q_0 values down to 100? (Assume $f_0 = 1000$ kc.) (d) What will be the gain at the center frequency if the tube used has a g_m of 1800 micromhos? (Assume effective $Q_0 = 100$, $L = 200$ μh, $f_0 = 1000$ kc.) (e) Calculate the capacitance of the tuning capacitors if $f_0 = 1000$ kc and $L = 200$ μh.

37. Prove the statement of Eq. (7-132). To do this, assume the losses for primary copper, secondary copper, and core to be, respectively, $P/4$, $P/4$, and $P/2$. Assume a power efficiency η and input and output voltages E_1 and E_2/a, respectively. Starting with E_1, compute the currents and voltages for the equivalent circuit in terms of P, η, and E_1; in this way determine an expression for E_2/a. From this determine a relation between E_2 and E_1 in terms of a and η. Hence obtain Eq. (7-132) as an approximation.

38. Derive Eq. (7-52) for the fall in output voltage in a video-frequency amplifier caused by charging of the coupling capacitor.

CHAPTER 8

AUDIO-FREQUENCY POWER AMPLIFIERS

8-1. Introduction. In this chapter we shall study vacuum-tube amplifier circuits which give a large output voltage of substantially the same waveform as that of the input voltage, accompanied by a large current of like waveform, so that there is considerable power delivered to the load. The operation really is a power-conversion process, as in most cases all of the output power derives from the plate d-c source, under the control of the signal voltage applied to the grid. In general the treatment can often be made similar to that used for voltage amplifiers, except that usually the operation is not linear and therefore graphical or experimental procedures give most accurate results.

Except for details, a power tube resembles a voltage-amplifier tube. Power tubes vary in size depending upon the amount of output power desired and the efficiency of operation. For a small amount of output power a voltage tube may be used. The tubes with the higher power ratings need larger anodes to dissipate by radiation the heat derived from the energy of the impinging electrons, without making too high a temperature necessary. For tubes of very high power rating the heat may be carried away by convection by an air or a water stream. A second requirement for a power tube is a sufficiently large cathode to furnish the necessary electron emission. This also means that a relatively large amount of heating power will be needed. Small power tubes use oxide-coated cathodes, while those of medium power employ thoriated tungsten, which will withstand somewhat more ionic bombardment. High-power tubes with their very high voltage drops use pure-tungsten filamentary cathodes. Tubes of small and moderate ratings may be triodes, tetrodes, or pentodes, but for the very high power ratings there is a problem of cooling the screen grid of the latter two kinds, which makes the triode usually preferred.

In making a choice of tube, the class of operation, which is related to the efficiency, must be taken into account. If the very best quality of output is required, as in some measurements circuits, it is well to use class A_1 operation as this renders the output most free of distortion. In this connection it may also be advisable to choose a tube of higher power rating than actually required, so that the grid drive may be made small compared with the maximum allowable and thus decrease the distortion.

If the distortion requirements are not so stringent, class A_1 operation with maximum allowable drive applied may be quite satisfactory. For sound-reproduction purposes there is usually no need for very low distortion, and a more efficient class of operation such as AB_1 or AB_2 may be good enough, when used with a balanced amplifier of the push-pull type. When the a-c power requirement is great enough so that large costly tubes are needed, it may be important to operate with the high efficiency of class B. Class C operation does not in general render an output waveform which is a replica of the input-voltage waveform and hence is useless for the purposes dealt with here.

The tube and class of operation having been decided upon, the operating point and the load must be chosen. Practically all power tubes require quite specific loads for proper operation, and these usually are not the ones available. Hence transformer coupling between tube and load is almost universal, for lower-frequency operation anyway, although other impedance-matching devices are available, particularly at higher frequencies. An extra advantage of transformer coupling is the low d-c load for the tube, making the plate-supply-voltage requirement but little more than the tube's operating voltage. The operating point is usually chosen together with the load since they are mutually dependent. The class of operation, the allowable distortion, and the plate operating voltage all have importance here.

8-2. Circuit Type and Efficiency Using Class A Operation. If one neglects important features for the sake of simplifying an analysis to a point where it may be completed with only moderate difficulty, the results are bound to be incomplete. They can also be misleading unless one realizes that simplifications have been made and that adjustments are required in interpreting the results for practical application. This must be kept in mind in the idealized analyses which follow.

For the ordinary power amplifier there is a choice between only two basic circuits, one with directly connected load and the other with inductively coupled load, ordinarily using a transformer. It has already been stated that the latter is the almost universal arrangement, but let us give the matter a bit of study, using idealized class A operation in the example.

The plate-circuit efficiency is defined as

$$\eta_p = \frac{P_{ac}}{P_{bb}} \tag{8-1}$$

Its value depends upon the type of circuit, the class of operation, the load, the operating point, and the amount of grid drive applied. For the amplifier of Fig. 8-1a with its direct-connected load, the load line is drawn in Fig. 8-1*b*. If no signal is applied, $P_{ac} = 0$, $P_{bb} = E_{bb}I_{bo}$, and $\eta_p = 0$. If a grid signal is applied sufficient to make the operation between points A and B and no distortion is assumed, the formulas of Art. 6-16 may be used to obtain

$$P_{ac} = E_p I_p = \frac{e_{b_{\max}} - e_{b_{\min}}}{2\sqrt{2}} \frac{i_{b_{\max}} - i_{b_{\min}}}{2\sqrt{2}} \tag{8-2}$$

and

$$P_{bb} = E_{bb}I_{bo} = E_{bb} \frac{i_{b_{\max}} + i_{b_{\min}}}{2} \tag{8-3}$$

[Note that $I_{bo} = (i_{b_{\max}} + i_{b_{\min}})/2$ only in the distortionless case.] Hence

$$\eta_p = 100 \frac{(e_{b_{\max}} - e_{b_{\min}})(i_{b_{\max}} - i_{b_{\min}})}{4E_{bb}(i_{b_{\max}} + i_{b_{\min}})} \% \tag{8-4}$$

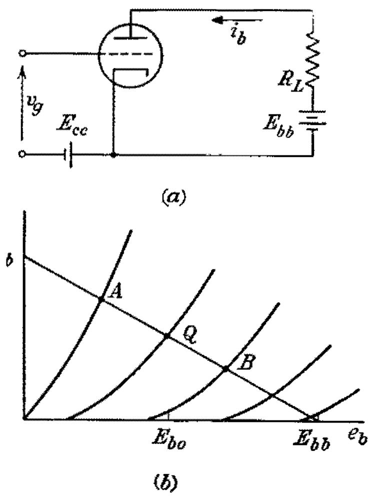


FIG. 8-1. (*a*) A power amplifier with direct-connected load. (*b*) The load line for this amplifier.

The greatest theoretical grid signal which could be applied for class A operation would move A and B to positions on the axes. Under these extreme conditions

$$\eta_p = 100 \frac{(E_{bb} - 0)(i_{b_{\max}} - 0)}{4E_{bb}(i_{b_{\max}} + 0)} = 25\% \tag{8-5}$$

Since the grid drive used would swing the grid to a high positive value, it greatly exceeds practical limits, so that 25 per cent is the upper limit to the efficiency of this type of circuit.

Let us now consider the circuit with inductively coupled load as in either of the diagrams of Fig. 8-2. The load line for R_L is shown in Fig. 8-2*c*. For these circuits suppose $R_{dc} = 0$ and hence $E_{bo} = E_{bb}$.

The formula for P_{ac} is the same as that given by Eq. (8-2), but the input power is

$$P_{bb} = E_{bb}I_{bo} = E_{bo} \frac{i_{b_{\max}} + i_{b_{\min}}}{2} \tag{8-6}$$

and hence the efficiency is

$$\eta_p = 100 \frac{(e_{b_{\max}} - e_{b_{\min}})(i_{b_{\max}} - i_{b_{\min}})}{4E_{bo}(i_{b_{\max}} + i_{b_{\min}})} \quad \% \tag{8-7}$$

If a large grid signal is applied which puts both A and B on the axes, then $i_{b_{\max}} = 2I_{bo}$ and $e_{b_{\max}} = 2E_{bb}$ and the theoretical upper limit to the efficiency becomes

$$\eta_p = 100 \frac{(2E_{bo} - 0)(2I_{bo} - 0)}{4E_{bo}(2I_{bo} + 0)} = 50\% \tag{8-8}$$

Thus higher efficiencies are obtainable by using inductively coupled loads, though in practice class A efficiencies are not nearly 50 per cent because the load required may not give enough power and low enough distortion. With pentodes and beam power tubes, the characteristic curve for $e_c = 0$ is closer to the vertical axis than in the case of triodes, and hence these tubes surpass triodes in practical plate-circuit efficiency.

For any of the circuits of Fig. 8-1 or 8-2 suppose E_{bb} and E_{cc} to be adjusted to make the quiescent plate dissipation of the tube not exceed the allowable limit. If an a-c signal e_g is now applied, power will be delivered to the load. If there is no distortion, the average current will be unchanged so that the power P_{bb} from the plate supply will be unchanged. Since R_L is receiving more power than before by the amount of P_{ac}, it is obvious that the tube must be dissipating just that much less power. Hence the plate dissipation decreases upon the application of a grid signal, its value with signal being

$$P_p = E_{bo}I_{bo} - P_{ac} \tag{8-9}$$

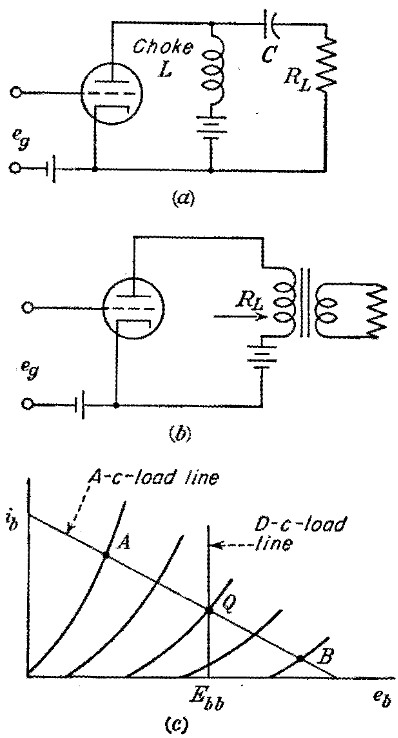


FIG. 8-2. (a) Shunt feed is used; $X_L \gg R_L$. (b) The effect is the same with a transformer. (c) Their load line.

In practice this theoretical relation is somewhat in error since the average current usually differs somewhat from I_{bo} because of distortion. And of course it applies only to class A operation.

8-3. Efficiency with Other Classes of Operation. If the circuit of Fig. 8-2*b* is operated with E_{cc} high enough to just cut off the plate current with no signal applied, the operation is class B (see Fig. 8-3). With sinusoidal signal applied, the grid-voltage, plate-current, and plate-voltage waveforms are shown in Fig. 8-4; the latter two are assumed to be half sinusoids. To be practical, a balanced-amplifier arrangement is needed to supply the other half cycle, or a tuned load might be used. For the purpose of this discussion let us consider only the power related

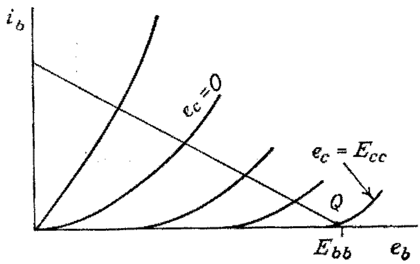


FIG. 8-3. Class B operation.

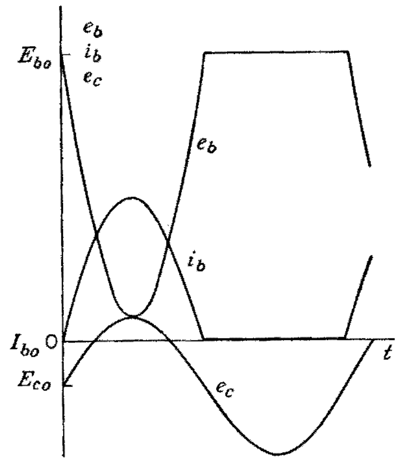


FIG. 8-4. Waveforms for class B operation.

to one tube of the balanced amplifier.

$$P_{ac} = \frac{1}{2} E_p I_p = \frac{1}{2} \frac{E_{pm} I_{pm}}{\sqrt{2}\,\sqrt{2}} = \frac{1}{2} \frac{(E_{bb} - e_{b_{\min}})(i_{b_{\max}} - 0)}{\sqrt{2}\,\sqrt{2}}$$

$$= \frac{(E_{bb} - e_{b_{\min}}) i_{b_{\max}}}{4} \tag{8-10}$$

$$P_{bb} = E_{bb} I_{ba} = E_{bb} i_{b_{\max}} \frac{1}{\pi} \tag{8-11}$$

$$\eta_p = 100 \frac{\pi}{4} \frac{(E_{bb} - e_{b_{\min}}) i_{b_{\max}}}{E_{bb} i_{b_{\max}}} = 100 \frac{\pi}{4} \frac{E_{bb} - e_{b_{\min}}}{E_{bb}} \tag{8-12}$$

If $e_{b_{\min}}$ is made to approach zero, the upper limit to the theoretical efficiency of a class B amplifier is obtained.

$$\eta_p = \frac{\pi}{4} \times 100 = 78.5\% \tag{8-13}$$

This is most easily approached by using a very high load impedance, in which case $i_{b_{\max}}$ will be small and the power output small. Hence in practice it is not desirable to obtain this high efficiency by using a high load impedance. Rather, it is usual to drive the grid highly positive to the limit of the allowable dissipation of the grid or of the plate or the limit of the power of the driver (the signal supply), whichever is least, and thus obtain satisfactory power output and high efficiency.

From the discussion of this article it is seen that class A operation involves low efficiency which cannot exceed 50 per cent and usually is much less. Class B operation cannot give efficiencies exceeding about 78 per cent and can be made to exceed 50 per cent in practice. The efficiency in class AB operation is intermediate between these values.

8-4. Optimum Load and Bias for a Class A_1 Amplifier. The load impedance which yields the most power output depends upon the conditions imposed upon the operation. Under light grid-drive conditions, when a high-impedance load is not needed to limit the distortion, one may determine one answer; when heavy grid drive makes it necessary to watch distortion, one obtains perhaps another answer. If E_{bb} is one of the parameters, still another result can be obtained. We shall discuss what appears to be the more important possible circumstances.

We shall first suppose that *the operating point and the available grid-signal voltage are fixed and that there is no appreciable distortion.* Then the linear-equivalent-circuit theorem may be used, and the results are easy to obtain. There is even some practical use for these results. For perfectly linear performance r_p and μ must be constant throughout the cycle of operation. This implies straight, parallel, and equispaced characteristic curves in the plate diagram. In a small enough region of the diagram this can usually be assumed to be the situation, and hence, if the grid-signal voltage is small, the results obtained by this method will be fairly reliable.

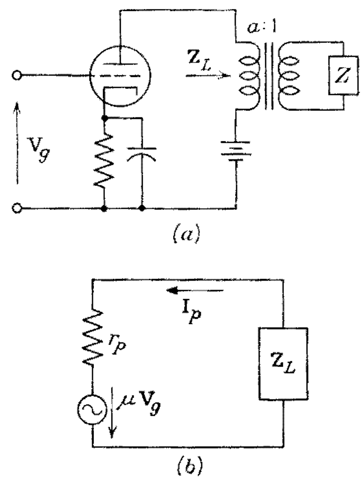


FIG. 8-5. (a) A power amplifier. (b) Its linear equivalent circuit.

Figure 8-5 shows a power amplifier and its linear equivalent circuit. $\mathbf{Z}_L$ is the load looking into the primary, and if the transformer is considered to be ideal, this will be $a^2\mathbf{Z}$.

$$\mathbf{I}_p = \frac{\mu \mathbf{V}_g}{r_p + \mathbf{Z}_L} = \frac{\mu \mathbf{V}_g}{r_p + Z_L \cos\theta + jZ_L \sin\theta} \tag{8-14}$$

since

$$\mathbf{Z}_L = R_L + jX_L = Z_L \cos\theta + jZ_L \sin\theta \tag{8-15}$$

$$I_p = \frac{\mu V_g}{[(r_p + Z_L \cos\theta)^2 + Z_L^2 \sin^2\theta]^{1/2}} \tag{8-16}$$

Hence

$$P_{ac} = I_p^2 R_L = \frac{\mu^2 V_g^2 Z_L \cos\theta}{(r_p + Z_L \cos\theta)^2 + Z_L^2 \sin^2\theta} \tag{8-17}$$

P_{ac} will be a mathematical maximum only if

$$\frac{\partial P_{ac}}{\partial Z_L} = 0 \qquad \text{and} \qquad \frac{\partial P_{ac}}{\partial \theta} = 0 \tag{8-18}$$

The first of these conditions yields

$$Z_L = r_p \tag{8-19}$$

and the second condition gives

$$\theta = 0 \tag{8-20}$$

Hence $\mathbf{Z}_L$ should be a pure resistance equal to r_p. This, of course, is no

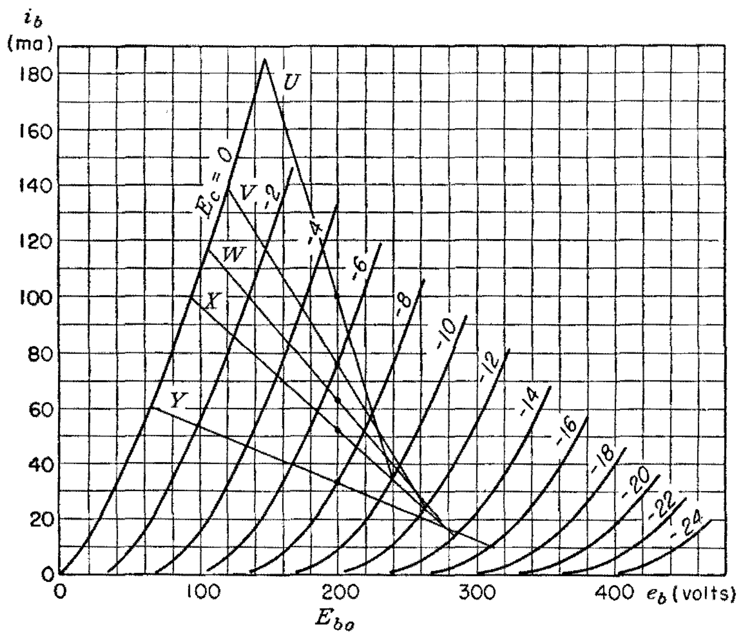


FIG. 8-6. Operation of a power amplifier with various loads, bias, and drive.

surprise as it is the usual condition for maximum power delivery for a fixed-voltage generator with purely resistive internal impedance.

Under the conditions of (8-19) and (8-20) the power delivered to $Z_L = r_p$ becomes, upon substituting into Eq. (8-17),

$$P_{ac_{\max}} = \frac{\mu^2 V_g^{\,2}}{4r_p} = \frac{\mu g_m V_g^{\,2}}{4} \tag{8-21}$$

which indicates that the power is proportional to the square of the grid-signal voltage and to the product μg_m. Thus μg_m may serve as a figure of merit for power tubes.

The conditions imposed in the analysis just completed are usually not very practical ones. In general it is desirable to design a power amplifier to give the largest possible power output within the set limitations of allowable plate dissipation, permissible harmonic distortion, and the available plate-supply voltage. The approximate rule yielded by Eq. (8-21) shows that high grid drive is important and hence the resulting harmonic distortion must become an essential element in the design. Distortion due to the flow of grid current through the impedance of the grid-voltage source can be obviated by using class 1 operation. Drive to zero grid voltage is usually considered to be the limit in this connection, and although a small amount of grid current flows for somewhat negative values of the grid voltage, there is only a small error made in not limiting slightly the grid drive.

Let us first consider the specific problem of a triode amplifier with E_{bo} held at 200 volts while the load resistance, grid bias, and drive are changed. In Fig. 8-6 five different bias values are used, and, with grid-signal voltage always sufficient to drive to zero, the load in each case is adjusted to give 5 per cent harmonic distortion. Table 8-1 shows the load

TABLE 8-1

Load line	E_{co}	R_L	P_{ac}, watts	Distortion, per cent
U	-5.0	$0.9r_p$	1.64	5
V	-6.0	$1.4r_p$	2.03	5
W	-6.5	$1.95r_p$	2.03	5
X	-7.0	$2.6r_p$	2.01	5
Y	-8.0	$4.2r_p$	1.53	5

in terms of r_p and the full-drive power output. It can be seen that the load giving maximum power output is apparently somewhat less than $2r_p$. Evidently under these conditions the load should not be r_p.

It is desirable to obtain a general rule for determining the best load for the amplifier. Since the problem is quite complicated, we again resort to some idealization of the tube characteristics and hence the solution obtained will be only approximate.

Let us write down our assumptions:

1. The tube is a triode.
2. E_{bo} is fixed in value.
3. It is noticeable in Fig. 8-6 that, although the maximum values of plate current vary widely, the minimum value is not greatly different from 20 ma. Therefore, to limit the distortion to a certain amount, the value of $i_{b_{min}}$ will be considered fixed.
4. Grid bias, grid drive, and R_L are variable.
5. To simplify the analysis, we shall make a further assumption that the characteristic curves are straight, parallel, and equispaced in the region above $i_{b_{min}}$.

Figure 8-7 gives us a geometric construction to work with. E_A is the plate voltage corresponding to $i_{b_{min}}$ and $e_c = 0$. Recognizing that

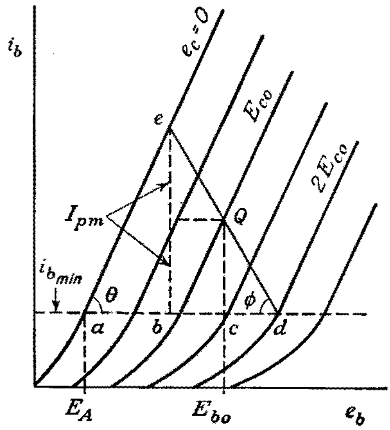


FIG. 8-7. Slightly idealized characteristic curves are used.

$$r_p = \frac{ab}{be} = \cot\theta \qquad (8\text{-}22)$$

and

$$R_L = \frac{bd}{be} = \cot\phi \qquad (8\text{-}23)$$

we have the geometrical relation

$$E_{bo} - E_A = ab + bc = 2I_{pm}r_p + I_{pm}R_L \qquad (8\text{-}24)$$

or

$$I_{pm} = \frac{E_{bo} - E_A}{R_L + 2r_p} \qquad (8\text{-}25)$$

Then

$$P_{ac} = I_p{}^2R_L = \frac{(E_{bo} - E_A)^2}{2(R_L + 2r_p)^2}R_L \qquad (8\text{-}26)$$

If dP_{ac}/dR_L is set equal to zero, the condition for maximum power is obtained. This process yields

$$R_L = 2r_p \qquad (8\text{-}27)$$

The corresponding power is

$$P_{ac_{max}} = \frac{(E_{bo} - E_A)^2}{16r_p} \qquad (8\text{-}28)$$

The optimum grid bias may be estimated by using a formula derived as follows: An approximate formula for the amplification factor is

$$\mu \approx \frac{\Delta e_b}{\Delta e_c} \quad \text{for constant } i_b\text{, and with small incremental values} \qquad (8\text{-}29)$$

Hence

$$\mu = -\frac{e_{b_{max}} - E_A}{2E_{co}} = -\frac{2I_{pm}(R_L + r_p)}{2E_{co}} = -\frac{(E_{bo} - E_A)(R_L + r_p)}{(R_L + 2r_p)E_{co}}$$
$$= -\frac{(E_{bo} - E_A)3r_p}{4r_pE_{co}} = -\frac{3}{4}\frac{E_{bo} - E_A}{E_{co}} \qquad (8\text{-}30)$$

from which

$$E_{co} = -\frac{3}{4}\frac{E_{bo} - E_A}{\mu} \qquad (8\text{-}31)$$

The value of E_A is not generally known. Trials with a 2A3 tube show its approximate value to be between $\frac{1}{11}$ and $\frac{1}{8}$ E_{bo} for 5 per cent distortion. Use of the value $0.1E_{bo}$ for E_A gives a formula which is approximately

$$E_{co} \approx -0.7\frac{E_{bo}}{\mu} \qquad (8\text{-}32)$$

It is perhaps important to repeat that these formulas are for a triode, they are not very accurate, and they should not be used at all if the plate dissipation exceeds the allowable limit thereby. For more reliable results a laboratory or even a graphical determination is to be preferred.

8-5. Graphical Determination of the Best Operation for a Class A_1 Triode Amplifier. Reference to a tube manual will usually determine the allowable plate dissipation for a given triode. For quiescent operation

$$P_{po} = E_{bo} I_{bo} \tag{8-33}$$

If the locus is plotted on e_b, i_b axes, it is a rectangular hyperbola. If the Q point for the tube is located above this hyperbola, the plate will overheat.

Let us suppose E_{bo} is specified, as it is usually possible to estimate its value. If use of formula (8-32) locates Q below the plate-dissipation curve, the operating point obtained can be safely used and the results obtained by using a load equal to $2r_p$ should be fairly satisfactory, except that the amount of distortion produced may exceed the desired value. If, on the other hand, the rule places Q above the curve, it should be lowered to a position on the curve, keeping E_{bo} the same value. Trial load lines may now be drawn through Q and determinations of harmonic distortion made using formulas (6-42) and (6-41), pursuing a trial-and-error method until a satisfactory line is obtained.

An alternate and generally more satisfactory method of drawing the load line for a given amount of distortion is to use a so-called distortion rule. The theory supporting this method, which is valid for class A_1 triode circuits only, depends upon the statement that the allowable second-harmonic current can be a certain percentage H of the fundamental current. Hence, again using formulas (6-42) and (6-41), we obtain

$$\frac{i_{b_{\max}} + i_{b_{\min}} - 2I_{bt}}{2(i_{b_{\max}} - i_{b_{\min}})} = \frac{H}{100} \tag{8-34}$$

One can clear this of fractions and group terms to obtain the relation

$$(1 - 0.02H)(i_{b_{\max}} - I_{bt}) = (1 + 0.02H)(I_{bt} - i_{b_{\min}}) \tag{8-35}$$

or

$$\frac{i_{b_{\max}} - I_{bt}}{I_{bt} - i_{b_{\min}}} = \frac{1 + 0.02H}{1 - 0.02H} \tag{8-36}$$

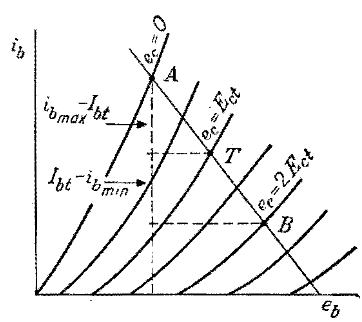


FIG. 8-8. Distortion occurs if AT is not equal to TB.

In Fig. 8-8, if $AT/TB = (1 + 0.02H)/(1 - 0.02H)$, relation (8-36) is satisfied and full-drive operation of the amplifier will produce H per cent of second-harmonic distortion in the plate current. Specifically, if $H =$

5 per cent, the ratio $AT/TB = 1.1/0.9 = {}^{11}\!/_{9}$. To recognize easily this relation between segments of the load line, as the load line is turned on the T point as an axis, one may make a rule as shown in Fig. 8-9, called a 9-11 rule or a 5 per cent distortion rule. From a point labeled 0, mark off to the right and label successive $\frac{9}{16}$ in. (or any other convenient unit). To the left mark off and label successive $\frac{11}{16}$ in.

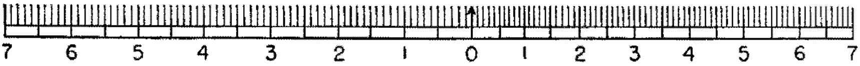


FIG. 8-9. A 5 per cent distortion rule.

If this rule is used as a straightedge for drawing a load line, with the 0 placed on T and the larger spaced end of the rule toward higher plate currents, the rule can be rotated until the number falling on the $e_c = 0$ curve is the same as the number falling on the $e_c = 2E_{ct}$ curve. When this occurs, the ratio of lengths will be 11:9, which makes the distortion 5 per cent.

Upon applying a signal the operating point for a triode rises to a new T point (Arts. 6-7, 6-17), which may be above the plate-dissipation curve. However, the plate dissipation actually decreases upon application of the grid signal (Art. 8-2), and so no trouble ensues, and unless one wishes in the interest of greater accuracy to reevaluate the distortion using the new T point, its determination may be omitted. After all it must be remembered that the characteristic curves used do not very accurately represent any individual tube, and there is no use in being particular to the point of absurdity.

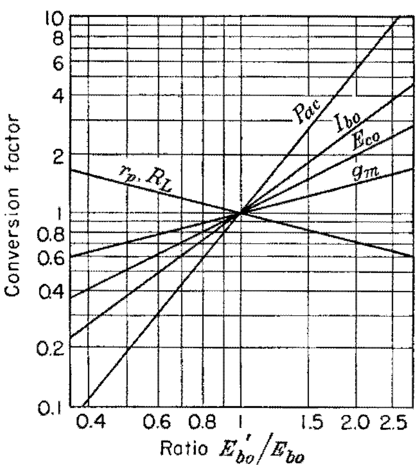


FIG. 8-10. A conversion chart for power tubes. (*Adapted from RCA Receiving-tube Manual, by permission.*)

After the load line is drawn, the values of I_{p_1} and E_{p_1} may be calculated, using formula (6-41) and its voltage paraphrase. From these the a-c load resistance and the power output may be computed. By calculating I_{ba} with the use of Eq. (6-40), the input power may be computed and, from that and P_{ac}, the plate-circuit efficiency obtained.

8-6. Designing a Class A_1 Triode Power Amplifier Using a Tube Manual. For power tubes in common use, a tube manual may be consulted for advice on proper operating conditions. If it is convenient to obtain the exact value of plate voltage, the required circuit values can be easily computed. If, however, the value of E_{bo} must differ somewhat from the published value, the manual may have a conversion chart which gives the new operating values. A typical chart is shown in Fig. 8-10.

The theory underlying this chart is not too rigorous, and hence the results are only approximate. By Art. 4-5 the plate current for a triode follows an approximate three-halves-power law,

$$i_b = k(e_b + \mu e_c)^{3/2} \tag{8-37}$$

At any point

$$r_p = \frac{\partial e_b}{\partial i_b} \tag{8-38}$$

and since

$$\frac{\partial i_b}{\partial e_b} = \frac{3}{2} k(e_b + \mu e_c)^{1/2} \tag{8-39}$$

$$r_p = \frac{2}{3k} (e_b + \mu e_c)^{-1/2} \tag{8-40}$$

If we use Eq. (8-32) as an approximation to the proper bias E_{co}, then as E_{bo} is changed, E_{co} also changes so that, at the operating point Q,

$$r_p = \frac{2}{3k} (E_{bo} - 0.7E_{bo})^{-1/2} = KE_{bo}{}^{-1/2} \tag{8-41}$$

or the values of r_p under the new and the published conditions, respectively, are related,

$$\frac{r'_p}{r_p} = \left(\frac{E_{bo}}{E'_{bo}}\right)^{1/2} \tag{8-42}$$

The load resistance bears an approximate linear relation to r_p, and hence

$$\frac{R'_L}{R_L} = \left(\frac{E_{bo}}{E'_{bo}}\right)^{1/2} \tag{8-43}$$

The relation for I_{bo} is obtained from Eq. (8-37) as

$$\frac{I'_{bo}}{I_{bo}} = \left(\frac{E'_{bo}}{E_{bo}}\right)^{3/2} \tag{8-44}$$

Since $g_m = \mu/r_p$ and μ is approximately constant,

$$\frac{g'_m}{g_m} = \left(\frac{E'_{bo}}{E_{bo}}\right)^{1/2} \tag{8-45}$$

The relation for grid bias is obtained from (8-32),

$$\frac{E'_{co}}{E_{co}} = \frac{E'_{bo}}{E_{bo}} \tag{8-46}$$

The output power is given approximately from Eq. (8-28) by neglecting E_A,

$$\frac{P'_{ac}}{P_{ac}} = \left(\frac{E'_{bo}}{E_{bo}}\right)^{5/2} \tag{8-47}$$

Let us continue with P_{ac} as an example.

$$\log \frac{P'_{ac}}{P_{ac}} = \frac{5}{2} \log \frac{E'_{bo}}{E_{bo}} \tag{8-48}$$

With the vertical axis uniformly calibrated in terms of log (P'_{ac}/P_{ac}) and the horizontal axis likewise calibrated in terms of log (E'_{bo}/E_{bo}) (but labeled in terms of P'_{ac}/P_{ac} and E'_{bo}/E_{bo}, respectively), the locus will be a straight line through the origin (log 1, log 1), and with slope $5/2$. To check the P'_{ac}/P_{ac} ratio for $E'_{bo}/E_{bo} = 2.5$, we evaluate $5/2$ log $2.5 \approx 0.995$, which then yields 9.9 as the value of the ratio P'_{ac}/P_{ac}. In a similar fashion the other graphs of Fig. 8-10 may be justified.

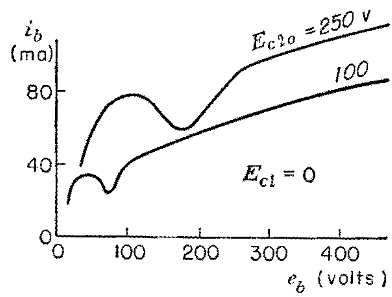


FIG. 8-11. There are better tubes than a tetrode for power amplifiers with an untuned load.

8-7. Class A_1 Amplifiers Using Pentode and Beam Power Tubes. Figure 8-11 shows hypothetical characteristic curves for a tetrode power tube. To avoid serious plate-current distortion, operation in the region of the "dip" should be avoided. A low screen voltage gives too low current to yield much power output. A high screen voltage forces operation for low distortion to be conducted with very high plate voltages, limiting the efficiency severely. This makes it appear that tetrodes of this type should not be used for power purposes unless the load is tuned to filter out distortion components, though with low screen voltages they might be fairly satisfactory as voltage amplifiers.

Tubes of the pentode type (including suppressorless beam tubes) with low screen voltages are likewise unsuitable for power use because of the low plate current. The characteristics of a type 6F6 with a high screen voltage appear in Fig. 8-12. High power output is now possible with a moderate plate-supply voltage. Superimposed on this graph are the $e_c = 0$ curves for a 2A3 triode and a 6V6 beam tube. Since the theoretical 50 per cent efficiency is approached more nearly as $e_{b_{\min}}$ approaches zero, the 6V6 is best in this regard and the 2A3 poorest. A pentode-type tube has a higher amplification factor than has a triode of the same g_m rating because the plate is shielded from the cathode while the control grid is not. Hence the pentode-type tubes require less grid drive to change the plate current from maximum to zero than does a triode working between the same limits. This quality is included in the expression called the

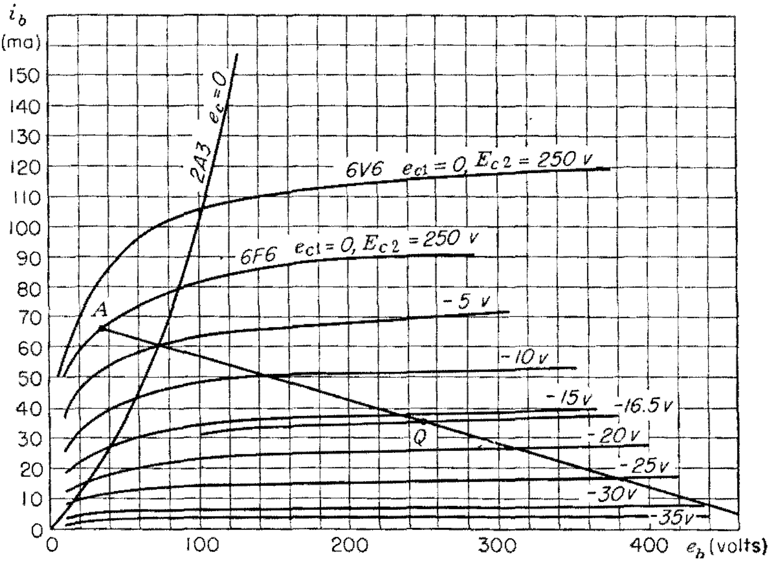


FIG. 8-12. Comparison of characteristics for tube types 6F6, 6V6, and 2A3.

power sensitivity, its definition being

$$S_p = \frac{P_{ac}}{E_g^2} \tag{8-49}$$

The optimum load for a pentode-type tube is somewhat more difficult to determine than that for a triode. The idealized plate characteristics may be drawn as in Fig. 8-13. The most power will be obtained if load line B is used. For this line

$$R_L = \frac{2E_{bo}}{i_{b_{\max}}} \tag{8-50}$$

This rough approximation is generally unsatisfactory and a laboratory determination or a graphical procedure using tube characteristics is preferable.

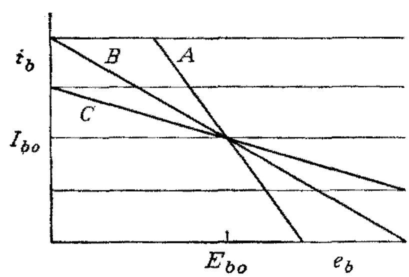


FIG. 8-13. Idealized curves for a pentode.

To determine the proper load to use, let us suppose the operating values E_{bo} and E_{co} are known. A possible Q point has been marked in Fig. 8-12 for the 6F6 tube. We may now draw several load lines through the Q point and for each compute the values of I_{p_1}, I_{p_2}, I_{p_3}, E_{p_1}, R_L, and P_{ac}, using the formulas of the (6-52) to (6-56) group and their voltage paraphrases. Figure 8-14 gives in graphical form the variation of P_{ac}, percentage second harmonic, percentage third harmonic, and percentage total harmonic distortion, vs. R_L. For low load resistance values the second-harmonic distortion is high because of the crowding together of the characteristic curves at the bottom. As the load is made higher, the segments at the two ends of the load line become more nearly equal, indicating less distortion. A handy criterion for the line position to give zero

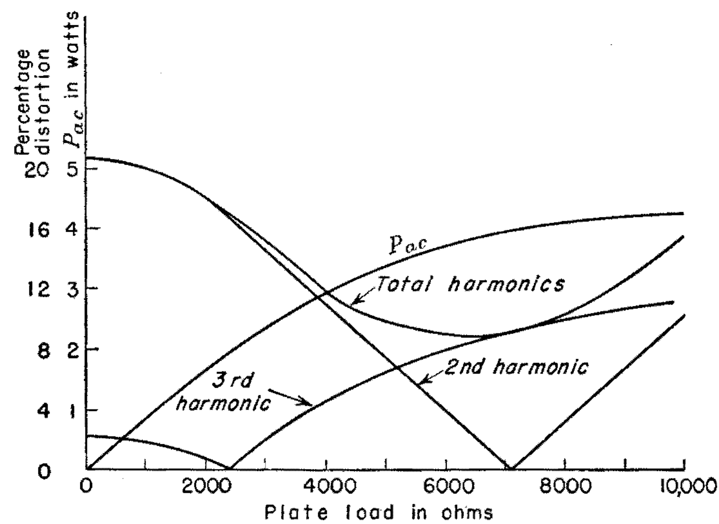


FIG. 8-14. How P_{ac} and distortion vary with load for a 6F6 pentode.

second harmonic is obtained by setting the expression for I_{p_2} [Eq. (6-54)] equal to zero, *i.e.*,

$$I_{p_2} = \frac{i_{b_{\max}} + i_{b_{\min}} - 2I_{bt}}{4\sqrt{2}} = 0 \tag{8-51}$$

or

$$i_{b_{\max}} - I_{bt} = I_{bt} - i_{b_{\min}} \tag{8-52}$$

As a first approximation I_{bt} can be taken equal to I_{bo}. Then when the two segments of the load line on opposite sides of Q are equal, the second harmonic approximately vanishes. Lines corresponding to higher load values cross crowded characteristics at their upper ends, and formula (6-54) will yield a negative answer. A negative coefficient in a Fourier series indicates 180° phase shift relative to terms having positive coefficients. If these numbers are plotted above the axis, the second-harmonic distortion curve takes on the appearance shown in Fig. 8-14. The other desired curves are also shown in the figure. It should be noted that to obtain a power output very near the maximum the total distortion is quite high. It is often judged better to sacrifice power in order to get more moderate distortion. In fact, the usual practice is to operate near the load value which gives minimum total distortion. Fortunately the sacrifice in power is not severe in the case of the tube whose characteristics are shown. To recognize the proper load to use without the labor just involved in this case is not too easy. However, it should be noted that the point of zero second harmonic is nearby, the total distortion for this load value being but slightly more and the power output being greater. Hence the load line which yields the relation (8-51) is commonly used in the absence of specific information.

If the value of E_{c1o} is not known, one may proceed as follows. In practice $i_{b_{\min}}$ is very close to zero, and $i_{b_{\max}}$ does not differ greatly from the value on the $e_{c1} = 0$ curve at the lower part of the knee (point A in Fig. 8-12). Hence I_{bo} is approximately $i_{b_{\max}}/2$. With the allowable plate dissipation known, E_{bo} may be chosen as any desired value up to the maximum limit imposed by this dissipation. With the Q point approximately located, one proceeds as above to find a suitable load line.

With the signal applied, the average plate current is given by Eq. (6-52). If this is not equal to I_{bo}, the T point moves away from Q in a manner determined by the method of obtaining grid bias. Should the change in current be considerable, the new T point should be determined (see Art. 6-17) and a retrial for proper load line made.

After the load has been determined, it is well to check the screen dissipation against the allowable value. The dissipation is given by a formula similar to that for plate dissipation.

$$P_{g2} = \frac{i_{c2_{\max}} + i_{c2_{\min}} + 2I_{c2t}}{4} E_{c2o} \tag{8-53}$$

The method given above for determining the proper load is laborious, and often the recommendations of the tube manual can be followed with satisfactory results. The conversion chart of Art. 8-6 is based upon triode theory, but the plate current for a pentode also follows an approximate three-halves power of the screen voltage; hence if the plate voltage, the screen voltage, and the control-grid voltage [see Eq. (8-46)] vary proportionally, the relative distribution of electrode currents in the tube is unchanged and the plate current follows the rule of Eq. (8-44). Hence the triode conversion chart is approximately applicable to pentode-type power tubes. To the formulas (8-42) to (8-47) must be added one giving the new required screen voltage.

$$\frac{E'_{c2o}}{E_{c2o}} = \frac{E'_{bo}}{E_{bo}} \tag{8-54}$$

Thus if one satisfactory set of operating conditions is found in a tube manual, another satisfactory set with a somewhat changed value of E_{bo} can be determined. *Note:* A check on the plate dissipation should always be made if E_{bo} is raised.

Besides the foregoing difficulties in determining the correct load to use, more troubles beset the user of a pentode-type power tube. One is caused by the fact that the load may not be constant at all frequencies, this being especially the case for loudspeaker loads. For a typical speaker the variation of impedance with frequency is shown in the graph of Fig. 8-15. This variation is largely due to the fact that a speaker is an electromechanical system which has electrical qualities of resistance, inductance, and a small amount of capacitance, in addition to dissipative, inertia, and spring actions, which to the electrical driver appear as more of the same electrical qualities, respectively. Resonances occur near the top left peak and near the bottom dip, these being the only points at which the equivalent impedance is entirely resistive. The output-transformer turns ratio is usually chosen to make the 400-cps (dip) value of speaker resistance appear to be the correct load to the tube. Thus throughout much of the frequency range the tube has a load impedance higher than the optimum value. For a triode this causes some loss in power (about 11 per cent for a twice normal load), but the distortion actually decreases below the normal value. However, for a pentode the distortion increases greatly, first because of greater nonlinear distortion, and second because different frequency components of the signal are not amplified the same amount, the high frequencies and those around 90 cps being favored the most. This means that low frequencies which have their harmonics near 90 cps will have an apparent increase in harmonic distortion and also h-f overtones in the output will be overemphasized, giving an unnatural sound to the output.

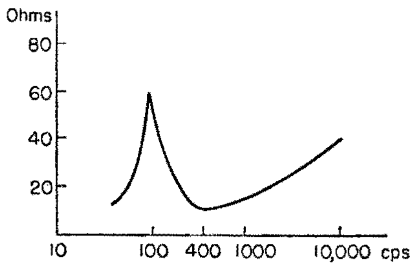


FIG. 8-15. Variation of impedance with frequency for a typical loudspeaker.

One treatment which tends to cure this effect of changing load is to shunt the output transformer primary by a network, usually R-C, whose impedance varies with frequency in a manner somewhat to offset the changes in speaker impedance. A second treatment is the use of negative feedback. As we shall see in a later chapter, this can cancel to a considerable degree the distortion produced in the tube by the nonmatching load along with distortion due to other causes.

For a speaker to give good reproduction of transients it needs considerable damping, *i.e.*, a way quickly to dissipate stored energy. In addition to that naturally possessed by a speaker there is the beneficial effect of the plate resistance of the power tube as seen by the speaker coil through the output transformer. The vibrating coil moves in a magnetic field and thereby becomes a generator which drives current through any impedance shunted across it. The flow of this current dampens the movement of the vibrating coil. Naturally, the lower the shunting impedance, the greater the damping effect. Figure 8-16 illustrates the effective impedance shunting the voice coil of a dynamic loudspeaker for a 2A3 triode and for a 6V6 beam tube in typical cases. The output impedance for a pentode-type tube is quite high, and the damping effect less satisfactory. However, the nominal loudspeaker impedance of 10 ohms, shown in the diagrams of Fig. 8-16, rises to many times this value at some frequencies, and this increases the relative damping. The use of negative voltage feedback reduces the equivalent output impedance of the tube and offers a satisfactory method of increasing the damping as well as giving other benefits.

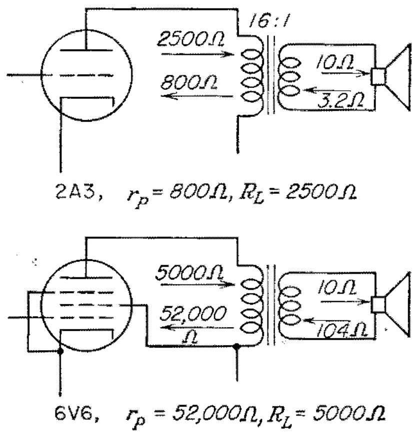


FIG. 8-16. A loudspeaker is damped by the power tube if the tube's output impedance is low.

8-8. Parallel Operation of Tubes for Greater Output Power. If the power output obtained from one tube is satisfactory in quality but insufficient in quantity, the latter may be doubled with a minimum outlay of expense and trouble by using parallel operation. If another power tube of approximately identical characteristics is added to the circuit and corresponding electrodes of the two tubes are connected together, the operation may be made satisfactory by changing the bias resistor, if cathode bias is used, and by changing the output transformer to a different turns ratio. For the same bias with double the cathode current, the value of R_k should be halved. The plate characteristics for the combination of the two tubes may be made by changing the values on the current axis for one tube to double their former amount. Thus the equivalent plate resistance will be halved, and the required matching load becomes half its former value; a new transformer turns ratio should be used to give this. The required grid drive is not changed, nor is the harmonic distortion reduced.

8-9. Push-pull Operation.[1] Under certain conditions more than twice the power output may be obtained with no increase in distortion, if push-pull operation is used. The circuit for a push-pull amplifier using triodes with fixed bias is shown in Fig. 8-17. The two grid signals are 180° out of phase and are usually furnished by a balanced amplifier or by a phase inverter of one of the types discussed in Art. 7-22. In all the discussion that follows the two tubes are considered to be identical.

The circuit being almost the same as that of a balanced amplifier studied in Art. 7-21, the equivalent-circuit method used there can be applied if the grid drive is small enough to assure approximately linear operation. This equivalent circuit is shown in Fig. 7-45a and is simplified to a more convenient arrangement in Fig. 7-45b, in which an equivalent tube with plate resistance $2r_p$ and amplification factor 2μ operates in a series circuit with the entire load for both tubes.

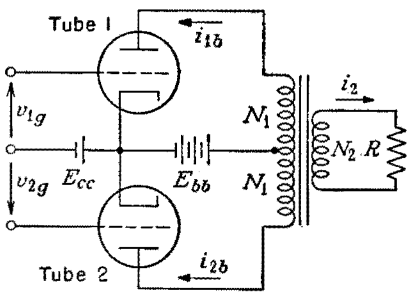


FIG. 8-17. A push-pull amplifier using fixed bias.

Since the purpose of a power amplifier is to obtain a large power output, the grids are usually supplied with large signals which render the operation far from linear. The equivalent-circuit method is then not very good for computing the power output and of course useless for determining the distortion, even if the proper load were known. A graphical method is needed. We already have studied successful graphical methods for single-sided (one tube or two parallel tubes) amplifiers. Let us see what we can do with this push-pull amplifier to convert it to an equivalent single-sided one.

The first thing to do with the amplifier circuit drawn in Fig. 8-17 is figuratively to pull both tubes from their sockets. Then we replace, say, the No. 1 tube by a new type of tube. If the new tube delivers the same power with the same current waveform to the load R as did the two original tubes, then we have successfully reduced the problem to a single-sided basis. It then remains to determine the characteristics of the new tube, or composite tube as it is called, and go about the familiar part of the problem—finding the correct load, the a-c power, the distortion, etc. Calculations pertaining to d-c quantities must be made from the characteristics of the real tubes since the composite-tube equivalence is for alternating quantities only.

In order to find the required characteristics for the composite tube, let us start with the output transformer, diagramed in Fig. 8-18. If it is of good quality, it may be considered as approximately ideal, and among other things the mmf drop due to the reluctance of the iron core may be neglected when the full-load currents are flowing. If we do this, the mmf drop (measured in ampere turns) around the magnetic path *abcda* is

$$N_1 i_{2b} - N_1 i_{1b} + N_2 i_2 = 0 \tag{8-55}$$

or

$$i_2 = \frac{N_1}{N_2}(i_{1b} - i_{2b}) = \frac{N_1}{N_2} i_b' \tag{8-56}$$

where

$$i_b' = i_{1b} - i_{2b} \tag{8-57}$$

i_b' is the current required to flow through half the primary turns N_1 in order to make i_2 flow from N_2 secondary turns through the load R. Since the composite tube feeds into the upper half of the primary turns and no

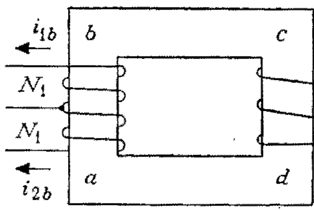


FIG. 8-18. The output transformer.

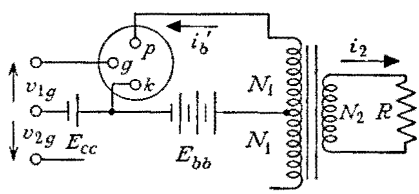


FIG. 8-19. Circuit with composite tube.

current flows into the lower half, i_b' is the required plate current for the composite tube.

The circuit for the composite tube is shown in Fig. 8-19, it being identical with that of Fig. 8-17 except that tube 2 is removed and tube 1 is replaced by the composite tube. The circuit is not like the linear equivalent ones, with which we are familiar, in that the d-c supplies are still in the circuit and the currents and voltages are total ones, not just alternating components.

Let us determine the operating conditions for the composite tube in terms of those of the original tubes, using primes to indicate composite-tube quantities and subscripts 1 and 2 to refer to the respective real-tube quantities. Comparing circuits, $E_{cc}' = E_{1cc}$, and $v_g' = v_{1g}$; therefore

$$e_c' = e_{1c} \tag{8-58}$$

If the waveform and amount of voltage applied to R are to be the same with the composite tube as with the original tubes, the flux in the winding of the secondary also must be the same as before. Hence the voltage induced in N_1 must be the same. This was e_{1p} originally and is now e_p'. Hence we may write $E_{bb}' = E_{1bb}$, $e_p' = e_{1p}$, and therefore

$$e_b' = e_{1b} \tag{8-59}$$

But it will be remembered that

$$i_b' = i_{1b} - i_{2b} \tag{8-57}$$

which completes the determination of the operating conditions for the composite tube.

The three relations of (8-57), (8-58), and (8-59) are sufficient to determine the characteristics of the composite tube, but it is helpful to develop better working relations. The operation of tube 1 at any moment is related to the operation of tube 2 at the same moment. It is possible to change the procedure of determining the characteristics to make the voltages on tube 1 alone, or on the composite tube alone, since these are the same, be a guide in choosing the currents to be subtracted according to Eq. (8-57). Since the plate current of a triode, or of a pentode-type tube with fixed screen voltage, depends upon the grid and the plate voltages, one may write

$$i_{1b} = f(e_{1b},e_{1c}) \qquad \text{and} \qquad i_{2b} = f(e_{2b},e_{2c}) \tag{8-60}$$

where

$$e_{1b} = E_{bo} + e_{1p} \qquad \text{and} \qquad e_{2b} = E_{bo} + e_{2p} \tag{8-61}$$

or since the flux in both halves of the primary is the same,

$$e_{2p} = -e_{1p} \tag{8-62}$$

and therefore

$$e_{2b} = E_{bo} - e_{1p} \tag{8-63}$$

Adding e_{1b} of (8-61) to e_{2b} of (8-63) and transposing yields

$$e_{2b} = 2E_{bo} - e_{1b} \tag{8-64}$$

which may be used to change the plate voltage of tube 2 to terms of tube 1. Also we may write

$$e_{1c} = E_{co} + v_{1g} \qquad \text{and} \qquad e_{2c} = E_{co} + v_{2g} \tag{8-65}$$

But

$$v_{2g} = -v_{1g} \tag{8-66}$$

and hence

$$e_{2c} = E_{co} - v_{1g} \tag{8-67}$$

Adding e_{1c} of (8-65) to e_{2c} of (8-67) yields

$$e_{1c} + e_{2c} = 2E_{co} \tag{8-68}$$

which is a handy relation between the instantaneous total grid voltages. Let us now substitute e_{1c} from (8-65) into i_{1b} of (8-60) and also substitute e_{2b} of (8-64) and e_{2c} of (8-67) into i_{2b} of (8-60).

$$i_{1b} = f(e_{1b}, E_{co} + v_{1g}) \qquad \text{and} \qquad i_{2b} = f(2E_{bo} - e_{1b}, E_{co} - v_{1g}) \tag{8-69}$$

Then

$$i'_b = f(e_{1b}, E_{co} + v_{1g}) - f(2E_{bo} - e_{1b}, E_{co} - v_{1g}) \tag{8-70}$$

is the formula which may be used to determine the characteristics of the composite tube. To do this, only the characteristics of one of the two identical tubes are needed. However, if two curve sheets are available, the work is somewhat simplified, particularly in determining the average current and the d-c power requirements.

Since formula (8-70) contains both E_{bo} and E_{co}, it is necessary to know these values before the characteristic curves for the composite tube can be drawn and of course a new set of curves is needed if the operating point is changed. In drawing a characteristic curve for a fixed e_c' value, several

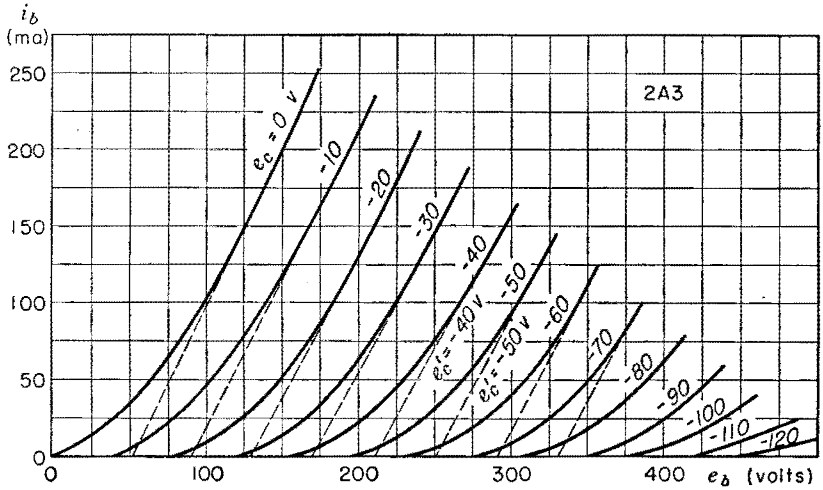


FIG. 8-20. Curves for the 2A3 tube and for the composite tube.

points are needed unless the curve is straight, which is in general not true. Such a curve shows composite-tube current i_b' for many values of e_b', and the current at each point on the curve is obtained by subtracting the current flowing in the No. 2 tube from that flowing in the No. 1 tube. The principal difficulty is in following Eq. (8-70) in matching values correctly. The procedure is best illustrated by examples.

Using 2A3 tubes, let us first carry out the process with but one sheet of characteristic curves (see Fig. 8-20). If we assume a value of E_{bo} equal to 250 volts, the highest safe value of I_{bo} is 60 ma, since the allowable plate dissipation is 15 watts. If we use a-c heating of the filaments, this operating point corresponds to a bias of about -44 volts. (The curves on the sheet are for d-c heating, in which case the grid already has 1.25 volts more bias than is labeled on the curves because the grid return is to the negative side of the filament battery.) In order to simplify our problem somewhat, we shall use 50 volts bias and neglect the error in curve labeling. We may first construct the curve for $e_c' = -50$ volts. The grid signal v_g' ($= v_{1g}$) required to drive to this value is zero, and hence by relation (8-70) the currents of the curve for $e_{2c} = -50 - 0 = -50$ volts are subtracted from the currents of the curve for $e_{1c} = -50 + 0 = -50$ volts. The No. 1 tube current may be taken at any plate voltage e_{1b}, and the No. 2 tube current to be subtracted from it must then be taken at a plate voltage $2E_{bo} - e_{1b} = 500 - e_{1b}$ volts. For $e_{1b} = 250$ volts as the first value, $i_{1b} = 31$ ma. The corresponding value of e_{2b} is also 250 volts with $i_{2b} = 31$ ma, making $i_b' = 0$, which we plot against $e_b' = 250$ volts on the same sheet. Moving now to $e_{1b} = 275$ volts where $i_{1b} = 60$ ma, and to $e_{2b} = 500 - 275 = 225$ volts where $i_{2b} = 13$ ma, we obtain $i_b' = 47$ ma, which is plotted vs. $e_b' = 275$ volts. If $e_{1b} = 300$ volts, $i_{1b} = 93$ ma, $e_{2b} = 200$ volts, $i_{2b} = 2$ma, and $i_b' = 91$ ma at $e_b' = 300$ volts. For $e_{1b} = 325$ volts, $e_{2b} = 175$ volts, and $i_{2b} = 0$; hence the value of i_b' is the same as i_{1b} and in the same position on the curve for $e_{1c} = -50$ volts. A smooth curve may now be drawn through the points obtained and continued exactly on the curve for $e_{1c} = -50$ volts after this curve is once reached. The new curve is now labeled $e_c' = -50$.

To draw the curve for $e_c' = -40$ volts, we take current values from the curves for $e_{1c} = -50 + 10 = -40$ volts and $e_{2c} = -50 - 10 = -60$ volts. For $e_b' = 250$ volts, $e_{1b} = 250$ volts, $i_{1b} = 76$ ma, and $e_{2b} = 250$ volts, $i_{2b} = 7$ ma, making $i_b' = 69$ ma. For $e_{2b} = 235 = 250 - 15$ volts, i_{2b} cuts off, and $e_{1b} = 250 + 15 = 265$ volts, $i_{1b} = 100$ ma. Thus $e_b' = 265$ volts, $i_b' = 100$ ma. In like manner, taking one plate voltage any amount less than 250 volts and the other plate voltage the same amount more than 250 volts, the currents may be rapidly determined and the curve for the composite tube drawn. Several of these curves are shown in the figure.

8-10. The Analysis of the Performance of a Push-pull Amplifier. As a second example let us draw in Fig. 8-21 the curves for a composite tube replacing two 2A3 tubes operating with $E_{bo} = 300$ volts and a fixed bias of $E_{co} = -60$ volts. Since the curve sheet for the No. 2 tube has been inverted and placed below that for the No. 1 tube with the axis markings for 300 volts matched, the values of e_b which are used together are always found together in a vertical line. For example, when $e_{1b} = 300 - 50 = 250$ volts, e_{2b} should be $300 + 50 = 350$ volts, and these values will be found to match so that it is necessary only to read current values above and below the value $e_b' = 250$ and subtract them to obtain i_b'. The arrangement also makes it easy to extend the characteristic curves for the composite tube to negative values of i_b'. We may now rapidly complete the drawing of the curves.

Examination of these characteristic curves for the composite tube in Fig. 8-21 shows the characteristics to be fairly straight and parallel, and the spacing is also much more uniform than for any real tube. Digressing a bit, if values of grid bias progressively more negative than -60 volts

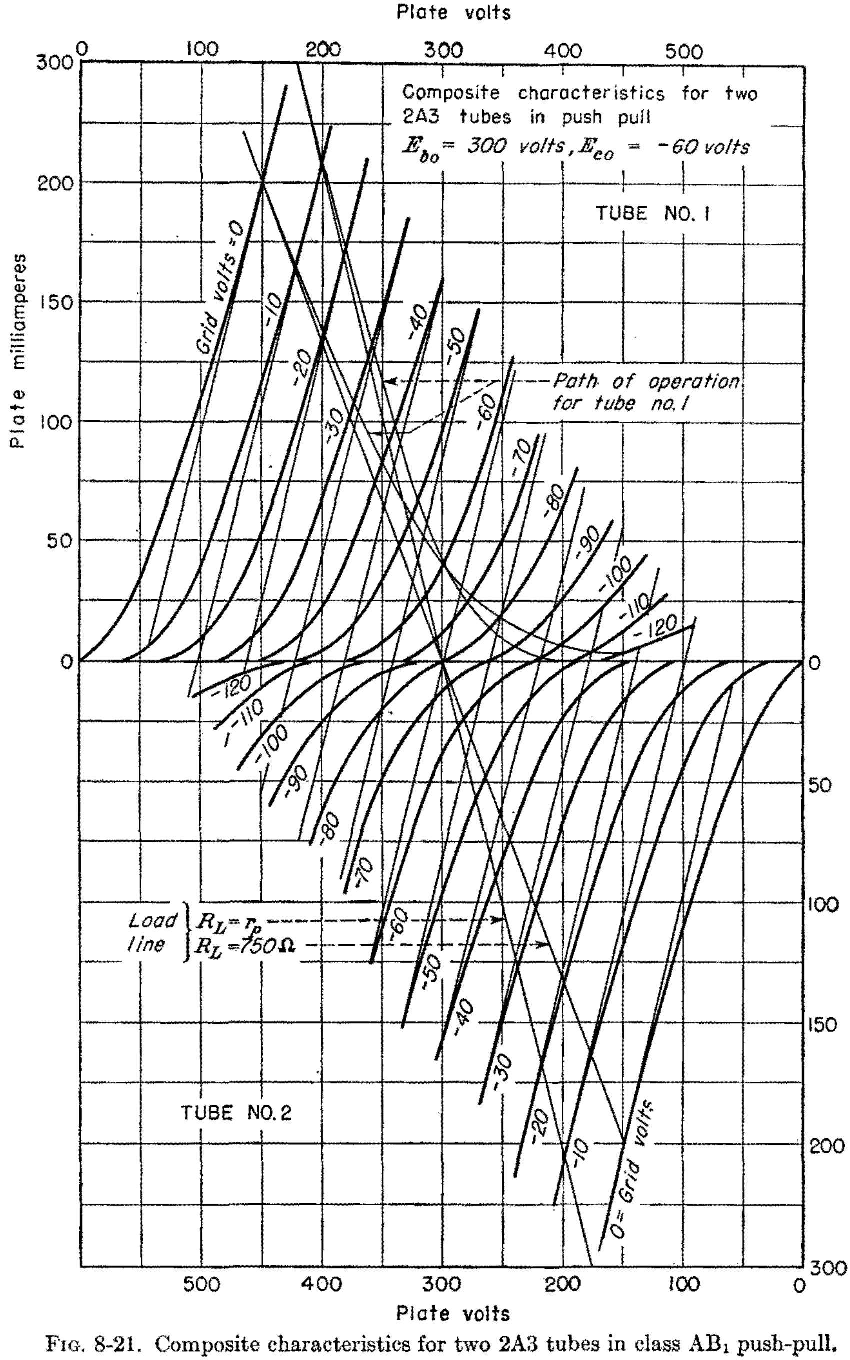


FIG. 8-21. Composite characteristics for two 2A3 tubes in class AB_1 push-pull.

are used, the characteristics become both increasingly bent and not so uniformly spaced. For a value of $E_{co} = -80$ volts, which is the bias for exact class B operation, the composite-tube characteristics are almost identically those of the two real tubes as shown, for example, by the original curves of Fig. 8-21.

Returning to our example, for amplifier operation with composite characteristics so nearly ideal as these, the distortion even for full class 1 drive is of course quite small. Since the curves are almost parallel, any load line drawn through the operating point, $E'_{bo} = 300$ volts, $E'_{co} = -60$ volts, will be cut in the same proportions by the characteristic curves and the distortion is almost the same for all loads. Furthermore, if the load line corresponding to a certain load crosses the curve for $e'_c = 0$ anywhere near its upper extremity, the position of this intersection is unchanged by moderate changes in bias. This is true because, if a new value of E'_{co} is adopted, not greatly different from -60 volts, new characteristic curves may be drawn but the one for $e'_c = 0$ will have the same position in its upper part. With $i'_{b\max}$ and $i'_{b\min}$ $(= -i'_{b\max})$ remaining the same and the load unchanged, the power output remains constant. Hence it appears that changing the bias (and corresponding drive) a moderate amount does not affect the power output, nor does changing the load affect the distortion greatly. An approximate rule for choice of load under these conditions is to let $R'_L = r'_p$ [Eq. (8-19)]. This means that the transformer turns ratio may be chosen to make the impedance looking into the upper half of the transformer appear to be equal to r'_p. A tentative load line corresponding to this amount has been drawn in Fig. 8-21.

This rule for choice of load line may be challenged, for it is important to check the plate dissipation with full drive applied against the allowable amount. In this case the first step is to obtain the a-c power output of the amplifier, which is that of the composite tube. From the graph of Fig. 8-21 we obtain $i'_{b\max} = 260$ ma, $i'_{b\min} = -260$ ma, $e'_{b\max} = 423$ volts, $e'_{b\min} = 177$ volts. Since the distortion is small, the simpler formula (6-41) and its voltage paraphrase suffice for this approximate determination.

$$P_{ac} = \frac{423 - 177}{2\sqrt{2}} \frac{0.260 + 0.260}{2\sqrt{2}} = 16 \text{ watts} \tag{8-71}$$

The path of operation of the composite tube is of course the load line. That for the No. 1 tube may be obtained by plotting points showing corresponding values of e_{1c} and e_{1b}, remembering that $e_{1c} = e'_c$ and $e_{1b} = e'_b$. This path of operation is shown in Fig. 8-21.

The approximate value of the average current through tube 1 may be obtained by using formula (6-52). Here $i_{b\max} = 260$ ma, $i_{1/2} = 132$ ma, $i_{-1/2} = 6$ ma, $i_{b\min} = 0$, and hence

$$I_{1ba} = \frac{260 + 2 \times 132 + 2 \times 6 + 0}{6} = 89 \text{ ma} \tag{8-72}$$

The power furnished by the plate supply to tube 1 and its load is

$$P_{bb} = E_{bb}I_{1ba} = 300 \times 0.089 = 26.7 \text{ watts} \tag{8-73}$$

This makes the plate dissipation of tube 1

$$P_{1p} = 26.7 - {}^{16}\!/_{2} = 18.7 \text{ watts} \tag{8-74}$$

which exceeds the allowable value of 15 watts.

In push-pull amplifiers with fixed bias the average current is often considerably more than the quiescent value, and an early step in a design should be a check on the plate dissipation with full grid drive applied. This applies to class A amplifiers as well as to the more efficient classes. As a consequence of this great variation in average current with amplitude of grid signal, power supplies with good regulation characteristics are required in all cases of fixed-bias operation.

We may now continue our example and choose a more suitable load. This is a trial-and-error process, and the results of using a load of 750 ohms for the composite tube are worked out here. The load line is drawn and labeled in Fig. 8-21. The corresponding path of operation of the No. 1 tube is also drawn. For the composite tube $i'_{b\max} = 200$ ma, $e'_{b\max} = 450$ volts, $e'_{b\min} = 150$ volts. Hence

$$P_{ac} = \frac{450 - 150}{2\sqrt{2}}\,\frac{0.200 + 0.200}{2\sqrt{2}} = 15 \text{ watts} \tag{8-75}$$

For the No. 1 tube $i_{b\max} = 200$ ma, $i_{1/2} = 108$ ma, $i_{-1/2} = 17$ ma, $i_{b\min} = 3$ ma. Hence

$$I_{1ba} = \frac{200 + 2 \times 108 + 2 \times 17 + 3}{6} = 75.5 \text{ ma} \tag{8-76}$$

and the power furnished the No. 1 tube and load by the plate supply is

$$\tfrac{1}{2}P_{bb} = 300 \times 0.0755 = 22.6 \text{ watts} \tag{8-77}$$

The plate dissipation of tube 1 is therefore

$$P_{1p} = 22.6 - {}^{15}\!/_{2} = 15.1 \text{ watts} \tag{8-78}$$

Accepting this as a satisfactory value, the composite-tube load of 750 ohms is retained. The plate efficiency is

$$\eta_p = \frac{P_{ac}}{P_{bb}} = \frac{15}{2 \times 22.6} = 33\% \tag{8-79}$$

The second-harmonic distortion is zero, since by Eq. (6-42)

$$I'_{p_2} = \frac{200 - 200 - 2 \times 0}{4} = 0 \tag{8-80}$$

A formula for the percentage of third-harmonic distortion is obtained by dividing Eq. (6-55) by (6-53). Upon substituting values we obtain

$$\text{3d harmonic} = \frac{200 - 2 \times 97 - 2 \times 97 + 200}{2(200 + 97 + 97 + 200)} \times 100 = 1.0\% \qquad (8\text{-}81)$$

This value will be increased somewhat in practice if the regulation of the plate power supply is imperfect.

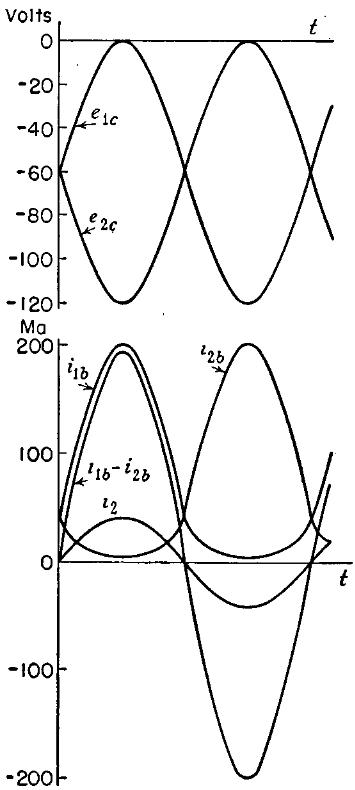


FIG. 8-22. Waveforms associated with the example of Art. 8-10.

Figure 8-22 shows waveforms of quantities associated with this example.

The 750-ohm load for the composite tube results in an effective plate-to-plate impedance looking into the whole transformer primary (with tubes removed from their sockets) of $4 \times 750 = 3000$ ohms. If one consults a tube manual for recommended operating values, the loading is usually given in plate-to-plate form. Sufficient turns ratio in the output transformer should then be used to give this impedance value. The actual impedance load on each individual tube varies widely during the cycle as can be seen by noting the variation in the slope of the operating curve. At the positive extreme of grid swing the load is approximately 750 ohms, when $e_c = -60$ volts the load is 1500 ohms, while when $e_c = -120$ volts the load is very high. The actual distortion in each individual tube is very high, but this fact is of no great importance.

8-11. Bias Values and Loads for Various Classes of Operation in Push-pull. In the example of Art. 8-10 the operation with the first load chosen, $R'_L = r'_p$, is class AB_1 since the plate current of each tube is cut off for a part of the cycle. This load value causes tube overheating, and a new load of greater ohmic value is chosen which gives satisfactory operation. Examination of the path of operation of the No. 1 tube in the latter case shows the operation to be class A_1. As far as distortion in the output is concerned, it makes little difference whether the operation is class A_1 or slightly AB_1 and the designer ordinarily needs to give the matter no great attention.

In cases where the rule $R'_L = r'_p$ can be safely used it is sometimes desirable to choose a bias value which results in a particular class of operation. The limiting bias value for class A_1 operation may be readily approximated. Assume the composite characteristics are straight and parallel and that the line for $e'_c = 0$ follows the curve for $e_{1c} = 0$ in the region where the load line crosses it. For $e_{1c} = 0$ it is approximately true that

$$i_{1b} = ke_{1b}^{3/2} \tag{8-82}$$

and therefore

$$\frac{\partial e_{1b}}{\partial i_{1b}} = \frac{2}{3k} e_{1b}^{-1/2} \tag{8-83}$$

which gives the value of r'_p if we let $e_{1b} = e_{b_{\min}}$.

$$r'_p = \frac{2}{3k} e_{b_{\min}}^{-1/2} \tag{8-84}$$

It is also true that

$$\frac{E_{bo} - e_{b_{\min}}}{i_{b_{\max}}} = R'_L = r'_p \tag{8-85}$$

or

$$\frac{E_{bo} - e_{b_{\min}}}{ke_{b_{\min}}^{3/2}} = \frac{2}{3k} e_{b_{\min}}^{-1/2} \tag{8-86}$$

whence

$$e_{b_{\min}} = \tfrac{3}{5}E_{bo} \tag{8-87}$$

It follows that

$$e_{b_{\max}} = 2E_{bo} - \tfrac{3}{5}E_{bo} = 1.4E_{bo} \tag{8-88}$$

This relation holds quite well for class A_1 push-pull amplifiers and even for class AB_1 as long as the characteristics are approximately ideal. It does not hold if $R'_L = r'_p$ is not the load used.

Under the restriction stated, the bias for class A_1 operation may be any value between that suitable for single-tube operation (see Arts. 8-4 and 8-5) and half the value of grid voltage which cuts off the plate current of the No. 1 tube at $e_{b_{\max}} = 1.4E_{bo}$ given by Eq. (8-88). Full grid drive will then never cause the plate current of either tube to be zero over any portion of the operating cycle. In the example of Art. 8-10 the highest bias for class A_1 operation is about -58 volts. Between this value and -80 volts, the value which cuts off the plate current at E_{bo}, the operation is class AB. Operation with bias higher than -80 volts is unsatisfactory because of the very high distortion produced, unless a filter such as a tuned load is used to eliminate the harmonics.

8-12. Push-pull Operation Using Cathode Bias. Cathode bias may be employed in push-pull operation as shown in Fig. 8-23. The advantages are threefold: the difficulty of arranging for a fixed-bias supply is avoided, and the need for a plate power supply of very good regulation is eliminated since the plate current does not vary greatly with signal amplitude. It is safer for the tube since the average current cannot rise greatly.

In order to operate with the same quiescent values as in the example of Art. 8-10, the value of cathode resistor needs to be

$$R_k = \frac{-E_{co}}{2I_{bo}} = \frac{60}{2 \times 0.040} = 750 \text{ ohms} \tag{8-89}$$

If a plate-to-plate load of 3000 ohms is used as before and it is first assumed that the average current is 75.5 ma with full grid drive applied, the bias rises to $(75.5/40) \times -60 = -113$ volts. With only 60 volts peak grid drive applied, the tube is cut off considerably more than half the time, showing that the average current cannot be 75.5 ma. One might guess that with full grid excitation it rises from the quiescent value of 40 ma to 45 ma. Then the bias rises from -60 volts to $\frac{45}{40} \times -60 = -67.5$ volts. With this new trial bias one may construct an entire new set of composite characteristics, calculate the average current, and check against the assumed value of 45 ma. Then one needs to check the plate dissipation and the harmonic distortion, probably discard the load for another more suitable one, and do all the work over again. To make a long story short, the process of graphically analyzing the performance for cathode bias is not easy, and it is perhaps better to determine the best load by laboratory methods. Generally speaking, the distortion becomes considerably higher, and to reduce it, a higher load should be used than for fixed bias. The amount of power output is usually reduced with the new load. If desired, a tube manual may be consulted for recommendations for cathode-bias operation, but unfortunately no data are given for many of the popular varieties of tubes despite the fact that cathode bias is the usual condition of operation.

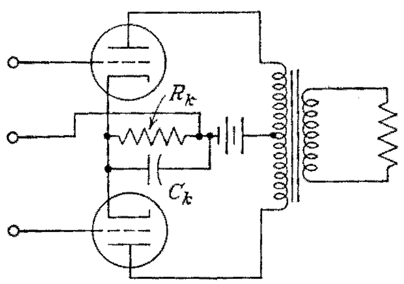


FIG. 8-23. A push-pull amplifier employing cathode bias.

8-13. Class B Amplifiers. When a large amount of power at audio frequencies is needed, as, for example, in modulating the output of a radio transmitter, class B operation is used. A properly designed amplifier of this class can deliver kilowatts of a-f power at efficiencies approaching the optimum of 78.5 per cent and with distortion as low as with ordinary class A amplifiers.

While any power tube useful in class A or class AB can be operated in class B, the tubes which perform best in this service are high-μ triodes. Some of these are designed to operate with zero bias, and while this operation is actually class AB_2, the quiescent plate current is so small that it is often labeled class B. Most larger class B amplifier tubes use negative grid bias sufficient to reduce the quiescent plate current to almost zero.

In general the procedure for class B design parallels that for any other push-pull a-f power amplifier, the principal difference being the necessity for furnishing grid power. Also there is no reliable rule for choosing a proper load except to use a trial-and-error method. Even though the distortion is low, the rule of Art. 8-4 for using $R'_L = r'_p$ does not apply. That rule was for the case of negligible distortion, a fixed operating point,

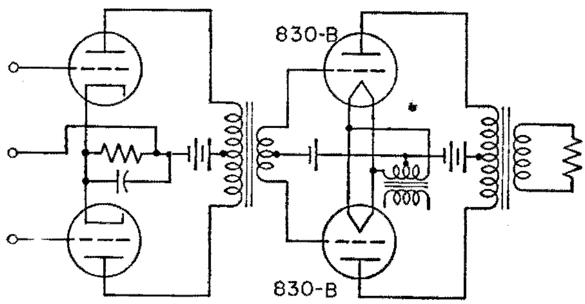


FIG. 8-24. A class B amplifier and its driver.

and a fixed drive. In this case, however, the operating point is fixed but the drive is limited by the relation of $e_{b_{\min}}$ to $e_{c_{\max}}$ ($e_{b_{\min}}/e_{c_{\max}}$ is called p, and this matter is further discussed in item 5 below). An analysis can be made, using idealized plate characteristics, which shows that the load should be

$$R'_L \approx \frac{p}{p+\mu} r'_p \tag{8-90}$$

for maximum power output. Use of this formula, however, generally gives values of plate dissipation which are too high. Since there is no simple relation for correcting the load to lower the plate dissipation to just the allowable amount, it is necessary to proceed by trial and error. In general, a load resistance greater than that given by formula (8-90) gives lower power output but less plate dissipation and higher plate efficiency.

The class B case will be further explained by means of an example. Figure 8-24 shows a circuit diagram for the class B output stage and its driver.

1. For the output stage let us use a pair of RCA type 830-B power triodes. For each tube $E_{ff} = 10$ volts, $I_{ff} = 2$ amp, $\mu \approx 25$. Maximum $E_{bo} = 1000$ volts, maximum $I_{ba} = 150$ ma, maximum $P_p = 60$ watts.

2. Let $E_{bo} = 1000$ volts.

3. Figure 8-25 shows the plate characteristics for a type 830-B tube. The composite-tube plate characteristic which passes through its operating point is not yet drawn. However, it will be approximately parallel to those for somewhat less negative grid voltages, which are identical for both the composite and the No. 1 tubes. Hence we can determine an approximate average plate resistance of the composite tube by determin-

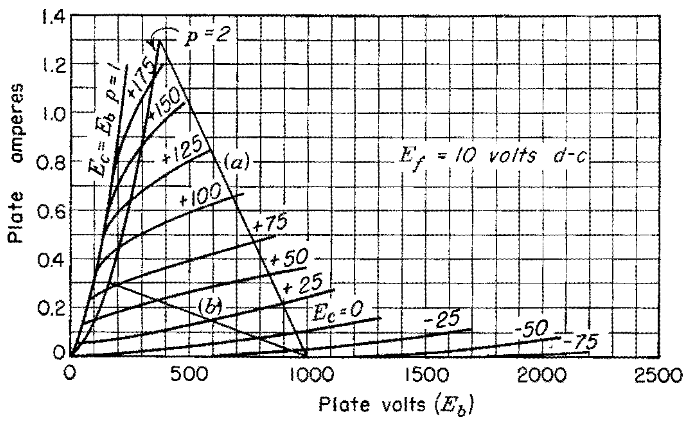


FIG. 8-25. Class B operation with an 830-B tube.

ing the reciprocal of the slope of, say, the $e_c = 0$ volt curve. This is approximately 6500 ohms in this case. By Eq. (8-90) the theoretical optimum load is $R_L' = \left[\frac{2}{(2 + 25)}\right] \times 6500 = 480$ ohms. We have assumed $p = 2$ for reasons explained in item 5.

4. We may now draw a load line (a) for $R_L' = 480$ ohms, through the point determined by $e_b = 1000$ volts, $i_b = 0$. This is half the operating curve for the composite tube. It is also the operating curve for the No. 1 tube during the active part of its cycle if the operation is exactly class B.

5. We now choose a value for the maximum drive point P. On the left side of the plate characteristics is drawn the locus of $e_b = e_c$, called the diode line. To the left of this line the grid voltage exceeds the plate voltage, and operation here is unsatisfactory. Even somewhat to the right of the diode line the spacing of the plate characteristics and the high grid current are likely to result in excessive distortion and grid heating. It is not uncommon to limit the drive to a condition where e_b equals two or more times e_c. The locus for $e_b = 2e_c$ ($p = 2$) is shown as the right-hand

limiting curve. We shall use this condition for this example, and the maximum drive point P then yields $i_{b_{max}} = 1.3$ amp, $e_{b_{min}} = 375$ volts, and $e_{c_{max}} = 187$ volts. Since the plate current is approximately a half sinusoid, we can compute for the No. 1 tube: $P_{bb} = E_{bb}I_{ba} = E_{bb}(i_{b_{max}}/\pi) = 1000(1.3/\pi) = 414$ watts. $P_{ac} = I_{p_1}{}^2R_L = (i_{b_{max}}/2)^2 R_L = (1.3/2)^2 \times 480 = 203$ watts. $P_p = 414 - 203 = 211$ watts, which greatly exceeds the allowable value of 60 watts.

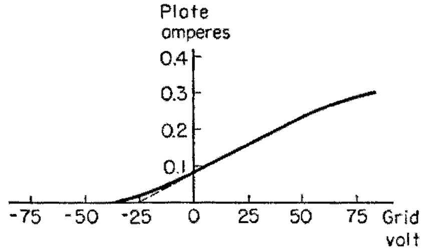


FIG. 8-26. Dynamic transfer characteristic for a type 830-B tube with $E_{bo} = 1000$ volts, $R_L = 1950$ ohms.

6. As a second trial load let us use $R'_L = 2800$ ohms, which is $0.43r'_p$.

7. A load line (b) for $R'_L = 2800$ ohms has been drawn in Fig. 8-25. Again using $p = 2$, we obtain $i_{b_{max}} = 0.3$ amp, $e_{b_{min}} = 160$ volts, $e_{c_{max}} = 80$ volts. For the No. 1 tube, $P_{bb} = 1000(0.3/\pi) = 95.4$ watts, $P_{ac} = (0.3/2)^2 \times 2800 = 63.0$ watts, and $P_p = 95.4 - 63.0 = 32.4$ watts, which is less than 60 watts. Hence we shall continue the computations for operation with this load.

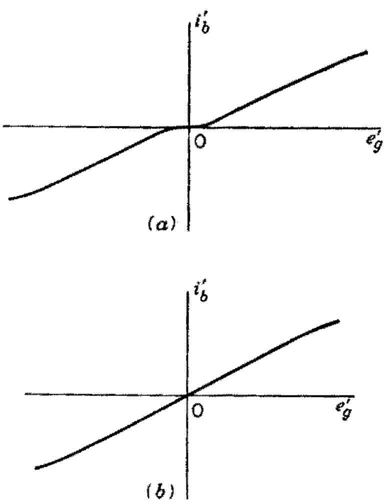


FIG. 8-27. The dynamic transfer characteristic for the composite tube (a) if exact cutoff bias is used, (b) if extended cutoff bias is used.

8. We can pick values from this operating path of the No. 1 tube and plot a dynamic transfer characteristic as in Fig. 8-26. The value of grid bias required to cut off the plate current is approximately -37 volts, and operation with E_{cc} equal to this value gives exact class B. Figure 8-27a shows the shape of the resulting dynamic transfer characteristic for the composite tube. It is readily seen that the distortion produced by such operation will be severe. On the other hand, if the grid bias is made somewhat less negative, the linearity of the dynamic transfer characteristic is improved. For $E_{cc} = -25$ volts, the value obtained by projecting the linear portion of the characteristic upon the grid-voltage axis (called the projected cutoff value), the shape of the dynamic transfer characteristic for the composite tube is shown in Fig. 8-27b. Note its excellent linearity. While operation with projected cutoff value of bias is actually

class AB, when high-μ tubes are employed it is near class B and is generally so called. We shall use $E_{co} = -25$ volts.

9. We obtain the following values for the operation of the composite tube: $i'_{b_{\max}} = 0.30$ amp, $i'_{b_{\min}} = -0.30$ amp, $i'_{1/2}$ (occurring at $e'_c = +27.5$ volts) $= 0.16$ amp, $i'_{-1/2} = -0.16$ amp. Hence by using formulas (6-55) and (6-53) we obtain

$$\text{Percentage of 3d harmonic} = \frac{100(0.30 - 0.32 - 0.32 + 0.30)}{2(0.30 + 0.16 + 0.16 + 0.30)} = 2.17$$

Reduced grid drive may result in less distortion although this is not a gen-

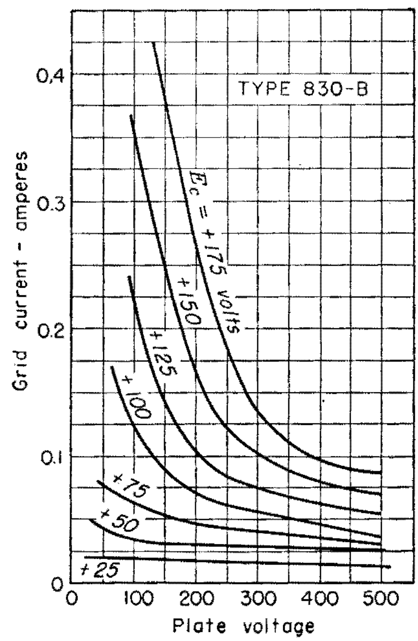


FIG. 8-28. Grid-current plate-voltage curves for type 830-B tube. *(Courtesy of RCA.)*

eral rule. However, in this case let us assume the value obtained is satisfactory.

10. Since the distortion is low, the assumption, used in item 7, that the plate-current waveform is sinusoidal is valid. Hence

$$P_{ac} = 2 \times 63.0 = 126 \text{ watts for two tubes}$$
$$P_{bb} = 2 \times 95.4 = 190.8 \text{ watts for two tubes}$$
$$P_p = 32.4 \text{ watts for one tube}$$
$$\eta_p = \frac{126}{190.8} = 66\%$$

This efficiency is quite good but P_p is low and probably a lower load resistance should be tried.

11. We next determine values of e_c corresponding to values of ωt during the cycle, assuming $e_{1g} = 105 \cos \omega t$ volts. Figure 8-28 shows the grid-current plate-voltage family of curves for this tube. From it we can determine values of i_c corresponding to the same values of ωt. These values are tabulated in Table 8-2.

TABLE 8-2

ωt	0°	15°	30°	45°	60°	75°	90°	105°	120°	135°	150°	165°	180°
e_g, volts	105	102	91	74	53	27	0	−27	−53	−74	−91	−102	−105
e_c, volts	80	77	66	49	28	2	−25	−52	−78	−99	−116	−127	−130
e_b, volts	160	180	262	380	570	700	1000						
i_c, amp	0.060	0.055	0.035	0.025	0.020	0	0	0	0	0	0	0	0
$e_g i_c$, watts	6.3	5.6	3.2	1.8	1.1	0	0	0	0	0	0	0	0
e_g/i_c, ohms	1750	1850	2600	2960	2650	∞	∞	∞	∞	∞	∞	∞	∞

The average of $e_g i_c$ over a half cycle is (by the trapezoidal rule)

$$\frac{6.3 + 2 \times 5.6 + 2 \times 3.2 + 2 \times 1.8 + 2 \times 1.1 + 0}{12} = 2.5 \text{ watts}$$

Some of this power is used to heat the grid of the No. 1 tube; some of it is wasted in the bias supply. The following half cycle the No. 1 tube draws no grid power, but the No. 2 tube driver power averages 2.5 watts. Therefore the driver must furnish an average power of 2.5 watts over the complete cycle.

The driver of our output stage must be capable of delivering a peak power of 6.3 watts to a load which is quite variable over the cycle (Table 8-2). The average power over the cycle is 2.5 watts, but if the amplifier in this case is capable of delivering an average power of 6.3 watts, the performance is generally more satisfactory.

It is generally desirable to have the load on the driver several times its r_p value throughout the cycle, and in order to obtain this high impedance, the driver-to-grid transformer usually needs a step-down ratio. This must be carefully chosen since this ratio makes it more difficult to deliver the high secondary voltage required.

Because the nominal driver load line is not the actual operating curve, the latter depending upon the instantaneous incremental resistance of the output-stage grid, the distortion is considerably increased over that obtaining for the nominal load line. While operation with a higher than normal nominal load lessens this distortion, it is sometimes quite severe and measures must often be taken further to decrease it. These measures often include the use of negative feedback around the driver stage, operation in push-pull parallel, and resistance loading of the driver transformer secondaries. The actual design of the driver stage is beyond the scope of this book.

REFERENCES

1. Thompson, B. J.: Graphical Determination of Performance of Push-pull Amplifiers, *Proc. IRE*, **21**, 591 (1933).
2. Cruft Laboratory, War Training Staff: "Electronic Circuits and Tubes," McGraw-Hill Book Company, Inc., New York, 1947.
3. Reich, H. J.: "Theory and Applications of Electron Tubes," 2d ed., McGraw-Hill Book Company, Inc., New York, 1944.

PROBLEMS AND QUESTIONS

1. A triode-connected 6L6GA is used in a power-amplifier circuit of the type shown in Fig. 8-1. The input voltage is sinusoidal with a peak value of 30 volts. $E_{bo} = 340$ volts, $E_{co} = -30$ volts, $R_L = 3200$ ohms. Compute P_{po}, P_{bb}, P_{ac}, P_p, and η_p.

2. Repeat Prob. 1 for the circuit of Fig. 8-2*a*. Assume the reactance of the choke to be very large and its resistance negligible. Also assume the reactance of the coupling capacitor to be very small.

3. A class B amplifier uses $E_{bb} = 300$ volts, $E_{co} = -80$ volts. On the positive peak of the grid drive, $i_{b_{max}} = 50$ ma, $e_{b_{min}} = 25$ volts. Determine the approximate power delivered to the load and the plate efficiency.

4. A 6B4 triode is used in a power amplifier of the type shown in Fig. 8-5*a*. The transformer primary has negligible d-c resistance. If $E_{bo} = 200$ volts, $E_{co} = -30$ volts, and the grid signal is 10 volts peak, determine (*a*) the values of R_k and E_{bb} needed, (*b*) the value of tube load which will absorb the greatest a-c power, and the transformer turns ratio needed if the secondary load is 10 ohms, (*c*) the power delivered to the transformer primary.

5. A 6B4 is operated with $E_{bo} = 200$ volts. It is desired to make a graphical determination of the best operating conditions for class A_1 if heavy grid drive is planned. (*a*) Using expression (8-32), determine the approximate grid bias to be used (round off to nearest multiple of 5). Is this a safe operating point? (*b*) Determine the theoretical "best" load for this tube. (*c*) If the grid is driven to 0 volts, what is the a-c power output? (*d*) What is the percentage distortion developed?

6. A 6B4 triode is used in a power amplifier with transformer coupling to the load. $E_{bo} = 240$ volts. (*a*) If the numerically least safe grid bias is used, determine E_{co} and I_{bo}. (*b*) In one application of this amplifier a small grid drive of $E_g = 7.07$ volts is used. What ohmic value of load on the tube should be used in order to obtain greatest power output? Show the construction on a plate-characteristic sheet. (*c*) In another application it is desired to obtain the greatest power possible with class A_1 operation and a limit of 5 per cent on the distortion. Use a 9-11 rule, and draw the optimum load line. Determine the power output and the percentage of harmonic distortion. (*d*) Twice the power output with the same distortion can be obtained by using tubes in parallel. Determine E_g and the load R_{ac}.

7. A 6B4 triode is used in a power-amplifier circuit of the type shown in Fig. 8-5*a*. $E_{bb} = 250$ volts, $R_k = 400$ ohms; C_k is very large. The transformer-primary winding resistance is 162 ohms. The total resistance in the secondary circuit is 11 ohms. The transformer ratio $N_1/N_2 = 10$. The grid is driven to 0 volts. (*a*) Draw the bias line and the d-c-load line. Determine I_{bo}, E_{bo}, and E_{co}. (*b*) Determine the tube's a-c load, draw the a-c-load line, and determine I_{p_1}, I_{p_2}, I_{ba}, P_{ac}, and P_p.

8. A 6B4 triode is used in a circuit like Fig. 8-5*a*. $E_{bb} = 300$ volts, $R_k = 727$ ohms. C_k is very large. The transformer-primary winding resistance is 490 ohms; secondary-circuit resistance is 10 ohms. Transformer ratio N_1/N_2 is 14. Determine the power output, the plate dissipation, and the percentage of harmonic distortion.

9. A linear class A amplifier has $r_p = 1000$ ohms and $\mu = 10$. The load is an impedance $R_L + j1000$ ohms. Determine the value of R_L needed in order that maximum power may be delivered to the load.

10. Design a rule (similar to the 9-11 rule) that can be used to determine the proper load for 4 per cent harmonic distortion for a triode in class A_1.

11. (*a*) Use a tube manual to obtain a satisfactory design for a power amplifier using a 6B4 triode with $E_{bo} = 250$ volts. Determine P_{ac} and P_{po}. (*b*) Use the method of Art. 8-6 to determine a new design for use with $E_{bo} = 200$ volts. Determine E_{co}, R_L, R_k, P_{ac}, P_{po} for the new conditions.

12. A 6V6 beam power tube (connected as a pentode) is used in a power amplifier with transformer coupling to the load. If $E_{bo} = 250$ volts, $E_{c2o} = 250$ volts, $E_{c1o} = -10$ volts, and the grid is driven to 0 volts, use a sheet of plate characteristics and determine the approximate best load for high power output with minimum total distortion. Also determine P_{ac}, P_p, and the percentage of third-harmonic distortion.

13. Repeat Prob. 12 without the restriction that $E_{c1o} = -10$ volts. Choose E_{c1o} to obtain high power with minimum total distortion. Neglect any change in the T point.

14. (*a*) Consult a tube manual to obtain a satisfactory design for a 6V6 pentode power amplifier, with $E_{bo} = 250$ volts. Specify all important circuit values. (*b*) Use the method of Art. 8-6 and Eq. (8-54) to determine a new design for use with $E_{bo} = 275$ volts. Determine E_{c1o}, E_{c2o}, R_L, R_k, P_{ac}, P_{po} for the new conditions.

15. Prove that a lightly driven triode with a load resistance double the optimum value delivers 11 per cent below maximum power, with some decrease in distortion.

16. A type 7C5 tube is to be used in a class A_1 pentode power amplifier. The grid drive is sufficient to make $e_{c_{max}} = 0$. The Q point is $E_{bo} = 300$ volts, $E_{c2o} = 250$ volts, $E_{c1o} = -12.5$ volts. Determine the load line for zero second-harmonic distortion and the values of R_L, P_{ac}, P_{po}. Estimate $i_{1/2}$, $i_{-1/2}$.

17. Starting with one sheet of plate-characteristic curves for a 6B4 tube, draw the curves for a composite tube representing push-pull operation. $E_{bo} = 200$ volts, $E_{co} = -40$ volts.

18. Two 6B4 tubes are operated push-pull with fixed bias of -70 volts and $E_{bo} = 300$ volts. (*a*) Draw the composite characteristics. (*b*) Assume a composite-tube load of 1000 ohms, and draw the load line. (*c*) Plot the path of operation of tube 1. (*d*) If the grid swing is between -10 and -130 volts, what is the a-c power output? (*e*) What is the plate dissipation per tube? (*f*) Determine the plate efficiency. (*g*) Determine the percentage of third-harmonic distortion.

19. Two 6B4 tubes are operated push-pull with $E_{co} = -60$ volts, $E_{bo} = 300$ volts. Draw the composite-tube curves, and determine the approximate optimum load line. Determine P_{po}, P_p, P_{ac}, and the percentage of third-harmonic distortion.

20. Two 6V6 tubes are used in a push-pull amplifier with $E_{bo} = 250$ volts, fixed grid bias of -15 volts. Refer to a tube manual for other recommended circuit values. Use a sheet of plate characteristics, and check the performance against that stated in the manual.

21. Replace the source of fixed bias in Prob. 20 by cathode bias, using $R_k = 200$ ohms and a very large C_k value. Use the same plate load, and determine the operation (*a*) with drive to 0 volts, (*b*) with drive sufficient to limit the distortion to 5 per cent.

22. Two type 830-B tubes are used in a class B a-f power amplifier. $E_{ff} = 10$ volts, $E_{bo} = 1000$ volts. As a trial load for the composite tube use 1900 ohms. Draw the dynamic characteristic for the No. 1 tube, and determine the projected cut-off grid voltage. Use this as the grid bias. Use as the maximum drive point P that given by $e_b = 2e_c$. Determine values of $i_{b_{max}}$, $i_{1/2}$, $e_{b_{min}}$, $e_{b_{max}}$, $i_{c_{max}}$, and $e_{c_{max}}$ for the composite tube. Compute P_{ac}, P_{bb}, η_p, P_{1p}, and percentage of third harmonic. Estimate the grid dissipation to determine the driver power required.

CHAPTER 9

POWER AMPLIFIERS USING TUNED LOADS

9-1. Operation for High Efficiency. In most cases the power supplied a radio-transmitter antenna or used in an industrial r-f heating process is obtained as the output of a power amplifier. Usually only a narrow band of frequencies is involved and a tuned load can be used as in the case of an r-f voltage amplifier. This makes it possible to use the otherwise troublesome tube and circuit capacitances to help tune the load. In addition, the plate-current waveform need not be sinusoidal since the

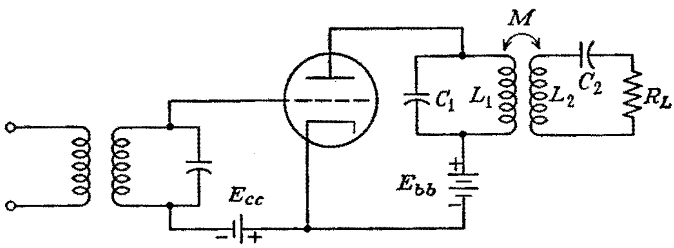


FIG. 9-1. Basic circuit for a tuned power amplifier. (Neutralization not shown.)

tuned circuit acts as a filter to attenuate the harmonics in the power output. This latter fact is important since it allows operation to be carried out in a more efficient manner than class A.

One simplified circuit for class B or class C operation is shown in Fig. 9-1. The load R_L in this case is inductively coupled to the coil of the tank circuit, and the capacitor C_2 may be adjusted to make the secondary circuit resonant at the frequency of operation. Hence, in effect, at the resonant frequency the plate load for the tube is a capacitor C_1 in parallel with a coil L_1 having a lower effective Q than its unloaded value.

For class B operation the value of E_{cc} is approximately that necessary to cut off the plate current. Hence, when a grid signal is applied, the plate current flows during only the positive half of the grid swing and it has a waveshape approximating a half sinusoid. The grid bias for a class C amplifier exceeds that for class B, usually being 1.5 to 3 times the cutoff value. Accordingly, the required grid-signal voltage must be high in order to drive the tube into conduction. With limitations, the plate-circuit efficiency increases as the grid drive is made larger, and hence both class B and class C amplifiers are nearly always class 2, drawing a considerable amount of grid current. For a class C amplifier the plate current flows in pulses, each pulse lasting less than a half cycle. Figure 9-2 shows approximate waveforms of e_b, e_c, i_b, and i_c for a class B amplifier, while Fig. 9-3 shows those for the class C case. In both of these it is sig-

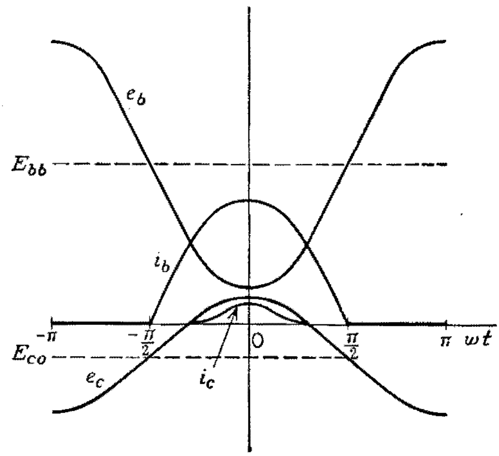


FIG. 9-2. Approximate waveforms for class B operation.

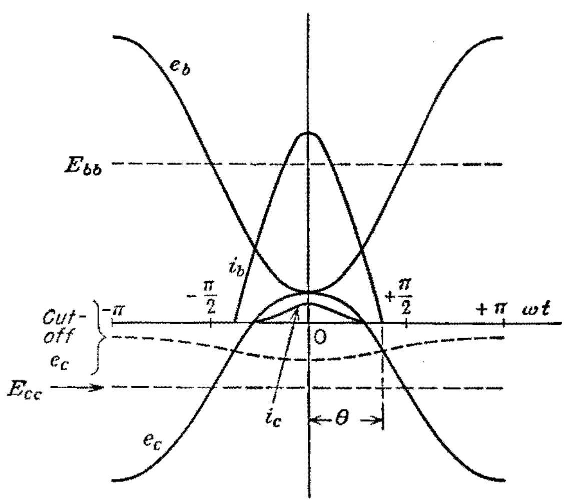


FIG. 9-3. Approximate waveforms for class C operation.

nificant that the plate current flows only during a period of low plate voltage and therefore the plate dissipation is small compared with that for class A operation. This is the reason for the high plate efficiencies for class B and class C operation.

9-2. The Plate-load Circuit. The plate load for a tuned power amplifier is usually not as simple as those employed with tuned r-f voltage amplifiers. In general the plate load for an r-f power amplifier must serve several purposes. First, it should present to the tube a proper impedance to render high power output with good efficiency at the signal frequency. (In some cases a harmonic of this frequency is chosen instead.) Second, the load should have a low impedance at harmonic frequencies to give good filtering action. Third, it must have a wide enough pass band to accommodate any modulation-frequency components. Fourth, the coupling network between the tank and the device which usefully consumes energy must often further filter the output of the tube to remove vestiges of harmonic components. Fifth, the efficiency of the coupling network, between its input and its output, must be high.

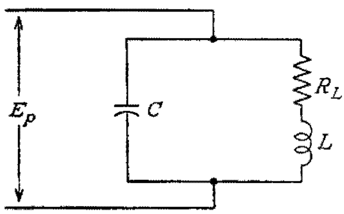


FIG. 9-4. A simple type of load.

The simplest type of load is that shown in Fig. 9-4, where the energy-consuming R_L is directly in series with the inductor L. In this case the impedance of the parallel-tuned circuit can be obtained from Eqs. (7-150) and (7-148) as

$$\mathbf{Z} = \frac{Q_L\omega_0 L}{1 + jQ_L[(2+\delta)/(1+\delta)]\delta} \tag{9-1}$$

for $Q_L \gg 1$.

The plate-voltage waveforms in Figs. 9-2 and 9-3 are both very nearly sinusoidal because of the filtering action of the plate-tank circuit. The plate current in either case is nonsinusoidal but periodic and hence may be represented by a Fourier series consisting of average current plus a fundamental-frequency component and harmonics. For example, in the case of the class B amplifier the current waveform approximates a half sinusoid of peak value I_{bm}, for which the Fourier representation is (see Prob. 14, Chap. 6)

$$i_b = \frac{I_{bm}}{\pi} + \frac{I_{bm}}{2}\cos\omega t + \frac{2I_{bm}}{3\pi}\cos 2\omega t - \frac{2I_{bm}}{15\pi}\cos 4\omega t + \cdots \tag{9-2}$$

For the fundamental frequency, $\delta = 0$, and

$$\mathbf{Z}_{f_0} = R_{f_0} = Q_L\omega_0 L \tag{9-3}$$

For the second harmonic, $\delta = 1$,

$$\mathbf{Z}_{2f_0} = \frac{Q_L\omega_0 L}{1 + j1.5Q_L} \tag{9-4}$$

and

$$R_{2f_0} = \frac{Q_L\omega_0 L}{1 + 2.25Q_L^2} \tag{9-5}$$

where R_{f_0} and R_{2f_0} are the resistance components of the impedances $\mathbf{Z}_{f_0}$ and $\mathbf{Z}_{2f_0}$. The power delivered to the tank circuit may be computed as

I^2R. For the fundamental frequency

$$P_{f_0} = I_{p_1}{}^2 R_{f_0} = \left(\frac{I_{bm}}{2\sqrt{2}}\right)^2 Q_L \omega_0 L \tag{9-6}$$

For the second harmonic the power is

$$P_{2f_0} = I_{p_2}{}^2 R_{2f_0} = \left(\frac{2I_{bm}}{3\pi\sqrt{2}}\right)^2 \frac{Q_L \omega_0 L}{1 + 2.25 Q_L{}^2} \tag{9-7}$$

The ratio of these powers is

$$\frac{P_{f_0}}{P_{2f_0}} = (1 + 2.25 Q_L{}^2) \frac{9\pi^2}{16} \tag{9-8}$$

Should Q_L be 10 or more, this ratio would equal or exceed $(1 + 2.25 \times 10^2)5.6 = 1266$. In the case of the class C amplifier the second harmonic

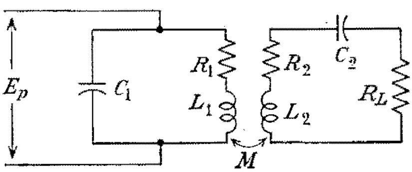


FIG. 9-5. Elements of a tank circuit.

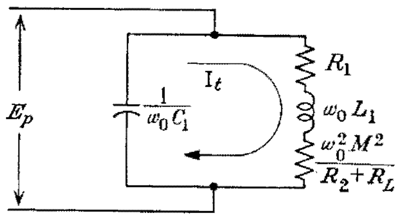


FIG. 9-6. The equivalent circuit for Fig. 9-5.

is relatively larger than for class B, but even so, the second-harmonic power developed in the load will be small if the value of Q_L as great as 10 is employed.

If the r-f energy from the plate tank is to be radiated, further filtering is usually necessary to meet FCC requirements. The filter, which serves also as a coupler between tank and load, becomes a part of the plate load. In this or any other ordinary case the simple tank representation of Fig. 9-4 seldom exactly applies, and the frequency characteristics of the plate load are usually no longer expressible by the simple equations just derived. The circuit of Fig. 9-5 shows one possible load configuration, but it too is somewhat too simple for representing very accurately the effect of a complex coupling circuit. Although the process of analysis in this case is similar to that for the double-tuned amplifier of Art. 7-37, the results are quite different because of the unsymmetrical circuit and the low value of Q for the secondary circuit. Upon analysis, it may be shown that the frequency-response curve somewhat resembles that of the simple parallel-tuned load considered earlier in this article, but the tuning is somewhat broader.

While a high value of Q for the tank circuit is desirable for suppressing harmonics, its value must be limited if power is to be delivered to a load.

To investigate the desirable conditions, we shall assume the load configuration to be that of Fig. 9-5. Assume L_2 and C_2 have equal reactances and that the whole tank is tuned to resonance. Then at the resonant frequency the simplified equivalent circuit of Fig. 9-6 can be used. The power delivered to the tank is absorbed by R_1 and $\omega_0{}^2M^2/(R_2 + R_L)$. The impedance represented by the latter term alone absorbs the power delivered to the secondary, of which only a part is load power. The ratio of load power P_L delivered to R_L to the total power P_t delivered to the tank by the tube is therefore

$$\frac{P_L}{P_t} = \frac{I_t{}^2[\omega_0{}^2M^2/(R_2 + R_L)][R_L/(R_2 + R_L)]}{I_t{}^2[R_1 + \omega_0{}^2M^2/(R_2 + R_L)]} \tag{9-9}$$

This equation can be rewritten as

$$\frac{P_L}{P_t} = \frac{\left(R_1 + \dfrac{\omega_0{}^2M^2}{R_2 + R_L} - R_1\right)\dfrac{R_L}{R_2 + R_L}}{R_1 + \dfrac{\omega_0{}^2M^2}{R_2 + R_L}}$$

$$= \left(1 - \frac{R_1}{R_1 + \dfrac{\omega_0{}^2M^2}{R_2 + R_L}}\right)\frac{R_L}{R_2 + R_L} \tag{9-10}$$

If we let Q_L represent the loaded Q of the tank,

$$Q_L = \frac{\omega_0 L_1}{R_1 + [\omega_0{}^2M^2/(R_2 + R_L)]} \tag{9-11}$$

and Q_p the Q of the primary alone,

$$Q_p = \frac{\omega_0 L_1}{R_1} \tag{9-12}$$

then

$$\frac{Q_L}{Q_p} = \frac{R_1}{R_1 + [\omega_0{}^2M^2/(R_2 + R_L)]} \tag{9-13}$$

and

$$\frac{P_L}{P_t} = \left(1 - \frac{Q_L}{Q_p}\right)\frac{R_L}{R_2 + R_L} \tag{9-14}$$

Thus we see that, if Q_p can be made very large compared with Q_L and if R_2 can be made small compared with R_L, the efficiency of transfer of power from tank to load can be made high. Since the value of Q_p can be made 100 or more for small coils and considerably greater than this for large transmitting coils, a value of Q_L near 12 would satisfy this requirement for high transfer efficiency.

Class C amplifiers may be used for producing a modulated wave, in which case the bandwidth for the tank must be great enough to handle the sideband frequencies. A very high value of Q_L must then be avoided.

If the plate-tank circuit is tuned to a harmonic of the grid-signal frequency, a class C amplifier may be used as a frequency multiplier. The waveforms for grid voltage, plate current, and plate voltage in the case of a frequency doubler are shown in Fig. 9-7.

It will be noticed that the plate-voltage wave loses amplitude between plate-current pulses because of loss in energy in the tank circuit. The same tendency exists in an ordinary class C amplifier. To show how the

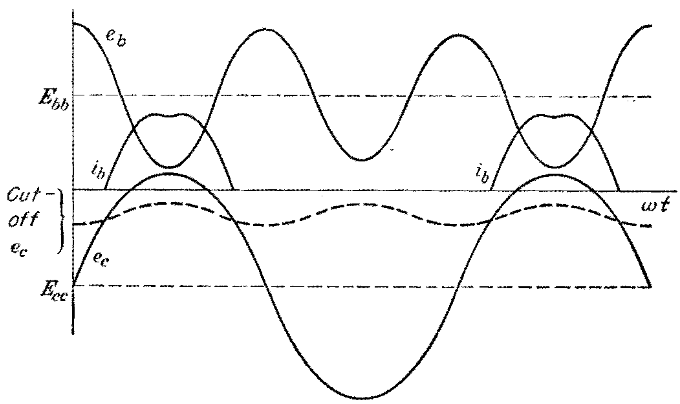


FIG. 9-7. Waveforms for a frequency doubler.

voltage amplitude falls for a class C amplifier, we may proceed as follows: At the beginning of the cycle when the value of E_{p_1} is a maximum, the energy stored in the tank is all in the capacitor and

$$W = \tfrac{1}{2}CE_{p_1m}^2 \tag{9-15}$$

If the drop in voltage is not great, then the approximate change in energy in one cycle is (power multiplied by time)

$$\Delta W = -\frac{E_{p_1}^2}{Q_L\omega_0 L}\frac{1}{f_0} = -\frac{1}{2}E_{p_1m}^2\frac{2\pi C}{Q_L} \tag{9-16}$$

The energy in the capacitor at the end of one cycle is (unless it is replenished)

$$W + \Delta W = \frac{1}{2}CE_{p_1m}^2 - \frac{1}{2}E_{p_1m}^2\frac{2\pi C}{Q_L} = \frac{1}{2}CE_{p_1m}^2\left(1 - \frac{2\pi}{Q_L}\right) \tag{9-17}$$

or, in terms of the new peak voltage E'_{p1m},

$$W + \Delta W = \tfrac{1}{2}CE'^2_{p1m} \tag{9-18}$$

Therefore,

$$\frac{1}{2}CE'^2_{p_1m} = \frac{1}{2}CE_{p_1m}{}^2\left(1 - \frac{2\pi}{Q_L}\right) \tag{9-19}$$

and

$$\frac{E'_{p_1m}}{E_{p_1m}} = \sqrt{1 - \frac{2\pi}{Q_L}} \tag{9-20}$$

The ratio of total stored energy to energy dissipated per cycle is

$$\frac{W}{\Delta W} = \frac{\frac{1}{2}CE_{p_1m}{}^2}{\frac{1}{2}E_{p_1m}{}^2(2\pi C/Q_L)} = \frac{Q_L}{2\pi} \tag{9-21}$$

Hence, to keep a good waveform, a high value of Q_L is desirable. The effect of the tank circuit in maintaining a good sinusoidal waveform depends upon keeping $Q_L/2\pi$ high, or in having a high ratio of total stored energy to energy dissipated per cycle. This is often called the flywheel effect.

9-3. An Algebraic Analysis of Class B Amplifier Operation. By means of a graphical analysis it is possible to determine closely the performance of any given class B or class C amplifier circuit. However, this supposes that all values of supply voltage, load, and drive are already chosen, in other words, that the circuit has already been designed. While such an analysis is important and will be treated in a later article, our concern at present is with the methods by which the correct load, operating voltages, and drive may be determined in order to obtain the usual objective: large output power with satisfactory efficiency and safe plate and grid dissipations. Such a problem is difficult for the class C case, and it is usually handled by a cut-and-try process. However, in the case of class B an approximate solution is not particularly difficult to obtain, and the results may serve as convenient guides in design although they do lack a high degree of accuracy.

To begin the algebraic analysis of a class B amplifier, it is necessary to make some assumptions, some of which are approximations. First, we shall consider the plate-characteristic curves for the tube to be straight lines, equally spaced. This amounts to stating that the exponent n in Eq. (4-3) is unity, so that

$$i_b = k(e_b + \mu e_c) \tag{9-22}$$

where μ and k are constants. If we differentiate in respect to e_c, $\partial i_b/\partial e_c = g_m = k\mu$ or $k = g_m/\mu = 1/r_p$. Therefore

$$i_b = \frac{1}{r_p}(e_b + \mu e_c) \tag{9-23}$$

provided, of course, that $i_b > 0$.

Second, we shall consider the tube to be biased to cutoff, so that when $e_b = E_{bb}$ and $e_c = E_{cc}$, then $i_b = 0$. Then

$$0 = \frac{1}{r_p}(E_{bb} + \mu E_{cc}) \tag{9-24}$$

and the cutoff bias is given by

$$E_{cc} = -\frac{E_{bb}}{\mu} \tag{9-25}$$

Third, we assume the tank is tuned to resonance for the fundamental frequency. Let its impedance for this frequency be $R_{f_0} = Q_L\omega_0 L$, where Q_L is the loaded value of Q. Furthermore, let us assume the impedance to the harmonics is zero. Since it is desirable to have i_b a maximum at the moment that e_b is a minimum, which calls for a resistance load, this resonant condition is actually the desired one.

Fourth, we shall assume the grid signal to be sinusoidal. The plate voltage will also be sinusoidal (because of the action of the tank) and 180° out of phase with the grid voltage.

$$e_c = E_{cc} + E_{gm}\cos\omega_0 t \tag{9-26}$$
$$e_b = E_{bb} - E_{pm}\cos\omega_0 t \tag{9-27}$$

We may now write i_b as a function of time, using Eqs. (9-23), (9-26), and (9-27),

$$i_b = \frac{1}{r_p}(E_{bb} - E_{pm}\cos\omega_0 t + \mu E_{cc} + \mu E_{gm}\cos\omega_0 t) \tag{9-28}$$

which reduces [using Eq. (9-24)] to

$$i_b = \frac{1}{r_p}(\mu E_{gm} - E_{pm})\cos\omega_0 t = I_{bm}\cos\omega_0 t \tag{9-29}$$

for values of $\omega_0 t$ in the fourth and first quadrants only, when the grid swing is positive. On the other hand

$$i_b = 0 \tag{9-30}$$

for values of $\omega_0 t$ in the second and third quadrants (when the grid swing is negative). Thus the waveshape of i_b is a half sinusoid, for which we may write the Fourier series

$$i_b = I_{ba} + I_{p_1m}\cos\omega_0 t + I_{p_2m}\cos 2\omega_0 t + \cdots \tag{9-31}$$

The coefficients in this series can be determined as follows,

$$I_{ba} = \frac{1}{2\pi}\int_{-\pi/2}^{3\pi/2} i_b\,d\omega_0 t = \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} I_{bm}\cos\omega_0 t\,d\omega_0 t = \frac{I_{bm}}{\pi} \tag{9-32}$$

and

$$I_{p_1m} = \frac{2}{2\pi}\int_{-\pi/2}^{3\pi/2} i_b \cos \omega_0 t \, d\omega_0 t = \frac{I_{bm}}{\pi}\int_{-\pi/2}^{\pi/2} \cos^2 \omega_0 t \, d\omega_0 t = \frac{I_{bm}}{2} \quad (9\text{-}33)$$

Hence, from Eqs. (9-29) and (9-33),

$$I_{p_1m} = \frac{\mu E_{gm} - E_{pm}}{2r_p} \quad (9\text{-}34)$$

or since

$$E_{pm} = I_{p_1m} R_{f_0} \quad (9\text{-}35)$$

(not a vector relationship and hence no minus sign is used)

$$I_{p_1m} = \frac{\mu E_{gm}}{2r_p + R_{f_0}} \quad (9\text{-}36)$$

and

$$\mathbf{I}_{p_1} = \frac{\mu \mathbf{E}_g}{2r_p + R_{f_0}} \quad (9\text{-}37)$$

since $\mathbf{I}_{p_1}$ is in phase with $\mathbf{E}_g$ by Eqs. (9-26) and (9-31).

As far as the fundamental-frequency quantities are concerned, we see that Eq. (9-37) may lead to an equivalent circuit, such as that shown in Fig. 9-8. We may write the voltage across the tank as

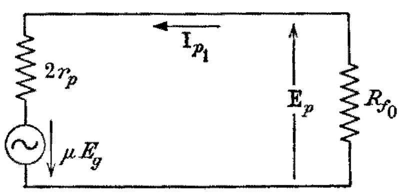


FIG. 9-8. An equivalent circuit useful for fundamental-frequency components.

$$\mathbf{E}_p = -\mathbf{I}_p R_{f_0} = -\frac{\mu \mathbf{E}_g R_{f_0}}{2r_p + R_{f_0}} \quad (9\text{-}38)$$

This equation shows that, in so far as the conditions of operation agree with our assumptions, the output voltage E_p is proportional to the signal voltage E_g. This makes the amplifier a class B linear one, useful for amplifying modulated signals, since their waveforms are preserved. In practice it is possible to adjust a class B amplifier so that this linear relationship holds very closely.

The power delivered to the tank is

$$P_{ac} = E_p I_{p_1} = \frac{\mu^2 E_g{}^2 R_{f_0}}{(2r_p + R_{f_0})^2} \quad (9\text{-}39)$$

The power furnished by the plate supply is [from Eqs. (9-32), (9-33), and (9-36)]

$$P_{bb} = E_{bb} I_{ba} = E_{bb}\frac{I_{bm}}{\pi} = E_{bb}\frac{2I_{p_1m}}{\pi} = \frac{2}{\pi} E_{bb} \frac{\mu E_{gm}}{2r_p + R_{f_0}} \quad (9\text{-}40)$$

and that dissipated by the tube (assume no d-c losses in the tank) is

$$P_p = P_{bb} - P_{ac} = \frac{2}{\pi} E_{bb} \frac{\mu E_{gm}}{2r_p + R_{f_0}} - \frac{\mu^2 E_{gm}{}^2 R_{f_0}}{2(2r_p + R_{f_0})^2} \quad (9\text{-}41)$$

The plate-circuit efficiency is

$$\eta_p = \frac{P_{ac}}{P_{bb}} = \frac{E_p I_{p_1}}{E_{bb} I_{ba}} = \frac{E_{pm}}{\sqrt{2}\, E_{bb}} \frac{I_{bm}/2\sqrt{2}}{I_{bm}/\pi} = \frac{\pi}{4} \frac{E_{pm}}{E_{bb}} = 78.5 \frac{E_{pm}}{E_{bb}} \quad \% \tag{9-42}$$

Equation (9-39) shows that the power output increases as the square of E_g, making it important to use a large value of grid drive. However, for drive heavy enough to make $e_{c_{\max}}$ exceed $e_{b_{\min}}$ in value, the relationship of (9-39) no longer closely holds. This is because, with the potential of the grid higher than that of the plate, the grid draws a large current of both primary electrons from the cathode and secondary electrons from the plate, with a dip in the plate current resulting at times when the plate voltage is a minimum. Peaks of plate current now occur at positions where the plate voltage is no longer very low, with the result that the plate dissipation increases rapidly, while that of the grid also rises together with the power required of the grid driver. We see therefore that grid drive should not exceed some value near that which makes $e_{c_{\max}} = e_{b_{\min}}$.

Having in mind the limitation in grid drive just discussed above, let us continue our development under the condition

$$\frac{e_{b_{\min}}}{e_{c_{\max}}} = p \tag{9-43}$$

where p is a constant (usually ranging between 1 and 3 in value). From this we may write

$$E_{bb} - E_{pm} = p(E_{cc} + E_{gm}) \tag{9-44}$$

and using Eqs. (9-25), (9-35), and (9-36), we obtain

$$E_{bb} - \frac{\mu E_{gm} R_{f_0}}{2r_p + R_{f_0}} = p\left(-\frac{E_{bb}}{\mu} + E_{gm}\right) \tag{9-45}$$

$$E_{gm} = \frac{E_{bb}[1 + (p/\mu)]}{p + [\mu R_{f_0}/(2r_p + R_{f_0})]} \tag{9-46}$$

This equation shows that a large value of grid drive E_{gm} (required for high power output) can be used if E_{bb} is made large.

E_{bb} should be near the maximum allowable direct plate voltage for the tube, as stated by the manufacturer. This matter of E_{bb} now settled, let us notice [by Eq. (9-46)] that E_g depends also upon the value of R_{f_0} used. Since R_{f_0} also determines the plate dissipation according to Eq. (9-41), we may eliminate E_{gm} [between Eqs. (9-41) and (9-46)] to give R_{f_0} in terms of E_{bb} and P_p, the allowable plate dissipation. If this is done, we obtain

$$R_{f_0} = \frac{E_{bb}^2(4 - \pi)}{4\pi P_p} - \frac{2pr_p}{p + \mu} + \sqrt{\left[\frac{E_{bb}^2(4 - \pi)}{4\pi P_p}\right]^2 + \frac{pr_p E_{bb}^2}{(p + \mu)P_p}} \tag{9-47}$$

The design procedure is to first use Eq. (9-47) to determine R_{f_0} from known values of allowable plate dissipation and E_{bb} and an assumed value of p. This value of R_{f_0} may be substituted into Eq. (9-46) to obtain E_{gm}, and then P_{ac} may be determined from Eq. (9-39). The plate efficiency may be checked by the use of Eq. (9-42). Making $p = 1$ will yield the highest but most unreliable answers for P_{ac} and η_p. For $p > 1$ the results are more conservative and more reliable. After the preliminary design is made by this approximate analytical method, it may be well to follow a graphical analysis to obtain further information about the amplifier.

9-4. Graphical Analysis of Class B and Class C Amplifiers. In principle, the graphical analysis of these tuned amplifiers is the same as that of the untuned resistance-loaded amplifiers of Chap. 8. On a sheet of tube characteristics an operating point is chosen, the path of operation drawn, the limits of grid swing indicated. Then values of current and/or voltage are read at convenient points on the path of operation, and these values are substituted into suitable formulas for approximating the quantities desired.

For the untuned amplifier the plate current and the plate voltage have identical waveforms as long as the load is constant for all frequency components. Thus $i_b = I_{bt} + i_p$ and $e_b = E_{bt} - R_L i_p$. Upon eliminating i_p between these equations we obtain a linear equation in i_b and e_b, which on i_b, e_b orthogonal axes plots into a straight load line.

Now for an amplifier with a tuned load and of the classes studied in this chapter, the grid voltage is sinusoidal, the plate voltage is very closely sinusoidal, while the plate current is far from being so. Hence on i_b, e_b axes the path of operation is not a straight line. However, since $e_c = E_{ct} + E_{gm} \cos \omega_0 t$ and $e_b = E_{bt} - E_{pm} \cos \omega_0 t$, elimination of $\omega_0 t$ yields a linear equation in e_c and e_b, proving that the path of operation is a straight line on e_c, e_b orthogonal axes. Hence we need tube characteristics drawn on these axes. Since i_b is the parameter used in drawing these curves, they are often called constant-current characteristics.

If the operating conditions E_{bo}, E_{co}, E_{gm}, and E_{pm} are known, there should be no difficulty in constructing the path of operation. We can then locate, on the path of operation, points which are spaced at equal time intervals. If desired, the values of plate current at these points may be used to draw a curve of time variation of plate current, although usually this step is unnecessary. Application of formulas to be stated later can then be made and values of I_{ba} and I_{p_1} computed. The curve sheet for constant-current characteristics usually has curves of constant grid current superimposed, and hence computations of I_{ca} and I_{g_1} can also be made.

For the class C case, if the operating conditions are not known, a cut-and-try process may be used to determine satisfactory values. In calculating the performance the considerations which determine these operating values are the power output required, the plate efficiency, the maximum allowable plate and grid dissipations, and the maximum allowable plate voltage and current. Also, the grid driving power is of considerable importance. The choice of values depends on both the tube employed and the service in which it is placed.

The power output depends upon the needs of the load and the efficiencies of the tank and the coupling network. High plate-circuit efficiency is desirable because this makes the plate dissipation and the d-c power less. However, high plate-circuit efficiency usually demands high plate-voltage operation and calls for considerable grid-driving power, and a lower efficiency may be more practical.

Since a blind cut-and-try process is extremely tedious, it is well to develop some guides to limit the necessary labor. The value of E_{bb} may be chosen at any value up to the safe value of E_{bo} stated in the tube manual. For high plate-circuit efficiency it is necessary to limit the plate-current flow to periods of low plate voltage, which means that a sharp, high pulse of plate current is preferred. Thus it is important to bias the grid well beyond plate-current cutoff, say 1.5 or more times this value. To obtain much a-c power output the short pulse of plate current must be very high, requiring grid drive to high positive values. This action requires considerable grid-circuit power, both to supply the grid losses and to charge the E_{cc} battery (the grid current enters the E_{cc} battery at its positive terminal).

The value of the plate load $R_{f_0}(= E_{p_1}/I_{p_1})$ must be a result of the calculations (because of I_{p_1}) and cannot be considered a known parameter in beginning a problem. This is unfortunate since it occasionally happens that a certain load value is desired. In the latter event it is necessary to make repeated trials until the correct value is obtained.

Many studies have been made to develop additional guides for choosing tentative operating conditions in order to save labor. The following method is not difficult to understand and should be adequate for triodes. Tetrodes and pentodes are handled somewhat similarly.

As a preliminary to the graphical solution, it is desired to formulate a relation among several quantities involved in the graphical solution. Figure 9-3 shows waveforms for class C operation. Let us assume that the waveform of the plate current is a portion of a cosine wave and that i_b cuts off at an angle θ, where θ does not exceed 90°. Hence we may show that the equation for the plate current may be written as (see Prob. 4 at end of chapter)

$$i_b = i_{b_{\max}}\left[\frac{1}{1 - \cos\theta}(\cos\omega t - \cos\theta)\right] \qquad \omega t \leq \theta \qquad (9\text{-}48)$$

and the cutoff angle θ is given by

$$\theta = \arccos \frac{E_{cc} + (E_{bb}/\mu)}{E_{cc} + (E_{pm}/\mu) - e_{c_{\max}}} \tag{9-49}$$

Then

$$\begin{aligned} I_{ba} &= \frac{1}{\pi} \int_0^\theta i_b \, d\omega t = \frac{1}{\pi} \frac{i_{b_{\max}}}{1 - \cos\theta} \int_0^\theta (\cos\omega t - \cos\theta) \, d\omega t \\ &= \frac{i_{b_{\max}}}{\pi(1 - \cos\theta)} (\sin\theta - \theta\cos\theta) \end{aligned} \tag{9-50}$$

or

$$\frac{i_{b_{\max}}}{I_{ba}} = \frac{\pi(1 - \cos\theta)}{\sin\theta - \theta\cos\theta} \tag{9-51}$$

$$\begin{aligned} I_{p_1m} &= \frac{2}{\pi} \int_0^\theta i_b \cos\omega t \, d\omega t = \frac{2}{\pi} \frac{i_{b_{\max}}}{1 - \cos\theta} \int_0^\theta (\cos^2\omega t - \cos\theta\cos\omega t) \, d\omega t \\ &= \frac{1}{\pi} \frac{i_{b_{\max}}}{1 - \cos\theta} (\theta - \sin\theta\cos\theta) \end{aligned} \tag{9-52}$$

The plate efficiency can be expressed as

$$\eta_p = \frac{P_{ac}}{P_{bb}} = \frac{I_{p_1}E_p}{E_{bb}I_{ba}} = \frac{I_{p_1m}E_{pm}}{2E_{bb}I_{ba}} \tag{9-53}$$

and hence from (9-53), (9-52), and (9-50)

$$\frac{E_{pm}}{\eta_p E_{bb}} = \frac{2I_{ba}}{I_{p_1m}} = \frac{2(\sin\theta - \theta\cos\theta)}{\theta - \sin\theta\cos\theta} \tag{9-54}$$

In Eq. (9-51) the ratio of $i_{b_{\max}}$ to I_{ba} is expressed in terms of the parameter θ. In Eq. (9-54) the ratio of E_{pm} to $\eta_p E_{bb}$ is expressed in terms of the same parameter. Figure 9-9 shows graphically the relation between $E_{pm}/\eta_p E_{bb}$ and $i_{b_{\max}}/I_{ba}$. Values of θ used in obtaining this graph are also indicated.

The designer can usually estimate the values of I_{ba} and the required plate efficiency from knowledge of P_{ac} desired, E_{bb} and the tube type available. Thus for any given point on the curve of Fig. 9-9, E_{pm} and $i_{b_{\max}}$ can be computed and a P point located on the constant-current characteristics of the tube. If the required efficiency is low, it is usually best to choose a point toward the left end of the curve of Fig. 9-9, since here $i_{b_{\max}}$ can be kept low without making E_{pm} excessively high. Thus plate-current distortion and grid current are kept to lower values. If the required efficiency is high, a point toward the right end is better since only in this way can E_{pm} be kept down in value. A P point which makes $p(= e_{b_{\min}}/e_{c_{\max}})$ between 1 and 3 is usually satisfactory.

A guide for grid bias can be obtained from Eq. (9-49).

$$E_{cc} = -\frac{(E_{bb}/\mu) + [e_{c_{\max}} - (E_{pm}/\mu)] \cos \theta}{1 - \cos \theta} \tag{9-55}$$

A straight load line can now be drawn through P and Q. Only the upper end of this path of operation need be drawn since for these classes of operation the plate and grid currents are both zero during the negative grid swing.

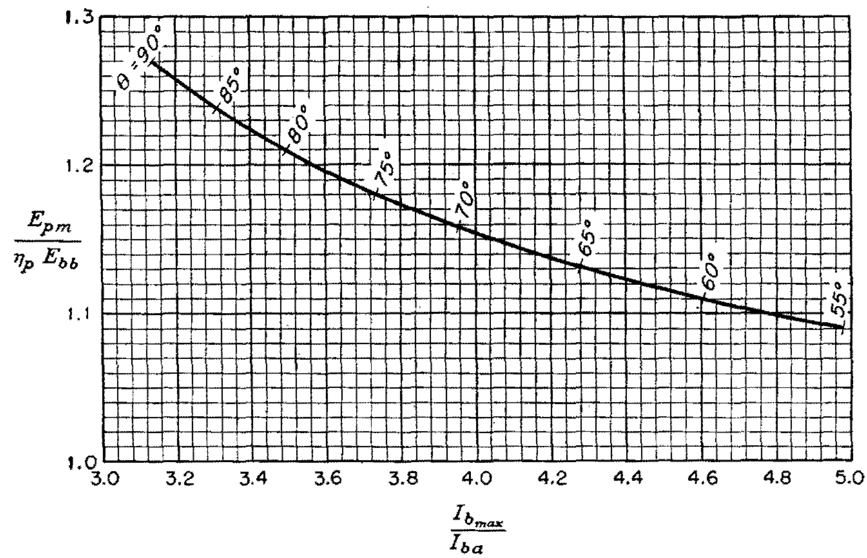


FIG. 9-9.

9-5. An Example of the Graphical Analysis of a Class C Amplifier. Let us assume that a class C amplifier is to deliver 400 watts of r-f power to its plate tank circuit. An Eimac type 100TH tube is available for which the maximum ratings are $E_{bo} = 3000$ volts, $I_{ba} = 225$ ma, $P_p = 100$ watts, $P_g = 20$ watts. The average value of μ is 38, obtained as the reciprocal of the slope of the constant-current characteristics.

For our case let us use $E_{bb} = 3000$ volts. If we assume P_p to be the maximum allowable, then $P_{bb} = 400 + 100 = 500$ watts and $\eta_p = 400/500 = 80$ per cent. $I_{ba} = P_{bb}/E_{bb} = 500/3000 = 167$ ma, which is less than 225 ma. Eighty per cent efficiency is fairly high. Hence we can use a value of θ near 64° where $i_{b_{\max}}/I_{ba} = 4.35$ and $E_{pm}/\eta_p E_{bb} = 1.125$. Then $i_{b_{\max}} = 4.35 \times 167 = 727$ ma, $E_{pm} = 0.80 \times 3000 \times 1.125 = 2700$ volts, and $e_{b_{\min}} = 3000 - 2700 = 300$ volts. If we locate this P point on the characteristics (Fig. 9-10), we find $e_{c_{\max}} = 185$ volts. $p =$

300/185 = 1.62, which is reasonable. We can use Eq. (9-55) to determine the bias value. Upon substitution we obtain $E_{cc} = -230$ volts.

On the tube characteristics of Fig. 9-10 we can now locate the Q point and draw the load line. This permits a check on the operation and allows

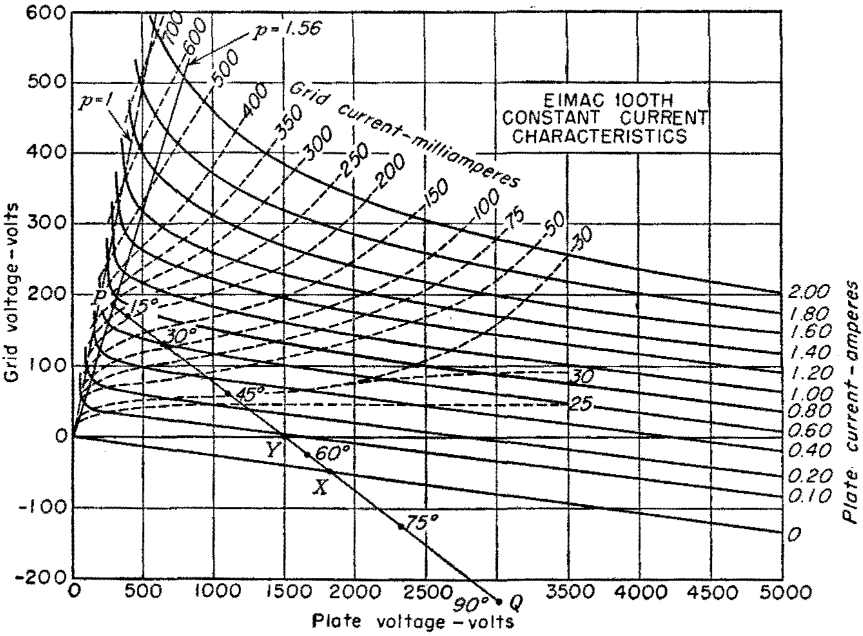


FIG. 9-10.

us to determine in addition the grid power, etc. We may now locate sampling points at equal time divisions along the path of operation, using increments of ωt of, say, 15°. More and shorter intervals will increase the accuracy of the results, and 10° intervals are often used instead. Table 9-1 may now be completed. An explanation of the items in it follows the table.

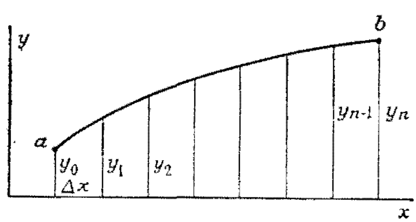


FIG. 9-11. Divide the area into trapezoids.

We assume $e_c = E_{cc} + E_{gm} \cos \omega t = -230 + 415 \cos \omega t$ and use computed values of e_c to locate the sampling points. The values of i_b and of i_c are now read at these points.

By formula (9-32), I_{ba} is the time average of i_b over one cycle. Similarly, I_{ca} is the time average of i_c over one cycle. Formula (9-33) shows that I_{p_1m} is *twice* the time average of $i_b \cos \omega t$ over one cycle, and similarly for I_{g_1m}. A convenient formula for approximating averages is pro-

TABLE 9-1

ωt	cos ωt	e_c, volts	i_b, ma	Computations for I_{ba}	i_c, ma	Computations for I_{ca}	$i_b \cos \omega t$, ma	Computations for I_{p_1m}	$i_c \cos \omega t$, ma	Computations for I_{g_1m}
0°	1.000	185	727	363	295	147	727	363	295	147
15°	0.966	172	715	715	210	210	692	692	203	203
30°	0.866	130	600	600	120	120	520	520	104	104
45°	0.707	64	300	300	32	32	212	212	23	23
60°	0.500	−23	40	40	0	0	20	20	0	0
75°	0.259	−123	0	0	0	0	0	0	0	0
90°	0	−230	0	0	0	0	0	0	0	0
		Sum		2018		509		1807		477
		1/12 × sum		168		42				
		1/6 × sum						301		80

vided by the so-called trapezoidal rule, used by most students when studying the calculus. This rule may be readily deduced from Fig. 9-11. The area under the curve from a to b is divided into n segments, each of equal width Δx. An approximation to the area under the curve is obtained by adding the areas of the trapezoids shown.

$$\begin{aligned}\text{Area} &= \frac{y_0 + y_1}{2}\,\Delta x + \frac{y_1 + y_2}{2}\,\Delta x + \frac{y_2 + y_3}{2}\,\Delta x + \frac{y_3 + y_4}{2}\,\Delta x + \cdots \\ &\qquad + \frac{y_{n-1} + y_n}{2}\,\Delta x \\ &= \Delta x\left(\frac{y_0}{2} + y_1 + y_2 + y_3 + \cdots + y_{n-1} + \frac{y_n}{2}\right) \qquad (9\text{-}56)\end{aligned}$$

This yields the average value if we divide by $n\,\Delta x$,

$$\text{Average } y = \frac{1}{n}\left(\frac{y_0}{2} + y_1 + y_2 + y_3 + \cdots + y_{n-1} + \frac{y_n}{2}\right) \qquad (9\text{-}57)$$

where n is the number of intervals, $n + 1$ being the number of ordinates.

The numbers in the I_{ba} and I_{ca} columns of Table 9-1 are obtained by following the plan of Eq. (9-57). Their sum is indicated at the foot of the column, and the sum is then divided by 12 since there is this number of 15° intervals in 180°, or one half cycle. The other half cycle averages the same.

The numbers in the I_{p_1m} column are derived from those in the column headed $i_b \cos \omega t$, and similarly for I_{g_1m}. The results of these computations are as follows: $I_{ba} = 168$ ma, $I_{p_1m} = 301$ ma, $I_{ca} = 42$ ma, $I_{g_1m} = 80$ ma.

We can now calculate the following pertinent quantities:

$$\text{Power output} = P_{ac} = E_p I_{p_1} = \frac{2725}{\sqrt{2}} \frac{0.301}{\sqrt{2}} = 410 \text{ watts}$$

$$\text{Plate dissipation} = P_p = E_{bb} I_{ba} - P_{ac} = 3000 \times 0.168 - 410 = 504 - 407 = 97 \text{ watts}$$

$$\text{Plate efficiency} = \eta_p = \frac{P_{ac}}{P_{bb}} = \frac{410}{504} = 81\%$$

$$E_{gm} = e_{c_{\max}} - E_{cc} = 185 + 230 = 415 \text{ volts}$$

$$\text{Grid driving power} = P_{\text{in}} = E_g I_{g_1} = \frac{415}{\sqrt{2}} \times \frac{0.080}{\sqrt{2}} = 16.6 \text{ watts}$$

$$\text{Grid dissipation} = P_g = P_{\text{in}} - (-E_{cc} I_{ca}) = 16.6 - 230 \times 0.042 = 16.6 - 9.7 = 6.9 \text{ watts}$$

$$\text{Power amplification} = \frac{P_{ac}}{P_{\text{in}}} = \frac{410}{16.6} = 24.7$$

$$\text{Plate-load impedance} = R_{f_0} = \frac{E_{pm}}{I_{p_1m}} = \frac{2725}{0.301} = 9070 \text{ ohms}$$

$$\text{Angle of flow of plate current} = 2 \arccos \frac{QX}{QP} = 2 \arccos \frac{182}{415} = 2 \times 64^\circ = 128^\circ$$

$$\text{Angle of flow of grid current} = 2 \arccos \frac{QY}{QP} = 2 \arccos \frac{230}{415} = 2 \times 56^\circ = 112^\circ$$

Note that the graphical computations check quite well with the assumptions.

9-6. The Operation and Adjustment of a Tuned Power Amplifier. The plate-bias voltage for a tuned power amplifier is usually supplied by a rectifier although in some applications a generator is employed. It is even possible to operate with raw alternating voltage as the plate supply e_{bb}, in which case the tube functions as an amplifier only during the part of the cycle when e_{bb} is positive. The r-f output then rises from zero to a maximum and back to zero again as e_{bb} passes through this positive half cycle. The bandwidth of the output is very wide, which makes the arrangement impractical in most cases.

In Fig. 9-1 the source of grid bias is shown as a battery. This arrangement or a generator or rectifier substitute is commonly used when a class C amplifier is used to produce an a-m wave (see Art. 12-5). However, for an ordinary tuned power amplifier, it may be preferable to use grid-leak bias in one of the arrangements shown in Fig. 9-12. The direct component of the grid current passes through R_g to make the required bias, while the alternating components pass through C_g. If the grid drive increases for any reason, greater grid current and grid bias result, and thus $e_{c_{max}}$ is only slightly greater than before. Hence the output of the plate circuit is little affected by changes in grid excitation. If the grid excitation fails, the bias becomes zero and the plate current will probably become so great that damage to the tube will result, unless the plate circuit is equipped with a circuit breaker or a cathode resistor is employed.

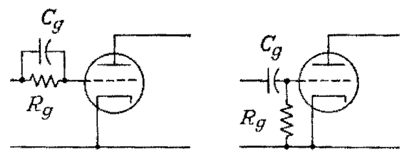


FIG. 9-12. Grid-leak-bias arrangements.

Cathode bias can also be used, but the usual high value of grid bias together with the large plate current makes a considerable power loss. Except as an auxiliary to grid-leak bias to limit the plate current when grid excitation is lost, cathode bias is not widely used for this type of circuit.

Since in the case of triodes considerable energy is fed from the plate circuit to the grid tank by way of C_{gp} for frequencies lower than the resonant frequency of the plate tank (see Art. 7-5), self-oscillation is likely to occur. This can be avoided if some form of neutralization is employed. Two methods for doing this are diagramed in Figs. 12-5 and 12-9. The general principles of neutralization involve either increasing the impedance of the plate-to-grid path by tuning C_{gp} to resonance with a parallel inductor or feeding to the grid tank another current, the effect of which is to cancel that caused by the current through C_{gp}.

The plate-current pulse in a tuned power amplifier flows at the time when e_b is near the value $e_{b_{min}}$. Hence if $e_{b_{min}}$ is not small, the plate current can be dangerously high. In adjusting a class B or class C amplifier for proper operation, if the plate tank is not yet adjusted to resonance at the signal frequency, its impedance is low and the alternating plate voltage is small. Hence $e_{b_{min}}$ is high, and the plate dissipation is likely to be excessive. The safest practice is to lower the value of E_{bb} until the tank adjustments are approximately correct, as indicated by a minimal reading of the average-plate-current meter. The grid tank is likewise adjusted to resonance by noting a maximum reading of the average-grid-current meter. For neutralizing, E_{bb} is made zero, and some device (such as a wavemeter or a small incandescent lamp coupled to the plate tank) is employed to indicate plate-tank r-f current. Since the only path for energy from the grid tank to the plate tank should be through C_{gp} and the neutralizing circuit, adjustments can then be made until the plate-tank current is a minimum. Some direct electromagnetic coupling between the grid and plate tanks may make it impossible to obtain zero tank current. Since these adjustments detune the grid and plate tanks to some extent, it is well to retune these circuits and then repeat the neutralizing adjustments until optimum conditions are obtained.

Except at very high frequencies, tetrodes and pentodes require no neutralization since the energy fed through the small plate-to-grid capacitive coupling is usually less than the circuit losses. Frequency multipliers, even when triodes are used, do not need neutralizing since the frequency of the current fed back differs from the resonant frequency of the grid tank and self-oscillation is not induced.

REFERENCES

1. Eastman, A. V.: "Fundamentals of Vacuum Tubes," 3d ed., McGraw-Hill Book Company, Inc., New York, 1949.
2. Seely, S.: "Electron-tube Circuits," McGraw-Hill Book Company, Inc., New York, 1950.
3. Ryder, J. D.: "Electronic Fundamentals and Applications," Prentice-Hall, Inc., New York, 1950.
4. "The Radio Amateur's Handbook," American Radio Relay League, Inc., West Hartford, Conn. (Annual.)

PROBLEMS AND QUESTIONS

Characteristic curves for several types of tubes are contained in Appendix A.

1. Prove that the circuit of Fig. 9-6 is equivalent to that of Fig. 9-5 if the secondary is tuned to series resonance.

2. The loaded Q_L of a tank depends upon the energy-storage capacity compared with the rate of energy dissipation. For a given coupled-in load the Q_L can be adjusted to the desired value by properly choosing C_1. Starting with Eq. (9-16), and assuming E_{pm} to be approximately E_{bb}, prove that $C_1 \approx \frac{2Q_L I_{ba} \eta_p}{\omega_0 E_{bb}}$. This relationship is useful in adjusting the tank circuit.

3. An 830-B triode is used in a single-tube class B r-f amplifier (to be used to amplify a modulated r-f signal). Maximum ratings are $E_{bo} = 1000$ volts, $I_{ba} = 100$ ma, $P_p = 60$ watts. Use $E_{bo} = 1000$ volts and a parallel-tuned plate tank with $Q_L = 12$, tuned to the carrier frequency $f_0 = 1$ Mc. Approximate values of the tube parameters are $\mu = 25$, $r_p = 8000$ ohms. Assume $p = 2$. Determine R_{f_0}, E_{gm}, E_{cc}, $e_{c_{\max}}$, $e_{b_{\min}}$, E_{pm}, I_{p_1m}, I_{ba}, E_p, I_t, P_{ac}, P_p (check), η_p, L, and C. If this operation is for the greatest r-f amplitude for 100 per cent modulation, what is the value of E_{gm} for the unmodulated carrier?

4. Derive expressions (9-48) and (9-49). Note that the value of the grid voltage for plate-current cutoff depends upon the instantaneous plate voltage at the instant of cutoff.

5. An Eimac type 100TH triode is used in a class C amplifier. $E_{bb} = 2000$ volts, and the desired plate power output is 235 watts. Make a design and compute the values of the same quantities as were determined in the example of Art. 9-5. Assume 70 per cent efficiency.

6. Sketch circuit diagrams showing how neutralization of a triode tuned amplifier can be effected by the methods suggested in Art. 9-6.

CHAPTER 10

FEEDBACK AMPLIFIERS

10-1. Distortion and Stability. It is well known that the output of an amplifier does not have a voltage waveform exactly a replica of the input signal. Because of the presence of inductance and capacitance in the tubes and circuits, the various frequency components of the signal are affected by different amplifications and phase shifts. Also because of nonlinear action of the tube itself, the output voltage is not exactly proportional to the input voltage. Again, if one or more of the bias voltages E_{co}, E_{bo}, etc., varies, the operating point shifts and the resulting changes in μ, r_p, and g_m affect the amplification. The latter effect is sometimes useful as a volume control, or it may be the basis for a method of changing the waveform of the output in an easily controlled manner. The latter process, called modulation, sometimes takes place when not wanted, as, for example, when a poorly filtered plate-supply voltage is used. Whether this occurs in an a-f amplifier or in an r-f amplifier, the result may be noticeable in a loudspeaker as hum. The change in gain with change in operating point may be very objectionable for another reason: If a battery-operated amplifier has a certain normal gain, this gain may increase upon renewal or recharge of the batteries. This may not seem serious at first thought, but suppose we consider a long-distance telephone system. The energy given by the microphone alone may be sufficient for satisfactory local operation, but it normally needs building up by repeater (amplifier) units two or more times every 100 miles of line. Since these repeaters are distant from one another, they operate independently and a rise in gain for one is not automatically compensated for by a lower gain in the next. Hence a distant repeater may find itself with a signal far too weak or maybe too strong. What is needed for each repeater is a stabilizer which will make the amplification independent of small changes in operating biases.

10-2. Interference.[1,2] In addition to the distortion and modulation products discussed in the preceding article, the output of an amplifier contains interference and noise components which have frequencies that are unrelated to the input signal. Plate-supply hum in a-f amplifiers may be considered as interference. This hum may be removed by more complete filtering, or it may be eliminated by the feedback processes discussed in this chapter. Other interference sources will now be considered.

1. Pickup due to magnetic and electrostatic fields is interference. For example, the region surrounding a power transformer contains a strong magnetic field varying at the power frequency. Any circuit conductor, as, for instance, a grid lead, will gain an induced voltage which may cause serious hum either by direct amplification or by the modulation process. It is well to place a power transformer at some distance from such conductors or, if placed close through necessity, to use a magnetic shield and orientate the transformer axis to obtain minimum effect.

Electrostatic fields cause difficulty. For example, if a grid lead is near a 60-cps ungrounded power lead, a current may flow through the stray capacitance between them. If the entire circuit has low impedance, no harm is done, but in this case the impedance between grid and cathode is high and the flow of current through a series circuit consisting of the power generator, the power line, the stray capacitance, the high grid impedance, ground, and the return to the generator causes an appreciable grid signal which appears in the output as hum. Hence one should keep the grid leads away from other a-c leads.

2. Radio-frequency interference often arises from ignition systems in automobiles, neon signs and fluorescent lights, sparking commutators in motors, and the like. The energy of this interference is spread through a wide band of frequencies and is often difficult to filter out of the r-f portion of a radio. Once the interfering voltage is in these amplifiers, modulation takes place in the nonlinear parts of the circuit and an audible output results. Suppressors at the source of this interference, which eliminate the r-f generation, are the best cure for this difficulty.

3. Alternating-current heating of cathodes is an important source of hum. If the cathode is of the filament type, the grid and plate return should be to the center tap on the heater-transformer winding or to a center-tapped resistor connected across the heater terminals. This makes the return to a point of average cathode potential, often reducing the hum in high-level tubes to a satisfactory value. For tubes using cathodes of large thermal capacity, indirectly heated with alternating current, this problem is largely eliminated, but not entirely so. The magnetic field surrounding the heater, the emission from the heater, the capacitance between heater and cathode—all are small but sometimes important sources of noise. If the cathode is grounded, it is usually best to connect one side of the heater or, better still, the center tap of the transformer winding to the cathode. Sometimes it is necessary to make the heater 10 or more volts positive relative to the cathode to keep heater emission from affecting the grid.

Even with d-c heating of the cathode there is some irregularity in cathode emission, which causes noise. One cause of this in heater-type cathodes is believed to be the movement of the heater within the cathode sleeve. This source of noise is of very low frequency content and is an important item in limiting the smallness of the grid signal which can be used with d-c amplifiers.

4. Microphonic noises arise from vibration of the tube elements. Removal of the cause, mounting of the tubes on rubber, placing heavy caps on the tubes—all are helpful in reducing this noise effect.

5. Noise due to random motion of electrons in a vacuum tube is always present and is a final limitation on the amount of amplification that can be made. The emission of the cathode is random, and this tends to make the operation noisy. However, the space charge present with normal operation makes the output considerably more noise-free. If there is a screen grid, the choice between screen or plate as the target for an electron seems to be random and hence pentodes are noisier than triodes.

6. Thermal noise is that due to the voltage which appears between the terminals of a resistor and is caused by the random motion of the free electrons in the resistor. The noise voltage is greater for carbon resistors than for wire-wound ones. The effect increases with temperature rise. The components of this noise voltage lie in a very wide band of frequencies.

To a radio listener hum is unmistakable. The noises due to thermal agitation and random electron motion in tubes appear as a background hiss when amplifiers with high gain are employed. Much study has been made of the problems relating to distortion and noise in an effort to raise the workable limit of amplification or to lower the limit on the size of signal that can be successfully amplified. In this chapter we shall study the effects produced by the process known as feedback.

10-3. How Negative Feedback Helps. If a voltage taken from the output of an amplifier is returned to the input, where it is combined with the normal input signal, a process termed feedback exists and the amplifier output is thereby modified from its nonfeedback value. The returned voltage may either increase or decrease the actual grid voltage for the first tube. In the latter event the feedback is termed negative.

The action of feedback is somewhat complicated, but even a partial understanding of its effect is helpful before a more complete study is attempted. Let us study a simple example.

Figure 10-1*a* shows in block-diagram form an amplifier having a gain of $10\underline{/180^\circ}$ and in which a noise voltage appears at a point somewhere between input and output. In order to obtain a signal output of 10 volts, it is necessary to apply an input signal of 1 volt. As shown in the figure, the signal-to-noise ratio in the output is 50. The distortion produced in the amplifier is assumed negligible for this amplitude of input signal.

Now, if it is desired to increase the signal-to-noise ratio to, say, 250, we may increase the input voltage as shown in Fig. 10-1*b*. Since noise is independent of the magnitude of the signal, the noise output remains the same. The difficulty with this plan is that a fivefold increase in signal greatly increases the distortion in the amplifier and may even overdrive it so that the operation is no longer class A.

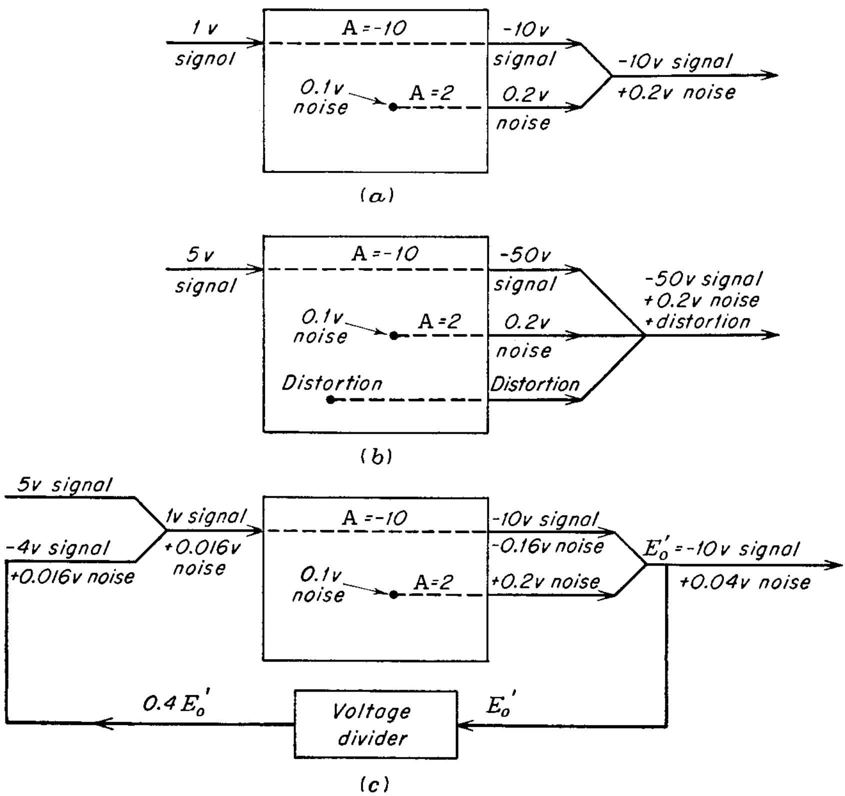


FIG. 10-1. (*a*) An amplifier with output signal-to-noise ratio equal to 50. (*b*) With increased input signal the output signal-to-noise ratio is increased to 250, but distortion results. (*c*) With increased input signal and with negative feedback, the output signal-to-noise ratio is 250 and the distortion is negligible.

An alternate method for achieving the desired signal-to-noise ratio of 250 is shown in Fig. 10-1*c*. If a voltage equal to 0.4 of the output voltage (signal, noise, and distortion) is fed back into the input in such phase that the resultant input signal is decreased, the results shown are obtained. The total input grid signal being again 1 volt, the distortion is reduced to at least its original negligible state. The signal output is again at its original desired level, and the signal-to-noise ratio has been raised to 250.

10-4. Feedback and Its Effect on Amplification. Let us begin our detailed study of feedback by determining its effect on the amplification. The simplified block diagram of Fig. 10-2 represents an amplifier having an output voltage $\mathbf{E}'_o$ due to the combined action of an amplifier with amplification **A** and of a feedback circuit with the ratio β, where $\beta = \mathbf{E}_{fb}/\mathbf{E}'_o$. β and **A** are both complex quantities. $\mathbf{V}_g$ and $\mathbf{E}_{fb}$ are actually combined at the input, but it is somewhat more convenient to show the outputs combined. Except for second-order effects, which are ignored in these discussions, the results are the same in both cases. From the diagram of Fig. 10-2

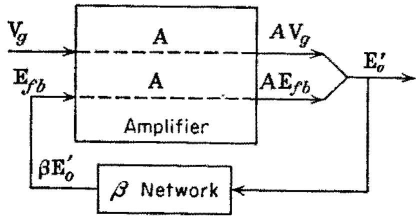


FIG. 10-2. Block-diagram representation of a feedback amplifier.

$$\mathbf{E}'_o = (\mathbf{V}_g + \mathbf{E}_{fb})\mathbf{A} = \mathbf{V}_g\mathbf{A} + \beta\mathbf{E}'_o\mathbf{A} \tag{10-1}$$

or

$$\mathbf{E}'_o = \frac{\mathbf{A}\mathbf{V}_g}{1 - \beta\mathbf{A}} \tag{10-2}$$

The over-all gain $\mathbf{A}'$ with feedback is therefore

$$\mathbf{A}' = \frac{\mathbf{E}'_o}{\mathbf{V}_g} = \frac{\mathbf{A}}{1 - \beta\mathbf{A}} \tag{10-3}$$

Hence the gain **A** of the amplifier alone is affected by a factor $1/(1 - \beta\mathbf{A})$ upon the application of feedback.

From Eq. (10-3) it is seen that A' exceeds A if $|1 - \beta\mathbf{A}| < 1$. Under this condition the feedback is called positive, or regenerative. If $1 - \beta\mathbf{A} = 0$, A' becomes infinite even though A is finite and hence output can be obtained from the amplifier with no input. This condition leads to oscillation in the circuit.

If $|1 - \beta\mathbf{A}| > 1$, A' is less than A and the feedback is termed degenerative, or negative. It is this type of feedback which holds our interest in this chapter.

10-5. Effect of Negative Feedback on Stability. Let us consider the effect of a change in the gain of the amplifier alone on the over-all gain of a feedback amplifier. To do this, consider **A** to suffer a small increment $d\mathbf{A}$. Then from Eq. (10-3)

$$d\mathbf{A}' = \frac{(1 - \beta\mathbf{A})\,d\mathbf{A} + \beta\mathbf{A}\,d\mathbf{A}}{(1 - \beta\mathbf{A})^2} \tag{10-4}$$

or

$$\frac{d\mathbf{A}'}{\mathbf{A}'} \times 100 = \frac{1}{1 - \beta\mathbf{A}}\frac{d\mathbf{A}}{\mathbf{A}} \times 100 \tag{10-5}$$

(The rules for the differentiation of complex variables are similar to those for real variables.) Hence the percentage change in the over-all gain of an amplifier with feedback is equal to the percentage change in the gain of the amplifier alone times a factor $1/(1 - \beta\mathbf{A})$.

If $1 - \beta\mathbf{A}$ is very large, then $\beta\mathbf{A}$ is very large and

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \beta\mathbf{A}} = \frac{\beta\mathbf{A}}{1 - \beta\mathbf{A}}\frac{1}{\beta} \approx -\frac{1}{\beta} \tag{10-6}$$

showing $\mathbf{A}'$ to be independent of the tube parameters and their variation. Also if $\beta\mathbf{A}$, called the feedback factor, is made very large, the amplification can have any desired characteristic by building the β circuit to have the inverse of that characteristic.

10-6. Effect of Negative Feedback on Distortion and Noise. If distortion and noise can be considered as changes in the amplification of the

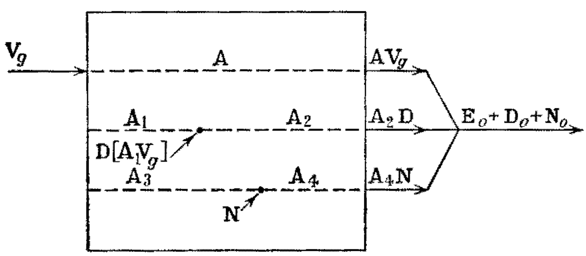


FIG. 10-3. Block diagram of a nonfeedback amplifier.

amplifier without feedback, the above discussion shows that these undesirable effects are reducible by a factor $1/(1 - \beta\mathbf{A})$ by negative feedback.

This general statement needs some qualifications, however, as it is not always applicable. Hence it is desirable to investigate the matter further. To do so, we may begin with Fig. 10-3, the diagram for a nonfeedback amplifier with a single-frequency signal $\mathbf{V}_g$ and amplification $\mathbf{A}$. For simplicity we assume $\mathbf{A}$ to be the same for all frequencies. The subscripted $\mathbf{A}$'s represent amplifications between input or output and the points of distortion, or noise introduction. $\mathbf{D}$ is the distortion of the signal such as that due to nonlinear action of some tube. If we consider one frequency component only, some harmonic of V_g, we may use complex representation for D. The value of $\mathbf{D}$ depends upon the signal level at the point of production, and hence it is sometimes written as $\mathbf{D}[\mathbf{A}_1\mathbf{V}_g]$, that is, $\mathbf{D}$ is a function of $\mathbf{A}_1\mathbf{V}_g$. Actually $\mathbf{D}$ depends upon the total voltage level at the point of production, but for simplicity it will be considered to depend upon the signal level alone. Noise is usually of an impulse type which cannot be approximated by a finite number of harmonics. It suits our convenience to pretend that it is cyclic. Hence $\mathbf{N}$ is one frequency component of a cyclic disturbance occurring at a point somewhere in the amplifier. The noise $\mathbf{N}$ is independent of the signal.

We may now write

$$\mathbf{E}_o + \mathbf{D}_o + \mathbf{N}_o = \mathbf{A}\mathbf{V}_g + \mathbf{A}_2\mathbf{D}[\mathbf{A}_1\mathbf{V}_g] + \mathbf{A}_4\mathbf{N} \tag{10-7}$$

From Eq. (10-3) it is seen that the application of negative feedback decreases the over-all gain of an amplifier. As it is generally undesirable to decrease the output voltage, there is a choice between two procedures. One is to increase the input voltage by a factor $1 - \beta\mathbf{A}$, and the other is

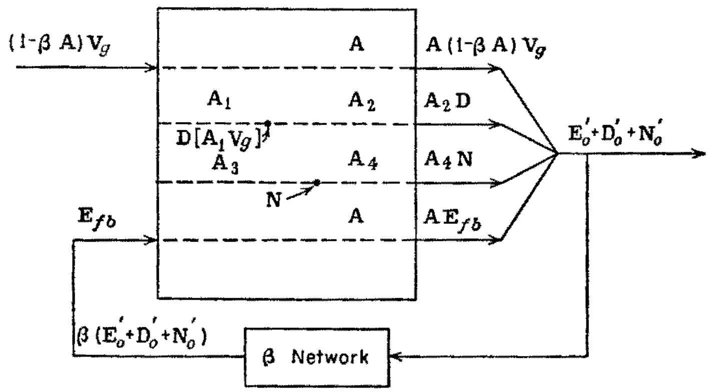


FIG. 10-4. A feedback amplifier with increased signal.

to increase the amplifier gain sufficiently to make up for the loss due to feedback.

Figure 10-4 shows an amplifier with feedback. The amplification of the amplifier alone is $\mathbf{A}$, the same as that of the nonfeedback amplifier of Fig. 10-3. The input voltage is $(1 - \beta\mathbf{A})\mathbf{V}_g$, making $\mathbf{E}_o'$ equal to $\mathbf{A}\mathbf{V}_g$, the same as the output of the nonfeedback amplifier. The signal level at the point of distortion production may be called $\mathbf{E}''$, and it follows that

$$\mathbf{E}'' = (1 - \beta\mathbf{A})\mathbf{V}_g\mathbf{A}_1 + \mathbf{E}_o'\beta\mathbf{A}_1 \tag{10-8}$$

or since $\mathbf{E}_o' = \mathbf{A}\mathbf{V}_g$,

$$\mathbf{E}'' = (1 - \beta\mathbf{A})\mathbf{V}_g\mathbf{A}_1 + \mathbf{A}\mathbf{V}_g\beta\mathbf{A}_1 = \mathbf{A}_1\mathbf{V}_g \tag{10-9}$$

Hence $\mathbf{D} = \mathbf{D}[\mathbf{A}_1\mathbf{V}_g]$ at the point of distortion production.

We may now write

$$\mathbf{E}_o' + \mathbf{D}_o' + \mathbf{N}_o' = \mathbf{A}(1 - \beta\mathbf{A})\mathbf{V}_g + \mathbf{A}_2\mathbf{D}[\mathbf{A}_1\mathbf{V}_g] + \mathbf{A}_4\mathbf{N} + \beta\mathbf{A}(\mathbf{E}_o' + \mathbf{D}_o' + \mathbf{N}_o') \tag{10-10}$$

or

$$\mathbf{E}'_o + \mathbf{D}'_o + \mathbf{N}'_o = \mathbf{A}\mathbf{V}_g + \frac{\mathbf{A}_2\mathbf{D}[\mathbf{A}_1\mathbf{V}_g]}{1 - \beta\mathbf{A}} + \frac{\mathbf{A}_4\mathbf{N}}{1 - \beta\mathbf{A}} \tag{10-11}$$

By comparing Eqs. (10-7) and (10-11) we may draw some conclusions. Noise is affected by a factor $1/(1 - \beta\mathbf{A})$ upon the application of negative feedback. If the signal output level is maintained constant by increasing the input, this results in an improvement in the signal-to-noise ratio by a factor $1 - \beta\mathbf{A}$. Note that a noise voltage is least disturbing in the output if it is produced in the output plate circuit where $\mathbf{A}_4 = 1$.

If distortion occurs in the output circuit, $\mathbf{A}_1 = \mathbf{A}$, $\mathbf{A}_2 = 1$, and hence the distortion in the output of the nonfeedback amplifier is $\mathbf{D}[\mathbf{A}\mathbf{V}_g]$, while

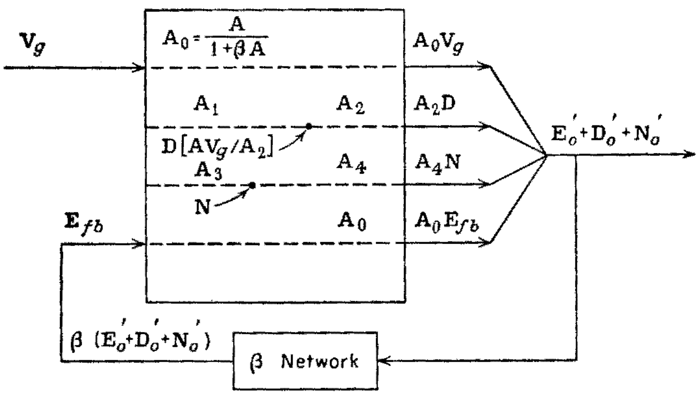


FIG. 10-5. A feedback amplifier with increased amplifier gain.

for the feedback amplifier it is $\mathbf{D}[\mathbf{A}\mathbf{V}_g]/(1 - \beta\mathbf{A})$, showing a lowering in percentage distortion $(1 - \beta\mathbf{A})$ times.

If distortion occurs in the grid circuit of the first tube, the distortion in the output of the nonfeedback amplifier is $\mathbf{A}\mathbf{D}[\mathbf{V}_g]$, while for the feedback amplifier it is $\mathbf{A}\mathbf{D}[\mathbf{V}_g]/(1 - \beta\mathbf{A})$, showing again a lowering in percentage distortion $(1 - \beta\mathbf{A})$ times. Thus, if the input voltage is increased sufficiently to make the output voltage the same as without feedback, distortion and noise in the output are affected by a factor $1/(1 - \beta\mathbf{A})$, both absolutely and percentagewise.

We may now consider the second case. With the input signal the same as in the nonfeedback case, it is necessary to use a higher gain for the amplifier alone in order that $\mathbf{E}'_o$ again be equal to $\mathbf{A}\mathbf{V}_g$, the output of the nonfeedback amplifier. One can designate this new value of gain of the amplifier alone by $\mathbf{A}_0$. Figure 10-5 is a block diagram useful in explaining the action. The details may be worked out by the student (see Prob. 5 at the end of this chapter). The results can be summarized as follows:

The required gain of the amplifier alone is

$$\mathbf{A}_0 = \frac{\mathbf{A}}{1 + \beta\mathbf{A}} \tag{10-12}$$

The output with feedback becomes

$$\mathbf{E}'_o + \mathbf{D}'_o + \mathbf{N}'_o = \mathbf{A}\mathbf{V}_g + \frac{\mathbf{A}_2\mathbf{D}[\mathbf{A}\mathbf{V}_g/\mathbf{A}_2]}{1 - \beta\mathbf{A}_0} + \frac{\mathbf{A}_4\mathbf{N}}{1 - \beta\mathbf{A}_0} \tag{10-13}$$

A study of formula (10-13) shows that distortion is reduced by feedback partly by cancellation and partly by less generation because of smaller signal amplitude. Noise arising near the output is reduced, making the signal-to-noise ratio larger by a factor $1 - \beta\mathbf{A}_0$. However, this improvement decreases for noises generated nearer the input, and at the input grid no improvement is effected by negative feedback.

10-7. Feedback in Amplifiers with Low-level Input. If the signal voltage available from a pickup or other device is very small, it may be necessary to amplify it. Of the noises generated in the amplifier, that produced in the first tube is the most disturbing. We have seen that in this case the signal-to-noise ratio in the output cannot be improved by making a feedback-amplifier loop include the first stage, although it will be of benefit for noises originating in later stages and for nonlinear distortion occurring in any stage. Also Eq. (10-6) shows that frequency distortion is decreased if it is caused by the amplifier circuit. The use of gain preceding the feedback-amplifier loop is beneficial. If a vacuum-tube amplifier is used to produce this gain, the noise generated in it becomes an increased problem. However, it is sometimes possible to use an input transformer, which in itself is relatively noise free. This step-up turns ratio results in an increased signal on the first grid, thus giving improved signal-to-noise ratio. The benefits do not extend to noise originating in the pickup, for, of course, to the amplifier that is indistinguishable from the desired signal.

In d-c amplifiers a slow drift in the plate current of the first tube, due to any cause other than the signal, is exactly like the signal effect, and hence feedback is useless in stabilizing the d-c amplifier against this fault. For drift in later stages, which is less important anyway, feedback does have a stabilizing effect.

10-8. Voltage and Current Feedback. In feedback amplifiers a voltage related to the output voltage is fed back to the amplifier input. The manner in which this voltage is obtained determines to some degree the properties of the feedback amplifier. Two arrangements are of especial interest.

If the feedback voltage is obtained in such a fashion that β is independent of the load impedance, the term "voltage feedback" is used. If, on the other hand, the value of β is inversely proportional to the load impedance, the name "current feedback" is used.

To find the gain with feedback, we may use Eq. (10-3),

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \beta\mathbf{A}} \tag{10-3}$$

where $\beta = \mathbf{E}_{fb}/\mathbf{E}_o'$. Figure 10-6 shows the circuit for a simple amplifier employing voltage feedback. The load $\mathbf{Z}_L$ for middle frequencies is the parallel combination of R_b, R_g, and $R_1 + R_2$. The value of β in this case is $R_2/(R_1 + R_2)$, and if the load is varied by changing R_b and/or R_g (but not R_1 and R_2, which serve only as a high-impedance voltage divider), the value of β remains constant, as it should. For this amplifier we may use Eq. (7-3) to write

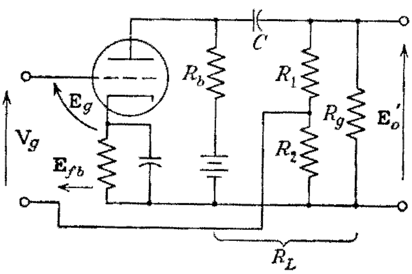


FIG. 10-6. An amplifier with simple voltage feedback.

$$\mathbf{A} = \frac{-\mu\mathbf{Z}_L}{r_p + \mathbf{Z}_L} \tag{10-14}$$

Substituting from Eq. (10-14) into Eq. (10-3) yields

$$\mathbf{A}' = \frac{-\mu\mathbf{Z}_L/(r_p + \mathbf{Z}_L)}{1 + \beta\dfrac{\mu\mathbf{Z}_L}{r_p + \mathbf{Z}_L}} = \frac{-\mu\mathbf{Z}_L}{r_p + (1 + \mu\beta)\mathbf{Z}_L} = \frac{-\mu/(1 + \beta\mu)}{\dfrac{r_p/(1 + \beta\mu)}{\mathbf{Z}_L} + 1} \tag{10-15}$$

Hence

$$\mathbf{E}_o' = \frac{[-\mu/(1 + \beta\mu)]\mathbf{V}_g}{\dfrac{r_p/(1 + \beta\mu)}{\mathbf{Z}_L} + 1} \tag{10-16}$$

Since β is independent of $\mathbf{Z}_L$, we see that if $1 + \beta\mu$ has considerable magnitude so that $|r_p/(1 + \beta\mu)|$ is small compared with Z_L, the output voltage becomes approximately independent of the load. Thus this amplifier tends to take on the character of a constant-voltage generator, whence the name "voltage feedback" is derived for this type of circuit.

For current feedback

$$\beta = \frac{k}{\mathbf{Z}_L} \tag{10-17}$$

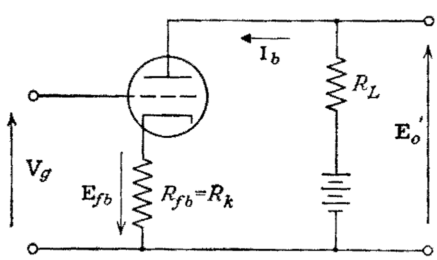


FIG. 10-7. An amplifier with simple current feedback.

where k is independent of $\mathbf{Z}_L$. Figure 10-7 shows an amplifier with simple current feedback. Here

$$\beta = \frac{\mathbf{E}_{fb}}{\mathbf{E}'_o} = \frac{R_k}{R_L} \tag{10-18}$$

The gain from grid to output is

$$\mathbf{A} = \frac{\mathbf{E}'_o}{\mathbf{E}_g} = \frac{-\mu R_L}{r_p + R_L + R_k} \tag{10-19}$$

Hence

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \beta\mathbf{A}} = \frac{-\mu R_L/(r_p + R_L + R_k)}{1 + \dfrac{R_k}{R_L}\dfrac{\mu R_L}{r_p + R_L + R_k}} = \frac{-\mu R_L}{r_p + (1 + \mu)R_k + R_L} \tag{10-20}$$

and

$$\mathbf{I}'_p = -\frac{\mathbf{E}'_o}{R_L} = \frac{-\mathbf{A}'\mathbf{V}_g}{R_L} = \frac{\mu\mathbf{V}_g}{r_p + (1 + \mu)R_k + R_L} \tag{10-21}$$

Since without feedback the value of $\mathbf{I}_p$ would be given by

$$\mathbf{I}_p = \frac{\mu\mathbf{V}_g}{r_p + R_k + R_L} \tag{10-22}$$

it is seen that the effect of this feedback is to make the plate current less dependent upon R_L. This constant-current character gives the name "current feedback" to this type of circuit.

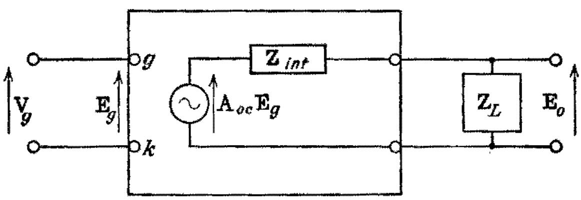


FIG. 10-8. A Thévenin's representation of a multistage amplifier.

10-9. Multistage Feedback Circuits. In many cases it is desirable to use more than one stage in the feedback amplifier. To compute the approximate performance of the circuit, it is sometimes convenient to represent the whole amplifier by a Thévenin's equivalent circuit. Such a representation is shown in Fig. 10-8 for an amplifier without feedback. To determine this circuit, one may imagine each stage of the amplifier to be represented in its ordinary linear-equivalent fashion. If $\mathbf{Z}_L$ for the last stage is removed and the output impedance obtained with the first-stage grid tied to the cathode, thus making all equivalent-generator voltages zero, the value obtained is $\mathbf{Z}_{\text{int}}$. With $\mathbf{E}_g$ again applied to the grid of the first tube of the amplifier, Thévenin's equivalent generator may be labeled with an equivalent emf $\mathbf{A}_{oc}\mathbf{E}_g$, where

$$\mathbf{A}_{oc} = \mathbf{A}_1\mathbf{A}_2\mathbf{A}_3 \cdot \cdot \cdot \mathbf{A}_{n-1}\mathbf{A}_{n_{oc}} \tag{10-23}$$

$\mathbf{A}_{oc}$ of an amplifier means the gain of that amplifier with infinite a-c load impedance.

The polarity of the equivalent generator is chosen with positive sign toward the output terminal designated as positive in labeling $\mathbf{E}_o$. Alternatively, of course, the emf may be expressed as $-\mathbf{A}_{oc}\mathbf{E}_g$ and the sense signs reversed, as has been done in preceding portions of this book. It should be noted that the new equivalent circuit resembles that for a single stage and hence simplifies the analysis.

Figure 10-9 shows an equivalent circuit for an n-stage amplifier with a feedback network. $\mathbf{V}_g$ and $\mathbf{E}_g$ are values for the first stage. The β net-

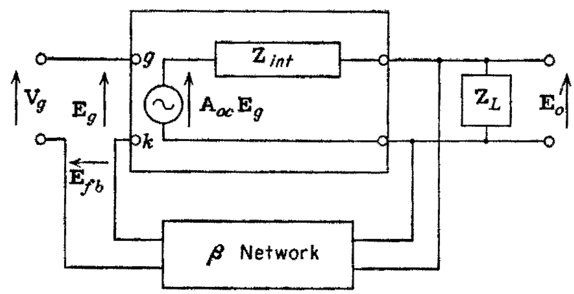


FIG. 10-9. A representation of a general voltage-feedback circuit.

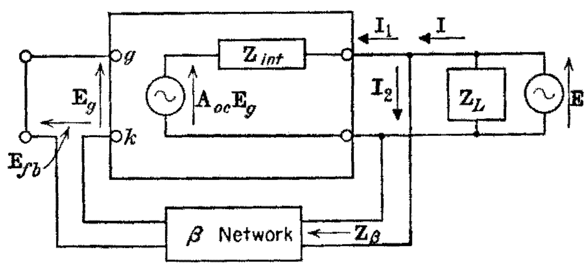


FIG. 10-10. Testing a voltage-feedback circuit for output impedance.

work is in general a four-terminal one and may range in complexity from direct connections to quite elaborate circuits.

Because of the fact that the source of $\mathbf{V}_g$ may have one side grounded and the load receiving $\mathbf{E}'_o$ may also be grounded, it may be impossible to use a circuit actually wired in the manner indicated in Fig. 10-9. In practical circuits the feedback voltage is introduced into the grid circuit of the first tube of the feedback loop in various ways, some of which are illustrated in later examples.

10-10. The Effect of Feedback on Output Impedance; Equivalent Circuits. Let us consider the general voltage-feedback circuit of Fig. 10-9, redrawn in Fig. 10-10. The voltage source for $\mathbf{V}_g$ is removed and the input shorted, while a voltage $\mathbf{E}$ is applied to the output terminals as shown. The current $\mathbf{I}$ which flows may be determined by calculating $\mathbf{I}_1$ and $\mathbf{I}_2$.

$$\mathbf{I}_1 = \frac{\mathbf{E} - \mathbf{A}_{oc}\mathbf{E}_g}{\mathbf{Z}_{\text{int}}} \tag{10-24}$$

Here

$$\mathbf{E}_g = \mathbf{E}_{fb} = \boldsymbol{\beta}\mathbf{E} \tag{10-25}$$

Hence

$$\mathbf{I}_1 = \frac{\mathbf{E}(1 - \mathbf{A}_{oc}\boldsymbol{\beta})}{\mathbf{Z}_{\text{int}}} \tag{10-26}$$

$$\mathbf{I}_2 = \frac{\mathbf{E}}{\mathbf{Z}_\beta} \tag{10-27}$$

Therefore

$$\mathbf{I} = \mathbf{E}\left[\frac{1 - \mathbf{A}_{oc}\boldsymbol{\beta}}{\mathbf{Z}_{\text{int}}} + \frac{1}{\mathbf{Z}_\beta}\right] \tag{10-28}$$

and

$$\mathbf{Z}_{\text{out}} = \frac{\mathbf{E}}{\mathbf{I}} = \frac{1}{\dfrac{1}{\mathbf{Z}_{\text{int}}/(1 - \mathbf{A}_{oc}\boldsymbol{\beta})} + \dfrac{1}{\mathbf{Z}_\beta}} \tag{10-29}$$

Thus $\mathbf{Z}_{\text{out}}$ is the internal impedance of the amplifier without feedback divided by $1 - \mathbf{A}_{oc}\boldsymbol{\beta}$, in parallel with the impedance of the feedback network as seen from the amplifier-output end. This does not include the load impedance, which is also a shunt. Hence the output impedance of an amplifier itself appears to be changed by a factor $1/(1 - \mathbf{A}_{oc}\boldsymbol{\beta})$ upon the application of voltage feedback.

The gain of the amplifier of Fig. 10-9 may be determined by using Eq. (10-3). In this case $\mathbf{A} = \mathbf{Z}_L\mathbf{A}_{oc}/(\mathbf{Z}_{\text{int}} + \mathbf{Z}_L)$, where $\mathbf{Z}_L$ includes the shunting effect of $\mathbf{Z}_\beta$. Hence

$$\mathbf{A}' = \frac{\dfrac{\mathbf{Z}_L}{\mathbf{Z}_{\text{int}} + \mathbf{Z}_L}\mathbf{A}_{oc}}{1 - \dfrac{\mathbf{Z}_L}{\mathbf{Z}_{\text{int}} + \mathbf{Z}_L}\mathbf{A}_{oc}\boldsymbol{\beta}} = \frac{\mathbf{A}_{oc}\mathbf{Z}_L}{\mathbf{Z}_{\text{int}} + \mathbf{Z}_L(1 - \mathbf{A}_{oc}\boldsymbol{\beta})} = \frac{\dfrac{\mathbf{A}_{oc}\mathbf{Z}_L}{1 - \mathbf{A}_{oc}\boldsymbol{\beta}}}{\dfrac{\mathbf{Z}_{\text{int}}}{1 - \mathbf{A}_{oc}\boldsymbol{\beta}} + \mathbf{Z}_L} \tag{10-30}$$

This result suggests the circuit of Fig. 10-11, which yields the same overall gain and which has the correct output impedance. Since $\mathbf{A}' = \mathbf{E}_o'/\mathbf{V}_g$,

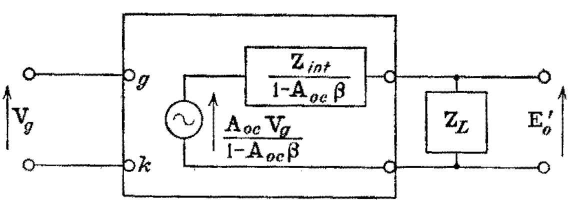


FIG. 10-11. A nonfeedback equivalent of a voltage-feedback amplifier.

the new internal emf can be expressed as

$$\frac{\mathbf{A}_{oc}}{1 - \mathbf{A}_{oc}\boldsymbol{\beta}}\mathbf{V}_g \tag{10-31}$$

showing that $\mathbf{A}_{oc}$ also has been affected by a factor $1/(1 - \mathbf{A}_{oc}\beta)$, while $\mathbf{V}_g$ replaces $\mathbf{E}_g$. Hence Fig. 10-11 is a nonfeedback equivalent of an amplifier with voltage feedback.

In the case of current feedback (Fig. 10-7) the output impedance may be determined as before by applying a voltage to the output. However, as a variation let us use the alternate method of obtaining a nonfeedback equivalent circuit. This may be done by paraphrasing Eq. (10-20) to obtain the amplification of a general current-feedback circuit of any

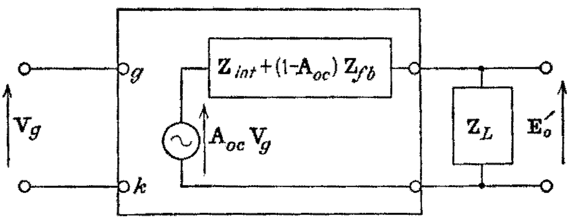


FIG. 10-12. A nonfeedback representation of a current-feedback amplifier.

number of stages, the feedback voltage being formed by the load current of the last tube passing through a feedback impedance $\mathbf{Z}_{fb}$. Then noting that the simple single-stage current-feedback amplifier has $\mathbf{A}_{oc} = -\mu$, $\mathbf{Z}_{\text{int}} = r_p$, and $\mathbf{Z}_L = R_L$, we may write

$$\mathbf{A}' = \frac{\mathbf{A}_{oc}\mathbf{Z}_L}{\mathbf{Z}_{\text{int}} + (1 - \mathbf{A}_{oc})\mathbf{Z}_{fb} + \mathbf{Z}_L} \tag{10-32}$$

Since Eq. (10-32) resembles the gain formula for a nonfeedback amplifier, we may draw Fig. 10-12 to represent a current-feedback amplifier. The effect of current feedback is that of increasing the output impedance of the amplifier.

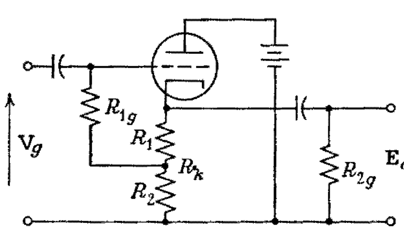


FIG. 10-13. A cathode-follower circuit.

10-11. Some Practical Feedback Amplifiers. 1. *The Cathode Follower.* The cathode follower studied in Arts. 7-24 and 7-25 is a feedback amplifier. Many of the important facts concerning this circuit have already been discussed in those articles, but it is interesting to treat this amplifier by the feedback methods. The circuit of Fig. 7-51a is repeated in Fig. 10-13. Figure 10-14 is the equivalent circuit of this amplifier, which shows how the feedback is obtained. R_{1g} is considered to be very high in value. $\mathbf{A}_{oc} = \mu$ [Eq. (10-23)]. Note that the plate is at ground potential in this circuit, and it is convenient to have the negative sense of $\mathbf{E}_o'$ on this side. $\mathbf{Z}_{\text{int}} = r_p$.

Figure 10-15 is another equivalent circuit for the cathode follower, which is obtained in the following manner: Since the equation $\boldsymbol{\beta} = \mathbf{E}_{fb}/\mathbf{E}_o' = -1$ does not contain $\mathbf{Z}_L$, this is voltage feedback. To obtain the internal impedance, one may employ Eq. (10-29). Since the impedance $\mathbf{Z}_\beta$ is infinite, the value of $\mathbf{Z}_{\text{out}}$ becomes

$$\mathbf{Z}_{\text{out}} = \frac{\mathbf{Z}_{\text{int}}}{1 - \mathbf{A}_{oc}\boldsymbol{\beta}} = \frac{r_p}{1 + \mu} \tag{10-33}$$

which is the new internal impedance. The equivalent emf for the circuit

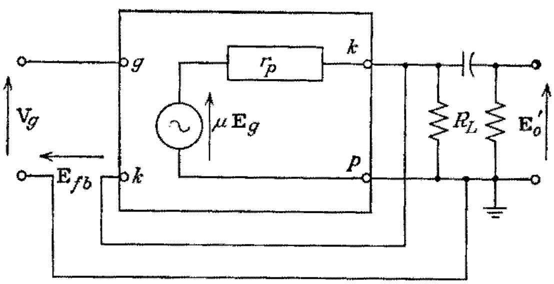


FIG. 10-14. An equivalent circuit for the cathode follower showing the feedback arrangement.

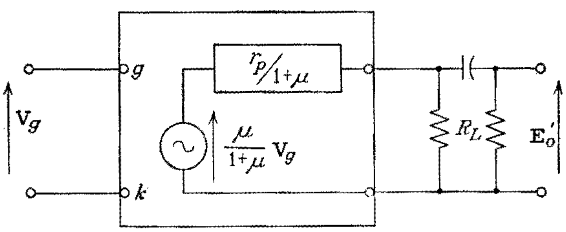


FIG. 10-15. A nonfeedback equivalent of the cathode follower.

can be expressed as

$$\frac{\mathbf{A}_{oc}\mathbf{V}_g}{1 - \mathbf{A}_{oc}\boldsymbol{\beta}} = \frac{\mu}{1 + \mu}\mathbf{V}_g \tag{10-34}$$

The over-all gain of the amplifier for middle frequencies is easily obtained from Fig. 10-15.

$$\mathbf{A}' = \frac{R_L}{r_p/(1 + \mu) + R_L}\frac{\mu}{1 + \mu} \tag{10-35}$$

which is identical with Eq. (7-89).

2. *Cathode Degeneration.* The use of a cathode resistor R_k to furnish grid bias causes a reduction in gain below the normal value at some frequencies. If a bypass capacitor C_k is not used, the gain will be the reduced value up to some high frequency where stray capacitance begins

to shunt R_k. If C_k of sufficiently high value is employed, the gain will be decreased appreciably only for frequencies below some low value. The main problem is to remove the transition from high to low gain from the operating region, and either the inclusion or the omission of C_k may be employed. Since high operating gain is usually desirable, the common solution to the problem is to place the transition at frequencies below the useful operating range.

The determination of the proper value of C_k to be employed is rather involved, and numerous schemes have been proposed for handling the matter. A common one is to use C_k sufficiently large that X_{C_k} is $0.1R_k$ at the l-f end of the operating range. This plan is also often followed in the case of C_d shunting R_d, the screen resistor, and certainly the method has the merit of simplicity. However, rather obviously, the degenerative effect of $\mathbf{Z}_k$, the parallel combination of R_k and C_k, depends not only upon its own value but also upon the tube parameters and the plate-load impedance, and the latter two items are ignored by the $0.1R_k$ rule. So while admittedly the results given by this simple plan are satisfactory in many instances, it is desirable to investigate the matter further.

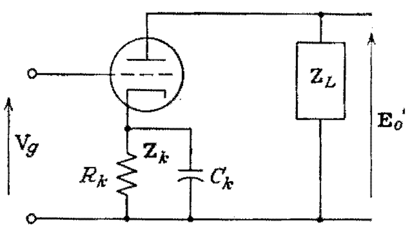


FIG. 10-16. R_k causes feedback if inadequately bypassed.

Let us refer to Fig. 10-16, in which $\mathbf{Z}_L$ represents the entire plate load, excluding $\mathbf{Z}_k$. E_{bb} is not shown since its position in the circuit is not definite and it is not important in our present consideration. Figure 10-17*a* is a linear equivalent circuit showing feedback connections. From it we see that $\beta = \mathbf{Z}_k/\mathbf{Z}_L$, and hence this is current feedback. Therefore, a more simplified equivalent circuit can be drawn, as shown in Fig. 10-17*b*. In obtaining the latter the diagram of Fig. 10-12 is used as a model. $\mathbf{A}_{oc} = -\mu$. Although the position of the actual output voltage is indefinite (depending upon the actual configuration of $\mathbf{Z}_L$), we can consider the output voltage as $\mathbf{E}_o'$.

The gain can be obtained by reference to Fig. 10-17*b* as

$$\mathbf{A}' = \frac{\mathbf{E}_o'}{\mathbf{V}_g} = \frac{-\mu\mathbf{Z}_L}{r_p + \mathbf{Z}_L + (1+\mu)\mathbf{Z}_k} = \frac{-\mu\mathbf{Z}_L/(r_p + \mathbf{Z}_L)}{1 + \dfrac{1+\mu}{r_p + \mathbf{Z}_L}\mathbf{Z}_k} = \frac{\mathbf{A}}{1 + \dfrac{1+\mu}{r_p + \mathbf{Z}_L}\mathbf{Z}_k} \qquad (10\text{-}36)$$

where $\mathbf{A}$ is the gain of the amplifier with perfect bypassing of R_k. By a somewhat involved procedure one can show that the following expression can be evolved,

$$\frac{\mathbf{A}'}{\mathbf{A}} = \frac{\mathbf{Z}(1 + j\omega C_k R_k)}{R_k(1 + j\omega C_k \mathbf{Z})} \tag{10-37}$$

where

$$\mathbf{Z} = R + jX = \frac{R_k[(r_p + \mathbf{Z}_L)/(1 + \mu)]}{R_k + [(r_p + \mathbf{Z}_L)/(1 + \mu)]} \tag{10-38}$$

It may be seen that $\mathbf{Z}$ is the effective impedance between cathode and ground, excluding C_k.

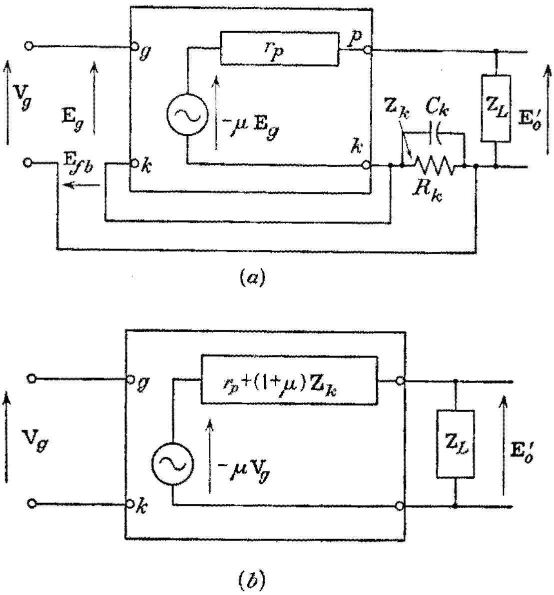


FIG. 10-17. Equivalent circuits for the amplifier of Fig. 10-16.

In the general case of any complex value of $\mathbf{Z}$, it is difficult to evolve an explicit expression for C_k. However, as an approximation the plate load $\mathbf{Z}_L$ for an R-C-coupled amplifier or for a transformer-coupled power amplifier is resistive in the operating-frequency range, and useful results can be obtained by assuming $X = 0$ so that $\mathbf{Z} = R + j0$. Then Eq. (10-37) can be changed to the form

$$\frac{\mathbf{A}'}{\mathbf{A}} = \frac{1 - j(1/\omega C_k R_k)}{1 - j(1/\omega C_k R)} \tag{10-39}$$

If we consider the transition point from low to high gain to be marked by the frequency f_t at which $A'/A = x$, this would mean

$$\frac{\sqrt{1 + (1/\omega_t^2 C_k^2 R_k^2)}}{\sqrt{1 + (1/\omega_t^2 C_k^2 R^2)}} = x \tag{10-40}$$

from which we can solve for X_{C_k} at the frequency f_l.

$$X_{C_k} = \sqrt{\frac{1 - x^2}{(x^2/R^2) - (1/R_k^2)}} \tag{10-41}$$

As an example of the use of this formula, consider the amplifier of Fig. 10-18. R_{low} is approximately 500,000 ohms, and the low end of the mid-frequency band is at $10f_1 = 10/2\pi R_{\text{low}} C = 159$ cps. Suppose it is desired to have the gain at 159 cps no less than $0.9A_{\text{mid}}$ on account of cathode degeneration. $R_L \approx 83{,}000$ ohms, and hence, by Eq. (10-38), $\mathbf{Z} = R = \dfrac{2500 \times 93{,}000/21}{2500 + 93{,}000/21} = 1600$ ohms. By formula (10-41), $X_{C_k} = 1100$ ohms, and it follows that $C_k = 0.91$ μf is the minimum value to be used.

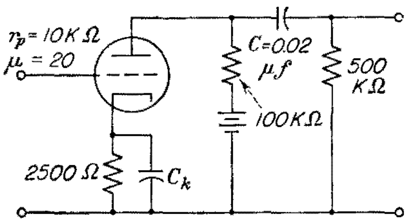


FIG. 10-18. What value of C_k should be used?

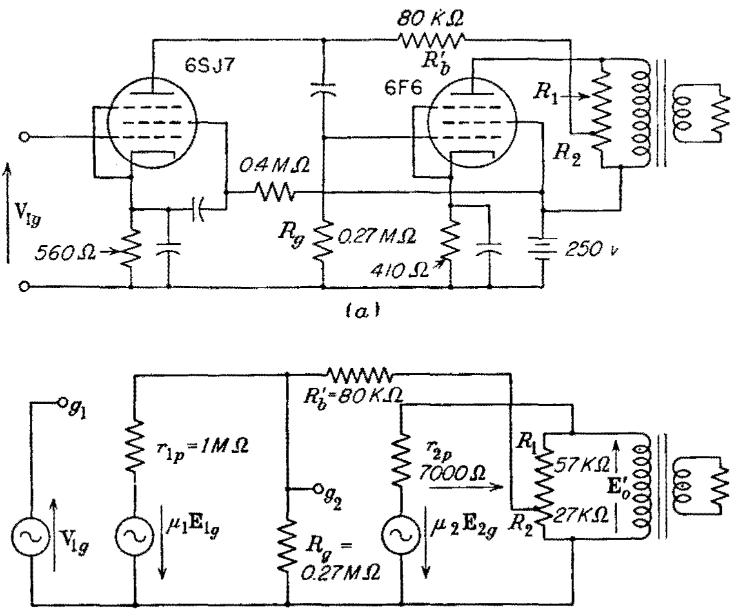


FIG. 10-19. An amplifier with a one-tube feedback loop (a) and its equivalent circuit (b).

3. *A Single-tube Voltage-feedback Arrangement.* Figure 10-19a shows a commonly used feedback arrangement, and Fig. 10-19b is its linear equivalent circuit useful for middle frequencies. For the 6F6 tube suppose tube-manual data are $R_L = 7000$ ohms, $r_p = 80{,}000$ ohms, $g_m = 2500$

micromhos, and $\mu = 200$. $P_{ac} = 3.2$ watts for a peak grid signal of 16.5 volts. There is 8 per cent harmonic distortion. It follows that $E_p = \sqrt{P_{ac}R_L} = 150$ volts and $\mathbf{A}_{\text{mid}} = -150\sqrt{2}/16.5 = -12.9$. (This estimate of $\mathbf{A}_{\text{mid}}$ is probably better than that obtained by using $-g_m R_{sh}$.)

Suppose it is desired to reduce the harmonic distortion to 2 per cent. Then $|1 - \boldsymbol{\beta}\mathbf{A}| = 8/2 = 4$, and therefore a value of $\boldsymbol{\beta} = 0.232\underline{/0°}$ is required. Since $\boldsymbol{\beta}$ depends upon the resistors in the interstage coupling circuit, it is necessary to consider the design of the 6SJ7 circuit.

The required peak grid drive for the 6F6 is 16.5 volts, and hence the output of the 6SJ7 must be $4 \times 16.5 = 66$ volts peak value, or 46.7 volts rms. The tube manual or Appendix B is consulted for satisfactory circuit values. The ones chosen are shown in the figure. In particular $R_g = 0.27$ megohm, $R_b = 0.1$ megohm, $A = 94$, and the stated percentage distortion is 4.2.

It should be noted that R_b is not represented by a single resistor. R_b is equal to R_b' in series with the parallel combination of R_2 as one branch and R_1 and the transformer primary as the other branch. If we choose R_b' as 80,000 ohms, then the combination of the other resistances named above must amount to approximately 20,000 ohms. It is desirable that $R_1 + R_2$ be quite high so as not appreciably to lower the load on the 6F6 tube below that given by the transformer.

The voltage drop across R_g is $\mathbf{E}_{2g}$. By superposition, this voltage drop is the sum of the voltage drops caused by the two generators $\mu_1\mathbf{E}_{1g}$ and $\mu_2\mathbf{E}_{2g}$. Since the drop caused by the former is the normal nonfeedback value of $\mathbf{E}_{2g}$, it follows that the drop caused by the latter is $\mathbf{E}_{fb}$. Since $\boldsymbol{\beta}$ is the ratio of $\mathbf{E}_{fb}$ to $\mathbf{E}_o$, we may write $\boldsymbol{\beta} = \dfrac{R_2}{R_1 + R_2}\dfrac{0.21}{0.21 + 0.08} = 0.232$, where 0.21 megohm is the parallel resistance of r_{1p} and R_g (assuming r_{1p} to be approximately 1 megohm). This assumes that R_b', R_g, and r_{1p} constitute a negligible shunt on R_2.

Now we can solve for R_1 to obtain $R_1 = 2.1R_2$. If we choose R_2 as 27,000 ohms, then $R_1 = 57,000$ ohms. A check on the parallel combination of the resistors mentioned above yields 19,000 ohms as the a-c resistance and 18,300 ohms as the d-c resistance. These values are satisfactory close to the desired value of 20,000 ohms.

The gain of the 6SJ7 is 94. Hence an rms input signal of $46.7/94 = 0.5$ volt is needed. The harmonic distortion of 4.2 per cent given by the tube manual applies here since the desired output voltage agrees with the tube-manual value. This percentage distortion in this case will roughly add to that of the 6F6 stage, making the total about 6 per cent. Thus we see that our purpose of reducing the distortion has been poorly accomplished. Increasing $\boldsymbol{\beta}\mathbf{A}$ may not improve the results as its effect is only on distortion originating in the 6F6 stage.

4. *A Two-tube Feedback Circuit.* Figure 10-20 shows an example of this type of feedback arrangement. The voltage fed back from the output is introduced into the cathode circuit, rather than the grid circuit, of the first tube to offset the phase reversal of the additional stage. Since a voltage $\mathbf{E}_{fb}$ in the cathode circuit of the 6SJ7 acts in both the grid and the plate circuits of this tube, its total effect is the same as $(\mu + 1)\mathbf{E}_{fb}$ volts in the plate circuit alone or $[(\mu + 1)/\mu]\mathbf{E}_{fb}$ volts in the grid circuit alone. μ being very high for a 6SJ7 tube, $(\mu + 1)/\mu$ is practically unity, and hence feedback to its cathode in this case is equivalent to feedback of the same voltage to its grid alone.

The circuit values for the 6F6 are those of the previous example: R_L = 7000 ohms, r_p = 80,000 ohms, g_m = 2500 micromhos, μ = 200. For the

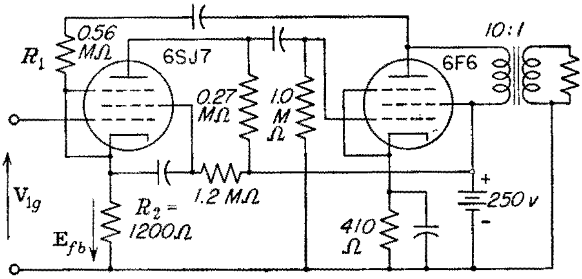


FIG. 10-20. A two-stage feedback circuit.

6SJ7 it should be noted that its plate current in passing through R_2 causes feedback, which reduces the gain for this stage and to some degree lowers its distortion. It is well to choose a 6SJ7 circuit design which will yield a gain somewhat higher than that of the previous problem, in order to compensate for the loss due to this current feedback. This design choice is somewhat a matter of trial and error, and we may tentatively choose the circuit values shown in Fig. 10-20. Here A = 200, and maximum rms output voltage is 50. The harmonic distortion for maximum drive is 4.7 per cent.

R_L, the plate load for this stage, is R_b in parallel with R_g, and it has a value of 213,000 ohms. The current feedback β is R_2/R_L = 1200/213,000 = 0.0056. The gain $\mathbf{A}_1'$ is $\mathbf{A}_1/(1 - \beta\mathbf{A}_1) = -200/(1 + 0.0056 \times 200) = -200/2.12 = -94.3$. If this value of gain for the 6SJ7 stage is used, the feedback from the 6F6 plate may be considered in effect to be introduced into the 6SJ7 grid circuit in series with the signal voltage $\mathbf{V}_{1g}$.

The output impedance of the amplifier alone, as seen by the transformer primary, is the plate resistance of the 6F6, 80,000 ohms. The internal emf is $\mathbf{A}_{oc}\mathbf{E}_{1g} = \mathbf{A}_1'(-\mu_2\mathbf{E}_{1g}) = -94.3 \times (-200\mathbf{E}_{1g}) = +18{,}860\mathbf{E}_{1g}$. An equivalent circuit for the two-stage amplifier is shown in Fig. 10-21.

The distortion for the 6SJ7 with current feedback is $D/|1 - \beta\mathbf{A}_1| =$

4.7/2.12 = 2.2 per cent, if the output of this stage is made 50 volts rms. However, with the desired output of $16.5 \times 0.707 = 11.7$ volts rms, the distortion is roughly $2.2 \times 11.7/50 = 0.5$ per cent. (This assumes that the variation of distortion with drive for a pentode is approximately linear, the same as for a triode. This approximation is fairly reliable for high-plate-voltage operation. See Art. 7-16.) The distortion in the output without voltage feedback is approximately $8 + 0.5 = 8.5$ per cent. If we wish to reduce this to 2 per cent, then $|1 - \beta\mathbf{A}| = 8.5/2 = 4.25$ and $\beta\mathbf{A} = -3.25$. Here $\mathbf{A} = (18{,}860 \times 7000)/(80{,}000 + 7000) = 1520$, and hence $\beta = -3.25/1520 = -0.00214$. Then $\beta = -R_2/(R_1 + R_2) = -0.00214$, and $R_1 = 467R_2 = 467 \times 1200 = 560{,}000$ ohms.

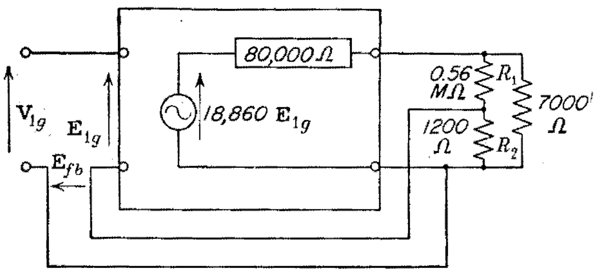


FIG. 10-21. An equivalent circuit for the amplifier of Fig. 10-20.

R_1 and R_2 constitute a negligible shunt on the 7000-ohm load. R_1 is so large that the assumption that R_2 is the feedback resistor for the current-feedback case is justified.

The required peak input voltage for the 6F6 is 16.5 as in the previous case, while the peak grid voltage for the 6SJ7 will be $16.5/200 = 0.083$ volt. Because of current and voltage feedback the input voltage V_{1gm} must be $0.083 \times 2.12 \times 4.25 = 0.75$, and $V_{1g} = 0.53$ volt rms.

In the case of the circuit of Fig. 10-19, the required rms input voltage was 0.5 volt, approximately the same as in this case. Hence both circuits perform somewhat the same, but only that of Fig. 10-20 successfully reduces the distortion.

We have introduced the feedback voltage into the cathode circuit of the 6SJ7 to compensate for phase reversal. It is often simpler to obtain the feedback voltage from a transformer secondary with the proper polarity to allow the feedback voltage to be introduced into the 6SJ7 grid circuit. This avoids the loss in gain due to current feedback. Furthermore, there is an additional advantage in that distortion caused by imperfect transformer performance is greatly reduced. In practice the phase shift caused by the transformer may result in a difficulty similar to that illustrated in Art. 10-12.

10-12. Oscillation in Feedback Amplifiers. In the examples of simple feedback circuits of the preceding article the amount of feedback varies somewhat with frequency because of series or shunt reactances in the circuits. In some cases this variation is so great that a circuit designed to give negative feedback at middle frequencies may give positive feedback in some other frequency range and may even break into oscillation.

In Art. 10-4 it was stated that $1 - \beta\mathbf{A}$ determines the character of the feedback. Hence it may be helpful if $\beta\mathbf{A}$ is plotted, say in polar coordinates. Figure 10-22*a* shows this plot for the simple amplifier of Fig. 10-6. Here $\beta = R_2/(R_1 + R_2)$, and as it is independent of frequency, the $\beta\mathbf{A}$ graph is similar to the $\mathbf{A}$ graph. $\mathbf{A}_{\text{mid}} = -g_m R_{sh}$, while

$$\mathbf{A}_{\text{low}} = \frac{\mathbf{A}_{\text{mid}}}{1 - j\dfrac{f_1}{f}} = \frac{\mathbf{A}_{\text{mid}}}{\sqrt{1 + \left(\dfrac{f_1}{f}\right)^2} \,\Big/ \arctan\left(-\dfrac{f_1}{f}\right)} = \frac{\mathbf{A}_{\text{mid}}}{\sqrt{1 + \tan^2\theta} \,\underline{/-\theta}}$$

$$= A_{\text{mid}} \cos\theta \,\underline{/180^\circ + \theta}$$

Here θ is the phase shift away from 180° and equals $\arctan(f_1/f)$. The locus of $\mathbf{A}_{\text{low}} = A_{\text{mid}} \cos\theta \,\underline{/180^\circ + \theta}$ is a semicircle in the third quadrant. Likewise the polar plot of $\mathbf{A}_{\text{high}} = \dfrac{\mathbf{A}_{\text{mid}}}{1 - j(f/f_2)}$ is also a semicircle, in the second quadrant. See also Arts. 7-12 and 7-13.

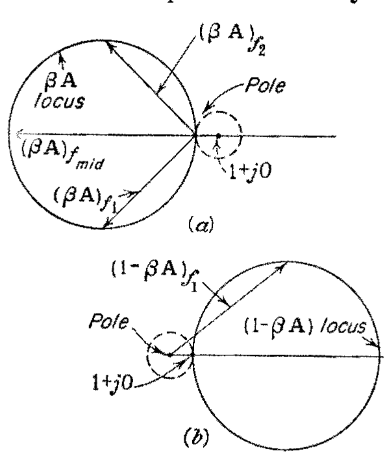


FIG. 10-22. Determination of the stability of the amplifier of Fig. 10-6. (*a*) The $\beta\mathbf{A}$ diagram. (*b*) The $1\text{-}\beta\mathbf{A}$ diagram.

Figure 10-22*b* shows a polar plot of $1 - \beta\mathbf{A}$ for the same amplifier. The point $1 + j0$ has been assumed to be in the position shown. Recalling that if $|1 - \beta\mathbf{A}| < 1$, the feedback is positive, it is seen that the vector $1 - \beta\mathbf{A}$ would need to terminate within the unit circle with center at $1 + j0$. If the $1 - \beta\mathbf{A}$ vector terminates at the pole, $\mathbf{A}'$ becomes infinite and oscillation occurs. In the case of this amplifier the vectors always terminate outside the unit circle except for zero frequency, and hence the feedback is negative for all other frequencies.

Some saving in time results if conclusions can be reached by drawing only a $\beta\mathbf{A}$ locus. Note that in Fig. 10-22*a* the feedback is always negative (except for zero frequency) because the $\beta\mathbf{A}$ vector never terminates within the unit circle with center at $1 + j0$. Termination at $1 + j0$ would result in oscillation.

Figure 10-23 shows the $\beta\mathbf{A}$ diagram for a single-stage amplifier with an interstage transformer in the feedback loop (see Art. 7-29). It is seen that in the h-f range the feedback becomes positive. The locus does not pass through $1 + j0$. Our discussion of the conditions for oscillation has not been complete as this is beyond the scope of this book. Nyquist[5] has shown, however, that if the $\beta \mathbf{A}$ locus encircles or passes through $1 + j0$, the circuit oscillates.

It may be readily perceived that in amplifiers with multistage feedback loops, the phase shift at high or at low frequencies may progress so rapidly that oscillation will occur in one of these bands. The design problem in such amplifiers is to reduce the gain rapidly enough so that the locus fails to encircle the $1 + j0$ point. For example, to design a three-stage feedback loop it is advisable to arrange two of the stages to have a much wider middle-frequency range than required for the working band. Then the phase shift for each of these will be low over this band. The third stage is then designed to have a sharp cutoff at the edges of the working band, and this reduces the over-all gain rapidly enough to satisfy the conditions for nonoscillation at any frequency.

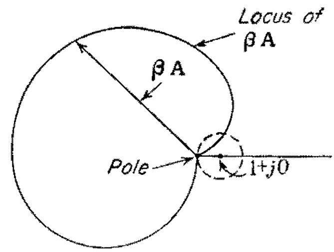


FIG. 10-23. The $\beta \mathbf{A}$ diagram for a transformer-coupled stage with feedback.

REFERENCES

1. Lawson, J. L., and G. E. Uhlenbeck: "Threshold Signals," McGraw-Hill Book Company, Inc., New York, 1950.
2. Goldman, S.: "Frequency Analysis, Modulation, and Noise," McGraw-Hill Book Company, Inc., New York, 1948.
3. Black, H. S.: Stabilized Feedback Amplifiers, *Elec. Eng.*, **53**, 114 (1934).
4. Bode, H. W.: "Network Analysis and Feedback Amplifier Design," D. Van Nostrand Company, Inc., New York, 1945.
5. Nyquist, H.: Regeneration Theory, *Bell System Tech. J.*, **11**, 126 (1932).

PROBLEMS AND QUESTIONS

1. An amplifier having a gain of $100\underline{/180^\circ}$ has an output of 100 volts at the signal frequency, together with 1 volt of noise. The noise originates at some point in the amplifier. If the signal-to-noise ratio is to be increased to 1000 without changing the magnitude of the signal output, discuss the ways in which this might be attempted. For one way, determine the fraction of the output voltage which must be fed back into the input, and also determine the required signal input voltage.

2. An amplifier having a gain of $100\underline{/180^\circ}$ has an output of 100 volts at the signal frequency, together with 5 per cent of second-harmonic distortion. Assume that the distortion voltage in the output varies as the square of the voltage on the first grid. (*a*) Draw box diagrams showing how the distorting voltage can be reduced to 1 per cent. (*b*) What input voltage is required and what fraction of the output voltage must be fed back if the output voltage remains constant at 100 volts?

3. An amplifier has a voltage amplification of -50 ± 4 per cent, the variation being due to changes in operating voltages. If negative feedback is added from output to input, with $\beta = 0.08$, what is the new voltage amplification? Determine its percentage variation. Obtain the answers by using Eq. (10-5) and also by drawing a box diagram in the manner of Prob. 2.

4. Draw a circuit with negative feedback and having a β network which causes the voltage output to rise at high frequencies.

5. If the output voltage of a feedback amplifier is maintained at the nonfeedback level by raising the gain of the amplifier in the feedback loop (see Fig. 10-5), prove the relationships expressed in Eqs. (10-12) and (10-13) and the conclusions concerning distortion and noise stated in Art. 10-6.

6. A single-stage amplifier has a gain $A = -6$ and an output of 120 volts with 6 per cent harmonic distortion. If negative voltage feedback is to be used to reduce the distortion to 2 per cent, what value of β should be used and what input voltage will be required if the output is to be maintained at 120 volts?

7. A one-stage amplifier with voltage feedback uses a tube with $\mu = 20$, $r_p = 12{,}000$ ohms. The load is $R_L = 100{,}000$ ohms. $\beta = 0.1\underline{/0°}$. What μ and r_p values would a tube need to have in a nonfeedback amplifier having the same R_L and the same gain?

8. Repeat Prob. 7 if the feedback amplifier has current feedback with $\beta = 0.1$ caused by a cathode resistor of 10,000 ohms.

9. What is the value of output impedance of the amplifier of Fig. 10-19, discussed in Example 3 of Art. 10-11, as seen when looking into the secondary of the output transformer (step-down ratio 10:1)?

10. Draw a nonfeedback equivalent circuit for the amplifier of Fig. 10-20, labeling all components. What is the output impedance as seen from the secondary of the output transformer (step-down ratio 10:1)?

11. See Fig. 10-24. For the tube $\mu = 20$, $r_p = 20{,}000$ ohms. $R_b = 100{,}000$ ohms, $R_g = 500{,}000$ ohms, tapped at 5000 ohms from the bottom end. C and C_k are very

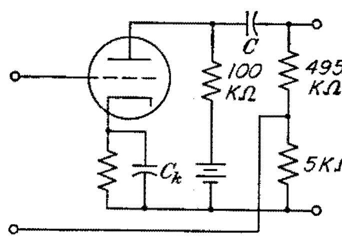


FIG. 10-24.

large. For middle frequencies determine (*a*) the gain of the amplifier, (*b*) its output impedance.

12. For the 6SJ7 design of Art. 7-16, determine the value of C_k needed to limit the degenerative loss in gain at f_1 to 1 per cent. An approximation can be made by assuming $(r_p + R_L)/(1 + \mu) \approx 1/g_m$. The value of g_m can be estimated from the formula $A = g_m R_{sh}$.

13. Remove the feedback connection from Fig. 10-24 (Prob. 11), and determine C to make the mid-frequency band extend down to 100 cps. Also determine the value of C_k needed to make the current feedback caused by Z_k ineffective at 100 cps. Assume $R_k = 1500$ ohms.

14. Redraw the circuit of Fig. 10-20 so that the feedback voltage is taken from the secondary of the transformer instead of the primary. Determine the correct transformer polarity. Determine the value of R_1 to be used to accomplish the results desired in Example 4 of Art. 10-11. What additional benefits are derived? What undesirable results, if any, might occur?

CHAPTER 11

OSCILLATORS

11-1. Types of Oscillators. In this chapter we shall study certain essentially electronic circuits which produce an a-c output, using only a d-c source of energy and no external a-c controlling voltage. Such devices are called oscillators. Being without external frequency control, an oscillator must have a self-contained frequency-determining portion. This portion takes various forms, the commonest being parallel L-C circuits and R-C networks. The latter are used in delay lines, in bridge circuits, and in tube circuits having two stable modes of operation. Some oscillators produce an output voltage which is essentially sinusoidal in form, others produce a saw-tooth wave, others a square wave, and so on. There are many kinds of oscillators, and a book could be written exclusively about them. We shall devote this chapter to selected topics considered to be of major interest. Ultrahigh-frequency oscillators are not treated, as the general techniques for such frequencies are outside the scope of this book.

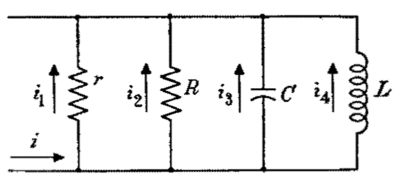


FIG. 11-1. An impulse is applied to an L-C tank.

11-2. The Parallel L-C Circuit as a Generator of Oscillations. Most of us have some acquaintance with a long pendulum having a heavy bob. Whether gently pushed to and fro in a class A manner, pushed from one side only in class B fashion, or just bumped at the right time in class C style, the pendulum bob swings back and forth in approximately simple-harmonic motion. Many of us are also aware of analogues between mechanical devices and electrical circuits, of the fact that a differential equation describing the action of a mass-spring-damper device is in the same form as the equation applying to some L-C-R electrical circuit.

With these things in mind, let us study the action of a coil in parallel with a capacitor. In Fig. 11-1 consider a tank circuit of L and C to have a loss resistance represented in parallel form by R. This loss may be due to energy dissipation in the inductor and in the capacitor or to energy delivered to a load. The parallel representation is only approximate, and an alternate method of handling the losses would be by a resistor in series with L. This representation being also approximate, we choose the more convenient form. Let us also assume an impulse of voltage of any waveform—pulse, sinusoidal, or otherwise—to be applied across the tank, and let r be the internal resistance of the impulse source. Figure 11-1 shows Norton's form of the circuit.

We may write the following circuit equations:

$$i_1 + i_2 + i_3 + i_4 = i \tag{11-1}$$

$$i_1 r = i_2 R = \frac{1}{C}\int i_3\,dt = L\frac{di_4}{dt} \tag{11-2}$$

From these equations we may solve for any of the currents i_1, i_2, i_3, or i_4 in terms of i and the circuit parameters. All the current solutions are of the same general form, and we may take that for i_4 as representative. Using Eqs. (11-2) to eliminate i_1, i_2, and i_3 from (11-1), we obtain

$$\frac{L}{r}\frac{di_4}{dt} + \frac{L}{R}\frac{di_4}{dt} + LC\frac{d^2 i_4}{dt^2} + i_4 = i \tag{11-3}$$

or

$$LC\frac{d^2 i_4}{dt^2} + L\frac{R + r}{Rr}\frac{di_4}{dt} + i_4 = i \tag{11-4}$$

The general solution of this second-order differential equation contains two arbitrary constants and is usually treated in two parts. One part, called the particular integral, depends upon the current i; it satisfies Eq. (11-4) and contains no arbitrary constants. If i is sinusoidal in waveform, there is no especial difficulty in finding the particular integral, for this is the value of i_4 obtained as the ordinary steady-state a-c solution of the circuit. If i is nonsinusoidal, there is considerably more difficulty. However, in the present case our immediate interest is in the second part of the general solution, called by mathematicians the complementary function—the part which contains the two arbitrary constants. The current which is the solution to this part has a waveform which is independent of that of the impulse i. We shall show that it is possible to obtain an approximate sinusoidal waveform for this current. This solution is obtained by solving Eq. (11-4) with i replaced by zero. Hence we wish to solve

$$LC\frac{d^2 i_4}{dt^2} + L\frac{R + r}{Rr}\frac{di_4}{dt} + i_4 = 0 \tag{11-5}$$

Substituting $i_4 = \epsilon^{mt}$ as a possible solution yields the requirement that

$$LCm^2 + L\frac{R + r}{Rr}m + 1 = 0 \tag{11-6}$$

or that

$$m = -\frac{L[(R+r)/Rr]}{2LC} \pm \sqrt{\frac{L^2[(R+r)/Rr]^2 - 4LC}{4L^2C^2}}$$
$$= -\frac{1}{2C[Rr/(R+r)]} \pm j\frac{1}{\sqrt{LC}}\sqrt{1 - \frac{L}{4C}\left(\frac{R+r}{Rr}\right)^2} \qquad (11\text{-}7)$$

There being two values of m which make ϵ^{mt} a solution, we may write these solutions as ϵ^{m_1t} and ϵ^{m_2t}. It is easily shown that the expression $i_4 = A\epsilon^{m_1t} + B\epsilon^{m_2t}$ also satisfies Eq. (11-5), and hence this is the complementary function. Substituting from (11-7) and factoring yields

$$i_4 = \exp\left\{-\frac{t}{2C[Rr/(R+r)]}\right\}(A\epsilon^{j\omega t} + B\epsilon^{-j\omega t}) \qquad (11\text{-}8)$$

where

$$\omega = \frac{1}{\sqrt{LC}}\sqrt{1 - \frac{L}{4C}\left(\frac{R+r}{Rr}\right)^2} \qquad (11\text{-}9)$$

ω is a real number if

$$\frac{L}{4C}\left(\frac{R+r}{Rr}\right)^2 < 1 \qquad (11\text{-}10)$$

i.e., if $\sqrt{\dfrac{C}{L}}\left|\dfrac{Rr}{R+r}\right| > \dfrac{1}{2}$, or $\dfrac{\left|\dfrac{Rr}{R+r}\right|}{\omega_0 L} > \dfrac{1}{2}$, or

$$|Q_{0_{eq}}| > \tfrac{1}{2} \qquad (11\text{-}11)$$

where $Q_{0_{eq}}$ is the equivalent circuit Q_0 [see Eq. (7-158)] and subscript 0 refers to conditions obtaining at the resonant frequency for which $\omega_0 L = 1/\omega_0 C$. Since

$$\epsilon^{jx} = \cos x + j \sin x \qquad (11\text{-}12)$$

we may change Eq. (11-8) to the form

$$i_4 = K\epsilon^{-\frac{t}{2C\frac{Rr}{R+r}}} \cos(\omega t + \theta) \qquad (11\text{-}13)$$

where K and θ are arbitrary constants used for convenience to replace combinations of the former ones, A and B. The expression $\cos(\omega t + \theta)$ is useful only if ω is real, which means if $|Q_{0_{eq}}| > \frac{1}{2}$. If this latter condition does not hold, the solution for i_4 reverts to the form (11-8) with ω either zero or a pure imaginary, making all of those exponentials real and the solution nonoscillatory. Now it is not at all difficult to make $|Q_{0_{eq}}|$ much greater than $\frac{1}{2}$, and if we do, Eq. (11-9) reduces to

$$\omega \approx \frac{1}{\sqrt{LC}} = \omega_0 \qquad (11\text{-}14)$$

and the inductor current i_4 is oscillatory with a frequency closely that of resonance. Likewise the capacitor current and the currents through R and r can be found to have the same waveform and frequency.

The amplitude of the oscillatory current i_4 is approximately

$$I_{4m} = K\epsilon^{-\frac{t}{2C\frac{Rr}{R+r}}} \tag{11-15}$$

This expression shows that if C, R, and r are positive quantities, the amplitude of the current decreases with time and eventually becomes inappreciable. If $2C[Rr/(R + r)]$ is large compared with T_0, the period of one cycle, the decay in current amplitude is slow. This condition is expressed by

$$\frac{2C\dfrac{Rr}{R+r}}{T_0} = \frac{2\omega_0 C\dfrac{Rr}{R+r}}{\omega_0(1/f_0)} = \frac{2\dfrac{Rr}{R+r}}{2\pi\omega_0 L} = \frac{2Q_{0_{eq}}}{2\pi} \tag{11-16}$$

In the study of tank circuits for class C amplifiers it is shown that $Q_0/2\pi$ is the ratio of total energy stored in the tank to that dissipated during one cycle. This means that $Q_{0_{eq}}$ should be large (the tank should exhibit a large flywheel effect) if good sinusoidal waveform of tank current is desired. Nevertheless, the oscillatory current in the tank dies out unless something is done to restore the lost energy. This requires that the impulse should be repeated with the correct magnitude and timing to offset the loss in energy.

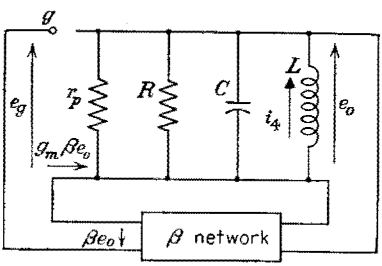


FIG. 11-2. The equivalent circuit for a regenerative-feedback oscillator.

There are two common ways in which this property of the L-C tank can be exploited to produce a good-quality sinusoidal output. The first is by the use of a regenerative feedback amplifier. The second is by the use of a negative-resistance device. We shall take up these methods in that order.

11-3. Regenerative Feedback Oscillators. Let us consider the equivalent circuit shown in Fig. 11-2 as one plan for accomplishing the purpose of supplying the tank with properly timed impulses of energy. This should be recognizable as a linear vacuum-tube amplifier in Norton's equivalent form and employing feedback. The current i is $g_m e_g = g_m \beta e_o$, where $\beta = e_g/e_o$. β is a positive or negative real number.

The differential equation to be solved for i_4 is adapted from (11-4),

$$LC\frac{d^2 i_4}{dt^2} + \left(\frac{L}{r_p} + \frac{L}{R}\right)\frac{di_4}{dt} + i_4 = g_m\beta e_o = -g_m\beta L\frac{di_4}{dt} \tag{11-17}$$

or

$$LC\frac{d^2i_4}{dt^2} + L\left(\frac{1}{r_p} + \frac{1}{R} + g_m\beta\right)\frac{di_4}{dt} + i_4 = 0 \tag{11-18}$$

which is the same as Eq. (11-5) if we substitute

$$\frac{1}{r} = \frac{1}{r_p} + g_m\beta \tag{11-19}$$

The complementary function is now the general or complete solution, since the right member of (11-18) is already zero. The solution for the equation is of the form of Eq. (11-13). Since β can be made either negative or positive, it is possible to make i_4 increase exponentially with time. This result occurs if $Rr/(R + r)$ [in Eq. (11-13)] becomes negative. Then $Rr/(R + r) < 0$ or $(R + r)/Rr < 0$. From this $(1/r) + (1/R) < 0$ or $(1/r_p) + g_m\beta + (1/R) < 0$, and hence $\beta < -1/g_m[Rr_p/(R + r_p)]$. Thus β is negative and $|\beta| > 1/g_m[Rr_p/(R + r_p)]$, or

$$|\beta| > \frac{1}{|A|} \tag{11-20}$$

Thus, if the circuit connections are made to reverse the phase between output and grid voltages to give positive feedback, and if the magnitude of β exceeds the reciprocal of the magnitude of the gain of the amplifier at the resonant frequency, the amplitude of i_4 will grow with time. There is nothing about the linear equivalent circuit to indicate that it will ever stop growing, but we know that it must eventually. We shall discuss reasons for this later.

Suppose the cathode of the vacuum tube to be heated and then the bias voltages applied. Either the sudden application of these voltages or any noise voltage in the tube or the circuit will supply an impulse to the plate tank. A small oscillatory current flows in the tank, and an output voltage appears across the tank inductor. A certain part of this voltage, with phase reversed, is applied to the grid. If the amount of voltage feedback and the amplification of the tube are sufficient to satisfy the relation (11-20), the energy in the tank is increased and the tank current increases in magnitude. Eventually, of course, the operation of the tube becomes nonlinear and the equivalent circuit becomes unreliable for studying its action. The current of sinusoidal waveform in the tank continues nevertheless, although harmonic currents, caused by harmonic components of i, may also flow. These, of course, may be considered, if desired, in the particular-integral part of the solution in this case. If the tank has a very low parallel impedance to harmonics of the resonant frequency, these harmonic tank currents are very small. As the tube and tank currents grow in magnitude, the average values of the tube parameters μ, r_p, and g_m change and finally we have equality between the members of (11-20). If $\boldsymbol{\beta} = \mathbf{E}_g/\mathbf{E}_o$, for steady-state operation it follows that

$$\boldsymbol{\beta} = \frac{1}{\mathbf{A}} \tag{11-21}$$

At this point the circuit produces oscillations of constant amplitude, and under favorable conditions the tank current is very nearly sinusoidal. The relation (11-21) for constant-amplitude oscillations is known as Barkhausen's criterion. It proves to be a very useful relation for studying feedback oscillators.

Actually we can see that positive feedback really produces the effect of making r in the circuit of Fig. 11-1 negative. In this sense, feedback oscillators are a form of so-called negative-resistance oscillators.

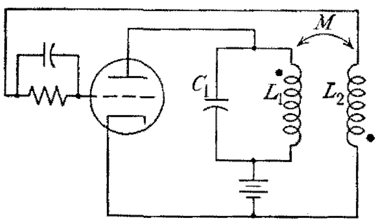


FIG. 11-3. A tuned-plate oscillator.

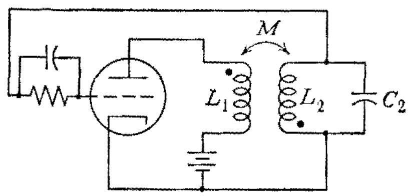


FIG. 11-4. A tuned-grid oscillator.

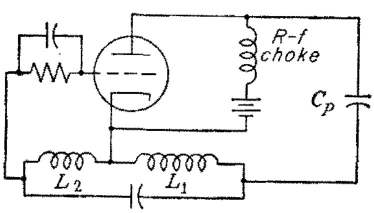


FIG. 11-5. A Hartley oscillator.

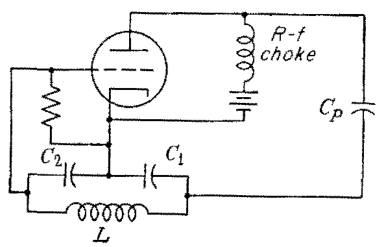


FIG. 11-6. A Colpitts oscillator.

11-4. Practical Feedback Oscillators. Numerous circuits have been devised which perform in the manner just described. Several of these are shown here.

In the tuned-plate oscillator of Fig. 11-3 the frequency-determining tank is in the plate circuit. The two dots on the coils show how the phase reversal to give positive feedback is accomplished. C_1 has a low impedance to harmonics of f_0, and hence the tank current has good sinusoidal waveform.

In the tuned-grid oscillator circuit of Fig. 11-4, the frequency of oscillation depends upon more than L_2 and C_2, since current flows in L_1 as well as in L_2. L_1 may have considerable impedance to the harmonics, and the waveform of the output voltage (across L_2) tends to be slightly inferior to that of the tuned-plate oscillator. The same effect occurs in the Hartley oscillator of Fig. 11-5, while the Colpitts of Fig. 11-6 again produces an excellent quality of waveform. However, the Hartley circuit is easy to adjust for satisfactory operation and is a very widely used device, while the two capacitors of the Colpitts circuit need simultaneous adjustment if the frequency of operation is to be changed very greatly.

Figure 11-7 shows an oscillator which uses two tuned circuits, one in the grid circuit and another in the plate circuit. The feedback in this case is through the grid-plate capacitance, and the frequency of oscillation must be slightly below that of resonance so that the plate load is inductive, making G_g (see Art. 7-5) negative. Energy is then supplied to the grid tank by the plate circuit, and a sinusoidal grid voltage results.

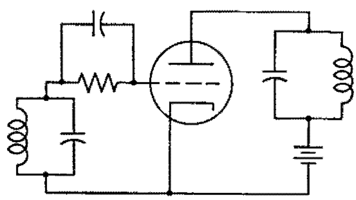


FIG. 11-7. A tuned-plate tuned-grid oscillator.

The grid tank of Fig. 11-7 may be replaced by a crystal (Fig. 11-8), the principle of operation remaining the same as before. The crystal has an equivalent circuit as shown in Fig. 11-9, where C_h is the capacitance of the crystal holder and R, L, and C are electrical equivalents of the mechanical crystal vibrational system. The crystal used is cut from a block of quartz to the form of a thin wafer, the planes of the cuts determining the vibrational frequency-temperature property. Quartz, like some other crystalline substances, possesses the piezoelectric property whereby compression along one axis produces a voltage between two opposite faces. The action is reversible so that application of a voltage between faces causes compression or expansion. The value of the equivalent Q_0 for the crystal

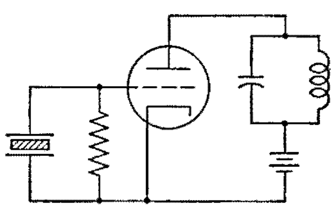


FIG. 11-8. A crystal oscillator.

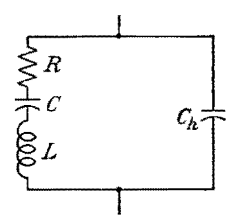


FIG. 11-9. The equivalent circuit for a crystal.

is very much higher than that for any actual coil, and with proper temperature control the frequency of oscillation can be held to a nearly constant value.

Note that some of these oscillator circuits use shunt plate feed, while others may employ series feed. In general shunt feed is preferable for any of the circuits as the total voltage across capacitors is less and the danger of d-c power arcs is lessened. This, of course, applies more to transmitting and high-power equipment, and for low-voltage apparatus it is often satisfactory to use series feed. The Colpitts circuit is one in which series feed cannot be employed.

11-5. The Analysis of a Tuned-plate Oscillator. Figure 11-3 makes it rather obvious that a tuned-plate oscillator is an amplifier with positive feedback from the output to the grid. It is entirely proper to write differential equations involving instantaneous voltages and currents in this circuit, as was done in Art. 11-3, and determine thereby the conditions for which oscillations occur. But it is considerably easier to assume that there will be oscillations of sinusoidal waveform and that these will be of constant amplitude if $\mathbf{A}\boldsymbol{\beta} = 1$, as expressed by Eq. (11-21) and also in Art. 10-4, where feedback amplifiers were being discussed. Another way of stating this Barkhausen criterion is to say that at any point in the circuit the loop gain must be unity. In the present case we shall use the Barkhausen method.

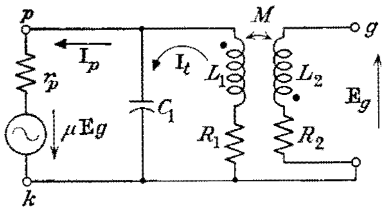


FIG. 11-10. Equivalent circuit for a tuned-plate oscillator.

Assume that the grid current is zero and that the operation is linear. Figure 11-10 shows an equivalent circuit in Thévenin's form. Since $\mathbf{I}_g = 0$,

$$\mathbf{E}_g = j\omega M\mathbf{I}_t \tag{11-22}$$

$$\mathbf{Z}_L = \frac{(R_1 + j\omega L_1)(1/j\omega C_1)}{R_1 + j[\omega L_1 - (1/\omega C_1)]} \tag{11-23}$$

The loop gain is unity if $\mathbf{E}_g$ reproduces itself.

$$\mathbf{E}_g = \frac{\mu\mathbf{E}_g \dfrac{(R_1 + j\omega L_1)(1/j\omega C_1)}{R_1 + j[\omega L_1 - (1/\omega C_1)]}}{r_p + \dfrac{(R_1 + j\omega L_1)(1/j\omega C_1)}{R_1 + j[\omega L_1 - (1/\omega C_1)]}} \frac{1}{R_1 + j\omega L_1} j\omega M \tag{11-24}$$

or

$$R_1 r_p + jr_p\left(\omega L_1 - \frac{1}{\omega C_1}\right) + (R_1 + j\omega L_1)\frac{1}{j\omega C_1} = \frac{\mu M}{C_1} \tag{11-25}$$

Since real numbers cannot equal imaginary numbers,

$$r_p L_1 - \frac{r_p}{\omega C_1} - \frac{R_1}{\omega C_1} = 0 \tag{11-26}$$

and

$$R_1 r_p + \frac{L_1}{C_1} = \frac{\mu M}{C_1} \tag{11-27}$$

From Eq. (11-26) it follows that

$$\omega = \frac{1}{\sqrt{L_1 C_1}}\sqrt{1 + \frac{R_1}{r_p}} \tag{11-28}$$

or

$$f = f_0 \sqrt{1 + \frac{R_1}{r_p}} \tag{11-29}$$

where

$$f_0 = \frac{1}{2\pi \sqrt{L_1 C_1}} \tag{11-30}$$

is the resonant frequency of the plate tank. Thus it is seen that the frequency of oscillation is near the resonant frequency if R_1 is much smaller than the plate resistance of the tube; this is the usual case.

Equation (11-27) may be changed to the form

$$g_m - \frac{L_1}{Mr_p} = \frac{R_1 C_1}{M} \tag{11-31}$$

By a little manipulation a graphical meaning to the expression $g_m - (L_1/Mr_p)$ can be obtained. Thus

$$g_m - \frac{L_1}{Mr_p} = \frac{\mu \mathbf{E}_g - \frac{L_1}{M}\mathbf{E}_g}{r_p \mathbf{E}_g} = \frac{\mu \mathbf{E}_g - \frac{L_1}{M} j\omega M \mathbf{I}_t}{r_p \mathbf{E}_g} = \frac{\mu \mathbf{E}_g - j\omega L_1 \mathbf{I}_t}{r_p \mathbf{E}_g}$$

$$= \frac{\mathbf{I}_p r_p}{r_p \mathbf{E}_g} = \frac{\mathbf{I}_p}{\mathbf{E}_g} \tag{11-32}$$

$\mathbf{I}_p/\mathbf{E}_g$ is the approximate slope of the secant line connecting the end points of the path of operation on i_b, e_c axes, and by Eq. (11-31) this equals R_1C_1/M. Thus, for the case of a tuned-plate oscillator having a fixed operating point, oscillations will begin if the slope of the tangent line at Q exceeds R_1C_1/M (see Fig. 11-11). The amplitude of oscillation will then increase until the slope of the secant line equals R_1C_1/M. This automatic amplitude limitation is caused by saturation effects, and it is readily seen that if good plate-current waveform is desired, the controllable circuit parameter (say M) should be adjusted to make the members of Eq. (11-31) almost equal. The steady-state oscillations will then be of small amplitude.

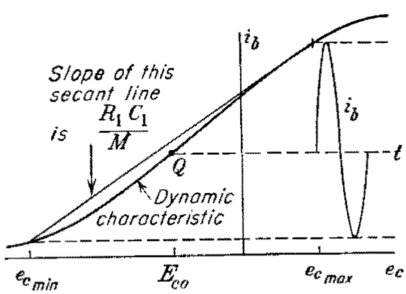


FIG. 11-11. Limitation of amplitude of oscillation by change in tube parameters.

Figure 11-12 shows vector diagrams for a tuned-plate oscillator. (*a*) makes it clear that the frequency of oscillation cannot be the antiresonant frequency of the plate tank, since $\mathbf{E}_g$ does not reproduce itself in phase. (*b*) shows vector relations at the true oscillating frequency, which is somewhat higher than the antiresonant frequency and where the tank is slightly capacitive. The exact positions for $\mathbf{I}_p$ and the dependent vectors depend upon r_p of the tube, and thus, if the tube operating voltages are allowed to vary, the frequency of oscillation changes to make adjustments in the vector positions. Some measure of stability can be attained by making r_p high compared with the resistance R_1. It is also beneficial to insert a capacitor in the plate lead as this makes $\mathbf{I}_p$ lead $\mu\mathbf{E}_g$ without the necessity of the plate tank acting capacitively.

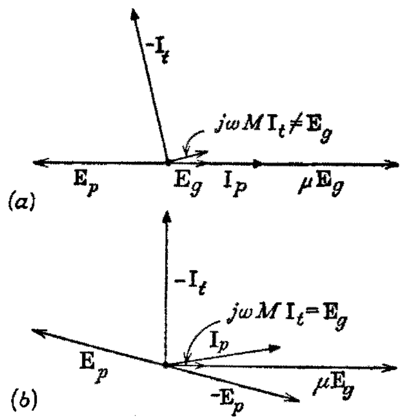


FIG. 11-12. Vector diagrams for a tuned-plate oscillator. (a) At the antiresonant frequency. (b) At the frequency of oscillation.

11-6. Power Oscillators and Voltage Oscillators. Feedback oscillators, like amplifiers, may be designed to give considerable power output directly to a load or to furnish only an alternating voltage. Power oscillators are usually operated class C_2 in order to obtain high plate-circuit efficiency and depend upon the filtering action of the tank to lower the harmonic content of the output. The design of these oscillators is essentially the same as that of amplifiers of the same class, the graphical approach being the most successful. Using constant-current characteristics for the tube, a Q point (E_{bo},E_{co}) and a P point $(e_{b_{\min}},e_{c_{\max}})$ are tentatively chosen and the straight-line path of operation drawn through them. Values obtained at selected points on the path of operation enable one to calculate I_{ba}, I_{ca}, E_{p_1}, I_{g_1}, and other quantities desired. Determination of P_{ac} and η_p and a check of P_p against the allowable value complete the test of the tentative design. This can be altered, if desired, in an attempt to achieve better results. Once this design is decided upon, a feedback circuit is developed to produce the desired grid voltage from the computed output voltage. The required load is E_{p_1}/I_{p_1}, and $R = Q_0\omega_0 L$ (for the tuned-plate case) must have this value.

The grid-bias resistor is usually a grid leak, and its value can be computed from the value of I_{ca}. The value of C_g should be great enough so that the bias will not decrease appreciably during the period of no grid-current flow. A value of R_gC_g equal to $3T$, where T is the period of the oscillatory voltage, gives a generally satisfactory value for C_g. A greatly too high value for R_gC_g should be avoided since intermittent oscillations may result. Normally, if the voltage output of the oscillator decreases, the grid bias quickly decreases, which causes the output to increase. If

R_gC_g is too high, the bias does not decrease quickly enough and the voltage output continues to decrease to zero. After C_g discharges through R_g, oscillations again begin and the action repeats.

For constancy of both frequency and amplitude of oscillations, the tube parameters must be maintained constant, and thus plate and screen bias voltages should not be allowed to vary. Also, connecting the load directly to the oscillator output has the disadvantage of causing the frequency of oscillation to change with changes in load impedance. It is obvious that any varying reactance in the load causes a frequency change. So too does a varying load resistance affect the frequency, as shown by Eq. (11-9), and also since the effect of the tube capacitances depends upon the load to some extent (Art. 7-5). Hence, if extreme frequency stability is desirable, as in radio transmitters, it is essential to isolate the load from the oscillator. This is usually accomplished by using a class A or a class C oscillator to drive a class 1 power amplifier, which in turn supplies the load, directly or by driving a more powerful class C_2 amplifier. There being no grid current for the class 1 stage, the latter acts as a buffer to prevent the undesirable reaction of a load on the oscillator.

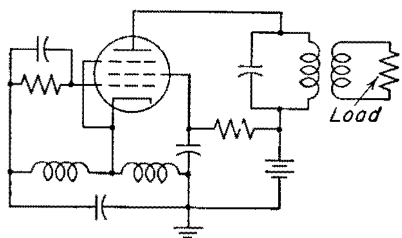


FIG. 11-13. An electron-coupled oscillator.

The same result is also achieved by a so-called electron-coupled oscillator, one circuit for which is shown in Fig. 11-13. The cathode, No. 1 grid, and No. 2 grid are associated with a Hartley oscillator, the No. 2 grid (the oscillator anode) being grounded for alternating voltages. Electrons which pass through the mesh of the No. 2 grid advance in a stream of varying density toward the suppressor grid just as though the cathode temperature fluctuated at a rate corresponding to the frequency of oscillation. To the suppressor grid and the plate, the oscillator portion appears to be a virtual cathode. The plate current and the screen current have the same waveshape and frequency. Any change in load in the plate circuit has little effect on the operation of the oscillator because of the shielding effect of the suppressor and the No. 2 grids.

Although fixed bias can be used for class A operation, it is unsuitable for class C since the oscillator will not be self-starting. In any case the use of a grid leak and capacitor is recommended. A reason is that additional amplitude control is obtained over that resulting from the saturation effect of Art. 11-5. Figure 11-14 shows the operation of a tuned-plate class C oscillator using grid-leak bias. Before oscillations start, a small amount of grid current flows through the grid leak to make the grid just slightly negative. At this Q point if Barkhausen's condition for increasing oscillation is satisfied, the alternating grid voltage rapidly increases in amplitude, ultimately reaching a value limited by a combined effect of change of tube parameters and the movement of the operating point to a more negative value because of greater average grid current.

11-7. Negative-resistance Oscillators. It will be recalled that the peak amplitude of the inductor current i_4 [expressed by Eq. (11-15)] can be made to increase if r in $2C[Rr/(R + r)]$ can be made negative. We have found that this can be accomplished by feeding back voltage from the output of the tank to the grid of a vacuum tube, which in turn feeds power into the tank. Thus in effect the tank is connected to a four-terminal network which acts in the required manner.

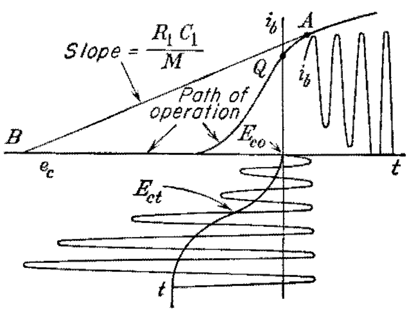


FIG. 11-14. Amplitude control in a class C oscillator with grid-leak bias.

It is possible to find two-terminal devices which also display a negative-resistance property. Briefly this means that the instantaneous alternating current flowing through the device is in a direction contrary to that of the voltage drop across it more than half the time, or that this device serves as a generator (see Art. 6-10). Figure 11-15 gives an equivalent circuit for an oscillator employing such a device. If the amount of the negative resistance changes with the amplitude of the voltage applied to the device, it is possible to have automatic voltage amplitude control similar to that of Art. 11-5. The equilibrium condition is reached when the average value of $|r|$ reaches the value R. Then $R(-R)/(-R + R) = \infty$, and i_4 in Eq. (11-13) has constant peak amplitude.

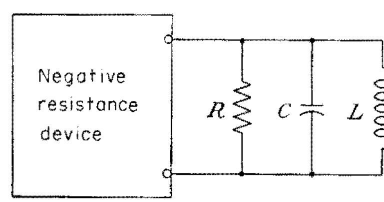


FIG. 11-15. Equivalent circuit for a negative-resistance oscillator.

One practical circuit of this type is the dynatron oscillator diagramed in Fig. 11-16. Figure 11-17 shows the property of negative plate resistance possessed by a tetrode in a partial region where $e_b < e_{c2}$ (see Art. 4-10). With Q placed in the negative-plate-resistance region, note that the waveforms of e_p and i_p are 180° out of phase; this would not be the case if the static operating point had been chosen at point Q', for instance.

With $R > |r_p|$ and with load line A as shown in the figure, the amplitude of oscillations increases until the average value of $|r_p|$ equals R. By careful adjustment of the tube biases, the value of $|r_p|$ at Q can be made very near to the value of R; the oscillations are then of small amplitude and of excellent sinusoidal waveform. For higher load impedances, as represented by line B, it can be shown that oscillations are still possible; however, since they have a nonsinusoidal waveform, this is usually not a desirable operating condition.

The dynatron oscillator is not too often used in practice, as tube aging changes the characteristics seriously and causes inconstancy of operation. Other circuits not having this drawback are possible, among them one utilizing the screen-plate transconductance $\partial i_b/\partial e_{c2}$ of a pentode, which is negative under some circumstances.

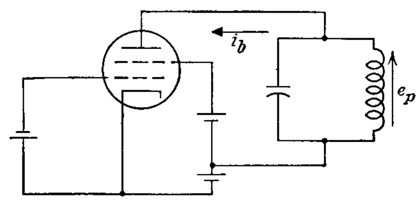


FIG. 11-16. The dynatron oscillator.

11-8. L-F Oscillators. Tuned-load oscillators are best adapted to moderately high-frequency use where the size of the inductor needed is not excessive. For low frequencies it is possible to use two h-f oscillators operating at somewhat different frequencies and feeding into a nonlinear amplifier, thus producing sum and difference frequency as well as harmonic components. A filter is used to eliminate all components except the difference one, which has a low frequency. This device is called a beat-frequency oscillator. A block diagram for one is shown in Fig.

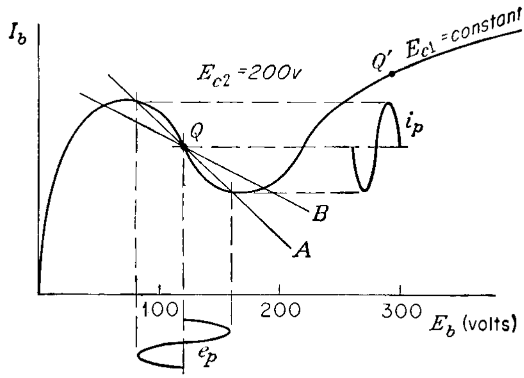


FIG. 11-17. The tetrode can be a negative-resistance device.

11-18. Another successful device is one which employs a phase-shift network to produce positive feedback. And still another uses a Wien-bridge arrangement. Both these latter devices are useful over the ordinary operating range of R-C-coupled amplifiers.

11-9. Phase-shift Oscillators. Figure 11-19 shows a prototype R-C phase-shift oscillator. Each C and R combination shifts the angle of the voltage somewhat less than 90° so that a minimum of three shift sections is required to furnish a total shift, including that of the tube, equal to 360°.

If we assume linear operation, we may use Norton's equivalent form shown in Fig. 11-20. R_1 represents the parallel combination of R_b and r_p. In order to lessen the shunting effect of each section on the resistor before

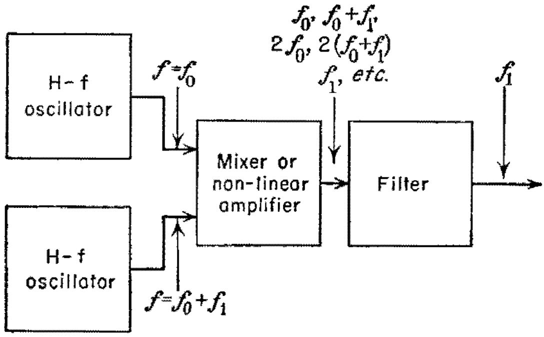


FIG. 11-18. Block diagram for a beat-frequency oscillator.

it, we shall increase the impedances progressively by a factor n as shown. Note that the capacitors are progressively decreased in size. The phase shifts of all sections are therefore approximately the same.

By assuming $\mathbf{E}_g = 1/\underline{0^\circ}$ volt, one can compute the voltages and currents back through the network and set the input current equal to $-g_m\mathbf{E}_g$.

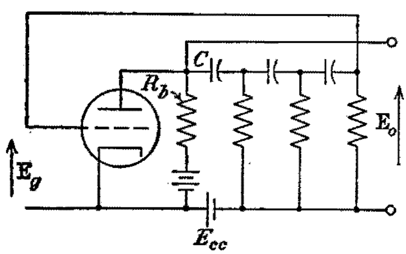


FIG. 11-19. A phase-shift oscillator.

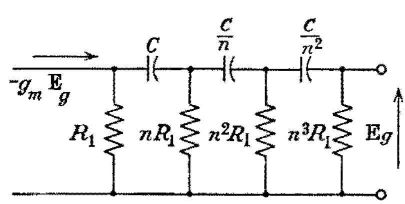


FIG. 11-20. Norton's equivalent of the phase-shift oscillator.

This procedure yields a relationship from which the frequency of oscillation is determined as

$$\omega = \frac{1}{\sqrt{3n^2 + 4n + 3}\, R_1 C} \tag{11-33}$$

and for oscillations to start, the required value of g_m for the tube is

$$g_m = \frac{8n^3 + 20n^2 + 20n + 8}{n^3 R_1} \tag{11-34}$$

The higher the value of n, the lower the value of g_m needed. However, there is little advantage in increasing n beyond 3 or 4 because there is some chance that the resistor in the shift section next to the grid may become excessively large, for this is the grid leak for the tube.

This analysis for the case of three shift sections supposes linear operation but serves as an approximate explanation of the oscillator performance. As oscillations begin, g_m is higher than the value given by (11-34); as the oscillations increase in amplitude, the average value of g_m decreases because of nonlinearity of the characteristics. With nonlinearity come harmonics in the plate current, but since the harmonics have a total phase shift of less than 360°, they are not built up.

11-10. Wien-bridge Oscillators. Figure 11-21 shows a simplified circuit for an oscillator of this type. It employs an ordinary two-stage amplifier and a bridge type of circuit to give approximately constant negative feedback and also positive feedback, which is sensitive to frequency changes. The R-C circuit elements Z_3 and Z_4 play an important part in the positive-feedback action. If the R's are equal and the C's are equal, it may be easily verified that the ratio $\mathbf{Z}_4/(\mathbf{Z}_3 + \mathbf{Z}_4) = \frac{1}{3}\underline{/0°}$ at the frequency for which $\omega = 1/RC$, while for other frequencies the ratio is less and has a positive or a negative angle.

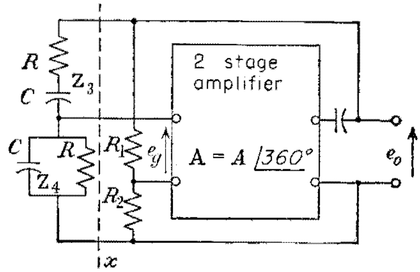


FIG. 11-21. Schematic diagram of a Wien-bridge oscillator.

We have seen in the analyses of other oscillators that sometimes the transient method is employed, involving the solution of differential equations, and sometimes a steady-state analysis is used, involving complex-number algebra—a considerably simpler process. The first method does not assume that the circuit will oscillate or that any oscillations will be sinusoidal. The second method assumes that for certain conditions there will be sinusoidal oscillations. In this case we shall employ the second method.

In order to oscillate, the circuit must have a loop gain of unity. Thus if we start with the output voltage $\mathbf{E}_o$, the gain of the loop must reproduce $\mathbf{E}_o$. Hence

$$\left(\frac{\dfrac{R/j\omega C}{R + 1/j\omega C}}{R + \dfrac{1}{j\omega C} + \dfrac{R/j\omega C}{R + 1/j\omega C}} \mathbf{E}_o - \frac{R_2}{R_1 + R_2} \mathbf{E}_o \right) \mathbf{A} = \mathbf{E}_o \quad (11\text{-}35)$$

where $\mathbf{A}$ will be taken to be $A\underline{/360°}$. Equation (11-35) then reduces to

$$(R_1 + R_2 + AR_2)\omega^2 R^2 C^2 - j(3R_1 + 3R_2 - AR_1 + 2AR_2)\omega RC - (R_1 + R_2 + AR_2) = 0 \quad (11\text{-}36)$$

Since the terms with the j coefficient must equal zero, it follows that

$$3R_1 + 3R_2 - AR_1 + 2AR_2 = 0 \tag{11-37}$$

or

$$R_1 = \frac{2A + 3}{A - 3} R_2 \tag{11-38}$$

The real part of Eq. (11-36) must also equal zero, and hence

$$(R_1 + R_2 + AR_2)(\omega^2 R^2 C^2 - 1) = 0 \tag{11-39}$$

from which we can obtain

$$\omega = \frac{1}{RC} \tag{11-40}$$

which gives the frequency of the oscillation.

Thus the output voltage is sinusoidal and of constant magnitude if Eq. (11-38) holds. Let us examine the negative-feedback amplifier circuit to the right of x in Fig. 11-21. The value of β is approximately $-R_2/(R_1 + R_2)$, and we may solve for the gain with feedback as

$$\mathbf{A}' = \frac{\mathbf{A}}{1 - \beta\mathbf{A}} = \frac{\mathbf{A}}{1 + [R_2/(R_1 + R_2)]\mathbf{A}} \tag{11-41}$$

which, upon application of the relation of (11-38), becomes

$$\mathbf{A}' = \frac{\mathbf{A}}{1 + \dfrac{\mathbf{A}}{[(2\mathbf{A} + 3)/(\mathbf{A} - 3)] + 1}} = 3\underline{/360^\circ} \tag{11-42}$$

for any value of **A**.

The gain of the whole circuit, including the positive-feedback loop for which $\beta = \mathbf{Z}_4/(\mathbf{Z}_3 + \mathbf{Z}_4)$, is

$$\mathbf{A}'' = \frac{\mathbf{A}'}{1 - \beta\mathbf{A}'} = \frac{3}{1 - [\mathbf{Z}_4/(\mathbf{Z}_3 + \mathbf{Z}_4)]3} \tag{11-43}$$

At the frequency $\omega = 1/RC$ this reduces to

$$\mathbf{A}'' = \frac{3}{1 - \frac{1}{3} \times 3} = \infty \tag{11-44}$$

and the circuit develops an output $\mathbf{E}_o$ with no input. At any other frequency, β for the positive-feedback circuit fails to meet these requirements.

Actually, if **A** does not have a phase shift of exactly 360°, the frequency of oscillation changes sufficiently to give positive-feedback β an angular value to compensate.

If the value of A is very high, the relation of (11-38) makes R_1 somewhat more than twice R_2. It is usual to use an incandescent lamp for R_2* and to adjust R_1 to a value which just permits oscillations to occur. The lamp acts as a stabilizer of oscillation amplitude since any increase in E_{om} results in a rise in lamp resistance, causing greater negative feedback and a reduction in E_{om}. Thus stabilization by saturation is not needed, and operation can be made linear class A, with the result that excellent waveform of output is obtained.

11-11. The Multivibrator.[1,3,5] A two-stage R-C-coupled amplifier with the output connected directly to the input (Fig. 11-22) will serve as a generator of oscillations. This is rather to be expected because the phase shift for proper positive feedback is provided and the capacitors serve to store energy. However, the character of the voltage waveforms differs from that found with previously studied circuits.

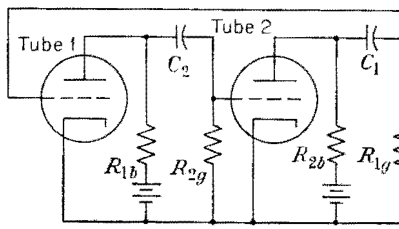


FIG. 11-22. A multivibrator circuit.

The multivibrator works on the principle that the circuit has two stable modes of operation, *viz.*, the No. 1 tube conducts, and the No. 2 tube does not, or vice versa. To show that this is the case, let us suppose that initially both tubes are conducting but that a noise voltage in the circuit causes a momentary small increase in the grid voltage of tube 2. Plate 2 voltage drops slightly, and with it grid 1 voltage falls, causing plate 1 voltage to rise. This causes grid 2 voltage to rise, further accentuating the original disturbance. The action is cumulative, and in a moment tube 2 is in full conduction, and tube 1 is cut off. Following this, if by some means grid 1 voltage can be made to rise above the cutoff value, tube 1 will conduct, while the other tube becomes nonconducting. If this action can be made automatic, an oscillator results. This is where the capacitors for energy storage come in. Although a coupling capacitor acts as a short circuit to a sharp rise in voltage, and this is because there is insufficient time allowed for the capacitor to gain a charge, after a brief interval the capacitor does charge and the shift in voltages necessary for switchover occurs. The rate of switching recurrence depends mostly upon the sizes of C and R_g used.

To make a more thorough study of this interesting circuit, we shall analyze the action of a specific case. It is often convenient to draw the multivibrator circuit in the manner shown in Fig. 11-23. The tube is a

* Actually, R_2 consists of the plate resistance and the plate load for the first tube as seen from the tube's cathode, that is, $(r_p + R_L)/(1 + \mu)$, in parallel with the incandescent lamp.

type 6SN7 double triode. Its plate and grid characteristics are shown in Fig. 11-24. Assumed component values are shown in the circuit drawing.

Let us suppose that the time t_1 approaches at which tube 2 goes into full conduction and tube 1 cuts off. C_1 is fully charged so that grid potential $e_{1c} = 0$. The current flowing through R_{1b} and R_{2g} to discharge C_2 is small compared with i_{1b}, since R_{2g} is very large. Hence we may draw a 40,000-ohm load line on the plate characteristics and find that $e_{1b} \approx 65$ volts and $i_{1b} \approx 5.9$ ma. The potential e_{2c} of g_2 is rising toward the cutoff value of -18 volts as capacitor C_2 discharges.

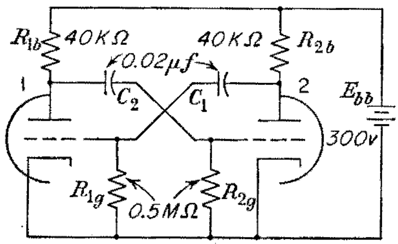


FIG. 11-23. An alternate representation for the multivibrator.

Time t_1 arrives. Figure 11-25 applies to events which immediately follow. The capacitor C_2 voltage is $65 + 18 = 83$ volts, with polarity as indicated. Tube 2 passes current, and tube 1 cuts off instantaneously. With the switch-over, e_{2c} suddenly becomes positive (as will be shown

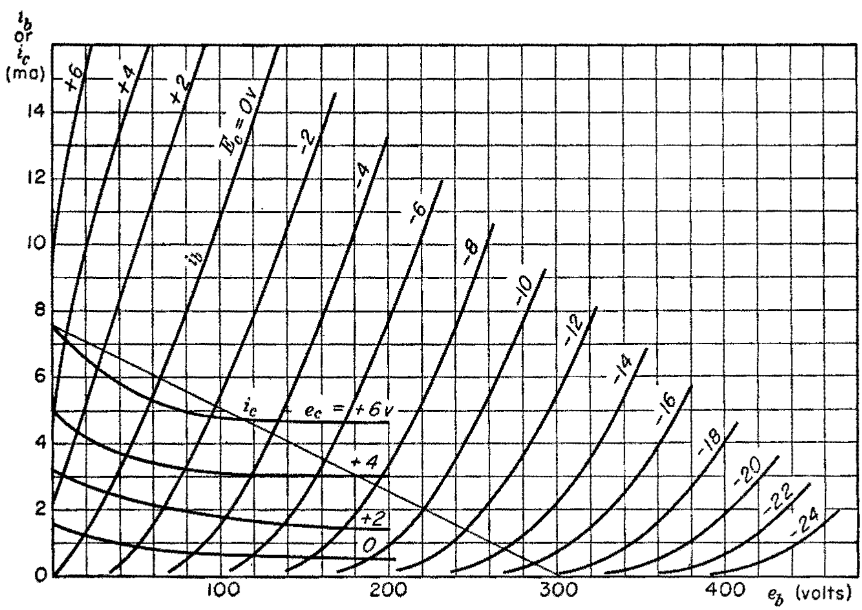


FIG. 11-24. Characteristics of a 6SN7 (one section).

later) so that the grid-cathode resistance R_{2c} shunts R_{2g}. We shall estimate (subject to later verification) the value of R_{2c} to be about 1000 ohms. Thus the current through R_{1b} is reduced immediately from 5.9 ma to $(300 - 83)/41{,}000 = 5.3$ ma. This makes the initial value of $e_{2c} \approx$

0.0053 × 1000 = +5.3 volts. For this value of grid voltage the plate voltage e_{2b} is reduced almost to zero, and $i_{2c} \approx 6$ ma. Thus R_{2c} is somewhat less than the 1000 ohms assumed, but close enough to suit our purpose.

At time t_1 when e_{2c} jumped from −18 to +5.3 volts, the plate voltage e_{1b} increased 23.3 volts also, taking a new value of about 65 + 23 = 88 volts.

Immediately following time t_1 the grid voltage e_{2c} begins to fall, the time constant (Fig. 11-25) being approximately $R_{1b}C_2 = 800$ μsec. Soon e_{2c} has a value of zero. Using the same time constant, the plate voltage e_{1b} rises from 88 volts to 300 volts as C_2 charges.

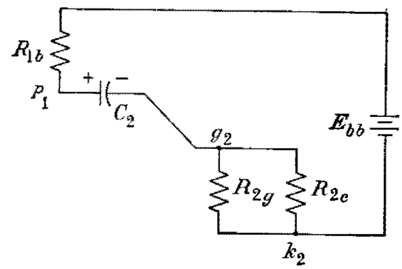


FIG. 11-25. Circuit applying to events soon after tube 1 cuts off.

Figure 11-26 shows the circuit which controls the events leading to the reestablishment of plate current in tube 1 at time t_2. For some time before t_2 the value of e_{2c} is practically zero so that we may determine the value of R_{2p} from Fig. 11-24 as approximately 11,000 ohms. The current to C_1 being small compared with i_{2b}, e_{2b} remains constant at 65 volts, the value to which it rose (with a time constant of only 800 μsec) after falling from +300 to almost zero after tube 2 began to conduct. Although this sudden drop of 300 volts carried e_{1c} down the same amount, the value of the latter rose to −235 volts shortly after time t_1 (using the same short time constant). Since that time C_1 has been discharging with a time constant of approximately $R_{1g}C_1 = 0.01$ sec (see Fig. 11-26), which is relatively long. As C_1 discharges, e_{1c} rises from its initial value E_{1c0} following the equation

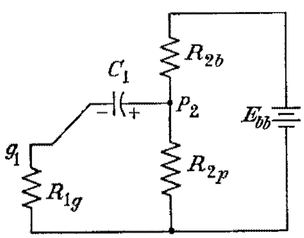


FIG. 11-26. Circuit applying to events just before tube 1 conducts.

$$e_{1c} = E_{1c0}\epsilon^{-t/R_{1g}C_1} \approx -235\epsilon^{-100t} \quad \text{volts}$$

The rate of rise is $de_{1c}/dt = -100 \times -235\epsilon^{-100t} = -100e_{1c}$ volts per sec. When e_{1c} reaches a value of −18 volts, tube 1 begins to conduct, while tube 2 is already conducting. Thereupon the rate of change of e_{1c} becomes dependent both upon the rate of capacitor C_1 discharge and upon the amplification of the two-tube-circuit loop. Hence $de_{1c}/dt = -100e_{1c} + (-g_{1m}R_{1sh})(-g_{2m}R_{2sh})(de_{1c}/dt)$, or $de_{1c}/dt = -100e_{1c}/[1 - (g_{1m}g_{2m}R_{1sh}R_{2sh})]$. When the conduction in tube 1 is sufficient so that g_{1m} makes the loop gain $g_{1m}g_{2m}R_{1sh}R_{2sh}$ equal to unity (and that is almost immediately), de_{1c}/dt becomes infinite and conduction snaps from tube 2 to tube 1. This explains the instantaneous changes in e_c and e_b which take place.

A similar study can be made to determine the changes in e_{2b} and e_{2c}. Figure 11-27 shows waveforms of plate and grid voltages for both tubes. The time between t_1 and t_2 can be computed by substituting -18 volts into the equation for e_{1c}. Thus $-18 = -235\epsilon^{-100(t_2-t_1)}$, and $t_2 - t_1 = \frac{1}{100} \ln \frac{235}{18} = 0.0257$ sec. With this symmetrical circuit, t_3 occurs 0.0257 sec after t_2. It should be noted that the oscillator period can be adjusted by varying the values of R_g and C.

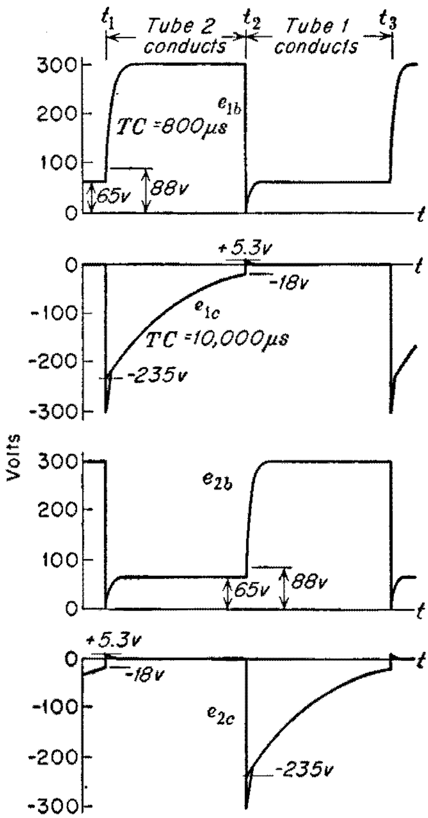


FIG. 11-27. Waveforms of plate and grid voltages for a symmetrical multivibrator.

Multivibrators are important for many reasons. The plate voltage of either tube is essentially a square wave, which may be useful as such, or because of the many harmonics which it contains. It is possible to feed an external voltage into the circuit and have the multivibrator operation synchronize either at the frequency of the external voltage or at a subharmonic of it. This principle is used in some types of electronic switches.

REFERENCES

1. Cruft Laboratory, War Training Staff: "Electronic Circuits and Tubes," McGraw-Hill Book Company, Inc., New York, 1947.
2. Eastman, A. V.: "Fundamentals of Vacuum Tubes," 3d ed., McGraw-Hill Book Company, Inc., New York, 1949.
3. Radar School Staff, MIT: "Principles of Radar," 2d ed., McGraw-Hill Book Company, Inc., New York, 1946.
4. Reich, H. J.: "Theory and Applications of Electron Tubes," 2d ed., McGraw-Hill Book Company, Inc., New York, 1944.
5. Seely, S.: "Electron-tube Circuits," McGraw-Hill Book Company, Inc., New York, 1950.

PROBLEMS

1. In Fig. 11-5 assume the effects of C_p, the r-f choke, and the coil resistances are negligible. Assume linear class 1 operation, and apply the Barkhausen criterion for steady oscillations. Determine the frequency of oscillation and the relation among the circuit parameters to make oscillation possible.

2. Repeat the procedure of Prob. 1, using the Colpitts circuit of Fig. 11-6.

3. Assume the crystal in Fig. 11-8 has $L = 35$ henrys, $C = 0.02\ \mu\mu f$, and $R = 3500$ ohms. The capacitance of the holder is 5 $\mu\mu f$, and the input capacitance of the tube is 3 $\mu\mu f$. What is the resonant frequency of the grid circuit? If the input capacitance of the tube changes to 5 $\mu\mu f$ because of a change in the plate load, what change in resonant frequency occurs? Determine the Q_0 of the crystal. *Note:* In determining the resonant frequency use the loop of Fig. 11-9, including C_g.

4. Figure 11-28 (redrawn from Fig. 11-5) shows the circuit for a Hartley power oscillator used to supply power to an r-f induction heater. It can be looked upon as a class C amplifier with a plate load of L_1 and R in parallel with the portion of C necessary to tune it to resonance. The tube is an Eimac type 100TH. $f_0 = 300$ kc. (*a*) Assume $E_{bb} = 3000$ volts and the desired load power to be 360 watts. The tank

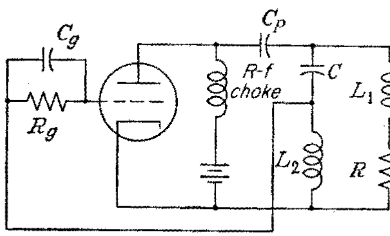


FIG. 11-28. A Hartley oscillator.

and coupling-circuit efficiency is 95 per cent, and the power amplification is approximately 25. Estimate values of P_{in}, P_{ac}, P_p, P_{bb}, η_p (see Art. 9-5). (*b*) Use constant-current curves for the tube and determine E_{cc}, I_{ba}, I_{ca}, I_{p_1m}, and I_{g_1m}. (*c*) Compute the values of P_{ac}, P_p, η_p, E_{gm}, E_{pm}, P_{in}, R_{f_0}. (*d*) Assume C_p and the r-f choke have negligible effect on the a-c performance. Assume $Q_L = 12$, and compute the value of $\omega_0 L_1$. (*e*) L_1 and L_2 constitute a voltage divider. Determine $\omega_0 L_2$. (*f*) Determine X_C and C. (*g*) Determine R_g, and estimate C_g by the rule on p. 332.

5. A phase-shift oscillator of the type shown in Fig. 11-19 uses a 6SJ7 tube and three shift sections with $n = 1$. If $C = 1000\ \mu\mu f$ and $R_b = 100{,}000$ ohms, determine the frequency of oscillation and the minimum required value of g_m. If $E_{bo} = 250$ volts and $E_{c2} = 100$ volts, consult a tube manual to determine the value of E_{c1o} required to allow oscillations to begin.

6. The Wien-bridge oscillator of Fig. 11-21 has resistors and capacitors of the following sizes: $R_1 = 1100$ ohms, $R_2 = 500$ ohms, $R = 10{,}000$ ohms, $C = 0.01\ \mu f$. What must be the value of A in order to have sustained oscillations? What is the frequency of oscillation?

7. A plate-coupled multivibrator uses the circuit of Fig. 11-23. The tube is a type 6SN7. $R_{1g} = R_{2g} = 0.5$ megohm, $R_{1b} = R_{2b} = 0.05$ megohm, $C_1 = C_2 = 0.01\ \mu f$, $E_{bb} = 200$ volts. Analyze the operation of the circuit, sketch the waveforms of e_{1c}, e_{2c}, e_{1b}, and e_{2b}, and determine the period of oscillation.

CHAPTER 12

MODULATION AND DETECTION

12-1. The Meaning of Modulation. The process by which some quality or qualities of one waveform (the carrier) are caused to vary according to the instantaneous amplitude of another waveform (the signal) is called modulation.

In telephony the signal is a sound-pressure wave which causes the carbon-grain microphone to change the waveform of a d-c carrier. Thus is formed an alternating current superimposed on a direct current. A transformer suppresses the direct current, and only the alternating component is transmitted to the distant receiver.

Carrier telephony makes it possible to transmit many conversations simultaneously over a single pair of wires. In addition to the original frequency band of 300 to 3000 cps, the process of modulation allows a second conversation to use a 4000 to 7000-cps band, a third conversation to occupy an 8000- to 11,000-cps band, and so on. Filters at the receiver end separate the outputs of the various channels, and the frequencies are then lowered to their original values by a process called demodulation, or detection.

Radio transmission must be carried on at high frequencies for several reasons. First is the impossibility of obtaining satisfactory radiation from an antenna with low frequencies. Second is the desirability of a great many frequency levels to allow simultaneous transmission of many programs without interference, as in the case of carrier telephony. Third, since filters are used to separate the desired from the undesired signals, smaller circuit components are needed at high frequencies than at low. Furthermore, the filtering action is more satisfactory at high frequencies than at audio frequencies. Thus it becomes necessary to use an h-f carrier and to modulate it with the a-f signal.

In most of the subsequent discussions in this chapter the carrier will be considered to be sinusoidal and of a frequency much higher than that of the modulating signal. Although the signal, or modulating waveform, varies widely in practice and is ordinarily nonrecurrent, generally satisfactory conclusions can be reached in a development by considering the modulating signal to be sinusoidal, as though it were a component of a cyclic waveform, and to have a frequency much lower than that of the carrier.

12-2. Types of Modulation. Let us consider an h-f alternating voltage,

$$e_k = E_{km} \cos \phi = E_{km} \cos \omega_k t \tag{12-1}$$

to be known as the carrier, and an l-f alternating voltage,

$$e_s = E_{sm} \cos \omega_s t \tag{12-2}$$

which will be referred to as the signal.

If E_{km} can be changed proportionally to e_s with no appreciable change in ω_k occurring in the process, the carrier is said to have been amplitude-modulated. Thus we shall make $\Delta E_{km} = K_a E_{sm} \cos \omega_s t$ (where K_a is a proportionality factor) and substitute for E_{km} in (12-1) the expression $E_{km} + K_a E_{sm} \cos \omega_s t$ to obtain the equation of the modulated voltage

$$e = (E_{km} + K_a E_{sm} \cos \omega_s t) \cos \omega_k t \tag{12-3}$$

or

$$e = E_{km}(1 + m_a \cos \omega_s t) \cos \omega_k t \tag{12-4}$$

where m_a, called the degree of modulation, or the modulation factor, is a more convenient symbol than its equal, $K_a E_{sm}/E_{km}$. It should be noted that the value of m_a increases if the amplitude of the signal is made greater.

A second way in which the modulation of the carrier by the signal can be effected is to change the angle ϕ by varying the value of ω, while keeping the value of E_{km} constant. This process is known as angle modulation; it has two forms which differ in the manner in which ω is varied.

Now ϕ for angular motion is like s for translation. If v represents a constant velocity, then $s = vt + s_0$; but if v is variable, then $v = ds/dt$ and $s = \int_0^t v\, dt$. Similarly, if ω is constant, $\phi = \omega t + \theta$, where the value of θ is arbitrary and may be made zero if desired, while if ω is variable, $\omega = d\phi/dt$ and $\phi = \int_0^t \omega\, dt$.

Let us consider first that the change in ω is proportional to the instantaneous value of e_s; thus $\Delta\omega = K_f E_{sm} \cos \omega_s t$, where K_f is a proportionality factor and $K_f E_{sm}$ is the maximum frequency deviation, and

$$\begin{aligned} \phi &= \int_0^t (\omega_k + K_f E_{sm} \cos \omega_s t)\, dt = \omega_k t + \frac{K_f E_{sm}}{\omega_s} \sin \omega_s t \\ &= \omega_k t + m_f \sin \omega_s t \end{aligned} \tag{12-5}$$

m_f, called the modulation index, is a convenient representation for $K_f E_{sm}/\omega_s$. Then

$$e = E_{km} \cos (\omega_k t + m_f \sin \omega_s t) \tag{12-6}$$

is the equation of the modulated voltage. This form of angle modulation is called *frequency* modulation.

Instead of varying ω in the manner just described, one can theoretically make the variation in almost any manner desired. For example, one can make the variation of ω in such a way that ϕ changes from its normal value (with constant frequency) by an amount proportional to the instantaneous value of e_s. Thus $\Delta\phi = K_pE_{sm} \cos \omega_st$, $\phi = \omega_kt + K_pE_{sm} \cos \omega_st$, and hence

$$e = E_{km} \cos (\omega_kt + m_p \cos \omega_st) \tag{12-7}$$

where $m_p = K_pE_{sm}$ is the modulation index in this type of angle modulation called *phase* modulation. For phase modulation $\omega = d\phi/dt = \omega_k - K_pE_{sm}\omega_s \sin \omega_st = \omega_k - m_p\omega_s \sin \omega_st$, and hence the maximum frequency deviation is $\Delta\omega_m = m_p\omega_s$. Hence for either type of angle modulation, $m(m_f$ or $m_p)$ is the maximum phase deviation from the normal angle ω_kt, and it is also $\Delta\omega_m/\omega_s$, the ratio of maximum frequency deviation to the audio frequency. In either case also the value of m varies linearly with the value of E_{sm}.

Comparison of the expressions (12-6) and (12-7) leads to a preliminary conclusion that, for a sinusoidal modulating waveform, the results are identical. This is not the case, however. The differences lie in the modulation indexes. m_f is inversely proportional to the frequency ω_s, while m_p is independent of this frequency. For other than a sinusoidal modulating voltage the differences are more pronounced, as may be readily perceived if we remember that for frequency modulation the change in frequency follows the modulating-voltage value; the change in angle is the integral of this voltage. For phase modulation, on the other hand, the change in phase follows the modulating voltage; the change in frequency is the derivative of this voltage. For a sinusoidal waveform the integral and the derivative are also sinusoidal.

It is of some interest to see how these types of modulation can be effected by using an alternator driven at constant speed. Certain of these machines have the stator built so that it can be rotated to a limited extent about the axis of the rotor in order to shift the phase angle of the generated voltage. If the stator is clamped in a fixed position and the strength of the d-c field is varied sinusoidally at a slow rate, the amplitude of the alternating voltage changes correspondingly and this voltage is amplitude-modulated. Next, suppose the field excitation to be constant, the stator clamp loosened, and the stator rocked slowly to and fro. Of course, the generated voltage will vary slightly, but if we ignore this small departure from constant amplitude, we have angle modulation. If the *angular velocity* of the rocking movement is made proportional to $\cos \omega_st$, frequency modulation results, while if the *amount of the angular movement* is made proportional to $\cos \omega_st$, we have phase modulation.

Since practical methods for producing modulated waves are, like other engineering processes, only approximately exact, angle modulation is seldom exactly either frequency modulation or phase modulation and usually involves degrees of both, together with a small degree of amplitude modulation.

12-3. Amplitude Modulation. A voltage with sinusoidal amplitude modulation can be represented mathematically by

$$e = E_{km}(1 + m_a \cos \omega_s t) \cos \omega_k t \tag{12-4}$$

In Fig. 12-1 the waveform from a to b is that of the unmodulated carrier. From b to c the amplitude of the r-f wave rises to a value 50 per cent greater than that of the unmodulated wave, and hence $m_a = 0.50$, or 50

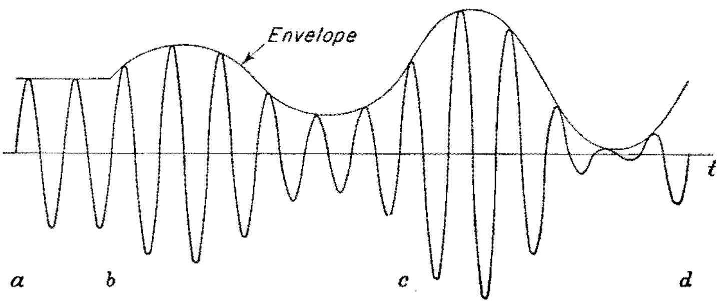


FIG. 12-1. Various degrees of amplitude modulation.

per cent. The waveform from c to d is 100 per cent modulated. This is the limit to the modulation if periods of no r-f voltage are to be avoided.

A smooth curve drawn through the tops of the r-f waves has an equation $E_{km}(1 + m_a \cos \omega_s t)$ and is called the envelope of the r-f waveform. The variable component of this expression has the waveform of the modulating signal. Note that the modulating-signal waveform (inverted) can also be observed as the envelope of the lower extremities of the r-f waveform.

If we multiply the factors contained in Eq. (12-4) and apply the identity $\cos x \cos y = \frac{1}{2} \cos (x + y) + \frac{1}{2} \cos (x - y)$, we obtain

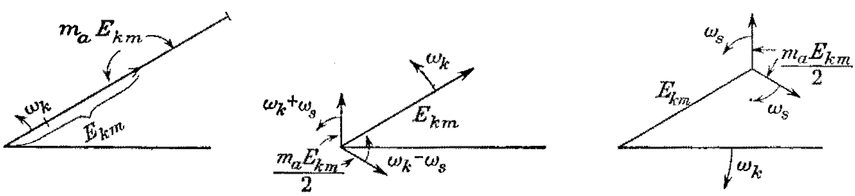
$$e = E_{km} \cos \omega_k t + \tfrac{1}{2}E_{km}m_a \cos (\omega_k + \omega_s)t + \tfrac{1}{2}E_{km}m_a \cos (\omega_k - \omega_s)t \tag{12-8}$$

This expansion shows that when a carrier is amplitude-modulated, the waveform of r-f voltage is no longer precisely sinusoidal (naturally not, since even the positive and negative amplitudes of one wave are not exactly equal) but consists of the sum of three r-f waveforms—that of the carrier, which is exactly the same as before modulation, and those of the side frequencies, one higher and the other lower than the carrier frequency by the amount of the modulating frequency. The amplitudes of the two latter components are the same and are equal to $m_a/2$ times the amplitude of the carrier.

The power delivered by an a-m voltage to a constant resistance can be computed from Eq. (12-8). Since the power varies as the square of the peak voltage, if the power for the carrier is unity, that from each side-frequency voltage is $m_a^2/4$. As an example, if the wave is 100 per cent modulated and the total power is 150 watts, the carrier power is 100 watts and each side-frequency power is 25 watts. Since all the signal transmission is in the side frequencies, it is well to make the modulation factor as near unity as practicable.

Because the usual a-f signal has a waveform which can be considered as the sum of many sinusoidal components, the modulated wave has many side frequencies above and below the carrier frequency. The totality of those above constitute the upper sideband; the others make the lower sideband. In Fig. 12-2 is shown graphically the frequency positions and

FIG. 12-2. The a-f spectrum is moved to a new position by modulation.



(a) (b) (c)

FIG. 12-3. Vector diagrams for an a-m wave.

the amplitudes of the components of a certain audio signal and also of the corresponding sidebands after modulation. The total bandwidth required for the modulated wave is double the frequency of the highest a-f component in the modulating wave. In ordinary a-m radiobroadcasts the highest audio frequency handled is about 7500 cps, requiring a bandwidth of 15 kc. Since the total r-f spectrum available for a-m broadcasting is divided into 10-kc channels, it has been advisable to make a geographical separation of stations on adjacent channels in order to avoid excessive interference.

Figure 12-3 shows vector diagrams for a waveform with sinusoidal modulation. In (a) a vector rotates at a constant angular velocity of ω_k radians per sec, while the vector length varies between the limits $E_{km}(1 - m_a)$ and $E_{km}(1 + m_a)$ in simple-harmonic fashion f_s times per second. Figure 12-3b shows the same occurrence explained with vectors representing the side frequencies. The sum of the side-frequency vectors together with the carrier vector varies in the same manner as in (a). Figure 12-3c shows the same motions as viewed by an observer traveling with the carrier vector. The reference axis then appears to rotate in the reverse direction, and the two side-frequency vectors rotate in opposite directions at a rate of ω_s radians per sec. In all three diagrams one can observe that the vector representing the total modulated wave rotates at a constant angular velocity relative to the reference axis. Since the actual instantaneous voltage is the projection of the rotating vector on the horizontal axis, it appears that the time intervals between zero values of the r-f voltage are constant for an a-m wave.

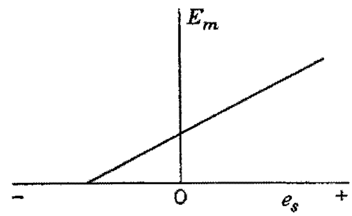


FIG. 12-4. Ideal modulation characteristic.

12-4. Methods of Amplitude Modulation. The ideal characteristic of a device for producing an a-m waveform is shown in Fig. 12-4, where E_m represents the peak value of the r-f wave and e_s is the instantaneous value of the modulating signal. We may note that the relationship is a linear one so that changes in the peak values are proportional to the changes in e_s.

There are many methods by which an a-m waveform can be produced. The most important ones are those which involve a class C amplifier. Of some importance also are those involving nonlinear circuit elements such as semiconductors and diodes and grid tubes which operate with a square-

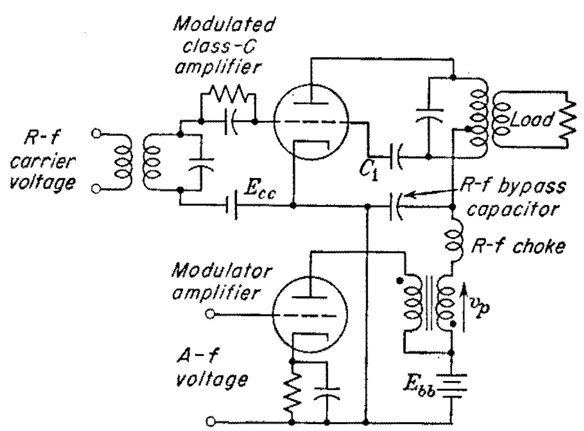


FIG. 12-5. A modulated class C amplifier and its modulator.

law characteristic. Of little importance are the old, crude methods such as the one in which a carbon microphone is placed in series with the antenna lead, where it effects modulation by its variable absorption of r-f energy.

12-5. Plate Modulation of a Class C Amplifier. The ideal modulation characteristic of Fig. 12-4 can be approached by applying the modulating signal in the plate circuit of a class C amplifier whose grid is driven by the r-f carrier voltage and whose tuned circuits are resonant at the carrier frequency. Figure 12-5 shows the schematic diagram for such a modulated amplifier, together with a modulator amplifier. The class C amplifier in this case employs a triode. To avoid self-oscillation, a neutralizing current is fed through C_1 to offset the current through C_{gp}, which would otherwise excite the tuned grid circuit. This method of neutralization is called Hazeltine-system, or plate, neutralization. The r-f output voltage to the load depends upon the effective plate-supply voltage, $e_{bb} = E_{bb} + v_p$, where v_p is the voltage output of the a-f power amplifier (often

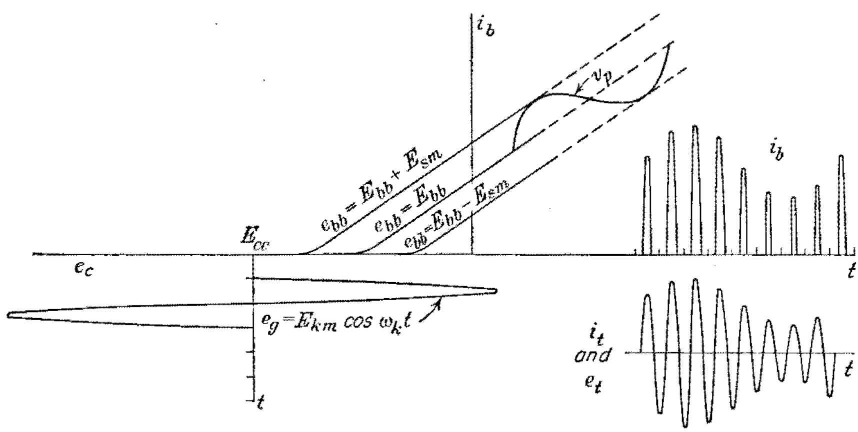


FIG. 12-6. Graphical explanation of the action of a plate-modulated class C amplifier.

called the modulator). The d-c magnetization of the a-f transformer core is reduced by properly polarizing the primary and secondary winding connections.

To show the modulating action, let us refer to Fig. 12-6, where dynamic characteristics for three values of e_{bb} have been drawn. These values of e_{bb} correspond to values of v_p equal to $+E_{sm}$, 0, and $-E_{sm}$, respectively. Since the impedance of the plate-tank coil is very low to an a-f current as well as to direct current, curves taken for static values of e_{bb} are sufficiently accurate for dynamic use.

With the class C tube biased well beyond cutoff by the effect of E_{cc} and a small amount of grid-leak action, the large r-f grid signal causes large pulses of plate current when $e_{bb} = E_{bb}$. These pulses last for somewhat less than a half cycle, just as in the case of an ordinary class C amplifier. Now as $v_p = E_{sm} \cos \omega_s t$ slowly changes e_{bb}, the action progresses along different dynamic characteristics, with the resulting pulses varying in amplitude and to some degree in conduction angle. These current pulses excite in the tank circuit a large circulating r-f current, which likewise varies in magnitude. If the effective Q of the tank is properly chosen, the pattern of tank current and of tank voltage will be as shown.

Figure 12-7 shows waveforms of grid voltage and of plate voltage for constant grid bias. The modulating voltage here has a peak value of E_{bb}. This causes 100 per cent modulation and is the greatest value that can be used if plate-current cutoff (overmodulation) is to be avoided. The pattern of plate voltage with the slow change filtered out will be the

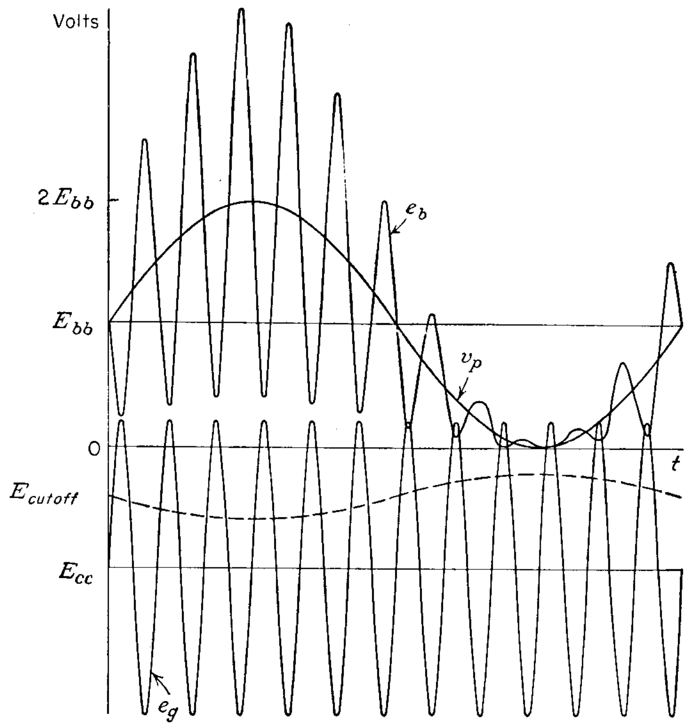


FIG. 12-7. Waveforms of grid voltage and of plate voltage with constant grid bias and 100 per cent plate-circuit modulation.

tank voltage and will resemble the waveform of e_t shown in Fig. 12-6. Although the class C amplifier was designed so that $e_{b_{\min}}$ was approximately $e_{c_{\max}}$, it may be noted that, at the positive crest of the modulation cycle, $e_{b_{\min}} = 2e_{c_{\max}}$. This departure from design conditions causes some nonlinearity in the modulation characteristic. If a grid leak is inserted for additional bias, the total bias will decrease at the peak of the modulation cycle because of the decrease in grid current and the waveforms of r-f grid voltage will slowly rise and fall with the modulation in such a way as to make $e_{c_{\max}}$ approximately equal to $e_{b_{\min}}$.

A graphical method for the design, which uses constant-current tube characteristics, follows that for a class C amplifier. Computations are made for instantaneous modulation voltages of, say, zero, positive maximum, negative maximum, and one-half each of these extreme values. It is well to choose E_{bb} as about one-half the normal class C amplifier value. Then for the zero-modulation case, $e_{bb} = E_{bb}$, and tentative values of E_{cc}, $e_{c_{max}}$ and $e_{b_{min}}$ are chosen. Computations of E_{p_1}, I_{p_1}, and $R(= E_{p_1}/I_{p_1})$, the antiresonant tank impedance, can now be made. For the case of the positive maximum of the modulation cycle, $e_{bb} = 2E_{bb}$ (for 100 per cent modulation), E_{cc} and $e_{c_{max}}$ are the same as before, and $e_{b_{min}}$ is somewhat greater than before, its value being determined by trial and error to maintain R constant. In like manner the values of E_{p_1} and I_{p_1} are made for the other modulating-voltage values and the value of R checked as being held constant.

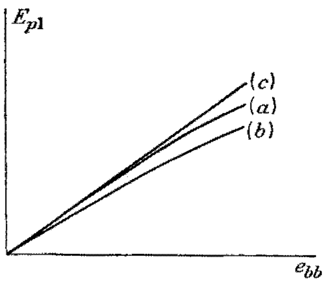


FIG. 12-8. Actual and ideal modulation characteristics for a class C amplifier.

If the values of E_{p_1} are plotted against those of e_{bb} and R has been well chosen, one has the modulation characteristic as shown by curve a of Fig. 12-8. Curve b shows the characteristic for a lower-Q tank. These curves should be compared with c, which represents the ideal. If the bias is made somewhat less as e_{bb} is increased, it will be found that curve a is somewhat straighter. Hence we see the advantage of a high impedance load and of the use of a grid leak for additional bias. The Q of the tank is largely dependent upon the amount of loading, and since a high impedance load and a high Q mean a low power output, compromise is obviously necessary. A further difficulty with a high-Q tank is its inability to change its stored energy rapidly and thus to respond to rapid variations in the modulating voltage. A Q value of 10 is usually satisfactory.

The fixed supply of E_{bb} and the modulator together furnish the power in the plate circuit. Since the frequency of the former is zero and that of the latter is ω_s, we need only these frequency components of the plate current to obtain the plate input power. These are both contained in the average current over the r-f cycle, which varies in value as the modulating voltage slowly changes. It is possible to determine the manner of this variation if we compute I_{ba} (in this case a function of time) for various values of e_{bb} in the same fashion that we computed E_{p_1} as described in a preceding paragraph. If I_{ba} is then plotted against e_{bb}, a curve similar to a of Fig. 12-8 will be obtained. Thus it is a reasonable assumption that I_{ba} varies linearly with the value of the modulating voltage. If I_{bao} means the average current with no modulation, then

$$I_{ba} = I_{bao}(1 + m_a \cos \omega_s t) \tag{12-9}$$

This expression for I_{ba} contains both the zero-frequency and the ω_s components of the plate current. Since the value of V_{pm} which gives 100 per cent modulation is E_{bb}, we may write, for a degree of modulation m_a,

$$e_{bb} = E_{bb} + v_p = E_{bb} + m_a E_{bb} \cos \omega_s t = E_{bb}(1 + m_a \cos \omega_s t) \qquad (12\text{-}10)$$

The total plate-circuit power input is

$$\begin{aligned} P_b &= \frac{1}{2\pi} \int_0^{2\pi} I_{ba} e_{bb} \, d(\omega_s t) \\ &= \frac{1}{2\pi} \int_0^{2\pi} I_{bao}(1 + m_a \cos \omega_s t) E_{bb}(1 + m_a \cos \omega_s t) \, d(\omega_s t) \\ &= E_{bb} I_{bao} \left(1 + \frac{m_a^2}{2}\right) \end{aligned} \qquad (12\text{-}11)$$

The power supplied by the source of E_{bb} is E_{bb} multiplied by the zero-frequency component of I_{ba}, and hence

$$P_{bb} = E_{bb} I_{bao} \qquad (12\text{-}12)$$

and therefore the power from the modulator must be

$$P_M = \frac{m_a^2}{2} E_{bb} I_{bao} \qquad (12\text{-}13)$$

For 100 per cent modulation the modulator must supply half as much power as that from the source of E_{bb}.

As modulation is applied, the fundamental components of plate voltage and current and the average plate current over an r-f cycle all vary approximately linearly with the modulating voltage. Thus both the power output and the power input vary as the square of the modulating voltage, and we see that the plate efficiency is approximately constant over a modulating cycle. We shall therefore assume the plate-circuit efficiency to be η_C, the same as for the class C amplifier with no modulation.

The power output of the class C amplifier may be computed and broken down as follows:

$$P_{ac} = P_b \eta_C = \eta_C E_{bb} I_{bao} \left(1 + \frac{m_a^2}{2}\right) = \eta_C (P_{bb} + P_M) \qquad (12\text{-}14)$$

The carrier power is the same with or without modulation, and hence $\eta_C P_{bb}$ is the power in the carrier output. It appears that $\eta_C P_M$ must be the added power contained in the sidebands. Thus the fixed plate power supply furnishes the carrier power; the modulator supplies the sideband power.

The plate dissipation in the class C amplifier is

$$P_{pC} = P_b(1 - \eta_C) = E_{bb}I_{bao}\left(1 + \frac{m_a^2}{2}\right)(1 - \eta_C) \qquad (12\text{-}15)$$

With no modulation, P_{pC} is only $E_{bb}I_{bao}(1 - \eta_C)$, while with 100 per cent modulation it becomes $1.5E_{bb}I_{bao}(1 - \eta_C)$. Hence it would appear that a plate dissipation only two-thirds of the normal allowable value should be assumed in making the modulated class C amplifier design. Since 100 per cent modulation is usually only a transient phenomenon and the degree of modulation under ordinary circumstances averages much less, the factor two-thirds is unnecessarily too low.

The power required for the modulation is given by Eq. (12-13). If the modulation transformer has an efficiency of η_T and the plate efficiency of the modulator tube is η_M, then the plate dissipation in the modulator tube is

$$P_{pM} = \frac{P_M}{\eta_T\eta_M}(1 - \eta_M) = \frac{m_a^2}{2\eta_T\eta_M}E_{bb}I_{bao}(1 - \eta_M) \qquad (12\text{-}16)$$

If we compare this value of plate dissipation with that for the class C tube [Eq. (12-15)], we see that for the case of 100 per cent modulation

$$\frac{P_{pM}}{P_{pC}} = \frac{1}{3\eta_T\eta_M}\frac{1 - \eta_M}{1 - \eta_C} \qquad (12\text{-}17)$$

Suppose the class C amplifier has a plate efficiency of 75 per cent and the modulation transformer has an efficiency of 80 per cent. If the modulator-tube plate efficiency is 20 per cent because of class A operation, the value of the ratio of P_{pM}/P_{pC} is 6.7. Thus for every 100 watts of allowable plate dissipation for the class C tube, 670 watts of plate dissipation must be provided for the modulator. This ratio can be greatly reduced if the modulator is made push-pull class B (see Art. 8-13). The efficiency of this class of operation can be made considerably higher than that for class A. If we assume 60 per cent, the ratio determined from Eq. (12-17) becomes only 1.1. Hence in the above example the total plate dissipation in the push-pull modulator need be only about 110 watts.

A matter of great importance is the load on the modulator. The modulator delivers P_M watts of power at a voltage V_p; hence the load resistance R_M can be computed from the formula $P_M = V_p^2/R_M$. For a degree of modulation m_a, $V_p = m_aE_{bb}/\sqrt{2}$, and $P_M = \frac{1}{2}m_a^2E_{bb}I_{bao}$. Hence

$$R_M = \frac{V_p^2}{P_M} = \frac{\frac{1}{2}m_a^2E_{bb}^2}{\frac{1}{2}m_a^2E_{bb}I_{bao}} = \frac{E_{bb}}{I_{bao}} \qquad (12\text{-}18)$$

Thus the secondary load for the modulator can be determined from instrument readings of E_{bb} and of the average plate current while the r-f grid voltage is applied, but with no modulation. The turns ratio of the modulator transformer is chosen to make this a suitable plate load for the amplifier.

Because of the fact that plate-modulated class C amplifiers can be easily adjusted to give excellent linearity, they are very popular. However, the large amount of modulating power required necessitates an expensive outlay for modulator tubes and transformers. For high-power transmitting stations the modulating transformer may weigh several tons. For mobile equipment the heavy weight and bulk of the modulator equipment are further drawbacks to this method of modulation.

12-6. Grid-bias Modulation of a Class C Amplifier. If a class C amplifier with an r-f input voltage has its grid bias varied at the rate of

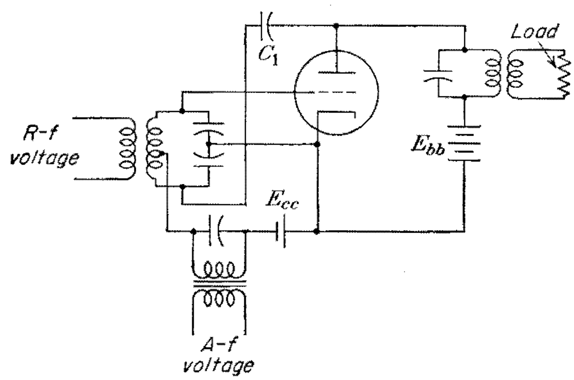


FIG. 12-9. A grid-bias-modulated class C amplifier.

the audio signal, the magnitude and conduction angle of the plate-current pulses vary in much the same manner as for plate modulation. Figure 12-9 shows a circuit which will produce a modulated wave in this fashion. This one employs a triode, but a tetrode or a pentode can be substituted if desired. The arrangement shown for neutralization with a triode utilizes a capacitor C_1 to feed a current into the grid tank at a point where its effect is opposite that of the current flow through C_{gp}. This method is known as Rice-system, or grid, neutralization.

Figure 12-10 shows the action in a grid-bias-modulated class C amplifier. The adjustment of the circuit should be such as to make the dynamic characteristic as straight as possible. The amount of grid bias E_{cc}, the amplitude of the r-f grid signal, and the amplitude of the modulating signal should be adjusted so that for 100 per cent modulation the following conditions hold:

1. On the positive peak of the modulation cycle the tube is usually made to operate class B (it must not be class C), and the grid should swing considerably positive in order to obtain good plate-circuit efficiency.

2. On the negative peak of the modulating voltage the plate current should be near zero, which means that the modulating voltage peak is one-half the r-f voltage peak.

3. It follows that the magnitude of E_{cc} should be equal to one-half the amplitude of the r-f voltage added to the magnitude of the cutoff voltage.

The actual modulation characteristic for the circuit can be determined by applying a fixed carrier-frequency grid voltage and zero modulating voltage, and determining the r-f tank current for various values of grid bias. Many runs can be made with different values of r-f grid voltage

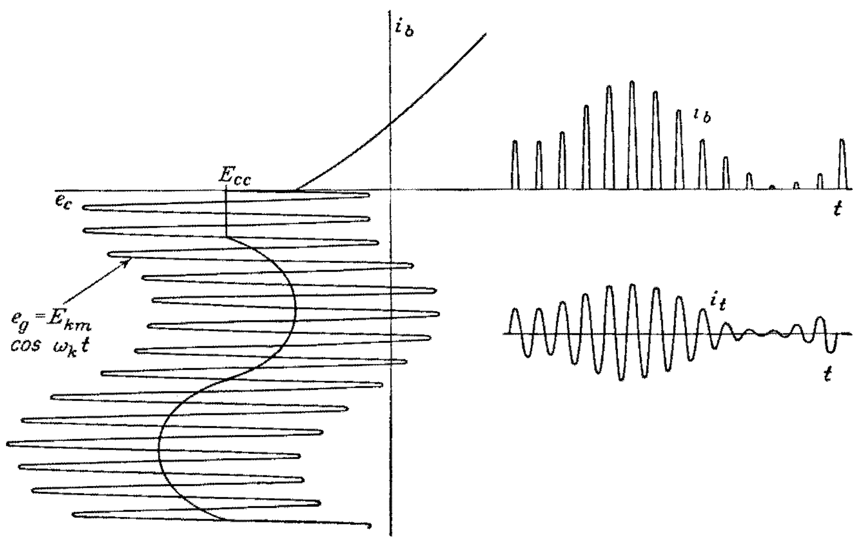


FIG. 12-10. Action of a grid-bias-modulated class C amplifier.

and of plate-load resistance. The graph of tank current I_t vs. E_{cc} which gives the greatest linearity with satisfactory power output is thus determined. It is difficult to obtain excellent linearity, the performance being generally somewhat inferior to that obtained with the plate-modulation system.

As in the case of plate modulation, the average plate current over an r-f cycle varies approximately as the modulation voltage, in this case as the change in grid bias. Since the grid voltage with modulation averages a constant amount E_{cc}, the plate current averaged over the modulation cycle is approximately constant regardless of the degree of modulation. On the other hand, since for a linear modulation characteristic the peak r-f tank current for 100 per cent modulation is double the value for no modulation, the peak r-f power is increased four times by applying the modulation voltage. Thus if the efficiency at the peak of the modulating cycle can be made, say, 80 per cent, that for zero modulation is only 20 per cent. For a given class C tube this makes the amount of average power obtainable from a grid-bias-modulated amplifier considerably lower than that from a plate-modulated one, which can have a practically constant high efficiency for any degree of modulation.

On the credit side, however, the grid-bias-modulated amplifier* needs little modulating power, and a low-power class A push-pull modulator is sufficient. For some applications the low bulk and weight of this type of modulating device make it preferable to the plate-modulated system. With the addition of negative feedback to the circuit, the distortion produced by the normal somewhat nonlinear modulation characteristic can be reduced to a satisfactory level.

12-7. Modulation of a Class A Amplifier. One of the early modulating devices (van der Bijl modulator) employed a class A amplifier with the r-f and a-f signals both applied to the grid circuit in the same manner as for the grid-bias-modulated class C amplifier of Art. 12-6. The action in the class A case, however, depends upon the curvature of the plate characteristics. The principle can be applied equally well to a diode, to a triode in which one signal is applied to the grid while the other is applied to the plate, or to a multigrid tube with the signals applied to the same or to different electrodes.

In any case an analysis can be made in the same fashion. Let us first assume the shape of the plate characteristics to be approximately parabolic in the operating region; thus

$$i_b = k(e_b + \mu e_c)^2 \tag{12-19}$$

this being an adaptation of Eq. (4-3). Another exponent value such as 1.5 can be used instead if desired.

The alternating plate current can be approximated by using the first few terms of Taylor's expansion (Art. 6-19). Thus [see Eq. (6-57)]

$$i_p = a_1 e + a_2 e^2 + a_3 e^3 + \cdots$$

where the a's are constants and e is the total effective signal in grid-circuit terms. Thus if two signals are both applied to the grid circuit, $e = e_1 + e_2$, while if one is applied to the grid while another is inserted in the plate circuit, $e = e_1 + e_2/\mu$, and so on. The results obtained in all cases are similar, the only differences being those caused by values of μ. Let us continue with the case in which both signals are applied in the grid circuit of a triode as in Fig. 12-11, where only the basic circuit is drawn.

* For further details of the grid-bias-modulated amplifier consult Samuel Seely, "Electron-tube Circuits," pp. 332–337, McGraw-Hill Book Company, Inc., New York, 1950.

Here

$$e = e_1 + e_2 \tag{12-20}$$

If we assume that the load is resistive and the same for all frequencies, then a_1 is given by Eq. (6-59),

$$a_1 = \frac{\mu}{r_p + R_L}$$

We can obtain r_p at T as the value of $\partial e_b/\partial i_b$ from Eq. (12-19). Thus

$$r_p = \frac{1}{2k(e_b + \mu e_c)}\bigg|_T = \frac{1}{2k(E_{bt} + \mu E_{ct})} \tag{12-21}$$

Hence a_1 is a constant dependent upon tube and circuit parameters. The value of a_2 is given by Eq. (6-60):

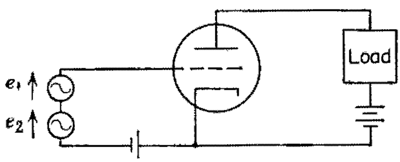


FIG. 12-11. Square-law triode with two signals applied in the grid circuit.

$$a_2 = \frac{-\mu^2 r_p}{2(r_p + R_L)^3}\frac{\partial r_p}{\partial e_b}$$

This equation contains $\partial r_p/\partial e_b$, which from formula (12-21) equals $-1/2k(e_b + \mu e_c)^2\big|_T = -1/2k(E_{bt} + \mu E_{ct})^2$ at the operating point. Thus a_2 will usually be small compared with a_1. Likewise the coefficient a_3 [formula (6-61)] is considerably smaller yet, and we shall assume that sufficient accuracy is obtained by using only the first two terms in the expansion, and the current for our application becomes

$$i_p = a_1(e_1 + e_2) + a_2(e_1 + e_2)^2 \tag{12-22}$$

It should be understood that the square-law form (12-22) holds even though the exponent in Eq. (12-19) can be chosen as 1.5 or otherwise (not unity). In other words, the square-law behavior is a good assumption when small signals are applied to a device with a curved dynamic characteristic.

If we assume e_1 to be a carrier voltage and e_2 to be the modulating voltage, then

$$e_1 = e_k = E_{km} \cos \omega_k t \tag{12-23}$$

and

$$e_2 = e_s = E_{sm} \cos \omega_s t \tag{12-24}$$

Hence

$$\begin{aligned} i_b &= a_1(E_{km} \cos \omega_k t + E_{sm} \cos \omega_s t) + a_2(E_{km} \cos \omega_k t + E_{sm} \cos \omega_s t)^2 \\ &= a_1(E_{km} \cos \omega_k t + E_{sm} \cos \omega_s t) + \frac{a_2 E_{km}^2}{2} + \frac{a_2 E_{km}^2}{2} \cos 2\omega_k t \\ &\quad + a_2 E_{km} E_{sm}[\cos (\omega_k + \omega_s)t + \cos (\omega_k - \omega_s)t] \\ &\quad + \frac{a_2 E_{sm}^2}{2} + \frac{a_2 E_{sm}^2}{2} \cos 2\omega_s t \end{aligned} \tag{12-25}$$

Now when i_b flows through a load of constant resistance, all the frequency components of (12-25) appear in the output. This is not a very useful result, and therefore let us replace the constant load resistance by an L-C tank circuit tuned to the frequency ω_k. The formulas for a_1, a_2, etc., used previously do not now apply, and instead we should obtain coefficients which depend upon frequency. Thus in (12-25) the form of the expansion remains the same, but the a_1 coefficients depend for their values upon the frequencies of the signals, and likewise for the a_2 values. If we are interested only in qualitative results, this makes no difference and hence formula (12-25) is still useful. In fact, in case the tank provides a high impedance $R_L = R_t$ to current components having frequencies near ω_k, and zero impedance to other components, the formulas (6-59) and (6-60) for a_1 and a_2 are correct. We shall assume this to be so. Thus the alternating plate voltage has an approximate form

$$e_p = R_t\{a_1E_{km} \cos \omega_k t + a_2E_{km}E_{sm}[\cos (\omega_k + \omega_s)t + \cos (\omega_k - \omega_s)t]\} \tag{12-26}$$

or

$$e_p = a_1E_{km}R_t\left(1 + \frac{2a_2E_{sm}}{a_1} \cos \omega_s t\right) \cos \omega_k t \tag{12-27}$$

which is an a-m wave since it is similar to Eq. (12-4). The modulation factor is

$$m_a = \frac{2a_2}{a_1} E_{sm} = \frac{-\mu r_p}{(r_p + R_t)^2} \frac{\partial r_p}{\partial e_b} E_{sm} \tag{12-28}$$

We see that m_a depends upon E_{sm} (but not upon E_{km}) and that m_a is small unless E_{sm} is made very large.

Square-law modulation as a substitute for the linear modulation previously discussed has few applications today. The action described here takes place inadvertently in amplifiers and causes hum and cross-modulation effects. In the case of the former, the hum-frequency voltage induced in the grid circuits of a tube operating nonlinearly modulates the signal and causes hum to appear in the output. In cross modulation a signal from an unwanted radio station is sometimes insufficiently filtered out from the circuit leading to the grid of a vacuum tube operating nonlinearly. The grid circuit rectifies this modulated wave and produces a modulation-frequency voltage in the grid circuit. This in turn modulates the wanted signal by square-law action.

Square-law modulation does have three very important applications, which are discussed in the following articles.

12-8. The Balanced Modulated Amplifier. The circuit for the balanced modulated amplifier (also called a balanced modulator) of Fig.

12-12 resembles those of the balanced amplifier of Art. 7-21 and the push-pull amplifier of Art. 8-9. An operating point (E_{bo},E_{co}) is chosen which gives each tube identical square-law characteristics. Thus from Eq. (6-57) and the discussion of the preceding article we can write the plate current of either tube as

$$i_p = a_1 e_g + a_2 e_g^2 \tag{12-29}$$

The signal voltages e_1 and e_2 can have any frequencies. For tube 1 the grid voltage is $e_{1g} = e_1 + e_2$, while for tube 2, $e_{2g} = e_1 - e_2$. Then

$$i_{1p} = a_1(e_1 + e_2) + a_2(e_1 + e_2)^2 \tag{12-30}$$

and

$$i_{2p} = a_1(e_1 - e_2) + a_2(e_1 - e_2)^2 \tag{12-31}$$

Now the voltages obtained from the plate-circuit transformers may be determined. e_3 will be proportional to the sum of i_{1p} and i_{2p}, while e_4 is

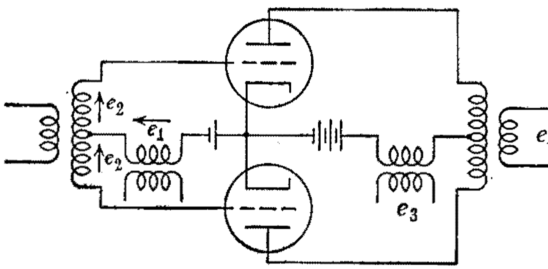


FIG. 12-12. Basic circuit for a balanced modulator.

proportional to $i_{1p} - i_{2p}$. Hence

$$e_3 = K_1(2a_1e_1 + 2a_2e_1^2 + 2a_2e_2^2) \tag{12-32}$$

and

$$e_4 = K_2(2a_1e_2 + 4a_2e_1e_2) \tag{12-33}$$

where K_1 and K_2 are constants of proportionality.

As an example of the applications let us suppose that $e_1 = E_{km} \cos \omega_k t$ and $e_2 = E_{sm} \cos \omega_s t$. Then e_3 contains components having frequencies ω_k, $2\omega_k$, and $2\omega_s$, while e_4 has frequencies ω_s, $\omega_k + \omega_s$, and $\omega_k - \omega_s$. Thus e_4 in this case contains the two sidebands of an a-m wave, but it lacks the carrier-frequency component. The balanced modulator can therefore be used in a system in which for any reason it is desired to suppress the carrier. Such a procedure is not usually desirable in radio transmission because of the difficulty at the receiver of reinserting the carrier with the correct phase relations required for proper recovery of the original audio signal. However, if one of the sidebands is filtered out, either before or after transmission, then the phase of the carrier is not important and a suppressed-carrier single-sideband system becomes practicable. This method is used in carrier telephony and in some radio-communication systems.

The student can readily determine the frequencies present in e_3 and in e_4 if other arrangements of the input voltages e_1 and e_2 are made.

12-9. Heterodyne Frequency Conversion.[2] In the circuit of Fig. 12-11 let us suppose an operating point Q has been chosen in the nonlinear region of the tube characteristics. Then a signal applied to the grid circuit produces a response representable as

$$i_p = a_1 e_g + a_2 e_g^2 \tag{12-34}$$

In this case let us make e_1 an a-m voltage

$$e_1 = E_{km}(1 + m_a \cos \omega_s t) \cos \omega_k t \tag{12-35}$$

and e_2 a sinusoidal voltage generated by a local oscillator,

$$e_2 = E_{2m} \cos \omega_2 t \tag{12-36}$$

where ω_2 is considerably higher than ω_k (considered so in this discussion only). Then since $e_g = e_1 + e_2$, it follows that

$$\begin{aligned} i_p = a_1[E_{km}(1 + m_a \cos \omega_s t) \cos \omega_k t + E_{2m} \cos \omega_2 t] \\ + a_2[E_{km}(1 + m_a \cos \omega_s t) \cos \omega_k t + E_{2m} \cos \omega_2 t]^2 \end{aligned} \tag{12-37}$$

There are many frequencies present in i_p, but if the plate load is tuned to a frequency of, say, $\omega_2 - \omega_k$, only terms having frequencies at or near this value are important and others may be omitted. Hence, if the plate-load impedance near this frequency is R_t,

$$\begin{aligned} e_p = R_t a_2 E_{km} E_{2m} \cos (\omega_2 - \omega_k)t + \frac{R_t a_2 E_{km} E_{2m} m_a}{2} [\cos (\omega_2 - \omega_k - \omega_s)t \\ + \cos (\omega_2 - \omega_k + \omega_s)t] = a_2 R_t E_{km} E_{2m}(1 + m_a \cos \omega_s t) \cos (\omega_2 - \omega_k)t \end{aligned} \tag{12-38}$$

Expression (12-38) for e_p is the equation for an a-m wave with a modulation frequency ω_s and a modulation factor m_a the same as those in the original modulated wave, but with a carrier frequency $\omega_2 - \omega_k$ replacing ω_k. Thus the action of this circuit has moved the intelligence-bearing sidebands from a position in the frequency spectrum where they were centered at ω_k to another position where they are centered about the frequency $\omega_2 - \omega_k$. The carrier and sideband frequencies in the new position are sometimes referred to as intermediate frequencies.

This relocation of the sidebands is called frequency conversion. The process is very commonly employed in superheterodyne receivers in order to shift the frequency band of any desired station to an i-f band. The fixed-frequency amplifier handling the latter is usually of the double-tuned type (Art. 7-37) and can give high gain and excellent selectivity without too much sideband clipping.

The oscillator producing the voltage e_2 may be separate from the square-law tube, in which case the latter is called a mixer tube. On the other hand, it is possible to have one tube act as both oscillator and mixer; this tube is then called a converter tube.

The application of the voltages e_1 and e_2 to the same grid causes some undesirable interaction between the circuits supplying these voltages. Except for very high frequencies it is more common to use multigrid tubes and apply the two signals to separate shielded grids. Figure 12-13 shows

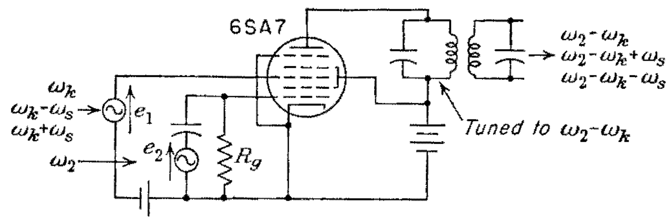


FIG. 12-13. A pentagrid mixer circuit.

a circuit for a mixer employing a pentagrid tube. Grids 2 and 4 are shielding electrodes, and grid 5 is a suppressor. The local-oscillator voltage e_2 is applied between grid 1 and cathode. The grid voltage e_2 swings just enough positive to make a small average grid current flow through R_g, thus biasing the grid to about the peak value of e_2. Since this bias value is usually beyond plate-current cutoff, the electron stream past grid 1 flows in pulses. Upon reaching grid 2, where the field is considerably dependent upon the grid 3 signal voltage e_1, the electrons either pass on to the plate or stop off at grid 2, each to a degree depending upon the relative phases of e_1 and e_2, which are constantly shifting. Both voltages hence determine the plate current. e_1 has but little effect on the space charge near the cathode and hence upon the current pulses passing through grid 1. Therefore, there is little interaction between the signal and the local oscillator. The plate voltage has little effect on the pulses passing through grid 3 because of the shielding effect of grids 4 and 5, and hence the tube has a high plate resistance. This is preferable to the low-r_p value for a triode which lowers the Q of the primary winding of the i-f transformer.

The alternating plate current caused by the signal e_1 can be computed as $g_m e_1$ (g_m of the third grid to plate) because r_p is very large. But since i_p varies cyclically at the rate of ω_2, g_m varies at the same rate. We can expand g_m in a Fourier series of cosine terms since the values of g_m for rising e_2 are repeated in reverse order as e_2 falls, making g_m an even function. Thus

$$g_m = b_0 + b_1 \cos \omega_2 t + b_2 \cos 2\omega_2 t + b_3 \cos 3\omega_2 t + \cdots \tag{12-39}$$

and hence if we consider only the carrier component of e_1 (other components can be treated similarly),

$$i_p = (b_0 + b_1 \cos \omega_2 t + b_2 \cos 2\omega_2 t + b_3 \cos 3\omega_2 t + \cdots) E_{km} \cos \omega_k t \tag{12-40}$$

If we tune the load to a frequency $\omega_2 - \omega_k$, we are interested only in the component of i_p having this frequency. It is not difficult to show that

$$i_{\omega_2-\omega_k} = \frac{b_1 E_{km}}{2} \cos (\omega_2 - \omega_k)t \tag{12-41}$$

where b_1 is the Fourier coefficient determined as

$$b_1 = \frac{2}{2\pi} \int_0^{2\pi} g_m \cos \omega_2 t \, d(\omega_2 t) \tag{12-42}$$

The ratio of the magnitude of the i-f carrier current to the magnitude of the carrier signal voltage on grid 3 is called the conversion transconductance g_c. We can see that this equals

$$g_c = \frac{\frac{1}{2} b_1 E_{km}}{E_{km}} = \frac{b_1}{2} = \frac{1}{2\pi} \int_0^{2\pi} g_m \cos \omega_2 t \, d(\omega_2 t) \tag{12-43}$$

If a graph of g_m (grid 3 to plate) values for various grid 1 voltages can be obtained, this formula enables one to make a graphical computation of g_c.

The use of the square-law modulator or the multigrid mixer for frequency conversion accomplishes what was done in the early days of superheterodyne receivers by using a certain type of detector circuit. Hence arose the custom of calling frequency conversion a detection process, and the converter stage is sometimes even yet called the first detector (as differentiated from a later tube in the circuit, the second detector, which is used to recover the audio signal).

12-10. Square-law Demodulation. Demodulation is the process of recovering the original audio signal from a modulated wave. If the modulated signal, assumed to have ω_k as its center frequency, is injected into the grid circuit of Fig. 12-11 as e_1, while the input e_2 is omitted, we are in a sense operating a heterodyne frequency converter for which the results are already stated in Eq. (12-37) if we make $e_2 = 0$. Thus

$$
\begin{aligned}
i_p &= a_1E_{km}(1 + m_a \cos \omega_s t) \cos \omega_k t + a_2E_{km}^2(1 + m_a \cos \omega_s t)^2 \cos^2 \omega_k t \\
&= a_1E_{km}(1 + m_a \cos \omega_s t) \cos \omega_k t + \frac{a_2E_{km}^2}{2}\Bigg[1 + \frac{m_a^2}{2} + 2m_a \cos \omega_s t \\
&\quad + \frac{m_a^2}{2} \cos 2\omega_s t + \left(1 + \frac{m_a^2}{2}\right) \cos 2\omega_k t + m_a \cos (2\omega_k + \omega_s)t \\
&\quad + m_a \cos (2\omega_k - \omega_s)t + \frac{m_a^2}{4} \cos (2\omega_k + 2\omega_s)t + \frac{m_a^2}{4} \cos (2\omega_k - 2\omega_s)t\Bigg]
\end{aligned}
\tag{12-44}
$$

If a plate load is chosen which offers high impedance to a-f components only, then the alternating plate voltage has the form

$$e_p = K\left(\cos \omega_s t + \frac{m_a}{4} \cos 2\omega_s t\right) \tag{12-45}$$

Thus the original a-f signal is recovered together with a second-harmonic distortion component. For 100 per cent modulation this distortion will amount to 25 per cent, and to hold it to even 10 per cent requires that the degree of modulation be kept down to 40 per cent. Since the average modulation for broadcasting may be near this figure, the distortion may be acceptable except on the high-volume peaks. However, an audio signal actually consists of numerous a-f components, and a further development following the method just employed would show additional distortion with frequencies evolved which are combinations of all these audio frequencies and also their harmonics. Since none of these components would have an amplitude greater than that of the second harmonic, however, the limitation just stated on the degree of modulation still applies.

At one time the square-law modulated-amplifier circuit was quite popular as a demodulator, but it has been superseded in most cases by the linear diode detector to be described in the next article, which also demodulates but with better results.

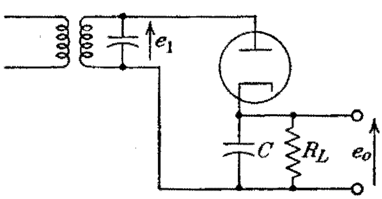


FIG. 12-14. A diode detector.

12-11. Linear Diode Detection. Most broadcast receivers today use diode linear detectors to recover the audio signal from the modulated r-f wave. Although the circuit of Fig. 12-14 will give an a-f output if C is omitted, the presence of C is useful in two ways. First, it acts as a filter to remove the r-f components from the output, and, second, it makes the output voltage higher than it would be otherwise. We shall consider only the case in which C is used.

Assume e_1 to be a modulated wave. Near the positive peak of the r-f cycle a pulse of current through the tube charges the capacitor C to a voltage value equal to that of the r-f peak less the amount of the tube drop. When the r-f voltage falls from its peak, the tube soon ceases to conduct because the capacitor holds the cathode potential almost at the peak value, where it remains except for a slow fall due to discharge through R_L,

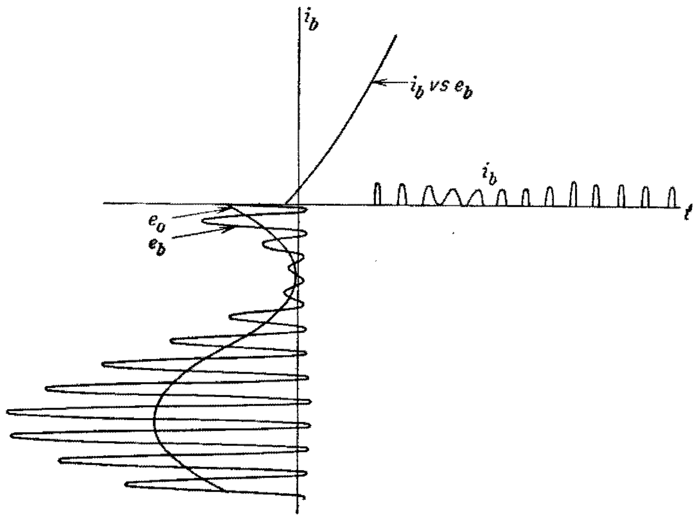


FIG. 12-15. Voltage and current waveforms for a diode detector.

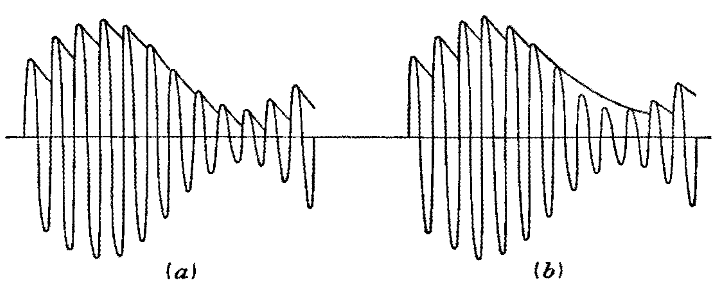


FIG. 12-16. e_o is the envelope of the modulated r-f waveform unless clipping occurs.

until the r-f voltage rises near its positive peak again. Thus in effect the capacitor voltage e_o is a bias as shown in Fig. 12-15, and the total plate voltage for the tube rises just slightly above zero at the positive peaks of the r-f cycle. The figure also shows the plate-current pulses. It should be noted that if the i_b vs. e_b curve for the tube rises steeply, denoting a low plate resistance, the excursions of e_b into the positive region are reduced and the curve of e_o is almost exactly the modulation waveform.

Figure 12-16a shows the manner in which e_o normally varies. Since there are actually many more r-f cycles than are pictured, the saw-tooth pattern is exaggerated in the drawing. Nevertheless, of necessity the voltage e_o must fall enough between r-f peaks so that some tube current always flows at the r-f peak voltage; otherwise when the envelope of the modulated voltage e_1 is falling, the value of e_o cannot fall fast enough to follow the envelope. The rate of fall of e_o is controlled by the value of R_LC, and it can be shown that $\sqrt{1 - m_a^2}/m_a\omega_s$ gives the greatest value of R_LC that can be used and avoid this distortion-producing phenomenon called diagonal clipping. Here ω_s refers to the highest frequency present in the modulating signal.

A graphical determination[3] of the action of the diode detector can be made if we use the tube's rectification characteristics, which are somewhat similar in appearance to triode plate characteristics. These diode characteristics are obtained (see the circuit of Fig. 12-17) by applying to the diode a succession of different fixed alternating voltages E_k in series with a variable direct voltage and measuring the direct current which flows. The alternating voltage need not be of radio frequency, but the bypass capacitor C must have negligible reactance at the frequency used. The characteristics obtained for a type 7A6 tube are shown in Fig. 12-18.

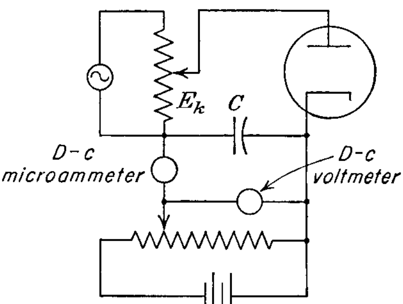


FIG. 12-17. Circuit for obtaining the rectification characteristic for a diode.

In Fig. 12-14, if the cathode of the diode is grounded, then the potential of the lower output terminal is negative relative to ground and it fluctuates at the modulation frequency. With further filtering to remove the modulation-frequency voltage and vestiges of r-f voltage, the resulting direct voltage will depend only upon the magnitude of the r-f carrier applied to the detector (see Fig. 12-15). Hence this direct voltage can be used for biasing the preceding r-f amplifier tubes of a radio receiver. If these tubes are of the remote-cutoff type, their gain will gradually decrease with increase in grid bias and hence the system will constitute an automatic volume control (AVC). To achieve an adequate degree of filtering the circuit of Fig. 12-19 might be used. In this case the load to direct current is approximately 1 megohm, while to a-f currents the load is reduced to approximately 680,000 ohms by the parallel 2-megohm resistor. The 100-$\mu\mu$f capacitors have low reactances for r-f currents only.

In Fig. 12-18 a d-c-load line for 1 megohm has been drawn. If we assume an unmodulated carrier input of 10 volts rms, the operating point is at Q. Through Q we now draw the a-c-load line for 680,000 ohms. If the r-f input voltage is modulated, say, 50 per cent, its rms value will range between 5 and 15 volts and the operation will be between points A and B, giving an a-f output voltage of 13.3 volts peak to peak. If the degree of modulation is made much greater than 50 per cent, the rectified plate current will cease to flow for some portion of the cycle near the nega-

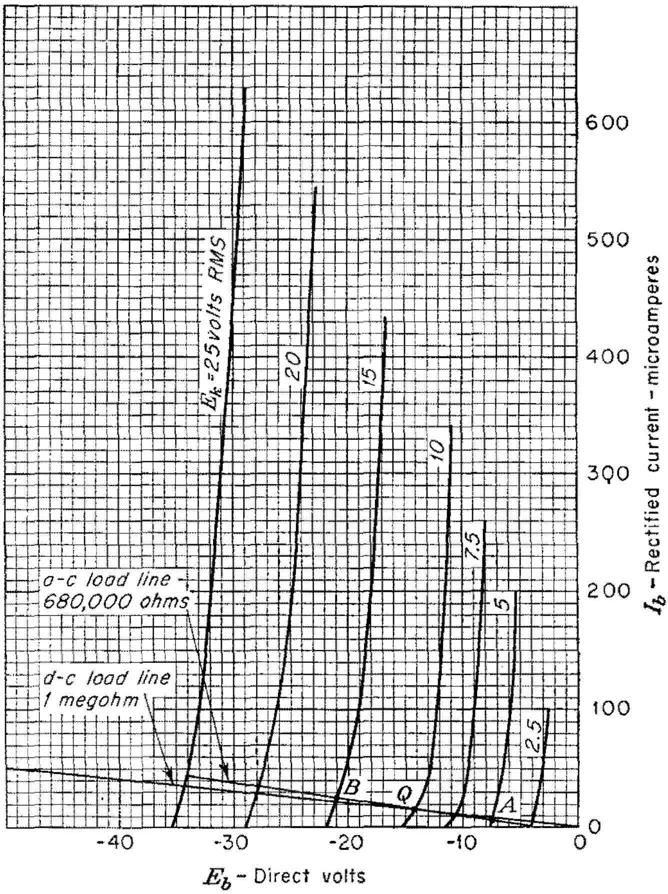


FIG. 12-18. Rectification characteristics for type 7A6 tube (one diode).

tive peak of the modulation envelope and serious distortion results. This occurrence is known as negative-peak clipping. The student should observe that the signal distortion resulting from the detection process can be computed from values taken from the rectification-characteristic curves in the same manner as for an amplifier.

Unlike the square-law demodulator of Art. 12-10, which operates with a negative grid and hence has a very high input impedance, the diode detector draws an appreciable amount of power from its r-f voltage source.

One important result is a broadening of the bandwidth of this preceding r-f stage. We can estimate the input impedance by using power relations. Let us assume in Fig. 12-14 that the plate resistance of the tube is very low so that in Fig. 12-15 the plate voltage swings negligibly positive. Also assume C large enough so that no r-f voltage exists across R_L. The power in the diode circuit is hence dissipated only in R_L, and the voltage across R_L has only two components, one a direct voltage of value E_{km} and the other of modulation frequency and effective value $m_a E_k$. Hence the power dissipated is

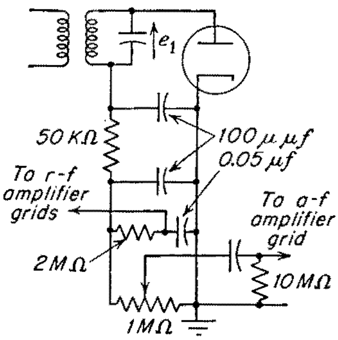


FIG. 12-19. A diode detector and AVC circuit.

$$P = \frac{E_{km}^2}{R_L} + \frac{m_a^2 E_k^2}{R_L} \tag{12-46}$$

The power input to the detector can be computed from the modulated r-f voltage e as given by Eq. (12-8). Each of the three frequency components renders power according to the square of its effective value. Thus

$$P_{\text{in}} = \frac{E_k^2}{R_{\text{in}}} + \frac{E_k^2 m_a^2}{4R_{\text{in}}} + \frac{E_k^2 m_a^2}{4R_{\text{in}}} \tag{12-47}$$

where R_{in} is the resistance seen by the r-f voltage source. Equating powers and simplifying yields

$$\frac{E_k^2}{R_{\text{in}}}\left(1 + \frac{m_a^2}{2}\right) = \frac{2E_k^2}{R_L}\left(1 + \frac{m_a^2}{2}\right) \tag{12-48}$$

from which we obtain the relation

$$R_{\text{in}} = \frac{R_L}{2} \tag{12-49}$$

12-12. Interference in Radio Reception. Interference in the reception of a-m signals is caused by other signals on the same or on adjacent channels and by natural and man-made static. In order to eliminate much of this difficulty, the FCC regulations specify very close tolerances for the carrier frequency, and thus two stations on the same channel have very nearly the same center frequency. If they did not, there would be heard from the receiver a beat note of the difference frequency.

The principal difficulty in interstation interference is the effect of a highly modulated weak carrier on a lightly modulated strong one. Since the frequencies of the two carriers may differ by a small amount, there are times when these carriers are in phase and other times when they are out of phase, these situations occurring at the difference-frequency rate.

As an extreme case let us assume the strong, lightly modulated signal has strength x with zero modulation, while the interfering station has strength y with 100 per cent modulation; $x \gg y$. When the carriers are in phase as in Fig. 12-20a, we see that the total degree of modulation is $y/(x + y) \approx y/x$. Somewhat more than a quarter of a cycle of the difference frequency later, the situation is as shown in Fig. 12-20b, where the resultant signal has the same length as x and hence the modulation is zero. In (c) is shown the situation when the carriers are out of phase, and now the degree of modulation is $y/(x - y) \approx y/x$ again. The effect of all this to the listener is the sound of the modulating signal of the weak station rising to a value proportional to y/x and falling to zero at a rate double the difference frequency, perhaps a few times a second. Since the desired signal when fully modulated would give a constant output proportional to $x/x = 1$, the ratio of the interfering sound to the desired one is y/x, the same as the ratio of the r-f signals. Hence, if it is desirable to keep the interference in audio frequencies below, say, 1 per cent, the r-f carrier from the interfering station must be below 1 per cent of that from the desired station.

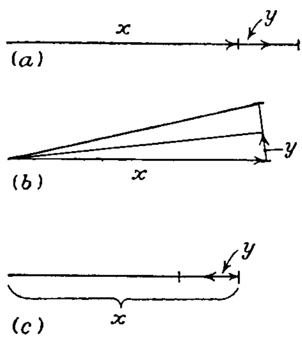


FIG. 12-20. Modulation of a strong unmodulated carrier by a weak a-m carrier.

The effect of random noise in an a-m receiver is much the same. The energy of static is quite uniformly distributed over the r-f spectrum, and hence the noise power picked up by the antenna and passed through the tuned circuits is proportional to the bandwidth of the latter. The noise-voltage amplitude is proportional to the square root of the bandwidth, and this noise voltage combines with the desired signal to add noise modulation to the desired modulation.

There are several proposals that have been made for solving the interference problem. The allocation of station frequencies is done in a manner to place stations operating on the same frequencies at great distances apart. Some are made to operate with directional antennas, which radiate the energy away from a station sharing the same channel. Narrowing the bandwidth by suppressing one of the sidebands is sometimes done, and this improves the signal-to-noise ratio by a factor $\sqrt{2}$. Radio stations could also employ what is called preemphasis. Since the high a-f components in a program are of relatively small amplitude compared with the lower-frequency ones, the former can be amplified before modulation to a greater extent than the latter without the danger of overmodulation occurring. With a device built into each receiver to provide deemphasis, the interference due to tube and circuit noises upon the weak h-f tones would be considerably lessened. This system is not used with a-m transmission since interference with adjacent channels would be increased and the detector would tend to clip more with the high degree of h-f modulation. Furthermore, many existing receivers would need to be fitted with a deemphasis device. The system is used, however, with f-m transmission.

However, a-m stations do follow the helpful practice of operating with as near 100 per cent modulation as practicable. The high-volume passages are lowered in intensity, those of low volume raised in intensity. The resulting compression should be compensated for in the receiver but usually is not. There is therefore some unnaturalness in the rendering.

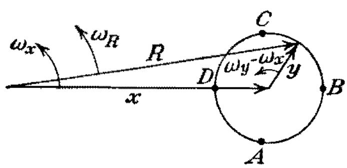


FIG. 12-21. Two unmodulated signals interfere very little in angle modulation.

A system of modulation which would yield signals that would be unaffected by static and by weak interfering modulated signals is fairly well realized in angle modulation. In Fig. 12-21 a desired unmodulated carrier is represented by vector x rotating at ω_x radians per sec, while the vector y representing an undesired carrier, on the same channel but with slightly higher frequency ω_y, rotates about the tip of x at the small difference frequency $\omega_y - \omega_x$ radians per sec. The resultant vector R deviates from the length and position of vector x, and this shows that R is both amplitude- and angle-modulated. Now a receiver designed for angle modulation can be made insensitive to amplitude changes by means of a so-called limiter, and so we need concern ourselves only with the effect of the angle modulation. As the tip of vector y describes a circle through points A, B, C, and D, the value of $\omega_R - \omega_x$ passes through a cycle of values of 0 at A, a maximum positive value of $(\omega_y - \omega_x)y/(x + y)$ at B, 0 again at C, and a minimum (negative maximum) of $(\omega_y - \omega_x)y/(x - y)$ at D. This cycle of values repeats $(\omega_y - \omega_x)/2\pi$ times each second. The greatest deviation in $\omega_R - \omega_x$ is the one occurring at D. The maximum drift in carrier allowed by FCC regulations is 2000 cps. If we suppose the magnitude of the undesired signal y is as great as one-half that of the desired signal x and the difference in carrier frequencies is 2000 cps, then the greatest frequency deviation is 2000 cps, occurring 2000 times each second. If the desired carrier is fully modulated to a deviation of 75,000 cps, then the interfering modulation will be about $\frac{2}{75}$ as strong as the desired one. This is not as bad an interference as would result for a lesser deviation than 75,000 cps, and hence we see the importance of using a high deviation. Since the interfering signal is seldom as strong or the carrier drift as great as suggested in the example, this type of interference is not very troublesome.

At one time it was believed that varying the frequency only a small amount at an audio rate would cut down interference by lowering the operating bandwidth, but as we shall see in the next article, the bandwidth even then is very wide, and no hope lies in that direction. However, with a receiver insensitive to amplitude changes and with a large frequency deviation made at an audio rate together with preemphasis for the higher audio frequencies, a very successful, relatively interference-free communication system is possible.

12-13. The Frequency Spectrum of an Angle-modulated Wave. From Eq. (12-6) we can write the equation for an angle-modulated voltage as

$$e = E_{km} \cos (\omega_k t + m \sin \omega_s t) \tag{12-50}$$

This equation can equally well represent either a frequency- or a phase-modulated wave since the difference in phasing between expressions (12-50) and (12-7) is unimportant. For frequency modulation $m = m_f = K_f E_{sm}/\omega_s$, while for phase modulation $m = m_p = K_p E_{sm}$. In either case m is the maximum deviation in phase angle from the normal value $\omega_k t$ or the ratio of the maximum frequency deviation $\Delta\omega_m$ to the audio frequency ω_s. It is perhaps well to write

$$m = \Delta\phi_m \text{ radians} = \frac{\Delta f_m}{f_s} \tag{12-51}$$

The constant K_f for an f-m transmitter is usually expressed in a practical way by giving the value of $K_f/2\pi$ in kilocycles per second deviation per signal volt. This makes it convenient to determine the maximum deviation for a given audio signal. The relation follows because $K_f/2\pi = m_f f_s/E_{sm} = \Delta f_m/E_{sm}$ kc per volt.

It should be noted that in an f-m system, if the voltage E_{sm} is made proportional to frequency, the result produced is a phase-modulated wave. Likewise in a phase-modulation device, if the value of E_{sm} varies inversely with frequency, the result is frequency-modulated. No matter which system is used, if m is the same, the result is the same.

To find the frequency spectrum of this wave, which obviously is not sinusoidal since its phase angle does not vary linearly with time, we may expand in the form

$$e = E_{km}[\cos \omega_k t \cos (m \sin \omega_s t) - \sin \omega_k t \sin (m \sin \omega_s t)] \tag{12-52}$$

The waveforms of cos (m sin $\omega_s t$) and of sin (m sin $\omega_s t$) are also not sinusoidal since the angles do not vary linearly with time. However, these waveforms are periodic since their values repeat when $\omega_s t$ increases in value by

2π. Therefore, we can expand these expressions by means of a Fourier series. These expansions can be found in books on advanced calculus,[1] and we can write

$$\cos(m \sin \omega_s t) = J_0(m) + 2J_2(m) \cos 2\omega_s t + 2J_4(m) \cos 4\omega_s t + \cdots \quad (12\text{-}53)$$

and

$$\sin(m \sin \omega_s t) = 2J_1(m) \sin \omega_s t + 2J_3(m) \sin 3\omega_s t + \cdots \quad (12\text{-}54)$$

where $J_n(m)$ is a Bessel function of the first kind and order n and defined by an infinite series in powers of m. This series itself is not important to

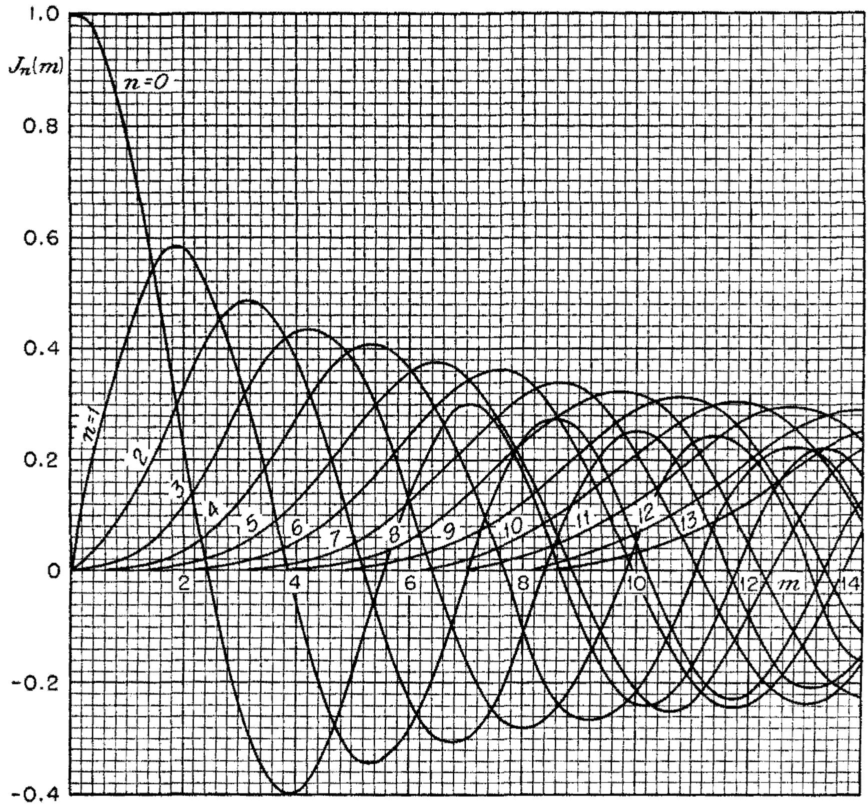


FIG. 12-22. Bessel functions $J_n(m)$.

us at this point since the coefficients can be evaluated from graphs of these Bessel functions given in Fig. 12-22.

It now follows that we can complete the expansion of e begun in Eq. (12-52).

$$e = E_{km} \cos \omega_k t[J_0(m) + 2J_2(m) \cos 2\omega_s t + 2J_4(m) \cos 4\omega_s t + \cdots] + E_{km} \sin \omega_k t[2J_1(m) \sin \omega_s t + 2J_3(m) \sin 3\omega_s t + 2J_5(m) \sin 5\omega_s t + \cdots] \quad (12\text{-}55)$$

By using the identities

$$\begin{aligned} -2 \sin x \sin y &= \cos (x + y) - \cos (x - y) \\ 2 \cos x \cos y &= \cos (x + y) + \cos (x - y) \end{aligned} \tag{12-56}$$

we obtain

$$\begin{aligned} e = E_{km}\{J_0(m) \cos \omega_k t \\ &- J_1(m)[\cos (\omega_k + \omega_s)t - \cos (\omega_k - \omega_s)t] \\ &+ J_2(m)[\cos (\omega_k + 2\omega_s)t + \cos (\omega_k - 2\omega_s)t] \\ &- J_3(m)[\cos (\omega_k + 3\omega_s)t - \cos (\omega_k - 3\omega_s)t] \\ &+ \cdots\} \end{aligned} \tag{12-57}$$

Equation (12-57) shows that the spectrum of e consists of a carrier and an infinite number of sidebands, all spaced ω_s apart. The amplitude of the carrier and of each sideband depends upon the modulation index.

The carrier and the important sidebands constitute the transmission band, and it is important to learn the bandwidth. This may be easily done if we examine Fig. 12-22, which shows graphically several of the Bessel functions. For $m = 1$, when the frequency deviation equals the modulating frequency, we see that $J_0(1) = 0.76$, $J_1(1) = 0.44$, $J_2(1) = 0.11$, while $J_3(1)$ is very small. This means that if the unmodulated carrier is taken as unity, then with a modulation index of 1 the carrier has an amplitude of 0.76, and the first pair of sidebands (frequencies $\omega_k \pm \omega_s$) has amplitudes of 0.44; for the second pair of sidebands (frequencies $\omega_k \pm 2\omega_s$) the amplitudes are 0.11, while higher-order sidebands are of negligible amplitudes. Thus the bandwidth is approximately $4\omega_s$. If we arbitrarily take a value of $J_n(m) > 0.1$ as denoting an important band, we see that for $m = 2$, BW $= 2 \times 3\omega_s = 6\omega_s$ and for $m = 3$, BW $= 2 \times 4\omega_s = 8\omega_s$. For $m = 10$, BW $= 2 \times 11\omega_s = 22\omega_s$. These examples furnish a rough rule that the bandwidth is

$$\text{BW} = 2(m + 1)\omega_s \text{ radians per sec} = 2(m + 1)f_s \text{ cps} = 2(\Delta f_m + f_s) \text{ cps} \tag{12-58}$$

Thus the bandwidth is somewhat more than twice the maximum frequency deviation.

For frequency modulation m_f is inversely proportional to frequency, and hence $\Delta f_m(= mf_s)$ is independent of frequency. Therefore, the bandwidth is approximately the same for all modulation frequencies (equal voltages). However, in the case of phase modulation, m_p is independent of frequency, which makes Δf_m and bandwidth vary linearly with the modulation frequency.

Suppose that the audio signal applied to an f-m transmitter contains a 500-cps component of 1-volt amplitude and a 10,000-cps component of amplitude 0.1 volt. Assume the constant of the transmitter is 75 kc per volt. Then for the 500-cps component, $\Delta f_m = 75 \times 1 = 75$ kc, $m_f = 150$, and BW $= 2(\Delta f_m + f_s) = 2(75 + 0.5) = 151$ kc. For the 10,000-cps component, $\Delta f_m = 75 \times 0.1 = 7.5$ kc, $m_f = 0.75$, and BW $= 2(7.5 + 10) = 35$ kc. If the same audio signal is applied to a phase-modulation transmitter to give the same maximum allowable frequency deviation of 75 kc, it is found by trial that the higher-frequency component determines the modulation indexes that can be used. For the 10,000-cps component, $m_p = \Delta f_m/f_s = 75/10 = 7.5$, BW $= 2(75 + 10) = 170$ kc. For the 500-cps component, m_p is 10 times greater because E_{sm} is 10 times larger; $m_p = 75$. Now $\Delta f_m = 75 \times 0.5 = 37.5$ kc, and BW $= 2(37.5 + 0.5) = 76$ kc.

An examination of the results of these two methods of modulation shows that for frequency modulation the amplitude of the small 10,000-cps component could be preemphasized 10 times before modulation in order to take advantage of the full 75 kc deviation permitted and without exceeding the allowable 200-kc bandwidth. This would greatly improve the signal-to-noise ratio in the output of the receiver. Of course, the preemphasis factor 10 suggested here would not always apply. No such preemphasis of the small h-f audio components is possible for the case of phase modulation. This explains why frequency modulation is preferable to phase modulation as a form for the transmitted wave.

Because a very wide band is used for angle modulation, there is a problem in tuned-circuit design unless the carrier frequency is made very high. For this reason it is common practice to use near 90 Mc and above as carrier frequencies for angle-modulation systems.

12-14. Preemphasis. In order to increase the relative energy in the high a-f spectrum to take better advantage of the 75 kc allowable deviation and thus better to outweigh the effects of tube and circuit noise, it is customary to build into the transmitter a preemphasis device and to provide for each receiver a corresponding deemphasis.

A circuit to provide preemphasis has a characteristic

$$\frac{E_2}{E_1} = \sqrt{1 + \left(\frac{\omega}{\omega_0}\right)^2} \tag{12-59}$$

where E_2/E_1 is the ratio of output to input voltage and ω_0 is a constant chosen on the basis of experience as being a satisfactory value. The value now used in the United States for carrier frequencies near 100 Mc is 2100 cps for f_0. This results in E_2/E_1 being $\sqrt{2}$ at 2100 cps and 10 at 21,000 cps. The effect on lower frequencies is much less; for 1050 cycles it is only 1.1.

12-15. Frequency Modulation Using a Reactance Tube. Probably the simplest way of producing an f-m wave is to vary the capacitance or the inductance of the frequency-determining tank of the oscillator. This can theoretically be done directly by using a condenser microphone or by having a phonograph needle move a capacitor plate or one coil of a mutual-inductor pair. A considerably more satisfactory method, however, is to use a variable-reactance tube in parallel with the tank. This tube can act either as a variable capacitance or as a variable inductor, and hence the same results can be obtained.

Figure 12-23 shows one form of reactance tube connected in parallel with the tank of a Hartley oscillator. The circuit to the right of the capacitor C_1 acts like a capacitor in parallel with C_1. Figure 12-24*a* is the equivalent circuit for the reactance tube alone, with the tank voltage **E**

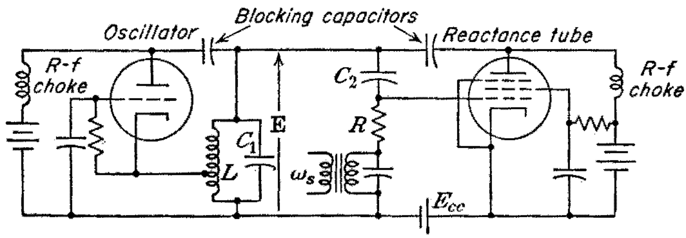


FIG. 12-23. A reactance tube modulating a Hartley oscillator.

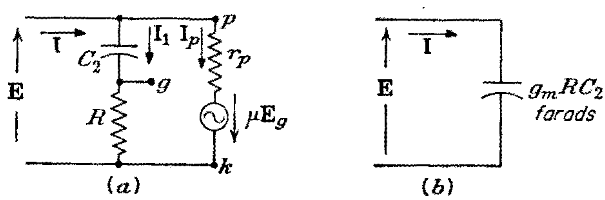


FIG. 12-24. Equivalent circuits for a reactance tube.

applied to it. The path through the supply of E_{bb} is omitted because of the r-f choke. The a-f transformer below R is used to introduce the audio-signal voltage. The secondary is bypassed by a capacitor of small reactance to r-f currents. The transformer is omitted from the equivalent circuit, which involves only r-f currents in the present discussion.

At the oscillation frequency let us suppose that

$$\frac{1}{\omega C_2} \gg R \gg \frac{1}{g_m} \tag{12-60}$$

Then

$$\mathbf{I}_1 \approx j\omega C_2\mathbf{E} \tag{12-61}$$

and

$$\mathbf{E}_g = j\omega RC_2\mathbf{E} \tag{12-62}$$

$$\begin{aligned}\mathbf{I} = \mathbf{I}_1 + \mathbf{I}_p &= j\omega C_2\mathbf{E} + \frac{\mathbf{E} + \mu\mathbf{E}_g}{r_p} = j\omega C_2\mathbf{E} + \frac{\mathbf{E} + j\mu\omega RC_2\mathbf{E}}{r_p} \\ &= j\omega C_2\mathbf{E} + \frac{\mathbf{E}}{r_p} + j\omega g_m RC_2\mathbf{E}\end{aligned} \tag{12-63}$$

If an audio-signal voltage is introduced in series with E_{cc}, the bias on the grid will rise and fall slowly (relative to the r-f variation). The tube parameters will vary with the bias. The first current component in (12-63) will be constant, however. The second will vary as r_p changes, and, being in phase with **E**, it will cause the amplitude of the oscillations to change. With r_p remaining very high (for a pentode), however, this current component and the amplitude modulation it causes are very small. The third component varies with the value of g_m, and since it leads **E** by 90°, we see that the reactance-tube equivalent circuit can be reduced to that of Fig. 12-24*b*, where the other two shunt paths have been neglected.

Now the angular frequency of oscillation for the Hartley oscillator is approximately

$$\omega = \frac{1}{\sqrt{LC}} = \frac{1}{\sqrt{L(C_1 + g_m R C_2)}} \tag{12-64}$$

The steady effect of the average value of g_m times RC_2 is of no consequence since it can be compensated for by adjustment of C_1. However, g_m is cyclic at the modulating frequency. The effect of this on ω can be computed as follows:

With no modulation, $\omega = \omega_0$. With modulation

$$d\omega = -\frac{1}{2}\frac{1}{\sqrt{LC}}\frac{dC}{C} = -\frac{1}{2}\omega_0\frac{dC}{C} \tag{12-65}$$

where

$$dC = d(C_1 + g_m R C_2) = 0 + RC_2\,dg_m \tag{12-66}$$

Hence

$$d\omega = -\frac{1}{2}\omega_0\frac{RC_2}{C}\,dg_m \tag{12-67}$$

If g_m varies linearly with e_s,

$$dg_m = kE_{sm}\cos\omega_s t \tag{12-68}$$

and

$$d\omega = -\frac{1}{2}\omega_0\frac{RC_2}{C}kE_{sm}\cos\omega_s t \tag{12-69}$$

where k depends upon the curve of variation of g_m with e_c. We see by Eq. (12-69) that the frequency changes linearly with the modulation and hence frequency modulation takes place.

An example may be helpful. Suppose it is desired to frequency-modulate a 90-Mc carrier with a deviation of 72 kc 5000 times a second. It is impractical to operate a reactance tube at this high radio frequency, and so we may use a series of frequency multipliers as shown in the block diagram of Fig. 12-25. These frequency multipliers are class C amplifiers, which have high amplitude-distortion components, and the plate tank is tuned to the harmonic desired. The nonlinear gain characteristic is not objectionable since any variation of the amplitude of the output is of no importance.

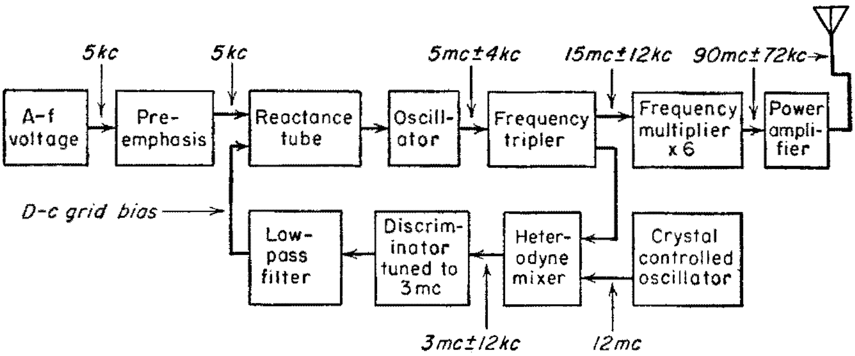


FIG. 12-25. A reactance-tube frequency-modulator system with a frequency stabilizer.

Let us suppose the oscillator and the reactance tube to be those of Fig. 12-23 and the latter tube to be a type 6SJ7. Figure 12-26 shows a graph of g_m vs. e_c for this tube. If we choose $E_{cc} = -3.5$ volts, it is seen that g_m will vary fairly linearly with e_s (which is an a-f e_g) according to the equation $g_m = (1400 + 430E_{sm} \cos \omega_s t) \times 10^{-6}$ mho, and the maximum change in g_m will be $430 \times 10^{-6}E_{sm}$. Hence from Eq. (12-67) we can write

$$d\omega = \frac{1}{2}\,\omega_0\,\frac{RC_2}{C}\,E_{sm} \times 430 \times 10^{-6} \tag{12-70}$$

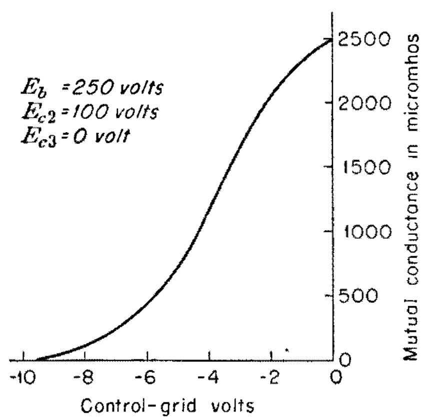


FIG. 12-26. Variation of g_m with grid bias for a 6SJ7.

In the reactance-tube circuit let us assume some values. Since $1/g_m = 1/(1400 \times 10^{-6}) = 714$ ohms, we can assume [see relation (12-60)] $R = 10{,}000$ ohms. $1/\omega_0 C_2$ should be much greater than R, and we can assume a value of, say, about 100,000 ohms, making $C_2 = 3\ \mu\mu\text{f}$. $C = 200\ \mu\mu\text{f}$ should be a suitable value.

Continuing from Eq. (12-70), we can solve for the necessary value of E_{sm} to produce a frequency deviation of 4 kc (see Fig. 12-25); thus

$$E_{sm} = \frac{2\,d\omega C \times 10^6}{430\omega_0 C_2 R} = \frac{2 \times 2\pi \times 4000 \times 200 \times 10^{-12} \times 10^6}{430 \times 2\pi \times 5 \times 10^6 \times 3 \times 10^{-12} \times 10{,}000} = 0.025 \text{ volt} \tag{12-71}$$

The oscillator which produces the f-m waveform cannot be crystal-controlled, and hence its carrier is too unstable to satisfy the legal requirements unless some sort of stabilizer is used. In Fig. 12-25 is shown the diagram of one form of stabilizer. The somewhat unstable 15 Mc output of the first frequency tripler is fed into a heterodyne mixer along with the very stable output of a crystal-controlled oscillator operating at some fixed frequency, say 12 Mc. The output of the mixer is 3 Mc plus or minus the unstable variation in the carrier, along with a 12-kc frequency deviation caused by the signal. The discriminator, which is to be studied later (Art. 12-19), is tuned to 3 Mc, and as long as the input frequency is centered at this value, the discriminator output will be 5000 cps, the audio modulation, but with zero d-c component. However, if the carrier frequency wavers from 3 Mc, the discriminator output then contains a

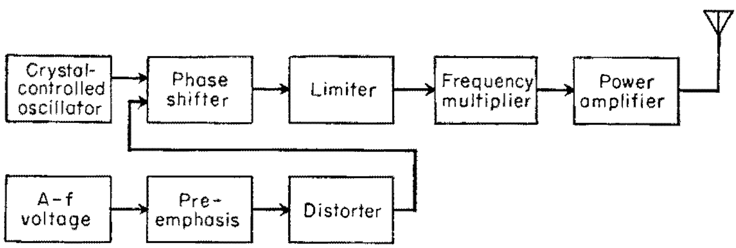


FIG. 12-27. Block diagram of an f-m transmitter using phase modulation.

d-c component with polarity depending upon whether the carrier frequency is too high or too low. A filter removes the a-f variations from the discriminator output, and the remaining direct voltage is applied as a bias in series with E_{cc} to the reactance-tube grid, its effect there being to change the average g_m enough to retune the oscillator nearer 5 Mc.

12-16. Frequency Modulation by Means of Phase Modulation. There are numerous methods by which the phase of a carrier can be shifted at an a-f rate. If then a carrier is generated under the control of a crystal and its phase advanced or retarded in accordance with the magnitude of a signal voltage, a phase-modulated wave is produced.

Equation (12-57) gives the frequency spectrum of either a frequency- or a phase-modulated voltage. It was pointed out that m_p is constant relative to signal frequency while m_f varies inversely as the signal frequency. With this sole difference between the two it is easy to predistort the audio input to the phase shifters of the preceding paragraph in a manner to make the a-f voltage vary inversely as ω_s, and then the r-f output of the shifter is frequency-modulated. Figure 12-27 gives a block diagram of this method. The limiter is discussed in Art. 12-18. It is used here to remove the effects of any amplitude modulation that may occur in the phase-shifting process. Since the phase shifter can usually vary the phase only a small amount compared with that desired in the output, the modulation process is carried out at a frequency much lower than the transmission frequency. Frequency multipliers then raise the carrier frequency and at the same time increase the frequency deviation in the same proportion. It should be noted that frequency multiplication does not change the frequency separation between the sidebands, which depends upon the rate at which the frequency change in the r-f voltage was made, *i.e.*, upon the audio frequency, and this is not altered by frequency multiplication.

12-17. F-M Receivers. Frequency-modulated receivers are usually of the superheterodyne type and are quite similar to those used for the reception of a-m signals. There are a few points of difference, some of them very important.

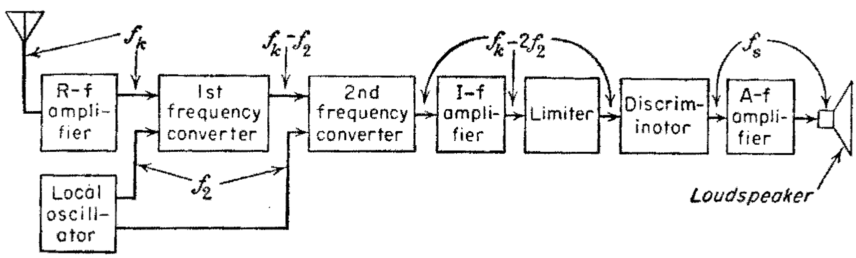


FIG. 12-28. A block diagram of an f-m receiver.

First, the freedom from interference means that much weaker signals can be successfully handled. Thus in the better receivers the i-f gain is usually much greater than for an a-m receiver. Since with increased gain comes more danger from positive feedback, it is common to use two intermediate frequencies because feedback of a different frequency does not result in oscillation. Thus two frequency converters may be used; part of the i-f amplification can take place at one frequency and part at another.

Second, the very great bandwidth of up to 200 kc makes it impossible to use the ordinary 456-kc intermediate frequency as in amplitude modulation (broadcast). It is usual to use near 5 Mc for this purpose. Then a 200-kc bandwidth would be similar to an a-m bandwidth of about 20 kc, which is not too difficult to obtain.

Third, the freedom from interference depends partly upon making the receiver insensitive to amplitude changes. The f-m receiver therefore has incorporated in it a limiter for producing a limiter output voltage of a constant amplitude.

Fourth, the f-m receiver needs a discriminator, which gives an a-f output corresponding to deviations in the radio frequency.

Figure 12-28 shows a block diagram of an f-m receiver.

12-18. Limiters. There are several types of amplitude limiters. Many of them are essentially triodes or pentodes operating with low bias voltages and with a high grid drive. On the negative grid swings the plate current cuts off, and on the positive swings it saturates.

A common type of pentode limiter which differs somewhat in its action from the above is shown in Fig. 12-29. The action depends upon the grid bias produced by the grid-current flow through R_g and is illustrated in Fig. 12-30*a*. The values of plate and screen biases are low so that the control-grid voltage required for plate-current cutoff is small. With a large r-f grid signal applied as with cycle 1, the grid voltage swings just enough in the positive direction to make the grid bias $R_g I_{ca}$ almost equal to the peak value of the r-f voltage. For an r-f wave of greater amplitude the situation is shown by 2, where the grid swing is slightly more positive to make the grid bias greater. Also shown are pulses of grid and of plate current under each circumstance.

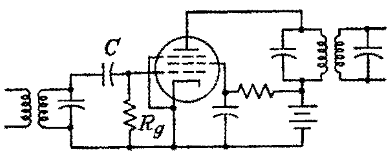


FIG. 12-29. A pentode limiter.

Let us observe the enlarged drawing in Fig. 12-30*b*, in which the plate-current pulses for cycles 1 and 2 are superimposed for comparison. As shown by the timing wave cos $\omega_k t$, each current pulse has a duration con-

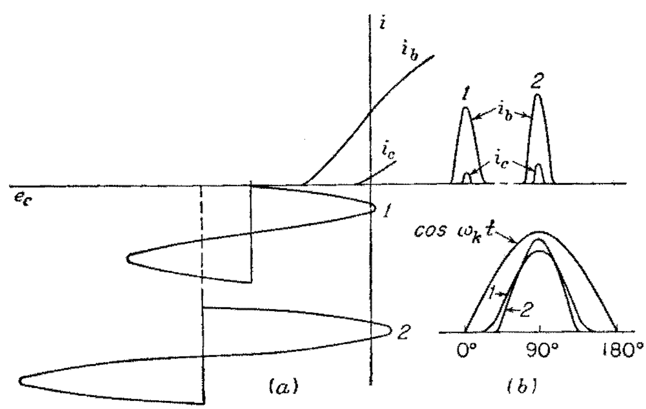


FIG. 12-30. Waveforms of e_c, i_c, and i_b for the limiter.

siderably less than 180°; that for wave 2 is higher but of shorter duration than that for wave 1. The fundamental component of current is given by

$$I_{p_1m} = \frac{2}{2\pi}\int_0^{2\pi} i_b \cos \omega_k t \, d(\omega_k t) \approx \frac{2}{2\pi}\int_0^{2\pi} i_b \, d(\omega_k t) \qquad (12\text{-}72)$$

since the cosine is near unity during the period when current flows. The last integral is the area under the current wave, and as both pulses tend to have nearly the same areas, the fundamental current through the plate load is about the same in both cases.

12-19. Discriminators. The circuit for a commonly used type of discriminator is illustrated in Fig. 12-31. An analytical explanation of its

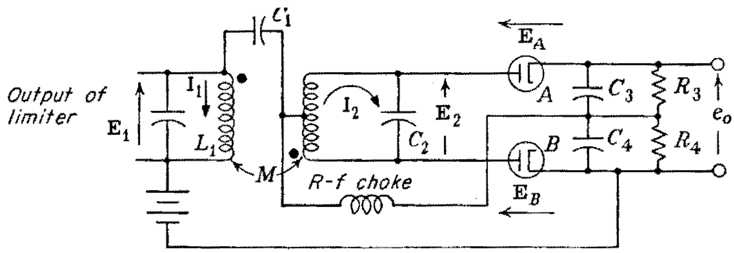


FIG. 12-31. A discriminator circuit.

action is quite involved, and we shall proceed with a graphical one. The limiter tank and the transformer secondary are each tuned to the intermediate frequency. The output voltage E_1 across the limiter tank is frequency-modulated and is centered at the intermediate frequency.

When the frequency of E_1 is exactly at the center value, the nature of the output voltage e_o can be determined as follows: $I_1 = E_1/j\omega L_1$. The voltage induced in the secondary with positive sense upward is $E' = -j\omega M I_1 = -(M/L_1)E_1$. The reactance of the r-f choke is large compared with X_{C_2}, and we assume, therefore, that the current I_2 flows through C_2 and is in phase with E' because of resonance. Hence E_2, the voltage drop across C_2, lags E' by 90° and leads voltage E_1 by 90°. The r-f voltage E_A applied to diode A is $E_1 + \frac{1}{2}E_2$ if the capacitors C_1, C_3, and C_4 are chosen to have small reactances at this frequency. Likewise the voltage E_B applied to diode B is $E_1 - \frac{1}{2}E_2$. These quantities E_1, E', I_2, E_2, E_A, and E_B are drawn in vector form in Fig. 12-32a. Since the r-f voltages E_A and E_B have like amplitudes, the two diode currents have equal average values and the direct voltages across R_3 and R_4 are equal in size but opposite in sense.

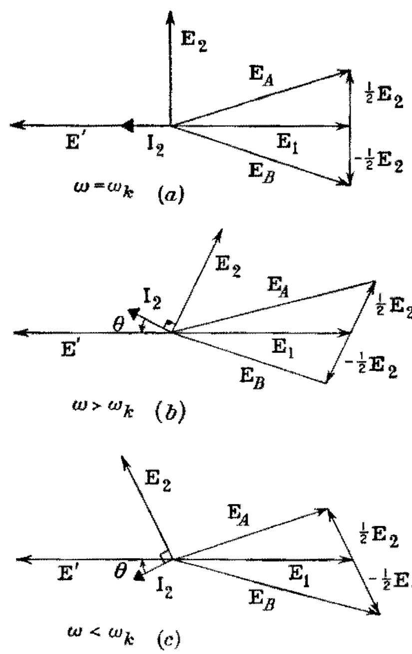


FIG. 12-32. Vector diagrams for the discriminator.

Hence there is no voltage developed between the output terminals, and $e_o = 0$.

Let us next consider the situation when ω is greater than the intermediate carrier frequency ω_k. The expression for $\mathbf{I}_1$ still holds, and $\mathbf{E}' = -(M/L_1)\mathbf{E}_1$ as before. However, the secondary circuit is now inductive, and hence $\mathbf{I}_2$ lags $\mathbf{E}'$ by some angle θ. θ depends upon the amount by which ω exceeds ω_k and upon the Q of the secondary circuit. It follows that $\mathbf{E}_2$, the voltage drop across C_2, lags $\mathbf{E}'$ by an angle $90° + \theta$. We can now determine the $\mathbf{E}_A$ and $\mathbf{E}_B$ vectors as shown in Fig. 12-32*b*. Since $\mathbf{E}_A$ exceeds $\mathbf{E}_B$ in amplitude, it follows that e_o is now positive.

When ω is less than ω_k, it can be shown that $\mathbf{I}_2$ leads $\mathbf{E}'$ by an angle θ and the vector diagram of Fig. 12-32*c* applies. Now $\mathbf{E}_B$ exceeds $\mathbf{E}_A$ in amplitude, and as a consequence e_o is now negative.

The variation of e_o is at the modulation frequency, and if correct circuit adjustments are made, e_o is a good reproduction of e_s.

It should be noted that if $\mathbf{E}_1$ is not modulated but its frequency varies from ω_k because of oscillator instability, the discriminator circuit gives a direct output voltage proportional to the amount of this variation and of polarity dependent upon whether the frequency is too high or too low. This is the application of the discriminator noted in the discussion of the frequency stabilizer in Art. 12-15.

REFERENCES

1. Woods, F. S.: "Advanced Calculus," p. 281, Ginn & Company, Boston, 1926.
2. Herold, E. W.: The Operation of Frequency Converters and Mixers for Superheterodyne Reception, *Proc. IRE*, **30**, 84 (1942).
3. Ballentine, S.: Detection of High Signal Voltages, *Proc. IRE*, **17**, 1153 (July, 1929).
4. Arguimbau, L. B.: "Vacuum Tube Circuits," John Wiley & Sons, Inc., New York, 1948.
5. Armstrong, E. H.: A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation, *Proc. IRE*, **24**, 689 (May, 1936).
6. Cruft Laboratory, War Training Staff: "Electronic Circuits and Tubes," McGraw-Hill Book Company, Inc., New York, 1947.
7. Eastman, A. V.: "Fundamentals of Vacuum Tubes," 3d ed., McGraw-Hill Book Company, Inc., New York, 1949.
8. Everitt, W. L.: Frequency Modulation, *Trans. AIEE*, **59**, 613 (1940).
9. Reich, H. J.: "Theory and Applications of Electron Tubes," 2d ed., McGraw-Hill Book Company, Inc., New York, 1944.
10. Roder, H.: Amplitude, Phase and Frequency Modulation, *Proc. IRE*, **19**, 2145 (December, 1931).
11. Ryder, J. L.: "Electronic Fundamentals and Applications," Prentice-Hall, Inc., New York, 1950.
12. Seely, S.: "Electron-tube Circuits," McGraw-Hill Book Company, Inc., New York, 1950.
13. Terman, F. E.: "Radio Engineering," 3d ed., McGraw-Hill Book Company, Inc., New York, 1947.

14. Wheeler, H. A.: Design Formulas for Diode Detectors, *Proc. IRE*, **26**, 745 (June, 1938).

PROBLEMS AND QUESTIONS

1. A carrier voltage has a peak amplitude of 10 volts and a frequency of 1×10^6 cps. (*a*) The amplitude of the r-f voltage is modulated by a sinusoidal 1000-cps signal to make the amplitude of the r-f wave vary between 7.5 and 12.5 volts. Write an equation for the modulated wave. Determine m_a. (*b*) The frequency of the r-f wave is varied sinusoidally between 0.925 and 1.075 Mc, 1000 cps. Write an equation for the modulated wave. Determine the maximum phase shift and m_f. (*c*) The phase of the r-f voltage is shifted sinusoidally from 75 radians leading to 75 radians lagging 1000 cycles each second. Write an equation for the modulated voltage. Determine the maximum and minimum frequencies and m_p.

2. The modulation processes of Prob. 1 are carried out using the same devices but with a 1000-cps modulating voltage of half the former amplitude. Determine m_a, m_f, and m_p.

3. The modulation processes of Prob. 1 are carried out using the same devices and with a modulating voltage of the same amplitude but of frequency 500 cps. Determine m_a, m_f, and m_p.

4. An antenna is supplied with 2160 watts of r-f power, 40 per cent amplitude-modulated. Determine the carrier power and the sideband power.

5. Show that there is a reduction in the d-c magnetization of the core of the modulation transformer of Fig. 12-5 if the marked polarity is used.

6. Draw the circuit diagrams for a plate-modulated class C amplifier using two parallel triodes and a push-pull triode modulator.

7. In the circuit of Prob. 6 all tubes are identical, and each has an allowable plate dissipation of 1 kw. If the plate efficiency of the class C tubes is 75 per cent, of the class B modulator tubes 60 per cent, and of the modulation transformer 75 per cent, what total r-f power, modulated 50 per cent, can be obtained from the class C amplifier?

8. The modulator in Prob. 7 is changed to class A push pull with 25 per cent plate-circuit efficiency. What r-f power output is now available?

9. In a certain transmitter the required plate-to-plate load for the push-pull modulator is 1600 ohms. The value of E_{bb} for both the modulator and the class C stage is 2000 volts. With grid excitation applied to the class C amplifier but with no modulation, the source of E_{bb} supplies 2.5 amp to the class C amplifier. What turns ratio should be used for the modulation transformer?

10. A square-law triode has the carrier voltage applied in the grid circuit, the modulating voltage applied in the plate circuit. The plate tank is tuned to the carrier frequency. Prove that the tank voltage is amplitude-modulated.

11. For the balanced modulator of Fig. 12-12, $e_1 = E_{km} \cos \omega_k t$ and $e_2 = E_{sm} \cos \omega_s t$. Determine the frequencies present in the output e_3.

12. Resketch the circuit of Fig. 12-13 to make a converter circuit. Use a Hartley-type oscillator with cathode, grid 1, and grids 2 and 4 used as cathode, grid, and plate, respectively. The oscillator plate operates at r-f ground potential.

13. The circuit of Fig. 12-13 is operated with $E_{bb} = 250$ volts, $E_{cc2} = E_{cc4} = 100$ volts, $E_{cc3} = -2.0$ volts. $R_g = 20{,}000$ ohms. The local oscillator voltage is adjusted to a value which makes the grid 1 direct current 0.5 ma. (*a*) If a carrier voltage of 0.1 volt is applied to grid 3, what is the i-f component of the plate current (use a tube manual)? (*b*) The carrier voltage to grid 3 rises to 0.2 volt, and AVC action shifts the bias of this grid to -6 volts. What is the new value of the i-f component of plate current?

14. Draw the circuit for a square-law modulator used as a detector of a-m signals.

Specify the type tube to be used, the bias voltage values, and the character of the load-impedance elements.

15. A certain diode detector circuit is like that of Fig. 12-19 except that the 50-kilohm resistor is omitted, a 100-kilohm volume control is used, and the AVC-filter resistor is 100 kilohms. The single diode is type 7A6. Use a tube manual. (*a*) Draw the d-c- and a-c-load lines. (*b*) For an r-f signal $e_1 = 15\sqrt{2}\,(1 + m_a \cos 2\pi \times 1000t) \cos 2\pi \times 10^6 t$ determine the greatest allowable value of m_a to avoid negative-peak clipping. (*c*) For this value of m_a determine the greatest a-f voltage output to the next amplifier grid. (*d*) Determine the direct voltage available for automatic volume control.

16. Compute the approximate value of R_{in} for the circuit of Fig. 12-19 if e_1 is 50 per cent modulated. Note that the d-c and a-c loads are unequal.

17. $e = 10 \cos (2\pi \times 10^8 t + 20 \cos 2\pi \times 10^3 t)$. (*a*) What is the maximum phase deviation? (*b*) What is the maximum frequency deviation, and at what rate does it deviate? (*c*) What is the bandwidth?

18. An audio signal consists of a 10-volt 1000-cps component and a 0.5-volt 5000-cps component. (*a*) Determine m for each component if phase modulation is used with a maximum frequency deviation of 75,000 cps. Find the bandwidth. (*b*) Repeat, using frequency modulation. What preemphasis factor can be used for the 5000-cps component without increasing the total bandwidth?

19. The voltage e in Prob. 17 is applied to the input of a heterodyne frequency converter. The local-oscillator frequency is 45 Mc. If the plate load is tuned to give an output voltage lower in frequency than that of e, what is the form of the equation for this output voltage?

20. The voltage e in Prob. 17 is applied to the input of a frequency doubler. (*a*) What is the form of the equation for the output voltage? (*b*) What is its bandwidth? (*c*) Is the side-frequency separation affected?

21. In Fig. 12-23 the positions of R and C_2 are interchanged. Show that the action of the reactance tube is that of a variable inductor. State the simplifying assumptions needed to obtain this result.

CHAPTER 13

CONDUCTION THROUGH GASES

13-1. Introduction. Thus far most of our discussions have treated tubes as though the vacuum were perfect. However, it is physically impossible to remove all the gas from within the tube envelope; in practice enough can be removed so that the residual has negligible effect on the tube characteristics. We call such tubes high-vacuum tubes. The pressures encountered in them may be as low as or even lower than 10^{-7} mm Hg (normal atmospheric pressure at sea level is measured as 760 mm Hg).

There are other types of electronic tubes in which the gas pressures are high enough so that the electrical characteristics are entirely different from those of the high-vacuum types. The pressures encountered in such tubes may range from several atmospheres (high-pressure arc lights) down through pressures of about 10 cm Hg (neon signs, fluorescent lights, and tungar rectifiers), on down through 0.1 mm Hg (mercury-arc rectifiers, thyratrons, and ignitrons), and finally to the low pressures encountered in high-vacuum tubes.

Our first discussion will concern the electrical characteristics of an electronic tube as the pressure is increased from that in a high-vacuum tube to that encountered in mercury arcs, thyratrons and ignitrons. We shall then take up the reasons for these characteristics.

There is always some ionization of the residual gas in a high-vacuum tube. In the normal high-vacuum diode the principal effect is to cause positive-ion bombardment of the cathode, which under certain circumstances may damage the emitting surface. If a small additional amount of gas were admitted to a high-vacuum thermionic diode, its characteristics would be considerably changed. The gas molecules would tend to interfere with the free passage of the electrons from the cathode to the anode. The collisions between the electrons and the molecules of the gas would result in an increased production of positive ions. These positive ions, being in the midst of a host of electrons, would tend to counteract the effects of the space charge. This would result in an increased anode or plate current for a given plate voltage. A second effect would be an increased bombardment of the cathode by the positive ions. Third, there would be a decrease in the inverse-peak voltage that the tube could withstand.

If the gas pressure were now gradually increased, these effects would become more pronounced. A point would finally be reached where there would be breakdown into a visible glow. The anode voltage would then be found to be constant at some low value and independent of the anode current. Thus the characteristics of such a tube are entirely different from those of a high-vacuum type.

13-2. Types of Gaseous Discharges. Three types of electrical discharges are encountered in gaseous tubes. The first, which usually emits no visible glow, is called a Townsend discharge and is characterized by a rather high voltage drop across the tube and a very small current flow (usually less than a few microamperes). It occurs in cold-cathode tubes and is the phenomenon which may lead to the second type. The latter is the glow discharge, which is characterized by a somewhat larger current, usually ranging from a few milliamperes up to hundreds of milliamperes. The tube drop, which is fairly constant for a given tube, may range from about 75 volts up to a few hundred volts. The third type is the arc-discharge. This type is characterized by currents which may range from a few milliamperes up to hundreds and even thousands of amperes. The tube drop will again be fairly constant and may range from about 7 volts up to about 50 volts.

13-3. Physical Properties of the Atom.[1] In order better to understand the reasons for the characteristics of gaseous tubes, we shall review a few of the physical properties of the atom. As explained in Chap. 2, an atom consists of a positive nucleus about which electrons are revolving in orbits. The more energy an electron has, the larger the radius of its orbit. Also, an electron never goes from one orbit to another without either absorbing or giving up energy. In the lowest energy level, which is closest to the nucleus, there is room for 2 electrons; in the next higher level there is room for 8, room for 18 in the next, etc. Each level is able to accommodate a certain maximum number and no more. In the normal atom each level is completely filled before any electrons appear in the next higher level. Thus in the normal atom all electrons are revolving in orbits as close to the nucleus as they can get, with the lowest-energy electrons closest to the nucleus.

13-4. The Excited Atom. If the right amount of energy is given to one of the higher-energy electrons, it may leave its normal orbit and go to one of the higher-energy levels. This amount of energy is called a quantum. After arriving at the higher level the electron may remain there for a very short period of time. Such an atom, with one or more electrons raised to a higher than normal energy level, is said to be excited. The minimum amount of energy, expressed in electron-volts, needed to raise an electron up to a higher-energy state is known as the minimum excitation potential. Experiments and theory prove that this state can exist for only about 1×10^{-8} sec, after which the excited electron must fall back to its normal energy level, giving up its added energy as electromagnetic or light radiation. The wavelength of this radiation depends on the difference in the energy levels of the two orbits. It is this process which produces the visible light in a "neon" sign.

Although the above situation is the usual one, some atoms appear to have certain energy levels to which electrons may be raised but from which they cannot return. Such an excited atom is said to be in the metastable state. The molecule containing this extra-energy atom floats about until perhaps it strikes some solid object, whereupon it becomes normal again and the energy releases an electron from the object or the energy is converted into heat. Or it itself may be struck by an electron and become ionized. In the latter case ionization has taken place because of two collisions, a phenomenon important in certain gas-tube characteristics.

13-5. Ionization. The complete removal of an electron from an atom is known as ionization. The only way in which an electron can be removed from an atom is for it to obtain from some external source sufficient energy to carry it beyond any of the allowable energy levels. The amount of energy necessary to separate the highest-energy electron is known as the minimum ionization energy. That part of the atom which remains after ionization occurs is known as a positive ion. It has a charge equal to that of an electron, and hence it will be acted upon by other charged bodies and by electric fields. The mass of the positive ion is practically that of the original atom since the mass of the electron is negligible when compared with the mass of an atom. Negative ions also exist. In these an extra electron attaches itself to the orbital system of an atom. These negative ions apparently play a very minor role in gaseous discharge.

13-6. Collision Processes in Gases.[2] When charged bodies in a partial vacuum are moving about because of an electric field, there are going to be collisions between the charged bodies and the gas molecules. Some collisions will not result in a transfer of energy, while others will. The first type is known as an elastic collision and the latter as an inelastic one.

The frequency with which these collisions occur will depend upon the pressure of the gas, which in turn depends upon the number of gas molecules per unit volume. Since the electric field will accelerate any charged body, the energy of the collisions will depend upon the distance each charged body travels before it takes part in a collision. When an inelastic collision occurs, there are three possible results. First, the charged body may give up its kinetic energy to the molecule of gas, which in turn may merely acquire more kinetic energy; second, it may be excited;

or, third, it may become ionized. When the latter action occurs to a sufficient extent, the gas breaks down into a visible glow discharge.

The molecules of any gas have a random distribution in space, and a free electron may be located at any point in the gas. The distance each electron travels between collisions may vary from almost zero up to a high value. The average of the distances traveled by a great number of electrons is called the mean free path; it will vary with the gas pressure—the higher the pressure, the shorter the mean free path. If the field is strong enough so that an electron gains sufficient energy to cause ionization in traveling a distance near the value of the mean free path, then copious ionization occurs and the gas may break down into a glow discharge. If the energy is not sufficient, then the current through the tube will increase slightly but there will be no detectable glow. The small amount of ionization that does take place will be due to the few electrons which travel a distance greater than the mean free path and thereby acquire sufficient energy to cause ionization. It is therefore necessary to have a relatively long mean free path, or to use a very high anode voltage, in order to establish a gaseous discharge. However, there is a limit to the mean free path. If it becomes greater than the distance between electrodes, the chance for ionizing collisions becomes very small and the tube will act as a high-vacuum tube. On the other hand, if the mean free path is too short, the collisions occur before the electrons have attained enough energy to ionize the gas.

13-7. Neutralization of Negative Space Charge by Positive Ions. One of the most important effects of ionization is the neutralization of the negative space charge by the positive ions. If at any instant the number of positive ions in the interelectrode space is equal to the number of electrons, the net space charge is zero and the characteristics of such a tube will differ greatly from those of a similar tube with a high vacuum.

Because of its larger mass the velocity of a positive ion is much lower than the velocity of an electron with the same amount of energy. This means that the positive ion will take a much longer time to travel a given distance than will the electron and hence will remain in the interelectrode space longer. As a result a relatively few positive ions can neutralize the space charge set up by a large electron current.

In addition to the cancellation of space charge, the positive ions have an important effect on the walls of the tube and on the electrode surfaces. The positive ions may be attracted toward these surfaces and are neutralized when they strike them. Their excitation and ionization energies, as well as their kinetic energy, are given up there. This energy may cause heating of the surface, or it may be given off in some form of radiation. It may cause sputtering, which is the removal of small particles of the cathode surface. Or it may cause the removal of electrons from the surface. At this point it might be well to mention that all the electrons causing current flow through a gas tube must come from the cathode, or at least for every electron that enters the anode one must be given off by the cathode. In the case of the cold-cathode tube these electrons are given the energy necessary for emission by the bombarding positive ions. Hence the tube drop must be at least large enough to give the bombarding ions the necessary energy. As a result the drop across a glow discharge will be 75 to several hundred volts.

13-8. Gaseous Discharges.[2] If a simple tube, such as pictured in Fig. 13-1, is constructed with two parallel plane electrodes, each a few centimeters in area and spaced a few centimeters from each other, and if

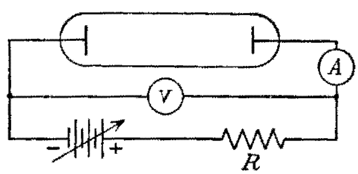


FIG. 13-1. Circuit for obtaining the volt-ampere characteristic of a cold-cathode gas discharge.

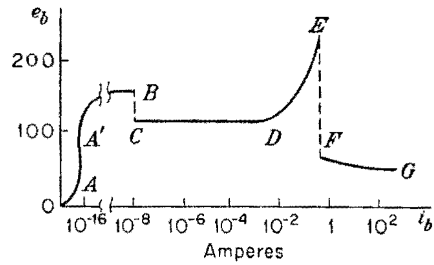


FIG. 13-2. Volt-ampere characteristics of a cold-cathode gas-discharge tube.

a variable voltage is applied between them through a resistor, a current-voltage characteristic such as shown in Fig. 13-2 will result. In the region OA the current increases with increasing voltage but soon reaches a saturation value where the current remains essentially constant over a considerable range of voltages. This current is so small that an ordinary galvanometer cannot be used to measure it. Beyond point A' the current again increases with increasing voltage. The rate of rise soon becomes so great that the characteristic becomes practically a horizontal line. The tube will then break down into a visible glow discharge. The region between A' and B represents what is known as the Townsend discharge, which is of importance because it is the mechanism by which a glow discharge is initiated in cold-cathode tubes.

At point C the tube has broken down into a visible glow. Between points C and D is the normal glow region, and here we see that the voltage is independent of the current. Between D and E the voltage across the tube increases with increasing current, and this is known as the abnormal glow region. At point E the tube breaks down into another type of discharge known as an arc which between F and G has the typical negative-resistance characteristic.

13-9. The Townsend Discharge.[3] Referring back to the region OA of Fig. 13-2 the question arises as to the source of the current which flows.

There are several possibilities. Any gas is in a state of partial ionization because of the bombardment of the gas molecules by cosmic rays. These rays originate somewhere in outer space and are continually bombarding the earth. They penetrate enormous thicknesses of most materials, and so it is practically impossible to prevent them from bombarding the gas molecules in a tube. Another possibility is the radiation from radioactive materials which may be present in very minute quantities in the electrodes or walls of the tube. Still another possibility is photoelectric emission from one of the electrodes. When voltage is applied to the tube, the free electrons, produced by one or more of the above processes, are attracted to the anode and hence result in current flow. However, this current is very small and soon reaches a saturation value. The movement of the electrons toward the anode and of the positive ions toward the cathode results in numerous collisions. But because of the low voltage applied no ionization by collision occurs. However, as the voltage across the tube is increased, the energies of the collisions increase, and a point will finally be reached where more positive ions and free electrons are produced. This results in an increased current and is the beginning of the Townsend discharge. As the voltage is increased still more, the collisions become more violent, resulting in an increased current.

One important fact about this discharge is that if the primary source of electrons (cosmic rays, etc.) is removed, the discharge ceases. In other words, this Townsend discharge is not self-maintaining. Now if the voltage of the battery in Fig. 13-1 is further increased, the current will finally increase very rapidly, the voltage drop across the tube will fall, and the region between the electrodes will emit a glow. It is interesting to examine the phenomena which are taking place. Let us suppose that a single electron is produced near the cathode, say by the action of a cosmic ray. The positive anode causes the electron to be accelerated, and if the gas pressure is high enough, the electron will collide with a gas molecule. For a Townsend discharge, the energy of this collision is sufficient to ionize the molecule, and thus one or more electrons are thereby freed to move on to the anode, perhaps colliding again and freeing more electrons, and so on. Meanwhile the positive ions remaining from these collisions progress toward the cathode, where they will eventually join with electrons on the surface to become neutral gas molecules again. They possibly collide with gas molecules on the way. If the energy of the moving ions is insufficient to knock free electrons from the cathode or to ionize molecules near the cathode surface, then it appears that cutting off the cosmic ray will end the discharge. On the other hand, suppose the energy of either type of collision is sufficient to free electrons which can be accelerated toward the anode. Then it is seen that the discharge perhaps can maintain itself. Whether this last action takes place or not depends partly upon the value of anode voltage, and thus as this voltage is raised, a self-maintaining discharge is finally attained and breakdown is said to have occurred. This breakdown point can be detected by the visible glow and by the sudden decrease in the anode voltage.

13-10. Breakdown.[4] If the breakdown characteristic is studied in more detail, several interesting facts come to light. Figure 13-3 shows the breakdown voltage as a function of the gas pressure. At very low pressures this voltage is extremely high. For a perfect vacuum it would be infinite since there would be no gas present to support a discharge. As the pressure increases, the number of gas molecules per unit volume of interelectrode space also increases. There are, then, more collisions of the free electrons with the gas molecules, hence more ionization by collision and more current for a given voltage. With the increased number of collisions, each collision need not produce as many free electrons; hence the breakdown voltage decreases. However, a point will finally be reached where the breakdown voltage will begin to increase with increasing pressure. This is because, with the high concentration of molecules in the tube, the mean free path is so short that the voltage applied to the tube must be increased in order to supply the necessary energy to the electron in its short travel.

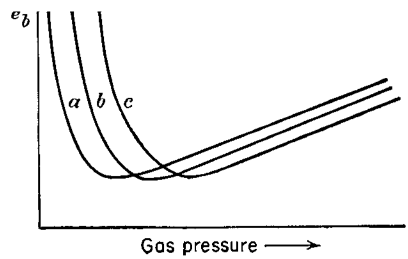


FIG. 13-3. Breakdown characteristics for a cold-cathode gas diode with parallel plane electrodes at various spacings.

If the distance between the electrodes is decreased, the resulting breakdown characteristic will be shifted horizontally with respect to the characteristic with larger spacing. This is shown by curves b and c in Fig. 13-3. From a family of such curves it can be easily seen that the breakdown voltage depends on the spacing between electrodes and the pressure of the gas. A third factor also governing the breakdown characteristic is the type of gas used. The breakdown voltages will differ among gases because of variation in the ease of removing electrons from the orbits, sizes of the gas molecules, etc. Another factor is the work function of the material used for the cathode, since all electrons must eventually be replaced by electrons from this source.

13-11. The Normal Glow.[4] Referring back to the normal-glow region in Fig. 13-2, we see that the drop across the tube is almost constant for quite a range of currents. This means that the tube has no control over the magnitude of the current, which must therefore be limited by some means external to the tube. This is the purpose of the resistor R in the figure. Presuming the supply voltage high enough for breakdown, if some means of limiting the current is not used, the latter will increase to such a high value that the tube will be damaged or even destroyed.

If the potential distribution between the two electrodes is plotted, a curve such as shown in Fig. 13-4 is the result. Here it can be seen that most of the tube drop occurs in the neighborhood of the cathode. This is because all the current flowing through the tube must come from the cathode. Hence the positive ions bombarding the cathode must have enough energy to cause secondary emission. The distance from point A to the cathode is approximately equal to the length of the mean free path, and hence it is in this region that the ions acquire the needed energy.

The region from A to the anode is called the positive column, and here the potential rises very slowly. The gradient is such that only enough ionizing collisions occur practically to cancel the negative space charge. If the anode is moved closer to the cathode, say to point B, the tube drop would fall slightly to a value approximately equal to E_{b2}. In fact, as the anode is moved closer to the cathode the tube drop decreases very slowly until point A is reached. Here the distance between the electrodes becomes less than the mean free path, and the glow will be extinguished and cannot be reestablished unless a higher voltage is used. From this discussion we see that the spacing has little to do with the tube drop but has considerable effect on the breakdown voltage.

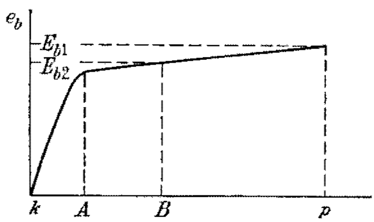


FIG. 13-4. Potential distribution in a cold-cathode gas diode with parallel plane electrodes.

13-12. The Abnormal Glow.[4] If a cold-cathode gas tube is made to break down into a normal glow and then the current through it is varied by means of the external resistor, some very interesting observations may be made. At low values of current the glow occurs over only a small portion of the cathode surface, and as the current increases, the area of the glow likewise increases while the tube drop remains constant. However, a point is finally reached where the glow completely covers the surface of the cathode. If the current is increased beyond this point, the glow increases in intensity and the tube drop begins to increase slowly. This is known as the abnormal-glow region and is indicated by region DE in Fig. 13-2. As the current through the tube is increased still further, the tube drop increases at an increasing rate, until the rate of rise becomes almost infinite.

13-13. The Arc Discharge.[4] As the tube drop increases in the abnormal-glow region, the heating of the cathode due to positive-ion bombardment likewise increases. A point may finally be reached where the cathode gets hot enough to cause thermionic emission. When this occurs, the high potential gradient near the cathode is no longer needed to supply electrons; hence the tube drop falls to a value comparable with the minimum ionization potential of the gas. This is an arc discharge and has the negative-resistance characteristic pictured in Fig. 13-2.

The difference between an arc and a glow is mainly in the method of obtaining electrons from the cathode. In a glow they must be removed

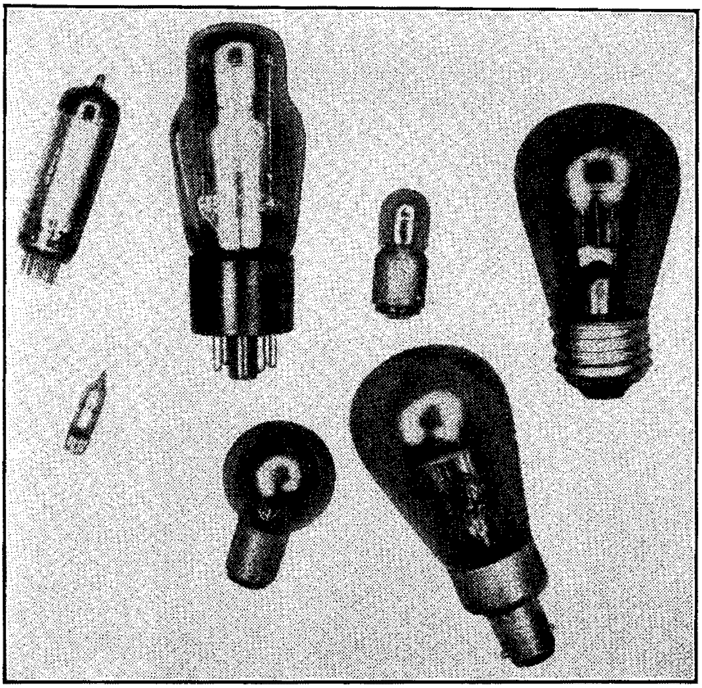


FIG. 13-5. Some common types of glow tubes.

by positive-ion bombardment, and hence, in order that these ions can have sufficient energy, they must fall through a considerable potential difference. In the arc discharge the electrons are produced by some other method, possibly by thermionic emission, and hence the potential rise in front of the cathode is not needed. As a result the tube drop is much lower and only great enough to produce sufficient ionization to neutralize the negative space charge.

13-14. Gaseous Diodes with Cold Cathodes. There are relatively few types of tubes which may be classified as cold-cathode gaseous diodes. Gas phototubes have light-sensitive cathodes and operate in the Townsend region (not a glow discharge). They are discussed in Art. 15-4. Possibly the most common gaseous diodes are the small glow lamps shown in Fig. 13-5. Either electrode may serve as a cathode, and hence they are bilateral conductors. They find their chief use as night lights or as indicators for electrical circuits, although occasionally they are used as voltage regulators in circuits such as will be described later in this article. Since the voltage drop in the glow region (Fig. 13-2) rises very little with increase in current, external resistance must ordinarily be used to limit the current flow; otherwise the operation may progress into the abnormal glow or the arc regions with resulting damage to the tube. This protection is unnecessary if the current is of very short duration, as in the case of the discharge of a capacitor. Some of these small glow lamps have a current-limiting resistor in the base, since the voltage drop is less than 100 volts and they are to be used on 115-volt circuits.

Another common type of cold-cathode gaseous diode is the voltage-regulator, or V-R, tube. The cathode is a cylinder about ¾ in. in diameter and 1 in. long. This cathode surrounds the anode, which is a straight wire along the axis of the cylinder. The tube is filled with one of the inert gases, usually argon, at such a pressure that the breakdown voltage is only a few volts greater than the drop across the tube after breakdown. In order to keep the breakdown voltage as low as practicable, a small wire projects out from the cathode to within a small fraction of an inch of the anode. The breakdown occurs between the end of this wire and the anode. When the current increases to about 5 ma, the glow transfers from this starter to the cathode cylinder. If the anode current is kept within the manufacturer's rating, usually between 5 and 40 ma, the tube operates with a glow ranging from normal to somewhat abnormal and the tube drop remains almost constant. Such tubes are made to have drops of 75, 90, 105, and 150 volts.

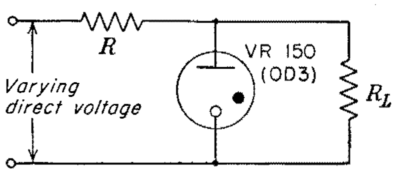


FIG. 13-6. Circuit for a voltage regulator using a V-R tube.

An example of a circuit for using a V-R tube is shown in Fig. 13-6. Here the voltage of the d-c source may not be perfectly constant; yet it is desirable to maintain the voltage applied to the load constant at 150 volts. Assume a 250-volt d-c source and a 5000-ohm load. The load draws a current of 150/5000 = 0.03 amp = 30 ma. If the circuit is so designed that normally a current of 20 ma passes through the V-R tube, then the resistor R passes 30 + 20 = 50 ma. Since the source voltage is 250 volts and the load voltage is 150, the drop across R must be 100 volts. The value of R must then be 100/0.050 = 2000 ohms. Now, if the source voltage increases, say to 270 volts, the drop across the V-R tube remains nearly constant, and hence the drop across R must increase. This requires an increase in the current through R, which means an increase in the current through the tube. Thus the fluctuation in source voltage is absorbed as increased IR drop across the resistor.

More complex circuits, using high-vacuum tubes along with V-R tubes, can be designed to regulate currents or voltages within closer limits.

13-15. Gas Diodes with Thermionic Cathodes.[5] A hot-cathode gas diode differs from a cold-cathode one in two ways. First, since the tube has a copious supply of free electrons, the tube has the characteristics of an arc and hence a low tube drop. Second, since the cathode emits electrons and the anode does not, the tube is unidirectional instead of bidirectional as is the cold-cathode tube. It may therefore be used as a rectifier, and here it finds its chief utilization.

13-16. Thermionic Cathodes Used in Gas Tubes. The cathode in a thermionic gas tube is usually oxide-coated and may be either of the filamentary or of the indirectly heated type. The main difference between the cathodes used in gas and in high-vacuum tubes is that heat shielding may be used in the former. In a high-vacuum tube the electrons have relatively high velocities, and their motion is continuous; hence they travel in approximately straight lines. Any attempt to shield the cathode, in order to cut down on the radiation of heat, will cause space charge to build up and hence results in cutoff. On the other hand, conditions in a gas tube are such that many collisions occur between electrons and gas molecules. Hence the velocities of the charged particles are very low. The directions of their motions can then be along the flux lines of the electric field. Also, since the space charge is practically neutralized, the plate current is not cut off.

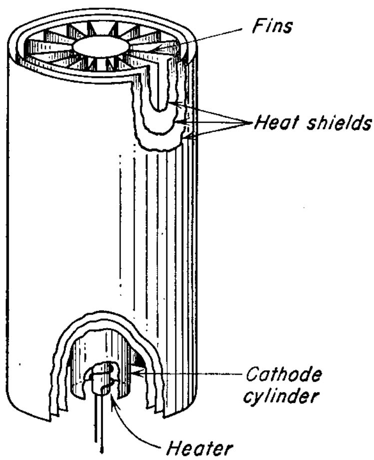


FIG. 13-7. A heat-shielded cathode.

Figure 13-7 is a drawing of a heat-shielded cathode. The cathode cylinder has several radial fins spot-welded to it. These fins in turn are spot-welded to another cylinder, which surrounds them. The cathode cylinder, the fins, and the outer cylinder are all coated with emitting material. These in turn are surrounded by two more bright nickel cylinders used solely to reflect radiated heat back into the emitting surfaces. Such a cathode may have an efficiency as high as 1500 ma per watt. However, the mass of a heat-shielded cathode is so great that it may require up to 15 min for it to arrive at the proper operating temperature.

Filamentary types of oxide-coated cathodes used in gas tubes are also heat-shielded, although not to such an extent as in the indirectly heated type. One example of the filamentary type might be a coated ribbon which is coiled into a helix and the whole structure surrounded by a cylinder of polished nickel. Each turn of the helix acts as a heat shield for the adjacent turn and the cylinder as the shield for the entire filament.

13-17. Gas- and Vapor-filled Tubes. Hot-cathode gas tubes may contain any of the inert gases (neon, argon, helium, krypton, or xenon), hydrogen, or mercury vapor. Strictly speaking, tubes containing mercury vapor are not gas tubes. However, their characteristics are so similar that they are included under the same general heading. Most commercial tubes contain mercury vapor; therefore much of our discussion will be concerning such tubes.

The hot-cathode gas tube must always be connected with a current-limiting resistor in the plate circuit. This resistor is usually the load. It must be of such size that it limits the anode current to a safe value. This statement brings up the question of what we mean by the term "safe value." When the ohmic value of the external resistor R is low enough so that the saturation-emission current of the cathode is passing through the tube, the cathode is incapable of supplying any more thermionic electrons. If more current is demanded by the circuit, the added current must come from the cathode. Since the value of R must be decreased in order to increase this current, the voltage drop across R decreases (if the current tends to remain constant); hence the drop across the tube must increase. As a result, the positive ions bombard the cathode with increased energy and remove more electrons by secondary emission. This, of course, results in an increased voltage drop across R, and the tube drop decreases, but not back to its former value. Of course, these actions actually occur simultaneously. As more and more current is demanded, the tube drop increases until a point is reached where the positive ions bombard the cathode with enough energy to knock off particles of the emitting material. The tube drop at which this destruction begins is known as the "disintegration voltage," and it has different values for different gases. For neon the disintegration voltage is about 27 volts, for argon 25 volts, and for mercury vapor 22 volts.

This brings up another precaution which must always be taken in using thermionic gas tubes. When the cathode cannot supply sufficient current for the external circuit to limit the tube drop to a value less than the disintegration voltage, the cathode will be seriously damaged. Hence it is always necessary to let the cathode heat to normal operating temperature before applying plate voltage to the tube.

13-18. Effects of Pressure on Operating Characteristics. When mercury vapor instead of an inert gas is used, the pressure in the tube will vary considerably if the temperature of the bulb is allowed to change. Figure 13-8 shows how the pressure in a mercury-vapor tube changes with increased temperature. A change in pressure will affect the characteristics in several ways. An increase in pressure will decrease the peak inverse voltage that the tube can withstand, *i.e.*, the voltage at which the tube breaks down into a glow when the plate is negative. When the gas pressure is too low, there are not enough gas molecules present to produce the necessary positive ions needed to neutralize the space charge. Therefore the tube drop will be high and might even be high enough to exceed the disintegration voltage.

If the pressure in the tube is too high, the peak-inverse-voltage rating may be too low. If the pressure is low, the peak-inverse-voltage rating may be high enough but the tube drop will also be high; hence destruction of the cathode may occur. We thus see that there are definite limits to the range of pressures and temperatures over which the tube can be operated.

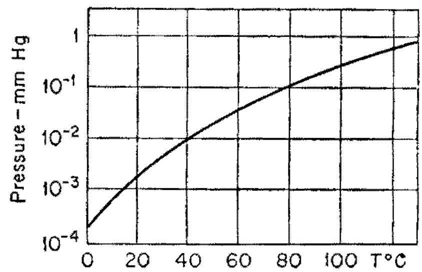


FIG. 13-8. Vapor pressure of mercury as a function of temperature.

Since the pressure in a gas tube is almost independent of the temperature, such tubes would seem to be more desirable than vapor tubes. This would be the case if it were not for the fact that during manufacture it is necessary to drive all gases out of the metal parts. Then when an inert gas is admitted to the tube, at some predetermined pressure, the electrodes start to absorb part of it. This process is rather slow and takes place over a long period of time. The tube will therefore gradually change pressure, with a corresponding change in characteristics occurring. On the other hand, vapor tubes always have a plentiful supply of mercury present, and hence the pressure depends only on the ambient temperature of the bulb. In other words, the characteristics of the gas-filled tube change as the tube ages because of the absorption of the gas, while mercury-vapor-tube characteristics do not change appreciably with time.

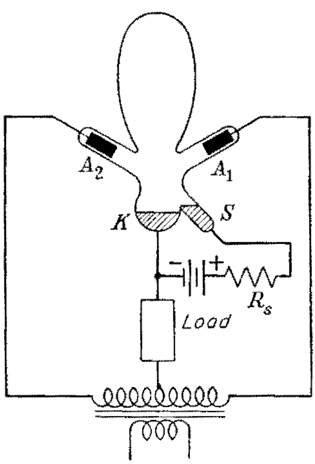


FIG. 13-9. An old-type mercury-arc tube and its circuit.

13-19. Mercury-pool Rectifiers.[6] One of the earliest forms of gaseous-discharge tubes was the mercury-pool, or mercury-arc, rectifier. Figure 13-9 shows a sketch of a tube which is an obsolete type but is used in this discussion because of its simple construction. This tube consists of a large evacuated glass bulb containing a pool of mercury which serves as a cathode, two graphite anodes, which are located in the two upper side arms, and a starter electrode, which is a mercury pool in the small lower side arm. At normal room temperature the pressure of the mercury vapor is of the order of 10^{-3} mm Hg. This is high enough to sustain an arc after it is once initiated. After the arc is established, the tube heats up and the pressure increases (see Fig. 13-8).

The mercury pool of the cathode is ordinarily incapable of supplying free electrons even if the anode voltage is applied, unless a cathode spot is caused to appear on the surface of the mercury. The starter electrode S is used for this purpose. The tube is tilted until mercury from the cathode pool runs into the starter side arm, thus connecting the starter to the cathode. A current, which is limited by the resistor R_s, begins to flow. When the tube is returned to its normal upright position, contact between the two pools is broken. A spark, or small arc, results and this causes a cathode spot to be formed. This means that electrons are being emitted from the cathode, and if at this time the anode is positive, an arc will be established between it and the cathode. The starter can then be open-circuited.

The cathode spot is easily recognizable since it is a small bright spot of light which continuously moves or dances over the surface of the mercury. Measurements seem to indicate that all the anode current passes through the cathode spot and that the current density may be as high as 10,000 amp per cm^2. When a high current is passed by the tube, two or more spots may form, with the current through each spot not exceeding 40 or 50 amp.

The theory of emission from the cathode spot has not been well established as yet. However, it has been proved fairly well that the temperature of the spot cannot be high enough to cause thermionic emission. The temperature of the spot can be calculated from the rate of evaporation of the mercury cathode. From studies of the drop across the tube it has also been shown that the positive ions do not have enough energy to knock electrons out of the cathode. Langmuir has suggested that the cloud of positive ions, which are traveling toward the cathode, is so close to it that a field is set up with an enormous potential gradient. This results in electrons being pulled from the mercury by high-field emission. To date this seems to be the most logical explanation.

In order to simplify the explanation of how this tube operates, let us first remove the connection to one anode, say A_2. A_1 is still connected to the a-c source. If the voltage applied to A_1 is positive when the tube is tilted and the arc started, the discharge will take place between A_1 and the cathode and will continue until the potential of A_1 falls to near zero. The anode current will then fall to zero, and the arc will be extinguished. If in the meantime the starter circuit has been opened, the cathode spot will have disappeared and the tube will have to be tilted again before it can start. If, however, the small d-c arc through the starter is allowed to continue, the next time the anode becomes positive the cathode spot will already exist and the arc will again occur between the anode and the cathode. The tube thus passes current on each positive half cycle, and we have a half-wave rectifier. This is, however, a rather inefficient method since the keep-alive circuit, which we have been calling the starter circuit, requires a current of several amperes in order to function.

Assume the load to be resistive. If A_2 is now connected back to the end of the center-tapped transformer, the voltage applied to A_2 is 180° out of phase with that applied to A_1. Now when the tube is tilted, the arc will occur between the cathode and one anode or the other, depending on which anode is positive. The arc will continue until the voltage of the supply and of the anode goes to zero; then the arc dies out. If in the meantime the starter circuit has been opened, the second anode will be unable to form a discharge when it becomes positive. If the resistance load is replaced by an inductive load, then anode A_1 will continue to fire after its supply voltage has fallen to zero. This is because the voltage induced in the inductive load is in such a direction as to oppose any change in the current, which is decreasing, *i.e.*, the anode voltage lags the supply voltage. The discharge may continue until the second anode becomes sufficiently positive, at which time the arc will transfer from the first anode to the second. Note that, regardless of which anode is firing, the current through the load is always in the same direction. The device thus works as a full-wave rectifier without a keep-alive electrode. But it does require mechanical motion to start it.

The drop across such a tube may be 20 to 50 volts. It is higher than the drop across a hot-cathode mercury-vapor rectifier since the energy for emission must come from the arc drop rather than from an external heating source. The tube also has a somewhat constant-voltage characteristic; hence a current-limiting resistor must always be used to limit the current to a safe value. Since bombardment of the mercury pool by positive ions can do no damage to the emitter, this tube can withstand enormous momentary overload currents. However, a sustained overload may cause overheating, with possible cracking of the glass where the lead-in wires enter the tube. Thus we see that the mercury-arc tube is a very rugged device that can be used where momentary overloads might damage the hot-cathode type. It does have some decided disadvantages. One of them is the fact that some mechanical method of starting the arc must be employed. Also, because of the relatively high tube drop and the large current needed to maintain this type of arc, the mercury-arc tube is not well suited to low-voltage low-current applications.

The large glass bulb used in this tube is merely for cooling the mercury vapor evaporated from the cathode. If the vapor pressure is allowed to become too high, an arcback (flow of current in reverse direction from normal) may occur; hence the bulb must be very large if a high voltage is to be rectified.

The glass mercury-vapor-arc tubes have been superseded by those using steel tanks for envelopes. Such devices may have up to 18 anodes for polyphase use. They may be built in this form up to about 8000 kw capacity and for voltages ranging from 200 to 3000 volts. The tanks of such tubes are surrounded by a jacket through which water is circulated in order to keep the vapor pressure inside down to a reasonable value. The anodes are large blocks of graphite with leads brought into the steel tank by means of insulating bushings. These bushings are not always vacuum-tight, so such tubes are always connected to vacuum pumps that are turned on when the pressure in the tube rises to a certain value.

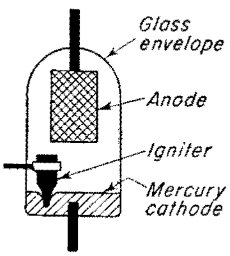


FIG. 13-10. A small sealed-off ignitron.

13-20. Ignitrons.[7] The ignitron (ig-nī′-tron) is one type of mercury-pool rectifier which overcomes some of the starting difficulties mentioned above. Figure 13-10 is a drawing of one of the smaller sizes of the tube. It differs from the ordinary mercury-arc tube only in the method of starting.

The cathode is the large pool of mercury in the bottom of the tube. The anode is the heavy piece of carbon supported by the lead wire coming out the top. The starter is the third electrode, which dips into the pool of mercury and is stationary. The starting is accomplished electrically and not by mechanical methods.

The starter electrode is shaped somewhat as shown in the figure and is made of a refractory semiconductor such as silicon carbide. Mercury will not wet this material and hence forms a meniscus at the point of contact. This seems to be one of the requirements for starting. When a surge of current is passed from the starter to the mercury, a cathode spot appears at the point of contact and if the anode is positive at this instant, the tube will fire. The exact reason for this cathode-spot formation is not thoroughly understood, but it is believed to be due to the high potential gradient set up between the unwetted starter electrode and the mercury. Only a few microseconds are required for ignition, and hence the tube can be fired each cycle. This would be impossible with mechanical starters. Thus the ignitron eliminates the need for keep-alive circuits.

If the time of firing is varied over the half cycle of voltage during which the anode is positive, the average or direct current can be varied from zero to some maximum value. Thus the tube becomes a control device as well as a rectifier. Figure 13-11 shows waveshapes of the output current when the tube is fired shortly after the anode becomes positive (*a*) and also when the anode voltage has reached its peak value (*b*). The methods of operating the igniter circuit will be studied later.

Ignitrons are usually made with water-cooled steel jackets instead of a glass envelope. A photograph of a typical tube with part of the steel jacket cut away is reproduced in Fig. 13-12. Still larger units are built, some of them being capable of rectifying thousands of amperes. These larger units must be continuously pumped in order to maintain a vacuum. To date the ignitron is a relatively low-voltage tube, but efforts are being made to develop high-voltage versions.

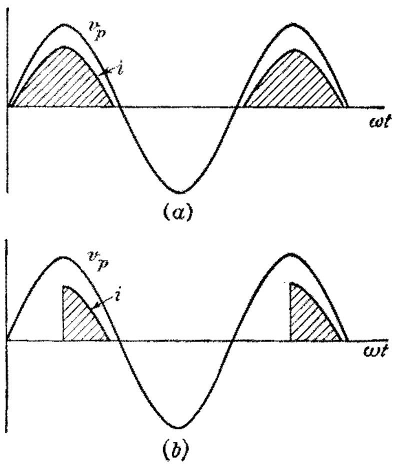


FIG. 13-11. Waveshapes of the current through an ignitron for two firing angles.

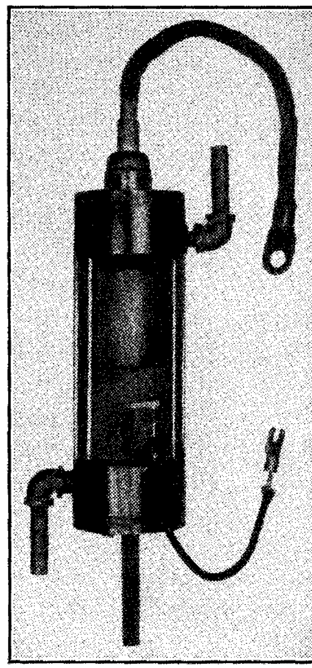


FIG. 13-12. Cutaway view of a water-cooled sealed-off ignitron.

This tube is of great importance commercially since it can rectify and control such large currents. However, it does have some disadvantages, one of them being the fact that while the current is unidirectional it is pulsating. This can be remedied to a certain extent by using the tubes in polyphase circuits. But for applications where waveshape is unimportant and the control of large unidirectional currents is required, the ignitron has not been surpassed. This is true in spot-welding control, electrolytic processes, d-c motor control, etc.

13-21. Action of the Grid in a Hot-cathode Gas Triode.[8] A hot-cathode gas diode may be converted to a triode by inserting a grid between the anode and cathode in such a way as fairly well to shield the cathode from the field of the anode. This tube will have entirely different characteristics from those of a high-vacuum triode. If the grid of this triode is made highly negative and then a positive anode voltage is applied, no anode current flows. If the negative grid voltage is gradually reduced, eventually a small anode current will begin to flow. If the grid voltage is further reduced, the anode current increases. A point will finally be reached where enough electrons pass through the grid and are accelerated sufficiently by the anode to cause ionization and breakdown into an arc. Up to this point the action of the grid in a gas tube has been practically the same as the action of the grid in a high-vacuum tube. However, the current in the former is very small.

After breakdown the tube drop is about the same as that of a thermionic gas diode since the negative space charge is neutralized by the positive ions produced by the ionization process. Hence the anode circuit must contain enough impedance to limit the current to some safe value.

If the potential of the grid is now varied, it will be found to have no effect upon the magnitude of the anode current. When the grid is negative, positive ions are attracted to it, and since they are gas molecules that have lost one or more electrons, they cannot enter the grid metal and pass out to the external circuit. They can do one of two things, however. They may strike the grid and remove an electron from it, thus becoming neutral molecules again, or they may form a cloud of ions about the grid structure. This cloud is called a positive-ion sheath. Positive ions will continue to come into the sheath until their charges neutralize the negative charge of the grid. The only effect of a more negative grid is to increase the thickness of the sheath. Theoretically the discharge could be choked off if the grid were to be made negative enough so that the passageways were covered by overlapping sheaths. In practice this is not an acceptable method of extinguishing the arc because of the very high negative voltage needed. The only practical way to stop the discharge is to remove the positive anode voltage.

Before breakdown the grid current consists of the few electrons which penetrate its negative field and strike the grid structure. As the grid is made less negative, a few electrons pass through the grid and are accelerated by the anode. They may strike gas molecules and produce positive ions. These in turn will be attracted to the negative grid and will be neutralized if they strike it. According to our convention, this means a negative grid current flows. Since more positive ions are formed as the grid voltage becomes less negative, the negative grid current increases as the grid approaches zero potential.

The grid is necessarily in the path of the discharge, and also it differs in potential from the other electrodes; hence it must be protected by means of a series ballast resistor. If this precaution is not taken, a discharge may be established between the grid and one of the other electrodes and thus damage to the tube may result.

13-22. Thyratrons.[8] The gas triode which we have been discussing is known as a thyratron. However, the construction of the usual thyratron is quite different from the construction of a high-vacuum triode. Figure 13-13 shows the plan for typical small thyratrons. The cathode may be either filamentary or indirectly heated. As in a gas diode, it may be of the high-efficiency heat-shielded type. The grid is merely a cylinder, with a baffle between the cathode and the plate. The hole in the baffle is to allow passage of electrons and positive ions. The anode is either a nickel disk or a block of graphite, depending on the current rating of the tube. Note that the grid structure almost surrounds the anode. This is because of the effect of charges which may collect on the walls of the tube. Suppose that the grid did not go up beyond the anode and that negative charges had collected on the walls of the tube. When the anode was positive, the grid would have difficulty making the tube fire because the charge on the walls would not allow electrons to pass between the cathode and the anode. Of course, if the grid were made positive enough, the tube would eventually break down. Since the amount of the charges collected on the walls is random, the firing voltage would be variable and the action of the tube would not be consistent. With the grid structure extending up beyond the anode, the charges on the walls have no effect on the firing voltage.

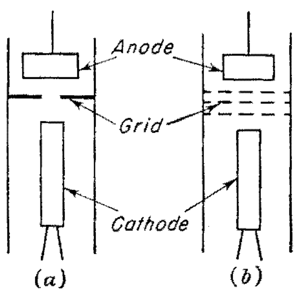


FIG. 13-13. Typical electrode structure for (*a*) negative-grid thyratrons, (*b*) positive-grid thyratrons.

An examination of the firing characteristics of different thyratrons would show that some will fire with the grid negative, while others will not fire unless the grid is positive. The only difference between these two types is the completeness of the shielding done by the grid. The negative-grid thyratron will usually have only a single baffle with one rather large hole in it as part of the grid structure. The positive-grid thyratron will have one or more baffles with a large number of small holes, thus affording better shielding.

The grid voltage which just prevents the thyratron from firing is known as the critical grid potential. The curve which shows the relationship between the anode voltage and the critical grid potential is known as the firing characteristic. A family of such curves is shown in Fig. 13-14. Note that the firing characteristic changes with temperature. This is because the tube is a mercury-vapor type, and hence the pressure is dependent on the temperature. This may be a decided disadvantage for certain applications, and hence some thyratrons are filled with an inert gas instead of mercury vapor. But as in gas diodes, the pressure may change in time because of the absorption of the gas by the tube elements. This is not desirable, since it results in changes in the firing characteristics. One solution may be to use mercury tubes in a temperature-controlled enclosure.

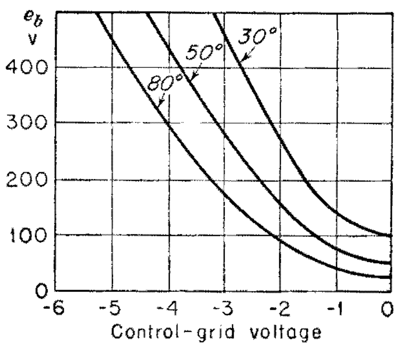


FIG. 13-14. Firing characteristics for a type FG27A thyratron showing the effects of temperature change.

Since the arc in a thyratron cannot be extinguished without removing the anode voltage, it is usual to employ these tubes with an a-c supply for the anode. Thus this voltage is reduced from positive values to zero once every cycle, and the arc is extinguished. For this reason it is sometimes desirable to draw a curve showing the critical grid voltage on the same set of axes as is the alternating anode-supply voltage. Such a curve and the graphical method of obtaining it are shown in Fig. 13-15, which is self-explanatory. However, in many thyratrons, using large alternating supply voltages, the critical-grid-voltage curve comes so close to the zero-voltage axis that practically no error is introduced by assuming the critical grid voltage to be along this axis.

13-23. Shield-grid Thyratrons. One disadvantage of the single-grid thyratron is that, in order to have the grid control the firing, it must almost completely surround the cathode. This necessitates the grid being rather large physically, which means that the grid current will be considerable. This may be quite a drawback when the tube is to be fired from a high-impedance source. Also, the capacitance between the grid and the cathode will be rather large, with a resulting relatively low input impedance.

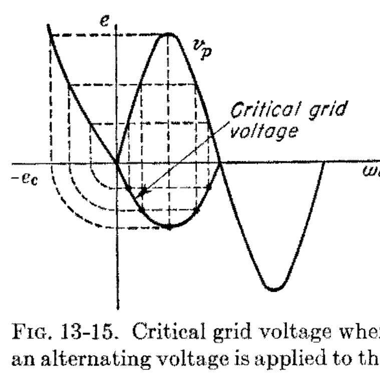


FIG. 13-15. Critical grid voltage when an alternating voltage is applied to the plate of a thyratron.

In order to increase the input impedance and reduce the grid current, the shield-grid thyratron was developed. Figure 13-16 shows a cross section of such a tube. The shield grid is the large structure which surrounds the cathode and can be made either positive or negative with respect to the cathode. The control grid is the small cylindrical structure between the two baffles of the shield grid. The large shield grid, being held at a fixed potential and doing most of the shielding, will collect most of the current. The small control grid will collect very little of it. The size of the control grid also helps reduce the grid-to-cathode capacitance, and hence the input impedance is much higher than it is for the single-grid type. Figure 13-17 shows the firing characteristics of a shield-grid thyratron for various values of shield-grid voltage. Note that these curves point out another advantage of the double-grid type over the single-grid one. The firing characteristic can be shifted by varying the shield-grid

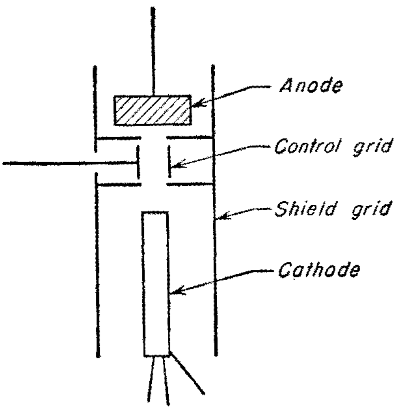


FIG. 13-16. Typical electrode structure for a shield-grid thyratron.

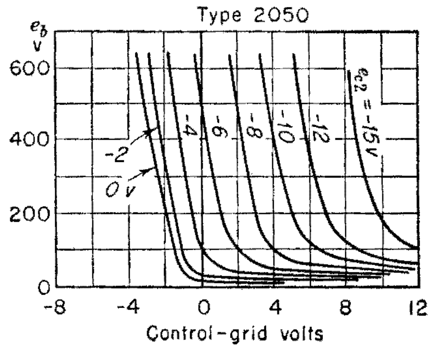


FIG. 13-17. Firing characteristics for a shield-grid thyratron.

voltage. In the gas-type thyratron this shift can compensate for the effect of change in gas pressure.

13-24. Ionization and Deionization Times. As mentioned above, the ionization time for a thyratron will be a matter of only a few microseconds. By ionization time is meant the time it takes the tube to break down into a self-maintaining arc after the voltage becomes such that the tube will fire. Because of the small interval of time required, the ionization time is not of too much interest to us. On the other hand, after the anode voltage is removed, it requires an appreciable time for all the positive ions to drift over to one of the electrodes or to the walls of the tube and pick up electrons in order to become neutral molecules again. If the anode voltage is applied before this neutralization process is complete, the grid will not have regained control and the tube will fire regardless of the potential of the grid. The time required to accomplish this neutralization is known as the deionization time and may be as long as 1000 μsec. Thus we see that the frequency of the anode voltage must be low enough so that there is time for the tube to deionize during the negative half cycle and for the grid to regain control.

REFERENCES

1. Tonks, L.: Electrical Discharges in Gases—Ionization and Excitation, *Elec. Eng.*, **53**, 239 (1934).
2. Compton, K. T., and I. Langmuir: Electrical Discharges in Gases; Part I, Survey of Fundamental Processes, *Revs. Modern Phys.*, **2**, 171 (1930).
3. Townsend, J. S.: "Electricity in Gases," Oxford University Press, New York, 1915.
4. Loeb, L. B.: "Fundamental Processes of Electrical Discharges in Gases," John Wiley & Sons, Inc., New York, 1939.
5. Cobine, J. D.: "Gaseous Conduction," McGraw-Hill Book Company, Inc., New York, 1941.
6. Prince, D. C., and F. B. Vogdes: "Principles of Mercury-arc Rectifiers and Their Circuits," McGraw-Hill Book Company, Inc., New York, 1927.
7. Slepian, J., and L. R. Ludwig: A New Method of Initiating the Cathode of an Arc, *Trans. AIEE*, **52**, 693 (1933).
8. Hull, A. W.: Hot Cathode Thyratrons, *Gen. Elec. Rev.*, **32**, 213 (1929).
9. E. E. Staff, MIT: "Applied Electronics," John Wiley & Sons, Inc., New York, 1943.
10. Millman, J., and S. Seely: "Electronics," 2d ed., McGraw-Hill Book Company, Inc., New York, 1951.
11. Reich, H. J.: "Theory and Applications of Electron Tubes," 2d ed., McGraw-Hill Book Company, Inc., New York, 1944.

PROBLEMS AND QUESTIONS

1. State the approximate pressure of the gas in high-vacuum tubes; in thyratrons.

2. Name the three types of discharges discussed in this chapter. State the order of magnitudes of the current through and of the voltage across the tube in each case.

3. What is an excited atom; a metastable atom?

4. For how long a time can an excited atom exist; how long can the metastable atom exist?

5. What is meant by ionization of a gas? What is a positive ion? What is a negative ion?

6. How are ions produced?

7. What is an elastic collision; an inelastic collision?

8. List the factors which determine the frequency of collisions between charged particles and gas molecules in gas which is in an electric field.

9. What is the meaning of the term "mean free path"?

10. What effect does gas pressure have on the length of the mean free path?

11. Explain how a few positive ions in a gaseous discharge can neutralize the effects of a large number of electrons and hence result in practically zero space charge.

12. What effects might positive ions have on the surfaces of electrodes and on the walls of a tube?

13. Why is the drop across a cold-cathode discharge high?

14. Name the possible primary ionizing sources in a Townsend discharge. Is this type of discharge self-maintaining?

15. What is the difference between a self-maintaining and a non-self-maintaining discharge?

16. Explain why the breakdown voltage of a gas varies as shown in Fig. 13-3.

17. What is the type of discharge which occurs immediately after breakdown from a Townsend discharge?

18. What are the characteristics of the normal glow?

19. Sketch the potential distribution through a tube which is operating in the normal-glow region.

20. Explain what happens to the tube drop as the electrodes, in a tube with a normal glow, are moved closer and closer together.

21. What is the abnormal glow? How can you recognize the point where it begins in a tube where you can vary the discharge current?

22. What are the characteristics of an arc discharge?

23. Why is the drop across an arc less than that across a glow discharge?

24. Name two types of tubes using cold-cathode glow discharges.

25. Why must a resistor always be placed in series with a cold-cathode gas diode?

26. A type OC3 (VR150/40) is a cold-cathode voltage-regulator tube whose voltage drop remains constant at approximately 150 volts as the current through it is varied between 5 and 40 ma. Design a circuit, similar to that of Fig. 13-6, to keep the voltage across a 5000-ohm load constant at 300 volts as the power-supply voltage varies between 400 volts and 450 volts.

27. Does the hot-cathode gas diode operate with an arc or a glow type of discharge?

28. Why is it possible to use heat-shielded hot cathodes in gas tubes and not in high-vacuum tubes?

29. Why do hot-cathode gas tubes always use a resistor in series with the anode?

30. What happens if the current through a hot-cathode gas tube becomes too high?

31. What precaution must we always take when putting a hot-cathode gas tube into operation?

32. Name the effects of gas pressure on the operating characteristics of a gas diode.

33. What is meant by the term "peak-inverse-voltage rating"?

34. Why might it be more desirable to use an inert gas in a tube rather than mercury vapor?

35. Explain briefly why a load that is partially inductive permits the use of the circuit of Fig. 13-9 where the auxiliary electrode is used for starting only and there are no keep-alive electrodes.

36. Could the circuit of Fig. 13-9 be used if the load was a pure resistance and the leakage reactance of the transformer large?

37. Why are mercury-pool rectifiers able to withstand very large momentary overload currents?

38. How does the ignitron differ from the ordinary mercury-arc rectifier?

39. Why are ignitrons superior to the ordinary mercury-pool rectifiers in many applications?

40. What effect does a grid have on the anode current of a gas triode before the tube breaks down? After breakdown?

41. What is a positive-ion sheath?

42. Why is a current-limiting resistor used in the grid circuit of a gas triode?

43. What is a thyratron, and what are its characteristics in so far as grid control is concerned?

44. Does the value of grid-to-cathode voltage of a thyratron appreciably affect the value of the plate voltage for which an arc is formed?

45. Does the value of grid-to-cathode voltage of a thyratron appreciably affect the value of the plate voltage for which an arc in the tube is extinguished?

46. Give the advantages of a shield-grid thyratron over the single-grid type.

47. What do we mean when we speak of the ionization time for a gas; the deionization time?

48. Thyratrons are inherently rather low frequency devices. Why?

CHAPTER 14

RECTIFIERS

14-1. Introduction. A rectifier is a device which converts one or more alternating voltages into a unidirectional one. This definition covers the use of electronic tubes as well as many devices such as barrier-layer rectifiers (exemplified by copper-oxide and iron-selenium rectifiers), crystal detectors, certain electrolytic cells, and mechanical rectifiers. Rotary converters and motor-generator sets are not usually classified as rectifiers, although they often serve the same purposes.

In the conversion of large amounts of power from the alternating to the direct form it is usually advantageous to employ a polyphase source, whereas for moderate and small power requirements a single-phase source is satisfactory. In this text we shall confine our discussions to single-phase rectifiers employing electronic tubes. However, much of the material can be applied to other types.

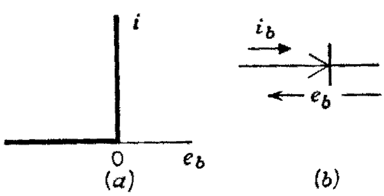


FIG. 14-1. The volt-ampere characteristic (a) and symbol (b) for an ideal rectifier.

Usually the matters of principal interest with a particular circuit may include the direct current (or voltage), the alternating components (ripple) of the current, the peak plate current, the tube's plate dissipation, the rectification efficiency, and the d-c magnetization of the transformer core.

14-2. The Ideal Rectifier. The ideal rectifier would be one in which the drop across it while conducting would be zero. In other words, it would have zero resistance in the forward direction. It would also have infinite resistance in the nonconducting, or reverse, direction. The volt-ampere characteristic of such a device is pictured in Fig. 14-1. The symbol which we shall use to represent the ideal rectifier is also shown in the same figure. Such a rectifier would have no voltage drop across it while conducting and hence no power loss. When not conducting, the current would be zero so again there would be no power loss.

14-3. The High-vacuum Diode as a Rectifier. When using a high-vacuum diode as a rectifier we find that, while the device conducts in only one direction, it is not ideal. The current-voltage relationship is given by the Child-Langmuir equation ($i_b = Ke_b^{3/2}$) during conduction and by

$i_b = 0$ during the nonconducting time. The volt-ampere characteristic for a type 6H6 tube is pictured as curve a in Fig. 14-2. If we were to work with this more or less exact relationship, the solution of rectifier problems would be extremely difficult. It is therefore desirable to make some approximations and simplifications.

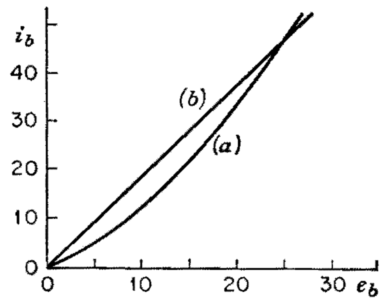


FIG. 14-2. The volt-ampere characteristic for a 6H6 tube.

Using this same figure, let us first draw a straight line which seems to approximate the actual characteristic. This is shown as curve b. Note that there is not a great deal of variation between these curves. If we now draw the volt-ampere characteristic of the series combination of the ideal rectifier and a fixed resistor equal in magnitude to the d-c plate resistance R_p of the tube, we see that this too yields curve b of Fig. 14-2. In our equivalent circuits we shall use the idealized tube as an approximation for the actual one.

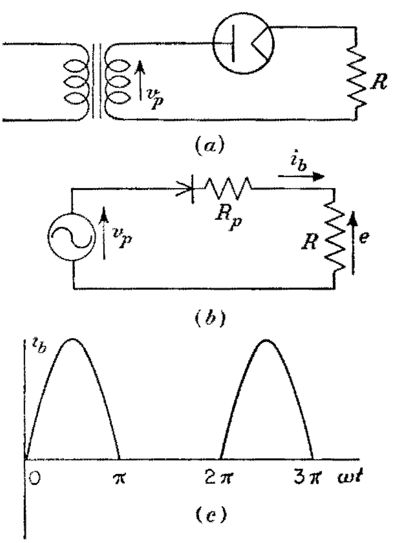


FIG. 14-3. (a) Circuit diagram for a half-wave rectifier with resistance load. (b) Equivalent circuit for (a) using a high-vacuum diode. (c) Waveform of load current.

14-4. The Half-wave High-vacuum Rectifier with a Resistance Load. Figure 14-3 shows the circuit diagram and the equivalent circuit for a half-wave high-vacuum rectifier with a resistance load. Except for the ideal rectifier, all circuit elements are linear. We can assume the supply voltage to be sinusoidal; thus $v_p = V_{pm} \sin \omega t$.

During the time when the tube conducts, the tube-and-load current is given by

$$i_b = \frac{v_p}{R + R_p} = \frac{V_{pm} \sin \omega t}{R + R_p} \qquad 0 < \omega t < \pi \tag{14-1}$$

This, of course, assumes that the cathode emission exceeds $i_{b_{\max}}$. During the remaining part of the cycle

$$i_b = 0 \qquad \pi \leq \omega t \leq 2\pi \tag{14-2}$$

The waveform of i_b is shown in Fig. 14-3c.

To find the average, or direct, value of the current, we can use the ordinary method of integrating under the curve and dividing by 2π. This results in the following:

$$I_{dc} = I_{ba} = \frac{1}{2\pi}\int_0^{2\pi} i_b\, d\omega t = \frac{1}{2\pi}\frac{V_{pm}}{R + R_p}\int_0^{\pi} \sin \omega t\, d\omega t$$
$$= \frac{1}{\pi}\frac{V_{pm}}{R + R_p} \tag{14-3}$$

From this it follows that the direct voltage across the load is

$$E_{dc} = I_{dc}R = \frac{R}{\pi}\frac{V_{pm}}{R + R_p} = \frac{1}{\pi}\frac{V_{pm}}{1 + (R_p/R)} \tag{14-4}$$

Examination of this equation discloses that if the load resistor R is decreased in value, the value of E_{dc} decreases unless R_p is negligible compared with R. For good voltage regulation this latter relationship must be the case. The maximum value of direct voltage that can be obtained from this circuit is V_{pm}/π, obtained when R is infinite.

Since the purpose of a rectifier is to supply direct current, the useful output or load current will be taken as the direct component. How well the rectifier performs may be expressed by the rectification efficiency, which is the ratio of d-c power delivered to the load to the total power furnished by the source. The useful power in the load is

$$P_{dc} = I_{dc}^2 R = \frac{V_{pm}^2 R}{\pi^2 (R + R_p)^2} \tag{14-5}$$

The a-c components also pass through the load, but in this case their only effect is to produce heat, which must be considered as power loss.

The total heating effect of the current, both on the tube and on the load, may be computed from its effective value I_b. This may be calculated as

$$I_b = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} i_b^2\, d\omega t} = \sqrt{\frac{1}{2\pi}\frac{V_{pm}^2}{(R + R_p)^2}\int_0^{\pi} \sin^2 \omega t\, d\omega t}$$
$$= \frac{1}{2}\frac{V_{pm}}{R + R_p} \tag{14-6}$$

Because the actual rectifier is represented by an ideal rectifier in series with a resistor of value equal to the d-c resistance of the tube, the power dissipated in the tube can be found from the relation

$$P_p = I_b^2 R_p = \frac{V_{pm}^2 R_p}{4(R + R_p)^2} \tag{14-7}$$

The total power supplied by the source is

$$P_{\text{in}} = I_b^2 (R + R_p) = \frac{V_{pm}^2}{4(R + R_p)} \tag{14-8}$$

Thus the rectification efficiency is

$$\eta = \frac{P_{dc}}{P_{\text{in}}} = \left(\frac{2}{\pi}\right)^2 \frac{R}{R + R_p} = \frac{40.5}{1 + (R_p/R)} \quad \% \tag{14-9}$$

Hence the maximum efficiency is 40.5 per cent. A curve of efficiency as a function of the ratio R/R_p is shown in Fig. 14-4. Note that when $R = 10R_p$, the efficiency is approximately 37 per cent. Hence, for many of the loads encountered in electronic applications, the rectifier would be operating in the neighborhood of its theoretical maximum efficiency.

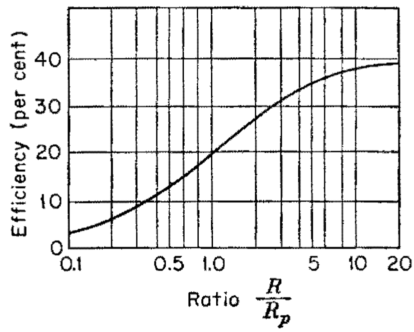


FIG. 14-4. Efficiency of a half-wave rectifier as a function of the load R.

Another measure of the effectiveness of a rectifier in supplying direct current is the ripple factor. This quantity γ is defined as the ratio of the effective value of all the alternating components constituting the ripple to the direct, or average, value of the current.

Since we do not have the equation for the ripple current alone, we cannot find its effective value directly. Here the Fourier-series representation would be convenient. This series would have the form (see Prob. 14*b*, Chap. 6)

$$i_b = I_{dc} + I_{1m} \sin \omega t - I_{2m} \cos 2\omega t + \cdots \tag{14-10}$$

and of the coefficient values, only that of I_{dc} has yet been determined. The other terms on the right collectively constitute the ripple i_{ac}. Figure 14-3*c* shows that the cycle of i_b repeats every 2π radians. Hence the term $I_{1m} \sin \omega t$ is an important component of the ripple. This is true of all half-wave rectifiers, either with or without filters.

We may again compute the effective value I_b, this time as

$$\begin{aligned} I_b &= \left\{\frac{1}{2\pi}\int_0^{2\pi} i_b^2\, d\omega t\right\}^{1/2} \\ &= \left\{\frac{1}{2\pi}\int_0^{2\pi} (I_{dc} + I_{1m}\sin\omega t - I_{2m}\cos 2\omega t + \cdots)^2\, d\omega t\right\}^{1/2} \\ &= \left\{\frac{1}{2\pi}\left[\int_0^{2\pi} I_{dc}^2\, d\omega t + 2\int_0^{2\pi} I_{dc}(I_{1m}\sin\omega t - I_{2m}\cos 2\omega t + \cdots)\, d\omega t \right.\right. \\ &\qquad \left.\left. + \int_0^{2\pi} (I_{1m}\sin\omega t - I_{2m}\cos 2\omega t + \cdots)^2\, d\omega t\right]\right\}^{1/2} \\ &= \{I_{dc}^2 + 0 + I_{ac}^2\}^{1/2} \end{aligned} \tag{14-11}$$

where I_{ac} is the effective value of the ripple current i_{ac}. From (14-11) we may solve for the ripple factor as

$$\gamma = \frac{I_{ac}}{I_{dc}} = \frac{\sqrt{I_b^2 - I_{dc}^2}}{I_{dc}} = 1.21 \tag{14-12}$$

This numerical value is obtained by substituting for I_b from (14-6) and for I_{dc} from (14-3). Thus for this half-wave rectifier we see that the efficiency is low and that the effective value of the ripple current is greater than that of the direct current. Because of these drawbacks the use of this rectifier circuit is rather limited.

One application of this circuit is for use as a voltmeter. Since I_{dc} is proportional to V_p, a d'Arsonval type of voltmeter inserted in the series circuit can be calibrated in volts. However, in order to render the input impedance of the circuit high, as is desirable, a high-sensitivity d'Arsonval movement is required, resulting in a somewhat expensive instrument. The ripple can be passed through the meter, as it causes no deflection. Since such an instrument is usually calibrated in rms volts on a sinusoidal source, its indications on nonsinusoidal voltages will be in error.

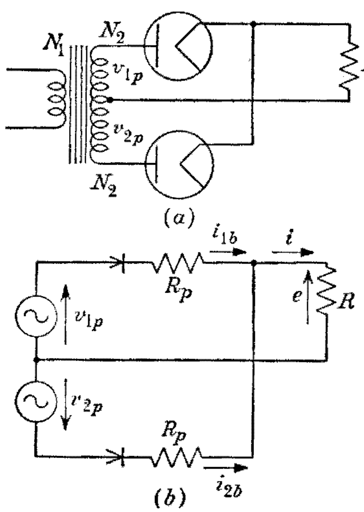


FIG. 14-5. (a) Circuit diagram for a full-wave rectifier with resistance load. (b) Equivalent circuit for (a).

14-5. Full-wave High-vacuum Rectifier with a Resistance Load. The full-wave rectifier can be looked upon as consisting of two half-wave rectifiers feeding a common load. The two half-wave sections are supplied from two sources with voltages 180° out of phase with each other. Such a circuit and its equivalent are pictured in Fig. 14-5.

If the assumption is made that the two voltages v_{1p} and v_{2p} are also equal in amplitude, we may write

$$v_{1p} = V_{pm} \sin \omega t \tag{14-13}$$

and

$$v_{2p} = -v_{1p} = V_{pm} \sin (\omega t + \pi) \tag{14-14}$$

During the half cycle when tube 1 is conducting, tube 2 will be cut off and when tube 2 conducts, tube 1 will be cut off. The current passed by both tubes goes through the load resistor R, and so we may write

$$i = i_{1b} + i_{2b} \tag{14-15}$$

The waveforms of the tube currents are half sinusoids, making the load-current waveform that shown in Fig. 14-6. The cycle of current repeats every π radians, and therefore the lowest frequency present for the alternating components of the load current is double the source-voltage frequency. This is a decided advantage over the half-wave rectifier, as we shall see when we take up the study of filters to be used with these circuits.

Inspection of the path of the circuit involving the power transformer shows that the primary ampere turns equal $N_1 i_1$, while the secondary ampere turns can be expressed as $N_2(i_{1b} - i_{2b})$, where N_2 is the number of turns in half the secondary. If the mmf loss in the core is negligible, we may equate the two expressions and obtain

$$i_1 = \frac{N_2}{N_1}(i_{1b} - i_{2b}) \tag{14-16}$$

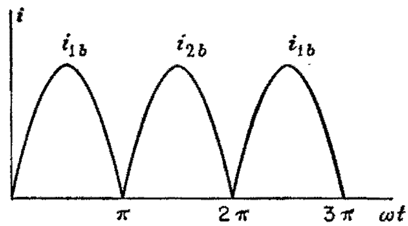


FIG. 14-6. Waveshape of the current through the load of a full-wave rectifier.

If the current of the second tube is subtracted from that of the first, it can be readily seen that i_1 has a sinusoidal form. Hence the d-c and all harmonic components must be missing from the primary current. Likewise, the secondary ampere turns $N_2(i_{1b} - i_{2b})$ contain no d-c magnetizing component. The same is true of the primary ampere turns. Therefore there is no d-c magnetization of the core. Since less iron is needed, the full-wave rectifier has this decided advantage over the half-wave one.

The average and effective values can be easily calculated as they were for the half-wave circuit. They are given by

$$I_{dc} = \frac{2}{\pi}\frac{V_{pm}}{R + R_p} \tag{14-17}$$

and by

$$I_b = \frac{1}{\sqrt{2}}\frac{V_{pm}}{R + R_p} \tag{14-18}$$

From these equations the values of E_{dc} and E can be computed if desired.

The d-c power delivered to the load can now be calculated as

$$P_{dc} = I_{dc}^2 R = \left(\frac{2}{\pi}\right)^2 \left(\frac{V_{pm}}{R + R_p}\right)^2 R \tag{14-19}$$

If we reexamine Eq. (14-5), we see that the d-c power output in the full wave case is just four times that for a half-wave rectifier using tubes of the same type and the same source voltage V_p.

The power dissipated by the plates can be written as

$$P_p = I_b^2 R_p = \frac{1}{2}\left(\frac{V_{pm}}{R + R_p}\right)^2 R_p \tag{14-20}$$

As a result of the increased power output and less a-c power loss in the load, the full-wave rectifier is more efficient than is the half-wave rectifier. The total power supplied to the circuit is

$$P_{\text{in}} = I_b^2(R + R_p) = \frac{V_{pm}^2}{2(R + R_p)} \tag{14-21}$$

The rectification efficiency is then found to be

$$\eta = \frac{P_{dc}}{P_{\text{in}}} = \frac{8}{\pi^2}\frac{1}{1 + (R_p/R)} = \frac{81}{1 + (R_p/R)} \quad \% \tag{14-22}$$

Its maximum value is just twice that found for the half-wave rectifier.

The ripple factor is found to be less than one-half that of a half-wave rectifier and is calculated as

$$\gamma = \frac{\sqrt{I_b^2 - I_{dc}^2}}{I_{dc}} = 0.48 \tag{14-23}$$

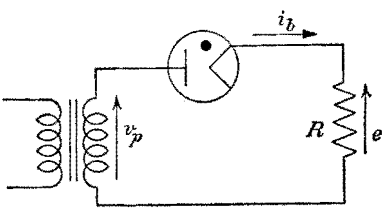


FIG. 14-7. Circuit diagram of a half-wave gas-diode rectifier with a resistance load.

We may summarize the advantages of a full-wave over a half-wave rectifier. For the same d-c power supplied to a load, the former has double the rectifying efficiency, and it may be shown that the required total plate dissipation is only half that for the half-wave case. Also the ripple factor is less for the full-wave rectifier, and the transformer may be smaller.

14-6. Half-wave Gas-diode Rectifier with a Resistance Load. A hot-cathode gas diode may be connected as a half-wave rectifier as shown in Fig. 14-7. In order to draw an equivalent circuit, it is again desirable to use an idealized tube. The volt-ampere characteristic for a type 83 tube, which is a hot-cathode mercury-vapor rectifier, is shown in Fig. 14-8. Because of its constant-voltage characteristic, a fairly good approximation can be made by drawing it as a straight vertical line at the average value E_d as shown by curve b. This is the same characteristic as is possessed by a battery in series with an ideal rectifier. The polarity of the battery representing the tube drop must be in opposition to current flow through the rectifier as indicated in Fig. 14-9. Note that this representation is quite different from that for the high-vacuum diode.

Writing a mesh voltage equation for the equivalent circuit during the conducting period, we obtain the relation

$$v_p = V_{pm} \sin \omega t = i_b R + E_d \quad (14\text{-}24)$$

where E_d is the voltage of the equivalent battery. Solving for the current yields

$$i_b = \frac{V_{pm} \sin \omega t - E_d}{R} \quad (14\text{-}25)$$

When the tube is blocking, the current is obviously zero.

An examination of the equivalent circuit shows that the tube blocks until v_p becomes positive and greater than E_d. The tube then breaks down into a gaseous discharge. The angle θ_1 at which the breakdown or ignition occurs can be found by solving Eq. (14-25) for $\omega t (= \theta_1)$. Here $i_b = 0$, and hence

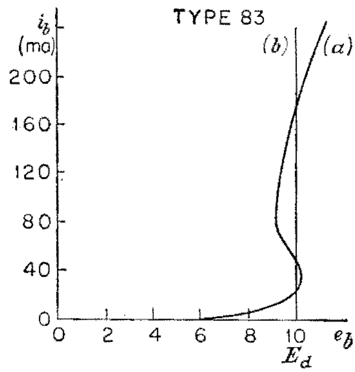


FIG. 14-8. Volt-ampere characteristic for a hot-cathode mercury-vapor rectifier.

$$\theta_1 = \arcsin \frac{E_d}{V_{pm}} \quad (14\text{-}26)$$

The tube will continue to conduct until the applied voltage again becomes less than E_d. Hence from symmetry we may express the extinction angle as

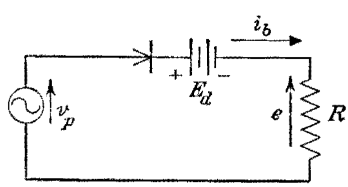


FIG. 14-9. Equivalent circuit for a half-wave mercury-vapor rectifier with a resistance load.

$$\theta_2 = \pi - \theta_1 \quad (14\text{-}27)$$

The conduction angle then becomes

$$\theta_c = \theta_2 - \theta_1 = \pi - 2\theta_1 \quad (14\text{-}28)$$

The load current is sinusoidal during the conduction period, which is less than one-half cycle. Figure 14-10 shows the shape of the current wave. A sketch of v_p serves as a timing wave.

Complete expressions for the tube current are

$$i_b = \frac{V_{pm} \sin \omega t - E_d}{R} \qquad \theta_1 < \omega t < \theta_2 \quad (14\text{-}29)$$

and

$$i_b = 0 \qquad \theta_2 < \omega t < 2\pi + \theta_1 \quad (14\text{-}30)$$

We may calculate the average current as

$$I_{dc} = \frac{1}{2\pi} \int_0^{2\pi} i_b \, d\omega t = \frac{1}{2\pi} \int_{\theta_1}^{\theta_2} \frac{V_{pm} \sin \omega t - E_d}{R} \, d\omega t \tag{14-31}$$

or

$$I_{dc} = \frac{V_{pm}}{2\pi R} \left[-\cos \theta_2 + \cos \theta_1 - \frac{E_d}{V_{pm}} (\theta_2 - \theta_1) \right] \tag{14-32}$$

The direct voltage across the load resistor can be found by multiplying I_{dc} by R.

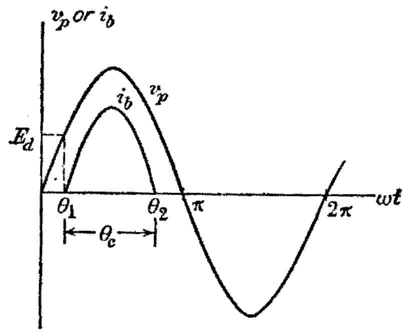


FIG. 14-10. Waveform of the plate current in a gaseous rectifier.

If the magnitude V_{pm} of the source voltage is large compared with E_d, $\theta_1 \approx 0$ and $\theta_2 \approx \pi$. Hence the average current becomes

$$I_{dc_{\max}} \approx \frac{V_{pm}}{\pi R} \tag{14-33}$$

The peak current through a gas tube is of importance because of resulting damage to the cathode if this current becomes too high. The peak current flows at the time when the supply voltage is at its maximum value. This occurs when $\omega t = \pi/2$. We may then write from Eq. (14-29)

$$i_{b_{\max}} = \frac{V_{pm} - E_d}{R} \tag{14-34}$$

The average power supplied to the circuit can be calculated as

$$\begin{aligned} P_{\text{in}} &= \frac{1}{2\pi} \int_0^{2\pi} v_p i_b \, d\omega t = \frac{1}{2\pi} \int_{\theta_1}^{\pi - \theta_1} (V_{pm} \sin \omega t) \frac{V_{pm} \sin \omega t - E_d)}{R} \, d\omega t \\ &= \frac{V_{pm}^2}{2\pi R} \left(\frac{\pi - 2\theta_1}{2} + \frac{\sin 2\theta_1}{2} - 2 \frac{E_d}{V_{pm}} \cos \theta_1 \right) \end{aligned} \tag{14-35}$$

When E_d is small compared with V_{pm}, this reduces to

$$P_{\text{in}} \approx \frac{V_{pm}^2}{4R} \tag{14-36}$$

The power lost in the tube can be found from the equation

$$\begin{aligned} P_p &= \frac{1}{2\pi} \int_0^{2\pi} e_b i_b \, d\omega t = \frac{1}{2\pi} \int_{\theta_1}^{\theta_2} \frac{E_d(V_{pm} \sin \omega t - E_d)}{R} \, d\omega t \\ &= \frac{E_d}{2\pi} \int_{\theta_1}^{\theta_2} \frac{V_{pm} \sin \omega t - E_d}{R} \, d\omega t \end{aligned} \tag{14-37}$$

since E_d is constant. By comparison with Eq. (14-31), this can be recognized as being

$$P_p = E_d I_{dc} \tag{14-38}$$

The rectification efficiency can be calculated from the relation

$$\eta = \frac{P_{dc}}{P_{\text{in}}} = \frac{I_{dc}{}^2 R}{P_{\text{in}}} \times 100\% \tag{14-39}$$

Substitution from Eqs. (14-32) and (14-35) into (14-39) yields

$$\eta = \frac{1}{2\pi} \frac{[\cos\theta_1 - \cos\theta_2 - (E_d/V_{pm})(\theta_2 - \theta_1)]^2 \times 100}{\dfrac{\pi - 2\theta_1}{2} + \dfrac{\sin 2\theta_1}{2} - 2\dfrac{E_d}{V_{pm}}\cos\theta_1} \quad \% \tag{14-40}$$

This rather complex formula points out one fact to us—the efficiency is independent of the load resistance. The efficiency does increase with increased values of V_{pm}, and in order to investigate the maximum possible efficiency, let us again assume that V_{pm} is very large compared with E_d. Equation (14-40) then reduces to the approximate form

$$\eta \approx \frac{4}{\pi^2} \times 100 = 40.5\% \tag{14-41}$$

and we see that the theoretical maximum efficiency is the same as that found for the half-wave rectifier using a high-vacuum tube.

The ripple factor can be determined if I_b, the effective value of i_b, is calculated. This involves a not too difficult integration, which will not be carried out here. If E_d is small compared with V_{pm}, as is usually the case, the ripple factor will be approximately the same as for the high-vacuum-tube circuit.

14-7. Half-wave Rectifier with a Capacitor Filter. When rectifiers are to be used as sources of d-c power for equipment such as amplifiers, radio receivers, or radio transmitters, they must be able to supply direct voltages with a very small ripple. Hence some type of filtering must be used.

The filter is a frequency-selective network connected between the rectifier and the load terminals, which tends to allow the direct current to pass through the load while stopping or diverting the a-c components from reaching it. Filters are combinations of L, C, and R elements, the simpler ones being a single capacitor shunted across the load or a single choke in series with the load. While these very simple arrangements are relatively ineffective, when compared with more complex arrangements, they are occasionally used and warrant some study.

Figure 14-11*a* shows a rectifier using a simple capacitor filter. Sufficient accuracy for an estimate of performance is obtained by replacing the rectifier tube by an ideal one; the resulting equivalent circuit appears in Fig. 14-11*b*.

During the portion of the cycle when v_p is positive, if the voltage across C is lower than v_p, current flows through the rectifier tube to the load and to the capacitor, charging the latter to the voltage V_{pm}. If R is very high in value, the capacitor voltage remains near the value V_{pm} as the voltage v_p recedes and reverses, because the charge on C is trapped by the unilateral rectifier. There is, of course, a small current flow through R, and a slight loss in capacitor and load voltages results. On the next positive swing of v_p this small loss is restored, and the load voltage remains practically constant at a value V_{pm}.

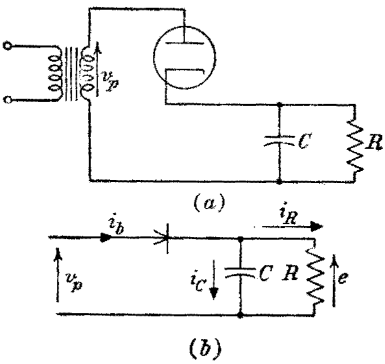


FIG. 14-11. (*a*) Half-wave rectifier with a simple capacitor filter. (*b*) Its approximate equivalent circuit.

Now to be practical we know that in many applications R cannot be very large. Hence the very simple analysis just made often cannot be applied. On the other hand, it is just as practical to insist that R should not be very small either. The reason for this is that if R is small, the capacitor will discharge through R almost as fast as v_p recedes, until e_C reaches zero. There it remains until later in the cycle, when v_p becomes positive again and recharges the capacitor. The resulting waveform of load voltage is practically as ripple-laden as though no filter were used. Actually it is not the size of R alone which counts, as an increase in capacitor size or in frequency will offset a decrease in the value of R. Hence ωRC is the quantity which should be kept reasonably large.

Let us continue the analysis. First we assume

$$v_p = V_{pm} \sin \omega t \tag{14-42}$$

During the portion of the cycle when the tube is conducting

$$i_R = \frac{V_{pm}}{R} \sin \omega t \tag{14-43}$$

$$i_C = C\frac{dv_p}{dt} = \omega C V_{pm} \cos \omega t \tag{14-44}$$

and

$$i_b = i_R + i_C = \frac{V_{pm}}{R} \sin \omega t + \omega C V_{pm} \cos \omega t \tag{14-45}$$

The tube ceases to conduct when $i_b = 0$, which we shall say occurs at the extinction angle $\omega t = \theta_2$. From Eq. (14-45) we can show that

$$\theta_2 = \arctan(-\omega RC) \tag{14-46}$$

Figures 14-12 and 14-13 show waveforms represented by these equations for i_R, i_C, and i_b. The portions actually used are drawn with heavy lines.

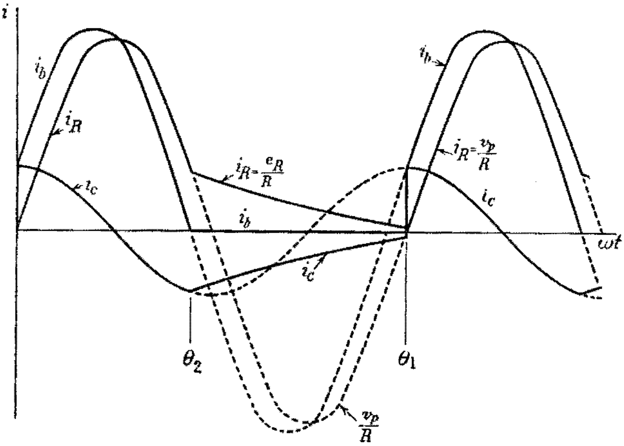


FIG. 14-12. Waveforms for i_R, i_C, and i_b; RC small.

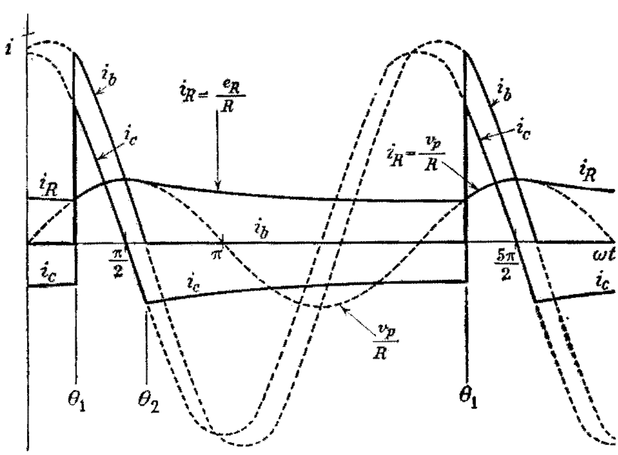


FIG. 14-13. Waveforms for i_R, i_C, and i_b; RC moderately large.

These two sets of curves differ in that one has ωRC small, while for the other ωRC is moderately large.

Beginning at θ_2, the capacitor discharges through R in its characteristic exponential manner until such time later in the cycle when v_p again becomes positive and equals e_C and e_R. At this time, say when $\omega t = \theta_1$, called the ignition angle, the capacitor begins to recharge and the currents i_R, i_C, and i_b again take their sinusoidal waveforms described above.

Between $\omega t = \theta_2$ and $\omega t = \theta_1$,

$$i_b = 0 \tag{14-47}$$

$$i_C = -A\epsilon^{-t/RC} \tag{14-48}$$

and

$$i_R = +A\epsilon^{-t/RC} \tag{14-49}$$

where A is an undetermined constant. Graphs of these equations are inserted in Figs. 14-12 and 14-13 to complete the waveforms of operation.

Examination of the waveforms of Fig. 14-12 shows i_b to be practically a half sinusoid. i_R has a waveshape which shows the presence of a high percentage of ripple. As indicated earlier, a small value of RC gives results little better than those obtained without a filter.

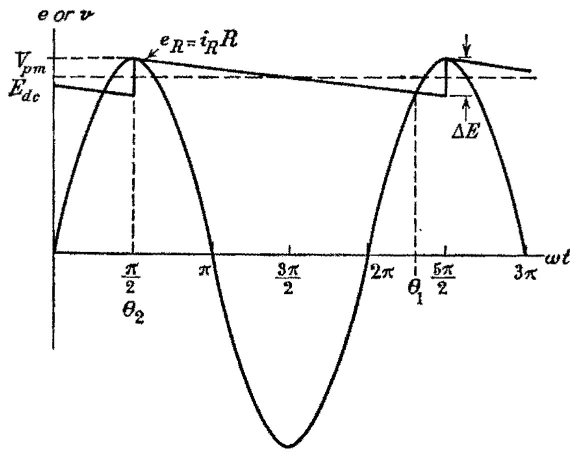


FIG. 14-14. An approximate saw-tooth representation of e_R in Fig. 14-13.

The practical case is that illustrated in Fig. 14-13 for a moderately large value of RC. The items of principal interest are values of I_{dc}, of the peak tube current $i_{b_{max}}$, and of the ripple factor γ. All these quantities may be readily determined if the ignition angle θ_1 and the extinction angle θ_2 are known. An expression for the latter has already been obtained in Eq. (14-46). To obtain θ_1 by simultaneous solution of the equations written previously is quite possible, at least graphically. However, for the practical case it is much easier to make further simplifying assumptions. Let us replace the exponential curve of i_R between θ_2 and θ_1 by an approximate linear one extending from $\theta = \pi/2$ to $\theta = 5\pi/2$. θ_2 is slightly greater than $\pi/2$, while θ_1 is somewhat less than $5\pi/2$, as shown in Fig. 14-14. Because of the sag in the exponential curve, the value of i_R at θ_1 is approximately the same as the minimum value on the straight line at $\theta = 5\pi/2$. The straight-line locus for i_R implies an assumption of a linear drop in capacitor and resistor voltage from a maximum value of

V_{pm} to a minimum value which we may denote as $V_{pm} - \Delta E$. A drawing of this assumed voltage waveform is given in Fig. 14-14. We can compute the average voltage by writing

$$E_{dc} \approx V_{pm} - \frac{1}{2}\Delta E = V_{pm} - \frac{1}{2}\frac{\Delta Q_C}{C} = V_{pm} - \frac{1}{2C} I_{dc} \frac{1}{f}$$
$$= V_{pm} - \frac{1}{2C}\frac{E_{dc}}{R}\frac{1}{f} \quad (14\text{-}50)$$

from which

$$E_{dc} \approx \frac{V_{pm}}{1 + (1/2fRC)} = \frac{V_{pm}}{1 + (\pi/\omega RC)} \quad (14\text{-}51)$$

I_{dc} can be obtained by Ohm's law.

Since

$$i_{\min} = 2I_{\text{ave}} - i_{\max} = 2I_{dc} - i_{\max} \quad (14\text{-}52)$$
$$i_{\min} = 2\frac{V_{pm}}{R[1 + (\pi/\omega RC)]} - \frac{V_{pm}}{R} = \frac{V_{pm}}{R}\frac{1 - (\pi/\omega RC)}{1 + (\pi/\omega RC)} \quad (14\text{-}53)$$

Assuming the current at θ_1 is the same as $i_{\min}$, we can substitute this value into Eq. (14-43) to determine the value of $\omega t = \theta_1$, suitable for large values of RC. Thus

$$\theta_1 \approx \arcsin \frac{1 - (\pi/\omega RC)}{1 + (\pi/\omega RC)} \quad (14\text{-}54)$$

θ_1 is an angle in the first quadrant.

We can now calculate the approximate value of the peak plate current. Examination of Fig. 14-13 shows that when ωRC is large, $i_{b_{\max}}$ is less than the peak value of the sinusoidal wave. Its value is obtained from Eq. (14-45) as

$$i_{b_{\max}} \approx V_{pm}\left(\frac{1}{R}\sin\theta_1 + \omega C\cos\theta_1\right) \quad (14\text{-}55)$$

For a fixed source frequency the value of $i_{b_{\max}}$ is increased by lowering R or by raising C. Its value should be checked against the specified peak-current rating of the tube as given by the manufacturer. For a gas tube, damage to the cathode results if this rating is exceeded. For a high-vacuum tube no damage results, but the direct-voltage output is lower than that calculated because of the inability of the cathode to supply the required peak charging current. In other words, the internal resistance R_p of the tube may be too high to be ignored in deriving the equations for performance.

In order to determine the ripple factor, one may first determine the effective current by integration. However, in the practical case sufficient accuracy is obtained by assuming the waveform to be saw-tooth. It may be easily shown that the effective value of the alternating components of a saw-tooth wave is $a/\sqrt{3}$, where $2a$ is the peak-to-peak value. Using this formula, we can write an expression for the ripple voltage across the load.

$$E_{ac} \approx \frac{V_{pm} - E_{dc}}{\sqrt{3}} = \frac{1}{\sqrt{3}}\left[V_{pm} - \frac{V_{pm}}{1 + (\pi/\omega RC)}\right] = \frac{V_{pm}}{\sqrt{3}}\frac{1}{1 + (\omega RC/\pi)} \qquad (14\text{-}56)$$

From this and Eq. (14-51), the ripple factor becomes

$$\gamma = \frac{E_{ac}}{E_{dc}} \approx \frac{1}{\sqrt{3}}\frac{1 + (\pi/\omega RC)}{1 + (\omega RC/\pi)} = \frac{\pi}{\sqrt{3}\,\omega RC} \times 100\% \qquad (14\text{-}57)$$

If a full-wave rectifier is substituted for the half-wave one, the waveforms of Figs. 14-12 and 14-13 must be amended to allow for current components furnished by the second diode unit. The exponential i_R will drop less than before because the second diode recharges the capacitor almost a half cycle earlier than a single diode did. In deriving the formulas the same method as before can be employed, the only essential difference being the use of time $1/2f$ instead of $1/f$. Thus for a full-wave rectifier with a smoothing capacitor across the load,

$$E_{dc} \approx \frac{V_{pm}}{1 + (\pi/2\omega RC)} \qquad (14\text{-}58)$$

$$\theta_1 \approx \arcsin\frac{1 - (\pi/2\omega RC)}{1 + (\pi/2\omega RC)} \qquad (14\text{-}59)$$

where θ_1 is a first quadrant angle for the tube being considered. The formula for $i_{b\max}$ is again given by Eq. (14-55). The ripple factor for the full-wave case is

$$\gamma = \frac{\pi}{2\sqrt{3}\,\omega RC} \times 100\% \qquad (14\text{-}60)$$

Example. A full-wave rectifier with $V_p = 230$ volts at 60 cps supplies direct current to a 10,000-ohm load. Find the size smoothing capacitor needed to reduce the ripple to 5 per cent. Also find E_{dc}, I_{dc}, and the peak tube current.

Solution. $\gamma = 0.05 = \pi/2\sqrt{3}\,\omega RC$. Therefore

$$\frac{\pi}{2\omega RC} = 0.0866 \qquad \text{and} \qquad C = \frac{\pi}{0.0866 \times 754 \times 10{,}000} = 4.8\ \mu f$$

$$E_{dc} = \frac{\sqrt{2} \times 230}{1 + 0.0866} = 300 \text{ volts} \qquad I_{dc} = \frac{300}{10{,}000} = 30 \text{ ma}$$

$$\theta_1 = \arcsin\frac{1 - 0.0866}{1 + 0.0866} = 57.5°$$

$$i_{b\max} = \sqrt{2} \times 230\left(\frac{\sin 57.5°}{10{,}000} + \frac{\pi}{1732}\cos 57.5°\right) = \sqrt{2} \times 230\,(0.0842 + 0.975) \times 10^{-3} = 344 \text{ ma}$$

A 5Y3 double diode with a peak-current rating of 375 ma per plate is satisfactory. The allowable peak inverse plate voltage is 1400 volts, while the applied value is approximately $\sqrt{2} \times 230 \times 2 = 650$ volts.

14-8. Half-wave Rectifier Using a Series-inductor Filter. A series inductor serves as a filter in rectifier circuits because it stores energy when the current is rising and releases it as the current falls. Thus there is a tendency to maintain the current constant in value. The actual inductance may be that intentionally inserted to serve this purpose, or it may be inherently part of the circuit, as a component of the load, as the leakage inductance of a power transformer, or as an element in the power-generating system.

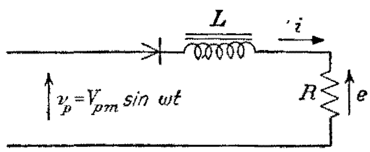


FIG. 14-15. Equivalent circuit for a rectifier with a smoothing inductor.

If we neglect the tube drop or the tube resistance, we may draw the equivalent circuit in the form shown in Fig. 14-15. For this circuit the loop voltage equation with sinusoidal applied voltage may be written as

$$L \frac{di}{dt} + Ri = V_{pm} \sin \omega t \tag{14-61}$$

valid for only positive values of i. The complementary function in the solution for the current is

$$i = A\epsilon^{-Rt/L} \tag{14-62}$$

where A is an arbitrary constant. The particular integral is the familiar expression

$$i = \frac{V_{pm}}{\sqrt{R^2 + \omega^2 L^2}} \sin (\omega t - \theta) \tag{14-63}$$

where

$$\theta = \arctan \frac{\omega L}{R} \tag{14-64}$$

The complete solution becomes

$$i = A\epsilon^{-Rt/L} + \frac{V_{pm}}{\sqrt{R^2 + \omega^2 L^2}} \sin (\omega t - \theta) \tag{14-65}$$

The sinusoidal component of this current can become zero and reverse after a half cycle, but the exponential component is always positive. Hence there is a question whether the total current does or does not become zero during the cycle. As L is made larger, R remaining fixed, the more slowly the exponential component dies, and thus for higher values of L it seems that it may be possible for the current to flow continuously. As a matter of fact this is the case for full-wave rectifiers, but for half-wave ones the current always cuts off. Since this is the case, at the end of the cutoff period, voltage is applied to the R-L circuit under the condition that there is no stored energy in the inductor. Consequently the current begins to rise as the voltage rises, without the lag which occurs in the sinusoidal case. Figure 14-16 shows sketches of waveforms of i for various values of $\omega L/R$.

We may now proceed to determine the cutoff angle θ_2. This will be used as a limit in the integration necessary to determine I_{dc} and the ripple.

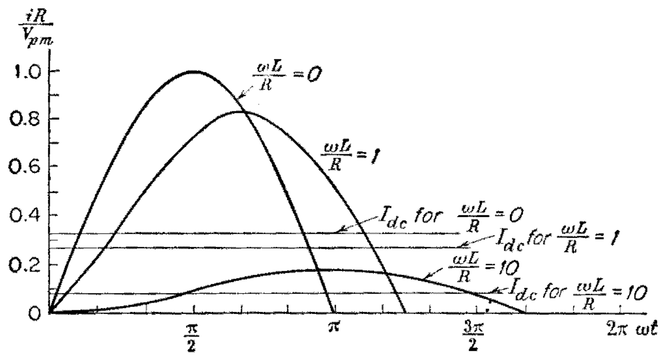


FIG. 14-16. Waveforms of current for a half-wave rectifier with a smoothing inductor.

Assume $i = 0$ when $\omega t = 0$. We may then evaluate A in Eq. (14-65) as

$$A = \frac{V_{pm}}{\sqrt{R^2 + \omega^2 L^2}} \sin \theta = \frac{V_{pm}\omega L}{\sqrt{R^2 + \omega^2 L^2}\,\sqrt{R^2 + \omega^2 L^2}} \tag{14-66}$$

We may substitute this value for A and arctan $(\omega L/R)$ for θ in Eq. (14-65):

$$i = \frac{V_{pm}}{\sqrt{R^2 + \omega^2 L^2}} \left[\frac{\omega L}{\sqrt{R^2 + \omega^2 L^2}}\, \epsilon^{-Rt/L} + \sin\left(\omega t - \arctan \frac{\omega L}{R}\right)\right] \tag{14-67}$$

At $\omega t = \theta_2$ suppose the current again becomes zero. Substituting into Eq. (14-67) and rearranging, we obtain the relation

$$-\frac{\omega L/R}{\sqrt{1 + (\omega L/R)^2}}\, \epsilon^{-\theta_2/(\omega L/R)} = \sin\left(\theta_2 - \arctan \frac{\omega L}{R}\right) \tag{14-68}$$

This equation should yield the value of θ_2 if $\omega L/R$ is known, but it is impossible to obtain an explicit expression for θ_2. We can, however, plot the relationship of θ_2 vs. $\omega L/R$ by the tedious process of assuming one value of $\omega L/R$, plotting the left member vs. θ_2, plotting the right member vs. θ_2, and finding the intersection of these two graphs and hence the value of θ_2 corresponding to the assumed value of $\omega L/R$. This is repeated for other values of $\omega L/R$. The results are shown in the finished graph of Fig. 14-17. Note that the conduction period does increase as $\omega L/R$ increases, whether because of an increase in L or a decrease in R.

A formula for I_{dc} can be derived in the usual manner by averaging the current over a cycle. For this purpose i may be expressed by using either Eq. (14-67) or Eq. (14-61). Investigation proves that the latter proce-

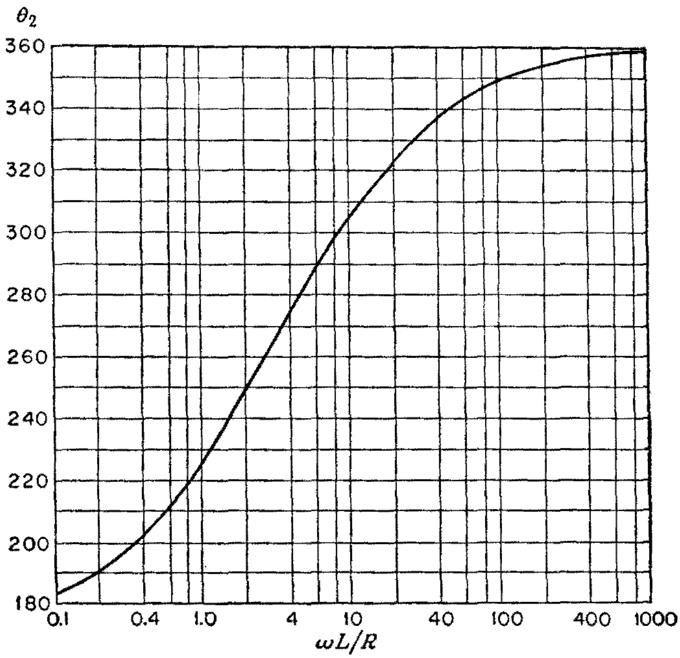


FIG. 14-17. Variation of extinction angle θ_2 with $\omega L/R$ for a half-wave rectifier using a smoothing inductor.

dure is shorter. Hence, by the use of (14-61),

$$
\begin{aligned}
I_{dc} &= \frac{1}{2\pi}\int_0^{2\pi} i\,d\omega t = \frac{1}{2\pi}\int_0^{\theta_2}\left[\frac{V_{pm}}{R}\sin\omega t - \frac{L}{R}\frac{di}{dt}\right]d\omega t \\
&= \frac{1}{2\pi}\left[-\frac{V_{pm}}{R}\cos\omega t - \frac{\omega L}{R}\,i\right]_0^{\theta_2} = \frac{1}{2\pi}\left[\frac{V_{pm}}{R}(1-\cos\theta_2) - 0\right] \\
&= \frac{V_{pm}}{2\pi R}(1-\cos\theta_2) \qquad (14\text{-}69)
\end{aligned}
$$

Values of I_{dc} for various values of $\omega L/R$ have been drawn in Fig. 14-16. Examination of these values together with the corresponding waveforms of instantaneous currents shows that the ratio of peak current to I_{dc} is not particularly high. The highest ratio is 3.14, obtained for $\omega L/R = 0$. Thus peak current is not an important item in rectifier operation with a smoothing inductor.

From Eq. (14-69) it follows that

$$E_{dc} = \frac{V_{pm}}{2\pi}(1 - \cos \theta_2) \tag{14-70}$$

The value of the ripple factor at the load is given by

$$\gamma = \frac{I_{ac}}{I_{dc}} = \frac{\sqrt{I^2 - I_{dc}^2}}{I_{dc}} = \sqrt{\frac{I^2}{I_{dc}^2} - 1} = \sqrt{\frac{(1/2\pi)\int_0^{2\pi} i^2 d\,\omega t}{I_{dc}^2} - 1} \tag{14-71}$$

The work involved in this development is long and tedious. The result is

$$\gamma = \sqrt{\frac{\pi \cos \theta}{(1 - \cos \theta_2)^2}[\theta_2 \cos \theta - \sin \theta_2 \cos (\theta + \theta_2)] - 1} \tag{14-72}$$

Values of γ are plotted against $\omega L/R$ in Fig. 14-18. This graph shows that as $\omega L/R$ is increased, the value of γ first decreases slightly and then

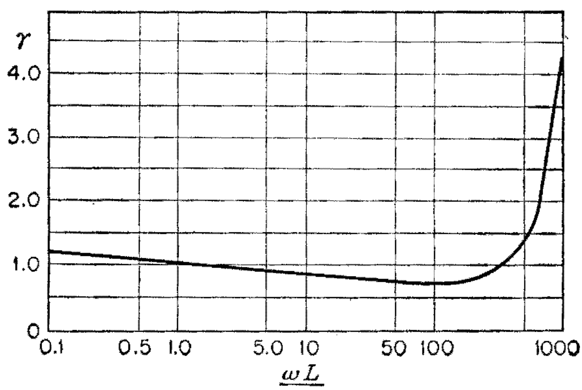


FIG. 14-18. Variation of the ripple factor γ with $\omega L/R$ for a half-wave rectifier using a smoothing inductor.

increases. This last rise is due to the fact that while both the direct and the ripple currents decrease, the former decreases faster than does the ripple. The filtering efficiency of a smoothing inductor with a half-wave rectifier is on the whole very poor, and this type of filter is not recommended for applications requiring a good quality of direct current. However, a smoothing inductor is useful with a full-wave rectifier (see Prob. 17 at the end of this chapter) and with polyphase rectifiers.

14-9. Full-wave Rectifier with an L-section Filter. So far none of the filters that we have studied would prove to be very satisfactory sources of direct current for communications equipment or for other applications where a ripple-free output is desirable.

A more satisfactory device for such uses is the full-wave rectifier with an L-section filter. The circuit diagram and its equivalent circuit are shown in Fig. 14-19. Note that an ideal tube is used in the equivalent circuit in order to simplify the analysis.

Under normal operating conditions the current through the inductor never falls to zero at any time during the cycle. If the current to the filter flows at all times, one or the other of the tubes must always be passing current and hence the voltage e_f applied to the filter input must be the rectified transformer voltage, which is a series of half sinusoids, as was shown in Fig. 14-6. The equation for this voltage from 0 to π can be written as $e_f = V_{pm} \sin \omega t$, which would be useful in determining I_{dc}. However, the filter is a complex network of elements, some of which have frequency-dependent impedances, and it becomes necessary to have a Fourier analysis of e_f so that the effect of each frequency component can be found. In this case the Fourier series takes the form (see Art. 6-16 and Prob. 15 in Chap. 6)

$$e_f = \sqrt{2}\, V_p \left(\frac{2}{\pi} - \frac{4}{3\pi} \cos 2\omega t - \frac{4}{15\pi} \cos 4\omega t - \cdots \right) \tag{14-73}$$

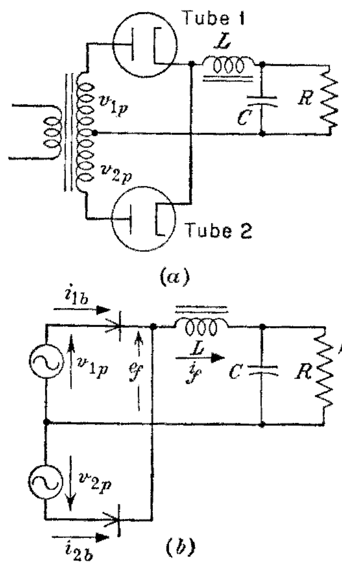


FIG. 14-19. (a) A full-wave rectifier with an L-section filter and a resistance load. (b) Equivalent circuit for (a).

It is now a relatively simple matter to consider each voltage component as being applied separately to the input of the filter and to calculate the currents which flow and the voltages across the load.

The first term of the series is the d-c input to the filter. It is also the d-c output voltage since we are assuming for the present that there is no d-c drop in the filter or in the tube. Thus

$$E_{dc} = \frac{2\sqrt{2}\, V_p}{\pi} \tag{14-74}$$

The direct current flowing through the d-c path is then given by the equation

$$I_{dc} = \frac{2\sqrt{2}\, V_p}{\pi R} \tag{14-75}$$

We may calculate the second-harmonic current flowing into the filter by dividing E_{f_2} (the effective value of the second-harmonic voltage) by the filter's input impedance $\mathbf{Z}_{f_2}$ for this frequency. This yields

$$\mathbf{I}_{f_2} = \frac{4}{3\pi}\frac{\mathbf{V}_p}{\mathbf{Z}_{f_2}} \tag{14-76}$$

where

$$\mathbf{Z}_{f_2} = j2\omega L + \frac{1}{(1/R) + j2\omega C} \tag{14-77}$$

In the well-designed filter the capacitor is large in order to bypass the a-c components around the load. Also the impedance of L will be large in order to offer a high impedance to these same components of current. If we specify the condition at this second-harmonic frequency,

$$X_L \gg X_C \ll R \tag{14-78}$$

Eq. (14-77) simplifies to

$$\mathbf{Z}_{f_2} \approx j2\omega L \tag{14-79}$$

Substituting this into Eq. (14-76), we obtain

$$I_{f_2} \approx \frac{4}{3\pi}\frac{V_p}{2\omega L} \tag{14-80}$$

Since the reactance X_C of the capacitor is very small compared with the resistance of the load, practically all the a-c components of current will pass through C instead of R. Hence we may write

$$E_2 \approx I_{f_2}X_C = \frac{4V_p}{3\pi(4\omega^2LC)} \tag{14-81}$$

as the magnitude of the second-harmonic voltage across the load.

If the relation expressed in (14-78) is true for the second harmonic, it is an even better approximation for the fourth. The fourth-harmonic voltage across the load can be found by a process similar to that for the second. Such a process yields

$$E_4 = \frac{4}{15\pi}\frac{V_p}{16\omega^2LC} \tag{14-82}$$

If we now find the ratio of the fourth-harmonic voltage to the second-harmonic voltage, many of the factors cancel out, yielding

$$\frac{E_4}{E_2} = \frac{1}{20} = 5\% \tag{14-83}$$

Thus the ripple voltage is predominantly second harmonic, and for the sake of simplicity we shall assume it consists entirely of this harmonic, The ripple factor can be written as

$$\gamma = \frac{E_2}{E_{dc}} = \frac{\sqrt{2}}{12\omega^2 LC} \tag{14-84}$$

Note that the ripple factor is independent of the load resistance. A graph of the ripple factor as a function of $\omega^2 LC$ is shown in Fig. 14-20, curve *a*.

Earlier in this article the statement was made that under usual operating conditions the current through the inductor does not fall to zero in this type of rectifier and filter. This is not necessarily always true, as can easily be proved by open-circuiting the load and observing the direct output voltage. Under these circumstances, since no charge is being removed from the capacitor, the latter will charge up to the peak value of the a-c supply voltage. E_{dc} will then equal V_{pm}, and no tube current will flow. As the circuit is loaded, the output voltage decreases rather rapidly until E_{dc} reaches a certain value. The output voltage then becomes rather constant and approximately equal to the value calculated from Eq. (14-74). This means that when the load current becomes too small, the regulation of the circuit becomes very poor. The problem is to find how small a direct current can be supplied and yet have good voltage regulation. Steps may then be taken to prevent the load current from falling below this value.

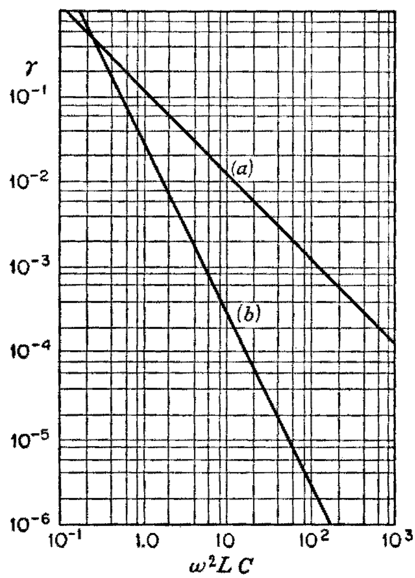


FIG. 14-20. Ripple factor as a function of $\omega^2 LC$ for a full-wave rectifier with (*a*) a single L-section filter, (*b*) a double L-section filter.

If the inductor current is observed by means of an oscilloscope, it can be seen that when the direct current through the inductor (which is also the load current I_{dc}) becomes smaller than the peak value of the ripple current through this same component, the value of E_{dc} begins to rise rapidly as the load resistor is increased in value. This occurrence is related to the fact that the tube current has ceased to be continuous. Our problem is to find some means of preventing I_{dc} from becoming less than the peak value of the inductor ripple current.

The assumption has already been made that the current input to the filter contained no higher harmonics than the second. We may then write the expression

$$i_f = I_{dc} + \sqrt{2}\, I_{f_2} \cos 2\omega t \tag{14-85}$$

As I_{dc} becomes smaller, the inductor current begins to cut off during a portion of the cycle if the negative peak value of i_{f_2} is numerically greater than I_{dc}. The critical condition is therefore

$$I_{dc} = \sqrt{2}\, I_{f_2} = I_{f_2m} \tag{14-86}$$

Substituting values for I_{dc} and I_{f_2} from Eqs. (14-75) and (14-80) into (14-86), we obtain

$$\frac{2\sqrt{2}\, V_p}{\pi R} = \frac{4\sqrt{2}}{3\pi}\frac{V_p}{2\omega L} \tag{14-87}$$

Solving for R, we have the relation

$$R = 3\omega L \tag{14-88}$$

If the frequency of the supply voltage is 60 cps, this becomes $R = 1131L$. An approximate form of this may be written as

$$R \approx 1000L \tag{14-89}$$

where R is in ohms and L is in henrys.

In some practical circuits the load current may vary all the way from full load to zero, and the regulation of the circuit will be poor unless something is done to prevent I_{dc} from falling below $\sqrt{2}\, I_{f_2}$. This is easily done by permanently connecting a bleeder resistor across the output of the filter. The resistance of this bleeder should be not greater than $1000L$ ohms. Now if the external load is removed entirely, the bleeder will still continue to draw sufficient current to ensure good regulation.

The use of the bleeder permits filter capacitors with lower voltage ratings to be used. Without the bleeder the capacitor should have a d-c rating of at least V_{pm} volts, while with a bleeder it need be only as large as the direct-voltage output of the circuit, which is $2/\pi$ times as large as V_{pm}. The bleeder also satisfies a safety requirement. Unless there is always a d-c path between the terminals of the filter, the capacitor may hold its charge for a considerable length of time after the a-c input voltage is removed, and hence there is danger of electrical shock to anyone assuming the circuit to be dead.

The formulas already developed can easily be changed to adapt them to the case where the plate resistance of the tube and the resistance R_{ch} of the choke are to be taken into consideration. Equation (14-75) changes to

$$I_{dc} = \frac{(2\sqrt{2}/\pi)V_p}{R + R_p + R_{ch}} \tag{14-90}$$

and Eq. (14-74) changes to the form

$$E_{dc} = \frac{2\sqrt{2}\,V_p}{\pi}\,\frac{R}{R + R_p + R_{ch}} \tag{14-91}$$

The equation for the ripple voltage will remain approximately the same because R_p and R_{ch} are small compared with X_{ch} in well-designed circuits.

When gas tubes are considered, the tube drop acting as a battery merely serves to reduce the voltage input to the filter by a constant amount equal to E_d. Equations (14-75) and (14-74) then, respectively, take forms

$$I_{dc} = \frac{(2\sqrt{2}/\pi)V_p - E_d}{R + R_{ch}} \tag{14-92}$$

and

$$E_{dc} = \left(\frac{2\sqrt{2}}{\pi}\,V_p - E_d\right)\frac{R}{R + R_{ch}} \tag{14-93}$$

The ripple voltage will again remain about the same since E_d does not affect the harmonic content of the filter input voltage.

It is sometimes desirable to use two L sections in cascade instead of a single section. This arrangement results in much better filtering. The harmonic output voltage for the second section can easily be found by using the harmonic output of the first section as the input to the second. By means of the assumptions made in (14-78) we can write the output second-harmonic voltage as

$$E_2 = \frac{4}{3\pi}\,\frac{V_p}{4\omega^2L_1C_1}\,\frac{1}{4\omega^2L_2C_2} \tag{14-94}$$

The ripple factor then becomes

$$\gamma = \frac{\sqrt{2}}{48}\,\frac{1}{\omega^2L_1C_1}\,\frac{1}{\omega^2L_2C_2} \tag{14-95}$$

A graph showing the ripple factor for a two-section filter has also been plotted in Fig. 14-20. In order to use the same units for the abscissa as were used for the single section, the provision was made that the two sections have equal LC values. These curves make it quite obvious that a two-section filter is far superior to a single-section one if small values of γ are considered.

14-10. Rectifiers with a Π-section Filter. A capacitor-input, or Π-section, filter is often used where the voltage output of an L-section filter would be somewhat too low. The output voltage of the Π-section filter is almost equal to the peak value of the source voltage, while the output voltage of the L section is approximately equal to $2/\pi$ of the peak of the source voltage.

Figure 14-21 shows the circuit for such a filter used with a half-wave rectifier. Briefly, its operation is as follows: Capacitor C_1 draws a pulse of charging current once each cycle, charging up to a value slightly less than V_{pm}. Soon after the source voltage passes through its peak value and begins to decrease, the tube cuts off and the capacitor starts to discharge through the L-section filter that follows it. The voltage of C_1 continues to decrease while the source voltage passes through its negative half cycle and then again becomes positive and equal to the capacitor voltage. Thereupon the tube starts to conduct and to charge C_1 back to its former voltage. This is quite similar to the action encountered in half-wave rectifiers with capacitor filtering.

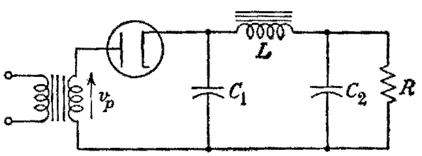


FIG. 14-21. A half-wave rectifier with a Π-section filter.

If the waveshape of the varying voltage across capacitor C_1 could be expressed as a Fourier series, the analysis of this circuit would be rather simple. Because of the fact that it is difficult to determine the time when the tube begins conduction and the time when it cuts off, the exact analysis becomes quite complicated. Therefore we shall use a rather simple approximate method which gives results good enough for most practical purposes.

We continue with the half-wave case. The voltage across C_1 varies in approximately the same manner as did e_R (Fig. 14-14) in the case of the simple capacitor filter. Thus, as a first approximation, $E_{dc} = V_{pm}$. However, if the current drawn by the load is at all appreciable, a better approximation is to use Eq. (14-51) to express the direct voltage at the input to the filter. Thus the direct voltage across the load becomes

$$E_{dc} = \frac{V_{pm}}{1 + (\pi/\omega RC)} \frac{R}{R + R_{ch}} \tag{14-96}$$

where R_{ch} is the total d-c resistance of the choke or chokes used.

To determine the ripple factor, we shall consider the Fourier analysis of the current entering the filter rather than that of the voltage across the filter. This current, which in this case is the same as the tube current, would be expected to have a waveshape somewhat similar to that found for the capacitor filter in Art. 14-7. However, when the waveshape is observed by means of an oscilloscope, it will be found to look more like that shown in Fig. 14-22. The reasons for this variation in shape are due to the inductances of the transformer and choke.

The constant, or d-c, term in the Fourier series is given by

$$I_{f_{dc}} = \frac{1}{2\pi} \int_0^{2\pi} i_f \, d\omega t \tag{14-97}$$

The peak value of the fundamental-frequency current is

$$I_{f_1m} = \frac{2}{2\pi}\int_0^{2\pi} i_f \cos \omega t \, d\omega t \tag{14-98}$$

Equations (14-97) and (14-98) are obtained from the general Fourier method of calculating the coefficients for the series (Art. 6-16). The current through the tube in a well-designed power supply flows for only a small portion of the cycle. During this time $\cos \omega t$ in Eq. (14-98) is near its peak value and approximately equal to unity. Hence we may replace $\cos \omega t$ in this equation by unity to obtain a convenient approximation

$$I_{f_1m} \approx \frac{2}{2\pi}\int_0^{2\pi} i_f \, d\omega t \tag{14-99}$$

But by comparing this with Eq. (14-97) we see that it may be written as

$$I_{f_1m} \approx 2I_{f_{dc}} = 2I_{dc} \tag{14-100}$$

or

$$I_{f_1} = \sqrt{2}\, I_{dc} \tag{14-101}$$

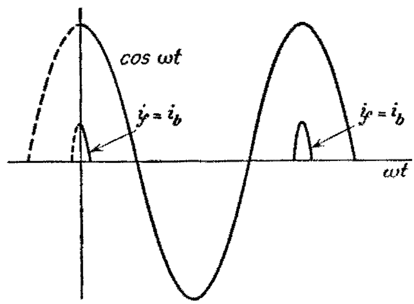


FIG. 14-22. Waveshape of the tube current for a half-wave rectifier with a Π-section filter.

In a practical filter C_1 is very large and hence has a very low reactance compared with the reactance of the L section that follows it. Therefore, most of the alternating current will flow through C_1. The fundamental voltage across the first capacitor is then approximately given by

$$E_{f_1} \approx \sqrt{2}\,\frac{I_{dc}}{\omega C_1} \tag{14-102}$$

If this is a well-designed filter, it is a reasonable assumption that the reactance X_{C_2} of the second capacitor is very small compared with R. Hence the fundamental-frequency voltage across C_1 will be applied to C_2 and L in series, and we may write for the inductor current

$$I_{ch} \approx \frac{E_{f_1}}{\omega L} = \frac{\sqrt{2}\, I_{dc}}{\omega^2 L C_1} \tag{14-103}$$

and the fundamental output voltage is

$$E_1 \approx I_{ch} X_{C_2} = \frac{\sqrt{2}\, I_{dc}}{\omega^3 L C_1 C_2} \tag{14-104}$$

The effect of each higher harmonic in the input current can be investigated by a slight modification of Eq. (14-98). For the nth harmonic this becomes

$$I_{f_nm} = \frac{2}{2\pi}\int_0^{2\pi} i_f \cos n\omega t \, d\omega t \tag{14-105}$$

The derivation is exactly the same as for the fundamental except that the reactances of the various components must be computed for the frequency of the harmonic. The final result for the nth-harmonic output voltage will then be

$$E_n \approx \sqrt{2}\,\frac{I_{dc}}{(n\omega)^3 L C_1 C_2} \tag{14-106}$$

For the case of the second harmonic we may compare Eqs. (14-106) and (14-104) to see that the ratio of the magnitude of the second harmonic to that of the fundamental is

$$\frac{E_2}{E_1} \approx \frac{1}{8} \tag{14-107}$$

This is a justification for assuming that the ripple voltage across the load is entirely of the fundamental frequency.

The ripple factor can be calculated from the ratio of E_1 to E_{dc}. Thus

$$\gamma \approx \frac{E_1}{E_{dc}} = \frac{\sqrt{2}\,I_{dc}}{\omega^3 L C_1 C_2}\,\frac{1}{I_{dc}R} = \frac{\sqrt{2}}{\omega^3 L C_1 C_2 R} \tag{14-108}$$

If it is desired to use a Π-section filter with a full-wave rectifier, the foregoing procedure can be followed in order to find the expressions for E_{dc} and the ripple factor. From the analysis for the case of the simple capacitor filter (full wave) we can write the direct output voltage across the load as

$$E_{dc} = \frac{V_{pm}}{1 + (\pi/2\omega RC)}\,\frac{R}{R + R_{ch}} \tag{14-109}$$

Because of the two identical pulses of current per cycle, the lowest-frequency a-c component is the second harmonic. We may compute this second-harmonic ripple by using Eq. (14-106) and hence obtain the ripple factor as

$$\gamma \approx \frac{\sqrt{2}}{8\omega^3 L C_1 C_2 R} \tag{14-110}$$

This result is one-eighth that found for the half-wave rectifier.

14-11. Applications of Diode Rectifiers. We have now completed a theoretical study of diode rectifiers using various filter arrangements. In general it has been found that the unilateral characteristic of the rectifier has caused considerable difficulty in the process of analysis, even when this has been simplified by making suitable approximations. Without this tedious work, however, we would not be in our present advantageous position of knowing the characteristics of the various circuits. We can now better judge the applications to which the devices can be put.

Uses of the high-vacuum rectifier with a resistance load and without a filter are few. It is possible to employ it in electrolytic processes, but it would usually be better to use a gas diode instead. For hot-cathode gas-diode battery chargers the battery emf adds to the E_d of the tube. The current is limited by the leakage reactance of the transformer and perhaps by a series resistor. The theory is a modification of that of Art. 14-6.

Rectifiers used with X-ray or other high-voltage equipment often draw only a small current, and the filtering needs are adequately met by the use of a simple capacitor filter. The series-inductor filter is seldom deliberately used, but the filtering action described in Art. 14-8 occurs inadvertently in some circuits because of an inductive load or the leakage inductance of the transformer or of the inductance of the power-generating equipment.

The L- and Π-section filters are very effective when used with either high-vacuum or gas diodes. Between the two filters the latter produces a higher d-c output voltage for the same a-c supply voltage, but the regulation is poorer with changing load resistance. Furthermore, the high peak tube currents associated with the Π filter make it necessary that some impedance be placed in series with a gas tube to limit this current value. This impedance may be supplied in some cases by the power transformer in the form of resistance and leakage reactance. Other considerations may also be involved, but the choice between the two is often on an economic basis.

If the direct current to a constant load impedance is to be variable, the secondary of the power transformer may be provided with taps or alternatively the primary may be supplied with an alternating voltage of controlled value. It is often more satisfactory to replace the diode rectifier by one employing triodes, generally of the gas type.

14-12. Grid-controlled Rectifiers. When diodes are used as rectifiers, the usual methods for controlling the amount of direct current to a fixed load are either to vary the amplitude of the alternating source voltage or to place a variable resistance in series with the load. For many applications these methods would prove to be very unsatisfactory. A far better system might be to use thyratrons or ignitrons and control the firing angle.

Since the critical grid voltage for a thyratron can usually be taken as zero and since the breakdown voltage is usually small compared with the plate-supply voltage, no great error is introduced if we assume that the tube can have a maximum conduction angle of 180°. The average, or direct, current in the plate circuit of a thyratron which fires at an angle θ_1 and is supplying a resistance load R_L can easily be calculated if we assume the tube to have zero voltage drop across it. Hence

$$I_{dc_{\theta_1}} = \frac{1}{2\pi}\int_0^{2\pi} i_b \, d\omega t = \frac{1}{2\pi}\int_{\theta_1}^{\pi} \frac{V_{pm}}{R_L} \sin \omega t \, d\omega t = \frac{V_{pm}}{2\pi R_L}(1 + \cos \theta_1) \tag{14-111}$$

We may obtain the expression for the maximum obtainable direct current by setting $\theta_1 = 0$.

$$I_{dc_{\max}} = \frac{V_{pm}}{\pi R_L} \tag{14-112}$$

Thus we may vary the d-c output of the thyratron from the maximum value $I_{dc_{\max}}$ to zero by merely varying the angle θ_1 at which the tube fires. Ratio wise,

$$\frac{I_{dc_{\theta_1}}}{I_{dc_{\max}}} = \frac{1 + \cos\theta_1}{2} \tag{14-113}$$

14-13. D-C Control of Thyratrons. The circuit diagram for a thyratron connected for d-c control is shown in Fig. 14-23. Its firing characteristic is shown in Fig. 14-24. If the grid of the tube is biased to the voltage E_c shown in (*a*), the tube will not fire. As the negative grid bias is reduced toward zero, nothing happens until it reaches the value of E_c'. The grid bias then equals the critical value just as the plate voltage attains a maximum. The tube will then start to conduct and will pass current until the plate voltage falls to zero. Hence the plate current is represented by quarter sinusoids as shown in Fig. 14-24*b*. As the bias is reduced still further, the tube fires earlier in the cycle and the average plate current increases. The ignition angle may be advanced to the point where the tube conducts for 180°. This results in the maximum direct current that the tube can pass when using a particular source-voltage-and-load combination.

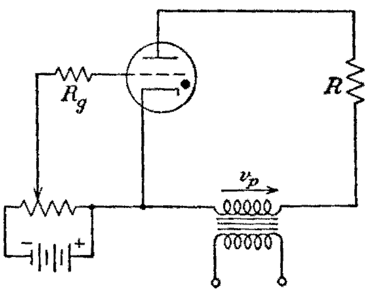


FIG. 14-23. Direct-current control of a thyratron.

This is a rather unsatisfactory type of control, not only because the ignition angle can be varied only from 0 to 90°, but also because the operation may be rather erratic. The actual firing characteristic is very close to the zero-voltage line and almost parallel to it. Hence a very slight drift in E_c may cause the firing angle to change considerably.

14-14. On-Off Control. Because of the above shortcomings, d-c control is used only in applications where the thyratron can act as an on-off switch and where continuous control of the average current is not needed. The tube action is then similar to that of a relay and a rectifier.

The circuit of Fig. 14-25*a* illustrates on-off control. When the switch S is open, the tube is biased negatively to such a value that the discharge cannot be established at any time during the cycle. When the switch is closed, the grid potential becomes practically that of the plate. Hence the tube will fire when the supply voltage becomes positive and will continue to conduct until the end of the half cycle. It fires at the beginning of every positive half cycle thereafter as long as the switch is closed. Figure 14-25*b* shows a similar circuit which fires as long as the switch is open but will not fire when the switch is closed.

14-15. Phase-shift Control. One of the most widely used and more satisfactory methods of varying the current through a thyratron is known

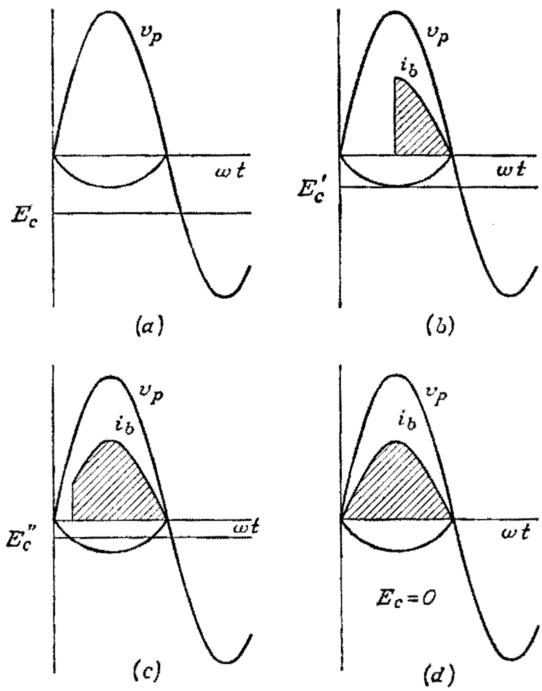


FIG. 14-24. Waveshapes illustrating d-c control of a thyratron.

as phase-shift control. This control is accomplished by applying an alternating voltage to the grid of the tube as well as to the plate. The phase angle between these two voltages is varied so that the grid voltage lags the plate voltage. An examination of Fig. 14-26 will disclose how changes in this phase angle affect the conduction angle of the thyratron. If the two voltages are in phase, the tube conducts for 180° and the direct current, which is the maximum obtainable, can be calculated by means of Eq. (14-112). If the grid voltage lags the plate voltage by an angle θ_1, the current can be calculated by using the relation (14-113). Should the grid voltage lead the plate voltage, the tube conducts for 180°, the grid having no control.

A simple way to shift the phase of an alternating voltage is to impress a polyphase set of voltages on the stator windings of a polyphase motor with a single-phase rotor. The phase of the voltage induced in the rotor will depend upon the position of the rotor, which may be rotated through 360°. However, such a phase shifter, even though available, is not

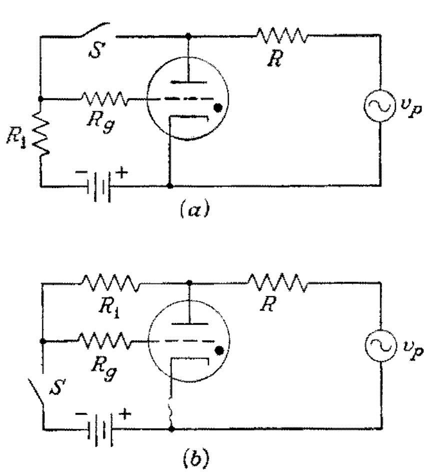


FIG. 14-25. On-off control of a thyratron.

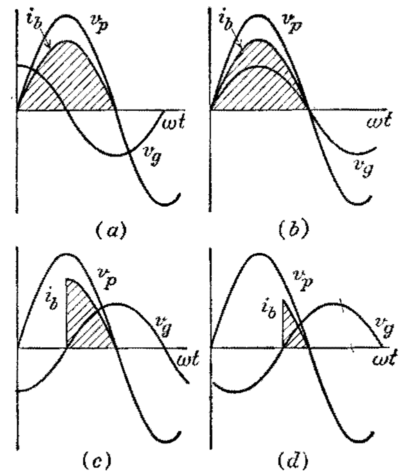


FIG. 14-26. Waveshapes illustrating phase-shift control.

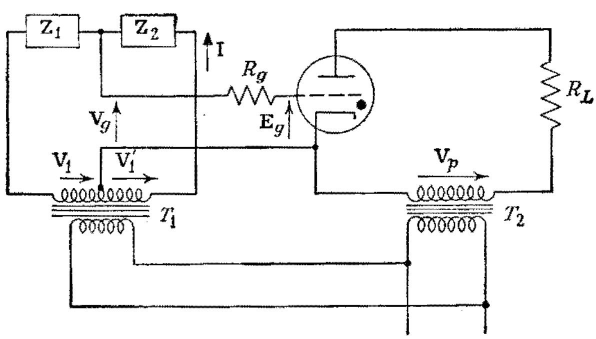


FIG. 14-27. A thyratron with phase-shift control.

always the most economical or the easiest to use; hence other methods are more common.

One much used method of obtaining a variable phase shift uses the circuit of Fig. 14-27. The transformer T_1 supplies two sinusoidal voltages $\mathbf{V}_1$ and $\mathbf{V}_1'$. These are equal in magnitude and are in phase. This transformer is used for control purposes only, and the power requirement is very small. T_2 is the transformer supplying the power to the load. Its secondary voltage $\mathbf{V}_p$ is in phase with $\mathbf{V}_1$.

The grid-supply voltage $\mathbf{V}_g$ is obtained between the center tap of T_1 and the junction between the two impedances $\mathbf{Z}_1$ and $\mathbf{Z}_2$. It is sinusoidal and is displaced by an angle θ from $\mathbf{V}_1$. The actual grid voltage e_g is equal to v_g if the tube is nonconducting, *i.e.*, during the period when the grid is in control. During the conduction period some grid current flows, and e_g is lower than v_g by the amount of the voltage drop in R_g. Thus R_g serves to limit the grid current to a safe value. Similarly, e_b equals v_p during the nonconducting period, although during the conducting period it is a constant equal to a few volts.

We can determine the control-circuit operation by considering the circuit before breakdown occurs. Since the voltages in the control circuit are sinusoidal during this period, we can draw a vector diagram. We shall first write mesh equations for Fig. 14-27. The voltage $\mathbf{V}_1 + \mathbf{V}_1'$ causes a current $\mathbf{I}$ to flow through the two impedances $\mathbf{Z}_1$ and $\mathbf{Z}_2$. Hence we may write

$$\mathbf{V}_1 + \mathbf{V}_1' = \mathbf{I}(\mathbf{Z}_1 + \mathbf{Z}_2) \qquad (14\text{-}114)$$

and

$$\mathbf{V}_g = \mathbf{I}\mathbf{Z}_1 - \mathbf{V}_1 = \mathbf{V}_1' - \mathbf{I}\mathbf{Z}_2 \qquad (14\text{-}115)$$

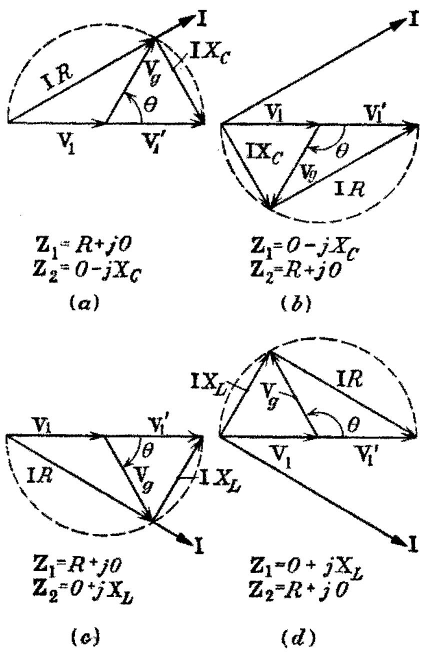


FIG. 14-28. Vector diagrams for the phase-shift circuit.

Figure 14-28 shows vector diagrams for various types of impedances $\mathbf{Z}_1$ and $\mathbf{Z}_2$. It should be recalled that if the grid voltage leads the plate voltage the grid has no control and the tube will conduct for 180°. If the grid voltage lags the plate voltage, the grid does have control and hence a variable direct current can be obtained.

In these diagrams it is assumed that when $\mathbf{Z}_1$ is a pure resistance, $\mathbf{Z}_2$ is a pure reactance; when $\mathbf{Z}_1$ is a reactance, $\mathbf{Z}_2$ is a resistance. If this is true, then $\mathbf{I}\mathbf{Z}_1$ and $\mathbf{I}\mathbf{Z}_2$ must always be 90° apart. Hence, since $\mathbf{V}_1 = \mathbf{V}_1'$, it can be proved that $\mathbf{V}_g$ is constant in magnitude and that its locus is a semicircle. Therefore $\mathbf{V}_g$ must equal $\mathbf{V}_1$.

When the voltage $\mathbf{V}_g$ leads the voltage $\mathbf{V}_1$, the angle of lead can be computed from the circuit constants. Referring to Fig. 14-29 we see that angle OAB is an inscribed angle measured by one-half arc BC, while angle θ is a central angle measured by all of arc BC. Therefore, angle

$OAB = \frac{1}{2}\theta$. We may then write

$$\tan\frac{\theta}{2} = \frac{IX_C}{IR} = \frac{1}{\omega RC} \tag{14-116}$$

or the angle of lead is

$$\theta = 2 \arctan \frac{1}{\omega RC} \tag{14-117}$$

Likewise for a lagging phase angle it can be shown that the angle of lag is

$$\theta = 2 \arctan \omega RC \tag{14-118}$$

From equation (14-118) we see that either R or C can be varied from zero to infinity to change θ from 0 to 180° lagging.

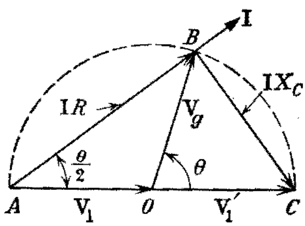


FIG. 14-29. Vector diagram for a phase-shift circuit for V_g leading V_1 by an angle θ.

These phase-shift circuits may be used in various ways. For instance, in the circuit of Fig. 14-27, Z_2 may be a resistor which can be adjusted manually, while Z_1 is a fixed capacitor. Or Z_2 may be an inductor and Z_1 a resistor, either one variable. Sometimes it may be more desirable to connect a circuit as shown in Fig. 14-30. Here the saturable-core reactor acts as a variable inductance. As the current through the d-c coil is varied by changing the grid voltage for the high-vacuum tube, the impedance of the a-c coil changes and hence shifts the phase of the thyratron grid voltage.

14-16. Bias Phase Control. For certain applications it would be very convenient to vary the firing angle of a thyratron over a limited range by means of a variable direct voltage. One method of doing this is known

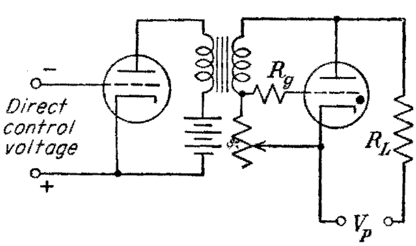


FIG. 14-30. A phase-shift circuit for controlling a thyratron by a direct voltage.

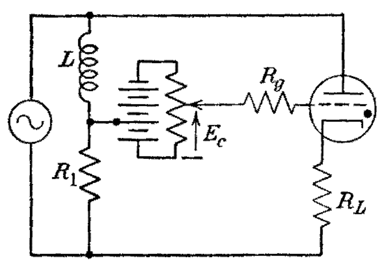


FIG. 14-31. Circuit for bias phase control.

as bias phase control. Figure 14-31 shows this type of circuit. Both alternating and direct voltages are applied to the grid of the thyratron. The alternating voltage is given a fixed phase angle relative to the plate voltage, by means of the inductor and resistor in series across the line. The drop across the resistor lags the source voltage by an angle θ which is determined by the relative magnitudes of R_1 and ωL. This a-c component is additive to the variable direct voltage. Variation of the direct voltage merely has the effect of raising or lowering the alternating voltage relative to the zero-voltage axis, as shown in Fig. 14-32. This in turn results in the grid-voltage curve crossing the critical-voltage curve at different times depending on the magnitude and polarity of the direct voltage.

This system finds limited use since the firing angle can be shifted reliably only over a range which is considerably less than 180°.

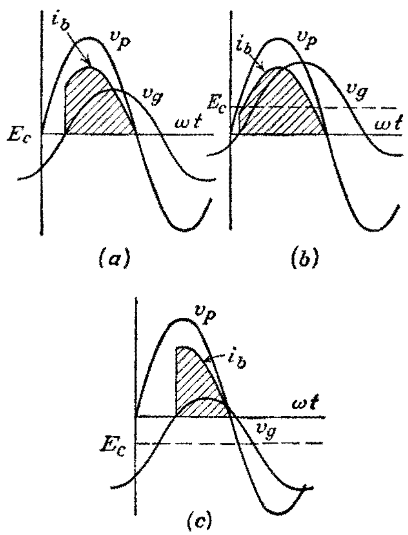


FIG. 14-32. Waveshapes for bias phase control.

14-17. The Ignitron as a Controlled Rectifier. The ignitron can also be used as a controlled rectifier if the igniter current can be supplied as a sudden high-peak pulse. Hence the ordinary phase-shift circuits cannot be used directly.

It is usually customary to use a thyratron to fire the ignitron. This gives the sudden pulse of igniter current. Figure 14-33 shows a circuit diagram of an ignitron which is fired by a thyratron. The thyratron in turn uses the common phase-shift circuit to vary its firing time. Thus the ignitron is controlled indirectly.

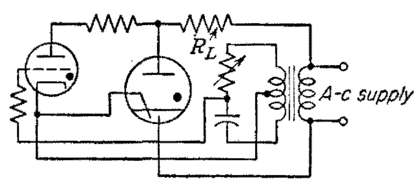


FIG. 14-33. Ignitron circuit using a thyratron for firing.

REFERENCES

1. Cage, J. M.: "Theory and Application of Industrial Electronics," McGraw-Hill Book Company, Inc., New York, 1951.
2. Cruft Laboratory, War Training Staff: "Electronic Circuits and Tubes," McGraw-Hill Book Company, Inc., New York, 1947.
3. E. E. Staff, MIT: "Applied Electronics," John Wiley & Sons, Inc., New York, 1943.
4. Millman, J., and S. Seely: "Electronics," 2d ed., McGraw-Hill Book Company, Inc., New York, 1951.

PROBLEMS AND QUESTIONS

1. Use the characteristic curve shown in Fig. 14-2. (*a*) Calculate the approximate d-c plate resistance for the type 6H6 diode. (*b*) Draw the equivalent circuit for a half-wave rectifier using one diode of a type 6H6 tube, a load resistor of 5000 ohms, and a voltage source $v_p = 200 \sin 377t$. (*c*) Calculate I_{dc}, E_{dc}, I_b, P_{in}, P_{dc}, P_p, and η for this circuit. (*d*) Are any of the tube-manual ratings exceeded?

2. One diode of a type 6X5 tube is to be used in a half-wave rectifier circuit as shown in Fig. 14-34. The d-c meters are of the d'Arsonval type, and the a-c meters are of the iron-vane type. Assume that the milliammeters have zero impedances

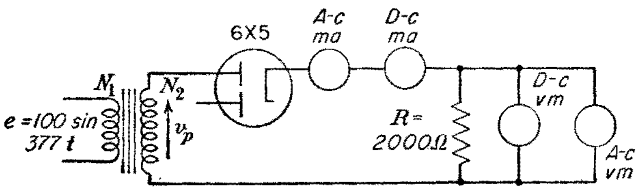


FIG. 14-34.

and that the voltmeters have infinite impedances. The transformer turns ratio is 1:5. Consulting a Sylvania Technical Manual we find that this type of tube has a drop of 22 volts when passing a current of 70 ma per plate. (*a*) What is the reading of the a-c milliammeter? (*b*) What is the reading of the a-c voltmeter? (*c*) What is the reading of the d-c milliammeter? (*d*) What is the reading of the d-c voltmeter? (*e*) How much d-c power is supplied the load? (*f*) How much power must the transformer supply? (*g*) What is the rectification efficiency?

3. A type 5Y3 tube is used in a full-wave rectifier circuit as shown in Fig. 14-35. The Sylvania Technical Manual yields the information that the tube drop is 60 volts when the plate current is 125 ma per plate. Calculate (*a*) I_{dc}, (*b*) E_{dc}, (*c*) I_b, (*d*) P_{in}, (*e*) P_p, (*f*) P_{dc}, (*g*) η, (*h*) $i_{b_{max}}$, the maximum steady-state peak current per plate. (*i*) Are any of the tube ratings exceeded?

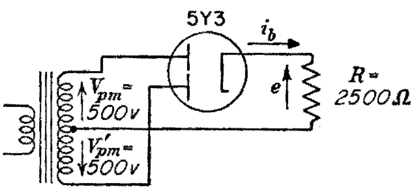


FIG. 14-35.

4. The connection to one of the plates of the 5Y3 used in the previous problem becomes detached. Calculate the quantities called for in that problem.

5. A type 5U4 tube is connected in a rectifier circuit to supply a direct current of 225 ma to a load. The transformer supplying the rectifier is such that $v_p = 500 \sin 377t$. The tube manual gives the information that the tube drop is 58 volts when the plate current is 225 ma per plate. (*a*) Calculate the plate power loss P_p if the tube is to be connected in a half-wave circuit with the two halves of the tube connected in parallel. (*b*) Calculate P_{dc} for part (*a*). (*c*) Calculate the plate power loss if the tube is connected in a full-wave circuit. (*d*) Calculate P_{dc} for part (*c*).

6. The 5U4 tube in Prob. 5 is to be used to supply a load with a resistance of 2000 ohms. The secondary of the transformer is such that $V_{pm} = 500$ volts. (*a*) Calculate the plate power loss in the tube when connected as a half-wave rectifier with both halves of the tube in parallel. (*b*) Calculate the plate power loss when the tube is connected as a full-wave rectifier.

7. The 5U4 tube of Prob. 5 is to be used in a rectifier circuit which must supply 50 watts of direct power to a load for which the resistance is 500 ohms. (*a*) Calculate the plate power loss in the tube when connected as a half-wave rectifier with both halves of the tube in parallel. (*b*) Calculate the plate power loss in the tube when connected as a full-wave rectifier.

8. A type 83 tube is used in the circuit drawn in Fig. 14-7. The tube drop is 15 volts when the tube is conducting. $v_p = 30 \sin 377t$, and $R = 100$ ohms. Find (*a*) the ignition angle θ_1, (*b*) the extinction angle θ_2, (*c*) the direct current I_{dc}, (*d*) the direct voltage E_{dc} across the load, (*e*) the power loss in the tube.

9. Repeat Prob. 8 except let $V_{pm} = 300$ volts.

10. A transformer, a tungar tube (a hot-cathode gas diode built for battery-charging service), and a 0.5-ohm current-limiting resistor are used in conjunction with a 110-volt 60-cps source, to charge a storage battery. If the secondary of the transformer supplies 20 volts rms and if the battery has a terminal voltage of 6 volts, calculate the direct current flowing through the battery. In order to simplify the problem, consider that the terminal voltage of the battery remains constant at 6 volts during the charging period. $E_d = 10$ volts.

11. Prove that the effective value of the alternating components of a saw-tooth wave is $a/\sqrt{3}$, where $2a$ is the peak-to-peak value.

12. A type 6X5 tube is used as a half-wave rectifier with capacitance filtering as shown in Fig. 14-36. Calculate (*a*) γ, (*b*) $i_{b_{\max}}$, (*c*) E_{dc}, and (*d*) I_{dc}. (*e*) Consult the tube manual, and find the maximum allowable peak plate current for the tube. (*f*) Is the tube being used within its ratings? If not, what remedy would you suggest?

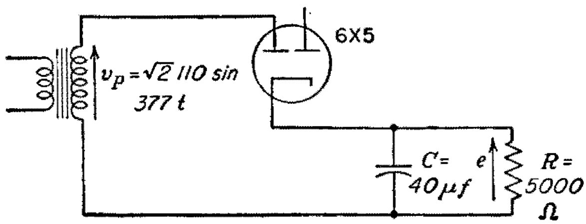


FIG. 14-36.

13. In Fig. 14-37*a*, $v_p = 100 \sin 377t$, $R = 1000$ ohms, $L = 26.5$ henrys. (*a*) Write the expression for i. A diode is inserted in the circuit as shown in Fig. 14-37*b*. (*b*) Write the expression for i_b, neglecting the tube drop. (*c*) Sketch the approximate waveforms of the two parts of the current expression of part (*b*), and add to obtain the waveform of i_b for the interval between $t = 0$ and $t = \frac{1}{60}$ sec. (*d*) Estimate the value of θ_2 from the above waveshape. (*e*) Compute the average value of current I_{dc}, using the appropriate formula. (*f*) Compute the effective value of the ripple current through the load.

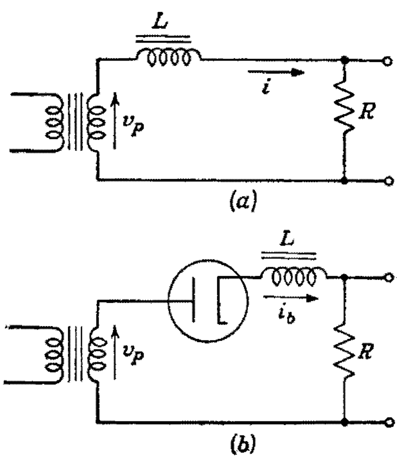


FIG. 14-37.

14. A full-wave rectifier circuit is shown in Fig. 14-38. $V_p = 500$ volts, $f = 60$ cps, and the tube drop is negligible. Compute I_{dc}, I_{f_2}, E_2, I_2, and γ.

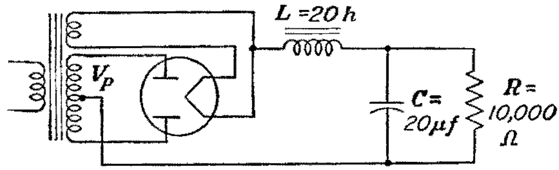


FIG. 14-38.

15. If the filter for the circuit of Prob. 14 is built in two sections with each choke having an inductance of 10 henrys and each capacitor having a capacitance of 10 μf, calculate the ripple factor.

16. If the choke coils used in Prob. 15 each had a resistance of 100 ohms, and if the tube was a type 83, calculate the direct voltage across the load. Assume the tube drop to be 15 volts.

17. A full-wave rectifier with simple inductance filtering supplies a resistance load as shown in Fig. 14-39. Calculate E_{dc} and the ripple factor γ. Assume the tube and

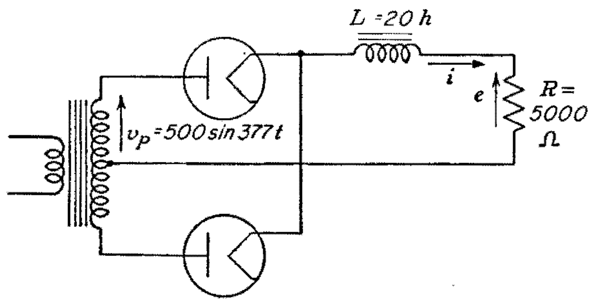


FIG. 14-39.

choke to be ideal and that the ripple consists of the two lowest-frequency components present in the output.

18. A full-wave rectifier with a single L-section filter employs a transformer which delivers a voltage of 400 volts rms from one end to the center tap. The filter capacitor has a reactance of 200 ohms at 60 cps, and the choke has a reactance of 5000 ohms at 60 cps and a resistance of 100 ohms. The circuit uses a type 83 tube. No bleeder is used. (*a*) What is the minimum load current if good voltage regulation is to be maintained? (*b*) What is the direct output voltage when the load current is 100 ma? Do not neglect the tube drop or the drop due to the resistance of the choke. (*c*) If a bleeder was to be connected in the circuit, what should be its maximum resistance?

19. It is desired to convert the L-section filter of Prob. 18 to a Π-section filter. Use a capacitor similar to that employed in the L-section filter. Estimate the output voltage of the filter if $I_{dc} = 10$ ma, by making the following calculations: (*a*) Estimate the change in voltage across the first capacitor by assuming linear discharge of that capacitor. (*b*) Approximate the direct voltage across the first capacitor. (Remember the tube drop!) (*c*) Calculate the approximate output direct voltage by subtracting the direct voltage drop across the choke from the direct voltage across the first capacitor. (Note that, at best, this result is only a reasonable approximation.)

20. A full-wave rectifier circuit shown in Fig. 14-40 has the following components: $C_1 = 10\ \mu f$, $L_1 = 10$ henrys, $C_2 = 10\ \mu f$, $L_2 = 10$ henrys, $C_3 = 10\ \mu f$. Assume the

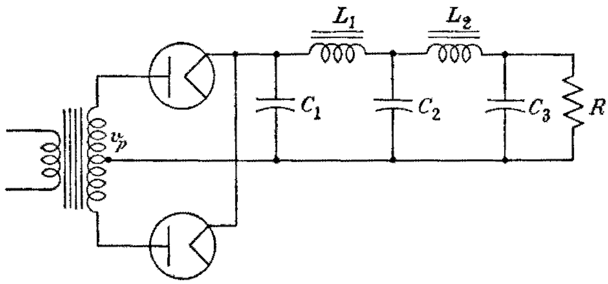


Fig. 14-40.

chokes to have zero resistance. (*a*) Write an expression similar to that of Eq. (14-106) for calculating the *n*th-harmonic output voltage. (*b*) If $I_{dc} = 100$ ma, calculate the ripple voltage E_2.

21. A type FG-17 thyratron at 40°C has a sinusoidal voltage of 2000 volts peak value applied to its anode. Sketch to scale one cycle of a sinusoidal voltage whose peak value is 2000 volts. Then, using information obtained from the firing character-

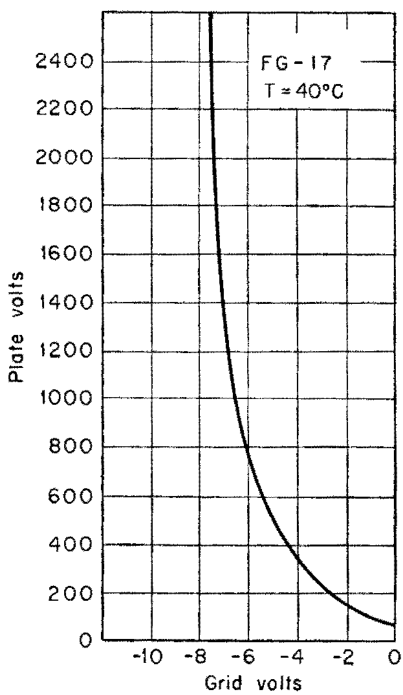


Fig. 14-41.

istic of an FG-17 (Fig. 14-41), sketch the critical grid voltage necessary to cause the tube to fire at any instant during the positive half cycle of the anode voltage.

22. Assume that the critical grid voltage for a certain thyratron is zero for all anode voltages. Find the average value of anode current when $v_p = 220\sqrt{2}\sin 377t$, $v_g = 220\sqrt{2}\sin[377t - (\pi/3)]$, and $R_L = 50$ ohms. The tube drop should be considered as constant at 15 volts.

23. A thyratron is connected in the circuit shown in Fig. 14-42. Calculate the firing angle θ_1 and the average load current. Neglect the tube drop, and assume that the critical grid voltage is zero for all anode voltages.

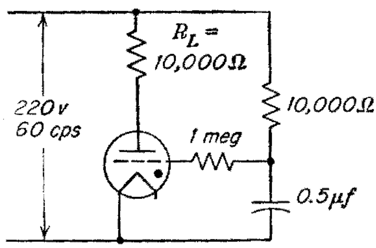


FIG. 14-42.

24. Figure 14-43 shows the circuit diagram for a grid-controlled rectifier using a type FG-17 thyratron whose firing characteristic is shown in Fig. 14-41. (*a*) What is the purpose of the time-delay relay? (*b*) Calculate the peak plate current if v_p is the square wave shown. (*c*) The grid signal voltage is a series of short pulses as shown. Plot the plate-current waveform. (*d*) Calculate the average plate current. (*e*) Is the bias voltage necessary if the grid signal is to control the firing angle? (*f*)

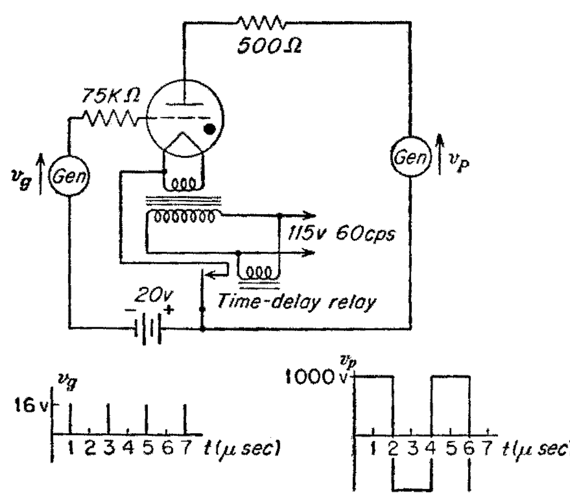


FIG. 14-43.

Suppose that v_g and v_p are now replaced by two sinusoidal voltages $v_g = -1000\cos 377t$ volts and $v_p = 1000\sin 377t$ volts. Estimate the peak grid current. (*g*) Calculate the peak plate current. (*h*) Plot the approximate plate-current waveform. (*i*) Is the bias voltage necessary?

25. The thyratron shown in Fig. 14-44 is to be phase-controlled by means of Z_1 and Z_2. A 10-henry inductor and a 10,000-ohm variable resistor are available. Is Z_1

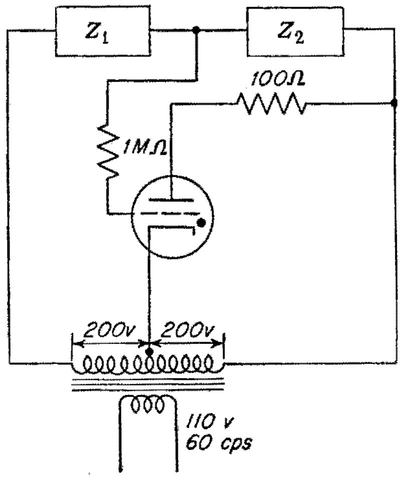


FIG. 14-44.

or is Z_2 the inductor? Justify your choice by means of a vector diagram. (*b*) If the critical grid voltage is assumed to be zero for all anode voltages and the tube drop is zero, plot a curve of average anode current vs. the value of the resistance as the latter is varied from 0 to 10,000 ohms.

26. Repeat Prob. 25, using a 1-μf capacitor in place of the 10-henry inductor.

27. Figure 14-45 is the circuit of a grid-controlled rectifier. Calculate the average value of the load voltage if the tube drop is zero and the critical grid voltage is zero.

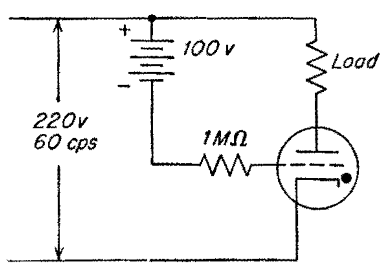


FIG. 14-45.

CHAPTER 15

PHOTOELECTRIC CELLS

15-1. Classification of Photoelectric Cells. The emission of electrons from solid matter under the influence of light was probably the first photoelectric effect observed by early scientists. It was first reported by Hertz in 1887 when he found that a spark gap illuminated by ultraviolet light would break down more readily than one not so illuminated. Elster and Geitel build the first practical photoelectric cells in 1889. They used the cells in a photometer. Since that time the photoelectric effect has received considerable attention from physicists, but it was not until the advent of talking motion pictures that the photoelectric cell came into widespread use.

The purpose of all photoelectric cells is to convert light energy into electrical energy. Hence, we may classify these devices according to the manner in which they accomplish this result.

Photoemissive Cells. This is the most widely used type of photoelectric cell. The sensitive surface (cathode) and the anode are enclosed in a glass or quartz envelope which may be highly evacuated or filled with an inert gas at a very low pressure. It operates on the principle that light falling on the surface of certain solids will cause the emission of electrons. The emitted electrons are then attracted to an anode in the tube and hence cause current to flow through the external circuit. The magnitude of this current is usually less than a few microamperes. The photoemissive cell is the type to be studied in this chapter.

Photovoltaic Cell. This is probably the second most commonly used type of photoelectric cell. It is the light-sensitive element in lightmeters used by photographers to measure light levels in order to calculate exposure times correctly. It operates on the principle that an emf is generated by the action of the incident light on the sensitive surface. The photovoltaic cell is also known as a barrier-layer cell. It will be discussed in more detail in Chap. 16.

Photoconductive Cell. This device is one whose resistance changes with variations in the intensity of the incident light. It consists of a very thin film of selenium or certain metallic oxides between two electrodes. For mechanical strength the electrodes may be coated onto a flat piece of glass as shown in Fig. 15-1. Selenium is then vaporized onto the glass in such a fashion that it bridges the gap between the two electrodes. The action of the photoemissive cell in an electrical circuit is that of a variable resistance whose magnitude depends upon the intensity of the light falling on it. This again is not an electronic tube, and it also will be treated in the next chapter.

15-2. Fundamental Theory of Photoemission.[1] The theory of photoemission is very complicated and by no means completely understood. However, for our purposes we may assume that the only difference between photoelectric emission and thermionic emission is the manner by which electrons near the surface of a solid are given the energy needed to free them from the boundary forces. In thermionic emission this energy comes from the heating process, while in photoelectric emission the energy comes from the light.

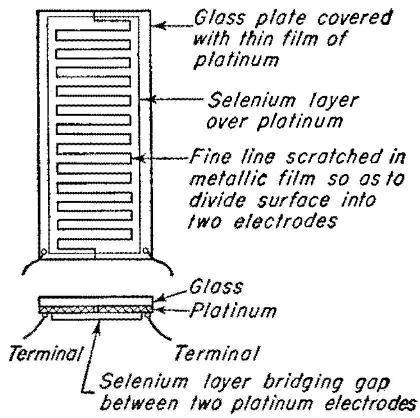


FIG. 15-1. Construction of a photoconductive cell.

The wave theory of light can be used to explain certain aspects of photoemission but runs into difficulties when used to explain others. For example, if we were to study the effects of varying the frequency (color) of the light falling on a photoemissive surface while keeping the intensity (energy) constant, we would reach the following conclusions: As the frequency of the light varies, the emission current changes. This effect can be explained by means of the wave theory of light. However, the velocities with which electrons are emitted is dependent on the frequency of the incident light and decreases with increased frequency. Likewise it can be shown that the velocity of emission is independent of the light intensity. These latter facts concerning velocity cannot be explained by the wave theory but can be explained by the quantum theory of light. Hence light is treated as a wave motion and also as consisting of small particles.

Quantum theory postulates that energy does not flow continuously but rather flows in discrete quantities or small packages. The energy in each package is known as a quantum. The amount of energy in a quantum is directly proportional to the frequency of the light and is given by the formula

$$W_Q = hf \tag{15-1}$$

where W_Q is the quantum of energy, h is Planck's constant and has a value 6.624×10^{-34} watt-sec per cycle, and f is the frequency of the light in cycles per second. Thus we see that the energy per quantum is not the same for different colors of light. For example, a monochromatic blue light with a wavelength of 5000 A (1 A = 10^{-10} m) has a frequency of 6×10^{14} cps, and the corresponding quantum has an energy of 39.75×10^{-20} watt-sec. This blue light can deliver energy only in bundles of this size. The total energy delivered over a given period of time must always be a multiple of the energy per quantum. Likewise a certain monochromatic red has a wavelength of 7500 A, a frequency of 4×10^{14} cps, and a quantum of 26×10^{-20} watt-sec. Thus we see that the energy

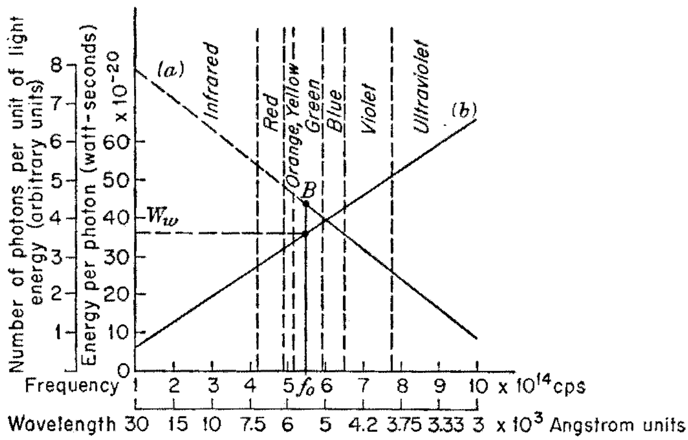


FIG. 15-2. Energy per photon as a function of the frequency of light.

of the quantum of red light is much less than the energy of the quantum of blue light.

A quantum of light energy is also known as a photon. If light is considered to be made up of a procession of photons, it can readily be seen that the number of photons in a given beam will depend on the total energy per second of the light and on its frequency. If the number of photons in a constant-energy beam is plotted as a function of frequency or of wavelength ($\lambda = c/f$, where λ is the wavelength of the light, c is the speed of light, and f is the frequency), the straight line a shown in Fig. 15-2 will result. The energy per photon for this same constant-intensity beam is shown by curve b.

If the work function W_w of a given material is expressed in watt-seconds, it can be indicated on the graph as shown in Fig. 15-2. Thus it would seem that light falling on the surface will cause no emission unless the frequency of the light is greater than 5.5×10^{14} cps. At this frequency, which is that of a green light, the energy per photon is just sufficient to cause the emission of electrons. Above the frequency for green light the energy per photon is more than sufficient to cause emission, and since the energy of the beam is constant, the number of photons is less. If we assume that each photon causes the emission of a single electron, the emission current will decrease with increasing frequency as shown by the solid portion of the curve *a* in Fig. 15-2. The lowest-frequency light which can cause photoemission from a given material is known as the threshold frequency for that material. It is directly related to the work function of the material.

While theoretically the photoemissive cell should have the response shown in Fig. 15-2, the actual response for several of the low-work-function elements is shown in Fig. 15-3. Note that these curves are con-

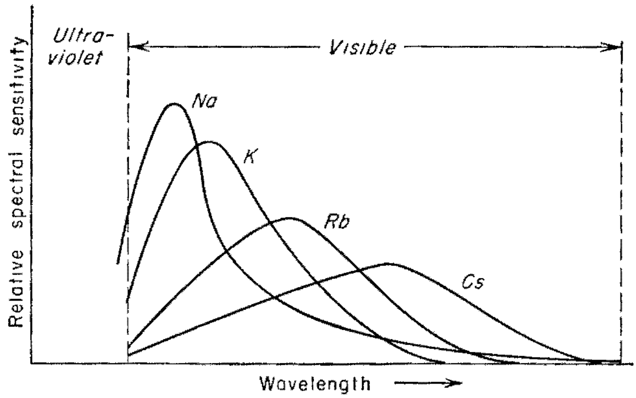


FIG. 15-3. Spectral-sensitivity curves for several pure metals.

siderably different from the predicted ones. There are several reasons for these variations. One reason is that our theory was based on absolutely pure material. The smallest amount of impurity may cause a considerable change in response, and since it is almost impossible to obtain pure materials, there is reason to expect considerable variation between actual and theoretical curves. A second reason is that two or more photons may give energy to a single electron, and hence emission starts at frequencies below the threshold value. A third reason, and a very important one, is that as the frequency of the incident light increases, so does the energy per quantum. This results in the freeing of lower-energy electrons.

Present-day photocells seldom utilize the simple cathodes used for the curves of Fig. 15-3; rather, they use much more complex surfaces which give them greater sensitivity and more desirable color response. Such a cell is the cesium–cesium oxide–silver cell. A typical response curve is shown in Fig. 15-4. Note that the cell has two peaks in its color-response curve, which seems to indicate that there are two threshold frequencies. This is quite possible because of the complex nature of the surface.

While this type surface is probably the most common one used today, there are other types of cells which give quite different frequency-response curves. By proper choice of emitting material and by certain activation procedures the maximum response can be located in the blue end, in the red end, or in the middle of the spectrum. Thus in selecting photoelectric cells, the response curve of the cells should be studied in order to select one which will perform satisfactorily with the color of light to be used.

15-3. The Vacuum Photoemissive Cell. Vacuum photocells or phototubes may vary considerably in physical size and appearance, but essentially they consist of a light-sensitive cathode and an anode to collect the

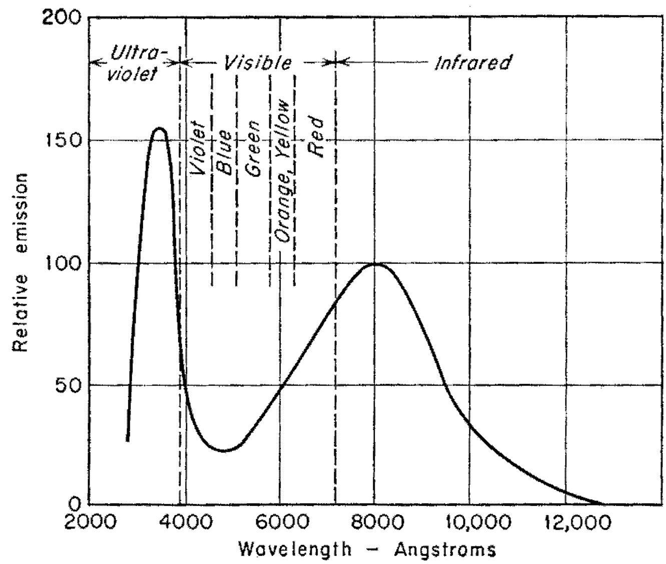


FIG. 15-4. Spectral-sensitivity curve for a typical cesium–cesium oxide–silver cell.

emitted electrons. Figure 15-5 shows a picture of several cells. In most photocells the cathode is a semicylindrical surface so arranged that light falls on its concave face. The anode is a straight wire running along the axis of the cylinder.

When light of an appropriate color falls on the cathode, electrons are emitted. If a small voltage is now applied to the anode, a small current flows. As the voltage is increased, the current increases quite rapidly. This is because, for small anode voltages, the current is space-charge-limited. However, on reaching a certain anode voltage it is found that the current levels off and remains approximately constant thereafter. Vacuum photocells are usually operated in this saturated region. Figure 15-6 shows a family of curves for the type 922 cell. Note that light flux in lumens is the parameter for this family. These curves correspond to the plate family for a triode. An examination reveals that the dynamic

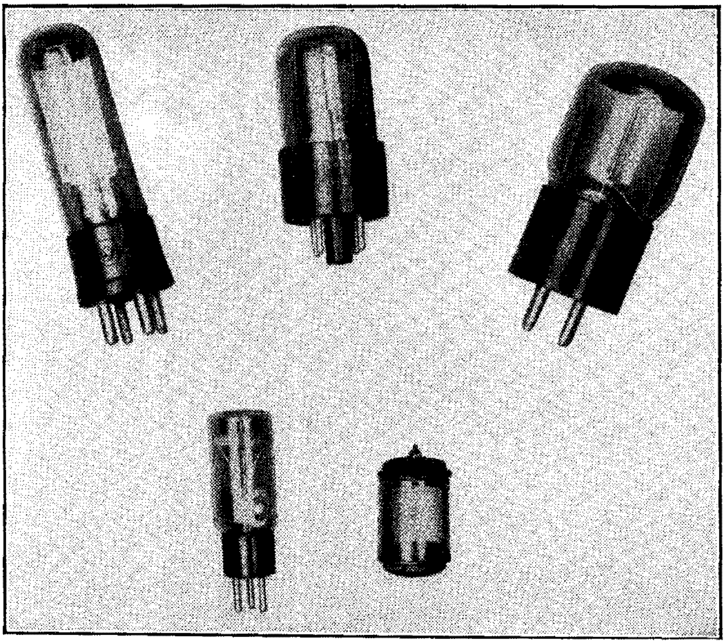


FIG. 15-5. Photoemissive cells.

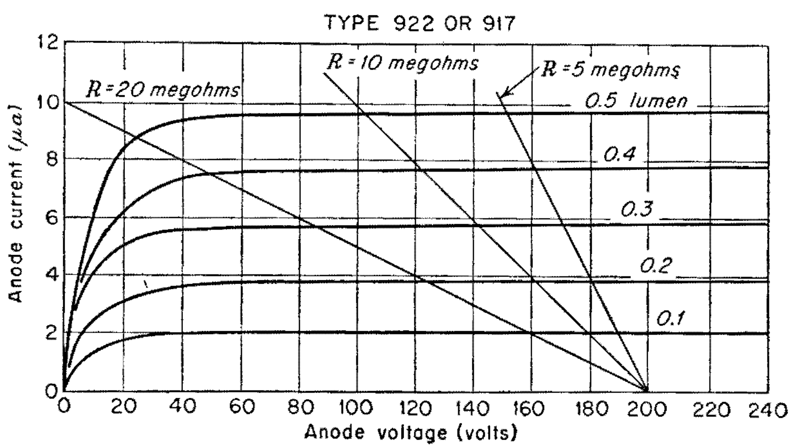


FIG. 15-6. Volt-ampere characteristic for a vacuum-type photocell.

plate resistance for such a phototube is very high, being of the order of hundreds of megohms. The flatness of the curves also indicates that the plate current is almost independent of the anode voltage as long as operation is above saturation. The vacuum cell may hence be considered as a constant-current generator whose output current is a function of the light intensity.

Data were obtained from the curves of Fig. 15-6 to plot Fig. 15-7*a*. Here the light flux is taken as the independent variable while the anode voltage was held constant at 90 volts. Note that the current output increases linearly with increasing light intensity. It is practically independent of the anode voltage as long as this remains above the saturating value.

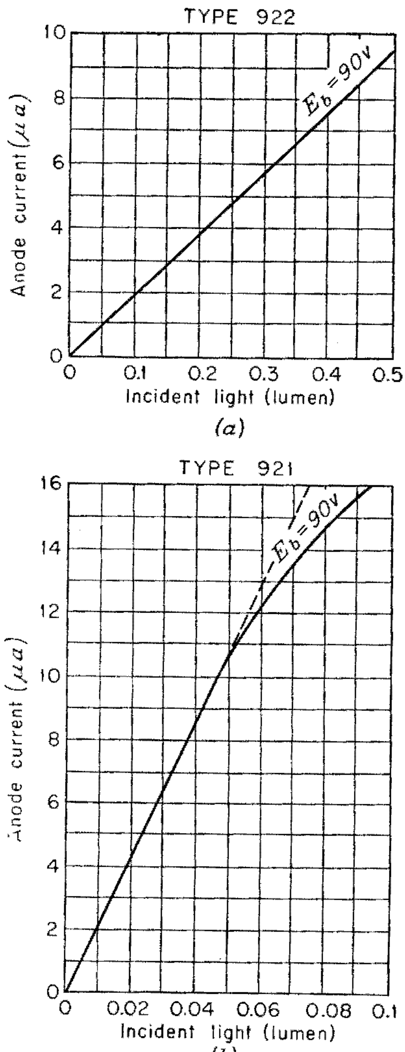


FIG. 15-7. Variations of emission current with light flux for (*a*) a vacuum-type cell, (*b*) a gas cell.

15-4. The Gas-filled Photoemissive Cell. By introducing the proper amount of inert gas, a photoemissive cell can be made to have quite different characteristics from those of the vacuum type. Figure 15-8 shows a family of plate characteristics for such a tube. When a constant light flux falls on the cathode, the anode current increases with anode voltage until it reaches saturation just as it did in the high-vacuum type. However, soon after reaching saturation the anode current again begins to increase with increased voltage. This is now a Townsend-type discharge. The increased current is caused by the ionization of gas molecules and by the bombardment of the cathode by positive ions. The latter process causes the freeing of electrons by secondary emission, which in turn causes an increase in anode current. In this manner it is possible to increase the anode current obtained from a photoelectric surface with a given light flux by a factor as great as 10. This process is also known as gas amplification. Efforts to obtain a gas-amplification factor of greater than 10 usually result in a glow discharge instead of a Townsend discharge, with resultant damage to the cathode.

Examination of the curves of Fig. 15-8 shows that the dynamic plate resistance of a gas-filled phototube is much less than that for a high-vacuum tube. Also an examination of Fig. 15-7*b* shows that the relationship between light flux and anode current is no longer linear. This is somewhat of a disadvantage but is not too serious since it remains approximately linear for small light variations.

The most serious disadvantage of the gas-filled phototube is its loss of sensitivity for h-f light variations. This loss is caused by the time lag

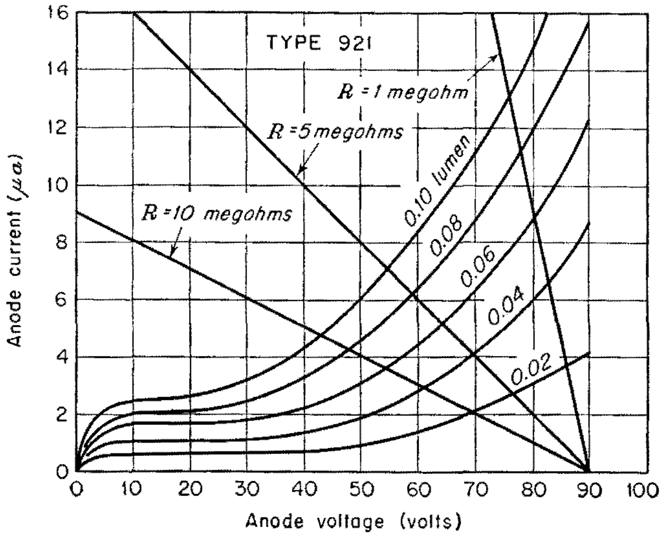


FIG. 15-8. Volt-ampere characteristic for a gas-type photocell.

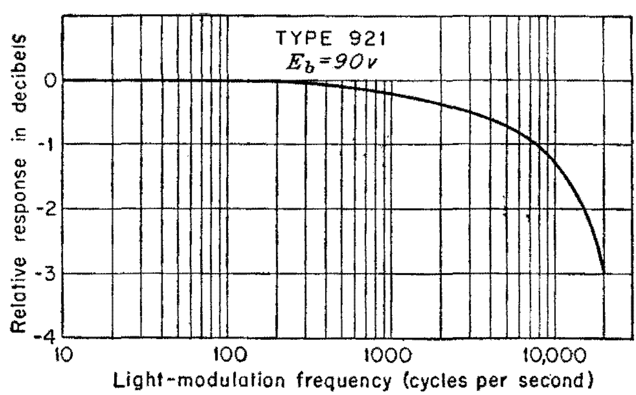


FIG. 15-9. Response of a gas photocell as a function of the light-modulation frequency.

between a change in light intensity and a change in current. This in turn is caused by the slowness of the positive ions in reaching the cathode to produce free electrons by secondary emission. The output of a gas phototube may begin to decrease for light-modulation frequencies above a few hundred cycles per second. Figure 15-9 shows the frequency response of a typical gas-filled phototube.

15-5. Sensitivity of Photoemissive Cells. The sensitivity of a photoemissive cell can be expressed in two ways. One is the spectral sensitivity, which is defined as the ratio of the total emission current to the total radiant power falling on the surface. This is usually stated in terms of microamperes per microwatt of incident light of a specified wavelength. The wavelength of the light will usually be such that it is somewhere near the peak response for the cell.

A second way to express the sensitivity of a cell is by the luminous sensitivity. Most photocells are used with incandescent lights. Hence it is customary to use a tungsten-filament lamp operating at a color temperature of 2870°K as a standard. The luminous sensitivity is then given in microamperes per lumen of the incident light. There are two luminous sensitivities. One is the static luminous sensitivity S_L, which is defined as the ratio of the direct anode current to the incident radiant flux L of constant value. In a similar manner we define dynamic luminous sensitivity s_L as the ratio of the varying anode current to the varying component ΔL of the incident light flux.

For vacuum cells the two luminous sensitivities are equal since the response of such a tube varies linearly with light intensity. For a gas-filled cell the dynamic luminous sensitivity is different from that of the static value. It may be found by calculating the slope of curve in Fig. 15-7*b* at a given point, while the static value is the slope of the line joining the point to the origin. An examination of Fig. 15-9 shows that the dynamic luminous sensitivity will fall off as the frequency of the light variation is increased.

15-6. Calculation of the Output of Photocells. The output for a high-vacuum phototube can be calculated by means of an equivalent circuit just as it was for the triode. However, since the phototube acts more like a constant-current generator, the parallel form of the equivalent circuit is customarily used. Because of the high internal impedance of the vacuum phototube, the equivalent circuit becomes merely a constant-current generator supplying a load resistor.

The gas-filled phototube, not being linear and having a varying dynamic luminous sensitivity, is usually treated graphically, as also can be the high-vacuum type. Figures 15-6 and 15-8 show load lines drawn on the plate characteristics for the vacuum and gas-filled phototubes. The treatment is identical to that for grid-controlled vacuum tubes, the only difference being that the light flux L is varied instead of the grid voltage.

15-7. Spectral Response of Commercial Photoemissive Surfaces. It is sometimes desirable to have phototubes whose peak response occurs in different parts of the spectrum. At times it is desirable to have a response similar to that of the human eye, and at others it may be desirable to have maximum response to red light or to blue light.

In order to simplify the specifications for different photoemissive surfaces, the Radio Manufacturers Association (RMA) has set standards for various spectral-response curves. Figure 15-10 shows the responses of four such surfaces. Note that the S-1 and S-2 surfaces respond very

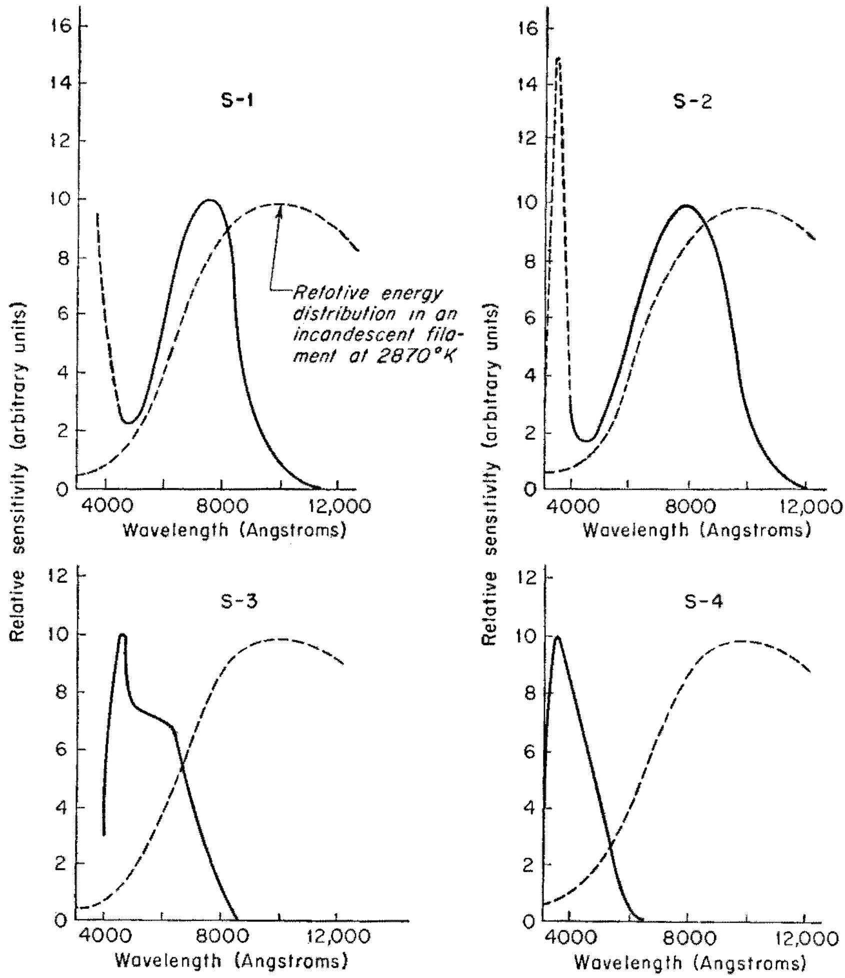


FIG. 15-10. Standard RMA spectral-response curves.

nicely to incandescent light. Surface S-3 has a peak response in the blue, while S-4 has its peak in the near ultraviolet.

In applications it is very important to select a cell which has high response in the spectral region in which the light source has high output. As an example, for use with an incandescent light, the S-4 surface would have very little response, while the S-1 would give good response.

15-8. Applications of Phototubes. The phototube, either the high-vacuum or the gas-filled type, has high internal impedance and supplies a very small output current. Hence it cannot be used to supply directly a low impedance load such as a meter or a relay. However, by the use of a high-resistance load a relatively high output voltage can be obtained which can be applied to the grid of an amplifier tube and the amplifier-plate output used to operate the meter or relay.

A typical circuit for a phototube used to close a relay is shown in Fig. 15-11. When the light falling on the cell reaches a certain intensity, the relay will close. The resistor R_g should be very high in value in order to obtain a high voltage output from the cell. However, it should not be too large, or stray electrons striking the grid will take so long to leak off

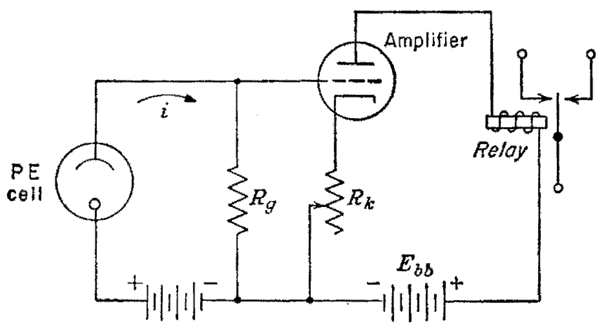


FIG. 15-11. A photocell and amplifier used to actuate a relay.

that the grid may charge up negatively enough to cause the tube to block. The tube may be a medium-μ triode, which will allow an R_g of 10 to 15 megohms. The operating point for the tube is so adjusted that the relay will remain open until the right amount of light falls on the phototube's cathode.

Phototubes may be used in conjunction with small thyratrons to actuate relays or other electrical equipment. Figure 15-12 is the circuit diagram for such an application. Note that the thyratron is the shield-grid type because of its low grid current. Alternating voltage is used instead of batteries as in Fig. 15-11. This allows the phototube to regain control of the thyratron when the incident light falls below a predetermined value.

15-9. Photomultiplier Tubes.[2] When phototubes are used at low light levels, the output of the tube is a very small current and may be a fraction of a microampere. If the tube is to be used to actuate a relay, a current amplifier must be used. A single stage may not be sufficient; so two or more stages may be necessary. If the light is slowly varying, the amplifier must be of the direct-coupled variety. This may become quite complicated since direct-coupled amplifiers are used only when they cannot be avoided. The photomultiplier tube was developed to eliminate the necessity of using direct-coupled amplifiers with phototubes.

Figure 15-13 shows a cross-sectional plan drawing of a photomultiplier tube. The cathode k emits electrons when light falls on it. These electrons are accelerated toward the electrode A_1 by an electric field. When they strike A_1, they cause several secondary electrons to be freed. These in turn are accelerated toward A_2, where each knocks out several more

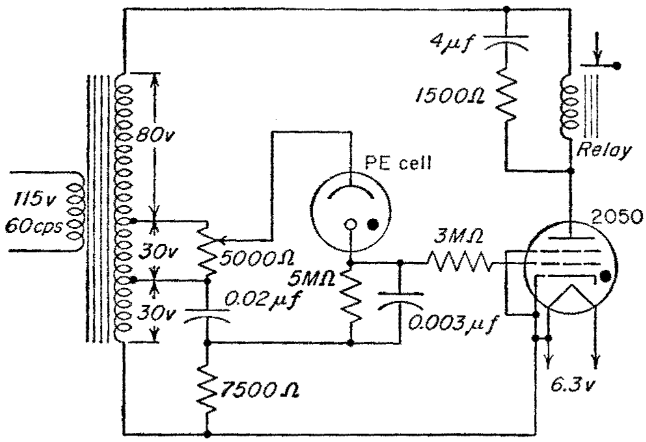


FIG. 15-12. Circuit diagram of the Westinghouse RQ Phototroller.

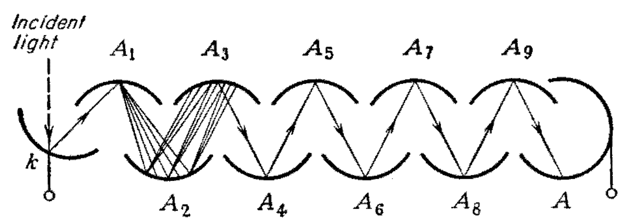


FIG. 15-13. Cross-section plan drawing of a photomultiplier tube.

electrons. Hence we can see that such a tube with nine dynodes (A_1, A_2, etc., are called dynodes), each having a secondary-emission ratio of about 5, will result in about 5^9 electrons arriving at the collector for each photoelectron. Thus such a tube gives a current amplification of approximately 2,000,000.

The smallness of the input to the photomultiplier is limited only by the noise or random fluctuations of the output current. This random fluctuation results because the electrons do not strike the dynodes and collector at a uniform rate but act rather erratically and cause minute fluctuations in the d-c output.

REFERENCES

1. Glover, A. M.: A Review of the Development of Sensitive Phototubes, *Proc. IRE*, **29**, 413 (August, 1941).
2. Zworykin, V. K., G. A. Morton, and L. Malter: The Secondary-emission Multiplier, *Proc. IRE*, **24**, 351 (1936).
3. Metcalf, G. F.: Operating Characteristics of Photoelectric Tubes, *Proc. IRE*, **17**, 2064 (1929).
4. Zworykin, V. K., and E. D. Wilson: "Photocells and Their Applications," John Wiley & Sons, Inc., New York, 1932.

PROBLEMS AND QUESTIONS

1. Name the three photoelectric effects.

2. What is a photon?

3. Calculate the energy per quantum for a red light having a wavelength of 7000 A.

4. What maximum work function can a surface have in order that a red light, whose wavelength is 7000 A, can cause photoelectric emission?

5. What is the luminous sensitivity of the 922 photocell, the characteristics for which are shown in Fig. 15-6?

6. The data for the curves of Fig. 15-6 were taken when light from an incandescent filament fell on the cell. Would the luminous sensitivity for daylight be the same as the value found in Prob. 5?

7. Calculate the static luminous sensitivity for the 921 cell for a light flux of 0.05 lumen and an anode voltage of 70 volts (see Fig. 15-8). Does the value differ from that for the dynamic luminous sensitivity for the same light flux?

8. A type 922 photocell is connected in series with a battery of 200 volts and a load resistor of 10 megohms. Find the change in anode current when the incident light is changed from 0.3 to 0.2 lumen. Repeat for a change from 0.3 to 0.27 lumen.

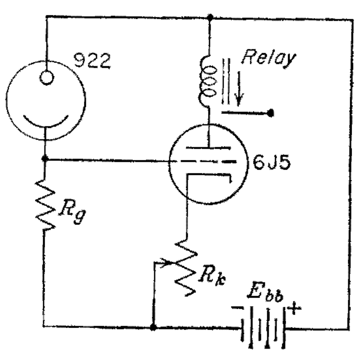


FIG. 15-14.

9. A type 921 gas photocell is connected in series with a 90-volt battery and a 10 megohm resistor. Find the change in anode current caused by a change in the incident light from 0.1 to 0.06 lumen.

10. The photocell in Fig. 15-11 is of the vacuum type and has a luminous sensitivity of 20 μa per lumen. The tube is a 6J5, $E_{bb} = 160$ volts, and R_k is adjusted so that $E_{cc} = -8$ volts. $R_g = 5$ megohms. The relay has a resistance of 8000 ohms and closes when the current through it is 4 ma. How much light must fall on the cell in order to energize the relay?

11. The circuit of Fig. 15-14 is to be used to control the illumination in a room. The photocell is a type 922 and the amplifier tube a type 6J5. $E_{bb} = 150$ volts, $R_g = 0.47$ megohms. The biasing resistor is adjusted so that the relay opens when the light falling on the cell is 0.1 lumen. The relay has a resistance of 5000 ohms and closes when the current through it is 5.0 ma. It opens when the current through it is 3.5 ma. (*a*) Find the value to which R_k must be adjusted. (*b*) How many lumens must fall on the cell to close the relay? (*c*) If the area of the window in the cell is 0.313 in.^{2}, calculate the variation in light intensity required to open and close the relay.

CHAPTER 16

SOLID-STATE ELECTRONICS AND REACTANCE AMPLIFIERS

16-1. Introduction. In recent years a number of devices have been developed to replace or to be used in conjunction with electronic tubes. These devices, while they do not involve the principles of conduction through gases or *in vacuo*, do involve theory similar to that for emission, have nonlinear characteristics, or exhibit other properties like those of the conventional electronic circuits. Some of them are classified under the heading of solid-state electronics. Among the devices to be discussed in this chapter are thermistors, blocking-layer rectifiers, and transistors. Magnetic amplifiers and dielectric amplifiers, while not usually classified under solid-state electronics, will also be discussed.

Solid-state electronics is not a new branch of the electronic art. It dates back to the early days of radio and even precedes the vacuum tube. The early crystal detector used to detect "wireless" waves was a form of a semiconductor which today we call a point-contact rectifier. It employed a fine wire, called a "cat whisker," in contact with a lump of galena (lead sulfide), silicon, silicon carbide, or iron pyrites. Rectification took place at the point of contact, and since there was a wide range of sensitivities at various points on the crystal, considerable time was usually spent hunting the most sensitive spot. For this reason it was soon superseded by the vacuum tube.

An early reference to another type of semiconductor was made by Fritt in 1883 when he observed the rectifying properties of a selenium-metal contact. In 1926 L. O. Grondahl[1] discovered the rectifying properties of copper–copper-oxide contact. These two devices were used for the rectification of currents greater than a few amperes and today are known as dry rectifiers.

16-2. Energy States in Solid Matter. In Arts. 2-2 to 2-5 we briefly discussed the structure of the atom and how some of the electrons became free. We also mentioned that if the energies of the free electrons could be measured at 0°K, they would range all the way from a value somewhat greater than zero up to a certain maximum value W_m. The distribution of the energy among the electrons has been studied quite extensively. Several distribution functions have been proposed, but the one which seems best to meet the physical requirements was developed independently by E. Fermi and P. A. M. Dirac and is known as the Fermi-Dirac distribution.[2] It is expressed in the form of an equation by

$$\rho_w = \frac{KW^{1/2}}{1 + \epsilon^{(W - W_m)/kT}}$$

where ρ_w is the number of electrons per unit volume per unit range of energy, k is a constant of proportionality equal to 6.82×10^{27} electrons per m^3 per electron-volt$^{3/2}$, W is the energy in electron-volts, K is Boltzmann's constant, T is the temperature in degrees Kelvin, and ϵ is the base of natural logarithms.

When the Fermi-Dirac distribution is plotted for 0°K, the solid curve of Fig. 16-1 results. Note that the curve indicates that there are no electrons with zero velocity and that there are no electrons with energy greater than W_m. It should also be noted that absolute zero does not denote the absence of all energy but rather indicates the condition of lowest energy content.

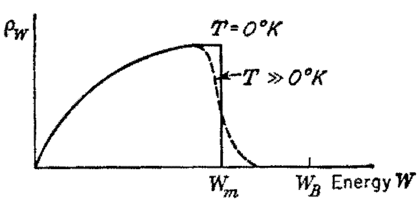


FIG. 16-1. Density of energy states in a solid at 0°K and at T°K.

As the temperature of the material is increased above 0°K, the higher-energy electrons are given additional energy and the distribution function takes the form shown by the broken-line curve in Fig. 16-1. Now there are electrons with energies greater than W_m.

The value of W_m is a characteristic of the material and is referred to as the Fermi level. Since it can easily be calculated for different materials, it is often used as a reference level in comparing energies. W_B is the energy that an electron must have in order to free itself from the forces at the surface of the material and thus become an emitted electron. The difference $W_B - W_m = W_w$ is the work function for that material.

In a conductor the energy distribution at 0°K is shown in Fig. 16-2a, where there is a band of empty allowed energy levels above W_m. If the material is heated, the electrons acquire additional energy and enter this allowed band. Or if the electrons are subjected to the action of an electric field, they likewise acquire the needed energy to enter the allowed band and become conduction electrons.

The energy-distribution curves for semiconductors and insulators are shown in Figs. 16-2b and c, respectively. In these two types of materials, it should be noted that at 0°K the substances have one or more low-energy bands, in which all available levels are filled, separated from a normally empty higher allowable-energy band by a third band of width W_f. Evidence seems to indicate that the latter is never occupied by electrons, and it is hence called a forbidden band. The difference between semiconductors and insulators is in the magnitude of W_f. For semiconductors W_f may be only a fraction of an electron-volt, while for insulators it may be several electron-volts.

In a semiconductor at 0°K there are no electrons in the higher-energy or unfilled band. Hence the application of a small electric field will not result in current flow. However, if the temperature of the material is gradually increased, nothing happens until the temperature reaches the point where some higher-energy electrons acquire an additional energy,

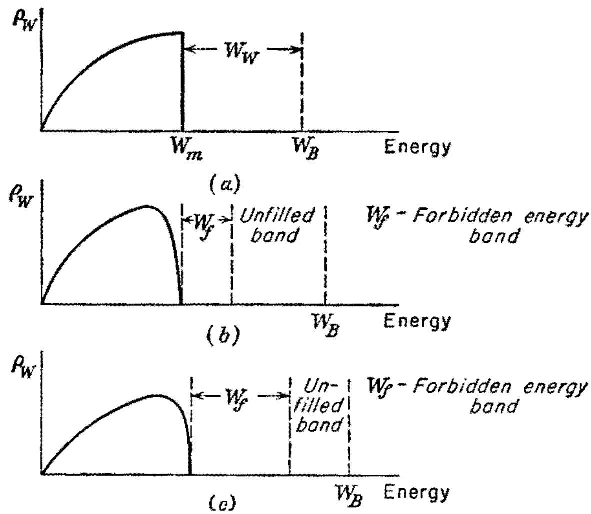


FIG. 16-2. Density of energy states in (a) a good conductor, (b) a semiconductor, (c) an insulator.

greater than W_f, which enables them to enter the unfilled band, where they become conduction electrons.

As the temperature is raised still more, the number of electrons available for conduction rapidly increases and hence more current may flow. Thus we see that, since conductivity increases with temperature, the temperature coefficient of resistivity must be negative. This is one of the characteristics of all semiconductors.

If the temperature of an insulator is increased sufficiently, it too acts like a semiconductor and becomes somewhat conducting. However, the temperature at which this action occurs is much higher than that required to cause conduction in a semiconductor.

16-3. *n*-type and *p*-type Semiconductors.[11] The conductivity of a semiconductor may be greatly affected by the presence of impurities. For instance, 1 part per 10,000 of certain impurities may increase the conductivity of silicon by a factor of 10^7. Other impurities in the ratio of 1 part per 100 may have practically no effect.

Some impurities have filled energy states which are slightly below the unfilled band of the semiconductor and hence fall in the forbidden band. These electrons are not actually in the forbidden band of the semiconductor but instead are in a filled band of the impurity. This band is at such an energy level that it falls between the filled and unfilled bands of the semiconductor. This condition is shown in Fig. 16-3*a*. Here it can be seen that an amount of energy equal to W_d can raise some of the

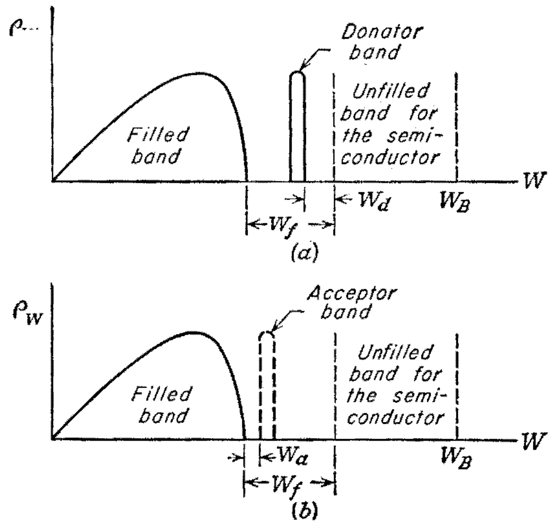


FIG. 16-3. Energy states in a semiconductor showing (*a*) donator level, (*b*) acceptor level.

electrons from the impurity up to the allowed band of the semiconductor. These electrons contribute to the current flow. The remainder of the impurity atom is a positive ion which is bound to a particular location in the solid material. Such an impurity is known as a donator impurity. The positive ion of the donator becomes neutral again when it absorbs another free electron. Hence the atom has a net negative charge. Because of this excess negative charge a semiconductor with such an impurity is known as an *n*-type semiconductor.

It has been found that some impurities have empty allowed bands between the filled and the unfilled bands of a semiconductor, as shown in Fig. 16-3*b*. Such unfilled bands may accept electrons from the filled bands of the semiconductor. Such impurities are called acceptor impurities. If the semiconductor is given sufficient thermal energy for an electron to be raised from its filled band to the acceptor unfilled band, the semiconductor is left with a "hole" in its normally filled band. This hole can be filled by an electric field causing the passage of an electron from a neighboring semiconductor atom. This then leaves the neighbor with a hole which is filled from another neighbor, etc. An atom with such a hole has a net positive charge. Hence we refer to the hole as having a positive charge. Such a semiconductor appears to be conducting by the motion of the positive charge, or hole; hence we call it a *p*-type, or hole, conductor.

Whether a semiconductor is of the *p*-type or the *n*-type is governed by the impurity used. The two most common types of semiconductors are silicon and germanium. These are elements in the fourth group of the periodic table. It has been found that the addition of an impurity from the third group, such as boron or aluminum, will give *p*-type conductivity. Addition of an impurity from the fifth group, such as phosphorus or antimony, gives *n*-type conductivity.

16-4. Thermistors.[3] Thermistors, or thermal-sensitive resistors, are semiconductors made in such a way that their negative coefficient of resistance is large throughout a desired range of temperatures. They find many uses in present-day electrical-engineering practice. These uses may include time-delay devices, protective apparatus, power-measuring devices, and detectors of small radiant power. In many applications thermistors are used because they are simple, small, and rugged, have a long life, and require little or no maintenance.

Any semiconductor may be used as a thermistor. However, in order to obtain a high negative temperature coefficient of resistance in a particular temperature range, they are usually made of a mixture of various metallic oxides. The oxides, which have first been extruded or pressed into a bead, rod, or disk, are heated at such a temperature that they sinter into a strong, compact mass. Contact wires are then fused or soldered on.

Thermistors may change resistance as much as 4.6 per cent per degree centigrade at room temperature. One particular rod-type thermistor has a temperature coefficient of resistance of -0.043 ohm per °C at 25°C. From 0 to 300°C the resistance changes from 145,000 to 85 ohms. The resistance in this case decreases by a factor of more than 1500. For comparison, the resistance of a typical metal such as platinum will be increased by a factor of only 2 over this same temperature range. Figure 16-4 shows how the resistance of a typical thermistor material varies with temperature.

If current is passed through a thermistor, there will be an I^2R loss in the device which will result in an increase in its temperature. This rise in temperature will in turn result in a decrease in the resistance of the thermistor and hence a decrease in the voltage drop across it. Figure 16-5 shows the volt-ampere characteristic for a typical small thermistor bead. As the current first increases from zero, the drop across the bead likewise increases. This is because the current through the bead is not causing sufficient power loss to heat it appreciably. As a result, the voltage increases almost linearly with the current. The temperature

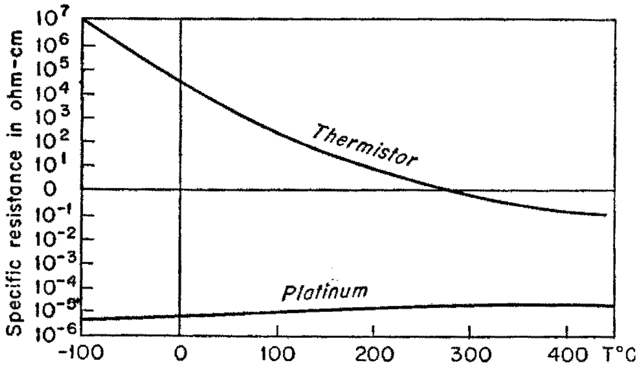


FIG. 16-4. Resistance-temperature curve of a typical thermistor and of platinum. (*Courtesy of Bell Laboratories.*[3])

of the bead is indicated by the numbers on the curve. A point is finally reached where the power loss in the bead is sufficient to cause considerable heating of it and the voltage commences to decrease with temperature.

Variable time delay for a small relay might be achieved by using a thermistor and an adjustable voltage source to actuate a relay as shown in Fig. 16-6. If the relay is so adjusted that it will close with a 5-ma current, it will be actuated in 0.25 sec if the applied voltage is 80 volts, in 1.0 sec if the voltage is 50 volts, and in 5 sec with a voltage of 30. If other delay characteristics are desired, one may select another thermistor from the many types available.

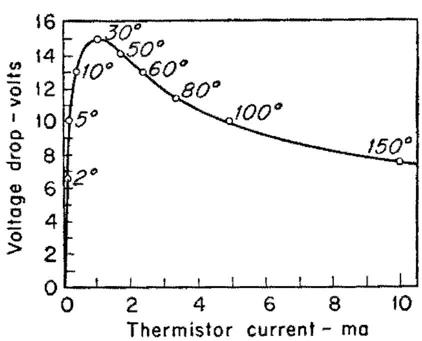


FIG. 16-5. Static volt-ampere curve for a typical thermistor. (*Courtesy of Bell Laboratories.*[3])

Another application for thermistors is shown in Fig. 16-7. This is an amplifier with negative feedback so arranged that the circuit provides constant output with negligible distortion for a wide range of signal input. The over-all gain of the circuit is regulated by the thermistor, which varies in resistance with output in such a way that the amount of feedback is varied to compensate for changes in input-signal level.

16-5. Point-contact Rectifiers.[4] When vacuum diodes are used in very high radio-frequency circuits, the interelectrode capacitances may cause considerable difficulty. Point-contact rectifiers, on the other hand, have very little capacitance between terminals; they can pass reasonable magnitudes of current and require no filament-power source. For these reasons the point-contact rectifier has replaced the vacuum tube in some electronic circuits.

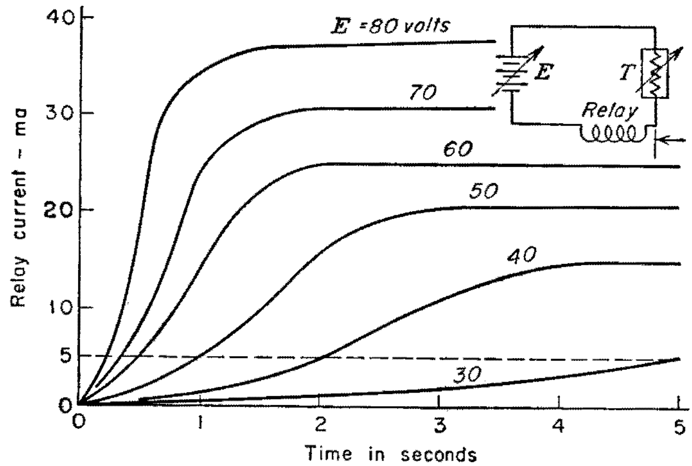


FIG. 16-6. Current vs. time curves for six values of the battery voltage in the circuit shown in the insert. (*Courtesy of Bell Loboratories.*[3])

The materials used for most point-contact rectifiers are either silicon or germanium contacted by a sharply pointed tungsten wire. The whole device is enclosed in a ceramic or glass shell as shown in Fig. 16-8. These devices are more commonly known as germanium or silicon crystals, crystal diodes, or crystal detectors.

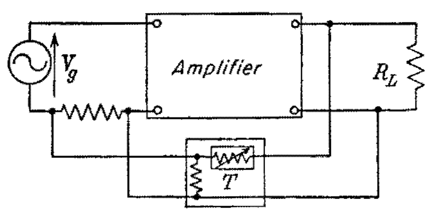


FIG. 16-7. Circuit showing the use of a thermistor in a feedback loop to provide automatic gain control.

In the manufacture of germanium crystals, germanium dioxide is first reduced by heating to an amorphous gray powdery form of the pure metal. This powder is then melted, along with the proper amount of impurity, to form the desired type of semiconductor. If a p-type is being made, the impurity might be aluminum, boron, or any other element from group III of the periodic table. If an n-type is being made, the impurity is any element from group V of the periodic table, such as arsenic or phosphorus. When the semiconductor has cooled, it is sliced into wafers about 2 or 3 mm square and 0.5 mm thick. These wafers are then ground and polished on one side to form the active element. The unpolished side is then soldered to a large-area base electrode. Contact is made to the polished side by means of the sharply pointed wire.

The complete theory of operation for point-contact rectifiers is too complex to give in its entirety. However, an approximate, and necessarily incomplete, explanation will help in understanding the characteristics of these useful devices.

When a metal and a semiconductor are brought into contact, a contact difference in potential will result, as explained in Art. 2-8. More electrons flow from the material with a low work function to the material with the high work function than flow in the reverse direction. This results in the high-work-function material attaining a negative potential relative to the low-work-function material.

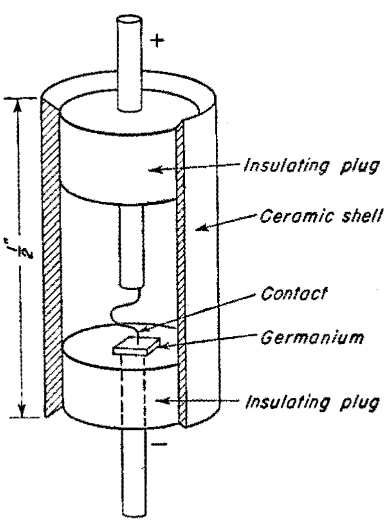


Fig. 16-8. Construction of a germanium diode.

In the case of an n-type semiconductor contacted by a tungsten wire, the tungsten has the higher work function and becomes negative relative to the crystal. However, unlike a metal, the semiconductor has a very limited number of free electrons available for current conduction. Hence, when the semiconductor reaches an equilibrium potential, there is a layer in the semiconductor, around the point of contact, from which all electrons have been removed. Other electrons from the semiconductor cannot enter this layer because of the repulsion of the negative charge accumulated on the tungsten point. This layer is known as the blocking layer and may be between 10^{-3} and 10^{-4} mm thick. It is this blocking layer which causes the crystal rectifier to have a unilateral characteristic.

When voltage is applied between the contact wire and the semiconductor, the device may pass current or it may block, depending upon the polarity and magnitude of the applied voltage. If the semiconductor is made negative with respect to the contact wire, electrons will flow from the tungsten to the external circuit and thus remove the negative charge from the contact metal, which repelled the free electrons in the semiconductor. This action wipes out the blocking layer, and hence a current can flow which is impeded only by the bulk resistance of the semiconductor. If the semiconductor is made positive with respect to the tungsten contact wire, electrons flow from the external circuit to the tungsten and hence increase the negative charge in this metal at the point of contact.

This charge has the effect of increasing the thickness of the blocking layer. Since there are no free electrons in the blocking layer, there can be no current flow through it unless electrons in the metal have enough energy to overcome the restraining forces and pass from that metal to the semiconductor. Since the thickness of the blocking layer increases with potential difference, the voltage required to cause conduction in the reverse direction is quite high. Thus point-contact rectifiers have relatively low resistance in one direction and an extremely high resistance in the reverse direction. Figure 16-9 shows the volt-ampere characteristic for a type 1N34, which is an n-type germanium diode.

If the contacting metal has a lower work function than the n-type semiconductor, the device will not rectify since no blocking layer can be established. For a p-type crystal the work function of the contact must be less than that of the crystal or no blocking layer will be set up.

When the germanium diode is used as a detector, the forward resistance is considerably lower than that of the usual vacuum tube. Its back resistance is not as great as that for the vacuum tube, and hence this latter fact may limit it to particular applications.

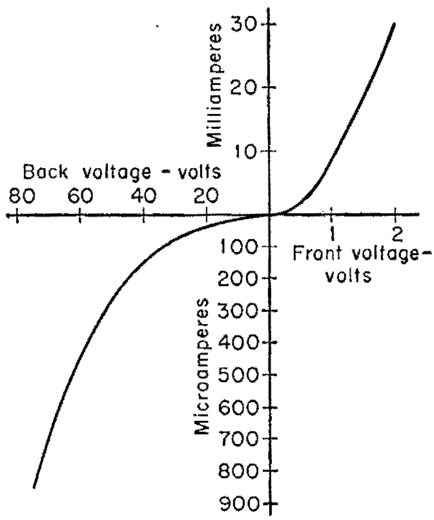


FIG. 16-9. Volt-ampere characteristic for a type 1N34 germanium diode. Note the changes in scale.

16-6. Large-area Rectifiers.[1,5] Large-area rectifiers, usually known as dry rectifiers, operate along the same principles as does the point-contact rectifier. They utilize the unilateral characteristic of a blocking layer developed in a semiconductor. They usually consist of a layer of semiconductor bonded onto a metal base with electrical contact being made to the semiconductor and to the metal base. The contact to the semiconductor is a large-area one under considerable pressure. Rectification depends upon the blocking layer established at the surface of the semiconductor when the correct polarity of voltage is applied.

One of the most common types of dry rectifiers is the copper–cuprous oxide cell developed by Grondahl in 1926. The cuprous oxide is formed on one surface of a copper washer by heating it to a high temperature in an oxidizing atmosphere. The oxidized washers are then stacked alternately with soft lead washers and then placed under pressure. Connections are made to the bottom copper washer and to the top lead washer as shown in Fig. 16-10.

A typical volt-ampere curve for a copper-oxide rectifier is shown in Fig. 16-11. Each disk can withstand about 8 to 10 volts in the reverse direction; hence, to rectify 100 peak volts, at least 10 disks would have to be stacked. One drawback of such a device is that it cannot operate at a temperature above about 45°C. For this reason it is built so that the copper disks extend out beyond the boundaries of the oxidized area in order to make the disks serve as cooling fins.

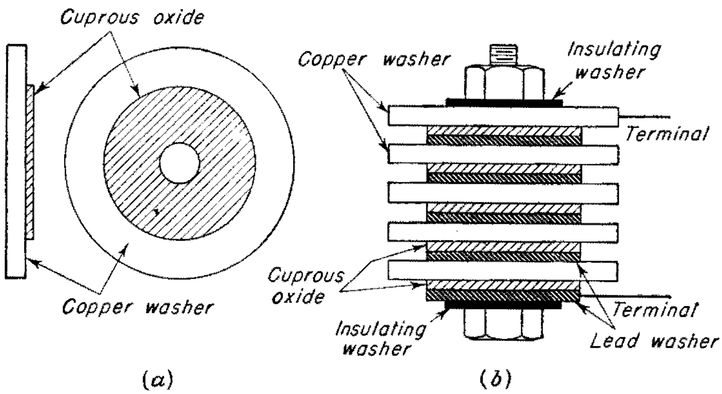


FIG. 16-10. Construction of a large-area copper–cuprous oxide rectifier. (a) A single disk. (b) The assembled unit.

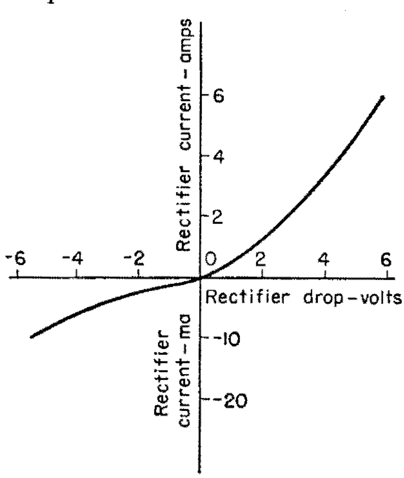


FIG. 16-11. Volt-ampere characteristic for a single copper–cuprous oxide disk.

The selenium rectifier, which is used in many small radios, is built on a similar principle. The selenium is deposited on iron disks which are then stacked with lead washers and placed under compression. The selenium rectifier can operate at temperatures up to 75°C and will withstand back voltages as great as 25 volts per disk. This allows them to be more compact, for a given electrical capacity, than is the copper-oxide rectifier.

16-7. Varistors. When four rectifiers, either the small-area or the large-area type, are connected in a bridgelike arrangement as shown in Fig. 16-12, they are known as a varistor. Among the various circuits in which they are used is the a-c milliammeter shown in Fig. 16-12a and the modulator shown in Fig. 16-12b.

16-8. Transistors.[6,11] The transistor is a semiconductor used in such a way that it can amplify voltage, current, or power without the use of vacuum tubes. One form of such a device, known as a type A transistor, is shown in Fig. 16-13. It consists of two adjacent metallic points making rectifying contacts with a small slab of germanium. The third contact is a large-area one which comprises the base electrode. The point contacts are called the emitter and the collector and are input and output terminals, respectively.

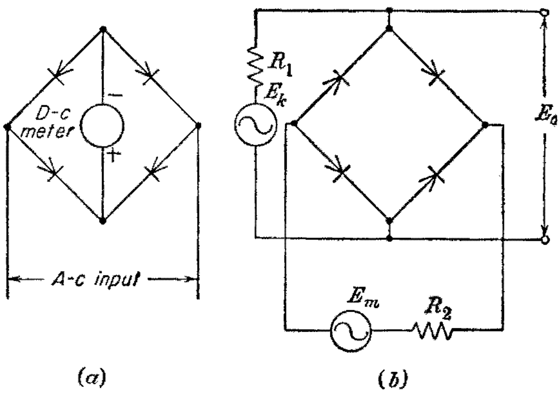


FIG. 16-12. Varistors used (a) to convert a d-c milliammeter for use with alternating current, (b) as a modulator.

Either n-type or p-type germanium can be used as the semiconductor. The only difference in the circuits for these two types is in the polarities of the biases for the emitter and collector. The emitter must always be biased in the conducting or forward direction, while the collector is biased in the nonconducting or reverse direction. The biasing arrangements for the two types are shown in Fig. 16-14. Only the n-type transistor will be discussed here, but the action of the p type is quite similar except that account must be taken of the fact that in the p-type crystal we have hole conduction while in the n type we have electron conduction.

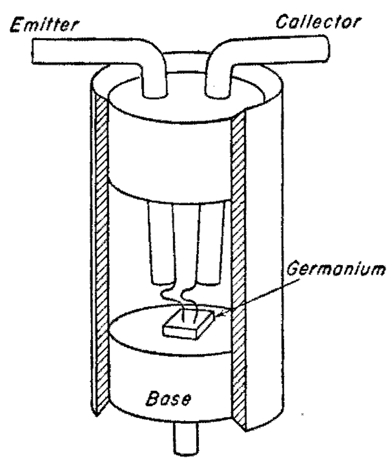


FIG. 16-13. Cutaway view of a transistor.

In the article on point-contact rectifiers the impression might have been given that conduction in an n-type crystal is entirely by electrons. This is not the case, because if it were, transistors could not work. Actually, when a point-contact rectifier is biased in the forward direction, the point injects a comparatively large number of holes into the crystal. These holes drift toward the negative terminal and may reach it, or they may recombine with an electron en route and again become neutral.

In the case of a transistor, the collector is biased in the reverse direction, and hence it may appear that it should not pass current. However, the emitter is biased in the forward direction and does pass current, injecting a considerable number of holes into the crystal. These holes, being positive in polarity, will drift over toward the collector, which is more negative than the base electrode. On arrival they remove electrons from the

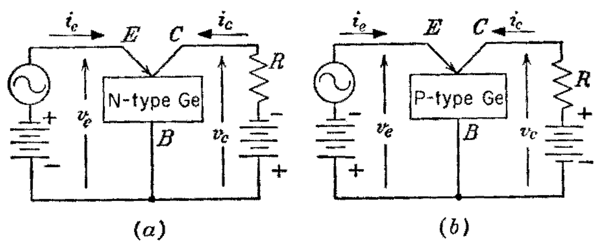


FIG. 16-14. Circuit arrangement for transistors when used as amplifiers.

collector, which results in collector-current flow. Since the number of holes injected by the emitter depends on the emitter current, the larger this current the larger will be the collector current. In addition to the hole current, the positive charge around the emitter results in the establishment of a positive space charge surrounding the collector. This space charge tends to pull electrons from the collector and hence to increase the collector current. The net result is that the collector current is greater than the emitter current, which means that we have current amplification. By the proper choice of load resistors, voltage amplification can be accomplished.

Another type of transistor which has been developed has characteristics which differ somewhat from those of the A type. This new transistor is made in two forms, one of which is known as the p-n-p type and the other as the n-p-n type. Both types are known as junction transistors.

The p-n-p type is made by sandwiching a very thin layer of n-type germanium between two pieces of p-type germanium. This is actually done by using a relatively long piece of p-type germanium and converting a very thin layer, midway along its length, into n-type germanium by nuclear bombardment. One of the p-type sections serves as the emitter, the other as the collector, and the n-type layer as the base. The emitter electrode is biased in the conducting direction. Under these circum-

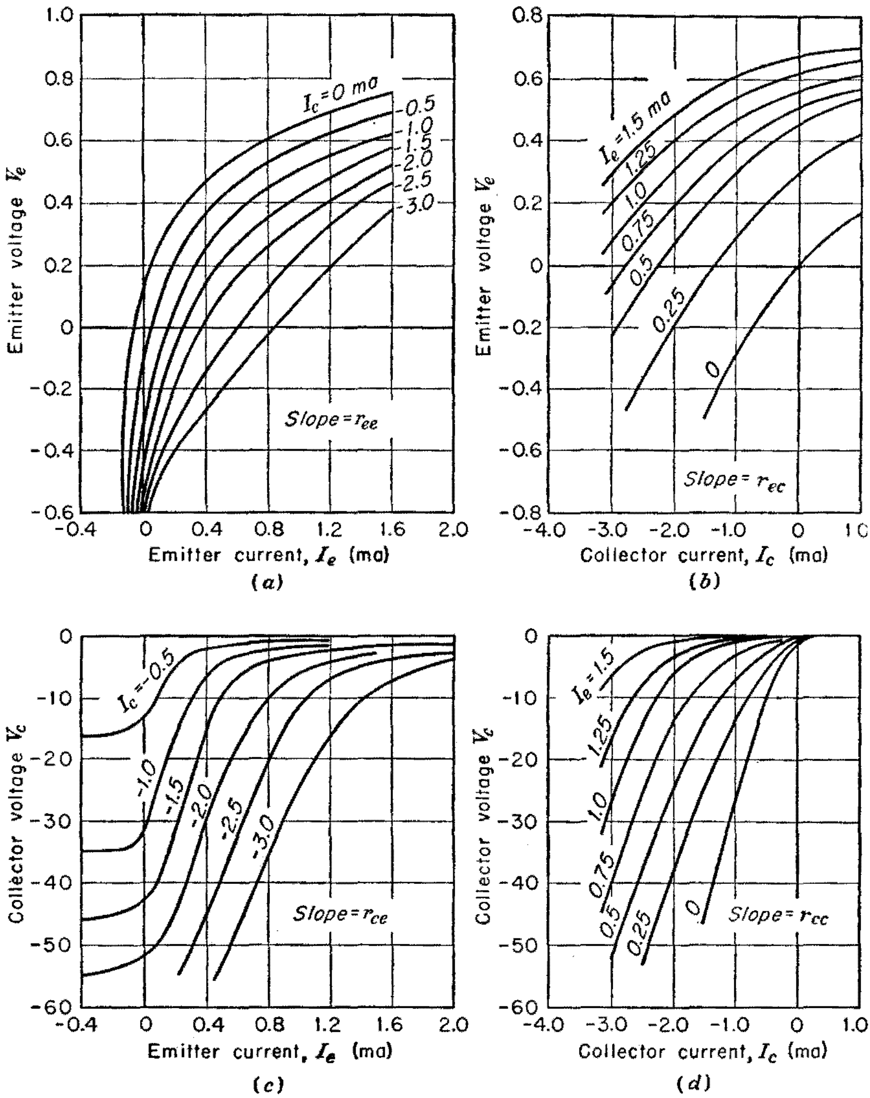


FIG. 16-15. Static characteristics for a typical transistor. (*W. Shockley, Bell Telephone Laboratories, Inc., "Electrons and Holes in Semiconductors," pp.* 38, 39, *D. Van Nostrand Company, Inc., New York,* 1950. *By permission.*)

stances it has been found that there is hole injection into the n-type germanium just as there was in the case of the type A transistor. The other section of p-type germanium is biased in the nonconducting direction, and hence it acts as the collector.

The n-p-n type transistor is similar in construction to the p-n-p type, the only difference being that the former consists of a layer of p-type germanium sandwiched between two sections of n-type germanium. The polarities of the biases must be the reverse of those for the p-n-p type.

Figure 16-15 shows families of static characteristics for a typical point-contact-type transistor. Note that, like electronic tubes, these are not linear devices, but when operated over limited regions, they can be considered approximately linear. Hence an equivalent circuit may be useful. In order to find such an equivalent circuit, we may express the equations for the static characteristics in functional form. Thus

$$V_e = f_1(I_e, I_c) \tag{16-1}$$

and

$$V_c = f_2(I_e, I_c) \tag{16-2}$$

where V and I signify steady values.

We may write the total differentials of these two functions in the approximate forms

$$\Delta V_e = \frac{\partial V_e}{\partial I_e} \Delta I_e + \frac{\partial V_e}{\partial I_c} \Delta I_c \tag{16-3}$$

and

$$\Delta V_c = \frac{\partial V_c}{\partial I_e} \Delta I_e + \frac{\partial V_c}{\partial I_c} \Delta I_c \tag{16-4}$$

But if we let the incremental quantities be the instantaneous alternating values of voltage and current and note that the partial derivatives are the slopes of the characteristic curves and have the dimensions of resistance, we may write Eqs. (16-3) and (16-4) as

$$v_e = r_{ee} i_e + r_{ec} i_c \tag{16-5}$$

$$v_c = r_{ce} i_e + r_{cc} i_c \tag{16-6}$$

where $r_{ee} = \partial V_e / \partial I_e$ is the slope of the curve in Fig. 16-15a at a given operating point. $r_{ec} = \partial V_e / \partial I_c$ is a mutual resistance and is the slope of the curve in Fig. 16-15b at the operating point. $r_{ce} = \partial V_c / \partial I_e$ is the slope of the curve of Fig. 16-15c, and $r_{cc} = \partial V_c / \partial I_c$ is the slope of the curve of Fig. 16-15d.

An equivalent circuit may be drawn from Eqs. (16-5) and (16-6). This has been done in Fig. 16-16a. Another form of equivalent circuit has been drawn in Fig. 16-16b. It can be easily shown that these two circuits give identical results if we make

$$r_e = r_{ee} - r_{ec} \tag{16-7}$$

$$r_b = r_{ec} \tag{16-8}$$

$$r_c = r_{cc} - r_{ec} \tag{16-9}$$

$$r_m = r_{ce} - r_{ec} \tag{16-10}$$

r_e is known as the emitter resistance, r_b is the base resistance, r_c the collector resistance, and r_m the mutual resistance. Typical values of these resistances are $r_e \approx 250$ ohms, $r_b \approx 300$ ohms, $r_c \approx 20{,}000$ ohms, $r_m \approx 35{,}000$ ohms.

When a transistor is connected as an amplifier, the equivalent circuit becomes quite simple. Let $\mathbf{V}_g$ be the open-circuit voltage and R_c be the internal resistance of the input-signal source. R_L is the resistance of the load. The equivalent circuit may then be drawn as in Fig. 16-17*a*.

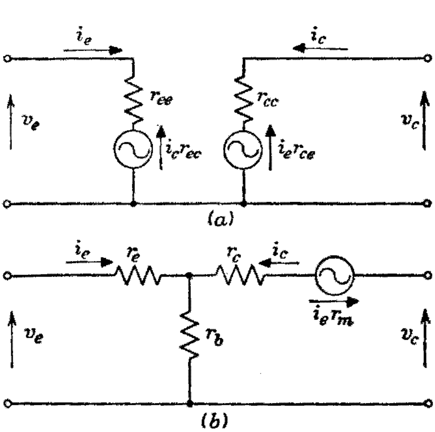


FIG. 16-16. Equivalent circuits for transistors.

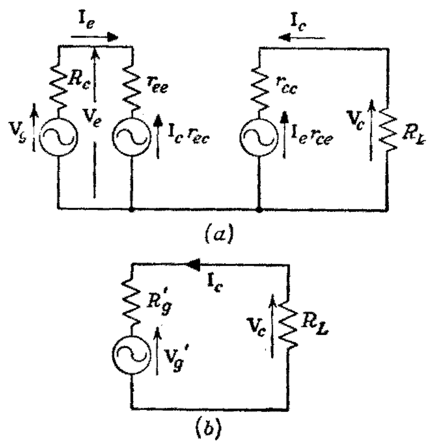


FIG. 16-17. Equivalent circuits for a transistor used as an amplifier.

We may use effective values of voltage and current and obtain loop equations:

$$\mathbf{V}_e = \mathbf{V}_g - \mathbf{I}_e R_c = r_{ee}\mathbf{I}_e + \mathbf{I}_c r_{ec} \tag{16-11}$$

$$\mathbf{V}_c = -\mathbf{I}_c R_L = \mathbf{I}_e r_{ce} + \mathbf{I}_c r_{cc} \tag{16-12}$$

Solving Eq. (16-11) for $\mathbf{I}_e$ gives

$$\mathbf{I}_e = \frac{\mathbf{V}_g - \mathbf{I}_c r_{ec}}{r_{ee} + R_c} \tag{16-13}$$

Substituting from Eq. (16-13) into Eq. (16-12) yields

$$\mathbf{V}_c = \mathbf{I}_c\left(r_{cc} - \frac{r_{ec} r_{ce}}{r_{cc} + R_c}\right) + \frac{r_{ce}}{r_{cc} + R_c}\mathbf{V}_g \tag{16-14}$$

or

$$\mathbf{V}_c = \mathbf{V}_g' + \mathbf{I}_c R_c' \tag{16-15}$$

where

$$\mathbf{V}_g' = \frac{r_{ce}}{r_{cc} + R_c}\mathbf{V}_g \tag{16-16}$$

and

$$R_c' = r_{cc} - \frac{r_{cc} r_{ce}}{r_{cc} + R_c} \tag{16-17}$$

If an equivalent circuit for Eq. (16-15) is drawn, Fig. 16-17*b* results. The current $\mathbf{I}_c$ can be written as

$$\mathbf{I}_c = \frac{\mathbf{V}'_g}{R_c + R_L} \tag{16-18}$$

Hence the power supplied to the load R_L is

$$P_{ac} = I_c{}^2 R_L = \left(\frac{V'_g}{R_c + R_L}\right)^2 R_L \tag{16-19}$$

16-9. Photoconductive Cells. The earliest photoelectric cells were of the photoconductive type. This cell consists of a thin layer or film of semiconductor between two electrodes. Such a cell is described in Art. 15-1. Selenium is the material ordinarily used as the photoconductor, although there are several of the metallic oxides which can also be used.

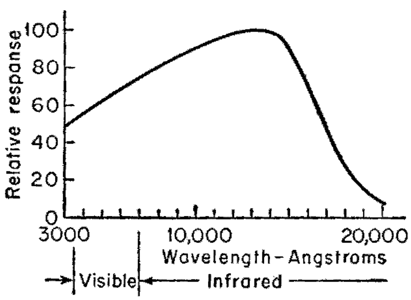


FIG. 16-18. Spectral response of a selenium photoconductive cell.

The theory of operation for these cells is the same as that for the conduction of current in a semiconductor. This was explained in Art. 16-2. Electrons in the filled band of the selenium are given energy by the radiant energy striking the surface. This energy raises some of the electrons from the filled band to the unfilled band, where they become conduction electrons. The more intense the radiant energy, the greater the number of conduction electrons in the unfilled band and hence the smaller the resistance becomes. The dark resistance (the resistance of the cell when there is no light energy falling on it) depends on the material used, the thickness of the film, the spacing between the electrodes, and the length of the electrodes; usually it is of the order of megohms. When light falls on the cell, the resistance may decrease by as much as 25 per cent.

The difference between the energy levels of the filled band and unfilled or allowed band of a semiconductor is on the order of 0.5 electron-volt. Hence the minimum quantum of energy needed to increase the conductivity of such a cell is less than the quantum of energy needed to cause emission in a photoemissive cell where the work function of the surface is about 1 electron-volt. Photoconductive cells can therefore be used with longer-wavelength light than can photoemissive cells. For this reason they find one of their chief uses when infrared light is to be detected. Figure 16-18 shows the response of a typical selenium cell.

A rather recent and interesting use of this type of photosurface is in the television camera tube known as the Vidicon. This tube is used in industrial equipment where slow response is not a drawback. This slow response is the chief reason why photoconductive cells are not more commonly used.

16-10. Photovoltaic Cells. The photovoltaic cell operates because a voltage is generated when light falls on the photosensitive surface. Figure 16-19 shows the construction of such a cell. A layer of semiconductor is either formed on or bonded to a metallic base electrode. The surface of the semiconductor is then coated with a very thin transparent metallic film or with a transparent conducting lacquer. Light passes through the transparent electrode and falls on the surface of the semiconductor, and, as in the photoconductive cell, electrons are raised from the filled band to the conduction band. At the same time a voltage appears between the semiconductor and the transparent electrode. This effect was explained in Art. 16-5, where the point-contact rectifier was discussed. If the two electrodes are connected together electrically, a current will flow which is proportional to the light energy falling on the cell.

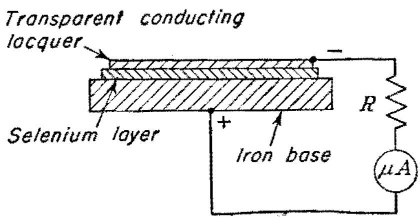


FIG. 16-19. Construction of a selenium-iron photovoltaic cell.

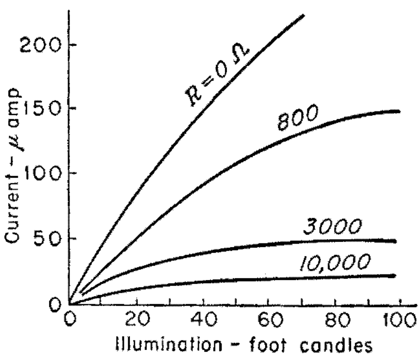


FIG. 16-20. Response of an iron-selenium photovoltaic cell as the external-circuit resistance is varied.

The potential developed between the two electrodes does not vary linearly with the intensity of the light falling on the surface. This voltage is on the order of a few hundred millivolts in bright light. The cell resistance is very small, and hence, if the external circuit has low resistance, an appreciable current may flow. The magnitude of the current may be as great as a few milliamperes in strong sunlight. The resistance of the semiconductor varies with the light intensity in such a manner that the short-circuit current which flows varies almost linearly with light intensity.

There are two common types of this cell in use today. One is the copper–copper oxide cell, and the other is the iron-selenium cell. Figure 16-20 shows the characteristics of a typical iron-selenium cell as the shorting resistance is varied. Since such a cell is more linear and has greater current output when used with a low-resistance short, it is advisable to use a low-resistance microammeter of the desired range as the short for the cell.

These cells find their greatest usage as foot-candle meters or as photographic exposure meters. In some instances they have been used to operate directly a sensitive relay which in turn may operate a larger relay. The latter then could be used to turn on or turn off large electrical loads. The cells cannot be used with modulated-light sources because of their extremely poor frequency response. The large electrodes

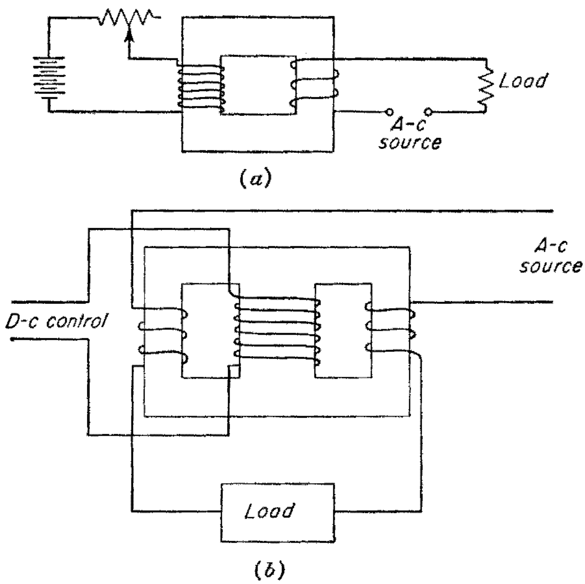


FIG. 16-21. Saturable-core reactor circuits.

form a closely spaced capacitor which shorts out any varying current generated by the cell. This shorting effect is so bad that a light modulated at a 60-cps rate will cause only half the current output that a constant light would cause.

16-11. Magnetic Amplifiers.[7] Saturable-core reactors, the heart of the magnetic amplifier, have been used in conjunction with electrical machinery since about 1895. However, they did not reach a high state of development until the Second World War. During this conflict the Germans set up an extensive program for the development of the magnetic amplifier. They spent millions of dollars to bring this device to the stage where it could be used in gun stabilizers, automatic pilots, servo control systems for their long-range rockets, blind-landing aids, etc. These devices also found many civilian applications such as in computing machines, electric brakes for locomotives, high-voltage power-line controls (where they could control up to 50,000 kva), streetcar controls, and many others. Since the war, engineers in this country have also found many uses for magnetic amplifiers and have made improvements on them.

A saturable-core reactor in its simplest form is shown in Fig. 16-21*a*. The secondary, or load, current is alternating and is controlled by the impedance of the secondary. This impedance in turn is controlled by the magnitude of the direct current flowing through the primary, or control, winding. When the direct current is zero, the impedance of the secondary is a maximum and the alternating current flowing through it is a minimum. If the direct current is increased sufficiently, the magnetic core becomes saturated and hence the permeability of the magnetic path becomes practically that of air. The incremental inductance then becomes about the same as though the coil had an air core. The reactance of the coil becomes very small, and hence the current increases accordingly.

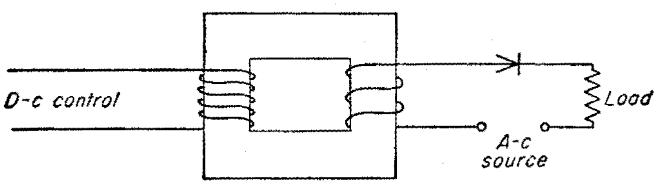


FIG. 16-22. A simple magnetic amplifier.

If the control winding has a large number of turns, only a small direct current is required to saturate the core. Hence a small direct current can be used to control a relatively large alternating current. This type of circuit has been used quite extensively for the control of lighting loads, particularly for stage lighting.

One drawback to this device as described is caused by the alternating flux from the secondary linking with the control winding and inducing a voltage in it. If the control winding has a large number of turns, this induced voltage may become quite high and may even cause breakdown of the insulation. A much more satisfactory arrangement for a saturable reactor is shown in Fig. 16-21*b*. In this circuit the net alternating flux linking the control winding is zero, and hence the induced voltage is zero.

The saturable reactor just described might be classified as a magnetic amplifier, although it is a very poor one. By inserting a rectifier as shown in Fig. 16-22 a much more efficient magnetic amplifier is obtained. The rectifier in series with the load results in a unidirectional load current which assists the control winding in saturating the core. Less control power is now needed, for a given amount of control, than if the rectifier had not been used.

In order to understand better how a magnetic amplifier works, let us first examine the magnetization curve for the magnetic core. Figure 16-23*a* shows such a curve. Actually the material goes through a hysteresis loop as the magnetizing current varies over a cycle, but for simplicity we shall assume that the loop is very thin so that it can be approximated by the curve given. Furthermore, we need not consider the part of the curve associated with negative values of B and H, because of the rectifier.

If the direct current is zero and the alternating voltage is of such a magnitude that H varies from zero to H_3, then the impedance of the

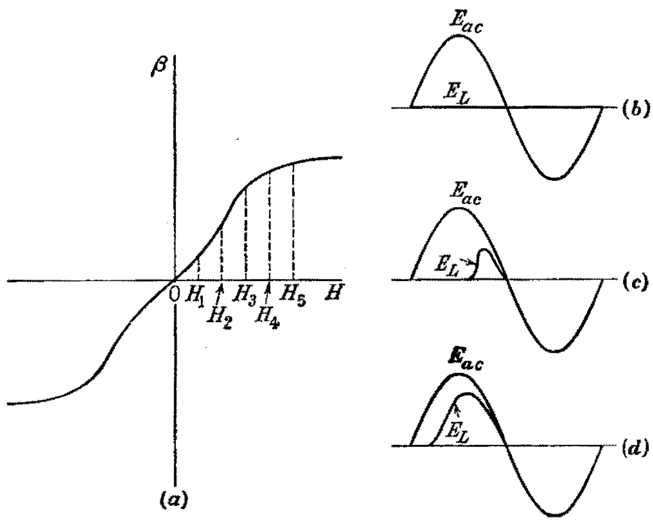


FIG. 16-23. Operation of a magnetic amplifier.

secondary will be large at all times and the drop across the load will be approximately zero as shown in Fig. 16-23*b*. If the direct current is now increased until the magnetizing force caused by it is H_1, then the rectified alternating voltage may cause current to flow such that the magnetizing force varies between H_1 and H_4. This results in a load-voltage waveshape as shown in Fig. 16-23*c*. A further increase in direct current might result in the magnetizing force varying between H_2 and H_5. This would result in the load voltage shown in Fig. 16-23*d*.

The simple magnetic amplifier just described performs very poorly because of voltage induced in the control winding. More complicated and far more satisfactory circuits are shown in Fig. 16-24. Circuit *a* is the basic circuit for a-c loads, while circuit *b* is for the control of d-c loads from an a-c source.

Feedback may be used with these amplifiers to change their characteristics. An additional coil, however, must be wound on the center leg of the core. Figure 16-25 shows two feedback circuits, where (*a*) is for d-c loads and (*b*) is for a-c loads. In both cases the output is rectified and fed back to the additional coil in such a fashion that its mmf either aids or opposes the mmf set up by the control coil. If this mmf aids, the feedback is positive and results in increased control. Gains of several million per stage can be attained using these circuits with positive feedback. However, with this high gain they may be very unstable. If

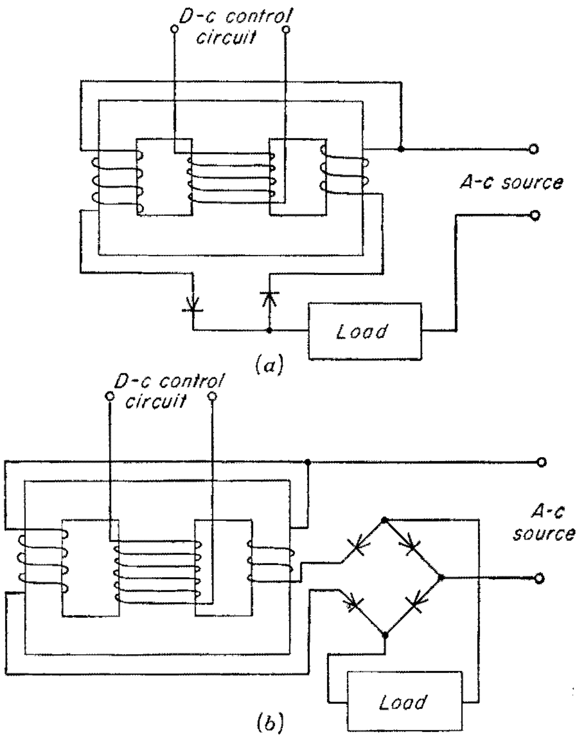


FIG. 16-24. Magnetic-amplifier circuit for (*a*) a-c loads, (*b*) d-c loads.

the rectified output is fed back in such a manner as to oppose the control-coil mmf, the feedback is negative and the gain is decreased but the response becomes more linear. Figure 16-26 shows the effect of feedback on the load voltage as a function of the control ampere turns as the feedback goes from negative to positive.

Up to this point only magnetic amplifiers using d-c control circuits and a-c or d-c loads have been discussed. These devices also serve very satisfactorily as a-c amplifiers for audio and even radio frequencies. When used as an audio amplifier, the audio signal replaces the d-c control source and the a-c source must have a frequency somewhat greater than the highest-frequency audio component to be amplified. Naturally the iron used in the core must be able to handle these higher frequencies without too much core loss. Provision must be made at the output of the amplifier to separate the unwanted frequency from the audio frequencies. Figure 16-27 shows the circuit of a simple a-f magnetic amplifier.

Multistage amplifiers have been constructed with power outputs of 500 watts and excellent linearity up to 7000 cps.

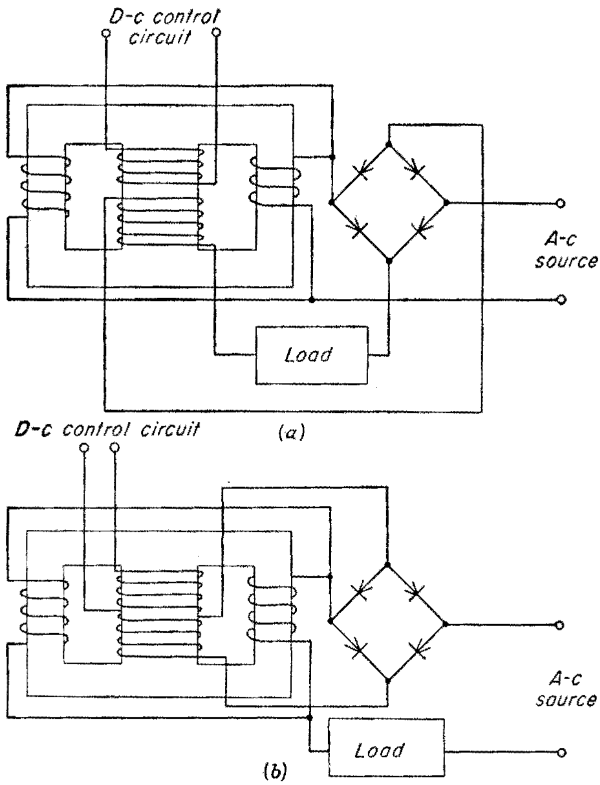


FIG. 16-25. Magnetic-amplifier circuits with feedback (a) for d-c loads, (b) for a-c loads.

If magnetic amplifiers are to be used at radio frequencies, particular care must be exercised in selecting a core material. Experiments have shown that some of the rolled magnetic alloys can perform with fair efficiency up to several hundred kilocycles. Above this frequency other core materials must be used. Powdered-iron cores do not make satisfactory magnetic amplifiers because of the relatively large air gaps caused by the separation between particles of iron. A satisfactory core can be made from magnetic ferrites. Metallic oxides are pressed or extruded into the desired shape and then fired at a temperature of about 1200°C in order to process them so that they have the desired characteristics. They then have a maximum permeability exceeding 1000 (compared with powdered iron with about 100) and they also have very high ohmic resistances. These characteristics make them very desirable for core materials.

Magnetic amplifiers have found a very definite place in modern electronics. They replace the vacuum tube in some applications and supplement it in others. A few of the applications are listed below.

Amplifiers, a-c and d-c for current, voltage, and power. As an example, they might be used with thermocouples and photocells to control

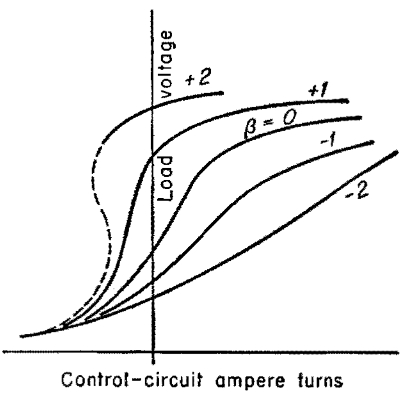


FIG. 16-26. Effects of negative and of positive feedback with magnetic amplifiers. ($|\beta| \propto$ No. feedback turns.)

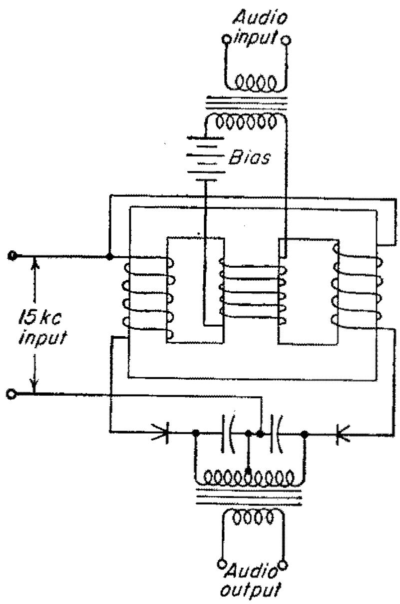


FIG. 16-27. A magnetic amplifier for use with audio frequencies.

motors directly. The magnetic amplifier is excellent with direct voltages since it is not subject to the drift difficulties often encountered in electronic-tube amplifiers.

Regulators. Control of current, voltage, and frequency of industrial power installations.

Motor starters and controls.

Servo systems. Complete systems have been built utilizing magnetic amplifiers as regulators, converters, and computers, and as a replacement for thyratrons and for a mechanical amplidyne system.

The magnetic amplifier has several advantages over the electronic-tube amplifier. It is more rugged and less affected by power-line variations.

It can carry larger overloads, has longer life, and requires no warm-up. As disadvantages, however, it has a long time constant and a relatively low input impedance; furthermore, it costs more than the electronic-tube amplifier.

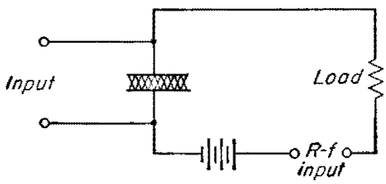


FIG. 16-28. Basic circuit for a dielectric amplifier.

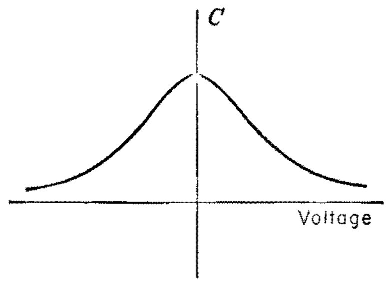


FIG. 16-29. Variation of capacitance with applied voltage for a dielectric-amplifier material.

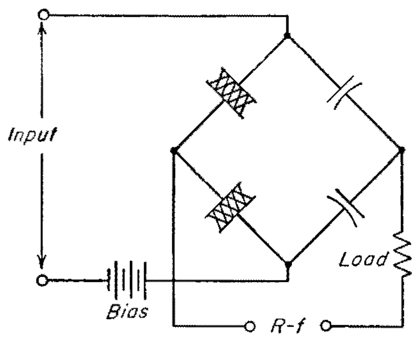


FIG. 16-30. A practical circuit for a dielectric amplifier.

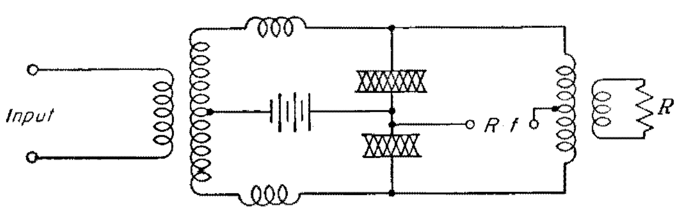


FIG. 16-31. A push-pull dielectric amplifier.

16-12. Dielectric Amplifiers.[8] A new and novel device, which is being investigated, is the dielectric amplifier. As yet it has found very few applications but shows promise of being very useful.

Its operation is very similar to that of the magnetic amplifier except that it works on the principle that the dielectric constant for certain insulators (such as barium zirconate) varies with the voltage across it. Figure 16-28 shows the basic circuit for such a device. Following a variation in the control voltage, the dielectric constant of the capacitor, the latter's capacitance, and the r-f load current all change. Figure 16-29 shows how the capacitance varies with the applied control voltage.

Figure 16-30 shows a much more satisfactory arrangement for a dielectric amplifier. This circuit has a high impedance and is balanced, with no r-f currents in the control leads. Figure 16-31 is a push-pull arrangement of a dielectric amplifier.

Applications of the dielectric amplifier parallel those of the magnetic amplifier. However, they have one advantage over these employing magnetic amplifiers because of their higher input impedances.

REFERENCES

1. Grondahl, L. O., and P. H. Geiger: A New Electronic Rectifier, *J. AIEE*, **46**, 215 (1927).
2. Millman, J., and S. Seely: "Electronics," 2d ed., McGraw-Hill Book Company, Inc., New York, 1951.
3. Becker, J. A., C. B. Green, and G. L. Pearsons: Properties and Uses of Thermistors, *Bell System Tech. J.*, **26**, 170 (1947).
4. Cornelius, E. C.: Germanium Crystal Diodes, *Electronics*, **19**, 118 (February, 1946).
5. Ramsey, G.: The Selenium Rectifier, *Elec. Eng.*, **63**, 425 (1944).
6. Becker, J. A., and J. N. Shive: The Transistor—A New Semiconductor Amplifier *Elec. Eng.*, **68**, 215 (1949).
7. Ogle, H. M.: The Amplistat and its Application, *Gen. Elec. Rev.*, (February, August, October, 1950).
8. Vincent, A. M.: Dielectric Amplifier Fundamentals, *Electronics*, **24**, 84 (December, 1951).
9. Department of the Navy, Bureau of Ships: "Magnetic Amplifiers," 1951.
10. Seitz, F.: "The Modern Theory of Solids," McGraw-Hill Book Company, Inc., New York, 1940.
11. Shockley, W.: "Electrons and Holes in Semiconductors," D. Van Nostrand Company, Inc., New York, 1950.
12. "Symposium on Electrical Properties of Semiconductors and the Transistor," American Institute of Electrical Engineers, New York, 1950.
13. Torrey, H. C., and C. A. Whitmer: "Crystal Rectifiers," McGraw-Hill Book Company, Inc., New York, 1948.

APPENDIX A

CHARACTERISTIC CURVES FOR SEVERAL TUBES

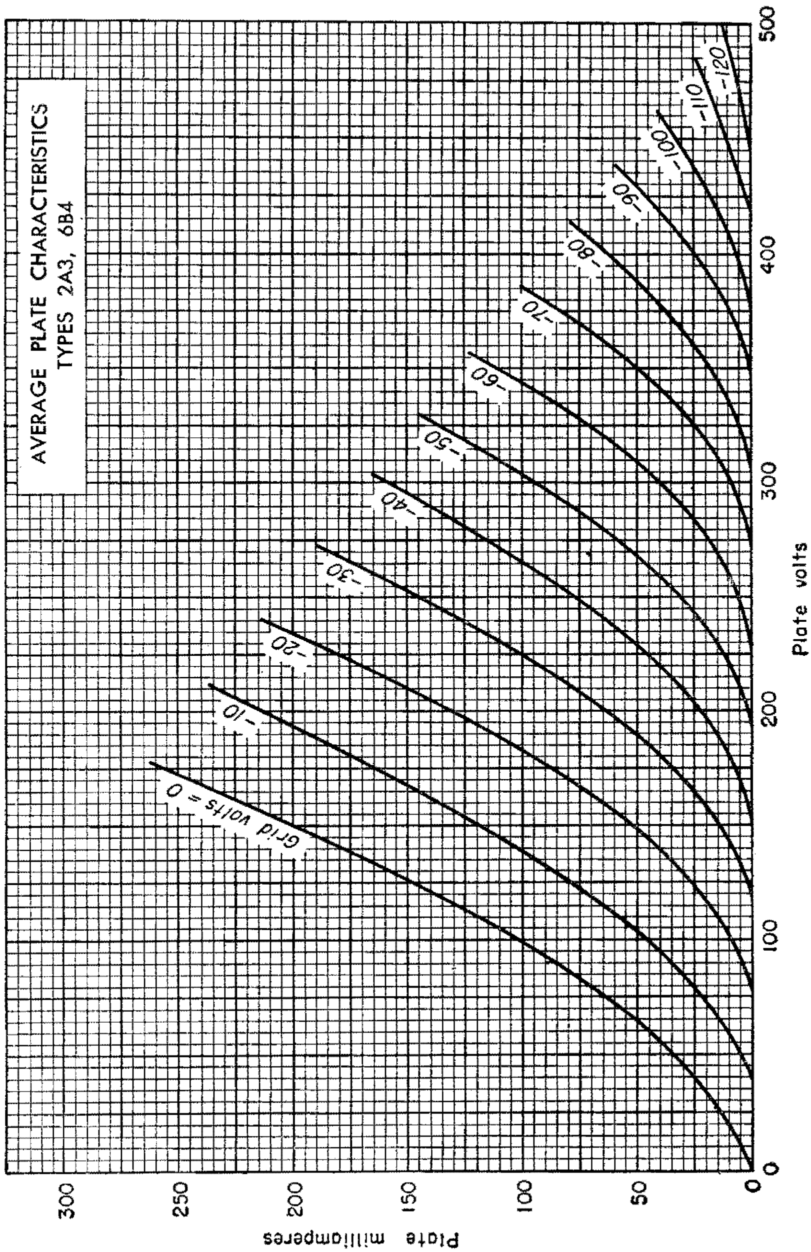


FIG. A-1.

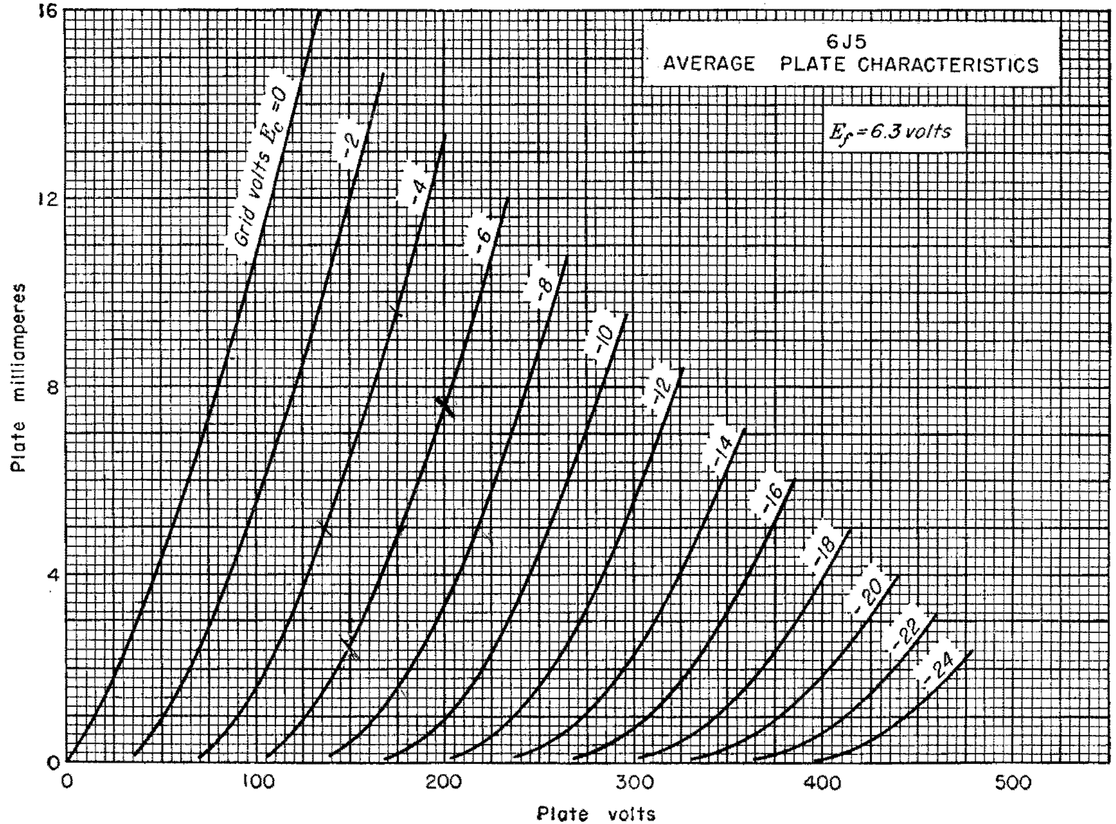


FIG. A-2. (*Courtesy of RCA.*)

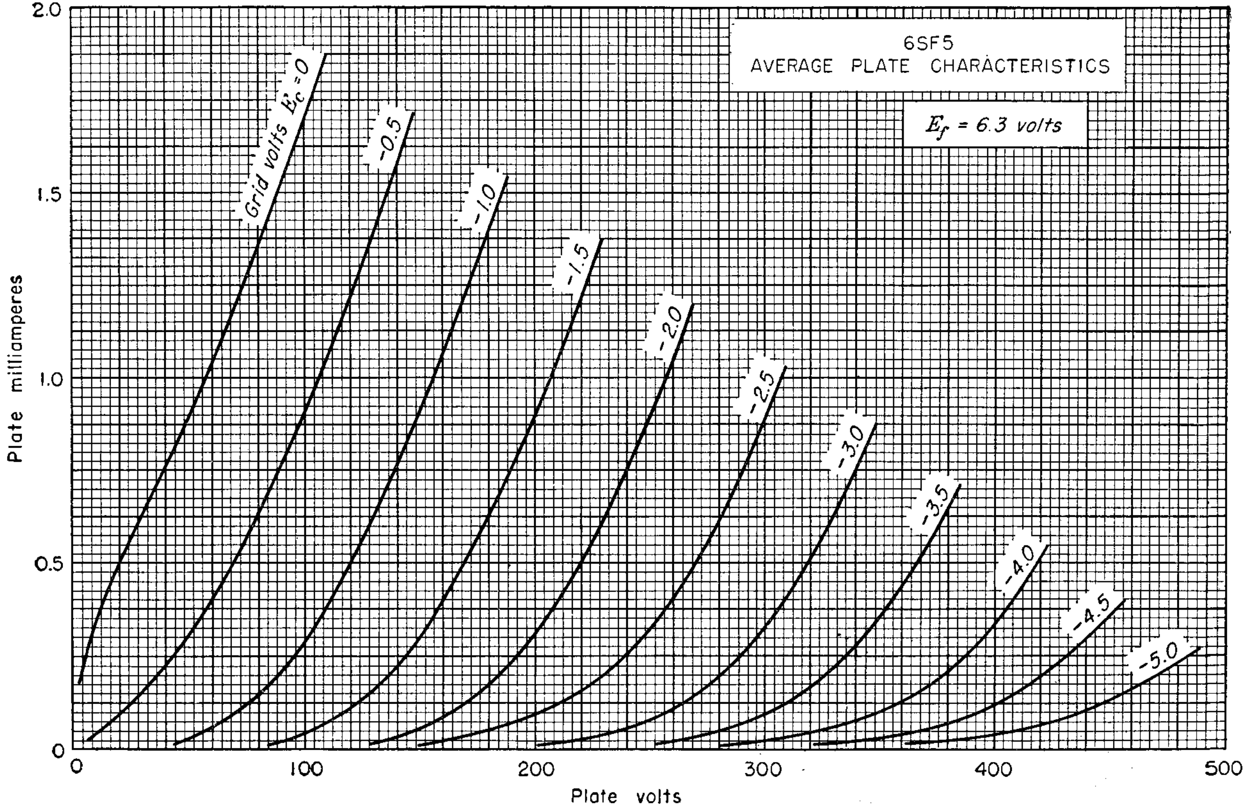


FIG. A-3. (*Courtesy of RCA.*)

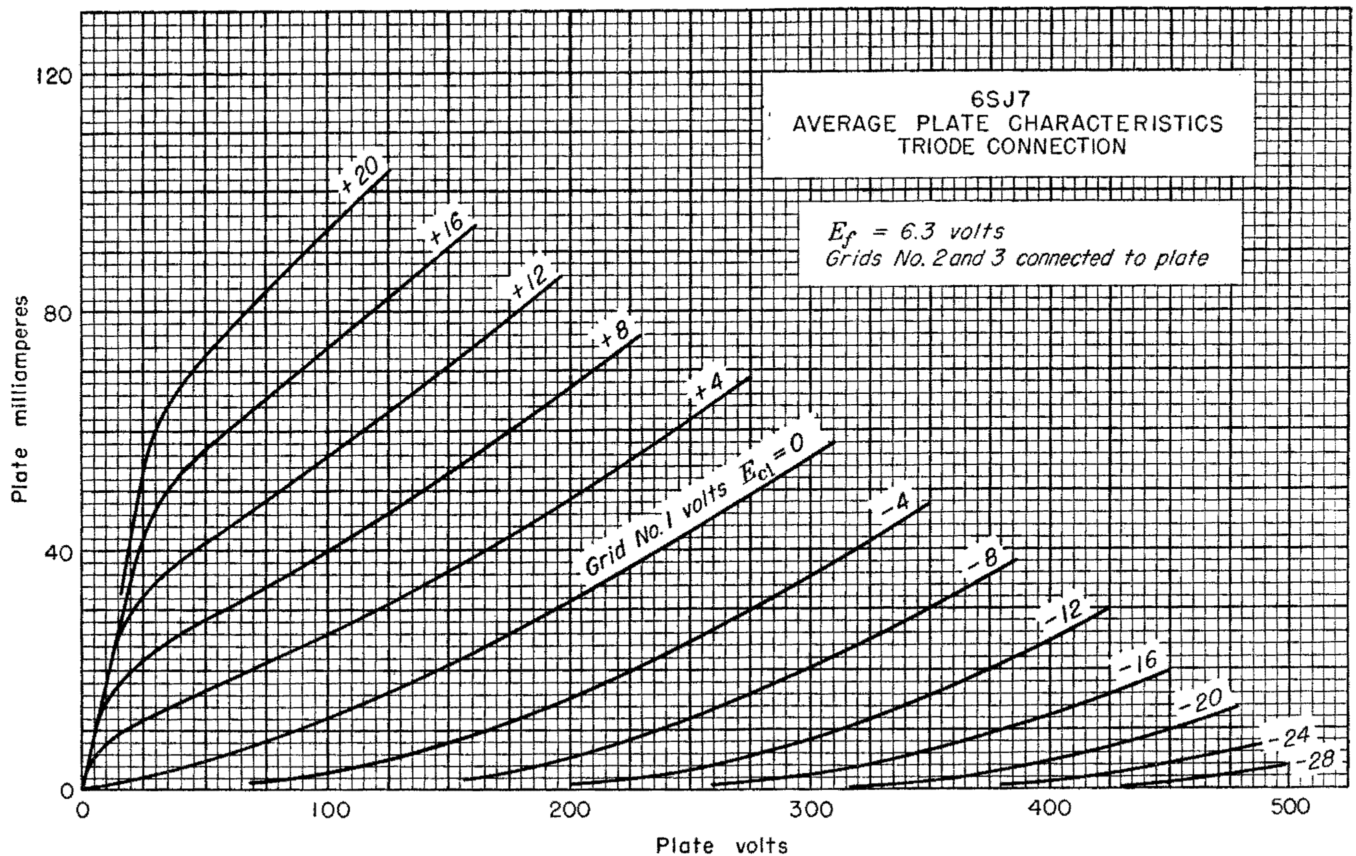


FIG. A-4. *(Courtesy of RCA.)*

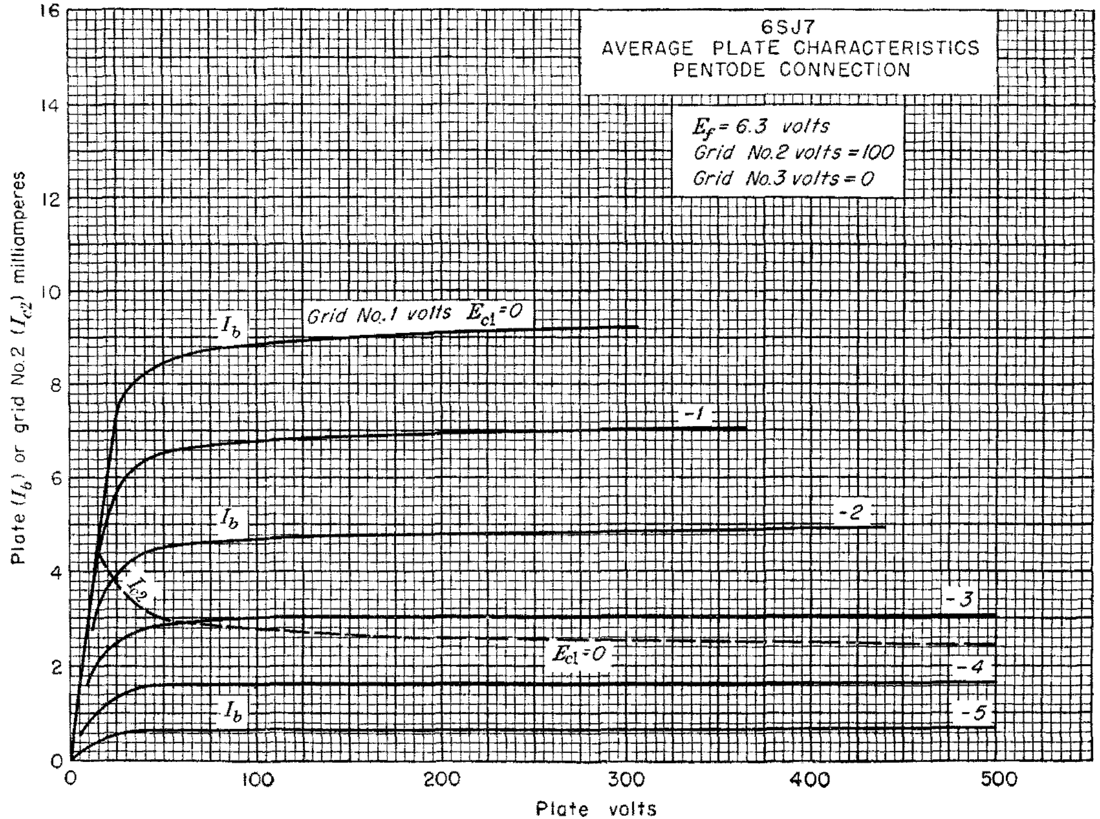


FIG. A-5. (*Courtesy of RCA.*)

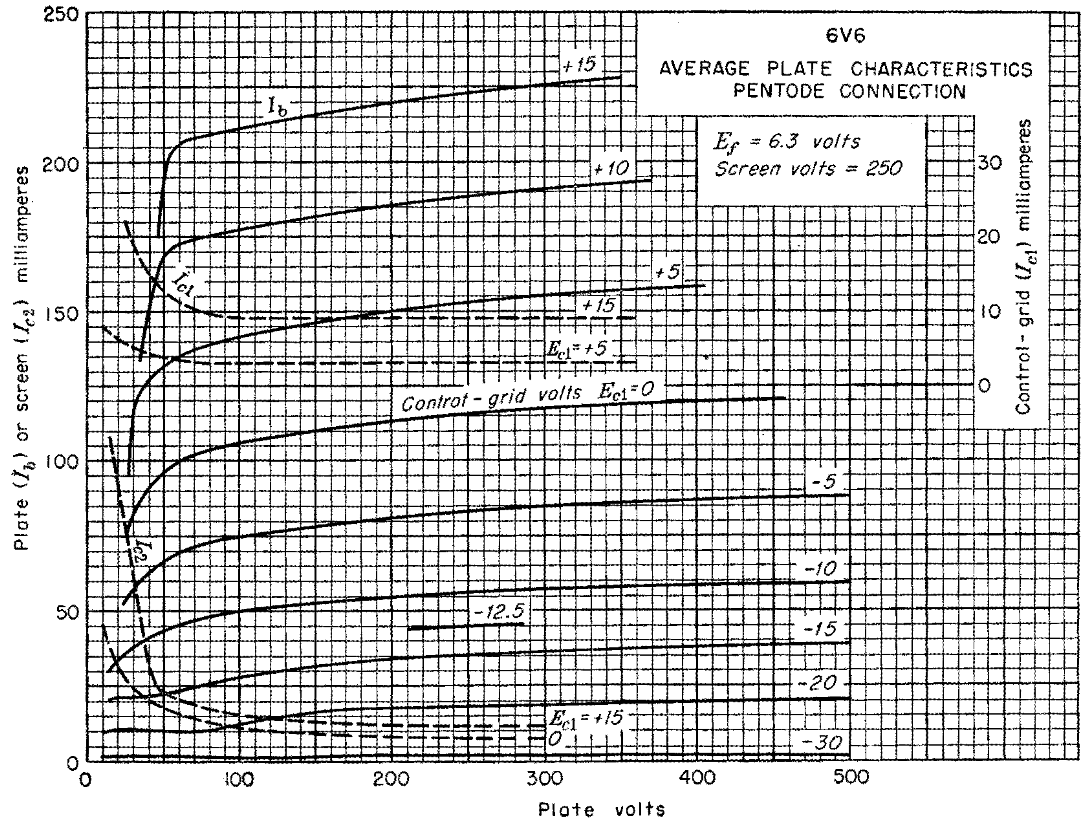


FIG. A-6. *(Courtesy of RCA.)*

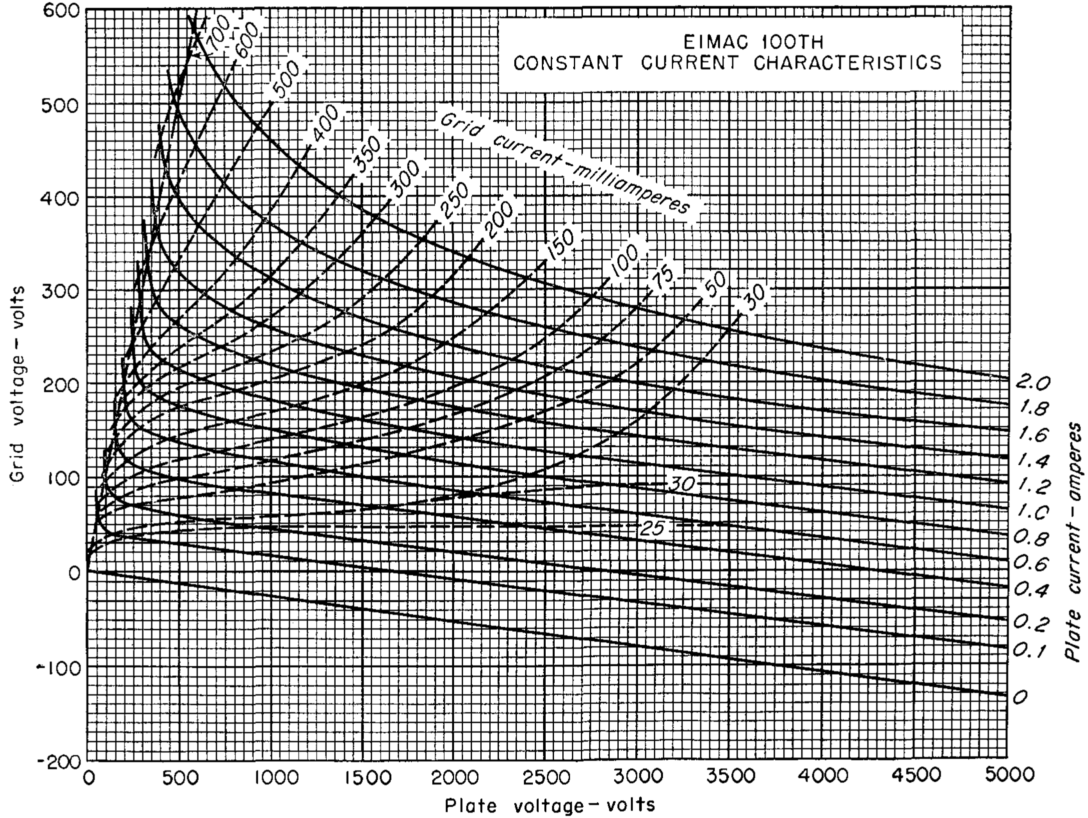


FIG. A-7. *(Courtesy of Eitel-McCullough, Inc.)*

APPENDIX B

R-C-COUPLED AMPLIFIER DESIGN CHART*

6J5, 6SN7, 7A4

E_{bb}, volts	100						250					
R_b, megohms	0.047		0.10		0.27		0.047		0.10		0.27	
R_g, megohms	0.1	0.27	0.10	0.47	0.27	0.47	0.10	0.27	0.10	0.47	0.27	0.47
R_k, ohms	1800	2200	3300	4700	8200	10000	1500	2200	2700	3900	6800	8200
I_{bo}, ma	1.05	0.97	0.57	0.50	0.24	0.22	2.79	2.40	1.49	1.31	0.61	0.58
E_{co}, volts	−1.9	−2.1	−1.9	−2.4	−1.9	−2.2	−4.2	−5.3	−4.0	−5.1	−4.2	−4.7
E_{bo}, volts	51	55	43	50	37	41	119	137	101	119	85	94
V_g, volts rms	0.5	0.5	0.5	0.5	0.5	0.5	1.0	1.0	1.0	1.0	1.0	1.0
E_o, volts rms	6.6	7.1	6.8	7.4	7.3	7.4	14.8	15.0	15.2	16.2	15.9	16.2
A	13.2	14.2	13.6	14.8	14.6	14.8	14.8	15.0	15.2	16.2	15.9	16.2
Dist., %	1.9	1.8	2.4	2.0	2.0	1.7	1.4	1.4	1.8	1.3	1.6	1.3
V_g, volts rms	0.95	1.13	0.95	1.30	0.95	1.20	2.70	3.50	2.55	3.30	2.64	3.05
E_o, volts rms	12.5	15.5	12.9	19.2	13.7	17.7	39.9	52.5	38.4	53.0	42.0	48.4
A	13.1	13.9	13.6	14.7	14.4	14.7	14.7	15.0	15.0	16.1	15.9	16.2
Dist., %	3.9	4.2	4.9	4.7	4.4	4.5	4.1	4.9	4.9	4.6	4.7	4.5

* Data in these charts are from "Technical Manual," Sylvania Electric Products Inc., Emporium, Pa., by permission.

R-C-COUPLED AMPLIFIER DESIGN CHART (*Continued*)
6SF5, 7B4

E_{bb}, volts	100							250						
R_b, megohms	0.10		0.27			0.47		0.10		0.27			0.47	
R_g, megohms	0.27	0.47	0.27	0.47	1.0	0.47	1.0	0.27	0.47	0.27	0.47	1.0	0.47	1.0
R_k, ohms	3900	3900	5600	5600	6800	8200	10000	1500	800	2700	2700	2700	3900	4700
I_{bo}, ma	0.22	0.22	0.14	0.14	0.13	0.10	0.09	0.84	0.76	0.44	0.44	0.44	0.30	0.27
E_{co}, volts	−0.9	−0.9	−0.8	−0.8	−0.9	−0.8	−0.9	−1.3	−1.4	−1.2	−1.2	−1.2	−1.2	−1.3
E_{bo}, volts	78	78	61	61	65	53	57	166	174	131	131	131	112	123
V_g, volts rms	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
E_o, volts rms	4.3	4.3	4.8	5.4	5.6	5.4	6.4	5.7	5.8	6.5	7.2	7.6	7.3	7.7
A	43	43	48	54	56	54	64	57	58	65	72	76	73	77
Dist., %	4.1	4.1	4.3	3.7	3.2	4.1	3.6	0.9	0.9	1.0	1.0	1.0	1.3	1.2
V_g, volts rms	0.12	0.12	0.10	0.10	0.13	0.10	0.15	0.47	0.54	0.39	0.39	0.39	0.33	0.45
E_o, volts rms	5.1	5.2	4.8	5.4	7.3	5.4	9.0	27	31	25	28	29	24	34
A	43	43	48	54	56	54	60	56	57	63	71	75	71	76
Dist., %	5.1	5.0	4.3	3.7	4.6	4.1	5.0	4.5	5.3	5.1	4.2	3.9	5.2	5.2

R-C-COUPLED AMPLIFIER DESIGN CHART (*Continued*)

6SJ7

E_{bb}, volts	100							250						
R_b, megohms	0.10		0.27			0.47		0.10		0.27			0.47	
R_d, megohms	0.39		1.20			1.80		0.39		1.20			2.20	
R_g, megohms	0.27	0.47	0.27	0.47	1.0	0.47	1.0	0.27	0.47	0.27	0.47	1.0	0.47	1.0
R_k, ohms	1200	1200	2700	2700	2700	4700	4700	560	560	1200	1200	1200	1800	1800
I_{bo}, ma	0.65	0.65	0.26	0.26	0.26	0.17	0.17	1.77	1.77	0.68	0.68	0.68	0.40	0.40
I_{c2o}, ma	0.18	0.18	0.07	0.07	0.07	0.05	0.05	0.50	0.50	0.18	0.18	0.18	0.10	0.10
E_{c1o}, volts	−1.0	−1.0	−0.9	−0.9	−0.9	−1.0	−1.0	−1.3	−1.3	−1.0	−1.0	−1.0	−0.9	−0.9
E_{c2o}, volts	30	30	19	18	18	19	19	55	55	31	31	31	26	26
E_{bo}, volts	36	36	30	30	30	23	23	73	73	68	68	68	61	61
V_g, volts rms	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
E_o, volts rms	6.9	7.8	8.2	10.2	12.5	10.2	13.1	10.2	11.5	13.6	17.9	21.6	19.5	25.6
A	69	78	82	102	125	102	131	102	115	136	179	216	195	256
Dist., %	0.6	0.7	3.4	2.6	2.3	2.8	3.2	0.7	0.8	2.2	1.8	1.5	3.1	2.4
V_g, volts rms	0.2	0.2	0.14	0.14	0.14	0.13	0.13	0.5	0.5	0.25	0.25	0.25	0.15	0.15
E_o, volts rms	13.2	14.9	11.1	13.9	17.2	12.8	16.6	47	54	33	42	50	28	37
A	65.8	74.5	79.4	99.5	123	98.5	128	94	108	132	168	200	187	247
Dist., %	3.0	2.9	5.1	4.3	3.7	4.6	5.0	4.2	5.0	5.2	4.4	4.7	4.5	3.7

APPENDIX C

A TABLE OF SYMBOLS*

Supply values	E_{ff}, E_{bb}, E_{cc}, E_{cc1}, E_{cc2}, etc.
Values for constant applied voltages	E_b, I_b, E_c, etc.
Instantaneous total values	e_b, i_b, e_c, i_c, e_L, e_{c2}, e_{2c}, etc.
Instantaneous incremental values	e_p, i_p, e_g, i_g, e_z, e_{g2}, e_{2g}, v_g, v_p, e_o, etc.
Effective values	E_b, I_b, E_p, I_p, E_g, I_g, E_{g2}, E_{2g}, V_g, V_p, E_o, etc.
Average values	E_{ba}, I_{ba}, I_{ca}, I_{c2a}, etc.
Quiescent values	E_{bo}, I_{bo}, E_{co}, I_{co}, I_{c2o}, etc.
Time-axis values	E_{bt}, I_{bt}, E_{ct}, etc.
Vector values	$\mathbf{E}_p$, $\mathbf{I}_p$, $\mathbf{E}_g$, $\mathbf{I}_g$, $\mathbf{V}_g$, $\mathbf{V}_p$, $\mathbf{E}_z$, $\mathbf{E}_o$, etc.
Peak values of sinusoidal quantities	E_{pm}, I_{pm}, I_{bm}, E_{gm}, I_{gm}, V_{gm}, V_{pm}, E_{zm}, E_{g2m}, I_{g_1m}, etc.
Special instantaneous values	$e_{b_{max}}$, $i_{b_{max}}$, $i_{b_{+max}}$, $i_{b_{-max}}$, $i_{1/2}$, $i_{-1/2}$, etc.
Harmonic components	i_{p_1}, i_{p_2}, I_{p_0}, I_{p_1}, I_{p_2}, I_{p_1m}, etc.
Applied signals	v_g, V_g, $\mathbf{V}_g$, v_p, V_p, $\mathbf{V}_p$, etc.

* Suffix numerals refer to the electrode number, for example, E_{cc2} is the direct supply voltage for the second grid.

Prefix numerals refer to the tube number, for example, e_{1c} is the total instantaneous grid voltage for the first tube in the circuit.

Subscript numeral refers to the order of the harmonic, for example, I_{p_2}.

Subscripts b and p refer to the plate.

Subscripts c and g refer to the grid.

Subscript k refers to the cathode or to a carrier.

Subscripts z and L refer to the load.

Subscript o refers either to a quiescent value or to the output.

For definitions of these symbols, including their positive sense, as well as the definitions of other symbols for power, resistance, reactance, capacitance, etc., refer to the text.

INDEX