PRINCIPLES OF ELECTRONICS

PHYSICAL SCIENCE TEXTS

General Editor
SIR GRAHAM SUTTON
C.B.E., D.Sc., F.R.S

Chairman of the National Environment Council and Director-General, Meteorological Office, Formerly Dean of the Royal Military College of Science, Shrivenham, and Bashforth Professor of Mathematical Physics.

Applied Mathematics for Engineers and Scientists
C. G. LAMBE,
B.A., Ph.D.

Techniques of Mathematical Analysis
C. J. TRANTER,
O.B.E., M.A., D.Sc.

A Compendium of Mathematics and Physics
DOROTHY MEYLER,
M.Sc.
and
O. G. SUTTON,
Kt., C.B.E., D.Sc., F.R.S.

Electron Physics and Technology
J. THOMSON,
D.Sc.
and
E. B. CALLICK,
B.Sc., A.M.I.E.E.

Principles of Electronics
M. R. GAVIN,
M.B.E., D.Sc.
and
J. E. HOULDIN,
Ph.D., F.Inst.P.

A Course in Applied Mathematics
(Covering B.A. and B.Sc. General Degrees)
D. F. LAWDEN,
M.A.

Physics for Electrical Engineers
W. P. JOLLY,
B.Sc., A.M.I.E.E.

Differential Equations for Engineers and Scientists
C. J. TRANTER,
O.B.E., M.A., D.Sc.
and
C. G. LAMBE,
B.A., Ph.D.

Elements of Dynamic Meteorology
A. H. GORDON,
M.Sc.

Applications of Nuclear Physics
J. H. FREMLIN,
M.A., Ph.D.(Camb.), D.Sc., F.Inst.P.

PRINCIPLES OF ELECTRONICS

M. R. GAVIN
M.B.E., D.Sc., F.Inst.P., M.I.E.E.
Professor of Electronic Engineering
University College of North Wales, Bangor

and

J. E. HOULDIN
B.Eng., Ph.D., F.Inst.P., A.M.I.E.E.
Senior Lecturer
Department of Physics, Chelsea College of Science and Technology

THE ENGLISH UNIVERSITIES PRESS LTD
ST. PAUL'S HOUSE WARWICK LANE
LONDON E.C.4

First printed 1959
Second impression 1960
Third impression 1961
Fourth impression 1963
Fifth impression 1965

*Printed in Great Britain for the English Universities Press, Limited,
by Richard Clay (The Chaucer Press), Ltd., Bungay, Suffolk*

GENERAL EDITOR'S FOREWORD

by

SIR GRAHAM SUTTON, C.B.E., D.Sc., F.R.S.

Chairman of the National Environment Council and Director-General, Meteorological Office, Formerly Dean of the Royal Military College of Science, Shrivenham, and Bashforth Professor of Mathematical Physics

THE present volume is one of a number planned to extend the Physical Science Texts beyond the Advanced or Scholarship levels of the General Certificate of Education. The earlier volumes in this series were prepared as texts for class teaching or self-study in the upper forms at school, or in the first year at the university or technical college. In this next stage, the treatment necessarily assumes a greater degree of maturity in the student than did the earlier volumes, but the emphasis is still on a strongly realistic approach aimed at giving the sincere reader technical proficiency in his chosen subject. The material has been carefully selected on a broad and reasonably comprehensive basis, with the object of ensuring that the student acquires a proper grasp of the essentials before he begins to read more specialized texts. At the same time, due regard has been paid to modern developments, and each volume is designed to give the reader an integrated account of a subject up to the level of an honours degree of any British or Commonwealth university, or the graduate membership of a professional institution.

A course of study in science may take one of two shapes. It may spread horizontally rather than vertically, with greater attention to the security of the foundations than to the level attained, or it may be deliberately designed to reach the heights by the quickest possible route. The tradition of scientific education in this country has been in favour of the former method, and despite the need to produce technologists quickly, I am convinced that the traditional policy is still the sounder. Experience shows that the student who has received a thorough unhurried training in the fundamentals reaches the stage of productive or original work very little, if at all, behind the man who has been persuaded to specialize at a much earlier stage, and in later life there is little doubt who is the better educated man. It is my hope that in these texts we have provided materials for a sound general education in the physical sciences, and that the student who works conscientiously through these books will face more specialized studies with complete confidence.

PREFACE

THE object of this book is to give a general introduction to the subject of electronics suitable for a first degree or diploma course in physics or electrical engineering. Emphasis is laid on the basic principles of operation of valves, transistors and other electron devices and of the circuits in which they are used. It is intended to provide the general common background which is essential to the physicist or the engineer prior to specialization in any particular branch of electronics. For most of the book the standard of mathematics required is no more than that of the Advanced Level of the General Certificate of Education.

After the introduction there are three chapters on the behaviour of free electrons and electrons in matter. These provide the physical background needed to explain the nature of the characteristics of the various types of vacuum and gas valves and transistors, which are the subject of the next two chapters. This concludes the study of electron devices. The following thirteen chapters are concerned primarily with the use of these devices in amplifiers, oscillators, rectifiers, switches, etc. In addition to the usual small signal analyses, problems are frequently discussed generally in terms of device characteristics and load lines based on circuit relations. The final chapter deals with the subject of noise. No attempt is made to cover specific applications of electronics such as radio, television, radar, computers, instrumentation, etc.

At the end of the book there is a collection of about 250 examples in groups corresponding to the chapters of the text. These examples form an important part of the book. Not only do they provide opportunity for testing the student's understanding but they are also used for further development and extension of principles. In some cases guidance is given to the solution. Many of the examples have been taken from recent examination papers of the Institution of Electrical Engineers and the Institute of Physics; the authors make grateful acknowledgment to these bodies.

Perhaps a note should be added regarding the use of symbols. In general these conform to the recommendations of the British Standards Institution. The small letters i_a, v_a, v_g, etc. are used, as is customary, to denote the instantaneous values of the varying components of the anode current, anode voltage, grid voltage, etc. Also, when these variations are sinusoidal, heavy type $\mathbf{I_a}$, $\mathbf{V_a}$, $\mathbf{V_g}$ is used to denote the corresponding complex or vector values. New symbols with capital suffixes, i_A, v_A and v_G, are introduced to represent total instantaneous values, including both steady and varying components. These symbols are used to represent the variables in static characteristics rather than I_a, V_a, and V_g. In this way the possibility is avoided of confusion, particularly in ordinary writing, between the capital letters which are used for both vector and total values.

In conclusion the authors wish to express their deep thanks to Mrs. Eveline Thorp who has typed the manuscript and to Mr. K. A. White who has drawn the diagrams.

CONTENTS

CHAPTER 6

CHAPTER 7

CHAPTER 8

CHAPTER 9

CHAPTER 10

CHAPTER 1

INTRODUCTION

1.1. Electronics

Electronics has been defined as " that field of science and engineering which deals with electron devices and their utilization ", where electron devices are " devices in which electrical conduction is principally by electrons moving through a vacuum, gas or semi-conductor ". The first six chapters of this book are concerned with the elucidation of this somewhat cryptic definition of electron devices. However, in these days of radio and television everyone is familiar to some extent with some of the devices such as valves, cathode-ray tubes and photo-electric cells. Yet at the beginning of the present century no electron devices as we know them existed. The nature of the electron itself, as a minute particle having mass and negative electric charge, had just been established. The Fleming diode, introduced in 1904, is usually considered to be the prototype electron device. As the name implies, this diode had two electrodes; one was a thin filament of wire (the cathode) which could be heated to incandescence by means of an electric current, and the other was a metal plate (the anode) close to the wire. The electrodes were enclosed in an evacuated glass envelope with wire leads through the glass. The diode acted as an electrical conductor when an e.m.f. was connected between the anode and the cathode in such a way that the anode was positive with respect to the cathode; when the polarity was reversed it acted as an insulator. This asymmetric or non-linear behaviour is typical to some extent of all electron devices. The one-way conduction in diodes is utilized in many ways for the rectification of alternating current.

In 1906 de Forest put a third electrode in the form of a wire grid between the cathode and anode. With this arrangement the current flowing to any electrode depends on the potentials of all three electrodes. It is found that under some conditions the grid potential acts as an effective control of the current to the anode without taking appreciable current itself; the grid controls large currents and power, without consuming much power. Thus three electrode valves or triodes can act as amplifiers of voltage, current or power. This ability to amplify opens up innumerable possibilities and is largely the reason for the importance of electronics to-day. After the triode, other multi-electrode valves with four, five and more electrodes were introduced. These give additional advantages of various kinds, but their operation generally depends on either their non-linearity or the amplifying property of a grid.

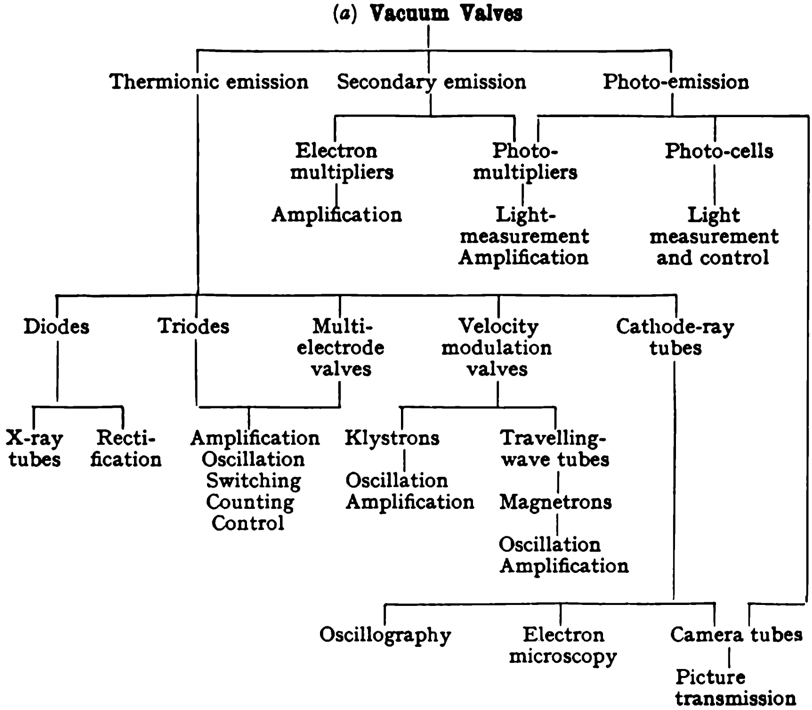
In the development of valves it was found that the presence of gas inside the envelope modified the properties considerably. There is a variety of gas diodes, triodes and multi-electrode valves.

From the early days of electron devices certain crystals, now called semi-conductors, were known to show non-linear conduction when used as diodes. In 1948 a major development occurred when it was found that the semi-conductor germanium could be used with three electrodes to give amplification. These new devices are called transistors. One electrode, the emitter, causes a flow of current to a second electrode, the collector. This current can be controlled readily by the potential difference between the emitter and the third electrode, the base, but the latter takes very little current. Thus, as in the triode, amplification can be obtained.

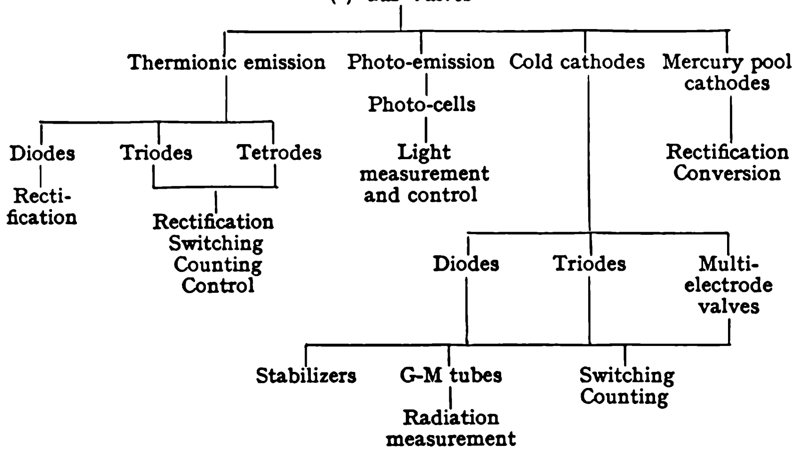
We have now had examples of electron devices with electrodes separated by vacuum, gas and semi-conductor. In all of these, as we shall see later, the currents are due almost entirely to the movement of electrons. Currents also flow in metals from the movement of electrons, but this current flow varies linearly with the potential difference across the metal, and such behaviour is of minor significance for electron devices.

For the conduction process there must be electron movement. Electron devices differ in the manner in which the electrons are made available. In some, the semi-conductors, electrons are available in the solid at

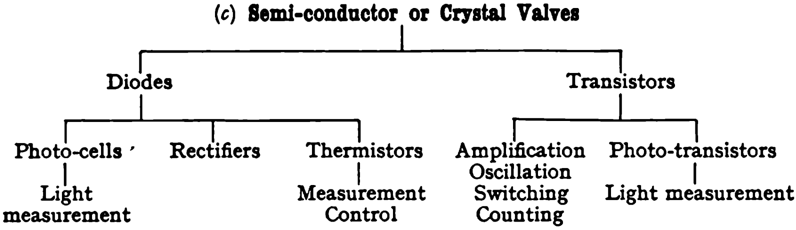
(*a*) **Vacuum Valves**



(*b*) **Gas Valves**



(*c*) **Semi-conductor or Crystal Valves**



ordinary temperatures. In vacuum tubes the electrons enter the vacuum from the cathode. The emission of the electrons from the cathode requires additional energy in some form. In thermionic emission the extra energy is given to the electrons by heating the cathode. The energy may also be supplied as radiation, when photo-emission occurs. Finally, the electrons may be knocked out of the cathode by impact with other fast electrons or ions. This is called secondary emission. Radiation may also be used to enhance the conductivity of semi-conductors. In gases, conduction electrons are produced by ionization of the gas atoms by means of radiation or fast particles. Positive ions are produced at the same time, and they may effect the conduction process, but their movement makes little contribution to the currents. In some gas valves electrons are also produced from a cathode by thermionic or photo-emission.

1.2. Electron Devices

The tables on pages 2 and 3 indicate some of the great variety of electron devices and their fields of use under three main headings: vacuum valves, gas valves and semi-conductor or crystal valves.

1.3. Diode Characteristics

The properties of any particular electron device can be described conveniently by means of characteristic curves, which show how the electrode currents vary with the potential differences between the electrodes.

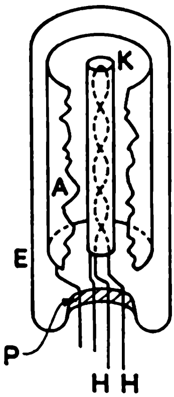


FIG. 1.1

One form of vacuum diode is shown diagrammatically in Fig. 1.1. The cathode K is a hollow cylinder which is heated indirectly by radiation from an insulated wire inside the cylinder. The heating current is supplied through the leads H. The anode A is another cylinder surrounding the cathode, and the whole system is enclosed in a glass envelope E, which is evacuated to a low pressure. The wire leads to the electrodes and the heater pass through a glass "pinch" P. The characteristic curves for such a diode are found using a circuit of the form shown in Fig. 1.2.*a*, which also shows the circuit symbol for the diode. The heater leads H are connected to a suitable electrical supply. Since there are only two electrodes in the diode, it follows that $i_A + i_K = 0$. Also the potentials may be measured from the cathode as zero. Thus there need

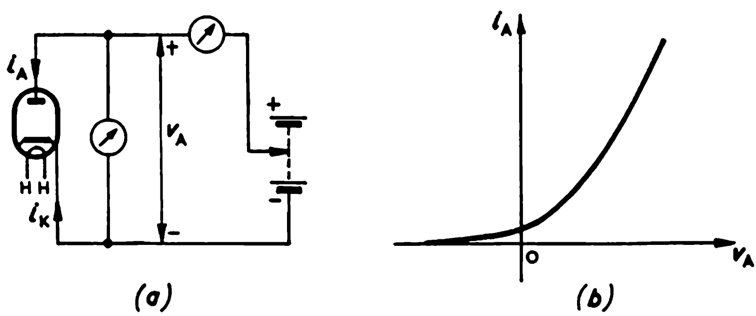


FIG. 1.2

be only two variables i_A and v_A, and the relation between them is shown in the characteristic curve in Fig. 1.2.*b*. It is seen that current flows when v_A is positive but little or no current when v_A is negative.

1.4. Triode Characteristics

A cylindrical thermionic triode is illustrated in Fig. 1.3. The grid is a wire cage surrounding the cathode. Now there are three electrode currents i_A, i_G and i_K and three potential differences v_{AK}, v_{GK} and v_{AG}, as shown in Fig. 1.4, which gives the circuit symbol for a triode. However, by Kirchhoff's Laws

$$i_A + i_G + i_K = 0$$

and

$$v_{AK} = v_{AG} + v_{GK},$$

and hence there are only four independent variables. If the potential differences or electrode voltages are measured from the cathode as v_G and v_A, then we may take i_G, i_A, v_G and v_A as the four variables. In many uses of the triode v_G is negative and no electrons flow to the grid, so that

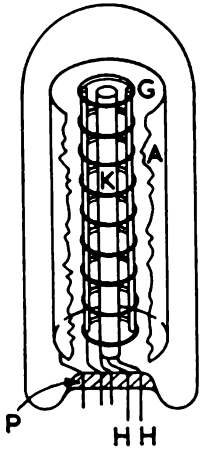


FIG. 1.3

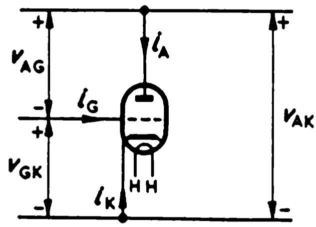


FIG. 1.4

$i_G = 0$. The relation between i_A, v_G and v_A may then be determined with the circuit of Fig. 1.5.*a*. The characteristics consist of the family of curves shown in Fig. 1.5.*b* or *c*. Each curve of the first set shows the

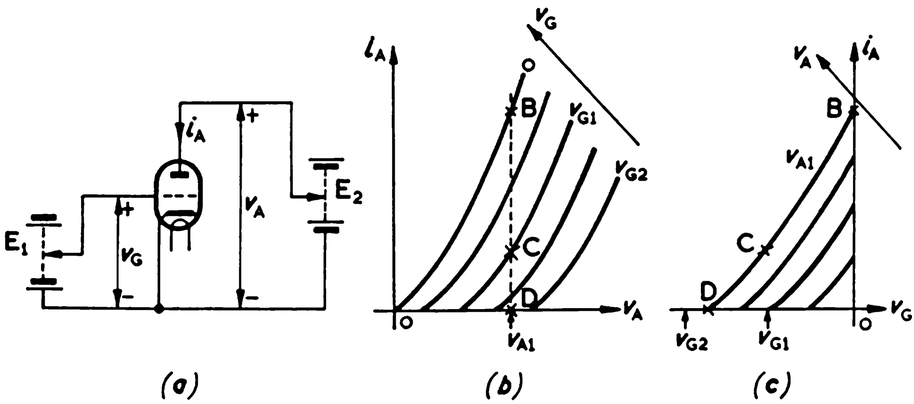


FIG. 1.5

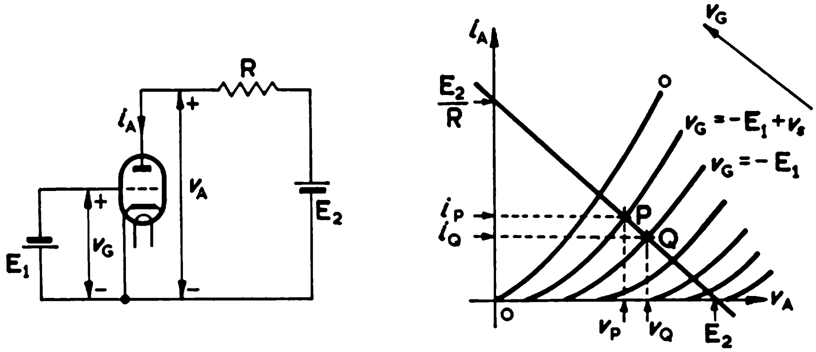
variation in anode current with anode voltage for a fixed value of grid voltage. Exactly the same information is given in different form in Fig. 1.5.*c*, where each curve shows the variation in anode current with grid voltage for a fixed value of anode voltage. Corresponding points are marked *B*, *C* and *D* in the two diagrams.

1.5. Triode Amplifier

In Fig. 1.6.*a* a resistance R is shown between the anode and the h.t. battery E_2. Now the electrode voltages are

$$v_G = -E_1 \text{ and } v_A = E_2 - Ri_A.$$

In the latter equation E_2 and R are constant, and the equation may be represented by a straight line in an i_A, v_A diagram as shown in Fig. 1.6.*b*. The straight line cuts the v_A-axis where $v_A = E_2$ and the



(a) (b)

FIG. 1.6

i_A-axis where $i_A = E_2/R$. Whatever the value of v_G, i_A and v_A must lie somewhere on this straight line. They must also lie on the valve characteristic curves which are drawn in the same diagram. Thus the actual values of i_A and v_A are found at the point Q on the characteristic corresponding to $v_G = -E_1$. If v_G changes by amount v_s, i_A and v_A must still lie on the straight line, and the new values are found at P. The change in anode current causes a change in voltage drop across R and a corresponding change in v_A. The anode voltage change is found from the diagram to be $v_P - v_Q$. If the magnitude of $v_P - v_Q$ is greater than v_s, then the valve has produced voltage amplification; the change v_s is frequently called a signal. The point Q which gives the initial or quiescent conditions before the change occurs is called the quiescent point or the operating point. The change in grid voltage may be produced by an alternator as shown in Fig. 1.7. Now the grid voltage and anode current vary sinusoidally and an alternating voltage appears across the resistance R and also between the anode and cathode.

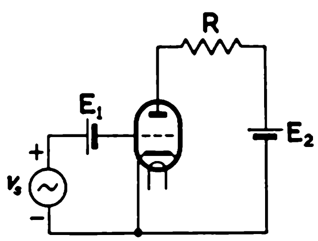


FIG. 1.7

1.6. Steady and Varying Values

In considering the currents and voltages in a valve there are several different conditions which have to be distinguished. Thus v_G and v_A are the actual values of the electrode voltages, however these arise. In the quiescent state

$$v_G = -E_1,$$
$$v_A = v_Q = E_2 - Ri_Q$$
and
$$i_A = i_Q,$$

where i_Q, and v_Q apply to the quiescent point. When the grid voltage is changed

$$v_G = -E_1 + v_s,$$
$$v_A = v_P = v_Q + (v_P - v_Q)$$
and
$$i_A = i_P = i_Q + (i_P - i_Q).$$

In each case there is a quiescent value, a change and a total value. Throughout this book we use capital suffixes to indicate total values (v_G, v_A and i_A) and small suffixes to indicate changes (v_g, v_a and i_a). In the case just considered

$$v_g = v_s,$$
$$v_a = v_P - v_Q$$
and
$$i_a = i_P - i_Q.$$

When the anode current increases the anode voltage decreases because of the greater voltage drop across R, and obviously

$$v_a = -Ri_a.$$
Also
$$v_G = -E_1 + v_g,$$
$$v_A = v_Q + v_a$$
and
$$i_A = i_Q + i_a.$$

In the absence of a signal the total values are equal to the steady or quiescent values.

When the signal is a small alternating voltage, say

$$v_s = \hat{v}_s \sin \omega t,$$
then
$$v_g = \hat{v}_s \sin \omega t = \hat{v}_g \sin \omega t,$$
$$i_a = \hat{i}_a \sin \omega t$$
and
$$v_a = -R\hat{i}_a \sin \omega t = -\hat{v}_a \sin \omega t.$$
Now
$$v_G = -E_1 + v_g = -E_1 + \hat{v}_g \sin \omega t,$$
$$i_A = i_Q + i_a = i_Q + \hat{i}_a \sin \omega t$$
and
$$v_A = v_Q + v_a = v_Q - \hat{v}_a \sin \omega t.$$

For sinusoidal alternating voltages and currents it is frequently convenient to use vector or complex values of these quantities. When this is done we use capital letters $\mathbf{V_a}$, $\mathbf{V_g}$, $\mathbf{I_a}$. Much confusion can arise in dealing with valves and valve circuits unless care is taken to distinguish carefully between quiescent values, total values, changes and vector values. The symbols suggested facilitate this distinction and are equally satisfactory in print or in writing. At the same time, the letters used indicate the nature of the quantity, voltage or current, and the electrode associated with it. The same notation may be easily extended to other devices. For example, i_E, i_e and $\mathbf{I}_e$ represent the total, change and vector currents respectively for the emitter of a transistor; i_R, i_r and $\mathbf{I}_r$ could represent the same currents in a resistor.

CHAPTER 2

ELECTRON MOTION

2.1. Electron Motion

Electrons may be considered as minute particles having mass and negative charge. As such, their movements in electric and magnetic fields can be determined by the application of the laws of electricity and magnetism and of mechanics. As the result of numerous experiments the following magnitudes have been established for the charge e and the mass m of an electron:

$$e = 1{\cdot}60 \times 10^{-19} \text{ coulomb}, \; m = 9{\cdot}11 \times 10^{-31} \text{ kilogram.}$$

This mass is $\frac{1}{1840}$ of the mass of a hydrogen atom. The above value for m is true only when the electrons are at rest or moving with velocities small compared with the velocity of light. According to the theory of relativity, the mass of a particle increases as its velocity approaches the velocity of light. The mass m at velocity u is related to the rest mass m_0 by the expression

$$m = \frac{m_0}{\sqrt{1 - \left(\frac{u}{c}\right)^2}},$$

where c is the velocity of light in free space. It may be verified from the Energy Equation (see Section 2.3) that the change in mass of an electron is negligible in devices operating below 1,000 V. The increase is about 1 per cent for electrons accelerated through a potential difference of 5,000 V. In most of this book the relativity correction is ignored.

2.2. Motion in a Steady Electric Field

The force F acting on a positive charge q in an electric field of strength E is, by definition,

$$F = qE$$

and the force is in the direction of the field. If v is the potential, then the field strength is related to the potential gradient by

$$E_s = -\frac{dv}{ds}.$$

This equation gives the component of the field strength in the direction of s. The component of the force in the same direction on the charge q is

$$F_s = -q\frac{dv}{ds}.$$

For an electron with negative charge the force is given by

$$F = e\frac{dv}{ds},$$

and is in the direction of increasing potential. If the electron is free to move, as it would be in a vacuum, the force does work on the electron. In moving between two points 1 and 2 the work done is

$$\int_1^2 F_s ds = ev_2 - ev_1,$$

where v_1 and v_2 are the potentials at the points 1 and 2 respectively. This work appears as kinetic energy of the electron. If u_1 and u_2 are the electron velocities at the two points, then

$$ev_2 - ev_1 = \frac{mu_2^2}{2} - \frac{mu_1^2}{2}.$$

This is to be known as the Energy Equation.

In the particular case where an electron starts from rest and moves through a potential difference v its final velocity u depends only on v, and is given by the relation

$$ev = \frac{mu^2}{2}.$$

The last two equations, which equate the gain in kinetic energy to the loss of potential energy, show how energy may be exchanged between electric fields and electrons, and they are of fundamental importance in the understanding of the operation of electronic devices. This energy exchange is considered in more detail in Chapter 14.

2.3. The Electron-volt

In dealing with electrons it is convenient to introduce a unit of energy called the electron-volt. This is defined as the amount of kinetic energy acquired by an electron in moving through a difference of potential of one volt. One electron-volt (eV) is equal to $1{\cdot}60 \times 10^{-19}$ coulombs $\times$ one volt $= 1{\cdot}60 \times 10^{-19}$ joules.

From the Energy Equation of Section 2.2 it follows that an electron starting from rest and moving through a potential difference of v V acquires a velocity given by $u = \sqrt{(2\,ev/m)} = 5{\cdot}94 \times 10^5\sqrt{v}$ m/s. If $v = 1$ V, then the velocity is $5{\cdot}94 \times 10^5$ m/s. An energy of 1 eV is extremely small, but it imparts an enormous velocity to an electron because of the small mass.

Although the electron-volt has been defined in terms of the electron, it may be used to measure any quantity of energy. For example, a proton, i.e., a hydrogen atom which has lost an electron, has a positive charge numerically equal to e and acquires a kinetic energy of 1 eV when it moves through a potential difference of 1 V. Since the proton has a much greater mass than the electron, an energy of 1 eV imparts to it a lower velocity. The Energy Equation, which applies to any charged particle, shows that the velocity varies inversely as $\sqrt{m}$, i.e., $\frac{1}{\sqrt{1840}}$ or $\frac{1}{43}$ in this case. Sometimes the " electron " is omitted from " electron-volt ". For example, a beam of 100 V electrons means a beam in which the electrons have started from rest and moved through a potential difference of 100 V. The energy of each electron is 100 eV or $1{\cdot}60 \times 10^{-17}$ joules.

2.4. Electric Fields

The user of an electron device controls the potentials of the electrodes by making suitable connections to the electrode leads. However, if the behaviour of electrons inside the device is to be studied the potential distribution in the space between the electrodes must be known. In the absence of appreciable charge in the space the potential must satisfy Laplace's Equation, which, in cartesian co-ordinates, is

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} = 0.$$

When the charge density ρ in the space is not negligible the potential obeys Poisson's Equation

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} = -\frac{\rho}{\varepsilon_0},$$

where ε_0 is the primary electric constant having the value $8{\cdot}86 \times 10^{-12}$F/m; it is sometimes called the permittivity of free space. Laplace's and Poisson's Equations follow directly from Gauss's Theorem. Thus the problem of finding the potential consists of solving one of these two partial differential equations using the electrode potentials as boundary conditions. For all but a few simple geometrical arrangements of the electrodes analytical solutions are impossible. Methods have been developed for dealing with special cases, and approximate numerical solutions of Laplace's Equation can be obtained experimentally by the use of analogues, particularly the rubber membrane, the electrolytic tank and the resistance network. These special methods are beyond the scope of this book, and we confine our attention to simple idealized cases which can be readily analysed. The results give some qualitative indication of what may be expected in the more complicated practical cases.

2.5. Electron Motion in a Uniform Electric Field

A uniform electric field, i.e., one in which the field strength is constant, can be realized between the plates of a parallel-plane condenser in regions remote from the edges, provided the dimensions of the planes are large compared with the distance between them. This case is illustrated in

Fig. 2.1, in which the y-axis is chosen at right angles to the planes and the x- and z-axes are parallel to the planes. The potential difference between the two electrodes, the cathode and the anode, is v_A. Obviously, except near the edges, the potential in the space between the electrodes

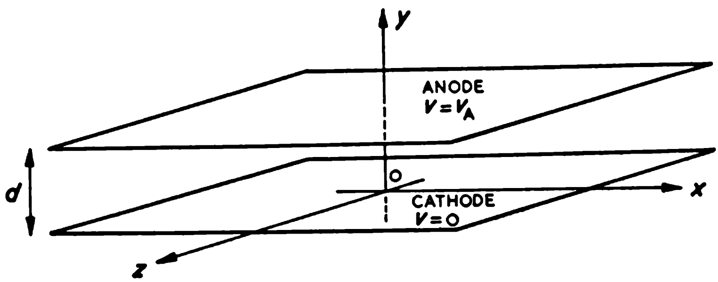


FIG. 2.1

does not vary in the x and z directions, and Laplace's Equation becomes one dimensional, i.e.,

$$\frac{d^2v}{dy^2} = 0.$$

Integration gives

$$\frac{dv}{dy} = A \text{ and } v = Ay + B,$$

where A and B are constants to be determined by the boundary conditions. The difference of potential between $y = 0$ and $y = d$ is v_A; $d =$ the distance between the planes. It follows that the field strength and the potential are given by

$$\frac{dv}{dy} = \frac{v_A}{d} \text{ and } v = \frac{v_A y}{d}.$$

These expressions confirm that the field is uniform. If there is an electron in the space it is acted upon by a constant force F given by

$$F = e\frac{dv}{dy} = \frac{ev_A}{d}.$$

If the space is evacuated the electron is free to move continuously under the influence of this force. In air the electron would still move, but its motion would be interrupted by impact with the air molecules. It is assumed below that the space is evacuated. Using Newton's Second Law of Motion

$$F = \frac{d}{dt}(mu) = m\frac{d^2y}{dt^2} = \frac{ev_A}{d}.$$

Hence

$$\frac{d^2y}{dt^2} = \frac{ev_A}{md},$$

and the acceleration is also uniform. By integration the velocity is found to be

$$u = \frac{dy}{dt} = \frac{ev_A t}{md},$$

where the constant of integration has been put equal to zero on the assumption that the electron left the cathode with zero velocity at time $t = 0$. The distance moved in time t is found by a further integration to be

$$y = \frac{ev_A t^2}{2md}.$$

The constant of integration is again zero. The time to travel a distance dy is equal to dy/u, and hence the time τ for the electron to move across the space is found from

$$\tau = \int_0^d \frac{dy}{u}.$$

From the Energy Equation and $v = v_A y/d$, it is found on integrating that

$$\tau = \frac{2d}{u_A} = \frac{2d}{(5{\cdot}94 \times 10^5\sqrt{v_A})} \text{ sec},$$

where u_A is the velocity of the electron at the anode. This expression could have been found by alternative reasoning. Since the motion is uniformly accelerated, starting from rest and finishing with velocity u_A, the average velocity is $u_A/2$. The transit time is the distance divided by the average velocity, i.e., $2d/u_A$.

2.6. Cathode-ray Tube with Electrostatic Deflection

Everyone is familiar with cathode-ray tubes for displaying electrical quantities visually. The name " cathode ray " is a relic of the early days of electron physics, when streams of electrons emanating from the cathode of a discharge tube were called cathode rays. One common form of cathode-ray tube is shown in the diagram of Fig. 2.2. A narrow beam of electrons is produced by an electron " gun " and passes between two pairs of deflecting plates. The beam finally strikes a luminescent screen and produces a bright spot of light. Details of the electron gun are considered later in this chapter. If the difference in potential between the cathode and the final anode is 1,000 V, then a fast beam of electrons leaves the final anode of the gun with the electron velocities corresponding to 1,000 eV. The deflecting plates are arranged parallel to the beam. Normally, the first pair, or Y-plates, are horizontal and the second pair, or X-plates, are vertical. The mean potential of the deflecting plates is also 1,000 V and the beam passes through the middle of the deflecting system and strikes the centre of the screen. The whole of the space from the final anode of the gun to the screen is at a potential of approximately 1,000 V, and the electrons move with constant velocity in this space. If a potential difference is applied to the Y-plates this produces

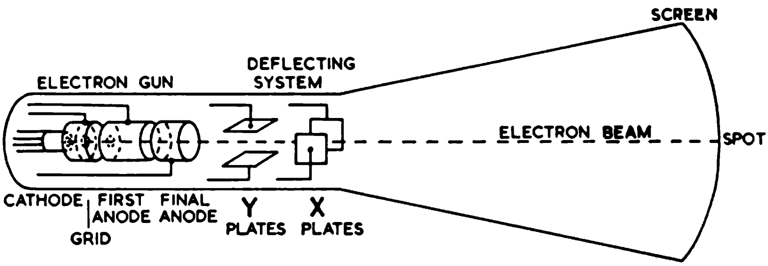


FIG. 2.2

a vertical deflection of the beam and the spot is displaced on the screen. Similarly, a potential difference across the X-plates causes a horizontal displacement of the spot. The size of these deflections can be determined approximately by using the results of the foregoing paragraphs. The passage of an electron beam through one pair of deflecting plates is represented in Fig. 2.3. The electrons reach O with horizontal velocity u_A, where $u_A = \sqrt{(2ev_A/m)}$, v_A being the potential of the final anode and also the mean potential of the deflecting plates and the screen. When

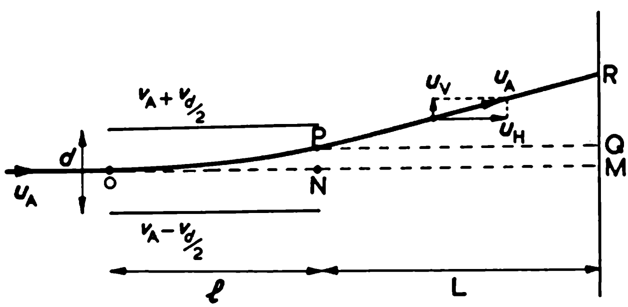


FIG. 2.3

there is no deflecting field the electrons move with constant velocity u_A along the path ONM. The deflecting plates are similar to the parallel-plane condenser of the previous section, and when a potential difference is maintained across the plates the electric field is uniform except near the edges of the plates. In order to simplify the calculations it is assumed that the deflecting field is uniform over the length l of the plates and is zero elsewhere. Under these conditions the electrons reach O with velocity u_A as before. Since the deflecting field is purely vertical, the horizontal velocity is unaffected. However, over the length l the deflecting field acts on the electrons so that, at the end of the deflecting system, they have acquired a vertical displacement NP and a vertical component of velocity, say u_V. The value of u_V can be found by the methods of the previous section, and is given by

$$u_V = ev_d t/md$$

where v_d = deflecting potential difference, d = distance between the deflecting plates and t = time spent in the deflecting plates. Obviously

$$t = l/u_A \text{ and } u_V = ev_d l/mdu_A.$$

From P onwards the electrons are free from field and move with velocity corresponding to potential v_A. However, they have now a vertical component of velocity equal to u_V and at the screen they have a vertical deflection QR equal to $u_V T$, where T is the time to travel from the deflecting plates to the screen. As long as

$$u_V \ll u_A,$$
$$T \simeq L/u_A,$$

where L = horizontal distance from the deflecting plates to the screen. Hence

$$QR = ev_d lL/mdu_A^2,$$

i.e.,

$$QR = lLv_d/2dv_A.$$

The total vertical deflection of the spot on the screen from the mean position should include NP. However, in most cathode-ray tubes $l \ll L$ and NP is negligible. Thus the deflection y of the spot is given approximately by

$$y = lLv_d/2dv_A.$$

The deflection sensitivity of the tube is defined as y/v_d. It is frequently quoted in millimetres per volt. The sensitivity is inversely proportional to v_A. However, the brightness of the spot and the sharpness of its focus increase with v_A, so that a compromise is necessary.

In deriving the formula for y several assumptions have been made. In particular, the deflecting field has been assumed to be confined to the length of the deflecting plates. This field must extend beyond these plates, and as a result deflection occurs over a length greater than l. Also, the horizontal component of the velocity over the length L is less than u_A. In this region, away from the deflecting plates, the potential is v_A. The Energy Equation relates the loss of potential energy to the total kinetic energy so that

$$ev_A = \frac{mu_H^2}{2} + \frac{mu_V^2}{2},$$

where u_H and v_V are used for the horizontal and vertical components. This equation shows that u_H must be less than u_A. However, u_V is frequently much less than u_H, and then u_H and u_A are nearly equal.

In some cathode-ray tubes the deflecting plates are flared as shown in

Fig. 2.4.*b*. With this arrangement the value of d varies along the plates, and obviously it gives greater sensitivity than parallel plates with a separation equal to the final value of d in the flared plates. The maximum deflection would be the same for both systems.

Mention was made in Section 2.3 of the enormous velocities which can be given to electrons by electric fields. The mass of the electron is so

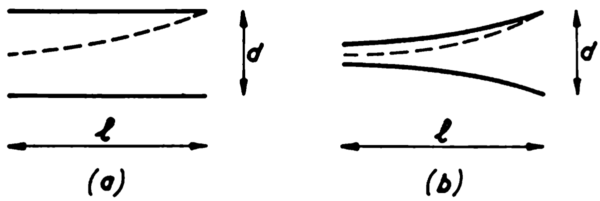


FIG. 2.4

small that its response to a change of field is practically instantaneous and the deflection takes very little energy from the field. In addition, the electrons move through the deflecting plates very rapidly. For 1,000 V electrons and deflecting plates 2 cm long the time taken is about 10^{-9} sec. Thus even for an electric field varying at a frequency of 10^6 c/s the field does not change appreciably during the passage of the electrons, and the formula for y may be used to give the instantaneous deflection.

2.7. Motion in a Uniform Magnetic Field

When an electron moves with velocity u in a magnetic field of flux density B the electron experiences a force F which is given by

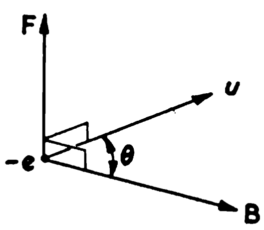


FIG. 2.5

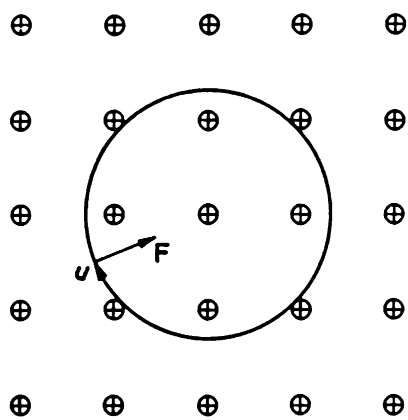


FIG. 2.6

$F = Beu \sin \theta$, where θ is the angle between the directions of B and u. This force is perpendicular to the plane containing B and u as shown in Fig. 2.5. The following interesting points may be established from the equation for F: (i) there is no force on a stationary electron in a magnetic field, (ii) the force is greatest when the electron moves at right angles to the magnetic field, (iii) there is no force on an electron moving parallel to a magnetic field, (iv) as the force is perpendicular to the velocity, an electron cannot gain kinetic energy from a magnetic field; the magnetic field may alter the direction but not the magnitude of the velocity.

In the particular case when the field and the velocity are at right angles the force is Beu and is perpendicular to both B and u. This case is represented in Fig. 2.6, where the magnetic field is uniform and is

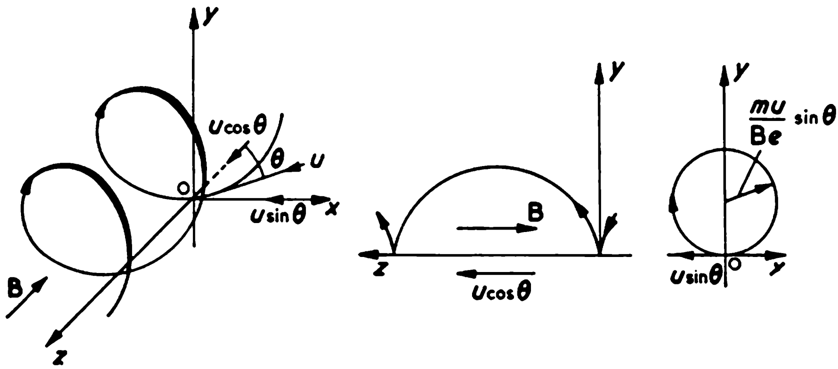


FIG. 2.7

directed into the paper. Here we have an electron moving with velocity u and subjected to a force Beu in a direction at right angles to u. This is a case of uniform motion in a circle. For such a motion the force towards the centre of the circle is mu^2/r, where r is the radius. Hence

$$F = Beu = mu^2/r \text{ and } r = mu/Be.$$

The time t for one complete circuit is $2\pi r/u$,

i.e., $$t = 2\pi m/Be$$

The time is thus independent of u and r; it depends only on the magnetic field. It may also be seen that the radius of the path varies with u. These results have importance in many devices, such as the cyclotron, the mass spectrograph, magnetic lenses, etc.

When the electron velocity makes an angle θ with the magnetic field the path is helical, as shown in Fig. 2.7. The velocity may be resolved into two components $u \cos \theta$ parallel to the field and $u \sin \theta$ perpendicular to the field. There is no force due to the former, and the latter gives rise to a circular motion of radius $\frac{mu}{Be} \sin \theta$. The resultant motion is due to this circular motion round the direction of the field and the uniform translational motion $u \cos \theta$ along the field. The radius of the helix may be controlled by varying B.

2.8. Motion in Crossed Electric and Magnetic Fields—The Magnetron

An important case of electron motion occurs in a combination of uniform electric and magnetic fields. This case is illustrated in Fig. 2.8, where a difference of potential is maintained between two parallel plates and there is a uniform magnetic field parallel to the plates. An arrangement of this type is called a planar magnetron. If a single electron starts

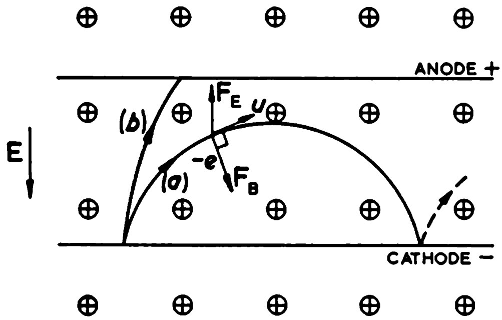


FIG. 2.8

from the cathode with zero velocity it experiences a force due to the electric field, but the magnetic field has no effect. The force due to the electric field gives an acceleration towards the anode. As a result, the electron moves towards the anode with increasing velocity. Now, because of its velocity, there is a force on it due to the magnetic field, and this force, which is always at right angles to the motion, increases as the velocity increases. The path, therefore, is curved. If the magnetic

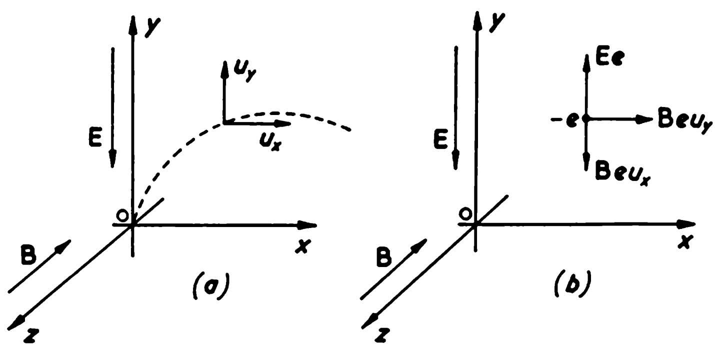


FIG. 2.9

field is sufficiently great the electron may be turned back towards the cathode as shown in curve (*a*), Fig. 2.8. Energy considerations show that at every point of the path the electron velocity depends only on the electric potential at that point, and after the electron is turned back its velocity decreases. As a result, the force due to the magnetic field decreases, until at the cathode the electron comes to rest again, and there is no force from the magnetic field. The electron is ready to start on another similar path as shown by the dotted line. For a weaker magnetic field the curvature of the path is less and the electron may move to the anode without being turned back as shown in Fig. 2.8, curve (*b*). The actual shape of the electron path is determined as follows. The problem is two-dimensional, and the co-ordinate axes are chosen as in Fig. 2.9. The electron's velocity at any point is resolved into two components u_x and u_y parallel to the axes. These two components, along with the magnetic field, give rise to forces Beu_y and $-Beu_x$ in the x- and y-directions respectively. Due to the electric field E there is a constant force Ee in the y-direction. From Newton's Second Law the equations of motion for the x- and y-directions are then

$$m\frac{d^2x}{dt^2} = Beu_y \quad \text{and} \quad m\frac{d^2y}{dt^2} = Ee - Beu_x,$$

i.e.,

$$\frac{du_x}{dt} = \omega u_y \quad \text{and} \quad \frac{du_y}{dt} = \frac{Ee}{m} - \omega u_x,$$

where $\omega = Be/m$. By differentiating the second of these equations and substituting from the first it follows that

$$\frac{d^2u_y}{dt^2} = -\omega^2 u_y.$$

The solution of this familiar differential equation is

$$u_y = a\cos\omega t + b\sin\omega t,$$

where a and b are constants to be determined. When the electron starts from the cathode,

$$t = 0,\ u_y = 0,\ \text{and}\ \frac{du_y}{dt} = Ee/m$$

Using these conditions, it is found that

$$a = 0 \text{ and } b = Ee/\omega m.$$

Hence

$$u_y = \frac{Ee}{\omega m}\sin\omega t.$$

Now by substitution for u_y in the original differential equation it follows that

$$u_x = \frac{Ee}{\omega m}(1 - \cos\omega t).$$

By integrating these two equations for u_x and u_y, and using the condition that $x = 0$ and $y = 0$ at $t = 0$, it is found that the position of the electron at time t after leaving the cathode is given by

$$x = \frac{Ee}{\omega^2 m}(\omega t - \sin\omega t)$$

and

$$y = \frac{Ee}{\omega^2 m}(1 - \cos\omega t).$$

These two equations represent a cycloid of the form shown in Fig. 2.8. The greatest distance reached from the cathode occurs when y has its maximum value. This obviously is when $\cos \omega t = -1$ and the actual distance is

$$y_{max} = 2Ee/\omega^2 m = 2Em/B^2 e.$$

If this distance is greater than d, the anode–cathode distance, then the electron is collected by the anode. The limiting case, where the electron just reaches the anode, occurs when $d = 2Em/B^2 e$. If v is the potential difference between the anode and cathode, then

$$E = v/d \text{ and } B = \sqrt{(2vm/ed^2)}.$$

This value of B is known as the critical magnetic field. For greater

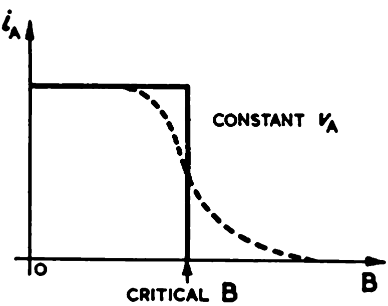


FIG. 2.10

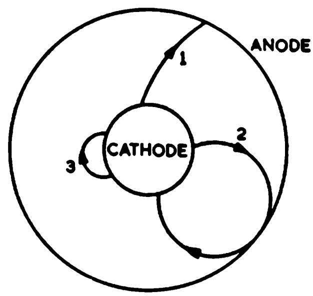


FIG. 2.11

values of B no electrons reach the anode, and for smaller values all the electrons which leave the cathode are collected by the anode. If the anode current of a planar magnetron is measured as the magnetic field is varied the current should vary as shown by the full line in Fig. 2.10. In practice, the current cut-off is more indefinite, as shown by the broken lines in the figure. Some of the factors contributing to this indefinite cut-off could be: (i) end effects of the finite electrodes, (ii) the electrons do not leave the cathode with zero velocity but have a velocity distribution (see Chapter 4), (iii) electron interaction, (iv) the magnetic field may not be uniform throughout.

Most magnetrons in use have the electrodes as co-axial cylinders with the magnetic field parallel to the axis. The exact analysis of this case is difficult, but the path of a single electron is rather similar to that of the planar magnetron. When the radius of the cathode is not much less than that of the anode, then the electron path is nearly cycloidal. When the cathode radius is much less than the anode radius the path is very nearly circular (see Exx. II). Typical electron paths for a cylindrical magnetron are shown in Fig. 2.11. Here the numbers 1, 2 and 3 correspond to increasing values of the magnetic field with a fixed anode voltage. Curve 2 occurs at the critical value of the magnetic field. These paths could also be obtained with fixed magnetic field and decreasing anode voltage.

In all these considerations of magnetrons so far the effect of space charge has been neglected, i.e., attention has been given to the behaviour of a single electron and any forces due to other electrons have been ignored. Space-charge effects complicate the problem enormously. They are considered again in Section 5.8.

2.9. Cathode-ray Tube with Magnetic Deflection

The deflection of an electron in a magnetic field is sometimes used in cathode-ray tubes as shown in the diagram of Fig. 2.12. Two pairs of coils XX' and YY' are arranged outside the tube. The XX' coils are

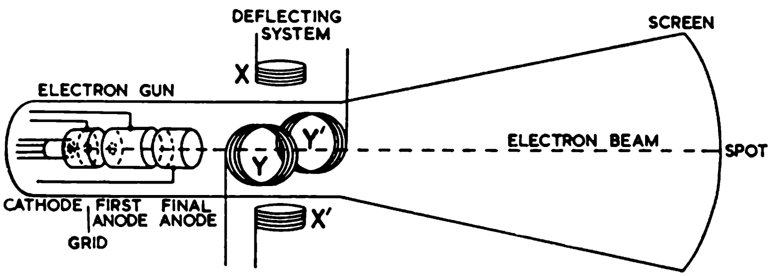


FIG. 2.12

normally placed horizontally on either side of the tube, and they are joined in series. When a current is passed through these coils a vertical magnetic field is produced inside the tube. This field, which is pro-

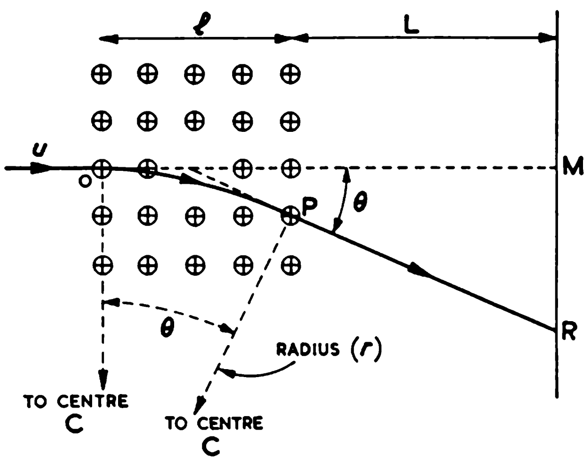


FIG. 2.13

portional to the coil current, causes a horizontal force on the electron beam and produces a horizontal displacement of the spot. The other pair of coils YY' give vertical displacement.

The deflection sensitivity of this type of cathode-ray tube is determined by making simplifying assumptions similar to those made for the electrostatic case. Over the length l of the coils it is assumed that the magnetic field is uniform and elsewhere zero (see Fig. 2.13). If the electron arrives at O with velocity u it describes an arc OP of a circle of radius r in the magnetic field. On leaving the field, it moves in a straight line with uniform velocity until it strikes the screen at R. The distance L from the coils to the screen is assumed to be much greater than l. With this condition $y/L \simeq \tan\theta$, where y is the deflection of the spot and θ is the angle between the horizontal line OM and the tangent to the circular arc at P. If C is the centre of the circle of which OP is an arc, then angle OCP is also equal to θ, and $\theta = \text{arc } OP/r$. The angle θ is usually small and $\tan\theta \simeq \theta$ and arc $OP \simeq l$. Thus $y = Ll/r$. But it has already been shown in Section 2.7 that $r = mu/Be$, and hence

$$y = LleB/mu$$

gives the deflection produced by the field B, i.e., the deflection is proportional to the coil current. The velocity u is found from the gun voltage and the Energy Equation.

2.10. Electron Optics

Electron optics is that branch of electronics which deals with the production of beams of electrons such as are used in cathode-ray tubes. The motion of the electrons is controlled by means of electric or magnetic fields, and the beams may be brought to a focus by suitable adjustment of the fields. The whole process has much in common with the control of light beams by means of apertures and lenses, as studied in geometrical optics. There are " electron lenses " of various types, and these have numerous applications, the most spectacular of which is the electron microscope, which can give a high magnification with greater resolution than the optical microscope. Some simple types of electron lens are considered in the next sections.

2.11. Magnetic Lens

The helical path of an electron moving in a uniform magnetic field is considered in Section 2.7, and this is the basis of magnetic lenses. In Fig. 2.14.a a cathode acts as a source of electrons which are accelerated by means of a high voltage anode with a small aperture. The electrons passing through the aperture form a diverging beam. An axial magnetic field B is provided by means of a solenoid. The electrons which pass normally through the aperture have no motion at right angles to B, and their motion is unaffected. A diverging electron, making an angle θ with the axis, describes a helix of radius $\dfrac{mu}{Be}\sin\theta$, where u is the velocity of the electrons at the aperture. As shown in Section 2.7, all the electrons complete one helical " pitch " in the same time, equal to $2\pi m/Be$. The distance travelled along the axis is then $\frac{2\pi mu}{Be}\cos\theta$, since $u\cos\theta$ is the axial component of the velocity. If θ is small, $\cos\theta$ is nearly unity, and hence all the electrons have approximately the same relative positions at I, at a distance $2\pi mu/Be$ along the axis. The paths of several electrons are shown in Fig. 2.14.*a*, and it is seen that they come together at I, where an image of the aperture is formed. This type of magnetic lens gives unit magnification.

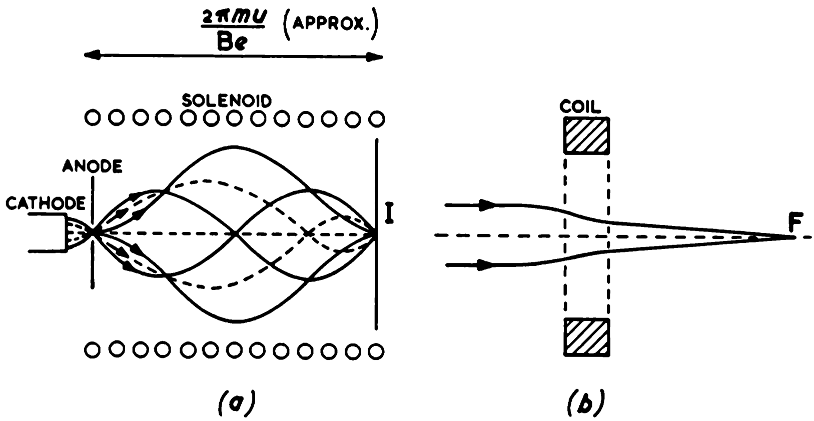


FIG. 2.14

A second type of magnetic lens uses a non-uniform magnetic field such as is formed by a short coil, Fig. 2.14.*b*. We consider an electron travelling from the left parallel to the axis but above it, and entering the field of the coil. There is an inward radial component of the field, and this gives the electron a velocity component sideways out of the plane of the paper. This velocity component is normal to the axial magnetic field, and gives rise to a component of force towards the axis. When the electron passes through the coil it is in a reversed radial field which gradually reduces the sideways component of the velocity. As long as there is any sideways velocity there is a force on the electron towards the axis. The electron therefore emerges from the lens with a velocity directed to the axis, and it ultimately intersects the axis at some point F. Since we started with a parallel ray, F is the focus of the lens.

2.12. Electrostatic Electron Optics

The force on an electron in an electric field is in the direction of increasing potential. Now consider an electron moving with velocity u_1 in a region where the potential is v_1 and which is separated by a small planar gap from a region where the potential is v_2 (see Fig. 2.15). There is no force on the electron except in the gap, where the force is normal to the boundary. After crossing the gap the electron moves with velocity u_2. There is no change in the electron velocity parallel to the gap, and hence

$$u_1 \sin i = u_2 \sin r,$$

i.e.,

$$u_2/u_1 = \sin i/\sin r,$$

where i and r are the angles between the beam and the normal to the boundary as shown. Using the Energy Equation, it follows that

$$\sin i/\sin r = \sqrt{v_2}/\sqrt{v_1}.$$

In the refraction of a light ray at the boundary of two media there is a similar formula of the form

$$\sin i/\sin r = n_2/n_1,$$

where n_1 and n_2 are the refractive indices of the two media. Thus $\sqrt{v}$

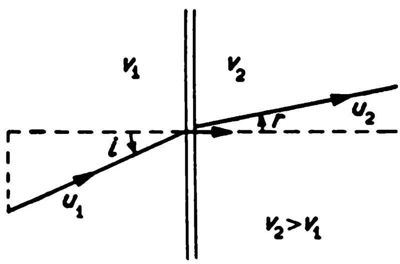


FIG. 2.15

in electron optics is analogous to n in geometrical optics. In electrostatic fields the potential does not change discontinuously, so that the strict optical analogy would require a medium with continuously variable refractive index.

2.13. Electrostatic Lenses

Fig. 2.16 represents a section through two equal co-axial cylinders separated by a small gap and maintained at potentials v_1 and v_2. The

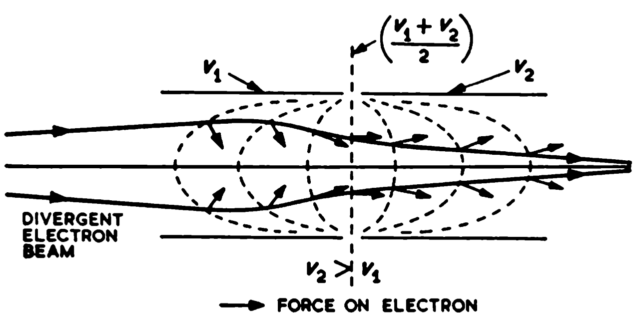


FIG. 2.16

sections through some surfaces of equal potential are also shown. Three representative rays of a divergent electron beam are shown entering the field. The force acting on an electron is in the direction of increasing potential and is perpendicular to the equipotential surfaces as shown by the small arrows. Thus up to the gap the electron paths converge towards the axis, and after the gap they diverge again. Although the field is symmetrical about the gap, the divergence is less than the convergence, since the electrons are moving with slightly greater velocity after the gap. In addition, the force towards the axis increases with the distance from

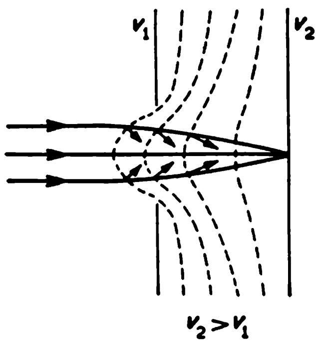


FIG. 2.17

the axis. This system therefore acts as a converging lens, and the "focal length" may be altered by varying the potential difference between the cylinders. Electrostatic cylindrical lenses are sometimes constructed with cylinders of unequal radius.

Various electrostatic lenses can be formed by means of apertures in diaphragms. The nature of the lens depends on the potentials of the diaphragm and the regions on either side of it. An aperture lens is shown

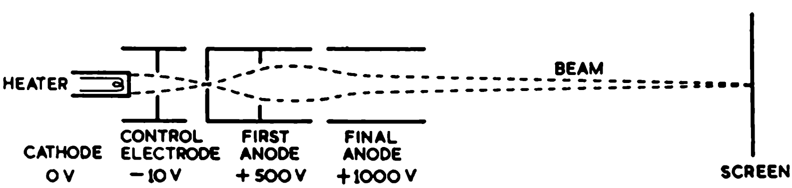


FIG. 2.18

in Fig. 2.17. An aperture also serves the purpose of limiting the size of the beam which passes through.

The electron gun in a cathode-ray tube is usually a combination of aperture and cylindrical lenses. An example is shown in Fig. 2.18, where there is an aperture lens and a cylindrical lens. The intensity of the beam or " brightness " is adjusted by varying the voltage of the control electrode, and the focus is controlled by means of the first anode voltage.

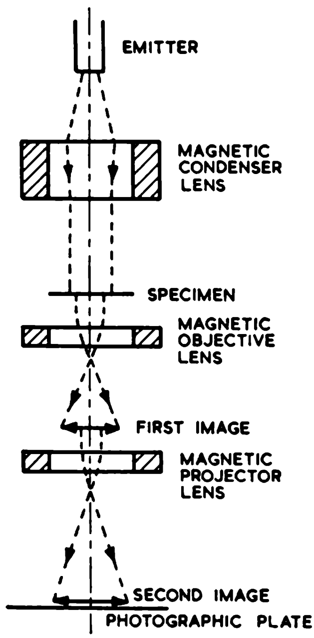


FIG. 2.19

In the electron microscope either magnetic or electrostatic lenses may be used, and the lenses correspond to the condenser, objective and projector of an optical projection microscope. One system is illustrated in Fig. 2.19.

CHAPTER 3

ELECTRONS IN MATTER

3.1. Electrons in Matter

In Chapter 2 the free electron is considered as a small negative charge having a certain mass. In order to explain some observed properties of electrons it is necessary to assume that an electron also has a magnetic moment, which may be considered as due to the spinning of the electron charge about an axis. This spin is significant only in a magnetic field, and the axis of spin has meaning only with reference to the direction of the field. The spin may be in one of two possible directions relative to the axis. Electron spin does not concern us to any great extent in electronics. It is introduced briefly in the present chapter, in which we consider the energies of electrons in association with matter.

3.2. Electrons in Atoms—Energy Levels

According to modern ideas, an atom of an element consists of a positive charge or nucleus surrounded by one or more negative charges or electrons.

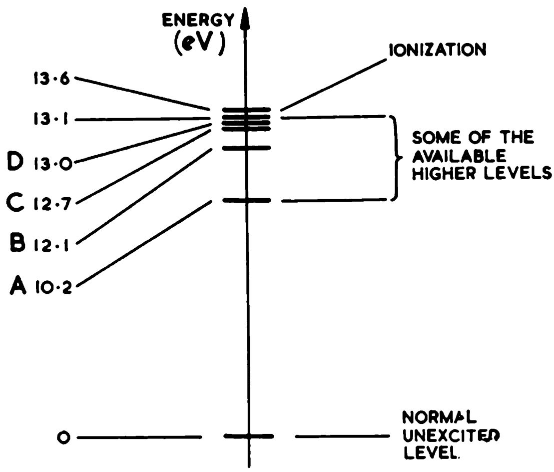


FIG. 3.1

Nuclei are much more massive than electrons. In a normal isolated atom the total negative charge of the electrons equals the positive charge of the nucleus, and the atom as a whole is electrically neutral. The simplest atom is that of hydrogen, which has a nucleus, called a proton, and a single electron. In order to remove this electron from the normal atom it is found that the minimum amount of energy required is always the same. It is concluded that the energy of the electron in the normal atom has one unique value and is the same for all isolated hydrogen atoms. The electron may acquire additional energy, say from impact with a fast moving particle, and the atom is said to be excited to a higher energy state or level. It is found, however, that only certain definite total energy values are possible; intermediate energy levels are not observed. This remarkable result has been established experimentally by the study of the radiation spectra of elements. The possible energy states for the isolated hydrogen atom are shown in Fig. 3.1, where the normal unexcited level is chosen arbitrarily as the zero of total energy. It is seen that excitation of the hydrogen atom requires at least 10·2 eV of energy from some outside source. Higher energies may excite the electron to levels such as B (12·1 eV), C or D. Electrons do not remain in these excited levels but return to the unexcited state in one or more steps, staying at each excited level for a short time of about 10^{-8} sec. During each step a photon of radiation is emitted in accordance with the familiar relation

$$hf = E_2 - E_1,$$

where f is the frequency, h is Planck's constant ($6{\cdot}6 \times 10^{-34}$ joule seconds) and E_1 and E_2 are the two levels of energy associated with the step.

When the electron is removed altogether the atom is ionized and is left with a positive charge. The energy required to just ionize a normal atom is the ionization energy, and for a hydrogen atom this is 13·6 eV; hydrogen is said to have an ionization potential of 13·6 V.

In the helium atom there are two electrons. In the normal isolated atom these two electrons have the same energy but they have opposite spins. In lithium with three electrons two have the same energy and opposite spins; the third electron has a greater energy. As we continue up the periodic table to atoms with higher atomic number and more electrons, it is found that only certain energy levels are possible and that each level may be occupied by two electrons, which always have opposite spin.

The atoms with greater atomic number can also be excited to higher energy states by external means. However, as there are now several electrons, there is greater complexity of energy levels and corresponding spectra. In Table 3.1 are given some excitation and ionization potentials for some of the elements commonly used in electronics; these are the least values required, firstly, to just excite the atoms, and, secondly, to just ionize an atom by releasing a single electron. In the latter case the electron involved is the one of highest energy in the normal atom.

Some atoms have excitation levels which are occupied for quite long times, 10^{-2} sec, compared with the normal excitation levels. Such atoms are said to be in a metastable state. In this state a further acquisition of energy may take the atom to a higher level of the normal short period type. The atom may then return to its unexcited state by the emission of radiation. The importance of the metastable states is that they permit excitation or ionization of atoms in two steps. Mercury has metastable

Atom.	*First excitation potential.* V	*First ionization potential.* V
Hydrogen	10·2	13·6
Helium	20·9	24·6
Neon	16·6	21·6
Sodium	2·1	5·1
Argon	11·6	15·8
Krypton	10·0	14·0
Xenon	8·4	12·1
Mercury	4·9	10·4

TABLE 3.1.—Excitation and Ionization Potentials of Various Atoms

states, and they play an important part in the flow of charge in mercury vapour (see Section 5.13).

3.3. Electrons in Gases

So far we have been considering electron energy levels in isolated atoms. We shall see in the next section that there are significant changes in the energy levels when atoms are brought close together. However, in gases the atomic spacing is so great that the changes are small, and the ionization and excitation potentials given in Table 3.1 are also approximately correct for gases.

For electrical conduction to occur in any substance there must be some free charges available. In a gas these charges are produced by ionization. When an atom is ionized electrical conduction may occur by movement of the electrons or the positive ions, or both. The conductivity of the gas, σ, is given by the expression

$$\sigma = q_i\rho_i\mu_i + e\rho_e\mu_e,$$

where q_i and e are the charges of the ion and the electron respectively, ρ is the number of charges per unit volume and μ is the charge mobility, i.e., the ratio of the average drift speed to the applied electric field. On account of the difference in mass, electrons have much greater mobility than ions and the conductivity of gases is due mainly to the electrons (see Section 5.15).

3.4. Electrons in Solids

In solids the atoms tend to arrange themselves in an ordered array or crystal lattice. The nuclei are more or less fixed in this array, and the spacing between them is such that the electrons of adjacent atoms intermingle to some extent. The effects of intermingling are greatest with the outermost or valency electrons which have the highest energies. By studying the X-ray spectra of solids and by relating the spectral distributions to the electron energies it is found that the atomic proximity modifies the energy levels. In the isolated atoms corresponding electrons all have the same energy, and in each atom only two electrons may have the same value of energy. In an assembly of N atoms corresponding electrons do not have identical energy. There are now N possible energy states close together, and again only two electrons of opposite spin may occupy the same state. Thus the N energy states may accommodate a

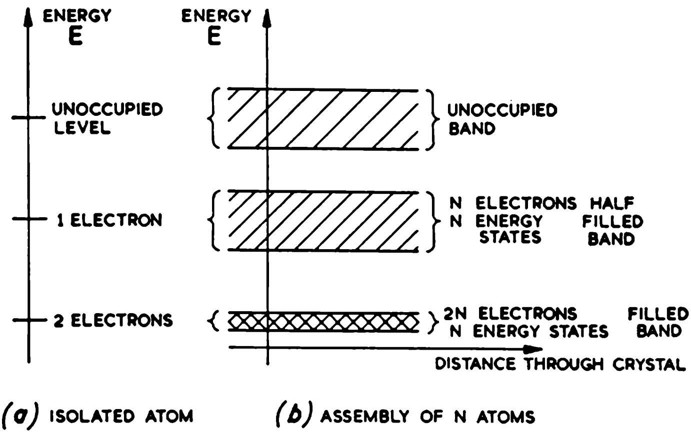


FIG. 3.2

maximum of $2N$ electrons. The single energy level of the isolated atom has become a band of energies in the solid. The individual energies within the band are so close together that the energy band may be considered to be continuous for many purposes. As an example we may consider the metal lithium, the simplest atom which forms a solid at ordinary temperatures. In the isolated atom we have seen that there are three electrons, two with the same energy value and one with higher value as shown in Fig. 3.2.*a*. In the solid the lower level forms a band of $2N$ electrons occupying N different energy states. The higher level forms a wider band also of N energy states; $2N$ electrons could be accommodated in this band, but as there are only N electrons available the energy band is half filled (see Fig. 3.2.*b*). There are also unoccupied excitation bands in the solid corresponding to the excitation levels of the isolated atoms. In some solids there is a gap between the occupied band and the first unoccupied band; in others adjacent energy bands overlap.

In addition to revealing the existence of electron energy bands in solids, the study of X-ray spectra can establish the actual distribution of electrons amongst the available energy states within a band. In the case of the half-filled band corresponding to the valency electrons in lithium it is found that the electron distribution follows a curve of the form shown in Fig. 3.3. Here the ordinate n_E is such that $n_E dE$ is the number of electrons per unit volume with energy lying between the limits E and $E + dE$. The area between the curve and the x-axis thus equals the total number of electrons per unit volume. At very low temperatures all the lowest energy states are occupied and there is a sharp limit to the distribution at W_F, which is called the Fermi level. If the solid is given some additional energy, say by heating, then some electrons occupy energy states above W_F, thereby leaving unoccupied states below W_F, as shown by the dotted line. There is continuous exchange of energy taking place between the electrons, and the higher energies are occupied only for a short time by individual electrons. On the average there are always some electrons in the higher levels. The Fermi level may be defined as the mean energy about which the higher electron energies are perturbed when the solid is heated above absolute zero.

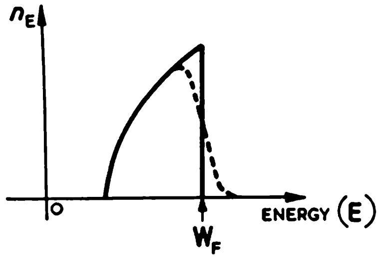


FIG. 3.3

Additional energy may be given to electrons by means of an external electric field. The random variations of energy and motion still occur, but the external field brings about an increase of average energy in a given direction, i.e., a current flows. It is only possible to give increased energy provided the highest occupied energy band is unfilled as in lithium. If the band is filled and there is an appreciable gap to the next unoccupied band the solid is an insulator. If the gap is not too great, then thermal energy may be sufficient for some electrons to bridge the gap, and then they can take part in the process of electrical conduction. Thus in insulators the number of conduction electrons increases with temperature, and so does the conductivity. The lowest unfilled energy band is frequently called the conduction band.

The atom beryllium has four electrons, which, in the solid, fill two energy bands. However, the upper of these bands overlaps the next vacant band so that beryllium is a good conductor.

We have just seen that the conductivity of insulators increases with temperature. The reverse is true in conductors. Increase of temperature does not affect the number of conduction electrons in a metal. The average energy and random motion of the electrons increase, but at the same time there is increased amplitude of vibration of the atomic nuclei about their fixed equilibrium positions. These vibrations interfere with the energy exchanges between the electrons, and as a result the conductivity of a metal decreases with temperature.

3.5. Carbon and the Semi-conductors

Carbon in its diamond form is of particular interest, since its crystalline structure is similar to that of the important semi-conductors, silicon and germanium. These three elements are tetravalent, and they form crystal lattices in which each atom is surrounded symmetrically by four other atoms forming a tetrahedral crystal as shown in Fig. 3.4. Each atom shares a valency electron with each of its four neighbours, thereby form-

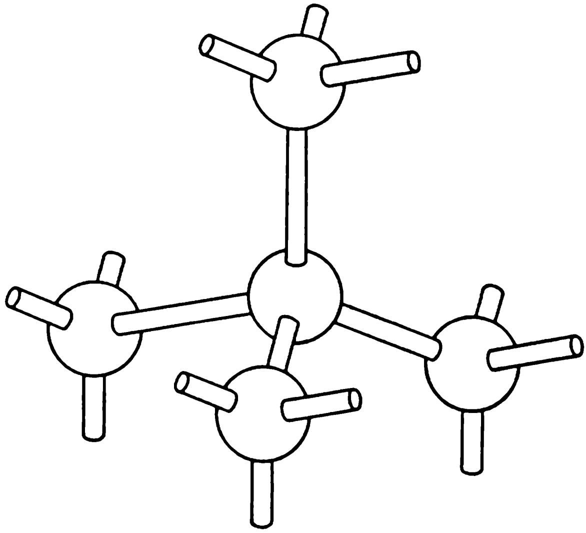


FIG. 3.4

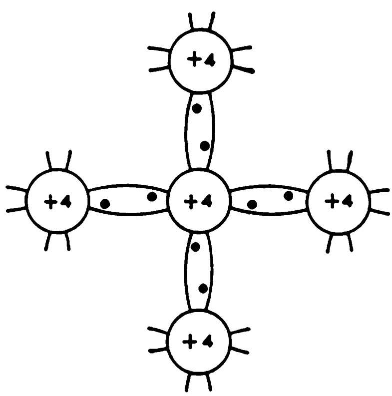


FIG. 3.5

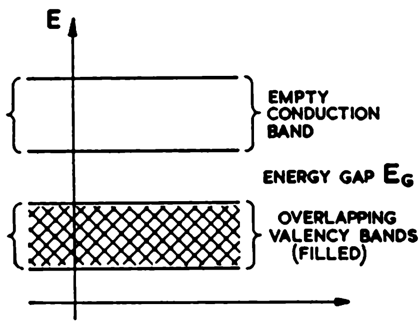


FIG. 3.6

ing a stable structure. The diamond lattice may be represented symbolically by the two-dimensional diagram of Fig. 3.5, in which the circles represent the atom cores consisting of the nuclei and inner electrons with resultant charge $+ 4$; the dots represent the valency electrons. In this arrangement the electrons are seen to be bound to the atoms. The shared electrons form co-valent bonds between the atoms. In terms of energy levels the valency electrons occupy filled energy bands which overlap, and there is a gap to the next unoccupied or conduction band as shown in Fig. 3.6. The width of the energy gap E_G is a measure of the strength of the co-valent bonds. The minimum additional energy required for conduction to take place is E_G. The differences in the electrical properties of carbon (diamond), silicon and germanium depend largely on their values of E_G, which are respectively 7, 1·1 and 0·7 eV. The large gap for diamond means that at all ordinary temperatures there are no electrons in the conduction band and diamond is a good insulator. For both silicon and germanium, thermal energies are sufficient to bridge the energy gap, and these materials have some conductivity at normal room temperatures. Their conductivities increase with temperature. When electrons acquire sufficient energy to bridge the gap E_G they leave vacant energy levels near the top of the valency band. Other electrons in that band may therefore increase their energy by moving into the vacant levels. Vacancies then occur at a lower energy level. These vacant levels due to the absence of electrons are called " positive holes ". Thus conduction is due to the movement of electrons with energies in the conduction band and an equal number of positive holes in the valency band. The conditions in the lattice are shown diagrammatically in Fig. 3.7. Materials of the type described in this section are called intrinsic semi-conductors.

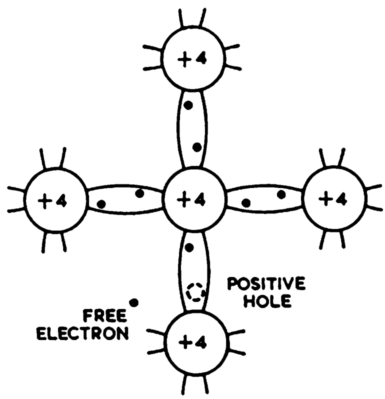


FIG. 3.7

3.6. Impurity Semi-conductors

The behaviour of intrinsic semi-conductors can be modified considerably by the presence of small quantities of impurities, particularly when the impurity atoms have either five or three valency electrons. In the former case the impurity atom may replace a germanium (or silicon) atom in the crystal lattice, contributing four electrons for the co-valent bonds and leaving one over. As long as the impurity atoms are far apart, the single electron behaves rather as it would in an isolated atom. The energy of this electron is just less than that of the conduction band of the germanium, and it may easily be excited by thermal energy into this band and so take part in conduction. An impurity atom of this type is known as a donor, as it gives an electron to the germanium conduction band. With a suitable proportion of donor atoms the conductivity may far exceed that of the intrinsic semi-conductor. The conduction is now mainly by electrons in the conduction band, and the material is called an n-type semi-conductor (negative carriers). Phosphorus and arsenic are impurities which give n-type conduction in germanium or silicon. When the donor atom gives an electron to the conduction band its core is left with a positive charge. However, this positive charge is fixed in the crystal lattice and cannot take part in the conduction process. The n-type semi-conductor is illustrated diagrammatically in the energy diagram of Fig. 3.8.*a* and the lattice diagram of Fig. 3.8.*b*. The Fermi level at room temperature of an n-type semi-conductor occurs in the energy gap but near to the bottom of the conduction band.

When an impurity atom having valency three replaces a germanium

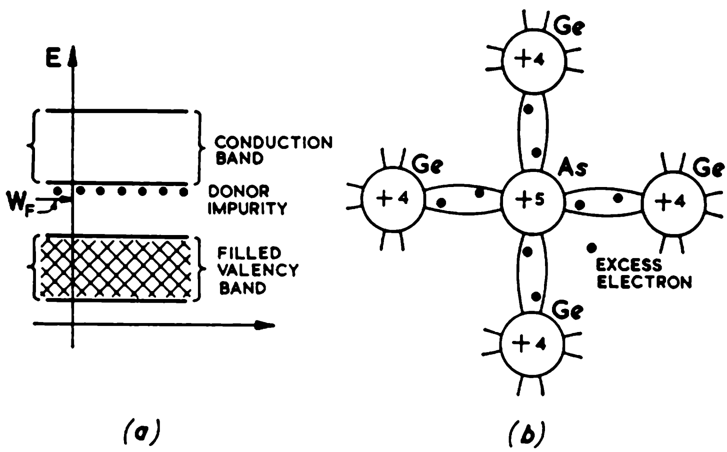


FIG. 3.8

atom in the lattice there is a positive hole available for electrons. The energy of the vacant level is just above the filled valency band of the germanium, and an electron may be easily excited thermally into the vacancy, leaving a positive hole in the valency band. Electrical conduction is now by means of positive holes, and we get p-type germanium. The impurity in this case is an acceptor, since it accepts electrons from the valency band of the parent material. Aluminium, boron and indium act as acceptor impurities with germanium and silicon. Fig. 3.9.*a* and *b* illustrates the p-type semi-conductor. The Fermi level in this case also occurs in the energy gap but near to the top of the valency band. Since the energies of the conduction and valency bands are the same for both types of material, the Fermi level for p-type germanium is less than for n-type.

The conductivity of impurity semi-conductors can be varied over a wide range by varying the concentration of impurity. When the temperature is raised intrinsic semi-conduction is increased and the proportion of majority carriers in either n- or p-type material decreases. It should be noted that even when conduction is due primarily to electrons or to positive holes the net charge of a semi-conductor is zero.

We may thus picture semi-conductors as having free charges consisting of electrons with energies in the conduction band and positive holes in

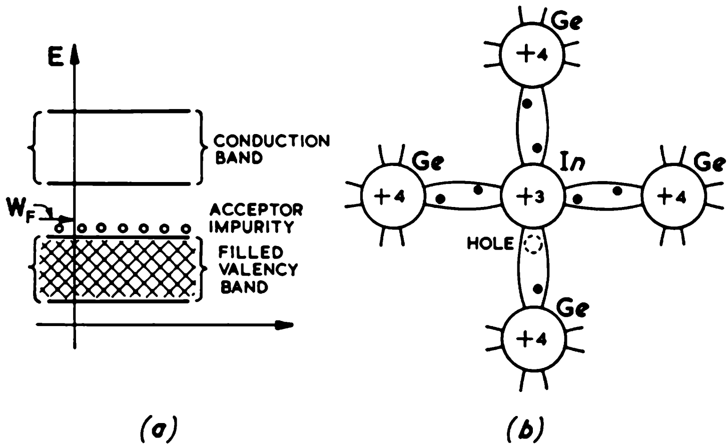


FIG. 3.9

the valency band. The electrons and holes have random motion through the crystal. When an electric field is applied there is superimposed on the random motion a drift of electrons in one direction and positive holes in the opposite direction. The process is very similar to that occurring in gases, where there are also negative and positive carriers, and the conductivity σ of a semi-conductor may be expressed in the form

$$\sigma = e(\rho_p\mu_p + \rho_n\mu_n),$$

where ρ and μ again give the density and mobility of the carriers. In many semi-conductors one or other type of carrier is predominant. The mobility of electrons is somewhat greater than the mobility of holes. These conditions should be compared with those in gases. One major difference between gases and semi-conductors is the method of producing the carriers. Thermal energies are sufficient to produce ionization in semi-conductors, but collision processes are necessary in gases, where the ionization energy is much greater.

The marked variation of conductivity with temperature in semi-conductors is put to practical use in a number of ways for measuring and controlling temperature. Materials prepared for this purpose are known as thermistors; they are characterized by a large negative temperature coefficient of resistance.

3.7. The *p–n* Junction

Due to the presence of donor impurities in *n*-type germanium, there are always electrons in the conduction band. Also due to thermal energies a few electrons are excited into the conduction band, leaving behind some holes in the valency band. Thus in n-type germanium there is an excess of conduction electrons but there are some holes in the valency band. Some electrons are always falling back into the holes as well as into the donor levels. The whole process is a dynamic one, with some statistical average distribution depending on temperature. Corresponding dynamic conditions exist in p-type germanium, where positive holes in the valency band are in the majority but there are always some electrons in the conduction band.

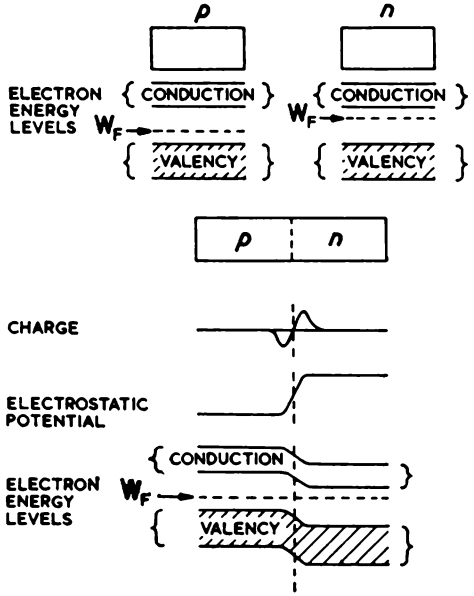


FIG. 3.10

When there is a transition from p- to n-type germanium in a single crystal a diffusion process occurs at the boundary. The excess of conduction electrons in the n-type material causes a diffusion gradient tending to drive electrons across the boundary from the n- to the p-material. Similarly, there is a diffusion of positive holes from p to n. Before diffusion occurs the materials are electrically neutral, and both diffusion processes result in the p-material becoming negatively charged and the n-material positively charged as shown in Fig. 3.10. It should be pointed out that the charged regions at the boundary are not due to excess, but to deficiency of carriers. Thus an electrostatic field is set up which tends to oppose the movement of the charges and there is a potential distribution across the boundary as shown in the diagram. Equilibrium occurs when the resultant average current of holes and electrons in both directions is zero. In this equilibrium electrons may be excited by thermal energy from n to p conduction levels and subsequently combine with holes in the valency band in the p-material. At the same time minority electrons created in the p-material can fall freely across the junction, giving a flow of negative charge in the opposite direction. Similar flow of holes may occur in both directions across the junction, but the net current in the equilibrium condition is zero. Equilibrium occurs when the Fermi level is the same in both materials. Thus the equilibrium energy levels in the p–n junction are as shown in the energy diagram of Fig. 3.10. At the same time there is the electrostatic difference of potential across the junction as described above. As shown in Chapters 5 and 6, these various effects explain the behaviour of junction diodes and transistors.

3.8. Contact Potential in Metals

In a metal the Fermi level lies close to the top of the filled part of the conduction band. The actual value varies from metal to metal. Likewise, the density of conduction electrons varies with the metal. When two metals are brought into contact there is again diffusion of electrons

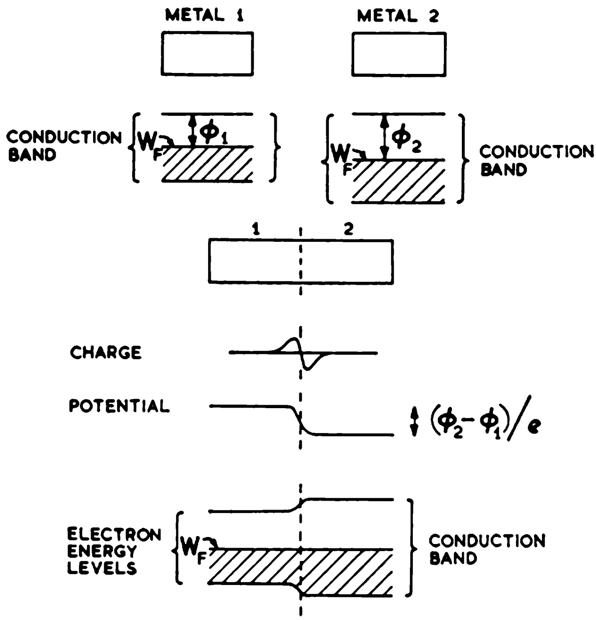


FIG. 3.11

due to difference in concentration, and equilibrium occurs with an electrostatic potential difference across the junction and alinement of the Fermi levels. However, the highest energy electrons can move equally readily in either direction across the junction. The conditions are illustrated in Fig. 3.11. This case should be contrasted with the p–n junction, where thermal energy is required to move electrons from n- to p-material but they may fall freely from p to n. This fundamental difference is due to the fact that conduction electrons are in the majority of the current carriers in n-type germanium but in the minority in p-type. In metals only conduction electrons are concerned.

The contact potential difference between two metals is not affected by the insertion of any number of other metals between them. Its value is equal to the difference in work function of the two metals (see Chapter 4). Contact potential difference is of importance in valves. For example, it may exist between the anode and cathode of a diode, and it is additive to any externally applied potential difference.

CHAPTER 4

ELECTRON EMISSION

4.1. Electron Emission

Many electronic devices depend for their operation on the movement of electrons across the space between two electrodes in a vacuum. This process involves the emission of electrons from one of the electrodes. In this chapter we consider some of the factors governing electron emission.

In a metal many electrons are free to move in a random manner amongst the atoms of the crystal lattice, and the electron-energy distribution is of the form shown in Fig. 4.1.*a*. This distribution assumes, amongst other things, that each electron moves in the metal in a region

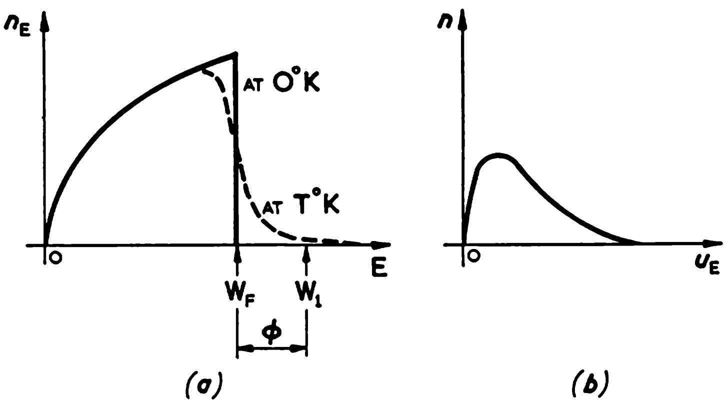


FIG. 4.1

of constant potential, and that there are no electric forces acting on the electron on the average. Since the forces are due to all the atoms and free electrons, then in the interior of the metal the assumption of zero force on the average is reasonable, since each electron has a very large number of atoms and electrons on all sides of it. Some of the electrons near the surface have, in their random motion, velocities directing them outwards from the metal surface. These electrons are no longer surrounded on all sides by charges. Fig. 4.2 represents an electron which has moved a distance x from the surface of the metal. As the electron has left the metal, the latter is positively charged by an amount $+e$. There is a force of attraction between the negative electron and the positive metal. The amount of this force may be determined by the method of images, and is equivalent to the force between two equal and opposite charges separated by a distance $2x$. Thus the retarding force is given by

$$F = \frac{e^2}{4x^2} \cdot \frac{1}{4\pi\varepsilon_0}.$$

The variation of this force with x is shown by the full line in Fig. 4.3.*a*. The above expression assumes that the metal has a continuous surface

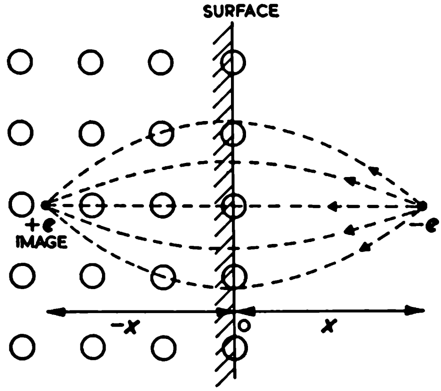


FIG. 4.2

and is therefore true only for values of x large compared with the atomic spacing. Inside the metal we have assumed that F is zero, and hence the effective force on the electron near the surface must take a form similar to the broken curve in Fig. 4.3.*a*. As an electron moves out from the surface, work must be done against the retarding force of amount

$$W = \int_0^x F dx.$$

W is thus the area under the curve of F and is asymptotic to some value W_1, which is the work done against the retarding field by an electron in

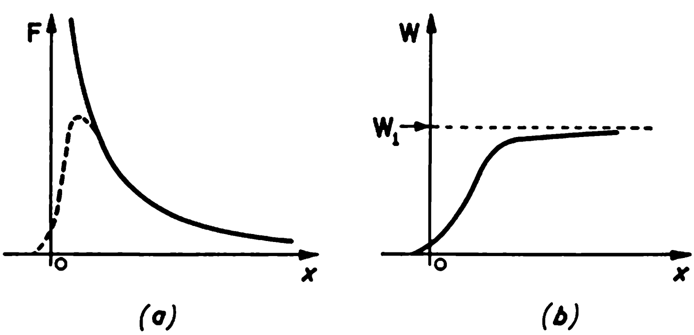


FIG. 4.3

escaping from the surface (see Fig. 4.3.*b*). W_1 represents a "potential barrier" which must be surmounted if an electron is to escape completely from the metal. In order to get over the barrier an electron must have kinetic energy satisfying the relation

$$\tfrac{1}{2}mu^2 > W_1,$$

where u is the component of velocity normal to the surface. Any electron with a value of u less than this limiting value is brought to rest before it has surmounted the barrier and returns to the metal. In Fig. 4.1.*a* only the relatively small number of electrons with energies greater than W_1 may be emitted. It should be noted that by no means all these electrons escape. The kinetic energy must be associated with a velocity at right angles to the surface. An electron with kinetic energy just equal to W_1

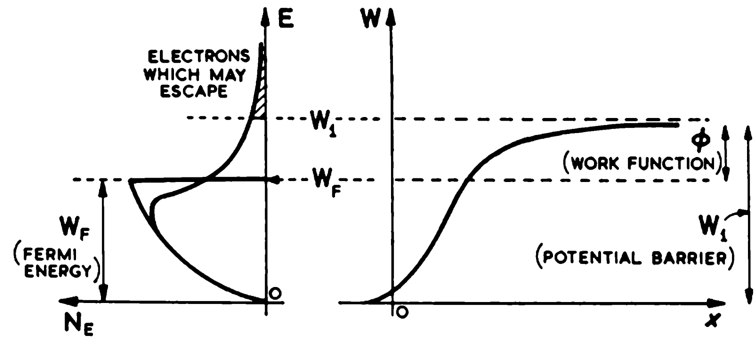


Fig. 4.4

is emitted with zero velocity. Those with greater energy are emitted with velocity u_E, where

$$\tfrac{1}{2}mu_E^2 = \tfrac{1}{2}mu^2 - W_1.$$

There is therefore a distribution of velocities of the emitted electrons as shown in Fig. 4.1.*b*. It might be thought at first that most electrons would be emitted with zero velocity. It is true that there are more electrons in the metal with energy equal to W_1 than with higher energies. However, the velocities of these electrons are distributed in all directions, and a negligible number are moving precisely normal to the emitting surface.

In a metal at absolute zero there are electrons with energies up to W_F, the Fermi level. The additional energy required for emission is then

$$W_1 - W_F = \phi.$$

This quantity ϕ is called the work function of the metal surface. It is usually measured in electron-volts. The work functions of various surfaces lie in the range 1 to 6 eV. The conditions for electron escape from the surface of a metal are summarized in Fig. 4.4. The energy distribution of the electrons in the metal is shown at the left, where Fig. 4.1.*a* has been turned through a right angle in order that the energies may be compared with the potential barrier.

If a suitable electric field is created near the surface of the emitter (or cathode) by having a nearby electrode (or anode) at a positive potential with respect to the emitter, then the escaping electrons are collected and a current flows from the cathode to the anode. The electrons which escape are amongst those with the greatest energies and the emission current causes a loss of energy from the cathode. Emission is therefore accompanied by a drop in temperature of the cathode. The actual drop is small, but it can be measured. The process of evaporation of electrons is closely analogous to the evaporation of molecules of liquid. The cooling of the cathode arises from the latent heat of evaporation of electrons.

If there is no collecting electrode, then the emitted electrons form a cloud of negative charge near the cathode, creating an opposing field which tends to drive the electrons back. An equilibrium condition is set up with electrons continuously entering the cloud and returning to the cathode. This equilibrium is also very similar to that occurring between a liquid and a vapour.

4.2. Types of Emission

At room temperatures very few electrons have sufficient energy to enable them to escape from a solid. The additional energy may be supplied in a variety of ways. When it is in the form of heat we have thermionic emission or primary emission. Alternatively, the electrons may be given the extra energy by the impact on the surface of the solid of fast electrons or positive ions, and then we have secondary emission. Thirdly, if the energy comes in the form of radiation there is photoelectric emission. There is another type of emission in which the electrons are not necessarily given additional energy but the potential barrier is reduced by means of a strong external electric field. This is known as field emission.

4.3. Thermionic Emission—Richardson's Equation

As the temperature of the cathode is increased more of the electrons have sufficient energy to escape. The relation between the temperature and the number of electrons emitted per second from unit area of the cathode has been established by Richardson in the form

$$J_S = AT^2\varepsilon^{-b/T},$$

where J_S = current per unit area, T = absolute temperature, and A and b are constants. It can be shown that b is related to the work function ϕ by the equation

$$b = \phi/k,$$

where k is Boltzmann's constant. If ϕ is in electron-volts and k in joules/° K, then

$$b = 11{,}600\,\phi,$$

and b is said to be the temperature equivalent of the work function, in degrees absolute. The validity of Richardson's Equation can be confirmed by plotting $\log_e J_S/T^2$ against $1/T$ for a diode, in which all the emitted current is collected at the anode for various cathode temperatures, which may be measured with an optical pyrometer. Curves for three commonly used thermionic cathodes are shown in Fig. 4.5. The straight

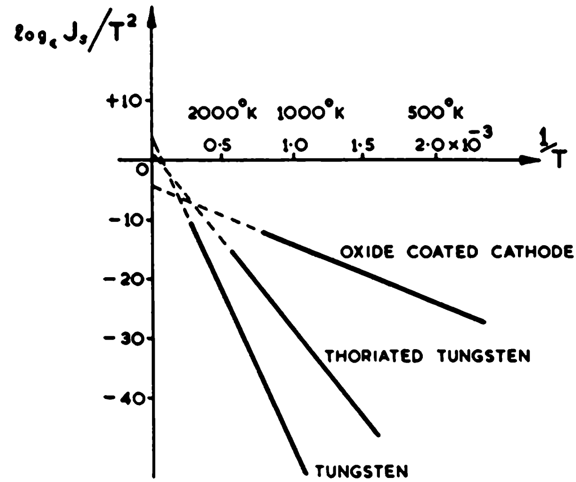


FIG. 4.5

lines confirm the validity, and they may be used for approximate determination of A and ϕ. It is found that A varies from 60 A cm^{-2} (° K)$^{-2}$ for many pure metals to 0·01 for a mixture of barium and strontium oxides. The work function ϕ ranges from 6 eV for platinum to 1 eV for the same oxide mixture. It is seen from Richardson's Equation and the curves

Substance	*A*, A cm^{-2} (° K)$^{-2}$	ϕ, eV.	*b*, ° K	*Working temperature*, ° K.	*Emission density*, mA/cm^2.	*Emission efficiency*, mA/W.
Tungsten . .	60	4·5	52,000	2,500	250	4
Thoriated tungsten	3	2·6	30,000	1,900	1,500	60
Barium oxide and strontium oxide on nickel	0·01	1·0	12,000	1,000–1,100	300	200

TABLE 4.1.—PROPERTIES OF THERMIONIC EMITTERS.

of Fig. 4.5 that the emission density J_S depends much more on b or ϕ than on A. J_S also increases rapidly with temperature, so that an important factor in determining the suitability of a material as an emitter is the maximum working temperature. Of the pure metals tungsten is the most widely used, largely on account of its high melting point. Values of A, b and ϕ for various materials used as emitters are summarized in Table 4.1.

4.4. Thoriated Tungsten

A mixture of two metals may have a lower work function than either of the pure metals alone. Thus a tungsten emitter with a small quantity of thorium has a work function of 2·6 eV, compared with 3 4 for thorium and 4·5 for tungsten. At the same time the thoriated tungsten may be used at a temperature above the melting point of thorium. Thus thoriated tungsten has the advantage of a reduced work function and a high working temperature. During operation the thorium gradually evaporates from the surface but more atoms diffuse from the core, maintaining the conditions at the surface. The working temperature must be restricted to prevent excessive evaporation of thorium. The thorium is introduced in the form of the oxide thoria (about 1 per cent).

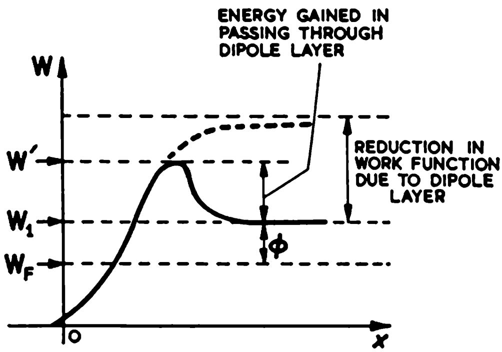


Fig. 4.6

It is believed that a single layer of thorium atoms is formed on the surface of the tungsten and serves to reduce the potential barrier. The atoms arrange themselves on the surface as dipoles with their positive sides on the outside. An electron passing through the dipole layer has a considerable outward acceleration. As a result, the potential barrier is reduced after the manner of Fig. 4.6. At first sight it appears that an electron requires to have energy W' for escape in order to get over the top of the barrier. However, in the case of a narrow barrier of the type shown in Fig. 4.6, it is possible for an electron to " tunnel " through the peak. Thus minimum energy needed for escape is W_1. The " tunnel effect " can be explained using the Principle of Uncertainty.

4.5. Oxide-coated Cathodes

Oxides of any of the alkaline-earth metals (calcium, strontium and barium) have very good emission characteristics. They have to be supported on a suitable metal base, which may affect the emission. The usual combination is a base of nickel, coated with a mixture of about equal parts of barium and strontium oxides. Oxide-coated nickel cathodes are used in practically all small thermionic valves. Although they have been in use and have been studied continuously for many years, their behaviour is still not fully understood, and widely differing values of the emission constants have been quoted.

4.6. Comparison of Various Thermionic Emitters

Pure tungsten, thoriated tungsten and oxide coatings all have their uses as thermionic emitters. The work functions are respectively 4·5, 2·6 and 1·0 eV, and the corresponding operating temperatures are 2,500° K, 1,900° K and 1,000° K. Of course the emission from any material increases rapidly with temperature, but the maximum temperature is limited to a value that gives a reasonable life. This is usually determined by the rate of evaporation of an essential constituent—tungsten, thorium and barium respectively in the cases under consideration.

Since oxide-coated cathodes give their emission at much lower temperatures than either of the tungsten cathodes, they require much less energy to heat them. The ratio of the emission current to the heating power, usually measured in mA/W, is called the emission efficiency. Another important parameter of an emitter is the emission current per unit area at the operating temperature (see Table 4.1). The values quoted are subject to wide variations, particularly for emitters other than pure tungsten. As already stated, the emission densities may be increased considerably by operation at higher temperatures, but this results in reduced life. The figures in the table are all based on continuous operation. Under certain intermittent conditions much higher emission currents may be obtained, especially from oxide coatings. Increases up to 20 times have been quoted for operation in pulses of 1 μs duration. Current densities of 100 A/cm² and more have been obtained from oxide cathodes in high-voltage pulsed operation, but some of this may be due to the large electric field causing enhanced emission.

In addition to emission density and emission efficiency, electrical and mechanical robustness are important properties of emitters. In any valve, no matter how careful the evacuation, there are always present large numbers of gas molecules which may form ions by impact with electrons when current flows. The positive ions move under the influence of the electric field and finally strike the cathode. If high voltages are used the cathode is subject to considerable bombardment and may be damaged, particularly in the case of oxide coatings. For this reason continuous operation with oxide cathodes is limited to less than 2,000 V. Thoriated tungsten cathodes are also liable, to some extent, to be damaged by ion bombardment, and their use is restricted to voltages of about 5,000 or less. Pure tungsten cathodes may be used at still higher voltages up to 20,000. The ability to withstand high-voltage operation may be called electrical robustness.

When mechanical robustness is considered the position is reversed and the oxide coating is distinctly superior to either of the tungsten cathodes. This is particularly true when the oxide cathode is used in the indirectly heated form which is described below. Tungsten cathodes, which are usually in the form of wire filaments, are inherently fragile.

The fields of use of the various types of cathode may be summarized roughly as follows. On account of the high emission efficiency, oxide-coated cathodes are used whenever possible. All receiving valves, small transmitting and amplifying valves and gas-filled valves have oxide cathodes. Pure tungsten is used in the largest valves, where high powers or high voltages are involved, as in large transmitters, high-frequency industrial heating sets and X-ray generators. Thoriated tungsten is used in some medium power valves. In high voltage pulsed applications, such as radar, oxide-coated cathodes are generally preferred.

4.7. Mechanical Form of Thermionic Emitters

Practically all tungsten and some oxide-coated cathodes are constructed from wire filaments. Filamentary cathodes are heated directly by passing current through them. There is therefore a potential drop along the cathode, and for some purposes this is undesirable. This difficulty is eliminated in the indirectly heated cathode, two forms of which are illustrated in Fig. 4.7. In each of these a nickel cylinder has an oxide coating on part of its outside surface, and it is heated with a separate wire inside the cylinder. The heater is usually of tungsten coated with alumina to insulate it from the cathode. Indirectly heated cathodes are normally used with oxide coatings where the nickel base may be formed readily in any convenient shape.

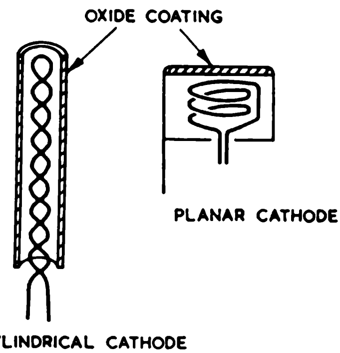


FIG. 4.7

In some directly heated cathodes, particularly in large transmitting valves, the magnetic field produced by the heating current may influence the electron flow.

4.8. Secondary Emission

When a beam of electrons or other particles strikes a surface with sufficient energy secondary electrons may be knocked out of the surface. The emitted electrons are called secondary to distinguish them from the incident or primary beam. The energy of the primary electrons must exceed some minimum value, and at first the number of secondaries increases with the primary energy. The ratio of the number of secondaries to the number of primaries is called the secondary emission coefficient, δ.

The value of δ depends not only on the energy of the primary beam but also on the nature of the surface and the angle of incidence of the beam. The variation of δ with the energy of the primaries usually follows a curve of the form shown in Fig. 4.8. As the primary energy increases δ rises fairly rapidly at first as more secondaries are knocked out. At some voltage, usually in the region of a few hundred volts, δ reaches a

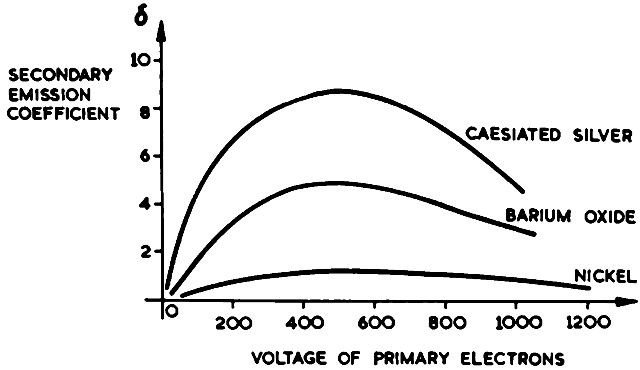


FIG. 4.8

maximum and then drops slowly. When a primary electron is moving very fast it may penetrate some distance into the metal before striking an electron, and this electron has less chance of escape than one released near to the surface; hence the maximum in the curve for δ.

In a diode with a positive anode the emission of secondary electrons is relatively unimportant, since they are immediately attracted back to the

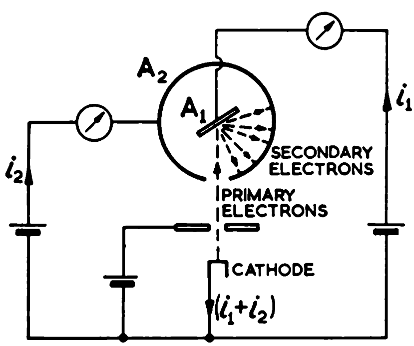


FIG. 4.9

positive anode. However, in a valve with two positive electrodes the secondary electrons emitted by one of them may be collected by the other. In Fig. 4.9 a primary beam strikes a target electrode A_1 and the emitted secondaries go to the surrounding collector A_2. The effect of the secondary emission is to reduce the target current i_1. If the secondary emission coefficient is greater than unity, then the secondary current exceeds the primary current and i_1 is negative. The total electrode current $i_1 + i_2$ is, of course, equal to the primary current, and is unaffected by secondary emission.

Electrodes with values of $\delta > 1$ may be used to give current amplification. By use of an arrangement such as that shown in Fig. 4.10 with several secondary emitting electrodes, large amplifications are possible. If δ is the coefficient for each electrode, then the current from the primary cathode is amplified at the final collector A by a factor of δ^n, where n is the number of secondary emitting electrodes. The primary beam in a device of this kind is usually produced from a photo-electric cathode and the tube is called a photo-multiplier. Current amplifications of the order of 10^6 may be obtained with tubes of this type. Satisfactory

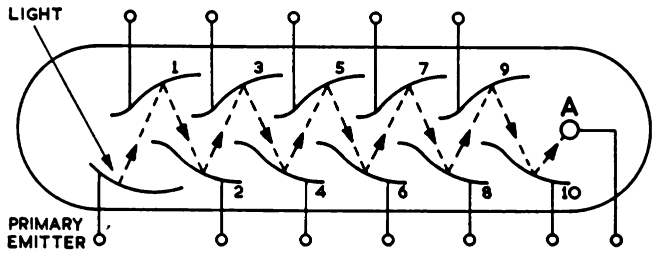


FIG. 4.10

operation of multiplier tubes depends on the production of surfaces with values of δ greater than unity by an appreciable amount. For clean metal surfaces δ ranges from about 0·5 to 1·5 at the optimum voltage. However, certain surface impurities consisting of thin layers of insulators or semi-conductors give values up to about 10 or even more. Caesium oxide on a base of silver with a value of δ about 6 is used in photo-multipliers. Satisfactory operation of multipliers depends not only on having high values of δ but also stable values of δ, if uniform performance is to be obtained. Since δ varies with the electron velocities, very stable voltage supplies are essential.

Secondary emission plays an important part in the operation of a cathode-ray tube. The phosphors which are used in the screen are insulators, and but for secondary emission the screen would develop a negative potential as the primary electrons are collected. The negative potential would increase until it exceeded the cathode potential, and then no more electrons would reach the screen. However, there is secondary emission from the screen, and the screen potential automatically adjusts itself to an equilibrium value nearly equal to the beam voltage.

Secondary emission occurs in many other electronic devices, and it may modify the working conditions considerably. Its effect on the characteristics of triodes and tetrodes is considered in Chapter 6. Where it is desirable to minimize secondary emission from an electrode, its surface is sometimes coated with carbon black, for which δ is 0·5.

4.9. Photo-electric Emission

Photo-electric emission occurs when electrons are emitted from a surface as the result of incident radiation falling on the surface. The energy for emission is supplied to the electrons by the photons of radiation. An essential condition for emission is that the energy of a photon must exceed the work function of the surface, i.e.,

$$hf > \phi$$

where ϕ is the work function. Thus the higher frequency (shorter wavelength) radiations are more effective in producing photo-electric emission. For a given surface the wavelength corresponding to $hf = \phi$ is called the threshold wavelength, and it may be verified that it is given by

$$\lambda = 12{,}400/\phi \text{ Å},$$

where ϕ is in electron-volts. Emission is not possible with radiation of longer wavelength. For photo-emission to occur in the visible spectrum

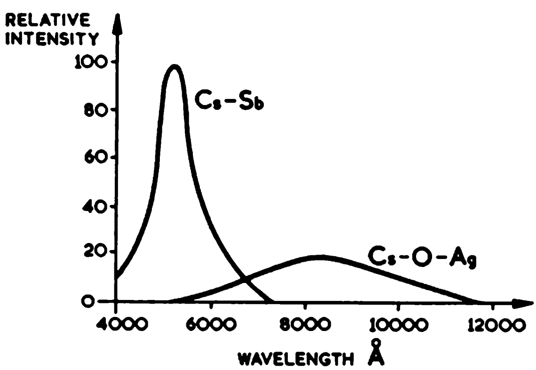


FIG. 4.11

with light of wavelength 6,000 Å, the work function of the emitter must be about 2 eV or less. Amongst the pure metals only caesium (1·9) satisfies this condition. The threshold wavelength is greater for certain composite cathodes. Amongst commercial cathodes caesium–antimony (Cs–Sb) and caesium–oxygen–silver (Cs–O–Ag) have threshold wavelengths of about 7,000 and 12,000 Å respectively. The latter wavelength is well into the infra-red region of the spectrum.

Another important property of a photo-cathode is the variation of the emission with wavelength or the spectral sensitivity. This depends in a complicated manner on many factors, such as the optical properties of the cathode surface. If all the radiation is reflected or transmitted there is no energy transferred to electrons. Where energy is absorbed the emission is greatest when the absorption is mainly at the surface. The spectral sensitivity curves for Cs–Sb and Cs–O–Ag photo-cathodes are shown in Fig. 4.11. In these curves the ordinates represent the emission current divided by the energy of radiation per unit bandwidth over a small band of wavelengths. It is found experimentally that for radiation of a given wavelength the current is proportional to the intensity of the radiation.

4.10. Schottky Effect and Field Emission

An external electric field may affect the emission of electrons from a surface. The external field combines algebraically with the potential barrier, and the emission is reduced or increased, depending on whether the field is retarding or accelerating. The conditions are illustrated in

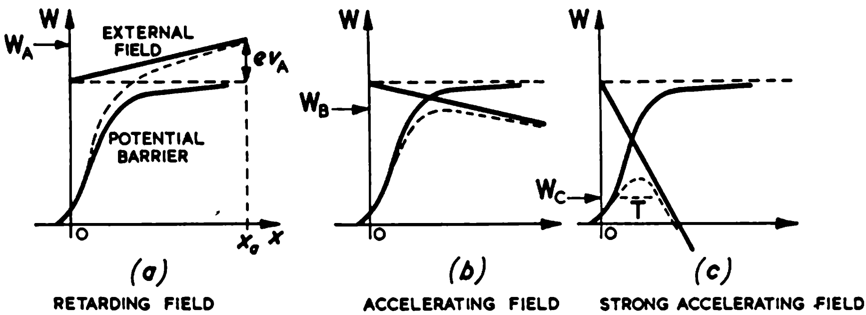


FIG. 4.12

Fig. 4.12. With a retarding field, Fig. 4.12.*a*, i.e., when a nearby electrode is at a negative potential with respect to the emitter, the potential barrier is effectively increased and only electrons with energy greater than W_A escape. When the external field is accelerating, Fig. 4.12.*b*, the effective potential barrier is reduced to W_B, and any electrons with energy greater than this may escape. This is known as the Schottky Effect, and explains the increase of thermionic emission in a saturated diode when the anode voltage is increased. When a very strong external field is set up near the emitter the potential barrier is narrowed as shown by the broken curve in Fig. 4.12.*c*, and now the tunnel effect may operate and large emission may be obtained (T). Under these conditions the many electrons with energy near to or below the Fermi level may be emitted, and the emission is little affected by temperature. Very intense fields are required to produce this Field Emission. Usually it is obtained only from sharp points on the emitter surface when the voltage gradient is of the order of 2 to 5 $\times$ 10^9 V/m. Emission densities as high as 1,000 A/cm^2 or more can be obtained.

A retarding field of the type shown in Fig. 4.12.*a* may be used with a diode to determine the numbers of electrons emitted with various velocities from zero upwards, for thermionic, photo-electric or secondary emission. The velocities are found from the Energy Equation when v is the retarding anode potential, and the numbers are proportional to the anode current.

CHAPTER 5

DIODE CURRENTS

5.1. Flow of Charge

In this chapter the flow of charge in the space between the two electrodes of a diode is considered. The diode is the simplest possible electronic tube, but even so, the evaluation of space charge flow is a complicated process and can be done exactly only in a limited number of idealized cases. Some of the factors affecting the flow are: (i) the nature of the medium between the electrodes, (ii) the size and relative positions of the electrodes, (iii) the electrical potentials of the electrodes, (iv) the availability of free charges in the space, (v) the physical and chemical nature of the electrode surfaces. External influences, such as radiation and magnetic fields, can also affect the current through the diode. Some or all of these factors may play a part either by deliberate action on the part of the valve designer or the user, or sometimes accidentally. It would be extremely difficult to try to take account of many of these factors at one time. In this chapter certain idealized cases are considered, and these may be used to give some qualitative explanation of the measured characteristic curves of actual diodes.

Three main types of diode may be distinguished by the medium between the electrodes: (i) vacuum diodes, (ii) gas-filled diodes, (iii) crystal diodes. Each type is dealt with in turn, and before attempting to explain the physical principles representative characteristic curves are given.

5.2. Characteristic Curves of Vacuum Diodes

For a flow of current to take place there must be charges available to act as current carriers. In a vacuum diode these charges are emitted from the cathode by using the thermionic effect or the photo-electric effect. In the thermionic diode the cathode is heated, and the number of electrons emitted per second may be controlled by varying the cathode temperature. If the anode is maintained at a constant positive potential v_A with respect to the cathode, the emitted electrons flow to the anode, and there is a current in the external circuit. At room temperatures the emission is negligible. A measurable current is first obtained at some temperature T_1, as shown in Fig. 5.1. The current

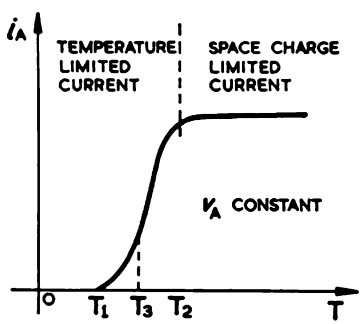


FIG. 5.1

increases with temperature up to T_2, after which it remains practically constant. If the cathode temperature is kept constant at some value such as T_3, then the anode current varies with the potential difference v_A after the manner shown in Fig. 5.2.*a*. This curve was obtained for a diode

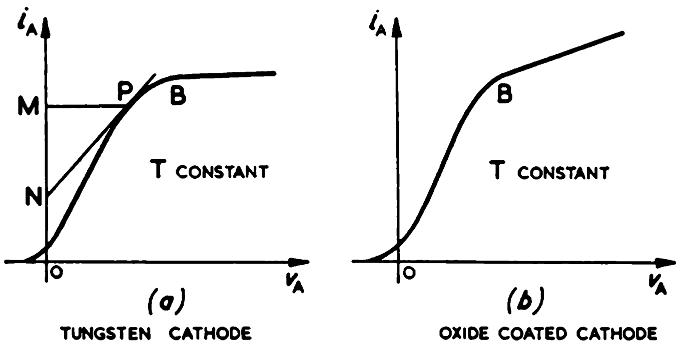


FIG. 5.2

with a pure tungsten cathode. Fig. 5.2.*b* shows the characteristic curve for a diode with an oxide-coated cathode. In both cases the current starts to flow at a small negative anode voltage and increases steadily as v_A is made more positive up to a point *B*. In the tungsten diode the current then increases only very slightly with v_A. In the oxide-coated

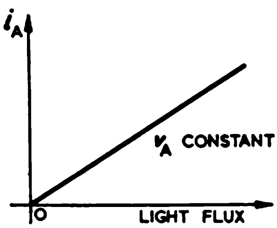


FIG. 5.3

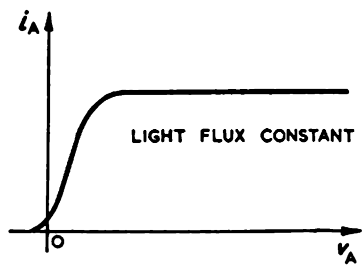


FIG. 5.4

diode the current continues to rise beyond *B* but at a slower rate. The rate of variation of the current with the voltage varies considerably. The quantity r_a, defined by

$$\frac{1}{r_a} = \frac{\partial i_A}{\partial v_A},$$

is called the diode slope resistance. Its value at any point on the curve may be found from the gradient of the tangent at that point. For example, in Fig. 5.2.*a*, $r_a = MP/NM$ gives the slope resistance at the point *P*.

In the photo-electric vacuum diode the electron emission is controlled by varying the quantity of radiation falling on the cathode as shown by the characteristic curve in Fig. 5.3. If the incident radiation is kept constant, then the anode current varies with the voltage across the diode as shown in Fig. 5.4. From these curves it is seen that for a given voltage the current varies linearly with the light flux, and for a fixed light flux the current is practically independent of the anode voltage after the initial rapid rise.

5.3. Physics of the Planar Vacuum Diode—Potential Distributions

The first idealized case that we consider is a thermionic diode consisting of two parallel conducting planes whose dimensions are large compared with the distance between them, and we confine our attention to

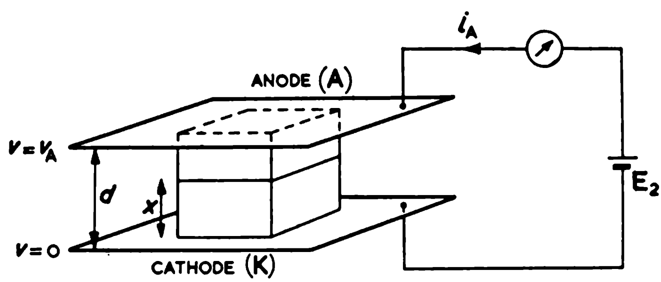
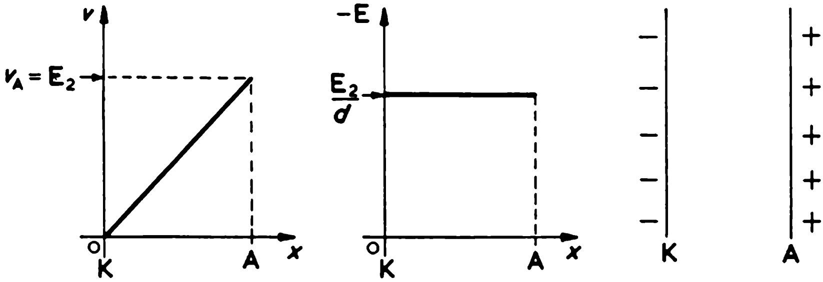
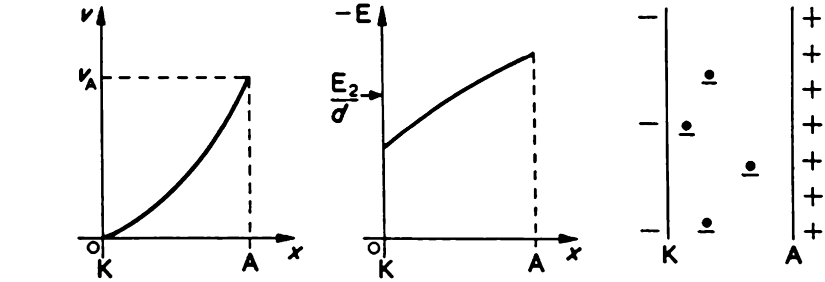


FIG. 5.5

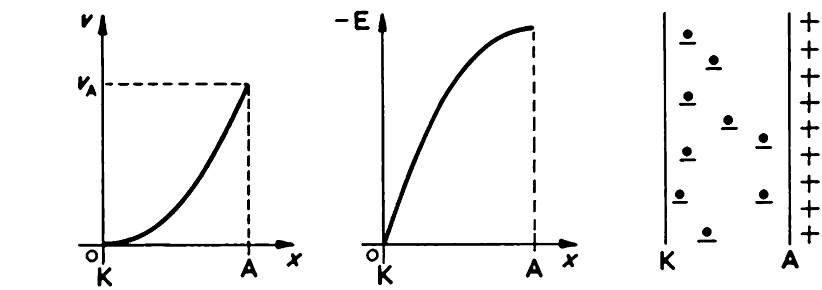
current flow in regions remote from the edges of the planes (Fig. 5.5). When the cathode is cold there is negligible charge in the space and the electric potential varies linearly from zero at the cathode to v_A at the anode, as shown in Fig. 5.6.*a*, where E_2 is the e.m.f. of the battery connected to the diode and $v_A = E_2$. The slope of this curve, $\dfrac{dv}{dx}$, gives the magnitude of the electric field strength, which is constant across the diode (Fig. 5.6.*b*). The conditions are also illustrated in Fig. 5.6.*c* with equal positive and negative charges on the anode and cathode. When the cathode temperature is raised until there is some emission the electrons move to the anode under the influence of the electric field and constitute the current whose value is given by Richardson's Equation. However, there are now negative charges in the space, and these cause a reduction in potential throughout the space. The cathode and anode potentials are maintained at their previous values by the battery, and the potential distribution across the diode now takes the form shown in Fig. 5.7.*a*. The corresponding field strength is given in Fig. 5.7.*b*. The presence of the electrons has reduced the field strength near the cathode and increased it near the anode. The negative charge in the space induces positive charges on the cathode and anode. These induced charges add to those



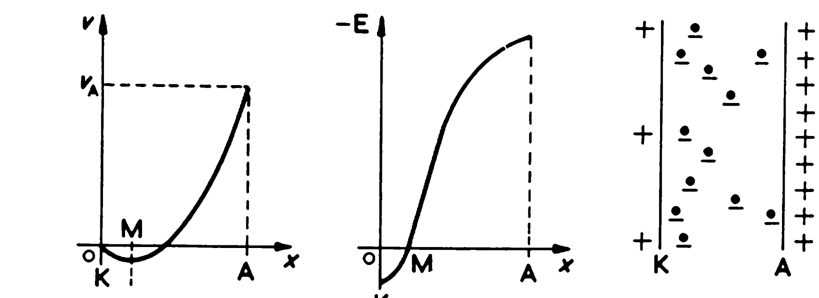
(a) FIG. 5.6 (b) (c)



(a) FIG. 5.7 (b) (c)



(a) FIG. 5.8 (b) (c)



(a) FIG. 5.9 (b) (c)

already in existence, and now the total positive charge on the anode is greater than the total negative charge on the cathode (Fig. 5.7.*c*). The actual charge density on the electrodes is proportional to the field strength just outside the surface. In considering these diagrams it must be realized that there is a steady flow of electrons across the space all the time. The total number of electrons in flight at any instant remains the same, and, although isolated charges are shown in Fig. 5.7.*c*, charge is distributed throughout the whole space. As the emission is increased further, the potential in the space continues to drop. Also, the field strength at the cathode decreases and may become zero as shown in Fig. 5.8, or negative as in Fig. 5.9. In the latter case the force acting on electrons in the region between the cathode and the potential minimum at M is such that it tends to oppose the emission. Now for an electron to reach the anode, it must be emitted from the cathode with a sufficiently high velocity to enable it to pass through the region of the retarding field. Electrons emitted with lower velocities are brought to rest between K and M and then return to the cathode. Under these conditions many of the electrons emitted by the cathode fail to reach the anode, and the current density is considerably less than the value given by Richardson's Equation. The current is limited by the retarding field set up by the electrons in the space and is called a space-charge-limited current. Where the emission is small as in Fig. 5.7 and the field strength at the cathode is accelerating, all the electrons emitted from the cathode reach the anode. Under these conditions there is a temperature-limited current, since the current is determined by the cathode temperature only. The limiting case between temperature limitation and space-charge limitation is shown in Fig. 5.8, where the field strength is zero at the cathode surface. This case would occur if all the electrons were emitted from the cathode with zero velocity. Then the density of the current would adjust itself to the value that would just reduce the electric field to zero at the cathode surface. Any tendency for the current density to increase would give a retarding field at the cathode, and the flow would drop, since the electrons have no initial velocity to overcome the retarding field. On the other hand, any tendency for the current to decrease would give an accelerating field at the cathode, and if the cathode emission is sufficient the current would immediately increase to give the equilibrium state with zero field at the cathode. Conditions such as those just described can be realized if v_A is varied. Increase in v_A raises the potential across the diode. The current density then increases to give zero field at the cathode again. Similarly, reduction in v_A gives a lower current density.

In the case of the photo-electric diode the electron currents involved are always very small, and they have practically no effect on the electric field except for very low values of v_A. There is therefore no space-charge limitation to the current, which depends almost entirely on the amount of radiation falling on the cathode and not on v_A as shown by Figs. 5.3 and 5.4. By analogy with the thermionic diode the photo-electric current is sometimes called " temperature-limited ", though strictly it would be more correct to call it " light-limited " Temperature-limited currents are frequently called saturated currents.

5.4. Planar Vacuum Diode—Space-charge Flow

In the case of space-charge limitation the current density varies with v_A, and is practically independent of the cathode temperature. In most applications valves are used under these conditions. A quantitative relation between the current density and v_A can be established in the following manner for the limiting case when the electrons leave the cathode with zero velocity. In order to ensure that space-charge limitation is maintained it is also assumed that there is an unlimited supply of electrons available from the cathode.

Any region with space charge must satisfy Poisson's Equation

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} = -\frac{\rho}{\varepsilon_0},$$

where v is the potential and ρ the charge density at the point (x, y, z). For the parallel-plane case considered in the last section the problem is one dimensional (see Fig. 5.5). If the x-axis is chosen perpendicular to the electrodes and distances are measured from the cathode, then Poisson's Equation becomes

$$\frac{d^2 v}{dx^2} = -\frac{\rho}{\varepsilon_0}.$$

The Energy Equation provides an additional relation

$$mu^2/2 = ev,$$

where u is the electron velocity at distance x from the cathode. This equation assumes zero velocity at the cathode where $v = 0$. A third relation connects the current per unit area J with ρ and u

$$J = \rho u.$$

This Continuity Equation merely affirms the continuous flow of charge analogous to the flow of liquid along a pipe. From the Energy Equation

$$u = \left(\frac{2e}{m}\right)^{1/2} v^{1/2},$$

and by substitution in the Continuity Equation

$$\rho = J\left(\frac{m}{2e}\right)^{1/2} v^{-1/2}.$$

Hence from Poisson's Equation

$$\frac{d^2 v}{dx^2} = -\frac{J}{\varepsilon_0}\left(\frac{m}{2e}\right)^{1/2} v^{-1/2},$$

i.e., $$\frac{d^2 v}{dx^2} = Av^{-1/2}, \text{ where } A = -\frac{J}{\varepsilon_0}\left(\frac{m}{2e}\right)^{1/2}$$

The equation may be integrated by multiplying both sides by $2\,\dfrac{dv}{dx}$, giving

$$\frac{d}{dx}\left(\frac{dv}{dx}\right)^2 = 2Av^{-1/2}\frac{dv}{dx}, \text{ i.e., } \left(\frac{dv}{dx}\right)^2 = 4Av^{1/2}.$$

The constant of integration is zero, since $\dfrac{dv}{dx}$ and v are both zero at $x = 0$. The first of these, $\dfrac{dv}{dx} = 0$, is the condition for space-charge limitation with zero initial velocities. If we now take the square root and rearrange the terms we find that

$$v^{-1/4}\,dv = 2A^{1/2}dx.$$

A final integration gives

$$v^{3/4} = \tfrac{3}{2}\,A^{1/2}x.$$

The constant of integration is again zero. On substituting for A we find for the current density

$$J = -\frac{4\varepsilon_0}{9}\left(\frac{2e}{m}\right)^{1/2}\frac{v^{3/2}}{x^2}$$

If numerical values are substituted

$$J = -\frac{2{\cdot}33 \times 10^{-6}\,v^{3/2}}{x^2}\ \text{A/m}^2.$$

The minus sign indicates that the positive direction of the current is from anode to cathode in the valve. At the anode $v = v_A$ and $x = d$, and then

$$J = -\frac{2{\cdot}33 \times 10^{-6}\,v_A{}^{3/2}}{d^2}\ \text{A/m}^2.$$

This equation is called the Child–Langmuir Equation or the "three-halves-power law" for space-charge flow. It may easily be verified from the above equations that $v \propto x^{4/3}$, $E \propto x^{1/3}$, $\rho \propto x^{-2/3}$ and $u \propto x^{2/3}$.

5.5. Effect of Space Charge on Electron Transit Time

In some uses of diodes the time of flight of an electron between the cathode and the anode is important. This is found from the equation

$$\tau = \int_0^d \frac{dx}{u},$$

where u is given by the Energy Equation. In Section 2·5 it is shown that the time of flight between parallel planes in the absence of space charge is $\tau = 2d/u_A$, where u_A is the electron velocity at the anode.

For space-charge limitation u varies as $x^{2/3}$, and so $u = \left(\dfrac{x}{d}\right)^{2/3} u_A$. With this value of u we find the transit time $\tau = 3d/u_A$. Thus the effect of the space charge is to increase the transit time by 50 per cent. The numerical value for τ with space charge can be determined from

$$\tau = \frac{3d}{5{\cdot}9 \times 10^5\sqrt{v_A}} \text{ sec.}$$

5.6. Space-charge Flow for any Geometry

The current flowing in a space-charge-limited diode is proportional to the three-halves power of the voltage for the case of parallel-plane electrodes. This relationship holds for a diode of any shape. Poisson's Equation, the Energy Equation and the Continuity Equation are universally true. Poisson's Equation shows that ρ is proportional to v provided the potential gradient and v are zero at the cathode. The energy equation gives u proportional to $v^{1/2}$. It follows therefore from the continuity equation that J is proportional to $v^{3/2}$. The constant of proportionality may vary considerably with the geometry.

5.7. Effect of Initial Velocities of the Electrons

In deriving the Child–Langmuir Equation for the current density under space-charge-limited conditions it is assumed that the electrons all leave the cathode with zero velocity. It is shown in Section 4.1 that the electrons are emitted with a range of velocities. This affects the value of the current and also, as explained in Section 5.3, there is a potential minimum at some position between the cathode and the anode. The potential minimum acts as a " virtual cathode " and the Child–Langmuir Equation is still applicable to some extent, provided the potentials and distances are measured from the virtual cathode. The current density is now given by an expression of the form

$$J = \frac{2{\cdot}33 \times 10^{-6}\,(v_A - v_M)^{3/2}}{(d - d_M)^2}\,(1 + K) \text{ A/m}^2,$$

where v_A and d have the same meanings as before, v_M = the difference of potential between the potential minimum and the cathode and d_M is the corresponding distance. The factor K is included to make allowance for the electrons passing the potential minimum with some velocity. All of these new factors increase the value of J (v_M is negative). The actual values of v_M, d_M and K depend in a complicated manner on the cathode temperature, the value of J and the total emission given by Richardson's Equation. The effect of initial velocities is greatest at low values of v_A. With an oxide-coated cathode operated at $v_A = 1$ V, J may be about double the value given by the normal Child–Langmuir Equation. Even at $v_A = 10$ V, the initial velocities may increase J by 25 per cent.

5.8. Space Charge in Magnetrons

In Chapter 2 the paths of individual electrons in magnetrons are discussed ignoring the effects of other electrons in the field. In practice, large currents may be involved, and then the space charge must modify the behaviour considerably. The analysis of the magnetron, taking space charge into account, is extremely difficult. However, in the case of the cylindrical magnetron at cut-off a possible state of equilibrium may be reached with all the electrons describing circular paths round the cathode. The greatest density of electrons occurs near the cathode and the density is just sufficient to reduce the field strength at the cathode to zero, i.e., to give the necessary condition for space-charge limitation. The force on each electron due to the magnetic field is towards the cathode. The resultant of this force and the forces due to the electrode potentials and the space-charge distribution is just sufficient to maintain the uniform circular motion. The presence of space charge does not affect the cut-off conditions. The critical magnetic field is derived from the Energy Equation and the condition of zero radial velocity at the anode, and hence it depends only on the potential difference between the electrodes.

5.9. Gas Diodes

When gas is introduced into a diode the flow of current may be modified considerably. This can be illustrated with reference to some of the diodes considered earlier in this chapter. For example, the i_A, v_A charac-

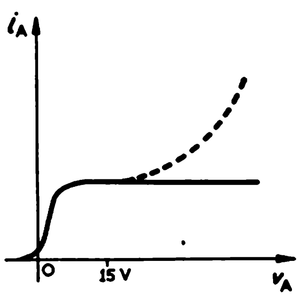


FIG. 5.10

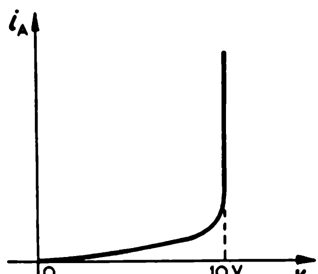


FIG. 5.11

teristic curve of a vacuum photo-electric diode is shown by the full line in Fig. 5.10. The current reaches its saturation value when the anode is a few volts positive with respect to the cathode. The introduction of a small quantity of gas into such a diode has little effect on the characteristic at lower voltages, but when v_A reaches about 15 V the current begins to rise again, and continues to increase fairly rapidly with v_A as shown by the broken line.

The thermionic diode is also affected by the presence of gas. Fig. 5.11 shows the current in a diode which has an oxide-coated cathode and contains some mercury vapour. As v_A is increased from zero the current rises slowly, and at about 10 V has reached 1 mA. When v_A is increased slightly above 10 V the current suddenly rises at an enormous rate. In practice, the characteristic curve of a gas-filled diode is obtained with a resistance in series to limit the current to a safe value. The circuit is shown in Fig. 5.12. When the current rises suddenly a glow appears inside the diode. The size and intensity of the glow increase as the current increases.

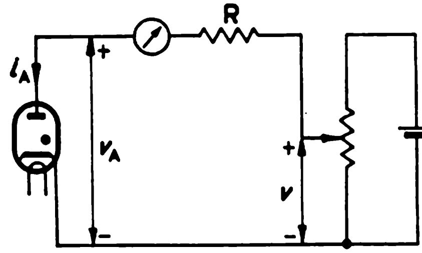


FIG. 5.12

For the flow of charge in a vacuum diode it is essential that the cathode should be heated or irradiated to produce electrons. In the presence of gas, however, there is no need to energize the cathode. If the voltage across the diode is gradually increased from zero no measurable current is obtained at first, but at some voltage v_S there is a sudden flow of current, as shown in Fig. 5.13; v_S is called the breakdown voltage. The value of v_S depends on many factors, such as the pressure of the gas and the geometrical arrangement of the electrodes. In air at atmospheric pressure breakdown occurs between

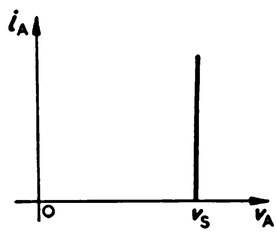


FIG. 5.13

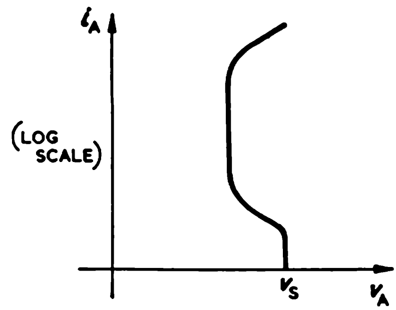


FIG. 5.14

parallel-plane electrodes when the voltage gradient is about 3×10^6 V/m. At lower pressures the breakdown occurs at much lower voltages. In order to investigate the i_A, v_A curve, the circuit of Fig. 5.12 is again used. A typical characteristic for a diode filled with neon at a pressure of about 1 mm Hg is shown in Fig. 5.14. Again it is found that at the breakdown voltage glow appears and the glow increases with current. Flow of current like that represented by Fig. 5.14 is frequently described as a "cold-cathode discharge".

There is one feature which is common to the characteristics of all these gas diodes. There is a marked increase in i_A at some value of v_A. In some cases this increase is so rapid that we say that breakdown occurs. Full explanation of the results is even more difficult than in the case of vacuum diodes. Some of the effects of gas cannot be adequately explained even qualitatively, and little has been achieved by way of quantitative explanation.

5.10. Electron Collisions with Gas Atoms or Molecules

When light shines on the cathode of a photo-electric diode some electrons are liberated and move towards the anode under the influence of the electric field due to the potential difference between the cathode and the anode. At low voltages some of these electrons are collected by the anode. Others escape from the field or are collected on the walls of the container. As the voltage is increased more of the electrons are collected and soon saturation occurs, corresponding to the horizontal part of the characteristic in Fig. 5.10 when all the liberated electrons reach the anode. The electrons in their flight may " collide " with the atoms or molecules of the gas. There are several possibilities, depending on the energy of the electron at the time of impact. A slow electron has an elastic collision with conservation of the total momentum. Since the atom is much heavier than the electron, the speed of the latter is barely affected, but its direction of motion may be altered. If the electron energy at the time of impact is equal to or greater than the excitation potential of the atom, then it may use some or all of its energy to excite the atom to a higher level, with subsequent emission of a quantum of radiation as the atom returns to its normal state. Finally, if the electron has sufficient energy it may cause ionization of the atom, resulting in the production of a positive ion and another electron. There are now three charges, instead of one, free to move in the electric field. The positive ion moves towards the cathode, the two electrons towards the anode and all three charges contribute to the current. When ionization takes place there is therefore an increase in the current. This is the explanation of the difference between the characteristics of the vacuum and the gas-filled photo-cells in Fig. 5.10. At about 15 V ionization occurs, causing the current to increase. This effect is known as " gas amplification ".

5.11. Breakdown

If the voltage across the photo-electric gas diode is several times the ionization potential, then one electron liberated at the cathode by radiation may produce ionization at a short distance from the cathode. The positive ion moves towards the cathode and two electrons now move towards the anode. Each of these may gain enough energy to cause further ionization, giving two more positive ions and two electrons. Now there are four electrons available to continue the process. It is obvious that very large current amplification is possible. Another effect may also occur due to the positive ions. It is possible for an electron to be ejected from the cathode due to bombardment by the ions, or for an electron to be produced close to the cathode owing to ionization by ions, although the latter is an inefficient process. If the ions arising from one initial photo-electron manage to produce at least one electron at the cathode by some method or other, then this electron initiates the whole process again, and the discharge once started maintains itself, even if the radiation is shut off. Breakdown has occurred, and the current increases enormously unless restricted by series resistance. Under these conditions there is a self-maintained discharge. The characteristic curve for the gas-filled photo-cell has now become extended, as shown in Fig. 5.15. The region of the characteristic from the commencement of ionization to the breakdown is called the Townsend discharge. After breakdown there is usually a visible glow, and this region is known as the glow discharge. The visible glow is due to the emission of radiation by those

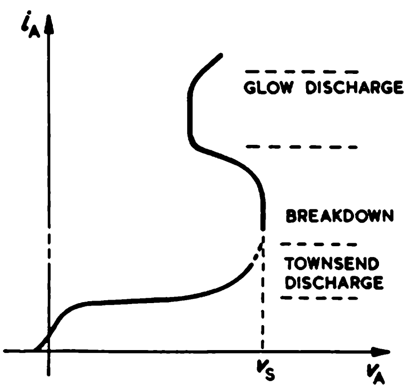


FIG. 5.15

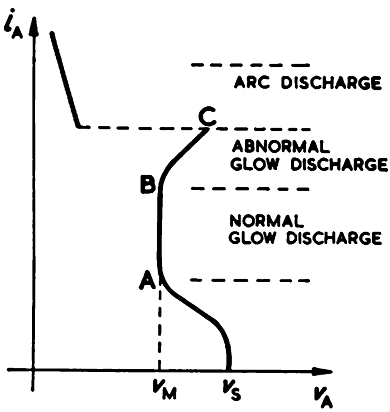


FIG. 5.16

atoms which have been excited to higher energy levels by electron collisions and are returning to their normal state.

In practice, gas-filled photo-cells are never used in the glow-discharge region, since the bombardment of the cathode by ions damages the photo-electric coating on the cathode. The limits of safe operation with reasonable gas amplification are specified by the maker.

5.12. Cold-cathode Discharge

There is always some ionization taking place in a gas. Cosmic rays, radioactive materials or some other source of radiation cause the liberation of electrons. Normally these electrons move through the gas until they recombine with positive ions. Equilibrium is set up when the rate of ionization by external causes equals the rate of recombination of ions and electrons. If, however, an electric field is maintained between electrodes in the gas some of the electrons and ions are collected at the electrodes. If the field is increased all the charges may be collected before there is any recombination, resulting in a saturated current. If now the field is increased further, ionization may occur, and obviously the process may pass through the Townsend discharge to a glow discharge exactly as in the case of the gas-filled photo-cell. The main difference is that the initial current and saturation may occur at a very low level of the order of 10^{-14} A. The actual characteristic of a cold-cathode diode, as determined by the circuit of Fig. 5.12, is therefore of the form shown in Fig. 5.16. After breakdown occurs at v_S the voltage across the diode drops to v_M. The value of v_M depends on the type of gas, the pressure and the nature of the electrodes. It varies from about 50 V to several hundred volts for different diodes. For a given diode it is found that in this condition of glow discharge the voltage drop across the diode is practically constant, even though the current changes considerably, as shown by the region AB in Fig. 5.16. The voltage v_M is called the maintenance voltage or sometimes the operating or burning voltage. The breakdown voltage v_S is sometimes called the striking or ignition voltage. Examination of a diode working in the region AB shows a distinct glow on the cathode surface, and the area of glow increases with the magnitude of the current. It is believed that the current density in the cathode glow remains constant and the change in current arises from the change in area of cathode glow. Under these conditions the voltage drop across the diode is practically constant. When the whole of the cathode surface is glowing, then further increase in current causes an increase in voltage drop across the diode, as shown by BC in Fig. 5.16. The region BC is sometimes referred to as the abnormal glow to distinguish it from the normal glow discharge over AB.

5.13. Arc Discharge

If the current in a glow discharge is increased through the normal glow to the abnormal glow it is found that at some point such as C in Fig. 5.16 there is a sudden change in the nature of the discharge. The voltage across the diode drops to the order of the ionization potential or even less. The mechanism of this new discharge, called an arc, is not fully understood. It is probable that the metastable energy levels play an important part in maintaining an arc discharge with a voltage drop lower than the ionization potential. Atoms may be excited to a metastable state and then receive additional energy to cause ionization. For example, mercury has an ionization potential of 10·4 V and it has a metastable level of 5·5 V. Thus if an atom of mercury collides with, say, a 6 V electron the atom may be raised to its metastable level. If this atom is now struck by another electron whose energy is 4·9 eV or more it may be ionized. Thus ionization may occur with voltages considerably less than the ionization potential. The possibility of successive stages of excitation to produce ionization is obviously much greater with atoms which have metastable levels of relatively long duration.

In addition to the low voltage drop, an arc discharge is characterized by an intense bright spot on the cathode which may cause evaporation of the cathode material. Another common feature of an arc discharge is a reduction in voltage across the diode as the current is increased, i.e., the arc has a negative slope resistance. The currents involved in arc discharges may be extremely high and frequently result in destruction of the cathode material. In many applications precautions must be taken to ensure that an arc discharge cannot occur. However, in some cases the high current densities are exploited, notably in the mercury-arc tube, in which a pool of mercury acts as the cathode, and in the carbon-arc lamp, in which the intense bright spot is used for illumination.

5.14. Effect of Pressure on Breakdown

For ionization an electron must have sufficient energy when it collides with the atom. The energy of an electron starting from rest at the cathode depends on the potential difference through which it moves before a collision. This potential difference in turn depends not only on the voltage applied to the diode but also on the distance the electron travels before the collision occurs. At the next collision the energy depends on the nature of the previous collision and the distance travelled. The important factors are obviously the applied potential difference and the average distance between the gas atoms, i.e., the " mean free path " of the atoms. The latter is known to be inversely proportional to the pressure when the temperature is constant. At atmospheric pressure the electrons collide with atoms before they have gained much energy from the field, and ionization is unlikely to occur. On the other hand, at very low pressures, although the electrons may acquire high energies from the field, the chances of striking an atom are small and again little ionization occurs. There is therefore an optimum pressure, and in the electron devices in common use this is usually around 1 mm of Hg.

5.15. Hot-cathode Discharge

The discharge in a cold-cathode diode is initiated by the few electrons produced by stray radiations. In the gas-filled photo-cell the electrons are produced by radiation falling on the cathode. In both cases the number of initial electrons is relatively small, and if breakdown is to occur these electrons must be involved in a number of ionizations. Thus the breakdown voltage is several times the ionization potential of the gas. In a gas diode with a thermionic cathode there is a plentiful supply of electrons immediately available, and it is found that there is a marked increase in current when the voltage across the diode is of the order of the ionization potential (see Fig. 5.11). However, the physical process involved is different from the cold-cathode case. The main cause is the neutralization of the negative space charge by the positive ions. We have already seen with the thermionic vacuum diode that the current density is limited by negative space charge. Partial neutralization of the space charge by positive ions allows the current density to be increased at the same anode voltage. With complete neutralization, i.e., with equal electron and ion densities, the full emission current density J_S could be drawn from the cathode. A negligible part of this current is due to the extra charges produced by ionization, as may be seen from the following considerations. The speed u of a charged particle at a position of potential v is given by the Energy Equation

$$u = \sqrt{(2qv/m)}$$

where m and q are the mass and charge of the particle which has moved from rest through a potential difference v without collisions. Thus for a positive ion and an electron with equal but opposite charges

$$u_e/u_i = \sqrt{(m_i/m_e)},$$

where the suffices e and i denote electron and ion respectively. In this equation it is assumed that the two charges have moved from rest through the same potential difference but in opposite directions. The current density due to electrons is

$$-J_e = \rho_e u_e,$$

where ρ_e is the density of the electrons. Similarly, the current density due to ions is

$$J_i = \rho_i u_i.$$

Taking the case of atoms of mercury,

$$m_i = 1{,}840 \times 201\ m_e;$$

and hence for $\rho_e + \rho_i = 0$

$$J_e = 600\ J_i.$$

Thus the space charge can be completely neutralized while the contribution of the ions to the current is negligible. A region with $\rho_e + \rho_i = 0$ is called a plasma. In practice, the flow of current is controlled by the resistance of the external circuit. It is important that the current should be limited to less than the total cathode emission.

We have explained qualitatively the sudden rise in the characteristic of the hot-cathode mercury-vapour diode given in Fig. 5.11. The actual voltage across the diode is less than the ionization potential, so that successive collisions and metastable energy levels must play a part in the process.

5.16. Potential Distribution in Hot-cathode Diode

In a parallel-plane diode with the cathode unheated the potential varies linearly from zero at the cathode to v_A at the anode, as shown by the broken line KQ in Fig. 5.17. When the cathode is heated and a gas discharge occurs the potential variation is as shown by the full line KPQ. Practically all of the voltage drop across the diode occurs in a short range from the cathode, KP. The region P to Q, where the voltage is nearly constant, corresponds to the plasma. In the region KP electrons move from the cathode to the plasma under the influence of the electric field and ions leave the plasma for the cathode. The electron current is much greater than the ion current, but on account of the much lower velocity of the ions the ion density is greater than the electron density. The region from the plasma to the cathode behaves rather like a diode with positive-ion flow obeying a three-halves-power law. The distance from K to P adjusts itself to give the ion flow needed to maintain the equilibrium conditions for the particular value of total current through the tube, as determined by the external circuit. The voltage across KP remains constant. The region KP round the cathode is known as a positive-ion sheath. The rest of the space is filled with plasma. On account of this, the geometrical shape of the electrodes in a thermionic gas diode is much less critical than in a vacuum diode. In the latter the distance between cathode and anode is an important factor in the current flow. In the gas diode the thickness of the positive-ion sheath adjusts itself automatically to suit the value of current required, and the rest of the space is filled with plasma. The voltage across the whole tube, which is practically equal to the voltage across the sheath, remains constant even though the current varies over wide limits.

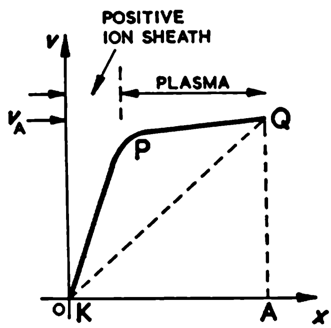


FIG. 5.17

As already mentioned, it is important in using a hot-cathode diode to keep the current below the total emission of the cathode. Otherwise, the voltage drop across the diode would rise and the increased energy of the positive ions would cause damage to the cathode, which is usually oxide coated.

When the anode of a hot-cathode diode is negative with respect to the cathode the device behaves like a cold-cathode diode, but the reverse breakdown voltage is much greater than the ionization potential. The value of the breakdown voltage varies with the gas pressure, and hence with the temperature. The latter has to be maintained within definite limits in a hot-cathode diode.

5.17. Ionization Counters

Certain types of gas-filled diode are used for the detection or counting of atomic, nuclear or radiation particles. A common form of diode counter is shown in Fig. 5.18. The cathode is a metal cylinder forming part of the diode envelope. The anode is a thin axial wire, and the particles to be counted pass through a thin window of mica or some light alloy. In one type of counter the anode voltage is well below the breakdown voltage. When a fast particle penetrates the window it causes ionization and the electrons flow towards the wire anode. The field is weak except near to the anode, where the electrons may acquire sufficient energy for ionization and so give an enhanced flow of charge. The total pulse of current flowing depends on the original ionization produced by the particle, and a diode of this type is called a proportional counter.

If a diode of the type shown in Fig. 5.18 is used at a higher voltage, then a fast particle may initiate a discharge which spreads along the whole length of the diode. Used in this way, the pulse of current through the diode is the same for all particles with energy great enough to trigger the discharge. As in the proportional counter, most of the ionization occurs near to the anode and the electrons are removed rapidly by the field, leaving the more massive positive ions as a space charge surrounding the anode. This space charge reduces the field near the anode, and the discharge ceases when the field is no longer sufficient to maintain it. Thus the initial particle produces a large pulse of current which is independent of the nature of the particle. Diodes of this type are known as Geiger counters. The rate of counting with these diodes is limited by the time taken for the positive space charge to disappear.

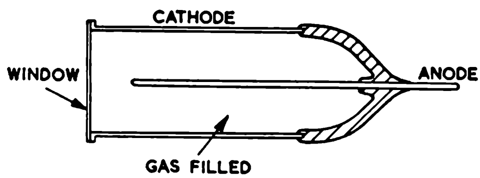


FIG. 5.18

Counters may be used for photons of radiation, in which case the ionizing process is initiated by the emission of photo-electrons from the cathode.

Electronic methods of counting the pulses of current are described in Chapter 18.

5.18. Crystal Diodes

The properties of semi-conducting crystals, such as silicon and germanium, which are discussed in Chapter 3, are used in certain types of diode. These crystal diodes may be divided into two main classes—point-contact diodes and junction diodes. In the point-contact diode a thin metal wire, or " cat's whisker ", makes contact with the surface of a slice of germanium or silicon crystal, as shown in Fig. 5.19.*a*. When the metal wire is made positive relative to the crystal it is found that the resistance to current flow is much less than with the opposite polarity. A typical characteristic curve is shown in Fig. 5.19.*b*.

A semi-conductor like germanium may be either *n*-type or *p*-type, depending on whether its free charges are mainly electrons or mainly holes.

Also *n*- and *p*-type materials may be combined in one crystal to form a *p–n* junction as shown in Fig. 5.19.*c*. Such a crystal junction has a non-linear characteristic of current against voltage of the form shown in Fig. 5.19.*d*. The nature of the *p–n* junction characteristic is explained qualitatively in the following section. The behaviour of the metal-semi-conductor point-contact diode is not fully understood. However, this type of diode is of considerable practical importance as a rectifier, particularly at very high frequencies, on account of its low capacitance of the order of 0·1 μμF.

Other semi-conductors are also found to give non-linear contacts with

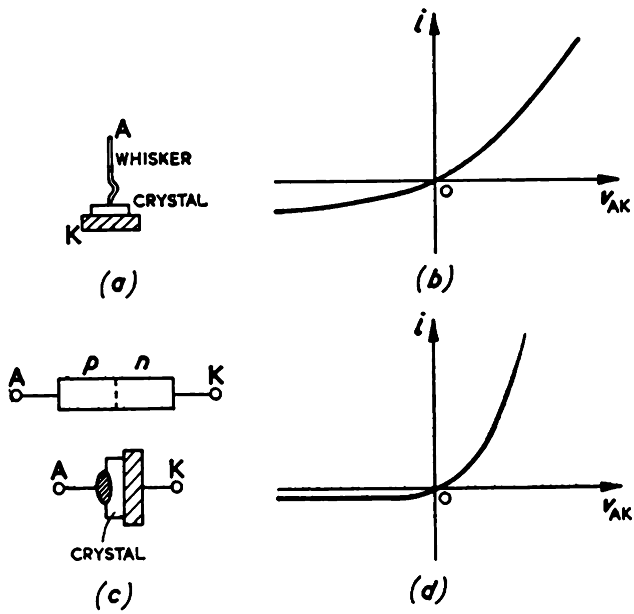


FIG. 5.19

metals. Selenium, copper oxide and copper sulphide are all used, as well as germanium and silicon.

The crystals of silicon or germanium used in diodes have to be prepared most carefully to a high degree of purity, with accurately controlled traces of impurities, in order to give the required type of semi-conduction. The connecting leads must make low-resistance contact with the crystal. As seen in Chapter 3, semi-conductors are highly sensitive to temperature changes. As a result, it is most important that crystal diodes are operated within their specified ratings.

5.19. The Junction Diode

It is shown in Section 3.7 that across the transition region of a *p–n* junction there is a potential difference and distribution of electron energy levels as shown in Fig. 5.20.*a*. Equilibrium is brought about by the charge concentrations and the resulting electrostatic field, which opposes the diffusion of the negative carriers from *n* to *p* and the positive holes from *p* to *n*. The flow of these majority carriers is limited to those with sufficient energy to overcome the potential barrier. There is also a continuous creation by thermal energy of minority carriers, i.e., electrons in the *p*-region and positive holes in the *n*-region. These minority carriers pass readily across the barrier. In the isolated junction in equilibrium there is equal flow of electrons in both directions across the junction. There is also equal flow of holes, and the resultant current is zero, as indicated by the arrows in the energy diagram in Fig. 5.20.*a*. Now con-

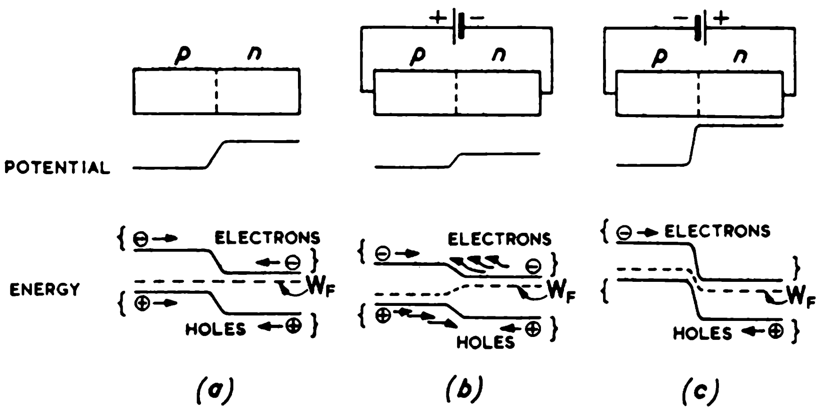


FIG. 5.20

sider the connection of a small potential difference across the crystal so that the *p*-region is positive with respect to the *n*-region, as in Fig. 5.20.*b*. This reduces the potential barrier at the junction so that more holes move across to the *n*-region and electrons to the *p*-region. The holes recombine with free electrons in the *n*-material. Similarly, the electrons recombine with holes in the *p*-material. Both of these movements give positive current flow from *p* to *n*. This current obviously increases with the applied potential difference. There are still minority carriers being created in both regions, and they cause a current in the reverse direction. However, their contribution is no greater than in the equilibrium case with no external potential difference. When the polarity of the applied potential difference is reversed the potential barrier is increased and it opposes further the flow of majority carriers (see Fig. 5.20.*c*). The movement of the minority carriers still occurs, and this now accounts for most of the current. Thus with reversed polarity there is a small saturated current. Since the number of minority carriers in both regions depends on the temperature, the saturated reverse current in a *p–n* junction varies with temperature.

TRIODES, MULTI-ELECTRODE VALVES AND TRANSISTORS

6.1. Triodes, Transistors and Amplification

In Chapter 5 we consider the nature of the current–voltage characteristics of various types of diode—vacuum, gas and semi-conducting. These characteristics vary considerably, but they show that the current depends in some manner on the potential difference between the cathode and the anode. In this chapter we consider in turn the characteristic curves of the triode and other multi-electrode valves, both vacuum and gas-filled, and the transistor. Some attempt is made to give physical explanations of the nature of these characteristics.

6.2. Characteristic Curves of Vacuum Triodes

At a fixed cathode temperature the current flowing to any electrode depends on the potentials of all the electrodes. If the potentials are measured from the cathode, which is taken as the zero of potential, then

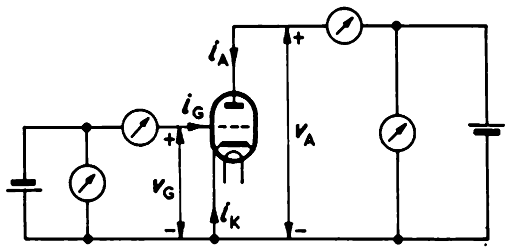


FIG. 6.1

the electrode currents i_A, i_G and i_K are functions of v_G and v_A (see Fig. 6.1) of the form

$$i_G = f_1(v_G, v_A),\ i_A = f_2(v_G, v_A),\ i_K = -(i_A + i_G).$$

The anode current is now a function of two variables, v_A and v_G, and any number of characteristic curves may be drawn. If v_G is taken as the independent variable, then a set of characteristics is obtained as shown in Fig. 6.2, where each curve corresponds to a fixed value of v_A. These i_A, v_G curves are called grid or mutual characteristics. A second set of curves relating i_A to v_A and v_G, but using v_A as the independent variable, is shown in Fig. 6.3, where each curve corresponds to a fixed value of v_G. These are known as anode characteristics. A third way of showing the relationship between the same variables is given in Fig. 6.4, where v_A is plotted against v_G with anode current constant for each curve;

these are constant current characteristics. All of these figures contain the same amount of information and, given any one set, the other two may be derived.

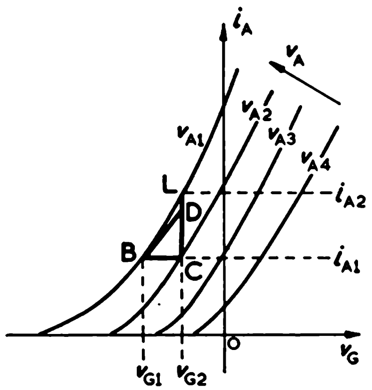


FIG. 6.2

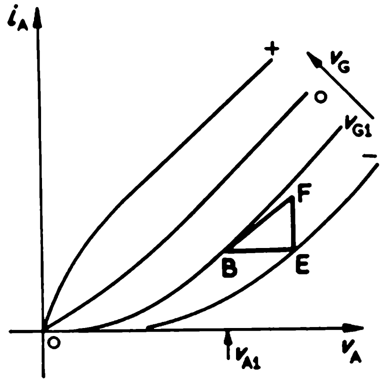


FIG. 6.3

The rate of variation of the anode current with the grid voltage when the anode voltage is constant is an important quantity for any triode. It is defined by the relation

$$g_m = \left(\frac{\partial i_A}{\partial v_G}\right)_{v_A \text{ constant}},$$

and is called the mutual conductance, or sometimes the transconductance, of the triode. It is given by the slope of the i_A, v_G characteristics, and its value varies with the point on the curve where it is measured, particularly when i_A is small. Hence in quoting the mutual conductance it is necessary to give the operating values of the anode current, anode voltage and grid voltage. In Fig. 6.2 a tangent has been drawn at the point B, and then $g_m = CD/BC$; g_m is usually measured in mA/V.

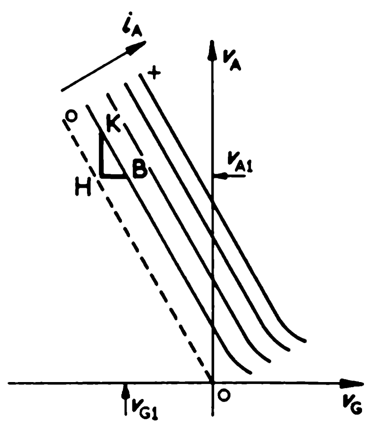


FIG. 6.4

The slope of the anode characteristics for constant grid voltage provides a second important triode parameter, defined by

$$g_a = \frac{1}{r_a} = \left(\frac{\partial i_A}{\partial v_A}\right)_{v_G \text{ constant}};$$

g_a is called the anode slope conductance and r_a the anode slope resistance. It is seen that r_a also varies with the value of the anode current, being greatest when i_A is least. Its value, at the same operating point B as was used for g_m above, is given in Fig. 6.3 by $r_a = BE/EF$.

Finally, the slope of the constant current characteristics gives a measure of the relative importance of the anode and grid voltages in controlling the anode current, and the relation,

$$\mu = -\left(\frac{\partial v_A}{\partial v_G}\right)_{i_A \text{ constant}},$$

defines the amplification factor μ. The minus sign is inserted, since the changes in v_A and v_G are usually in opposite directions if the anode current is to remain constant. Except when v_A is low, the constant current characteristics are parallel straight lines and μ is almost independent of the operating point over quite a wide range. The value of μ at the point B is given in Fig. 6.4 by $\mu = HK/HB$.

It follows from the definitions of μ, g_m and r_a that $\mu = g_m r_a$. There is no need to have three sets of characteristics in order to determine approximate values of μ, g_m and r_a. They may be found from any one set by taking the variation in any two of the variables i_A, v_G and v_A, whilst the third is kept constant. For example, we have seen already how to find g_m from the grid characteristics in Fig. 6.2. The same characteristics may be used to determine μ by using the line BC along which i_A is constant. Along BC, v_A changes from v_{A1} to v_{A2} and v_G from v_{G1} to v_{G2}, giving $\mu = (v_{A1} - v_{A2})/(v_{G2} - v_{G1})$ at an anode current of i_{A1}. Similarly, along CL in the same figure, v_G is constant and $r_a = (v_{A1} - v_{A2})/(i_{A2} - i_{A1})$ at about the same value of anode current. The values of μ and r_a obtained in this way are approximate, since finite increments of the variables are used instead of the infinitesimals implied in the formal definitions.

The curves considered so far have given the anode current in terms of the grid and anode voltages. Similar curves may be drawn relating the grid current to v_G and v_A, and corresponding grid current parameters could be determined. However, in many triode applications the operation is such that very little grid current flows. It can be seen from Fig. 6.2 that currents flow to the anode when the grid voltage is negative. Few electrons flow to the grid under these conditions although it acts as an effective control of the anode current. When the grid voltage is positive some of the electrons flow to it and the total space current is divided between the anode and the grid. Under most conditions of operation the grid current is much less than the anode current, except when the anode voltage is very low. The curvature at the lower ends of the constant anode current curves in Fig. 6.4 is due to an appreciable part of the total space current flowing to the grid. This is also the cause of the change of curvature from concave upwards to convex upwards in the anode characteristics when v_G is positive (see Fig. 6.3).

6.3. Valve Equation for Small Changes

Changes in anode current arising from small changes in both the grid and anode voltages may be determined approximately as shown in Fig. 6.5, in which two grid characteristics are drawn for anode voltages of v_{A1} and $v_{A1} + v_a$. The anode current is i_{A1} when the grid and anode voltages are v_{G1} and v_{A1}, and it becomes $i_{A1} + i_a$ when the voltages change to $v_{G1} + v_g$ and $v_{A1} + v_a$. In the figure, $BC = v_g$ and $CE = i_a$. Since BD

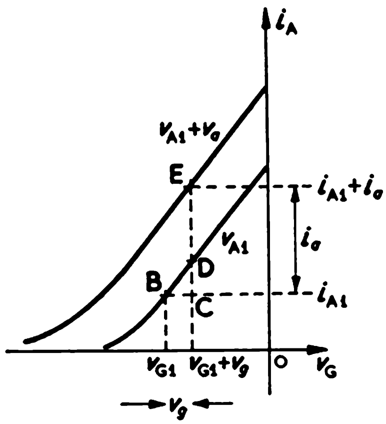


FIG. 6.5

is a grid characteristic and BD is not much different from the tangent at B, then $CD \simeq g_m v_g$. Also, along DE, v_G is constant and hence

$$DE \simeq \frac{\partial i_A}{\partial v_A} v_a = v_a/r_a.$$

Since $CE = CD + DE$,

$$i_a = g_m v_g + v_a/r_a.$$

It is implied above that g_m and r_a are constant over the range of variation. We obtain the important Valve Equation relating the change in anode current to changes in grid and anode voltages. This equation applies as long as the changes are sufficiently small for the valve parameters to remain constant. The grid or anode characteristics may be nearly parallel straight lines over an appreciable range of v_A and v_G (see Figs. 6.2 and 6.3). For that range, g_m and r_a are constant and the Valve Equation can be used.

6.4. Triode Ratings

The operating conditions for triodes vary considerably with their size and the purpose for which the valves are designed. Current may range from a few milliamperes to 100 A, voltage from a few volts to 10 kV, and powers from about a watt to over 100 kW. The current taken by any valve must usually be kept well within the saturation limit. If higher currents are required, then valves with larger cathode area must be used. Voltage limitations usually depend on the quality of the insulation, and on the possibility of an arc occurring between the electrodes. Frequently the voltage and current are not limited by any of these factors, but by the power capabilities of the valve. The product $v_A i_A$ represents the power dissipated by the electrons striking the anode, and its value must be restricted to keep the anode temperature within safe limits. In most small valves the anode gets rid of its heat by radiation through the glass envelope. The permissible anode dissipation may be increased considerably by making the anode part of the external envelope as shown in Fig. 6.6. The anode may now be cooled by natural or forced-air convection, or by circulating water. In cases where electrons flow to the grid another limitation is set by the maximum permissible grid temperature. This limit may be set not only to prevent damaging the grid by over-heating, but also to avoid thermionic emission from the grid. The maximum permissible grid dissipation depends on the nature, number, diameter and length of the grid wires, and on the grid supports, which may conduct or radiate heat from the wires.

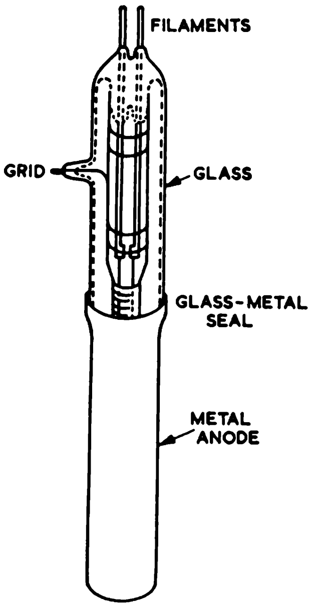


FIG. 6.6

In valves used at high frequencies a further limitation may arise from heating of the electrode leads and connections, due to the high-frequency currents flowing in them.

The various current, voltage and power ratings of any valve are closely interrelated. In use it is desirable to ensure that a limiting rating is not exceeded.

6.5. Physics of the Vacuum Triode

The triode characteristics described in Section 6.2 were all obtained with the cathode temperature sufficiently high to maintain space-charge limitation. If this condition is not satisfied the anode current curves show saturation effects similar to those in diodes when temperature limitation occurs.

In Section 5.3 it is shown that current flow in a diode is determined by the combined effects of the fields due to the positive anode and the negative electrons in the space. When initial electron velocities are neglected the equilibrium condition occurs when the field at the cathode surface produced by the space charge just cancels the field due to the anode potential. When the anode potential changes the current density adjusts itself to restore zero field at the cathode. Clearly the current density is closely related to the field at the cathode due to the anode potential alone. The same conditions apply in the triode, but now the grid and anode potentials both affect the field strength at the cathode. If the combined effect gives a positive potential gradient at the cathode in the absence of space charge, then when the cathode is heated to produce adequate emission, the current density assumes the value that just reduces this potential gradient to zero. These potential variations may be illustrated with reference to a planar triode of the type shown in section in Fig. 6.7.*a*. The grid consists of a series of equally spaced wires. Each of the electrodes is maintained at a definite potential by means of an external battery. The potential at any point in the space depends on the electrode potentials. For all of the diagrams, Fig. 6.7.*b* to *e*, the cathode is at zero potential and the anode at potential v_A. Each diagram corresponds to a different value of v_G. In *b* and *c* the potential gradient at the cathode is positive, and when the cathode emits, current flows at such a density that the potential gradient becomes zero as shown by the broken lines. In *d* the cathode field is zero, as a result of the combined effects of the positive v_A and negative v_G, and no current flows. This is the cut-off condition. In *e* the conditions are well beyond cut-off. The relation between v_A and v_G at cut-off is a straight line passing through the origin if emission velocities of the electrons are neglected (see curve for $i_A = 0$ in Fig. 6.4). Hence the amplification factor as previously defined in Section 6.2 is given by $\mu = -v_A/v_G$ at cut-off. These values of v_A and v_G produce zero field at the cathode in the absence of space charge. The amplification factor may therefore be defined in terms of a purely electrostatic system with no current flow. As defined in this electrostatic manner, μ may be calculated from the geometry of the electrodes. The actual calculations are beyond the scope of this book, but some of the factors affecting μ are self-evident. The further the anode is from the grid, the less is its effect on the field at the cathode. Thus μ increases with the grid–anode spacing. Also, increase of the number of grid wires per unit length, or of the wire diameter, increases the screening of the anode from the cathode, and so increases μ.

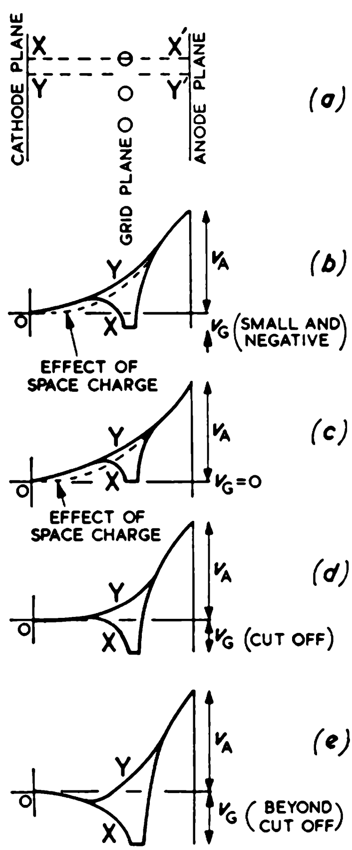


FIG. 6.7

It may be deduced from the electrostatic definition of μ, that $\mu = C_{gk}/C_{ak}$, where C_{gk} and C_{ak} are respectively the grid–cathode and anode–cathode capacitances between the electrodes. These capacitances include only the active electrodes and not any effects due to leads or other fields outside the electrode system.

6.6. Equivalent Diode

From Sections 6.2 and 6.5 it is seen that the grid is much more effective than the anode in controlling the current in a triode; in fact, the grid voltage is μ times as effective as the anode voltage. For some purposes it is convenient to imagine that the triode is equivalent to a diode with anode voltage equal to $v_G + v_A/\mu$. This anode may be placed at any position, but usually it is considered to coincide with the grid. It is shown in Section 5.6 that the anode current density of any diode varies as the 3/2 power of the anode voltage. Hence for the triode

$$i_A = k(v_G + v_A/\mu)^{3/2},$$

where k is a constant.

It follows then that

$$g_m = \tfrac{3}{2}\, k(v_G + v_A/\mu)^{1/2} = \tfrac{3}{2}\, k^{2/3}\, i_A{}^{1/3}$$

and

$$1/r_a = \tfrac{3}{2}\,\frac{k^{2/3}}{\mu}\, i_A{}^{1/3}.$$

These two relationships show that g_m and $1/r_a$ vary as the cube root of the anode current.

6.7. Effect of Space Charge in a Triode

In Fig. 6.7.*b* and *c* the effect of the space charge on the potential distribution is shown by the broken lines. The space charge lowers the potential appreciably between the cathode and the grid but has little effect near the anode. The reason for this is that the electrons are moving most rapidly at the anode. The current density J is constant across the valve and the Continuity Equation tells us that $J = \rho u$, where ρ = space-charge density and u is the electron velocity. Hence when u is greatest ρ is least, and the space charge has least effect on the potential. For many purposes, where valves are operated with negative grid and positive anode, it is permissible to neglect the effect of space charge on the potential in the grid–anode space. However, when the grid potential is positive and comparable

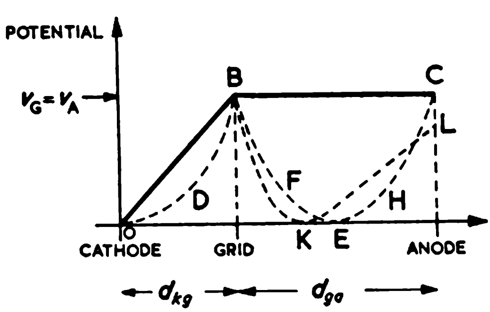


FIG. 6.8

with the anode potential, the space charge may have considerable effect on the conditions in the grid–anode space. This may be shown with reference to Fig. 6.8, which shows the potential distribution in a triode with the mean grid potential equal to the anode potential, and the grid–anode distance double the cathode–grid distance. The full line *OBC* shows the variation in potential across the valve in the absence of current. A mean grid potential is assumed and the variation around the grid wires is ignored. When the cathode emits and a space-charge-limited current flows the potential between the cathode and grid drops, as shown by the broken line *ODB* with zero slope at the cathode. We may assume in the first place that most of the electrons pass through the grid into the grid–anode space. The grid and the anode are maintained at equal potentials v_G and v_A by the external supplies. The space charge, however, depresses the potential between grid and anode. Minimum potential occurs mid-way between, at *E*. As a result, the electron velocities are least and the space-charge density is greatest at *E*. This concentration of space charge tends to depress the potential still further. With the particular arrangement shown, with d_{ga} equal to twice d_{kg}, it is seen from symmetry that the potential at *E* falls to zero if all the electrons leaving the cathode pass through the grid and move on to the anode. The potential and charge distributions for *EFB* and *EHC* are the same as for *ODB*. At *E* the potential is zero, the electron velocities are zero and a virtual cathode is said to be formed there. Although the electrons come to rest at the virtual cathode, it must be assumed that they all go on to the anode from *E*. Otherwise there would be more charge on the grid side than on the anode side of *E* and the symmetry would be upset. If v_A is reduced below v_G, the potential at *E* tends to drop below zero. This is obviously not possible, since no electrons can reach the negative potential region. In this case a zero potential is reached, but at some place nearer to the grid than *E*, as shown by curve *BKL* in Fig. 6.8. Only a fraction of the electrons now pass on to the anode. Those that return to the grid are responsible for the movement of the potential minimum towards the grid. If the grid–anode distance is increased the limiting case for all of the electrons to reach the anode occurs at some value of v_A greater than v_G. Similarly, with smaller grid–anode distance, equal grid and anode voltages produce a positive potential minimum as shown by curve 1 in Fig. 6.9. The limiting case now occurs at some lower anode voltage v_3, as shown by curve 3. For anode voltages greater than v_3 a positive potential minimum occurs as shown by curves 1 and 2. Thus the presence of space charge between the grid and the anode affects

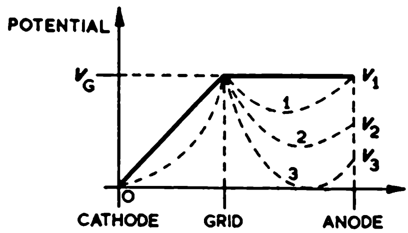


FIG. 6.9

the operation of a triode in the positive grid region. The effects are most noticeable with large grid–anode distances. These considerations are of importance in Class C operation of amplifiers and oscillators, in which the positive grid voltage is comparable with the anode voltage. Also, the existence of a potential minimum between a positive grid and anode is exploited in the beam tetrode (see Section 6.10).

6.8. Characteristic Curves of Tetrodes

In the triode the anode serves two purposes. It acts as the collector of the electron current, but it also controls the amount of current. It is possible to separate these two functions by introducing a fourth electrode between the grid and the anode. This additional electrode, called the screen, consists of a grid of wires which allows most of the electrons to pass through it. For a given value of grid voltage, the current leaving

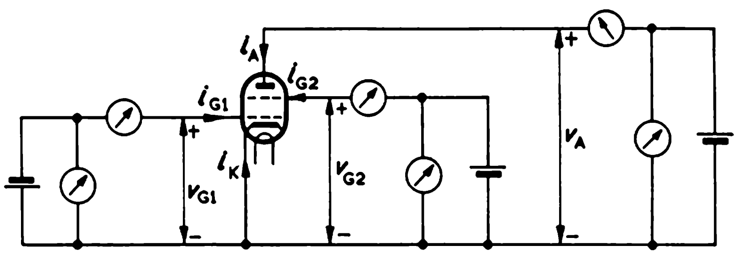


FIG. 6.10

the cathode depends mainly on the voltage of the screen and very little on the anode voltage. In order to distinguish between the two grids they are called the control grid or G_1 and the screen grid or G_2. In the four-electrode valve or tetrode the screen grid serves an additional purpose in high-frequency circuits by acting as a screen between the anode and the control grid.

In the tetrode there are possibilities for electron flow from the cathode to three electrodes, and the current division depends on the relative values of the electrode potentials. Many different sets of characteristics may be obtained using the circuit of Fig. 6.10, but the most interesting are those with the screen grid at a fixed positive potential and showing the variation in anode current with either the control grid voltage or the anode voltage. These may be referred to as control-grid characteristics and anode characteristics. One control-grid characteristic is shown in Fig. 6.11. The slope of this curve is again defined as the mutual conductance g_m of the tetrode, i.e.,

$$g_m = \frac{\partial i_A}{\partial v_{G1}} \text{ at constant } v_{G2} \text{ and } v_A.$$

The full line in Fig. 6.11 shows the anode current i_A and the broken line the cathode current i_K. The difference between these gives the screen current i_{G2}; i_{G1}, the control grid current, is negligible, since v_{G1} is negative for the curves shown. It may be seen from these results that i_{G2} is less than i_A. In most cases it is desirable to have i_{G2} as small as possible; the ratio of i_{G2} to i_A is usually about $\frac{1}{2}$ to $\frac{1}{10}$. The ratio depends not only

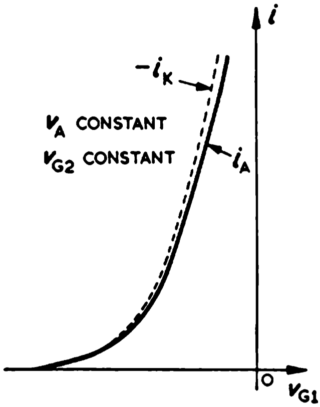


FIG. 6.11

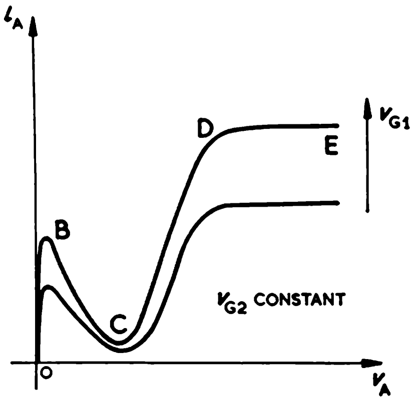


FIG. 6.12

on the electrode voltages but also on the number and size of the screen-grid wires, and on their alinement with the control grid.

Two anode characteristics are shown in Fig. 6.12. Each of these shows a peculiar "kink" with considerable variation in slope. Anode slope resistance r_a may be defined as before as

$$\frac{1}{r_a} = \frac{\partial i_A}{\partial v_A} \text{ at constant } v_{G1} \text{ and } v_{G2}.$$

Then r_a varies from a relatively low value at low anode voltages (OB) to a negative value (BC), then it is positive again (CD) and finishes at a fairly high value (DE). The negative resistance region BC is particularly interesting; over that range the current changes in the opposite direction to the voltage. This is equivalent to the valve acting as a source of power for changes in anode current. It is shown in Chapter 13 that negative resistance may be used to maintain self-oscillation in a circuit.

If the current to the screen is plotted as well as the anode current, it is found that the two currents are complementary, varying equally and oppositely, as shown in Fig. 6.13. The total cathode

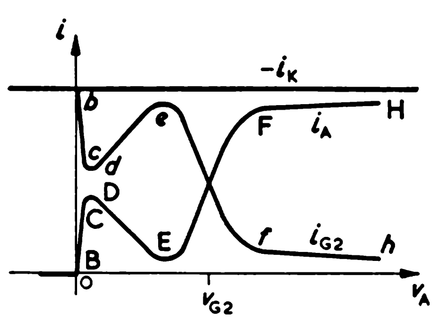


FIG. 6.13

current i_K is also shown in this figure; i_K is found to be free from "kink" and to be nearly constant, rising very slowly with anode voltage. In some cases it is found that the anode current actually reverses, the point C in Fig. 6.12 being below the v_A-axis. The nature of these characteristics may be explained in terms of the emission of secondary electrons from the anode.

6.9. Secondary Emission and Tetrode Characteristics

Electrons striking a surface with sufficient energy may knock out secondary electrons. Within limits, the higher the velocity of the impinging electrons, the greater is the number of secondary electrons (see Fig. 4.8). In a triode the electrons striking the anode may cause secondary emission. However, since the control grid is usually negative and the anode positive the emitted electrons are pulled back to the anode by the electric field, and the net effect on the anode current is negligible. In a tetrode, however, the screen is at a positive potential, and the secondary electrons emitted from the anode may go to the screen, thereby causing a net reduction in anode current. Whether a secondary electron returns to the anode or goes to the screen depends mainly on the relative values of the anode and screen voltages. This may be explained by reference to Fig. 6.14, which shows how the potential distribution in a tetrode varies with the anode voltage, when the control grid and screen voltages are kept constant. When $v_A > v_{G2}$ (curve 1) the potential gradient between the screen and the anode is such that the force on an electron in the screen–anode space is towards the anode. Under these conditions a secondary electron usually returns to the anode. When $v_A < v_{G2}$ (curve 2) the force on an electron is towards the screen, and secondaries from the anode are collected by the screen. The shape of the characteristics in Fig. 6.13 may now be explained as follows. Electrons emitted from the cathode pass through the negative control grid to the screen. Most of them pass through the screen with high velocity, and their subsequent behaviour depends on the anode voltage. When the anode voltage is zero the electrons gradually slow down and come to rest at the anode (if initial velocities of emission from the cathode are neglected). They are not collected by the anode but return to the screen and the screen current is large (b). When the anode is made slightly positive the electrons may reach it and i_A increases rapidly (BC) with a corresponding reduction in i_{G2} (bc). However, at v_A equal to about 5 V, some of the electrons reaching the anode produce secondary electrons.

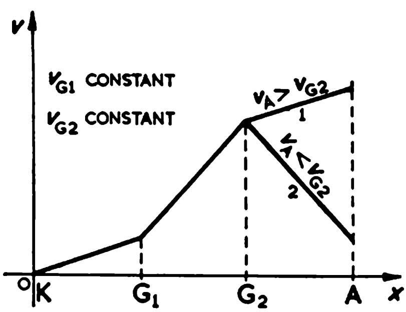


FIG. 6.14

Since $v_A < v_{G2}$, these secondaries go to the screen and reduce i_A, as shown by the reduction in the rate of increase of i_A (CD). As v_A is increased further, still more electrons reach the anode, but with greater energy and giving rise to a greater proportion of secondaries. When D is reached the rate of increase in primary electrons equals the rate of increase of secondaries. From D to E the increase in secondaries exceeds the increase in primaries, and i_A decreases. However, at E the anode voltage is becoming comparable with the screen voltage and some of the secondaries return to the anode. From E to F an increasing proportion of the secondaries return to the anode. When F is reached v_A exceeds v_{G2} sufficiently to ensure that no anode secondaries go to the screen. Throughout these variations of anode voltage, the initial velocities of the secondaries assist the electrons to reach the screen. At the same time the space charge due to primary and secondary electrons prevents some of the slow velocity secondaries from reaching the screen, even when the screen voltage exceeds the anode voltage. These factors account for the absence of sudden change when $v_A = v_{G2}$.

During the variation of i_A, i_{G2} varies in the opposite direction, confirming that the secondary electrons leaving the anode go to the screen. The total current i_K remains practically constant throughout, thus confirming that the anode voltage has very little effect on the current leaving the cathode. We see that the peculiar variations in the anode characteristics of a tetrode at low anode voltages are due to secondary emission. At higher anode voltages when the secondary emission effects are not noticeable the anode current varies very little with anode voltage (FH). The secondary emission coefficient δ, i.e., the ratio of the number of secondary to the number of primary electrons, may exceed unity. Thus it is possible in a tetrode for the anode current to be negative.

If we define the amplification factor μ_A in terms of

$$\mu_A = -\frac{\partial v_A}{\partial v_{G1}} \text{ at constant } i_A \text{ and } v_{G2},$$

then μ_A is high; μ_A is a measure of the relative effectiveness of the anode and control-grid voltages in controlling the anode current. Just as in the triode, μ_A may be interpreted electrostatically as the ratio of v_A to v_{G1} to give zero field at the cathode in the absence of space charge. Now there is an additional electrode, the screen, between the anode and the cathode, and the penetration of the anode field through to the cathode is very small. Obviously, μ_A for a tetrode is greater than μ for a triode. A second amplification factor μ_S may be defined for a tetrode as

$$\mu_S = -\frac{\partial v_{G2}}{\partial v_{G1}} \text{ at constant } i_A \text{ and } v_A.$$

μ_S measures the relative effectiveness of the control grid and the screen in controlling the anode current. Values of μ_S are similar to the values of triode amplification factors.

As in the case of a triode, the idea of an equivalent diode may be used with tetrodes. The anode voltage of the equivalent diode is given by $v_{G1} + v_{G2}/\mu_S + v_A/\mu_A$ and the anode current is given approximately by the expression

$$i_A = K(v_{G1} + v_{G2}/\mu_S + v_A/\mu_A)^{3/2},$$

where K is constant.

6.10. Effect of Space Charge in Tetrodes—Beam Tetrodes

In considering the potential distribution in the screen-anode space in a tetrode in Fig. 6.14 the effect of space charge is neglected. This assumption may be justified when the anode voltage is large, since the electrons are then moving with high velocity between the screen and the anode and the charge density is small. However, when the anode voltage is comparable with or less than the screen voltage the conditions are similar to those described for the grid-anode space of a triode in Section 6.7. There it is seen that a potential minimum may arise due to the space charge. Thus, instead of the potential distributions given in Fig. 6.14, it is possible to obtain conditions as shown in Fig. 6.15, particularly if there is a large distance between the screen and the anode.

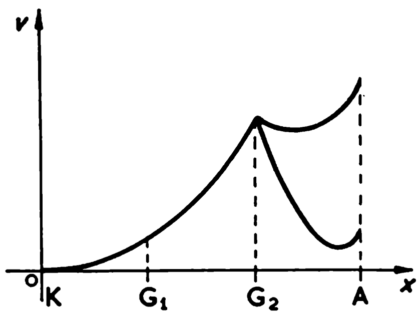


FIG. 6.15

Now the potential variation near to the anode, even at low values of anode voltage, gives a force on electrons towards the anode. As a result the secondary electrons return to the anode instead of going to the screen, and the "kink" is removed from the anode characteristics. The greater

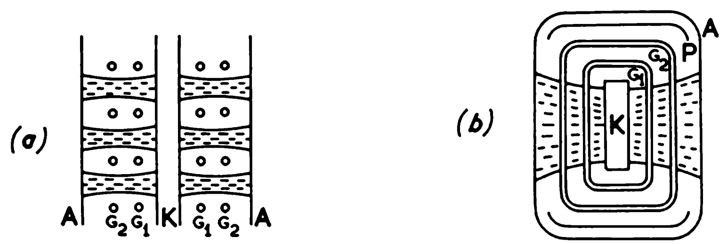


FIG. 6.16

the density of the space charge between the screen and the anode, the more effective is the "suppression" of secondary electrons. As well as utilizing large spacing to produce low potential minima, it is possible to gain the same effect by concentrating the electrons into high-density beams. This is achieved in the beam tetrode, which is shown diagrammatically in Fig. 6.16. The control grid and screen wires have identical pitch, and they are carefully alined so that the electrons are formed into beams as shown in the vertical section (Fig. 6.16.*a*). The width of the beams in the screen–anode space is restricted by two zero potential plates

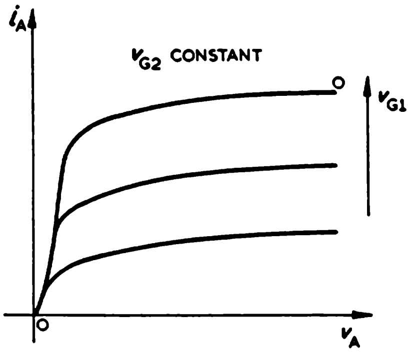


FIG. 6.17

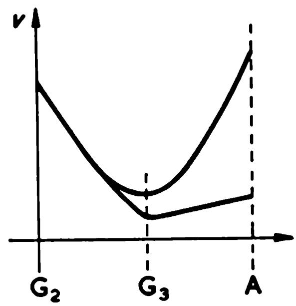


FIG. 6.18

P as shown in the horizontal section (Fig. 6.16.*b*). Anode characteristics for a beam tetrode are shown in Fig. 6.17, in which it can be seen that the " kinks " have been eliminated.

6.11. Pentodes

The effects of the secondary emission from the anode on the characteristics of a tetrode can be eliminated by the insertion of an additional grid between the screen and the anode, as is done in the pentode. This additional grid is usually maintained at zero potential, i.e., cathode potential, and then the potential distribution between the screen and the anode is as shown in Fig. 6.18. The space between the wires of this new grid is at a potential somewhat above zero when the screen and the anode are at positive potentials. With this arrangement the field outside the anode, even at low anode voltages, is such that secondary electrons return to the anode. The extra electrode, G_3, is usually called a suppressor grid. The pentode with five electrodes has a large number of possible characteristics. The most important are those showing the variation of i_A with v_{G1} and v_A, whilst v_{G2} and v_{G3} are constant, the latter usually being equal to the cathode potential. A typical set of anode characteristics for a pentode is shown in Fig. 6.19. These are very similar in shape to the anode characteristics of the beam

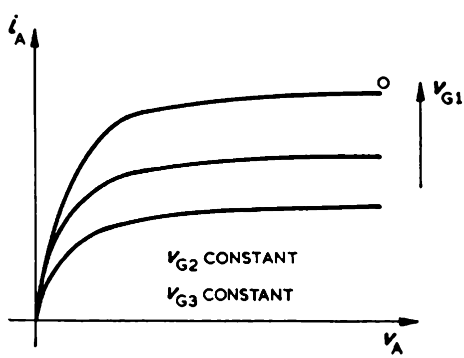


FIG. 6.19

tetrode. The two types of valve are almost interchangeable for many purposes. Careful comparison of the characteristics of pentodes and beam tetrodes of similar size shows that the "knee" of the anode characteristics is sharper and steeper with tetrodes. The pentode knee may sometimes be sharpened by operating the suppressor at a small positive voltage. Historically, pentodes were introduced before beam tetrodes.

Another type of pentode characteristic which is of some interest is shown in Fig. 6.20, which gives the variation of anode and screen currents with suppressor voltage. Normally the suppressor is maintained at zero

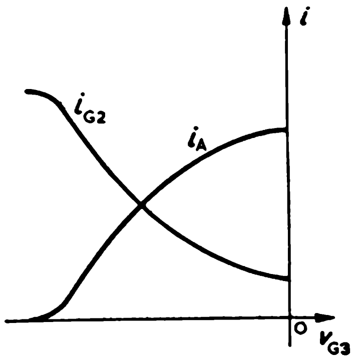


FIG. 6.20

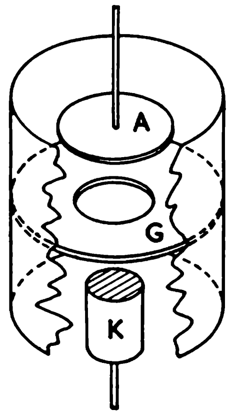


FIG. 6.21

potential relative to the cathode. Then the potential in the rest of the space between the screen and the anode is maintained positive by the combined effects of the screen, anode and suppressor potentials. If the suppressor is made negative a small region round each suppressor wire is at a negative potential and some of the electrons, which previously passed through the suppressor to the anode, now return to the screen. As the suppressor is made more negative the anode current decreases, and becomes zero when the whole space between the suppressor wires is reduced to negative or zero potential. It is seen from the figure that i_A and i_{G2} vary equally and oppositely and that their sum is constant. One feature of these curves is that i_{G2} varies in the reverse direction to v_{G3}, an increase of v_{G3} giving a decrease in i_{G2}, and vice versa. Thus the slope of the curve of screen current against suppressor voltage has a negative region (see Section 13.8).

6.12. Effects of Gas in Triodes—Thyratrons

In Chapter 5 it is shown that the presence of gas in diodes may modify the flow of current considerably. In particular, it is seen that there is an increase in anode current when the anode voltage approaches or exceeds the ionization potential of the gas. As the voltage increases further there is a sudden and very large increase in current. Similar effects are observed in the anode current of triodes with gas. At the same time grid currents are found to flow even when the grid potential is negative. The geometrical arrangement of the electrodes in gas triodes is usually different from the vacuum triodes which have already been considered. Fig. 6.21 shows one type of gas triode. The cathode

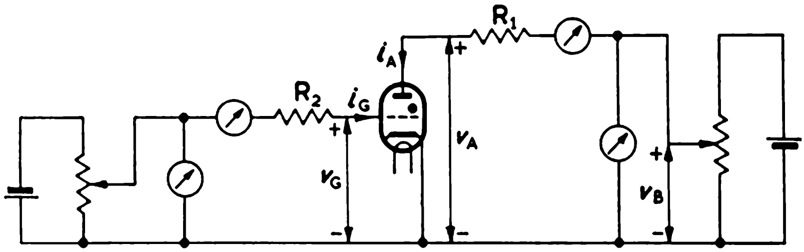


FIG. 6.22

and anode are approximately planar, but the "grid" is a cylindrical shield round the anode and cathode with a central disk between them. There is a hole in the disk. This structure differs considerably from the grid of wires of vacuum triodes, but it is obvious that the field at the cathode is determined mainly by the "grid" potential and to a small extent by the anode potential. It is a high μ structure. The behaviour of this structure may be determined with the circuit shown in Fig. 6.22. Consider the case where the cathode is heated, the anode voltage is, say, 100 V and the grid is sufficiently negative to cut off the anode current. As the negative grid voltage is reduced gradually, a point is reached where the anode current changes suddenly from zero to a large value. At the same time a glow discharge appears inside the valve. It is then found that further variation of the grid voltage has little effect on the anode current, even if the grid is made more negative. Once the discharge is struck, the grid loses control and the discharge may be extinguished only by reduction of the anode voltage to a low value less than about 10 V. If this experiment is repeated with a different initial value of the anode voltage, it is found that the valve strikes at some other value of grid voltage. The curve in Fig. 6.23 gives the corresponding values of anode and grid voltages at which the valve strikes. This curve is sometimes called the control characteristic.

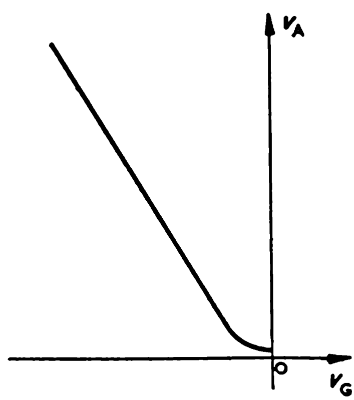


FIG. 6.23

If, on switching on, the values of v_G and v_A correspond to a point on or above the full curve in the figure the valve strikes. As with the diode, it is found that in the conducting state the voltage drop v' across the valve between anode and cathode remains practically constant at $v' \simeq 10$ V. The current that flows is given by $(v_B - v')/R$. The greater part of the control characteristic is a straight line. The slope of this line is called the control ratio; its value is usually about 20. The anode current after the discharge is struck is unaffected by the grid voltage. However, there is a current flowing to the grid itself, and this varies with the grid voltage, as shown in Fig. 6.24. At positive and low negative grid voltages this current flows in the positive direction, i.e., it corresponds to a flow of electrons from the gas to the grid. The current increases very rapidly as the grid voltage is made less negative. At some small negative grid voltage the grid current is zero. As the grid voltage is made more negative a reversed current flows to the grid, as shown.

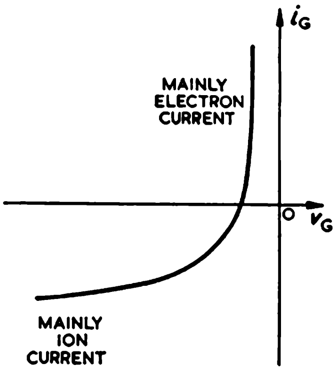


FIG. 6.24

The nature of these characteristics may be explained in the following manner. On switching on with the grid voltage sufficiently negative there is no anode current. The limiting value of v_G for cut-off, when the anode voltage is v_A, is $v_G = -v_A/\mu$. As soon as v_G is made less negative than the limiting value, electrons leave the cathode and move towards the anode. Normally, v_A is considerably greater than the ionization potential of the gas and breakdown occurs as described in Sections 5.15 and 5.16. The voltage between anode and cathode drops to about the ionization potential and most of the space is filled with plasma. A positive-ion sheath forms round the cathode just as in the diode. With the grid at a negative potential, positive ions are attracted to it from the plasma, and so a positive ion sheath is formed round the grid. The size of this sheath depends on the grid potential; the more negative the grid, the greater the number of positive ions surrounding it. Equilibrium is reached when the ion sheath just neutralizes the negative grid so that no more ions are drawn from the plasma. Under these conditions the main plasma is unaffected by the grid potential, and we see why the grid voltage does not affect the flow of current to the anode. However, the ion current to the grid varies with the grid voltage; in use, a resistor is always connected in series with the grid supply to limit this current to a safe value. If the grid voltage is reduced the ion current drops, and some electrons are collected at the grid because of their random velocities. When the grid is positive it collects electrons from the plasma and also some positive ions, as they too have random velocities. It might appear that no current should flow to the grid when it is at zero potential. However, the random velocities of the electrons in the plasma are greater than the ion velocities, and more electrons are therefore collected. This is why it is found that the total grid current is zero at a negative grid voltage (see Fig. 6.24).

From what was said above about the control characteristic it should be expected that its slope would give the electrostatic amplification factor, defined in terms of the ratio of v_A to v_G for cut off. On this basis the straight portion of the characteristic when extrapolated should pass through the origin. The actual extrapolation crosses the grid axis at a slightly negative value. This is probably due to the initial velocities of emission of the electrons from the cathode. Breakdown occurs at very low anode currents. Hence the extra energy of a small number of electrons emitted with appreciable velocity may assist breakdown. The actual characteristic cuts the v_A-axis near the ionization potential.

The main properties of a gas triode or thyratron are similar to those of a relay. The breakdown corresponds to the closing of the relay. The grid operating at a low voltage may be used to close the relay, and to control the flow of a large current. However, once the relay is closed the grid has no further control. Thus thyratrons have to be used in circuits where the anode voltage drops to a low or negative value when the relay is to be opened. When the anode voltage drops below the ionization potential, the positive ions and electrons recombine to form neutral molecules. This takes a definite time, depending on the ion mobilities. If the discharge is to cease completely so that the grid resumes control, it is important that the anode voltage should not change again until the process of recombination or de-ionization is complete. This factor sets a limit to the maximum frequency of operation. In the case of mercury-vapour thyratrons the frequency is limited to a few thousand cycles per second. Hydrogen or inert-gas thyratrons may be used at higher frequencies. As in the case of gas diodes, temperature affects the operation of thyratrons particularly when mercury vapour is used. It is necessary to control the temperature within limits.

Some gas tubes have a fourth electrode whose potential modifies the striking voltage of the control electrode. These gas tetrodes usually have a high control ratio.

So far, all the considerations on gas triodes have been based on tubes with heated cathodes. Gas diodes can operate with cold cathodes, where the ionization is initiated by cosmic rays or other external agency. In such tubes the discharge between two electrodes may be controlled by the insertion of a third electrode. When a discharge starts between one pair of electrodes it initiates a discharge to another pair. By biasing the first pair near to the firing voltage, a small voltage change may start a discharge to the second pair, thus giving switch or relay action as in the thyratron. Cold-cathode triodes and tetrodes are used in this way; they have the big advantage over thyratrons that they do not require any supply for heating the cathode. Multi-electrode cold-cathode tubes have been introduced for counting purposes (see Section 18.10).

6.13. Ionization Gauge

In valves which have been evacuated to a very low pressure there are still many molecules of gas present. Indeed, in the best vacua available the number of gas molecules usually exceeds the number of electrons even when large currents flow. Thus there is always some ionization taking place when the electrode potentials are sufficiently high. In valves operated with a negative grid, the positive ions produced by ionization flow to the grid. This positive-ion current gives a measure of the density of the gas molecules or the gas pressure. In this way a negative grid tube may furnish an indication of its own pressure. Frequently small triodes are incorporated in vacuum systems to indicate the pressure. Such triodes are called ionization gauges. At low pressures the ion current is proportional to the pressure.

6.14. Transistors

Crystal triodes or transistors are made from suitably prepared crystals of germanium or silicon. As in the case of crystal diodes, there are two main types—the junction transistor and the point-contact transistor.

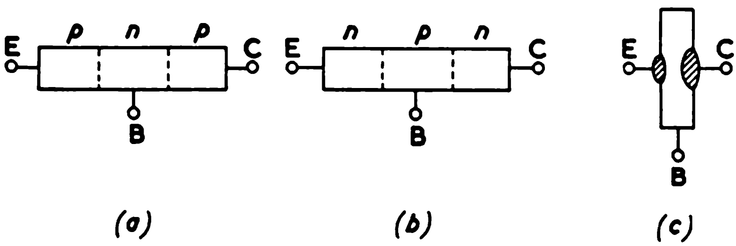


FIG. 6.25

The junction transistor consists essentially of two p–n junctions combined in one crystal as shown diagrammatically in Fig. 6.25.*a* and *b*, where it is seen that there are two main forms, depending on whether the middle section is p- or n-type material. In both cases the middle section is called the base, and it is made as thin as possible for reasons given in Section 6.17. The outside regions are called the emitter and the collector. One form of p–n–p junction transistor is shown in Fig. 6.25.*c*, in which two beads of indium are on opposite sides of a thin slice of n-type germanium. The p-type regions are formed at the boundary of the indium and the germanium. In the point-contact transistor, Fig. 6.26, the

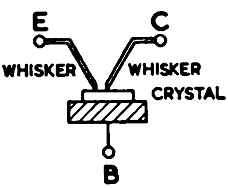


FIG. 6.26

emitter and collector are two metal wire cat's whiskers, very close together and making contact with the surface of a slice of germanium (or silicon) which acts as the base. Various symbols are used to represent transistors. Some of these are shown in Fig. 6.27.*a* to *f*; (*a*) and (*b*) are alternative methods of representing *p–n–p* junction transistors, and (*c*)

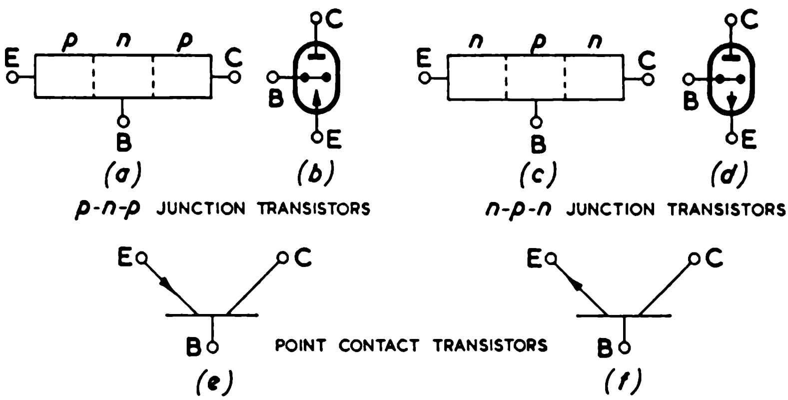


FIG. 6.27

and (*d*) are the corresponding symbols for *n–p–n* junction transistors. The two point-contact symbols shown at (*e*) and (*f*) are sometimes used for junction transistors also.

The transistor is a three terminal device, and its properties may be specified in terms of characteristic curves connecting the three currents and three voltages which are shown in Fig. 6.28. Since

$$i_E + i_B + i_C = 0$$

and

$$v_{BE} + v_{EC} + v_{CB} = 0$$

there are only four independent variables. If i_E, i_C, v_{EB} and v_{CB} are chosen, the two voltages are measured from the base, and these variables

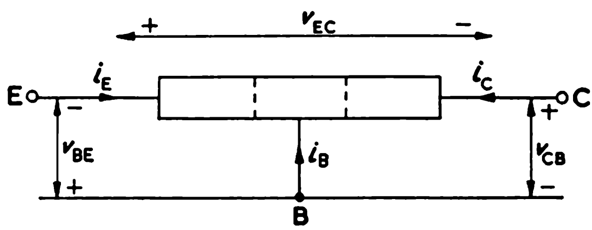


FIG. 6.28

belong to what is called a common-base arrangement. In the case of the triode there are also four independent variables i_A, i_G, v_A and v_G, but as i_G is zero in many applications, we represent the triode characteristics by a single equation $i_A = f(v_G, v_A)$ and a single family of curves. In the transistor there is no such simplification, and all four variables must be taken into account. One way of expressing the characteristics is

$$v_{EB} = f_1(i_E, v_{CB})$$

and

$$i_C = f_2(i_E, v_{CB}).$$

These are known as the hybrid characteristics. Two separate families of curves are required to represent the complete characteristics. Typical curves for a *p–n–p* junction-type transistor are shown in Fig. 6.29.*a* to *d*; (*a*) and (*b*) give in alternative forms families of curves for the first of the

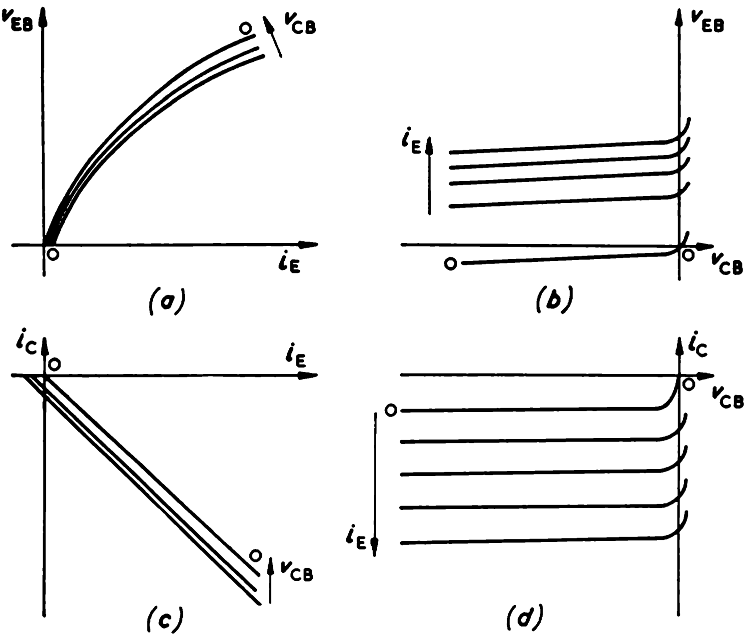


FIG. 6.29

above equations, and (*c*) and (*d*) do likewise for the second equation. The characteristics are completely specified by one pair of families, e.g., (*b*) and (*d*). Important and typical features of these curves are:

(i) i_E and i_C are related linearly, being practically equal and opposite. Actually, i_C is usually slightly less than i_E in magnitude. Consequently i_B is much smaller than either i_E or i_C.

(ii) i_E changes considerably with v_{EB}.

(iii) i_C is almost independent of v_{CB}.

(iv) v_{EB} changes slowly with v_{CB}.

From (i) and (ii) we see that the emitter–base voltage acts as an effective control of the emitter current and of the collector current, but at the same time the base current is small. This corresponds to some extent with conditions in a triode, where the grid acts as a control of the anode current without taking appreciable grid current. The emitter, base and collector are analogous to the cathode, grid and anode respectively.

Point-contact transistor characteristics are similar to those of the junction transistor, but they show greater variations from one sample to another, and the typical features mentioned above are not so marked.

In normal use with a *p–n–p* transistor v_{EB} is a positive voltage of one

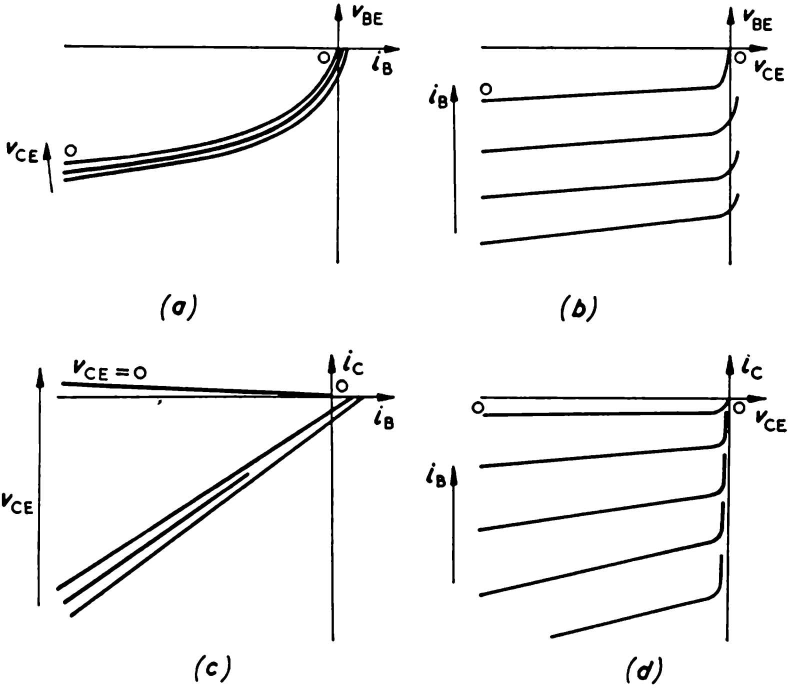


FIG. 6.30

volt or less, whereas v_{CB} is negative and of greater magnitude than v_{EB}. The emitter current is positive and the collector current is negative. All these polarities are reversed for an *n–p–n* transistor.

The common-base characteristics use the four variables i_E, i_C, v_{EB} and v_{CB} as independent. When v_{CE} and i_B are required they are determined from $i_E + i_C + i_B = 0$ and $v_{BE} + v_{EC} + v_{CB} = 0$, i.e., $v_{CE} = v_{CB} + v_{BE}$. Since i_C and i_E are approximately equal but of opposite sign, i_B is the small difference between two relatively large quantities. Thus characteristics based on i_C and i_E do not give the most accurate overall representation. Common-emitter characteristics based on i_B, i_C, v_{BE} and v_{CE} are frequently used. A set of such characteristics is shown in Fig. 6.30.*a* to d. Corresponding functional relationships may be expressed in the form

$$v_{BE} = f_3(i_B, v_{CE})$$

and
$$i_C = f_4(i_B, v_{CE}).$$

6.15. Transistor Parameters

As in the case of triodes and other electronic devices, the slopes of the various characteristics are important parameters. For the common-base connection the hybrid parameters are defined as follows:

$$h_{11} = \frac{\partial v_{EB}}{\partial i_E} \text{ at constant } v_{CB},$$

$$h_{12} = \frac{\partial v_{EB}}{\partial v_{CB}} \text{ at constant } i_E,$$

$$h_{21} = \frac{\partial i_C}{\partial i_E} \text{ at constant } v_{CB}$$

and
$$h_{22} = \frac{\partial i_C}{\partial v_{CB}} \text{ at constant } i_E.$$

The quantity h_{11} has the dimensions of a resistance. It is the resistance of the emitter to changes of the emitter–base voltage when the collector–base voltage is kept constant; it has a low value. It may be considered as a measure of the transistor input resistance when the output voltage is constant. The quantity h_{22} is the conductance between collector and base for changes in v_{CB} while the emitter current is constant. It may be thought of as the output conductance when the input current is constant. It has a very low value. It may be noted in passing that the i_C, v_{CB} characteristics are similar in shape to the i_A, v_A characteristics of pentodes, and the low value of h_{22} corresponds to the low value of $1/r_a$ of a pentode. The quantities h_{12} and h_{21} are dimensionless. The former is a measure of the fraction of the collector–base voltage change that exists between the emitter and base when the emitter current remains constant; h_{12} is usually small. Finally, h_{21} is the ratio of the change in collector current to the change in emitter current when the output voltage is constant. It is usually negative, and its value is just slightly less than unity. It is sometimes called the collector–emitter current amplification factor. Frequently h_{21} is replaced by $-\alpha_{ce}$ and then α_{ce} is usually positive. Point-contact transistors and certain types of junction transistor have values of α_{ce} greater than unity.

The slopes of the common-emitter characteristics may also be used as transistor parameters. They are defined as follows:

$$h_{11e} = \frac{\partial v_{BE}}{\partial i_B}, \quad h_{12e} = \frac{\partial v_{BE}}{\partial v_{CE}},$$

$$h_{21e} = \frac{\partial i_C}{\partial i_B} \text{ and } h_{22e} = \frac{\partial i_C}{\partial v_{CE}}.$$

Then h_{11e} is the transistor input resistance for the common-emitter connection; h_{22e} is the output conductance, h_{12e} is a measure of the voltage fed back from output to input; and h_{21e} is the current amplification factor from base to collector. This quantity is usually large and positive for a *p–n–p* junction transistor; it is sometimes denoted by the symbol α_{cb}.

When it is necessary to distinguish between the h parameters for common-base and common-emitter connections the former may be denoted by h_{11b}, h_{12b}, etc. Corresponding h values for both arrangements for a typical *p–n–p* junction transistor are shown in Table 6.1.

	Common-base.	*Common-emitter.*
Input resistance .	$h_{11b} = 30\ \Omega$	$h_{11e} = 1500\ \Omega$
Voltage feedback factor .	$h_{12b} = 7 \times 10^{-4}$	$h_{12e} = 8 \times 10^{-4}$
Current amplification factor	$h_{21b} = -\alpha_{ce} = -0{\cdot}98$	$h_{21e} = \alpha_{cb} = 50$
Output conductance	$h_{22b} = 1\ \mu$mho	$h_{22e} = 50\ \mu$mho

TABLE 6.1.

Mathematical relations between the h parameters may be established (see Exx. VI). In particular, it may be shown that $\alpha_{cb} \simeq \alpha_{ce}/(1 - \alpha_{ce})$ and $h_{11e} \simeq h_{11b}/(1 - \alpha_{ce})$. Since $\alpha_{ce} \simeq 1$, it follows that the common emitter arrangement gives a considerable increase in input resistance. The input resistance of a transistor in the common-emitter connection is much less than the input resistance of a thermionic triode operated with negative grid. Also, in contrast with the triode, there is always some reaction of the output on the input in the transistor due to h_{12}.

6.16. Transistor Equations for Small Changes

In dealing with the triode in Section 6.3 it is shown that small changes i_a, v_g and v_a in the variables i_A, v_G and v_A are related by the valve equation $i_a = g_m v_g + v_a/r_a$. With the transistor there are similar relations for small changes v_{eb}, v_{cb}, i_e and i_c in the variables v_{EB}, v_{CB}, i_E and i_C. From the equations $v_{EB} = f_1(i_E, v_{CB})$ and $i_C = f_2(i_E, v_{CB})$ it follows that

$$v_{eb} = \frac{\partial v_{EB}}{\partial i_E} i_e + \frac{\partial v_{EB}}{\partial v_{CB}} v_{cb} = h_{11} i_e + h_{12} v_{cb}$$

and
$$i_c = \frac{\partial i_C}{\partial i_E} i_e + \frac{\partial i_C}{\partial v_{CB}} v_{cb} = h_{21} i_e + h_{22} v_{cb}.$$

For the junction transistor h_{12} and h_{22} are small, and the magnitude of h_{21} is nearly unity, so that we get the approximate transistor equations

$$v_{eb} \simeq h_{11} i_e$$

and
$$i_c \simeq h_{21} i_e = -\alpha_{ce} i_e \quad \text{or} \quad i_c \simeq -i_e.$$

From these equations it is seen that v_{EB} is the important voltage; it may be called the driving voltage. A change in the driving voltage gives a change in i_E and an almost equal change in i_C. Changes in v_{CB} and v_{EC} are relatively unimportant in determining the currents flowing. This simplified picture is useful in determining the approximate behaviour of a transistor as long as the changes involved are small.

6.17. Physics of the Transistor

In an isolated *p–n–p* junction transistor there is a potential variation and distribution of energy levels of the form illustrated in Fig. 6.31. In the equilibrium state the Fermi energy level is constant throughout. This state is set up in a similar manner to that described in Chapter 5 for a single *p–n* junction, and involves zero resultant current of electrons and

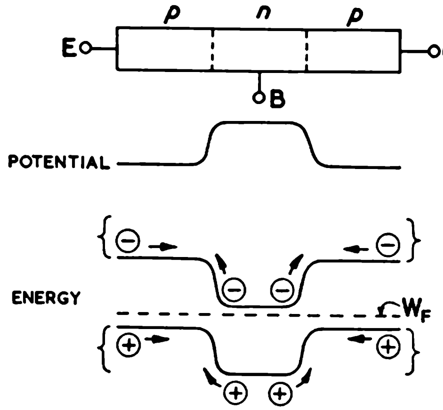


FIG. 6.31

holes across each junction. If the collector is joined directly to the base and a small positive voltage is applied between emitter and base, then a current flows just as with a *p–n* junction diode in the forward direction. Similarly, if the emitter is joined to the base and a negative voltage is applied between collector and base, the current flows as with a *p–n* junction diode in the reverse direction, showing the same saturation of collector current. We consider now what happens when these voltages with the same polarity are applied simultaneously. The potential and energy variations are shown in Fig. 6.32. The potential barrier between emitter and base is reduced and the flow of holes across the barrier is greatly increased. There is an increased flow of electrons from base to emitter, but the hole density is, by design, much greater in the *p*-region than the electron density in the *n*-region, so that the current may be considered as due mainly to holes. The holes enter the *n*-region and diffuse through it. There is a tendency for these holes to combine with the electrons in the region. However, if the *n*-region is sufficiently thin a large number of the holes reach the collector–base junction and very few arrive at the base terminal. At the collector–base junction the holes fall easily into the collector region on account of the field at the junction. Thus, we see that the collector current is very nearly equal to the emitter current and the base current is almost zero. The collector–base voltage has little effect on the current as long as its magnitude is above some minimum

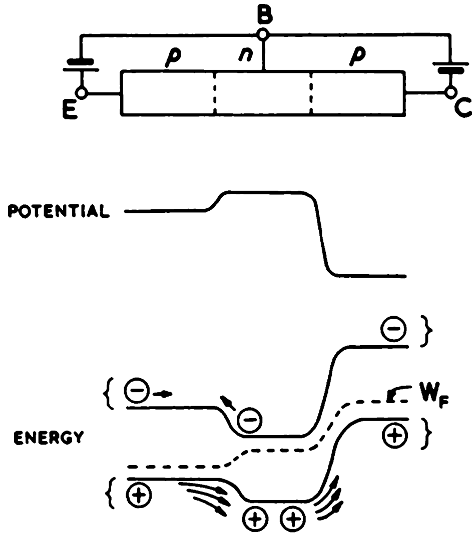


FIG. 6.32

value and the base region is sufficiently thin. A small change in emitter-base voltage causes a change in emitter current. This results in an almost equal change in collector current.

As with the junction diode, there are always some reverse currents due to minority carriers in the different regions. The minority carriers, and hence the reverse currents, increase with temperature. This temperature effect is one of the chief limitations to the use of transistors.

CHAPTER 7

VOLTAGE AMPLIFIERS

7.1. Valves and their Characteristics

In Chapters 5 and 6 the characteristics of various valves are discussed and it is shown that the electrode currents can be determined from the characteristics when the electrode voltages are known. In the use of valves the electrode voltages usually depend on a number of factors, such as the battery supplies and the circuits connected to the electrodes. Sometimes the voltages depend on one another and on the electrode currents. In order to determine the actual electrode voltages these factors must all be taken into account. In many applications of valves with control grids the grid–cathode voltage is negative, so that only a negligible amount of grid current flows. Then it is frequently possible

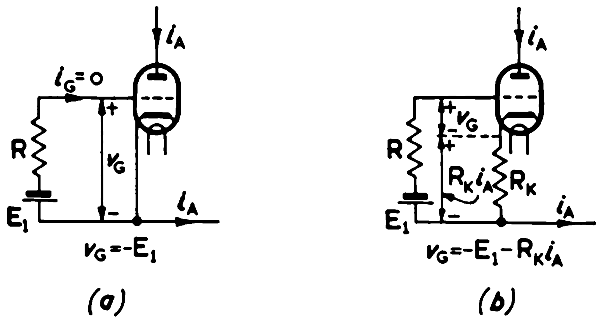


FIG. 7.1

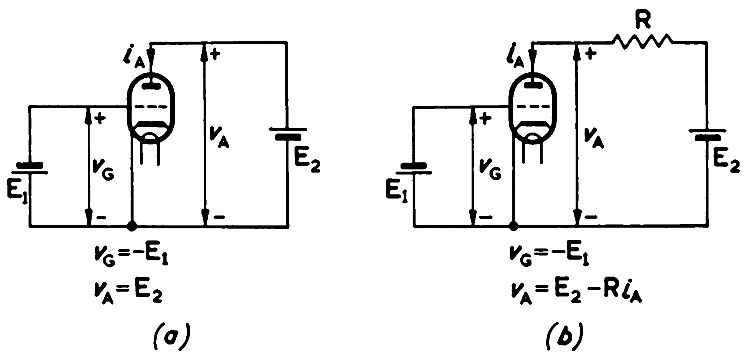


FIG. 7.2

o determine the grid–cathode voltage without having to take into con- ideration the valve itself. Such is the case in the circuit of Fig. 7.1.*a*. Iowever, in Fig. 7.1.*b* the grid–cathode voltage depends upon the value of anode current flowing through the resistance R_K, and hence upon the nature of the rest of the circuit through which the anode current flows. n this case it is not possible to write down directly the value of the grid voltage. Throughout this chapter it is assumed that the grid voltage is negative and the grid current zero.

From Fig. 7.2.*a* the values of v_G and v_A are known directly from the ircuit as $v_G = -E_1$ and $v_A = E_2$. This allows the anode current to be leduced from the static characteristics. In Fig. 7.2.*b*, however, where here is a resistance in the anode circuit, only the value of v_G is known

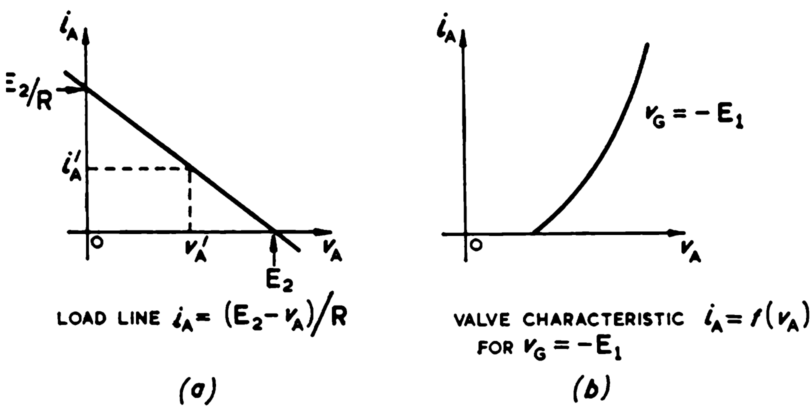


FIG. 7.3

lirectly from the circuit. The determination of v_A involves the anode current through the valve, which itself requires a knowledge of v_A before t can be determined. There are in fact two relations to be satisfied: the valve characteristic

$$i_A = f(v_A) \text{ for } v_G = -E_1,$$

und the circuit condition

$$v_A = E_2 - Ri_A \quad \text{or} \quad i_A = (E_2 - v_A)/R.$$

The values of i_A and v_A for the circuit of Fig. 7.2.*b* must satisfy these equations simultaneously.

7.2. The Load Line

We consider, firstly, the circuit equation

$$v_A = E_2 - Ri_A.$$

When $$i_A = 0,\ v_A = E_2.$$

The relation between i_A and v_A is a linear one, since E_2 and R are constant, and only one other point is required for drawing the graph of the equation. A suitable point is

$$v_A = 0 \text{ and } i_A = E_2/R.$$

Any other point, such as

$$v_A = v_A' \text{ when } i_A = (E_2 - v_A')/R,$$

could be used. The circuit relation is represented by the straight line in Fig. 7.3.*a*. This is known as the load line and R as the load resistance. The equation

$$i_A = (E_2 - v_A)/R$$

is called the Load Line Equation. The slope of the load line is $-1/R$. The static characteristic corresponding to $i_A = f(v_A)$ is shown in Fig. 7.3.*b*. The actual values of i_A and v_A are determined from the intersection at Q of the curves of Fig. 7.3*a*. and 7.3.*b* when they are plotted together as in Fig. 7.4.

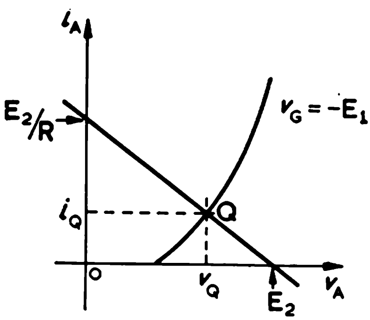


FIG. 7.4

Now let us consider a d.c. change v_g in the grid voltage so that $v_G = -E_1 + v_g$, as shown in Fig. 7.5. The load line is unchanged, but there is a new valve characteristic $i_A = f(v_A)$ for $v_G = -E_1 + v_g$. The anode voltage and current are changed to the values v_P and i_P, where the point P gives the intersection of the load line and the new characteristic. Thus, due to a change in the grid voltage, we have changes in the anode current and voltage of $i_a = (i_P - i_Q)$ and $v_a = (v_P - v_Q)$. It may be seen that v_a is, in this case, negative in value. In many cases v_a is much greater in magnitude than v_g, so that the triode and its resistance load

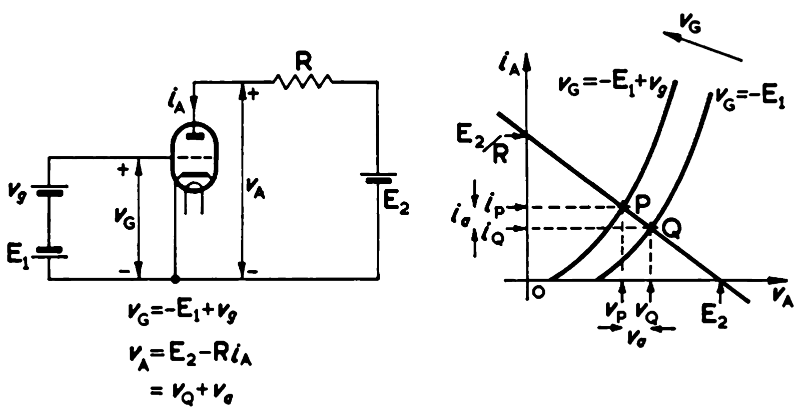


FIG. 7.5

act as a voltage amplifier. The voltage amplification is defined as $A = v_a/v_g$, and the magnitude of this voltage amplification is often referred to as the stage gain. It may be noted that this voltage amplification has been obtained without expending any power in the grid circuit, i.e., no current has been drawn from the source of the voltage v_g.

The point Q in Fig. 7.4 or Fig. 7.5 which gives the operating conditions before any changes are introduced is called the quiescent point, and i_Q and v_Q are the quiescent anode current and voltage respectively. When

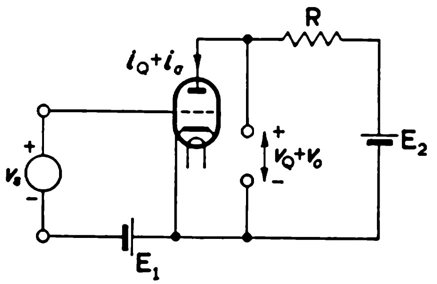
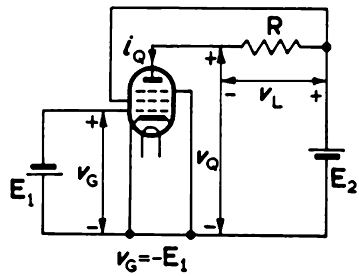
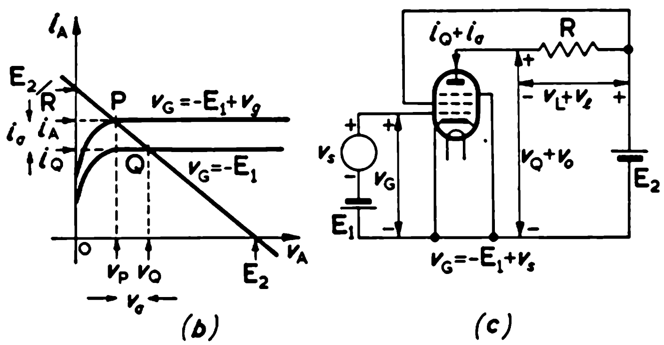


FIG. 7.6



(a)

FIG. 7.7



(b) (c)

FIG. 7.7

it is necessary to distinguish between the quiescent values of the anode voltage and grid voltage, the symbols v_{AQ} and v_{GQ} may be used. In the present case $v_{GQ} = -E_1$. The terminals across which the voltage change v_g is introduced are referred to as the input terminals, whilst those across which the amplified voltage change occurs are the output terminals. The input voltage change is sometimes called the signal voltage, and is given the symbol v_s. The output voltage change is given the symbol v_o, so that the voltage amplification may also be defined as $A = v_o/v_s$ (see Fig. 7.6).

What has been described above for a triode amplifier applies equally well to a pentode amplifier. The appropriate conditions are shown in Fig. 7.7.*a*, *b* and *c*.

7.3. Voltage Amplifier—Small Signal Theory

In the previous section no restrictions are placed on the size of the changes in voltages and currents, except that the grid voltage does not become positive. In this section only small changes are considered. The changes are sufficiently small for the values of g_m, r_a and μ to remain constant throughout the range of variation, so that the Valve Equation derived in Section 6.3 may be used. A change v_g is made in the grid voltage (see Fig. 7.8.*a* and *c*), and it gives rise to changes i_a and v_a in the anode current and voltage. Then the Valve Equation gives $i_a = g_m v_g + v_a/r_a$. The change in anode voltage may also be related to the change in

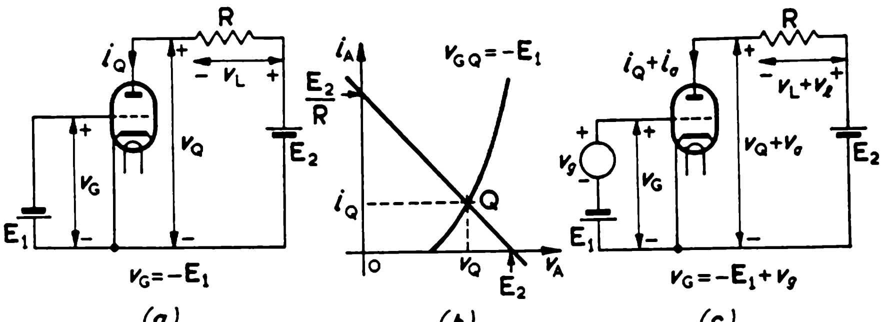


FIG. 7.8

anode current, using the Load Line Equation $i_A = (E_2 - v_A)/R$. In this case $i_a = -v_a/R$. The two equations for i_a yield

$$v_a = -\frac{g_m v_g}{(1/R + 1/r_a)}$$

and the voltage amplification is

$$A = -\frac{g_m}{(1/R + 1/r_a)}.$$

This expression may be written in the alternative forms

$$A = -\frac{g_m}{(1/R + 1/r_a)} = -\frac{g_m}{(G + g_a)} = -\frac{\mu R}{(R + r_a)},$$

where $G = 1/R$ and using $\mu = g_m r_a$. The values of μ, g_m and r_a used in these formulae must be determined for the actual Q-point in use, as shown in Fig. 7.8.*b*. The above formulae all apply to small d.c. signals for any valve amplifier. In the case of pentodes r_a is very large and frequently $r_a \gg R$, so that the voltage gain is simply

$$A = -g_m R.$$

In triodes r_a is usually comparable with R, and smaller voltage gain is obtained than with a pentode for the same mutual conductance and load resistance. It is to be noted that Ri_a is the change in voltage drop v_l across the load resistance, and as $v_a = -Ri_a$, then $v_a = -v_l$. The voltage change across the valve is equal and opposite to the voltage change across the load. This relation is obvious from Kirchhoff's Second Law but it is a useful check in determining the signs of the circuit voltages.

7.4 Voltage Amplifier—Small a.c. Signal

The discussion of the amplifier has so far been in terms of a d.c. signal. When the value of the signal is changing all the time, as with an a.c.

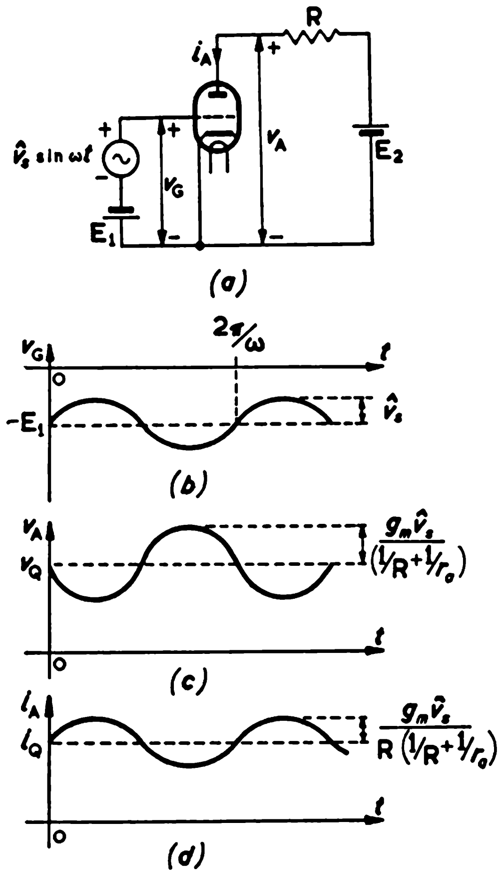


FIG. 7.9

voltage, then as long as the change is not too rapid it is found that the values of the valve parameters corresponding to the instantaneous values of the electrode voltages are identical with those found from the static characteristics. This implies that the static characteristics also apply to instantaneous values of applied voltages. When the time of transit of the electrons through the valve is comparable with the period of the a.c. signal, this is no longer true. However, this source of error can be ignored at frequencies below some tens of megacycles per second.

In the circuit of Fig. 7.9.*a* a sinusoidal alternating signal is applied in series with the grid bias supply, and the load is again a resistance R. The instantaneous value of the signal is $v_s = \hat{v}_s \sin \omega t$, and thus the instantaneous grid voltage is $v_G = -E_1 + \hat{v}_s \sin \omega t$. This is represented in Fig. 7.9.*b*, which also serves to define the polarity of the signal voltage with respect to the bias battery at any given time. When the signal voltage is zero, the anode current and anode voltage have their quiescent values, i_Q and v_Q, as given by the load-line construction. At any other time the instantaneous change of grid voltage is $\hat{v}_s \sin \omega t$. Provided $\hat{v}_s$ is sufficiently small, the changes of anode voltage and current at the same instant may be found from the Valve and Load Line Equations, as in the previous section. Thus

$$v_a = -\frac{g_m \hat{v}_s}{\left(\frac{1}{R} + \frac{1}{r_a}\right)} \sin \omega t = \frac{g_m \hat{v}_s}{\left(\frac{1}{R} + \frac{1}{r_a}\right)} \sin (\omega t + \pi)$$

and

$$i_a = \frac{g_m \hat{v}_s}{R\left(\frac{1}{R} + \frac{1}{r_a}\right)} \sin \omega t.$$

The actual anode voltage and current are

$$v_A = v_Q + v_a \quad \text{and} \quad i_A = i_Q + i_a.$$

These are shown in Fig. 7.9.*c* and *d*. It is seen from the equations above that there is a phase difference of 180° between the a.c. components of the anode current and anode voltage. The voltage amplification is again

$$A = -\frac{g_m}{(1/R + 1/r_a)}$$

as for the d.c. signal, and it is independent of the frequency of the signal.

7.5. Valve Equivalent Circuits for Small Signals

In Fig. 7.10.*a* is shown a valve amplifier with a load resistance R. When a signal is applied we are usually interested only in the changes that take place. We may therefore redraw the circuit as shown in Fig. 7.10.*b*, in which the d.c. supplies E_1 and E_2 have been omitted and only the varying components of voltage and current are included. It is assumed that E_1 and E_2 have negligible impedance to the varying currents. For small signals the Valve Equation gives

$$i_a = g_m v_g + v_a/r_a.$$

The circuit conditions give

$$v_a = -Ri_a.$$

In Fig. 7.10.*c* a circuit has been drawn in which a current generator $g_m v_g$ is in parallel with a resistance r_a and a load R. It may be seen that the same two equations hold for this circuit as for the amplifier. The circuit of Fig. 7.10.*c* is known as the current-generator equivalent circuit of the amplifier. The valve is equivalent to a current generator $g_m v_g$ in parallel with a resistance of r_a.

The Valve Equation for Fig. 7.10.*b* can be rewritten as

$$r_a i_a = (g_m r_a)v_g + v_a = \mu v_g + v_a,$$

whilst the circuit relation is still $v_a = -Ri_a$. If we study the circuit of Fig. 7.10.*d*, where a voltage generator μv_g is in series with a resistance

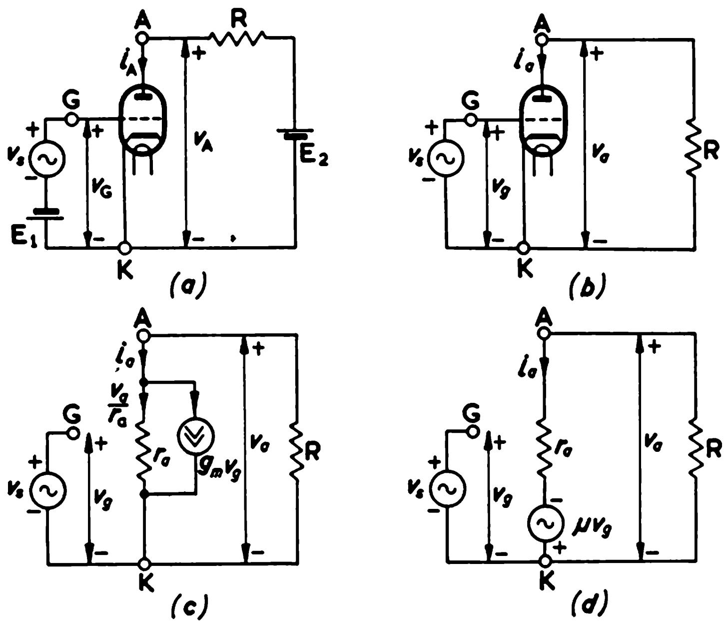
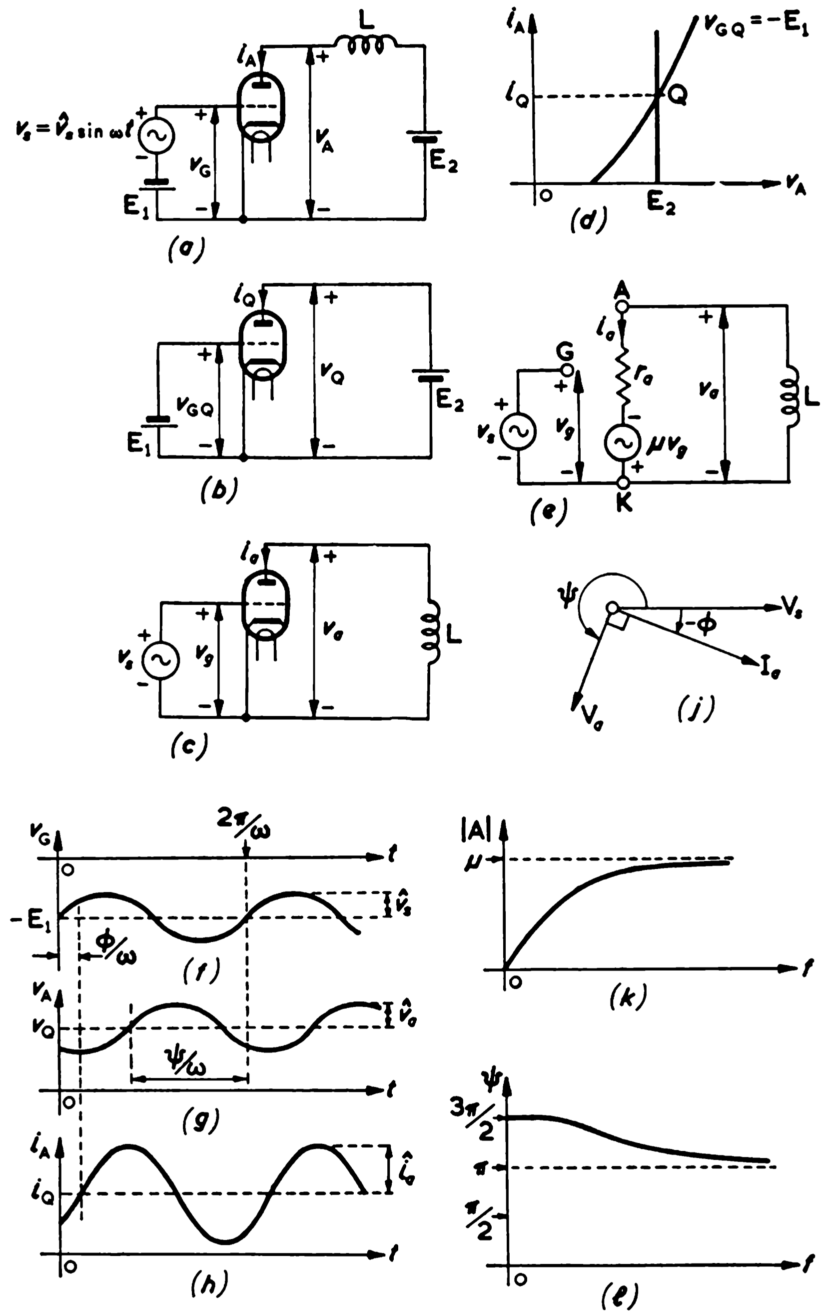


FIG. 7.10

r_a and the load R, it is seen that these two equations hold also for this circuit. The circuit of Fig. 7.10.*d* is known as the voltage-generator equivalent circuit of the amplifier. The valve is equivalent to a voltage generator which has an internal series resistance of r_a. These equivalent circuits are useful in solving amplifier problems, and they are used frequently throughout this book. It must be stressed that the equivalent circuits are only alternative expressions of the Valve Equation, and hence are restricted to small signals.

7.6. Voltage Amplifier—Inductive Load

With a resistance load, the amplification is independent of the frequency, and there is a constant phase difference between the anode current and



L
i_A
$v_s = \hat{V}_s \sin \omega t$
v_G
v_A
E_1
E_2
(a)
i_A
$v_{GQ} = -E_1$
i_Q
Q
v_A
E_2
(d)
i_Q
v_{GQ}
v_Q
E_1
E_2
(b)
A
G
i_a
r_a
v_s
v_g
v_a
μv_g
L
K
(e)
i_a
v_s
v_g
v_a
L
(c)
ψ
V_s
$-\phi$
I_a
V_a
(j)
v_G
$2\pi/\omega$
t
$-E_1$
$\hat{v}_s$
ϕ/ω
(f)
v_A
V_Q
$\hat{v}_a$
ψ/ω
t
(g)
i_A
i_Q
$\hat{i}_a$
t
(h)
$|A|$
μ
f
(k)
ψ
$3\pi/2$
π
$\pi/2$
f
(ℓ)

Fig. 7.11

the anode voltage of 180°. Frequently there is some inductance or capacitance associated with the load, and then the amplification and the phase difference vary with the signal frequency. In this section we consider an amplifier with an inductive load with zero resistance, as shown in Fig. 7.11.*a*. In this case the static load line equation is simply $v_A = E_2$ (Fig. 7.11.*b*). This gives a vertical line through E_2, as seen in Fig. 7.11.*d*. Also $v_Q = E_2$. If the grid voltage is given by

$$v_G = -E_1 + \hat{v}_s \sin \omega t$$

then we are primarily interested in the alternating components of the anode current and the anode voltage (Fig. 7.11.*c*). These take the form

$$i_a = \hat{\imath}_a \sin (\omega t + \phi) \quad \text{and} \quad v_a = \hat{v}_a \sin (\omega t + \psi).$$

To find $\hat{\imath}_a$, $\hat{v}_a$, ϕ and ψ in terms of v_s we may use the equivalent circuit with vector voltages and currents. The constant-voltage generator may be used as shown in Fig. 7.11.*e*. By applying Kirchhoff's Laws and using vector algebra we find

$$\mathbf{I_a} = \mu\mathbf{V_s}/(r_a + j\omega L) \quad \text{and} \quad \mathbf{V_a} = -j\omega L\mathbf{I_a},$$

i.e., $\quad \mathbf{I_a} = \mu\mathbf{V_s}(r_a - j\omega L)/(r_a^2 + \omega^2 L^2).$

Hence $\quad \tan \phi = -\omega L/r_a$

and $\quad \hat{\imath}_a = \mu\hat{v}_s/\sqrt{(r_a^2 + \omega^2 L^2)}$

Also, since $\quad \mathbf{V_a} = -j\omega L\mathbf{I_a}$, then

$$\psi = \phi + 3\pi/2$$

and $\quad \hat{v}_a = \mu\omega L\hat{v}_s/\sqrt{(r_a^2 + \omega^2 L^2)}.$

Figs. 7.11.*f*, *g*, *h* and *j* summarize the relations between the total and a.c. values of the various currents and voltages. It may be seen from the equations that the magnitude and phase of the anode voltage vary with the signal frequency. The stage gain is

$$|\mathbf{A}| = \hat{v}_a/\hat{v}_s = \mu\omega L/\sqrt{(r_a^2 + \omega^2 L^2)}$$

When $\omega L \gg r_a$, then $|\mathbf{A}| = \mu$ and is independent of frequency. The variation of $|\mathbf{A}|$ and ψ with frequency are shown graphically in Figs. 7.11.*k* and *l*.

7.7. Voltage Amplifier—Capacitive Load

This case is shown in Fig. 7.12.*a*, where a resistance load R is shunted by capacitance C. With no signal the equilibrium d.c. circuit is shown in Fig. 7.12.*b*, where it is seen that the circuit is exactly like that with resistance load only. The condenser is charged up to a voltage equal to the potential drop across R. The quiescent point Q is found as usual in Fig. 7.12.*d*, and then the appropriate values of g_m, r_a and μ may be determined. If the grid signal is given by

$$v_s = \hat{v}_s \sin \omega t$$

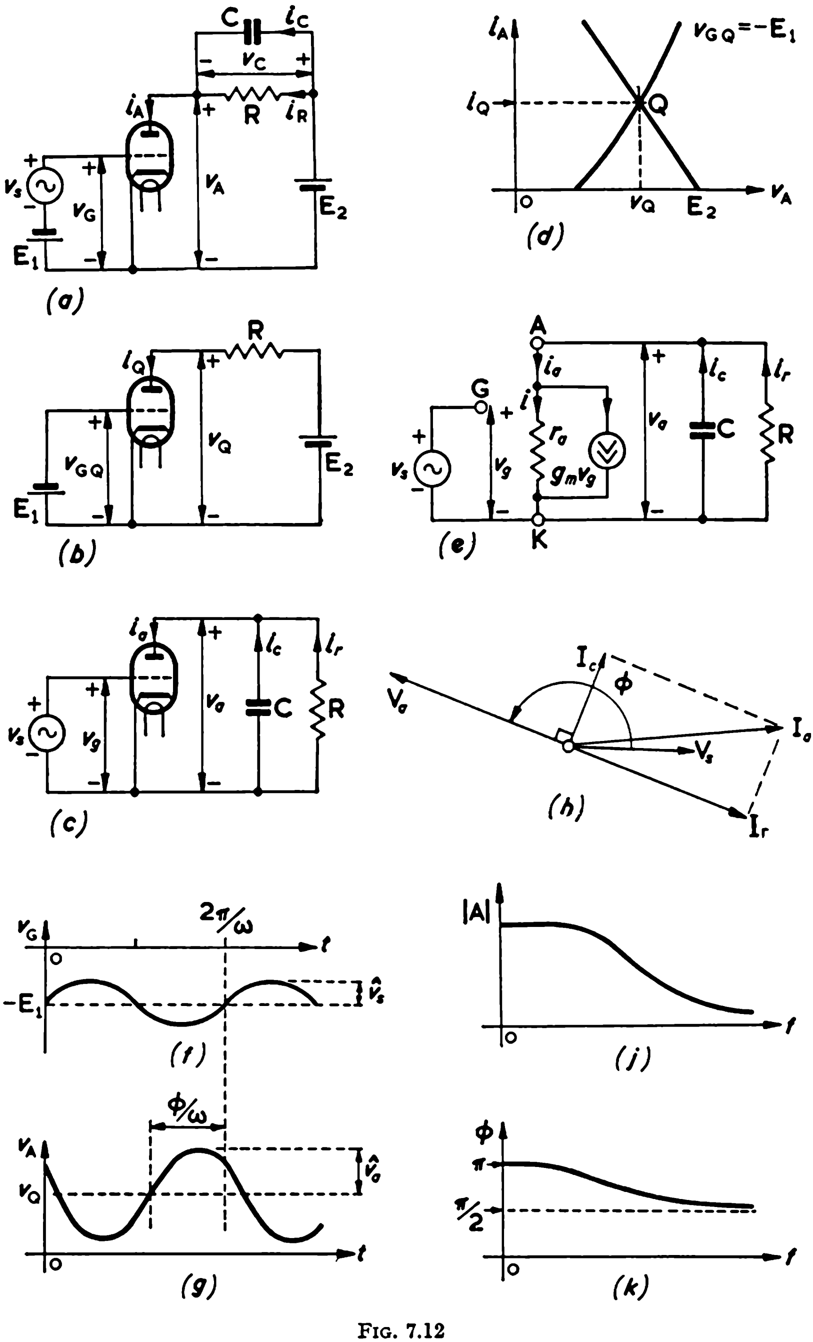

C
i_C
v_C
i_A
R
i_R
v_s
v_G
v_A
E_1
E_2
(a)
i_A
i_Q
Q
$v_{GQ}=-E_1$
0
v_Q
E_2
v_A
(d)
R
i_Q
v_{GQ}
v_Q
E_1
E_2
(b)
A
G
i_a
i
v_s
v_g
r_a
$g_m v_g$
v_a
i_c
C
i_r
R
K
(e)
i_a
v_s
v_g
v_a
i_c
C
i_r
R
(c)
I_c
ϕ
V_a
I_a
V_s
I_r
(h)
$2\pi/\omega$
v_G
t
0
$-E_1$
$\hat{v}_s$
(f)
ϕ/ω
v_A
v_Q
$\hat{v}_a$
0
t
(g)
|A|
0
f
(j)
ϕ
π
$\pi/2$
0
f
(k)

FIG. 7.12

then we assume that it gives rise to an alternating anode voltage of the form

$$v_a = \hat{v}_a \sin(\omega t + \phi).$$

To determine $\hat{v}_a$ and ϕ it is convenient to use the current-generator equivalent circuit, as shown in Fig. 7.12.*e*. With vector notation this gives

$$g_m\mathbf{V}_s = \mathbf{I}_c + \mathbf{I}_r - \mathbf{I}, \quad \mathbf{I}_c = -j\omega C\mathbf{V}_a,$$

$$\mathbf{I}_r = -\mathbf{V}_a/R \text{ and } \mathbf{I} = \mathbf{V}_a/r_a.$$

Hence $\quad \mathbf{V}_a/\mathbf{V}_s = -g_m\Big/\left(\dfrac{1}{R} + \dfrac{1}{r_a} + j\omega C\right),$

i.e., $\quad \mathbf{V}_a/\mathbf{V}_s = -g_m\left(\dfrac{1}{R} + \dfrac{1}{r_a} - j\omega C\right)\Big/\left\{\left(\dfrac{1}{R} + \dfrac{1}{r_a}\right)^2 + \omega^2C^2\right\}.$

Thus $\quad \tan\phi = -\omega C\Big/\left(\dfrac{1}{R} + \dfrac{1}{r_a}\right)$

and $\quad |\mathbf{A}| = |\mathbf{V}_a/\mathbf{V}_s| = g_m\Big/\sqrt{\left\{\left(\dfrac{1}{R} + \dfrac{1}{r_a}\right)^2 + \omega^2C^2\right\}}.$

The variations of the voltages, the stage gain and the phase angle with frequency and a vector diagram are shown in Fig. 7.12.*f* to *k*. The stage gain varies from $\dfrac{g_m}{(1/R + 1/r_a)}$ at very low frequencies to zero at very high frequencies. The gain drops to $1/\sqrt{2}$ of its low-frequency value, i.e., by approximately 3 dB, when $\omega C = 1/R + 1/r_a$.

7.8. Frequency Distortion and Phase Distortion

With reactive anode loads, the stage gain of an amplifier varies with the frequency of the signal. If two signals of equal magnitude but different frequency are applied to the input terminals the output magnitudes are not equal. The amplifier is said to introduce frequency distortion. Also, with a reactive load the phase difference between the input and the output voltages varies with frequency, and the amplifier may have phase distortion. When a signal consists of a combination of several sinusoidal waves of different frequencies, frequency distortion and phase distortion cause the output waveform to differ from the input.

When an amplifier has a resistor as its anode load there is always some capacitance in parallel with it. This may arise from the anode–cathode capacitance of the valve, or from stray capacitance to a metal chassis from the anode and the components attached to it. It is difficult to reduce the effective capacitance across the load resistor below about 10 $\mu\mu$F, and this sets a limit to the upper frequency range of the amplifier.

7.9. Tuned Amplifiers—Parallel Resonant Load

An amplifier with a load consisting of inductance L, capacitance C and resistance R, all in parallel, is shown in Fig. 7.13.*a*. Such a load is purely resistive and of value R at the resonant frequency when $\omega_0 L = 1/\omega_0 C$. At any other frequency the load has some reactance and its impedance is less than R. The performance of the amplifier may be analysed by

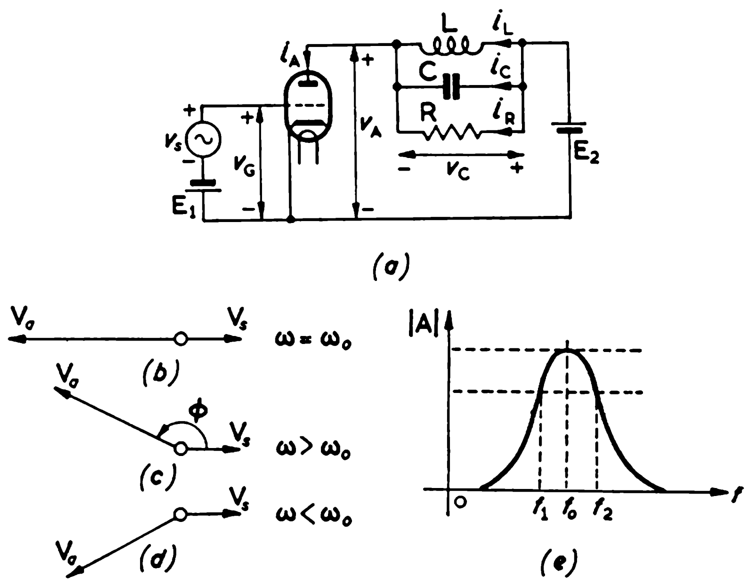


FIG. 7.13

the method used in Section 7.7, the only difference being the extra component L in parallel with the rest of the circuit. It is found that

$$\frac{\mathbf{V}_a}{\mathbf{V}_s} = \frac{g_m\{-(1/R + 1/r_a) - j(-\omega C + 1/\omega L)\}}{(1/R + 1/r_a)^2 + (-\omega C + 1/\omega L)^2}.$$

The output voltage is seen to lead the signal voltage by an angle ϕ where

$$\tan\phi = \frac{-(-\omega C + 1/\omega L)}{-(1/R + 1/r_a)}.$$

ϕ is an angle in the third quadrant as long as $1/\omega L > \omega C$. At the resonant frequency $f_0(= \omega_0/2\pi)$,

$$\omega_0 C = 1/\omega_0 L$$

and $$\mathbf{V}_a/\mathbf{V}_s = -g_m/(1/R + 1/r_a).$$

This confirms that the amplifier behaves at its resonant frequency as if it had a load resistance R. At frequencies below the resonant frequency the performance is similar to that of an amplifier with an inductive load. At higher frequencies it approximates to an amplifier with a capacitive load. The various cases are illustrated in Fig. 7.13.*b*, *c*, *d* and *e*. Such an amplifier may be used for the selective amplification of a signal covering a narrow band of frequencies around f_0. At a frequency f_1 below f_0 at which

$$-\omega_1 C + 1/\omega_1 L = 1/R + 1/r_a,$$

the stage gain is $1/\sqrt{2}$ of the value at resonance. The stage gain has the same value at a frequency f_2 where

$$\omega_2 C - 1/\omega_2 L = 1/R + 1/r_a.$$

The fraction $(f_2 - f_1)/f_0$ is a measure of the sharpness of the resonance curve or of the selectivity of the amplifier. When the voltage across the tuned circuit is $1/\sqrt{2}$ of the resonant voltage, the corresponding power ratio is 1/2. Thus $f_2 - f_1$ is the width of the resonance curve at the " half-power points ", i.e., when the power output of the stage is approximately 3dB below the value at f_0.

7.10. Amplifiers with Several Stages

So far only single valve amplifiers have been considered. Frequently, the voltage amplification required is greater than can be obtained from one valve, and then several stages have to be used. Obviously, the overall

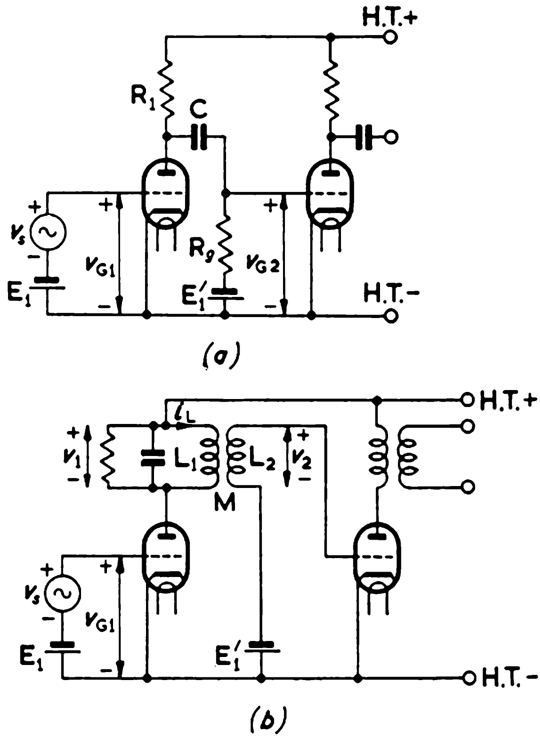


FIG. 7.14

stage gain is the product of the individual gains and the phase shift is the sum of the phase shifts of each stage. In coupling the output of one stage to the input of the next, care must be taken to ensure that the electrode d.c. voltages are correct. In a.c. amplifiers the stages are normally isolated from one another for d.c. by one of the two methods shown in Fig. 7.14. The condenser coupling of Fig. 7.14.*a* is usual in voltage amplifiers at low frequencies, whilst the mutual inductance coupling is commonly used at high frequencies with tuned amplifiers, and sometimes with low-frequency power amplifiers (see Section 8.5).

In the condenser-coupled amplifier the d.c. grid voltage for the second valve is obtained from the battery E_1' through the resistance R_g, which is called a grid leak. In normal use the grid voltage does not become positive and no d.c. current flows through R_g, so that its presence does not affect the grid bias voltage. Over the working frequency range of the amplifier it is desirable that $R_g \gg R_1$ and $1/\omega C \ll R_g$. If both of these conditions are satisfied, then the coupling arrangement does not affect the stage gain obtainable from the first valve. The voltage amplification of the stage is now $\mathbf{V}_{g2}/\mathbf{V}$ If $1/\omega C$ is comparable with R_g, then C and R_g act as a voltage divider across R_1. Thus some of the output voltage across R_1 is dropped across C and is not passed on to the grid of the second valve, with a resulting loss of gain. There is also some frequency and phase distortion. It may easily be shown that the reduction in gain in the coupling circuit due to C is given by

$$R_g/\sqrt{\{R_g^2 + (1/\omega C)^2\}},$$

and the phase change is given by the angle ϕ, where

$$\tan \phi = 1/\omega C R_g.$$

Thus the frequency and phase distortion are greater at lower frequencies. If $1/\omega C \ll R_g$ there may still be a loss of gain if R_g is comparable with R_1, since they are effectively in parallel for a.c. This loss is not accompanied by frequency or phase distortion.

It follows that resistance-loaded amplifiers with condenser-grid leak coupling are all subject to frequency and phase distortion at low frequencies and high frequencies (see Section 7.8). In designing amplifiers care must be taken to use components that keep these distortions within reasonable limits over the required range of frequencies.

In the circuit with mutual inductance coupling shown in Fig. 7.14.*b*, the a.c. flowing in the primary inductance L_1, which is part of the anode load of the first valve, induces an e.m.f. in the secondary L_2, and this is passed on to the grid of the next valve. If the current in L_1 is $\hat{\imath}_l \sin \omega t$ the induced e.m.f. is

$$v_2 = M \frac{di_l}{dt} = \omega M \hat{\imath}_l \cos \omega t \text{ or } \mathbf{V}_2 = j\omega M \mathbf{I}_l.$$

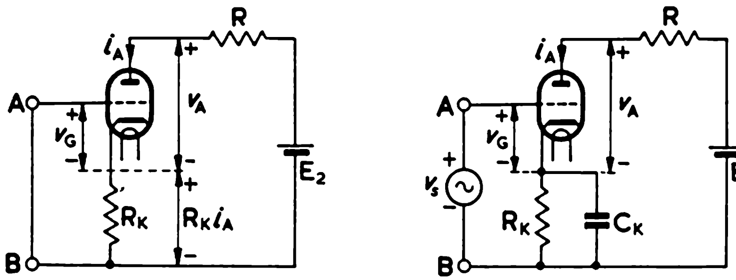
The value of M is related to L_1 and L_2 by the coefficient of coupling k, where $M^2 = k^2L_1L_2$ and $0 \leqslant k \leqslant 1$. The a.c. in the secondary winding is normally very small, since it is open-circuited apart from the input capacitance of the second valve. Hence $v_1 = L_1 \dfrac{di_1}{dt}$. It follows that

$$v_2/v_1 = \frac{M}{L_1} = k\sqrt{(L_2/L_1)}.$$

Thus if $L_2 > L_1$ it is possible to get additional voltage amplification from the mutual inductance coupling. It is shown in Section 8.5 that the mutual coupling also serves to give efficient operation in power amplifiers.

7.11. Automatic Bias

In the circuits considered so far in this chapter separate batteries are used for the anode and grid supplies. In most amplifiers only one battery is used even when there are several stages. The grid bias voltage is



(a) (b)

FIG. 7.15

obtained from the voltage drop across a resistor R_K as shown in Fig. 7.15.*a*. In the circuit it is assumed, in the first place, that the input terminals A and B are joined together and that current i_A is flowing through the valve, the anode resistance R and the cathode resistance R_K. This current produces a voltage drop R_Ki_A across R_K. One end of R_K is joined to the cathode and the other end to the grid, and it follows that $v_G = - R_Ki_A$. If the value of R_K is suitably chosen, then v_G gives the correct value of grid bias voltage for the required Q-point. When an a.c. signal is applied to the input terminals the anode current varies and the grid bias varies. In order to keep the bias voltage steady a capacitor C_K is connected in parallel with R_K as shown in Fig. 7.15.*b*. Then, if $1/\omega C_K \ll R_K$ at the frequency of operation, the a.c. voltage across R_K is much less than the d.c. voltage, and the grid bias is steady and equal to $- R_Ki_Q$, where i_Q is the quiescent anode current. The quiescent anode voltage is now given by $v_Q = E_2 - (R + R_K)i_Q$.

The steady voltage required for the screen of a pentode is also obtained from the E_2 battery, and its value may be adjusted by means of a series resistance in the lead to the screen. The arrangement is shown in Fig. 7.16. In the quiescent state the screen voltage v_{G2} is given by $v_{G2} = E_2 - R_S i_{G2} - R_K i_K$, where i_{G2}, is the quiescent screen current. When a signal is applied the screen current varies, but the screen voltage is kept constant by means of the capacitor C_S, provided $1/\omega C_S \ll R_S$ at the signal frequency.

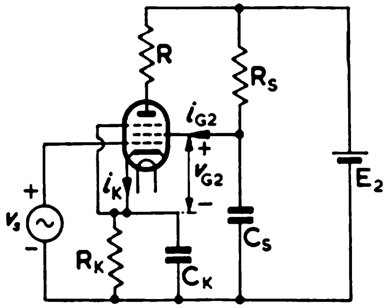


FIG. 7.16

A circuit for a two-stage pentode amplifier with resistance loads is shown in Fig. 7.17. In such a circuit it is essential to have a d.c. conducting path through the input source in order to give the first valve its grid bias.

7.12. Alternative Connections of Valve Amplifier

Since the triode valve has three electrodes, there are six possible ways in which the input could be applied between one pair of electrodes and the output could be obtained across another pair. In the conventional amplifier the input is applied between grid and cathode, and the output is obtained between anode and cathode. This may be referred to as

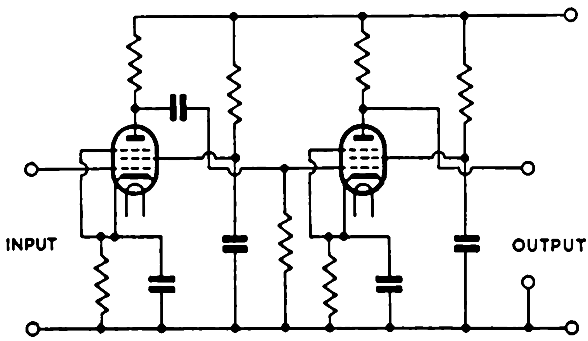


FIG. 7.17

the common-cathode connection. Two other connections are frequently used in which the grid and the anode are the common electrodes. The common-grid amplifier is shown in Fig. 7.18.*a* and the equivalent circuit in Fig. 7.18.*b*. From the latter it is seen that

$$\mathbf{V}_s(\mu + 1) = (r_a + \mathbf{Z})\mathbf{I}_a \text{ and } \mathbf{V}_o = -\mathbf{Z}\mathbf{I}_a.$$

The voltage amplification is given by $\mathbf{A} = \mathbf{V}_0/\mathbf{V}_s$

i.e., $\mathbf{A} = -\mathbf{Z}(\mu + 1)/(r_a + \mathbf{Z}) = -\mathbf{Z}\{g_m + 1/r_a\}/\{1 + \mathbf{Z}/r_a\}.$

The corresponding formula for the conventional or common-cathode amplifier is

$$\mathbf{A} = -\mu\mathbf{Z}/(r_a + \mathbf{Z})$$

so that the common-grid amplifier gives slightly higher voltage gain.

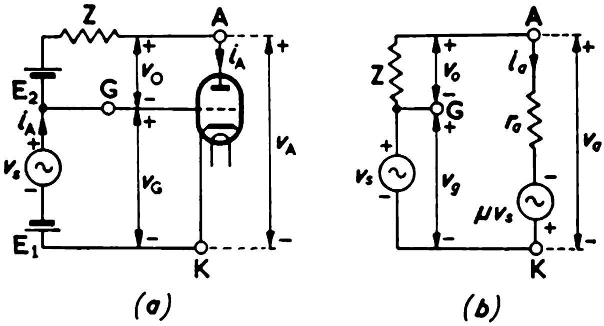


FIG. 7.18

This type of amplifier is sometimes used at very high frequencies (see Section 13.9).

The common-anode amplifier and its equivalent circuit are shown in Fig. 7.19.*a* and *b*. The input is connected between grid and anode and

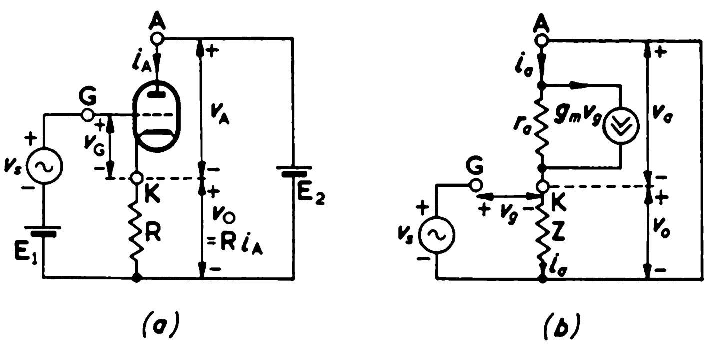


FIG. 7.19

the output is taken between anode and cathode. The voltage amplification in this case may be derived in the form

$$\mathbf{A} = g_m\mathbf{Z}/(1 + g_m\mathbf{Z} + \mathbf{Z}/r_a).$$

When the load $\mathbf{Z}$ is resistive, $|\mathbf{A}|$ is less than unity. Such an amplifier has certain special properties. It is usually called a cathode follower, and it is dealt with further in Section 10.10.

7.13. Input and Output Impedance of an Amplifier

When a signal generator is connected to an amplifier the current flowing in the input circuit depends on the input impedance of the amplifier, $\mathbf{Z}_s$, which is defined to be $\mathbf{Z}_s = \mathbf{V}_s/\mathbf{I}_s$ (see Fig. 7.20). In a conventional amplifier the input impedance depends on the inter-electrode capacitances of the valve and the leakage resistance between the grid and cathode electrodes. At low frequencies the input impedance can be very high.

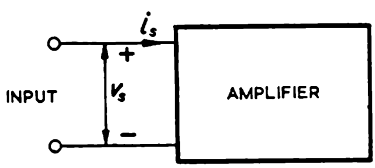


FIG. 7.20

On the output side of an amplifier the valve acts as a generator feeding a load. The internal impedance of the generator is sometimes called the output impedance of the amplifier. In the conventional amplifier the output impedance is r_a. Input and output impedances are considered further in Chapter 10. It is shown there that the common-grid amplifier has a low input impedance and the cathode follower has high input impedance and low output impedance.

7.14. Limitations of Small Signal Theory

In most of this chapter we have limited the signals to small values, and it may be worth while repeating and emphasizing the significance of this limitation. The use of static characteristics and the load line by themselves is subject to no limitation of amplitude. The true output voltage may always be determined in principle for any given input voltage as long as the characteristics are available for the range of operation. Large signals are given more detailed consideration in the next chapter. The small signal limitation has arisen from the use of the Valve Equation, which relates small changes of anode current to the changes in grid and anode voltages,

$$i_a = g_m v_g + v_a/r_a$$

or

$$\mathbf{I}_a = g_m \mathbf{V}_g + \mathbf{V}_a/r_a$$

when the changes are sinusoidal. The changes are limited to a range over which the valve parameters g_m and r_a are constant. These restrictions also apply to the valve equivalent circuits.

POWER AMPLIFIERS

8.1. Large Signals

In Chapter 7 the operation of amplifiers is confined to signal amplitudes over which the valve parameters remain constant. Such amplifiers, though dealing with small input and output voltages, may give large voltage amplification. These are the conditions which usually exist in the early stages of an amplifier. However, in the final stages it frequently happens that considerable power is required for some purpose, such as operation of a loudspeaker, energizing a transmitting aerial or activating

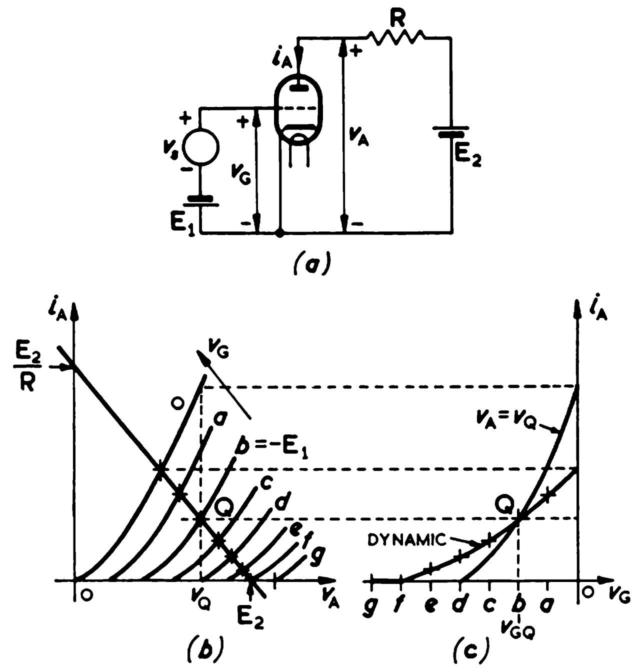


FIG. 8.1

some control mechanism. Then both large voltage and large current changes are required. The use of the Load Line Equation is not limited with regard to size of signal, and the load line is therefore the starting point in our consideration of power amplifiers.

8.2. Load Line and Dynamic Characteristics

A set of anode characteristics with a load line is shown in Fig. 8.1.*b* for the valve and circuit of Fig. 8.1.*a*. As v_G is varied and E_2 and R are kept constant, the values of v_A and i_A always lie on the load line, and the actual values may be found by picking the appropriate static characteristic corresponding to the instantaneous value of v_G. This diagram is based on the anode characteristics and gives the values of i_A in terms of v_A, as v_G is varied. The same information could be given directly in terms of

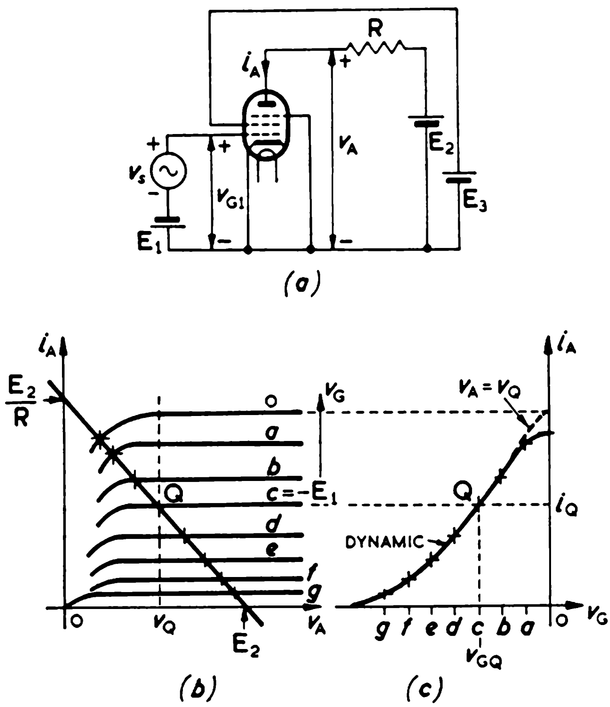


FIG. 8.2

v_G in a diagram based on the grid characteristics as shown in Fig. 8.1.*c*. This diagram, which shows the variation of the anode current with grid voltage when there is a load resistance is called the dynamic grid characteristic or merely the dynamic characteristic. The load line and the dynamic characteristic give the same information, but sometimes one is more convenient to use than the other.

Corresponding load line and dynamic characteristic for a pentode with resistive load are given in Fig. 8.2. The pentode dynamic characteristic coincides with the static grid characteristic, except when v_G approaches zero.

When a signal is applied to the grid of an amplifier the variation in anode current and anode voltage may be determined from the load line or the dynamic characteristic as shown in Fig. 8.3. The variation in grid voltage is confined to a region over which the dynamic characteristic is practically straight and the variation in anode current is a faithful replica

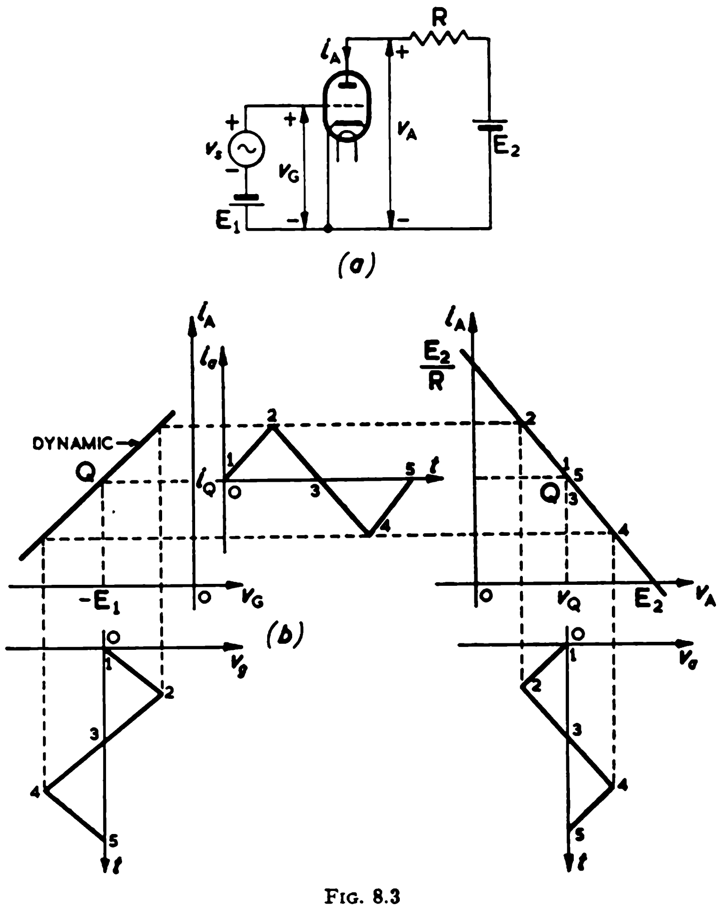


FIG. 8.3

of the grid signal. However, if a large signal is applied as in Fig. 8.4, then the anode current waveform differs from the signal waveform. When this happens the anode current is said to have non-linear distortion. The output voltage, which is proportional to the anode current, suffers similarly from non-linear distortion. This type of distortion is sometimes called amplitude distortion, since it increases with the amplitude of the signal. The subject of distortion is considered in more detail later in this chapter.

The conditions for small signals (constant g_m and r_a) imposed in Chapter 7 correspond in Fig. 8.1.*b* or 8.2.*b* to any range along the load line, on either side of Q, over which the anode characteristics are parallel and cut off equal intercepts on the load line. In Fig. 8.1.*c* or 8.2.*c* the corresponding condition is the portion of the dynamic characteristic about Q which is straight.

8.3. Power Output with Resistance Load

The power output is most easily determined from the anode characteristics and the load line. In Fig. 8.5 it is assumed that the signal is

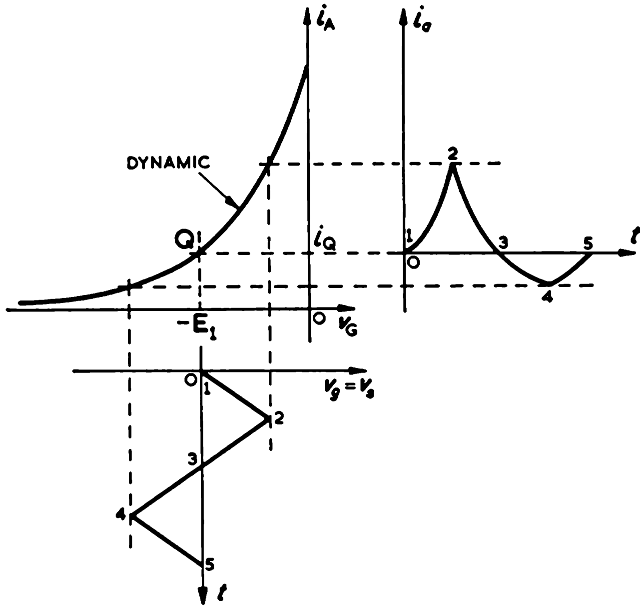


FIG. 8.4

sinusoidal and the effects of distortion are ignored. Characteristics are drawn for a triode and a pentode. As a result of the signal the anode current varies sinusoidally about its mean value i_Q up to i_{max} and down to i_{min}. The anode voltage varies similarly about v_Q down to v_{min} and up to v_{max}. These values give the current and voltage variations across the load resistance, and the power in the load is found from the product of the r.m.s. values of the current and voltage. Obviously $(i_{max} - i_{min})/2$ gives the amplitude and $(i_{max} - i_{min})/2\sqrt{2}$ the r.m.s. value of the a.c. component of the load current, and similarly $(v_{max} - v_{min})/2\sqrt{2}$ gives the r.m.s. voltage. The power output is therefore

$$P_O = (i_{max} - i_{min})(v_{max} - v_{min})/8 = R(i_{max} - i_{min})^2/8$$
$$= (v_{max} - v_{min})^2/8R.$$

This expression for P_O is derived on the assumption that the grid operating voltage v_{GQ} ($= - E_1$) is given and there is no distortion. This means that $\hat{v}_g$, the amplitude of the grid signal, is limited to the range on either side of v_{GQ} over which the static characteristics make equal intercepts on the load line. Frequently the range of operation of a power amplifier is

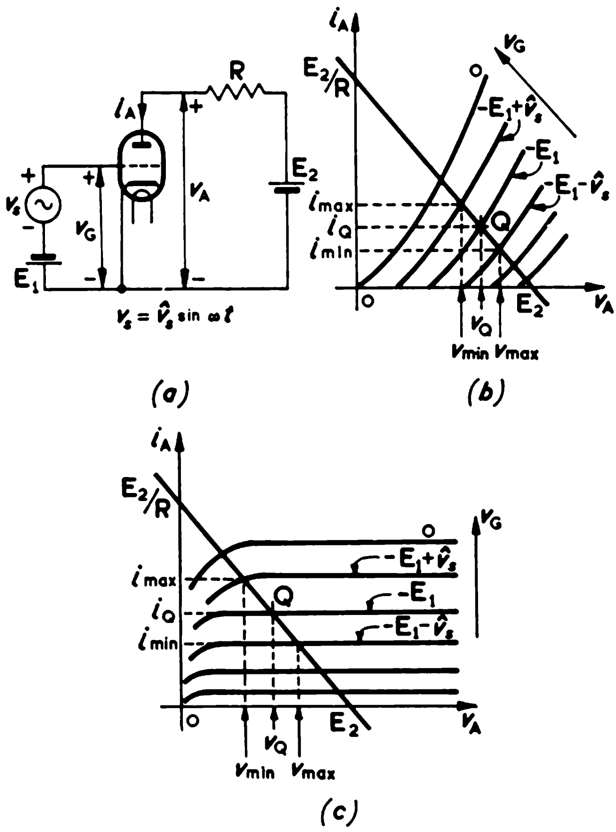


FIG. 8.5

limited to negative grid voltages, and this imposes a further limitation on $\hat{v}_g$, namely $\hat{v}_g \leqslant v_{GQ}$. We have assumed above that v_{GQ} is given. Usually in practice, E_2 and R are fixed and the diagram is studied for the range of grid voltage over which the distortion is negligible; v_{GQ} is chosen at the mid-point of this range. An ideal case would occur where there is no distortion over the whole length of the load line from $v_G = 0$ to the v_A-axis. For maximum output the anode current then varies from zero to i_{max}, the value for $v_G = 0$ (see Fig. 8.6.*a*); the anode voltage range is v_{min} to E_2. Under these conditions

$$i_{max} = 2\hat{i}_a = 2i_Q$$

and the power output is

$$P_O = i_Q(E_2 - v_{\min})/4.$$

An alternative expression for the power output in this ideal case may be obtained from the voltage-equivalent circuit (Fig. 8.6.*b*). Its use is justified, since we are assuming constant valve parameters over the

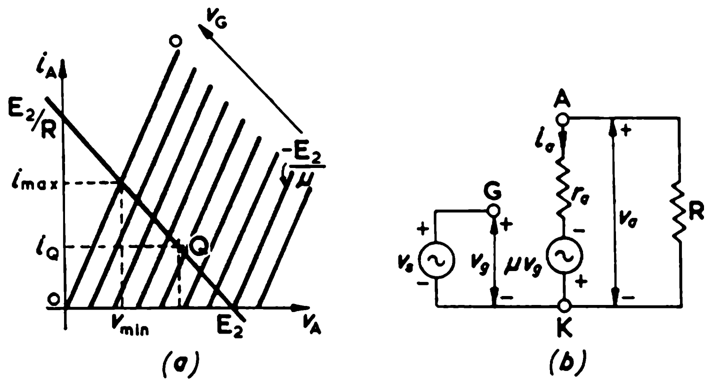


FIG. 8.6

whole working range. As the current is cut off at $v_A = E_2$, it follows that the full range of grid voltage is E_2/μ, and hence

$$v_{GQ} = -\hat{v}_g = -E_2/2\mu.$$

From the equivalent circuit

$$i_a = \mu v_g/(R + r_a)$$

and thus $$\hat{i}_a = E_2/2(R + r_a).$$

But $\hat{v}_a = R\hat{i}_a$, and so

$$P_O = \hat{v}_a\hat{i}_a/2 = RE_2^2/8(R + r_a)^2.$$

When R is varied the power output thus has a maximum value when $R = r_a$. This is in conformity with the usual relation that a generator gives maximum power output when the load resistance equals the internal resistance of the generator. The a.c. output power P_O is developed by the alternating anode current flowing in the load resistance. Although this current is controlled by the grid voltage, its source is the E_2 battery, and P_O is obtained by the consumption of energy in that battery. The total power P_I taken from the battery is its voltage multiplied by the mean current flowing through it, i.e.,

$$P_I = E_2 i_Q = E_2\hat{i}_a = E_2^2/2(R + r_a).$$

The efficiency of power conversion from d.c. to a.c. is given by $\eta = P_O/P_I$,

i.e., $$\eta = R/4(R + r_a).$$

Thus under the condition of maximum power output, $R = r_a$, the efficiency is 12·5 per cent. Greatest efficiency occurs when $R \gg r_a$ and the limiting value to η is then 25 per cent. Under this condition both the power output and the power input are negligibly small.

It is important to consider where all of the power taken from the battery is used. In addition to the a.c. component of the anode current there is also a d.c. component i_Q, which dissipates energy in the load at the rate

$$Ri_Q^2 = R\hat{i}_a^2 = RE_2^2/4(R + r_a)^2 = 2\eta P_I.$$

This is double the a.c. power output. The remainder of the battery power, $(1 - 3\eta)P_I$, is dissipated as heat at the anode of the valve, and

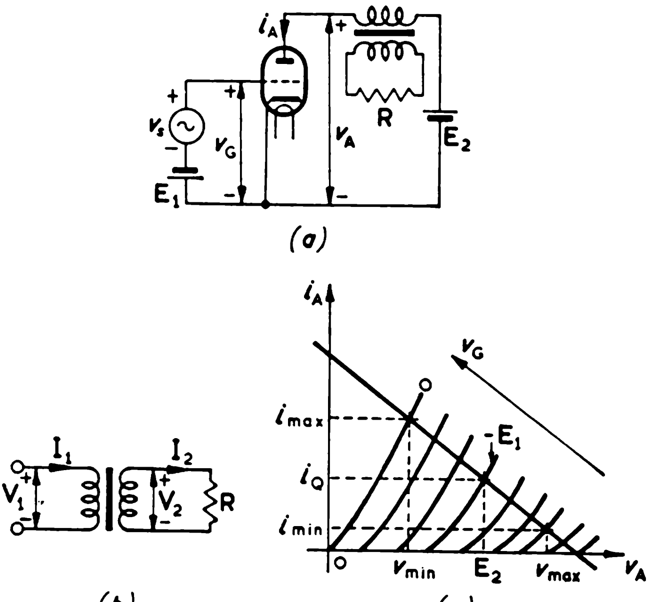


FIG. 8.7

represents the rate of loss of kinetic energy of the electrons striking the anode. This particular power loss is called the anode dissipation; it is considered further in Section 8.11.

8.4. The Transformer-Coupled Load

The steady component of the anode current flowing through the load resistance dissipates power equal to twice the useful power output. This wastage may be avoided if the load resistance is coupled to the valve through a transformer as shown in Fig. 8.7.*a*. This arrangement has the additional advantage that the transformer may be used for matching the actual load resistance R to the optimum load for the valve. It may be shown that, for an ideal transformer, the following relationships are satisfied (see Fig. 8.7.*b*):

$$\mathbf{V}_1/\mathbf{V}_2 = n_1/n_2,\ \mathbf{I}_1/\mathbf{I}_2 = n_2/n_1 \text{ and } \mathbf{V}_1/\mathbf{I}_1 = R(n_1/n_2)^2,$$

where n_1 and n_2 are the numbers of turns in the primary and secondary windings respectively. The ratio $\mathbf{V}_1/\mathbf{I}_1$ is the effective resistance to a.c. appearing across the primary terminals of the transformer. This resistance should equal the optimum load for the valve. If the optimum load is R_1, then the turns ratio of the transformer is chosen so that

$$n_1/n_2 = \sqrt{R_1/R}.$$

The transformer circuit then provides the valve with the optimum load resistance for a.c. However, for d.c. the load is the d.c. resistance of the

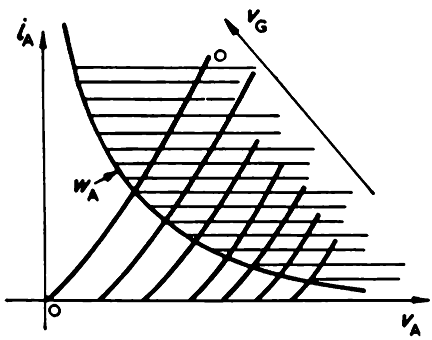


FIG. 8.8

primary winding of the transformer. This is usually very small, and the d.c. load resistance may be taken to be zero. In these circumstances there is negligible power loss due to the steady anode current, and $E_2 = v_{AQ}$. The load-line conditions are shown in Fig. 8.7.*c*. The power output is again given by the equation

$$P_O = (i_{max} - i_{min})(v_{max} - v_{min})/8.$$

Frequently a power amplifier has to cope with a wide range of signal, and then it gives maximum output when it is handling the largest signal. Under all other conditions the output is less. The power balance with transformer coupling is given by: power from battery = power output + anode dissipation. The power from the battery ($E_2 i_Q$) is constant, so that the anode dissipation is large when the output is small. There is always a maximum allowable anode dissipation (W_A), and in designing an amplifier where the signal level may fall to zero, a limiting condition is that the Q point must give

$$E_2 i_Q \leqslant W_A.$$

In Fig. 8.8 this means that the Q point must not lie in the shaded area.

In dealing with the question of power so far, only the anode circuit has been considered. When the grid voltage remains negative no appreciable current flows through the grid battery or the grid signal source. No power is therefore consumed in the grid circuit under these conditions. In any overall balance of power, account must be taken of the power required to heat the cathode.

8.5. Load Resistance for Maximum Power Output with Transformer Coupling

In Section 8.3 the power output and optimum load resistance are determined for an amplifier with a directly coupled load resistance and a given battery supply. With this arrangement and ignoring distortion,

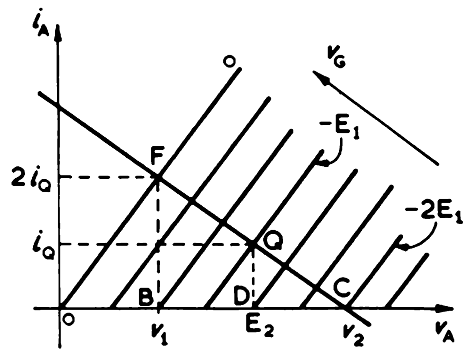


FIG. 8.9

the value of v_{GQ} is fixed at the mid-point between $v_G = 0$ and the E_2 point on the v_A-axis, and the maximum amplitude of the grid signal is equal to v_{GQ}. Under these conditions it is found that the maximum output is obtained when the load resistance is equal to r_a and the efficiency is then 12·5 per cent. The maximum possible efficiency is 25 per cent and occurs with $R \gg r_a$. When transformer coupling is used between the valve and the load, v_{AQ} is fixed. Then as R varies, v_{GQ} varies and so does the maximum signal amplitude. The conditions are quite different from those in Section 8.3, and we are now going to see that they give different values for the optimum load resistance and the efficiency. Again it is assumed that distortion is negligible over the whole range of the load line (see Fig. 8.9). Maximum power output is obtained when $i_{max} = 2i_Q$ and $i_{min} = 0$. Then $P_O = Ri_Q^2/2$. From the geometry of the figure,

$$v_2 - E_2 = E_2 - v_1$$

and from $\triangle QDC$, $\quad i_Q = (v_2 - E_2)/R = (E_2 - v_1)/R.$

Also from $\triangle FOB$, $\quad v_1 = 2i_Q r_a.$

By elimination of v_1 it is found that

$$i_Q = E_2/(R + 2r_a),$$

and hence

$$P_O = RE_2^2/2(R + 2r_a)^2.$$

For variation of R this has a maximum value when $R = 2r_a$. The power taken from the battery is $E_2 i_Q$ and the efficiency of power conversion is

$$\eta = R/2(R + 2r_a).$$

When $R = 2r_a$, the efficiency is 25 per cent. The maximum value of η again occurs when $R \gg r_a$ and the limiting value is 50 per cent. These figures confirm the improvement in performance which is obtained with a transformer-coupled load. Practically all power amplifiers use transformer coupling.

In this section and in Section 8.3 it is assumed that linear operation is obtained over the whole load line. In practice this is not true, and the range of operation has to be limited to keep the distortion to a reasonable value.

8.6. Non-linear Distortion

We have seen that, with large signals, the output waveform differs from the signal waveform. This distortion arises whenever the grid signal covers a range over which the dynamic characteristic is not straight. A curve may be represented by a suitable power series, and the dynamic characteristics of Fig. 8.1.*c* and 8.2.*c* can have equations of the form

$$i_A = A + Bv_G + Cv_G^2 + Dv_G^3 + \ldots$$

where A, B, C, D . . . are constants. The number of terms required depends on the shape of the curve. For the dynamic characteristic of a triode only the first three terms are required to give a fairly good approximation to the equation of the curve. In other words, the triode dynamic characteristic has approximately a square law. In the case of the pentode the dynamic characteristic requires a cubic term also (this is obvious from the shape of the curve which shows inflexion of the kind appropriate to a cubic law).

When a signal v_s is applied to the grid of an amplifier then $v_G = -E_1 + v_s$. Also $v_g = v_s$. The equation for the dynamic characteristic expresses the total anode current i_A in terms of the total grid voltage v_G. A similar equation relates the varying components i_a and v_g, viz.,

$$i_a = av_g + bv_g^2 + cv_g^3 + \ldots$$

where a, b, c . . . are a new set of constants. This equation is just as general as the previous equation, but it is more convenient for analysis. In the case of the triode the equation is

$$i_a = av_g + bv_g^2.$$

Now let $v_g = \hat{v}_g \sin \omega t$ as shown in Fig. 8.10.*a*. The resulting anode current variation is obviously not sinusoidal. On substituting in the equation, we find

$$\begin{aligned} i_a &= a\hat{v}_g \sin \omega t + b\hat{v}_g^2 \sin^2 \omega t \\ &= \frac{b\hat{v}_g^2}{2} + a\hat{v}_g \sin \omega t - \frac{b\hat{v}_g^2}{2} \cos 2\omega t, \end{aligned}$$

since $\sin^2 \omega t = (1 - \cos 2\omega t)/2.$

Thus the signal increases the mean component of the anode current by $\frac{b\hat{v}_g^2}{2}$. It produces a sinusoidal component $a\hat{v}_g \sin \omega t$ which is proportional to the signal, but it also produces a new component which has twice the

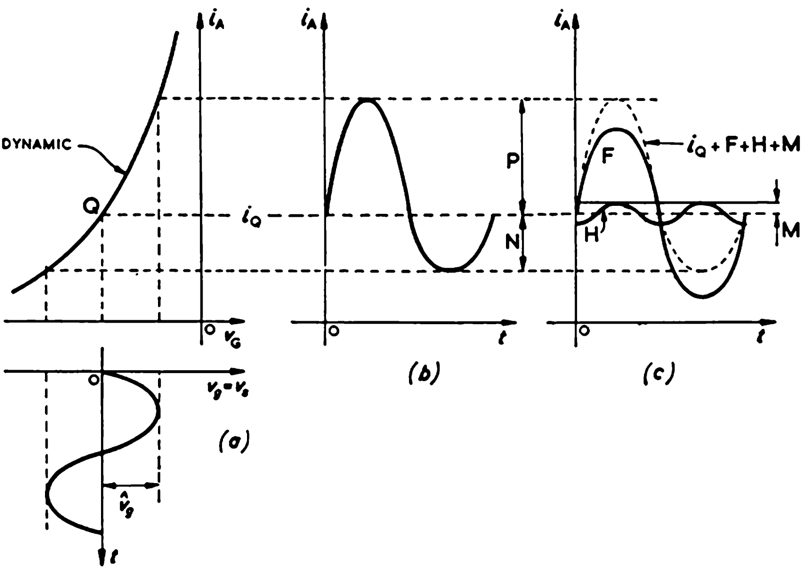


Fig. 8.10

frequency of the signal, i.e., a second harmonic of the signal is introduced. This effect is called harmonic distortion. It may be seen that the amplitude of the harmonic component equals the increase in the mean current. This provides a simple means of checking for harmonic distortion by using an ammeter to measure the mean anode current. A moving-coil meter is suitable.

The way in which the actual waveform is built up from its components is shown in Fig. 8.10.*b* and *c*. The expression for i_a may be written in the form

$$i_a = \frac{b\hat{v}_g^2}{2} + a\hat{v}_g \sin \omega t + \frac{b\hat{v}_g^2}{2} \cos (2\omega t - \pi).$$

The last two terms in this equation are represented by the curves F and H in Fig. 8.10.*c*. The positive and negative peaks, P and N, of the actual anode current are obviously given by $P = F + H + M$ and $N = F - H - M$, where F and H are now the amplitudes of the fundamental and harmonic components and M is the increase in the mean component. Since $H = M$, these expressions lead to $P + N = 2F$, $P - N = 4H$. Thus

$$H/F = (P - N)/2(P + N)$$

gives the fractional harmonic distortion. It is usually expressed as a percentage. In a straightforward triode amplifier operating at its full

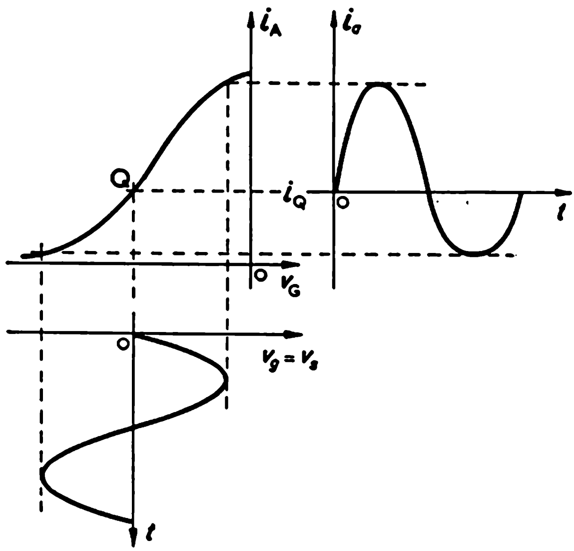


FIG. 8.11

output the second harmonic distortion may amount to 5 to 10 per cent.

A similar analysis may be carried out to determine the harmonic distortion introduced by a pentode amplifier. For a sinusoidal grid signal we get

$$i_a = a\hat{v}_g \sin \omega t + b\hat{v}_g^2 \sin^2 \omega t + c\hat{v}_g^3 \sin^3 \omega t.$$

The first two terms are the same as for the triode. The third term may be resolved using $\sin 3\omega t = 3 \sin \omega t - 4 \sin^3 \omega t$. The anode current is then

$$i_a = \frac{b\hat{v}_g^2}{2} + (a\hat{v}_g + \tfrac{3}{4} c\hat{v}_g^3) \sin \omega t - \frac{b\hat{v}_g^2}{2} \cos 2\omega t$$
$$- \frac{c\hat{v}_g^3}{4} \sin 3\omega t.$$

There is now also a third harmonic, and this is often of greater amplitude than the second harmonic. The anode current waveform is given in Fig. 8.11, and it shows flattening at both top and bottom. This is the effect of the third harmonic component. The unequal amplitudes are due to the second harmonic and the change in mean current. The waveform may be synthesized from its components as in the triode case.

8.7. Intermodulation

Non-linear characteristics may give rise to another type of distortion when the signal consists of more than one sinusoid. For example, let a grid signal

$$v_s = v_g = \hat{v}_1 \sin \omega_1 t + \hat{v}_2 \sin \omega_2 t$$

be applied to a triode. The anode current is then given by $i_a = a(\hat{v}_1 \sin \omega_1 t + \hat{v}_2 \sin \omega_2 t) + b(\hat{v}_1 \sin \omega_1 t + \hat{v}_2 \sin \omega_2 t)^2$. On expanding and using $\sin^2 \omega t = (1 - \cos 2\omega t)/2$ and $2 \sin \omega_1 t \sin \omega_2 t = \cos (\omega_1 - \omega_2)t - \cos (\omega_1 + \omega_2)t$, it is found that

$$\begin{aligned} i_a = \frac{b\hat{v}_1^2}{2} + \frac{b\hat{v}_2^2}{2} + a\hat{v}_1 \sin \omega_1 t + a\hat{v}_2 \sin \omega_2 t \\ - \frac{b\hat{v}_1^2}{2} \cos 2\omega_1 t - \frac{b\hat{v}_2^2}{2} \cos 2\omega_2 t + b\hat{v}_1\hat{v}_2 \cos (\omega_1 - \omega_2)t \\ - b\hat{v}_1\hat{v}_2 \cos (\omega_1 + \omega_2)t. \end{aligned}$$

Thus the two signals have produced not only output currents proportional to the signals but also an increase in the mean current, second harmonic components of the signal frequencies and two new a.c. components, whose frequencies are the sum and the difference of the signal frequencies. This new type of distortion is called intermodulation. In some respects it is more objectionable than harmonic distortion, particularly in audio amplifiers. Most sounds include some harmonic components of the fundamental frequency, and some increase of their amplitudes may be tolerated. However, the new components produced by intermodulation bear no harmonic relationship to the signals, and their presence is readily noticed. Intermodulation distortion may be emphasized when it occurs along with frequency distortion. For example, some loudspeakers show marked peaks in their frequency–output response curves. Correspondence of these peaks with the sum or difference intermodulation terms may explain the peculiar sounds that are sometimes heard.

Intermodulation distortion may also arise from the higher power terms in the equation of the dynamic characteristic.

8.8. Non-linear Devices

It should be noticed in passing that non-linear distortion may arise from any device whose output is not directly proportional to its input. Although it has been dealt with here in terms of curved valve characteristics, the same type of analysis would apply to any non-linear device.

A few examples are transformers or chokes with iron cores, the moving coil of a loudspeaker and the human ear.

It is also important to note that in this chapter, where we are dealing with amplifiers, the distortions arising from non-linearity are being emphasized. The same non-linear effects are essential for many important electronic applications, such as rectification, modulation and frequency changing. These are dealt with later (see Chapters 16 and 17).

8.9. Push-Pull Amplifiers

When a power amplifier is to be used with reasonable freedom from non-linear distortion it is important that the operation should be confined to straight portions of the dynamic characteristic. When more power is required, then larger valves should be used. The power may

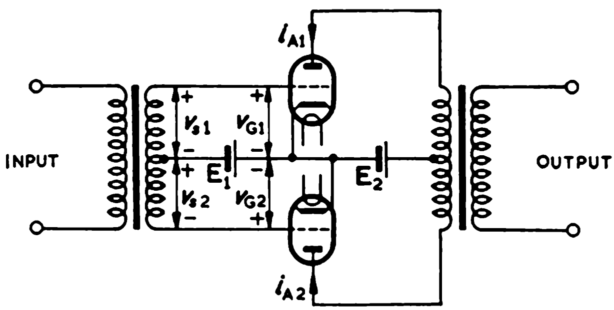


FIG. 8.12

also be increased by using a pair of valves. These can be connected in parallel, but there is an alternative method of connection which has considerable advantages. The circuit arrangement is shown in Fig. 8.12. Two identical valves and two centre-tapped transformers are used. The input signal is connected to the primary winding of one transformer, and the secondary voltage is divided equally between the grids of the two valves. As $v_{s1} = v_{s2}$, then $v_{g1} = -v_{g2}$ at any instant, and when one grid voltage rises the other drops. The primary winding of the second transformer is connected between the two anodes in such a way that the anode currents flow in opposite directions in the winding, and so produce opposing fluxes. The output is therefore proportional to $i_{a1} - i_{a2}$. An amplifier connected in this way is called a push-pull amplifier.

Since the two valves are identical they have the same dynamic grid characteristic and hence

$$i_{a1} = av_{g1} + bv_{g1}^2 + cv_{g1}^3 +$$

and
$$i_{a2} = av_{g2} + bv_{g2}^2 + cv_{g2}^3 +$$

Since
$$v_{g1} = -v_{g2},$$

$$i_{a2} = -av_{g1} + bv_{g1}^2 - cv_{g1}^3 +$$

$$i_{a1} - i_{a2} = 2av_{g1} + 2cv_{g1}^3 +$$

All the even powers have disappeared, and so there are no harmonics or intermodulation components arising from them. With triodes which have approximately square-law characteristics there is very little non-linear distortion. The push–pull circuit is less effective with pentodes, since they have an appreciable cubic term in the dynamic characteristic. The performance of the push–pull triode amplifier may be explained

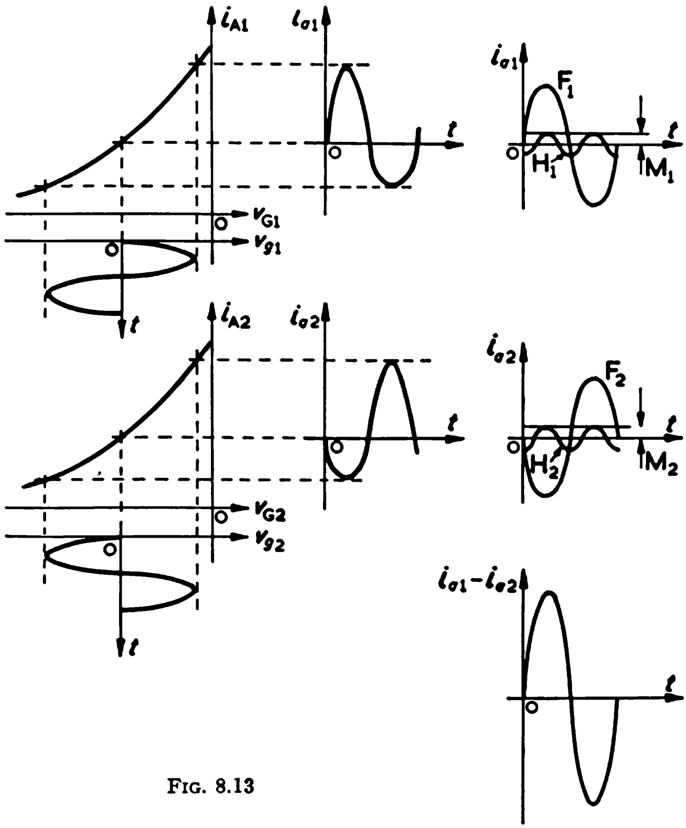


FIG. 8.13

graphically as shown in Fig. 8.13. Dynamic characteristics are drawn separately for the two valves, and i_{a1} and i_{a2} are determined for the instantaneous values of v_{g1} and v_{g2} arising from a sinusoidal input. The anode currents are then analysed into fundamental, second harmonic and mean components by the method used in Fig. 8.10. In obtaining $i_{a1} - i_{a2}$ it is seen that the harmonic and mean components disappear and the two fundamental components add together. The performance of push–pull circuits may also be determined conveniently by drawing a combined dynamic characteristic, as shown in Fig. 8.14. The two separate dynamic characteristics are placed together so that the anode currents may be found from a single input signal. The resultant combined dynamic characteristic is the sloping straight line through $- E_1$. Anode characteristics may be combined similarly and used with a common load line.

When a load resistance R is connected across the output terminals of a push–pull amplifier, then the effective a.c. anode-to-anode resistance presented by the transformer is $4R(n_1/n_2)^2$, where n_1 is the number of

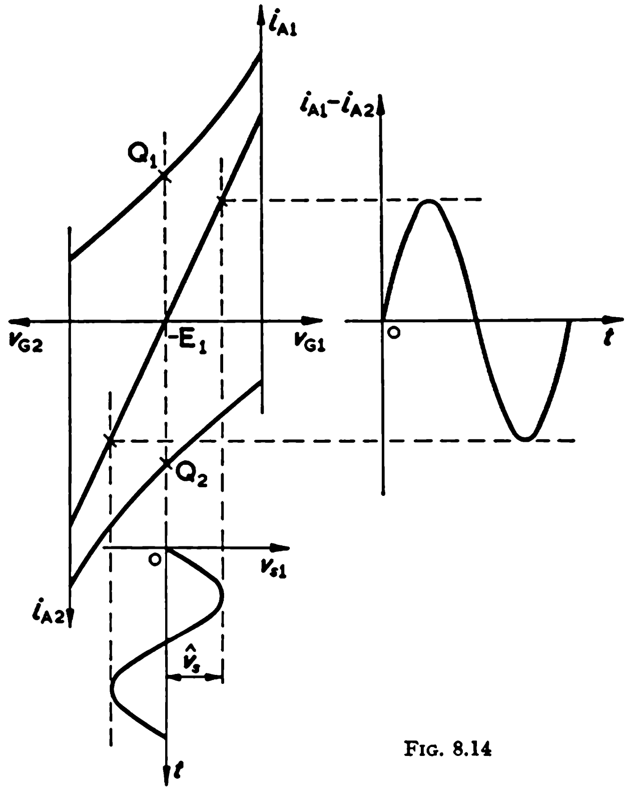


FIG. 8.14

turns on each half of the primary winding and n_2 is the total number of turns on the secondary. Each valve has therefore an effective load resistance of $2R(n_1/n_2)^2$.

In addition to reducing non-linear distortion, the push–pull amplifier has certain other desirable properties. Since the two halves of the mean anode current flow in opposite directions in the primary winding of the output transformer, there is no question of magnetic saturation of the core arising from the direct current; thus another source of non-linearity is eliminated. Also, and for the same reason, any variations in the mean anode current, due to hum or other variations in the power supply, produce no output. Finally, there are no components of the signal current in the power supply.

The push–pull circuits described so far have used transformers as essential components. Resistance loads may also be used in push–pull circuits, and examples may be found in Chapter 12.

8.10. Class B Amplifiers

The push–pull circuit can, under certain conditions, give almost distortionless output even when the operation is over ranges where the

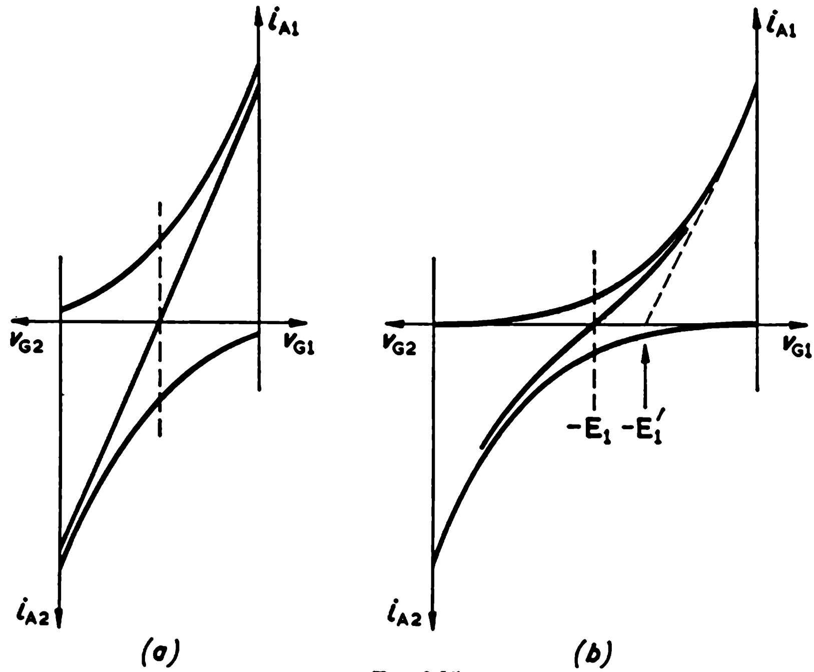


FIG. 8.15

individual dynamic characteristics show considerable curvature. The combined characteristic of Fig. 8.15.*a* illustrates this point graphically. It is possible to proceed even further, as shown in Fig. 8.15.*b*, where each valve is biased very nearly to cut-off. It may be seen that there is now appreciable curvature in the combined dynamic characteristic, and distortion would occur. The bias, $-E_1$, is too great. The greatest bias without large distortion is obtained by projecting the straight portion of the individual characteristic, as shown by the dotted line. The combined characteristic may now be redrawn with bias nearer to $-E_1'$.

When an amplifier is biased nearly to cut-off it is said to operate in the Class B state. Single-valve Class B operation is seldom used for an audio amplifier, as it would introduce excessive distortion. An amplifier which operates in the linear region of the valve characteristic is called a Class A amplifier. Practically all the amplifiers considered so far in this book are Class A. The main advantage of Class B amplifiers over Class A lies in their greater efficiency. Since they are biased nearly to cut-off, the quiescent anode currents are nearly zero. When a signal is received the mean anode currents increase. Thus very little power is consumed except when signals are received, and the consumption increases with signal size. This is a very desirable feature in an audio amplifier, where signal amplitudes vary considerably. The question of amplifier efficiency is considered in more detail in the next section.

A single valve may be used in a Class B amplifier at a high frequency

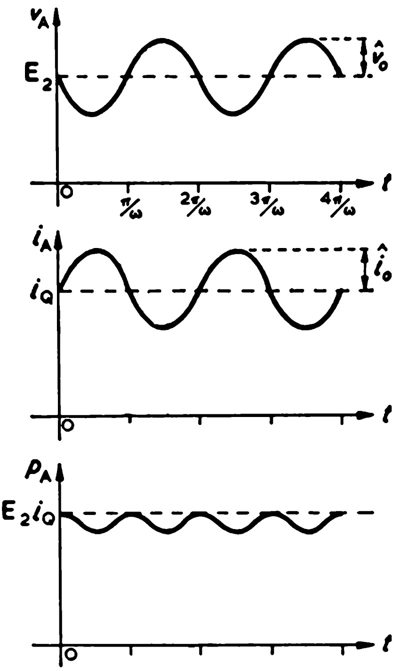


FIG. 8.16

with the load tuned to parallel resonance at the frequency of the signal. Then, although the anode current waveform is far from sinusoidal, it may be resolved into a series of components whose frequencies are multiples of the fundamental frequency. Since the load impedance is negligible at all frequencies except the fundamental, the anode voltage is practically sinusoidal.

8.11. Power Amplifier Efficiency

The efficiency of a power amplifier is given by the ratio of the useful a.c. power in the load to the total power taken from the anode battery. That power which is not useful a.c. power is dissipated as heat, either in circuit components or by electron bombardment of the anode. With a transformer-coupled load the power dissipation is practically all at the valve anode. We are now going to examine this dissipation more closely. In Fig. 8.16 the waveforms of v_A and i_A are shown for a Class A amplifier with resistance load and sinusoidal signal. As shown in Section 7.4, the a.c. components, v_a and i_a, differ in phase by 180°. We may therefore put for the instantaneous values of v_A and i_A

$$i_A = i_Q + \hat{\imath}_0 \sin \omega t \text{ and } v_A = E_2 - \hat{v}_0 \sin \omega t.$$

At an instant t an electron which leaves the cathode with zero velocity arrives at the anode with kinetic energy given by $ev_A = e(E_2 - v_0 \sin \omega t)$. On striking the anode, the electron gives up this energy as heat. Since i_A represents the rate of arrival of electrons at the anode, $v_A i_A$ gives the rate of loss of kinetic energy, i.e., the instantaneous power dissipation. The value varies with time and the mean power P_A dissipated at the anode is found by averaging $v_A i_A$ over one cycle T,

i.e., $$P_A = \frac{1}{T}\int_0^T v_A i_A dt.$$

On taking the product $v_A i_A$, all the terms in sines give an average of zero except the term in $\sin^2 \omega t$ and so

$$P_A = \frac{1}{T}\int_0^T \{E_2 i_Q - \hat{v}_0 \hat{\imath}_0 (1 - \cos 2\omega t)/2\} dt,$$

using $2 \sin^2 \omega t = 1 - \cos 2\omega t$. The cosine term gives an average of zero and then

$$P_A = E_2 i_Q - \hat{v}_0 \hat{\imath}_0 / 2.$$

In this equation $E_2 i_Q$ gives the total power taken from the battery and $\hat{v}_0 \hat{\imath}_0 / 2$ must represent the useful power output. This is confirmed when we remember that $\hat{v}_0/\sqrt{2}$ and $\hat{\imath}_0/\sqrt{2}$ are the r.m.s. values of the output voltage and current respectively. From the equation we see that the anode dissipation decreases as the output increases. Physical explanation of these conditions may be given as follows. With no signal applied, an electron in going from the negative end of the battery to the positive end takes energy eE_2 from the battery. The kinetic energy of the electron on arriving at the anode is also equal to eE_2. Thus all the energy is dissipated as heat at the anode. When a signal is applied the electrons arriving at the anode in the first half cycle have kinetic energy less than eE_2; those arriving in the second half cycle have kinetic energy greater than eE_2. Each electron takes energy eE_2 in going from the negative to the positive end of the battery. Since more electrons reach the anode during the first half cycle than during the second, the average dissipation at the anode is reduced on the application of a signal. The dissipation may be reduced still further by preventing electrons reaching the anode during the half cycle when the anode voltage is above E_2. This is what happens in the Class B amplifier. The anode current flows only during the negative half cycle of the anode voltage. The anode voltage and anode current waveforms for one valve are shown in Fig. 8.17. The anode current for one valve is far from sinusoidal, but if a tuned load is

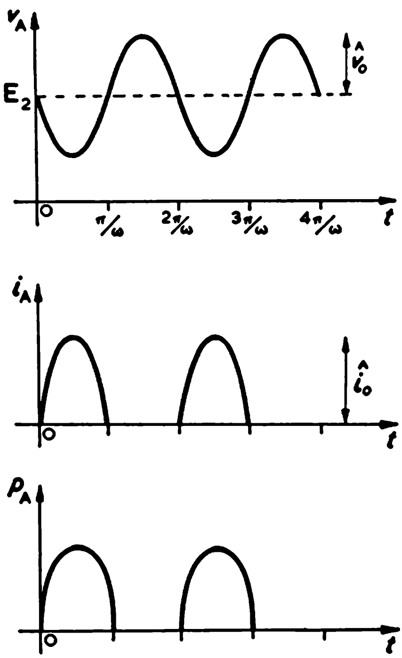


FIG. 8.17

used the anode voltage is as shown in the figure. The anode dissipation may be determined again from the instantaneous values,

$$i_A = \hat{i}_0 \sin \omega t \text{ from } 0 \text{ to } \frac{T}{2},$$

$$i_A = 0 \text{ from } \frac{T}{2} \text{ to } T$$

and $$v_A = E_2 - \hat{v}_0 \sin \omega t.$$

If we average $v_A i_A$ over a complete cycle we find for the anode dissipation in a Class B amplifier

$$P_A = E_2\hat{i}_0/\pi - \hat{v}_0\hat{i}_0/4, \text{ using } T = 2\pi/\omega;$$

$\hat{i}_0/\pi$ is the mean value of i_A and hence $E_2\hat{i}_0/\pi$ gives the power taken from the battery; $\hat{v}_0\hat{i}_0/4$ is the output power. This equation confirms that there is no dissipation when there is no signal. The efficiencies for Class A and Class B amplifiers are given respectively by the formulae

$$\eta_A = \hat{v}_0\hat{i}_0/2E_2 i_Q \quad \text{and} \quad \eta_B = \pi\hat{v}_0/4E_2.$$

The maximum possible values for $\hat{v}_0$ and $\hat{\imath}_0$ in a Class A amplifier may be seen from Fig. 8.16 to be $\hat{v}_0 = E_2$ and $\hat{\imath}_0 = i_Q$. These give for the limiting efficiency $\eta_A = 50$ per cent. The limiting value of $\hat{v}_0$ in the Class B amplifier is also E_2 and then $\eta_B = \pi/4$ or about 75 per cent. These

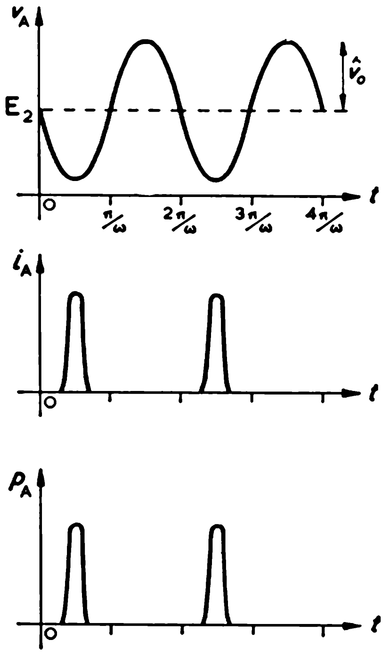


FIG. 8.18

limiting efficiencies are never reached in practice. Values of 25 and 50 per cent correspond more nearly to the highest values obtainable.

The efficiency may be increased still further by confining the electron flow to a narrow pulse at the time of the minimum anode voltage, as

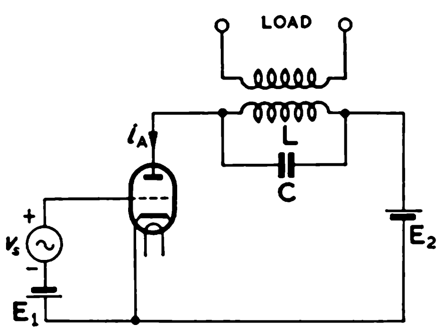


FIG. 8.19

shown in Fig. 8.18. In the limiting case when $\hat{v}_0 = E_2$, the anode dissipation would be practically zero and the efficiency would approach 100 per cent. Practical values of 70 to 80 per cent can be obtained with this type of operation. The anode current is now far removed from a sinusoidal waveform, and such amplifiers are used only at high frequency with tuned anode loads. They are known as Class C amplifiers. A suitable circuit arrangement is shown in Fig. 8.19. The LC circuit is

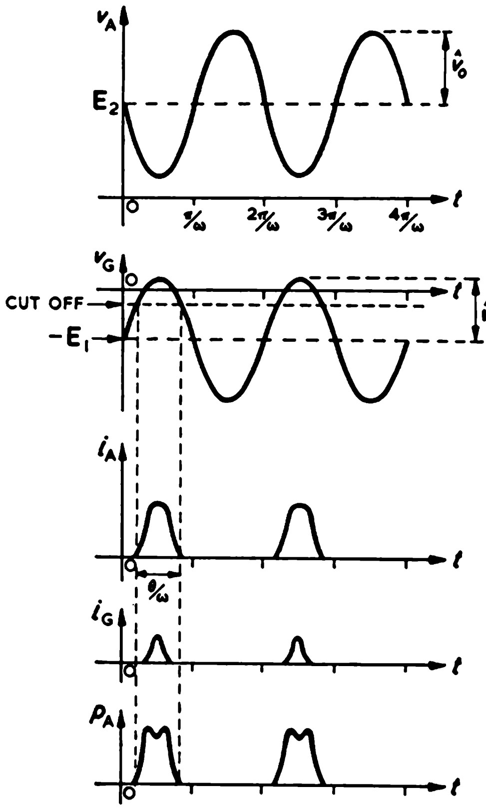


FIG. 8.20

tuned to resonance at the signal frequency, where $v_s = \hat{v}_s \sin \omega t$. The load, which may be an aerial or the input circuit of another power amplifier, is coupled to the inductance of the tuned circuit. The coupling is adjusted to give at the resonant frequency a suitable a.c. anode load for the amplifier. In order to achieve the Class C operation certain conditions are required. The grid bias must be adjusted well beyond the cut-off value for $v_A = E_2$, and the signal voltage $\hat{v}_s$ must be large enough for the grid voltage v_G to be positive for part of the operating cycle. Typical waveforms are shown in Fig. 8.20. The anode current pulse has appreciable duration, and its width is conveniently measured in terms of the angle of flow, $\theta = \omega t$, where t is the time duration of the pulse. The efficiency increases as θ is reduced; however, the power output also decreases, and some compromise is necessary. A width of about a quarter of the period of the oscillation ($\theta = \pi/2$) is commonly used. Since v_G is positive for part of the time, there is now some grid current, and power is consumed from the signal source. The Class C amplifier has departed a long way from the amplifiers which are considered at the beginning of this chapter. It is a highly non-linear device and does not readily permit of analytical study. The adjustment of the amplifier is usually carried out empirically.

The lines of demarcation between Class A, B and C amplifiers are not very clear, but it is usual to distinguish the three types by the amount of the grid bias. In Class A operation the bias is of the order of half the cut-off value, in Class B the bias is about cut-off, and in Class C much more than cut-off. The fields of use overlap to some extent. Class A amplifiers are used at any frequencies with fairly small input signals. Class B and Class C are both used as tuned amplifiers at high frequencies. Class B amplification is also possible at audio frequencies provided the push–pull circuit is used.

CHAPTER 9

TRANSISTOR AMPLIFIERS

9.1. Transistor Characteristics

The characteristics and some of the main features of transistors are described in Chapter 6. In some respects the transistor is analogous to a triode, and the emitter, base and collector correspond respectively to the cathode, grid and anode. In triode amplifiers there are three main types of circuit which are used; these are the conventional (or common-cathode) amplifier, the common-grid amplifier and the cathode follower (or common-anode) amplifier. Similarly, there are three useful types of transistor amplifier—common-emitter, common-base and common-collector. All three types are considered in this chapter. One important difference between triodes and transistors is that the latter always have some base current, whereas in triodes, grid current is frequently negligible.

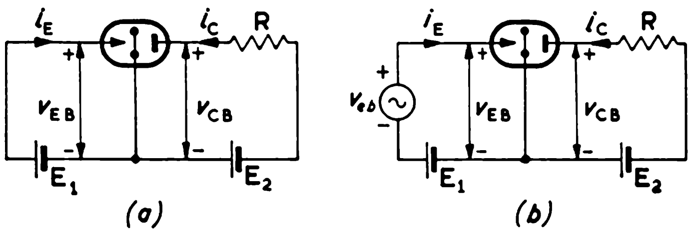


FIG. 9.1

This necessitates a slightly different procedure in establishing the performance of a transistor amplifier from the characteristics and the circuits.

In the following sections we consider mainly junction-type transistors. For these we show in Section 6.14 that the main factor controlling the transistor currents is v_{EB}, the driving voltage. Changes in v_{CB} and v_{EC} are relatively unimportant. As is done with triodes and other valves, we assume that the characteristics may be used to represent instantaneous values of the transistor currents and voltages. Thus we neglect the effects of diffusion rate or transit time of carriers. In practice, these effects may become important at lower frequencies in transistors than in vacuum valves (see Section 9.8).

9.2. Common-base Amplifier

The circuit of a common-base amplifier with a p–n–p transistor is shown in Fig. 9.1.a, where R is the load resistance and E_1 and E_2 are batteries. Then $v_{EB} = E_1$ and $v_{CB} = -E_2 - Ri_C$. In Fig. 9.1.b a signal v_{eb}, either d.c. or instantaneous a.c., is shown in the input circuit. This is the driving voltage, and it causes a change i_e in the emitter current and an almost equal but opposite change i_c in the collector current. Consequently, there is a voltage change across R and an equal but opposite change v_{cb} in the collector–base voltage, i.e., $v_{cb} = -Ri_c$. The voltage amplification is given by

$$A = v_{cb}/v_{eb} = -Ri_c/v_{eb}.$$

Thus, as long as $R > |v_{eb}/i_c|$, the stage gain is greater than unity. Since

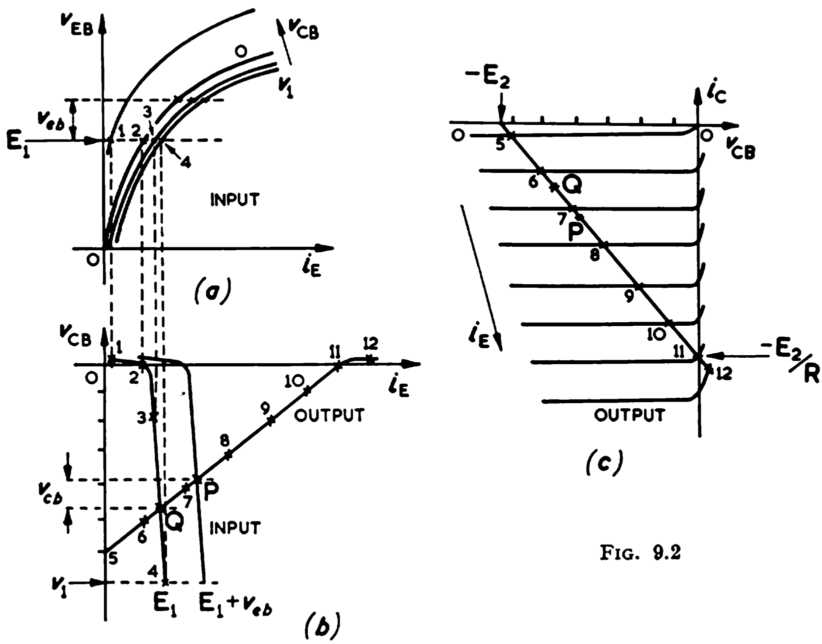


FIG. 9.2

i_c is approximately equal to $-i_e$, the condition becomes $R > |v_{eb}/i_e|$, which is usually of the order of 10 to 100 ohms, so that voltage amplification is readily obtained in the common-base circuit. Since v_{eb}/i_e varies with the signal magnitude, A also varies with the signal.

When, as in Fig. 9.1, the values of E_1, E_2 and R are given, the actual operating point and voltage gain may be found accurately by the following procedure. A start is made from v_{EB}, i_E characteristics with v_{CB} as parameter (see Fig. 9.2.a). From a constant E_1 line, relations are established between v_{CB} and i_E, and they are plotted in a v_{CB}, i_E diagram in Fig. 9.2.b, in which the numbers indicate corresponding points. In order to satisfy the input circuit conditions the value of v_{CB} must lie somewhere on this line. Its actual value is uniquely determined by the conditions in the output circuit. In Fig. 9.2.c i_C, v_{CB} characteristics are drawn for various values of i_E. Using the output circuit relation, $v_{CB} = -E_2 - Ri_C$, we can draw the load line as shown. The output relation between v_{CB} and i_E is determined, and this is also plotted in Fig. 9.2.*b*. The intersection of the two curves gives the operating point Q. The operating value of i_C is calculated from the Load Line Equation. When the signal v_{eb} is applied the conditions change in the input circuit, but the output relation between v_{CB} and i_E is unchanged. The new conditions are found at P as shown, and the output voltage v_{cb} is determined. In this case v_{cb} and v_{eb} are both positive quantities and the output voltage is in phase with the signal. The voltage gain is given by $A_v = v_{cb}/v_{eb}$.

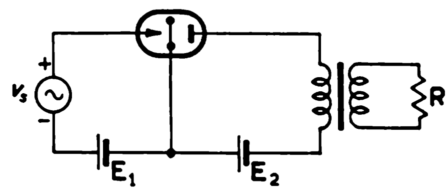


FIG. 9.3

The application of the signal v_{eb} changes i_E, and this means that there is a finite input resistance for the amplifier. Its value is given by $r_i = v_{eb}/i_e$. This may be of the order of 50 Ω or less. We have already seen that the common-base amplifier may be compared to the common-grid triode amplifier, which also has a low input resistance and can give high voltage gain.

The common-base amplifier may also give appreciable power gain. The ratio of output to input power is $v_{cb}i_c/v_{eb}i_e$. Thus the power gain A_p is given by

$$A_p = \frac{v_{cb}i_c}{v_{eb}i_e} = \frac{v_{cb}^2}{v_{eb}^2} \cdot \frac{r_i}{R}.$$

The numerical value of the power gain is normally slightly less than the voltage gain, since $|i_c/i_e|$ is less than unity. Note that the power gain in a Class A common-cathode amplifier is infinite, since ig is zero. In the common-grid case the power and voltage gains are equal.

The horizontal nature of the i_C, v_{CB} characteristics is maintained down to zero voltage. Thus the output current and output voltage in a transistor power amplifier may both vary down to zero. The value of the power output, ignoring distortion, is given by

$$P_O = (v_{\max} - v_{\min})(i_{\max} - i_{\min})/8,$$

as in the triode or pentode. The transistor may therefore approach the maximum theoretical efficiency of 25 per cent when the load resistance is in the collector lead. If the load is transformer coupled, as shown in Fig. 9.3, the limiting efficiency of 50 per cent may be approached. Of course, these efficiencies apply only to the collector circuit. Also, because of the nature of the characteristics, it is possible to design a push–pull amplifier with negligible quiescent current. Thus a Class B transistor audio amplifier consumes negligible power in the absence of a signal.

The graphical method in Fig. 9.2 of finding the output voltage is not limited as regards signal amplitude. Provided the appropriate characteristics are available, the method may be used for accurate determination of the output voltage for any given signal. Several lines for equally spaced values of v_{EB} are drawn in an i_E, v_{CB} plot in Fig. 9.4.*a*. Also in this diagram we plot, from Fig. 9.4.*b*, the output relations between i_E and v_{CB} for three values of the load resistance. The intercepts on any one output relation indicate the degree of distortion obtained with the corresponding load resistance. This distortion arises from the non-linear relationship between i_e and v_{eb}. As can be seen from Fig. 9.4.*a*, lines of

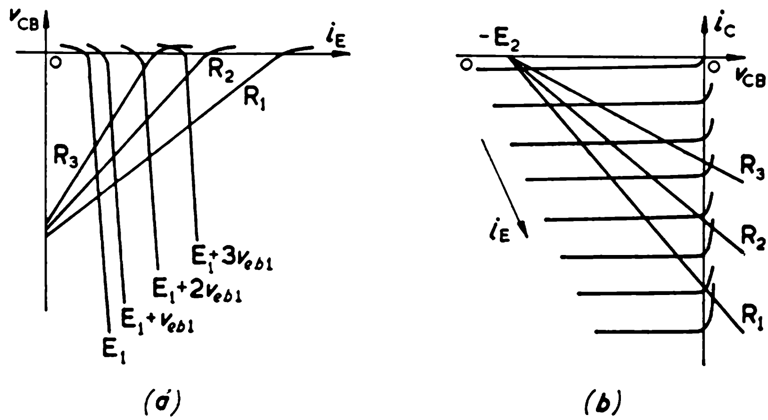


FIG. 9.4

equally spaced i_E would give very little distortion (except where v_{CB} approaches zero). A linear relationship between i_e and the signal voltage can be realized by placing in series with the input circuit a resistance large in comparison with the transistor input resistance.

9.3. Small-signal Theory of Common-base Amplifier

If the applied signal is limited to small amplitude, the transistor equations derived in Section 6.16 may be used, i.e.,

$$v_{eb} = h_{11b}i_e + h_{12b}v_{cb}$$

and

$$i_c = h_{21b}i_e + h_{22b}v_{cb}.$$

The load line relation for changes gives $v_{cb} = -Ri_c$. From these equations it can be shown that

$$A = \frac{v_{cb}}{v_{eb}} = \frac{-h_{21b}R}{h_{11b} + R(h_{11b}h_{22b} - h_{21b}h_{12b})}$$

and

$$r_i = \frac{v_{eb}}{i_e} = h_{11b} - \frac{h_{12b}h_{21b}R}{1 + h_{22b}R}.$$

Using the approximate relations of Section 6.16, we may put $h_{12b} = h_{22b} = 0$ and $-h_{21b} = \alpha_{ce} \simeq 1$, and hence

$$A \simeq R\alpha_{ce}/h_{11b} \simeq R/h_{11b}$$

and $$r_i \simeq h_{11b}.$$

These results may be compared with the corresponding expressions for a common-grid pentode amplifier (see Sections 7.12 and 10.9):

$$A = -R(g_m + 1/r_a)/(1 + R/r_a) \simeq -Rg_m$$

and $$r_i = (1 + R/r_a)/(g_m + 1/r_a) \simeq 1/g_m.$$

9.4. Common-emitter Amplifier

The circuit of a common-emitter amplifier using a p-n-p junction transistor is shown in Fig. 9.5. The input circuit is as that for the common-

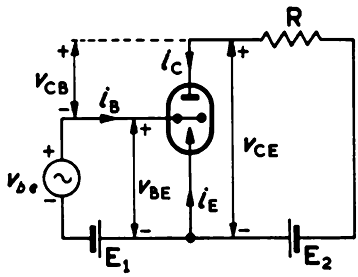


FIG. 9.5

base amplifier except for inversion of the terminals, but the output voltage is obtained between the collector and the emitter. The driving voltage v_{be} produces similar changes i_e and i_c as before, and the output voltage is given by $v_{ce} = -Ri_c$. Thus the voltage amplification differs very little from that obtained with the common-base circuit. The main difference between the two amplifiers lies in the value of the input resistance. In this case it is given by

$$r_i = v_{be}/i_b.$$

Since $|i_b|$ is much less than $|i_e|$, the common-~~base~~ em. amplifier has a higher input resistance, usually 10 to 100 times greater. As a result, the input circuit takes less power from the signal, and common-emitter amplifiers give higher power gain. Also, since $|i_c|$ is greater than $|i_b|$, the common-emitter circuit may be said to act as a current amplifier.

The graphical determination of the operating point and the output voltage is carried out in a similar manner to that used for the common base. In this case the characteristics used are v_{BE}, i_B with v_{CE} as parameter and i_C, v_{CE} with i_B as parameter, together with the Load Line Equation $v_{CE} = -E_2 - Ri_C$. The operating point is finally found from plots of i_B and v_{CE} (see Fig. 9.6).

When small signal theory is applied to the common-emitter amplifier we may use the parameters h_{11e}, h_{12e}, h_{21e} and h_{22e}, which were defined in Section 6.15. Then

$$A = \frac{v_{ce}}{v_{be}} = \frac{-h_{21e}R}{h_{11e} + R(h_{11e}h_{22e} - h_{21e}h_{12e})}$$

$$\sim \frac{-R\alpha_{cb}}{h_{11e}}$$

and
$$r_i = \frac{v_{be}}{i_b} = h_{11e} - \frac{h_{12e}h_{21e}R}{1 + h_{22e}R} \simeq h_{11e}.$$

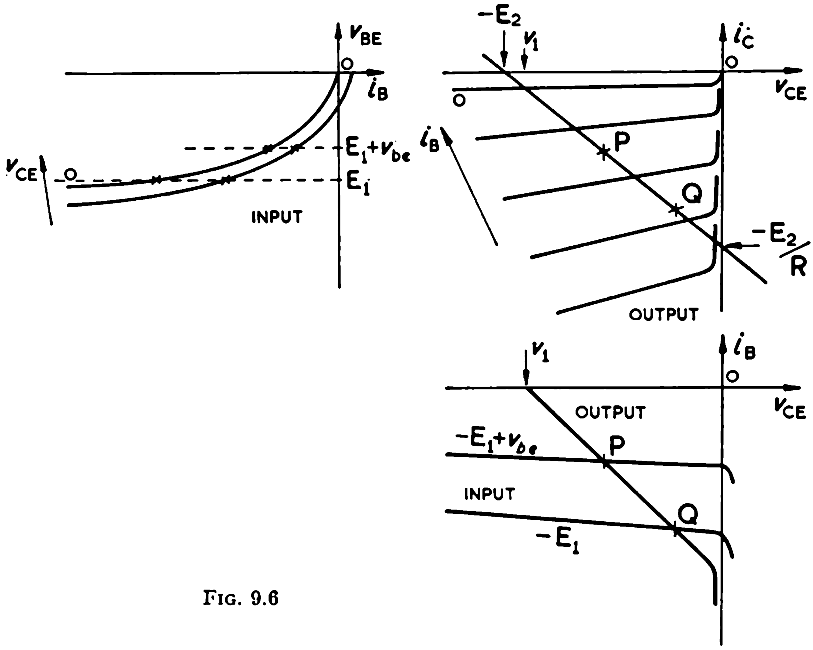


FIG. 9.6

Using the approximate relations between the common-base, common-emitter parameters given in Section 6.15 it follows that

$$A \simeq -\frac{R\alpha_{ce}}{h_{11b}}$$

and
$$r_i \simeq \frac{h_{11b}}{1 - \alpha_{ce}}.$$

The voltage gain is approximately the same as for the common-base amplifier, but the input resistance is considerably greater.

9.5. Common-collector Circuit

The common-collector amplifier, which is illustrated in Fig. 9.7, has the input signal v_s connected between base and collector, and the output circuit is between emitter and collector. The transistor amplifies the

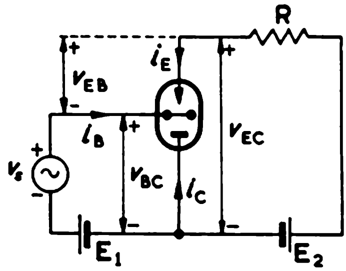


FIG. 9.7

driving voltage v_{eb} between its emitter and base. In this circuit $v_{eb} = v_{ec} + v_{cb} = -Ri_e - v_s$. The output voltage $-Ri_e$ is normally much greater than the driving voltage v_{eb}. Thus v_s and $-Ri_e$ are approximately equal. The voltage amplification $A = v_{ec}/v_{bc}$ is slightly less than unity. The conditions are very similar to those of the cathode follower, which is described in Sections 7.12, 7.13 and 10.10. Both amplifiers have high input impedance and low output impedance.

It may be shown that

$$A = \frac{1 - h_{12b}}{1 - h_{11b}/R} \simeq \frac{1}{1 - h_{11b}/R} \simeq 1.$$

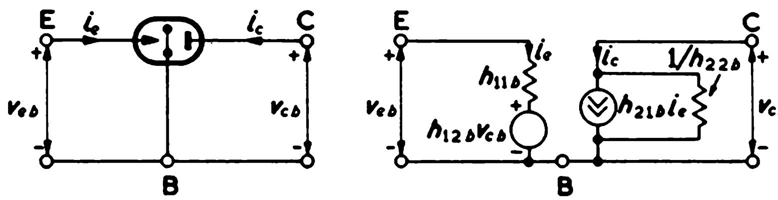


FIG. 9.8

9.6. Transistor Equivalent Circuits

The small signal transistor equations

$$v_{eb} = h_{11b}i_e + h_{12b}v_{cb} \quad \text{and} \quad i_c = h_{21b}i_e + h_{22b}v_{cb}$$

may be used to establish transistor equivalent circuits. It may be verified that the circuit shown in Fig. 9.8 gives the same two equations for v_{eb} and i_c. In this circuit $h_{12b}v_{cb}$ is a voltage generator and $h_{21b}i_e$ a current generator. The voltage generator has series resistance h_{11b} and the current generator parallel conductance h_{22b}. The equivalent circuit may be used to replace the transistor in actual circuits provided the changes of currents and voltages are sufficiently small.

Just as there are many ways of expressing transistor characteristics, so there are many different types of equivalent circuit. Some of these are given in Exx. IX.

9.7. Biasing Circuits

Separate battery supplies have been shown so far for the input and output circuits of transistor amplifiers. In practice, only one supply is used whenever possible. The circuit of Fig. 9.9.*a* shows one method of

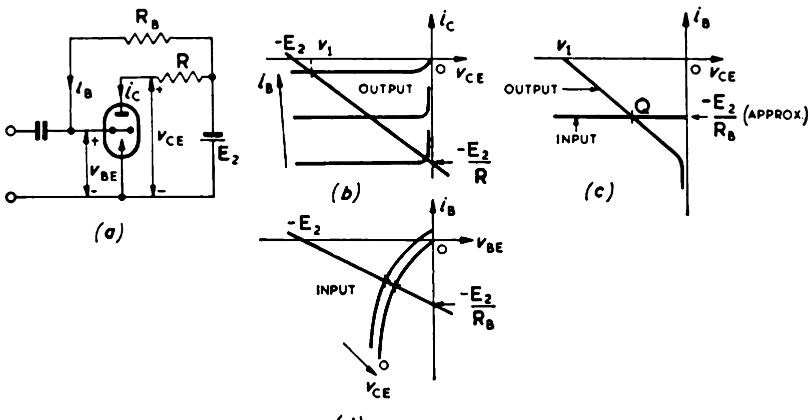


FIG. 9.9

biasing the base of a common-emitter amplifier. The battery E_2 is joined to the collector through the load resistance R as usual. The same battery is connected to the base through the resistance R_B, which controls the quiescent value of the base current and consequently the collector current. In the collector–emitter circuit there are two separate relations between collector current and collector voltage. These are the Load Line Equation

$$v_{CE} = -E_2 - Ri_C$$

and the Transistor Equation

$$i_C = f_4(i_B, v_{CE}).$$

These relations are plotted in Fig. 9.9.*b*, and from them we get the dynamic relation between i_B and v_{CE} for the output circuit; this is shown in Fig. 9.9.*c*. There are also two relations for the base–emitter circuit:

$$v_{BE} = -E_2 - R_B i_B$$

and $$v_{BE} = f_3(i_B, v_{CE}).$$

These are shown in Fig. 9.9.*d*, and from them we plot in Fig. 9.9.*c* the input circuit relation between i_B and v_{CE}. The point Q in this figure gives the quiescent conditions. In practice, v_{BE} is usually much smaller than E_2 in magnitude, so that

$$i_B \simeq -E_2/R_B.$$

Thus i_B is fixed for a given circuit and is independent of the transistor. This circuit is therefore said to have fixed bias. With a given transistor

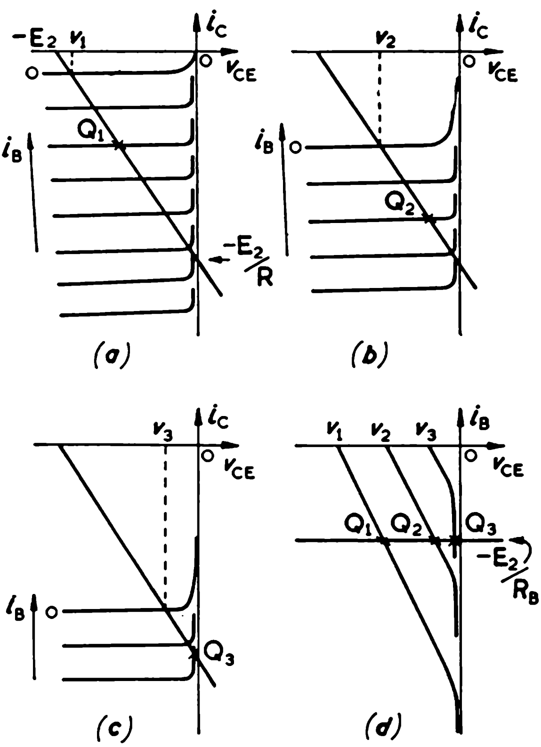


FIG. 9.10

the relation between i_C and v_{CE} at zero base current may vary appreciably with ambient temperature, so that

$$i_C \simeq \alpha_{cb} i_B + i_o$$

where α_{cb} remains almost constant but i_o is temperature dependent. The effect of this variation on the operating point is shown in Fig. 9.10.*a* to *d*, in which R, R_B and E_2 are assumed constant. In case (*c*) it is seen that the bias would be unsuitable for amplification.

An alternative bias circuit, which is less sensitive to these variations in the characteristics, is shown in Fig. 9.11.*a*. The base is fed from the collector through the resistance R_F. The base current now depends on the value of the collector–emitter voltage, which in turn depends on the collector current. This is an automatic bias circuit which operates in such a manner that any tendency for i_C to change is opposed by a change of bias. In most cases i_B is much less than i_C, so that the dynamic relation between i_B and v_{CE} for the output circuit can be determined in

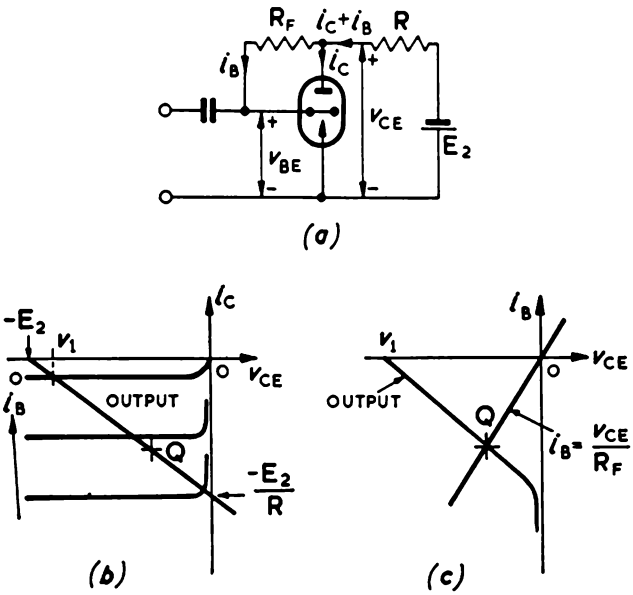


Fig. 9.11

the same manner as in the fixed bias circuit (Fig. 9.11.*b* and *c*). From the circuit we also see that

$$i_B = (v_{CE} - v_{BE})/R_F.$$

Usually $|v_{BE}| \ll |v_{CE}|$, and then

$$i_B \simeq v_{CE}/R_F.$$

This load line is drawn in Fig. 9.11.*c*, and the quiescent point is Q. When the transistor characteristics vary with temperature, Fig. 9.12 shows how the automatic bias circuit behaves. The variation in the operating point is much less than in Fig. 9.10. With the circuit of Fig. 9.11.*a* a.c. variations of collector–emitter voltage are transmitted to the base–emitter circuit. In order to prevent this a decoupling circuit is used, as shown in Fig. 9.13. The bias resistance is divided into two parts, and R_{F2} and C act as an a.c. filter.

The stabilizing effect of the automatic bias circuit on the Q-point depends on the load resistance R being sufficiently large for v_{CE} to vary appreciably with collector current. With a transformer-coupled load the d.c. resistance is negligible. Automatic bias may be obtained in this case with a resistance R_E in the emitter lead, as shown in Fig. 9.14.*a*. Although the bias voltage now depends on the collector current, its polarity is incorrect and it is necessary to provide a counteracting fixed bias by means of resistances R_1 and R_2. The Q-point can now be found

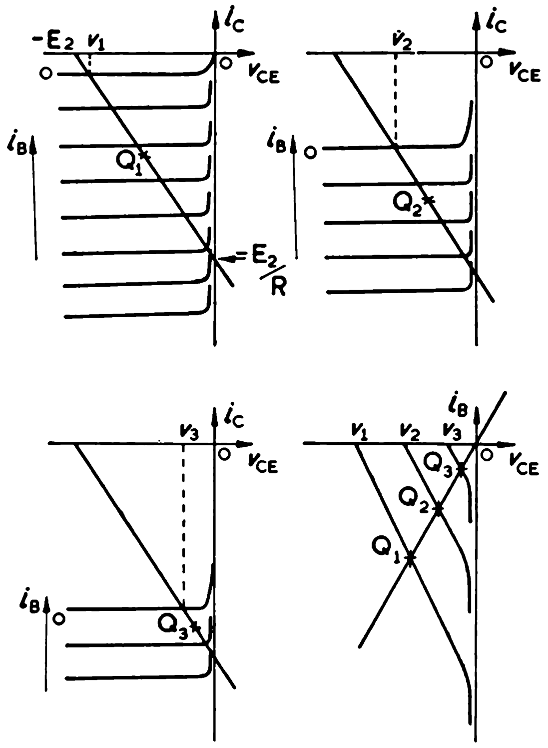


FIG. 9.12

as follows. Since $i_C \gg i_B$ in magnitude, the dynamic relation between v_{CE} and i_B can be found from the collector characteristics and the Load Line Equation

$$v_{CE} = -E_2 - R_E i_C,$$

as shown in Fig. 9.14.*b* and *c*. The voltages across R_1 and R_2 are of the same order as E_2, and normally $E_2 \gg |v_{BE}|$. Hence

$$R_E i_E \simeq R_1 i_1 \quad \text{and as} \quad i_E \simeq -i_C$$

then $$R_1 i_1 = -R_E i_C.$$

On substituting in the Load Line Equation we find

$$i_1 = (E_2 + v_{CE})/R_1.$$

From the circuit it follows that

$$E_2 - R_1 i_1 + R_2(i_B - i_1) = 0$$

and hence

$$i_1 = \frac{E_2 + R_2 i_B}{R_1 + R_2}.$$

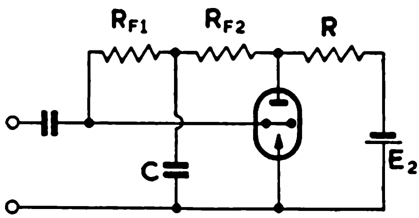


FIG. 9.13

The two expressions for i_1 lead to a second relation between v_{CE} and i_B,

$$v_{CE} = \frac{-E_2 R_2}{R_1 + R_2} + \frac{R_1 R_2}{R_1 + R_2} i_B.$$

This load line is also shown in Fig. 9.14.*c*, and its intersection with the previous v_{CE}, i_B curve gives the Q-point. To prevent the bias voltage

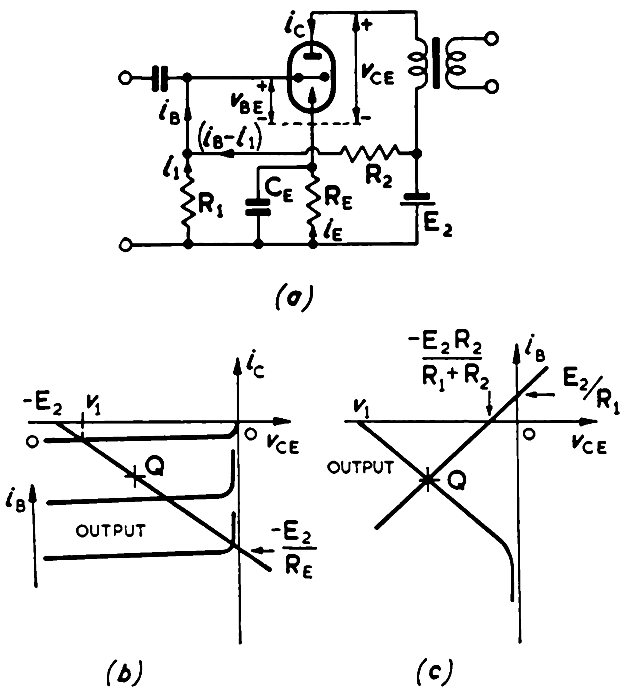


FIG. 9.14

containing a.c. components a capacitor C_E is connected in parallel with R_E.

Although this section has dealt entirely with the common-emitter amplifier, the biasing circuits for the other transistor amplifiers are similar in principle.

9.8. Transistor Amplifiers at High Frequencies

When transistors are operated at sufficiently high frequencies their inherent reactances become appreciable just as in triodes or other electronic devices. The effective capacitances across the *p–n* junctions, i.e., at the boundaries of the electrodes, must be taken into account.

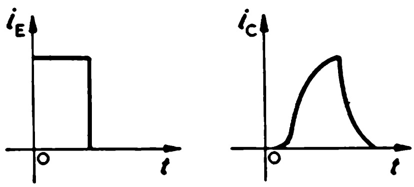


FIG. 9.15

Transistors also show effects due to the time of diffusion of the carriers from the emitter to the collector (holes in the *p–n–p* case). This time is not the same for all carriers owing to differences in diffusion path lengths for individual carriers. Thus if the emitter current is a square pulse, the collector current pulse is delayed and distorted as shown in Fig. 19.15.

For a.c. signals the effect of capacitances and of diffusion can be expressed in terms of variation of magnitude and phase angle of the transistor parameters. In particular, α_{ce} falls with increasing frequency.

CHAPTER 10

FEEDBACK

10.1. Feedback

In conventional amplifiers the output is much greater than the input signal which is connected between the grid and cathode of the first valve. A small part of this output may be transferred back to the input and put in series with the signal across the grid and cathode terminals. The effect on the amplifier then depends on the phase relation between the fed-back voltage and the signal. If the fed-back voltage is such that the grid–cathode voltage exceeds the signal, higher output and gain are obtained; alternatively, the signal may be reduced to give the same output as is obtained with no feedback. With these conditions it is said that positive feedback has been introduced. If the voltage fed back is exactly equal to and of the same phase as the grid–cathode voltage, then the signal may be reduced to zero, and the output is the same as without feedback. Such an amplifier is called a self-oscillator, as it provides an output with no external signal. If the fed-back voltage is such that the grid–cathode voltage is smaller than the signal, the output voltage drops and the gain is reduced. In this case there is negative feedback. In all cases it may be noted that the output of the amplifier depends only on the grid–cathode voltage. The valve amplifies the voltage appearing between its grid and cathode irrespective of how that voltage is produced. The inherent gain of the amplifier, i.e., v_o/v_g, is independent of feedback; it is v_o/v_s which is affected.

At first sight it might appear that negative feedback is undesirable, since it reduces the output for a given signal input. However, it may also have certain very desirable features which more than offset the loss of gain. Some of these features are: (i) greater stability against supply variations, (ii) independence of changes in valves, (iii) reduction in noise such as hum, (iv) reduction of frequency and phase distortion, (v) reduction in non-linear distortion, and (vi) possibility of achieving some particular frequency response. It is the main purpose of this chapter to study the effects of negative feedback. Positive feedback in oscillators is dealt with in Chapter 13.

10.2. Automatic Bias

In valve amplifiers the grid bias can be obtained automatically by putting a resistance in series with the cathode lead. We are now going to examine further some of the effects of automatic bias. In the first place we consider the circuit of Fig. 10.1.*a* with no grid battery and with the input terminals short-circuited. The output from such a circuit is normally taken from the terminals AB. It may be seen that $v_G = -R_K i_A$ and the output voltage is $v_O = v_A + R_K i_A$. Thus the grid voltage is part of the output voltage and we have a case of feedback. It is obvious that this circuit helps to stabilize the anode current against changes of the

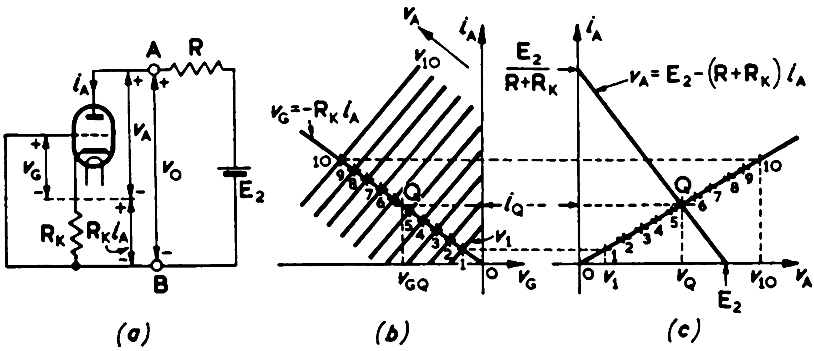


Fig. 10.1

supply voltage E_2. If, for example, E_2 increases, then i_A increases, and so does the grid bias, thus offsetting to some extent the original change.

The method of determining the Q-point is shown in Fig. 10.1.b and c. Besides being related by the grid characterstics, i_A and v_G satisfy the equation $v_G = -R_K i_A$. This is a grid-circuit load line, and it is drawn in Fig. 10.1.b. The grid characteristics are drawn as parallel straight lines for convenience. The points where the grid load line cuts each characteristic are then transferred to the accompanying i_A, v_A diagram (Fig.

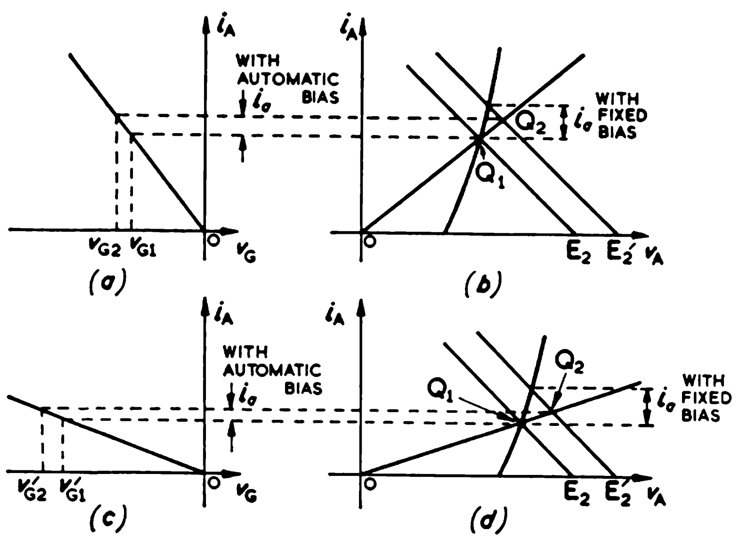
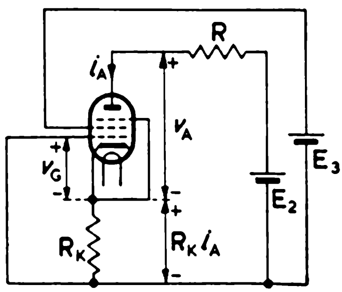


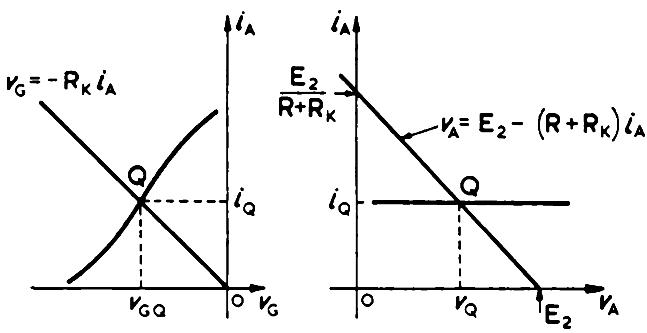
Fig. 10.2

10.1.c). Between i_A and v_A there is also the anode load line relation $v_A = E_2 - (R + R_K)i_A$. This, too, is plotted on the i_A, v_A diagram. The intersection of the two lines gives the Q-point for the given values of E_2, R and R_K.

If the supply voltage E_2 changes, the effect may be determined as shown in Fig. 10.2.a and b. The anode current and anode voltage both increase when E_2 changes from E_2 to E_2'. At the same time the bias voltage in-increases and i_a, the change in anode current, is smaller than it would have been if a fixed battery supply had been used for bias. The change that would have occurred with fixed bias may be found from the anode



(a)



(b)　　　　(c)

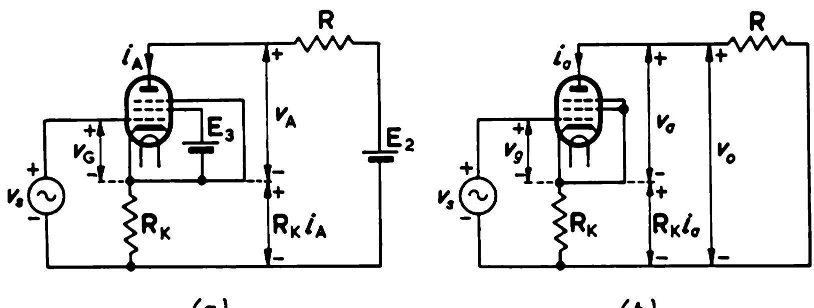
FIG. 10.3

characteristic through Q_1 as shown. The greater the value of R_K, the smaller the change in i_A for a given change in E_2. This is shown in Fig. 10.2.c and d, where the value of R_K has been doubled.

For a pentode the i_A, v_G characteristic is practically independent of v_A, except at very low values of v_A. The transfer from the i_A, v_G plot to i_A, v_A is therefore a horizontal straight line (see Fig. 10.3). The Q-point is found as before. Since the grid load line gives a horizontal line in the i_A, v_A diagram, it might appear that the cathode resistor does not contribute to stabilization of i_A against supply voltage changes. However, the screen electrode in a pentode corresponds to the anode of a triode, and hence the bias resistor stabilizes i_A against changes in the screen voltage.

10.3. Automatic Bias and Signal Feedback

The effect of automatic bias on a small signal may be determined with reference to Fig. 10.4.*a*, which shows the circuit of a pentode amplifier and the total currents and voltages at various points in the circuit. As



(*a*) (*b*)

FIG. 10.4

we are considering the effect of the signal, we can redraw the circuit as in Fig. 10.4.*b*, which shows only the varying components of the currents and voltages. The input to this amplifier is v_s and the output v_o, so that the voltage amplification is given by $A_f = v_o/v_s$. It may be found from the figure that $v_o = -Ri_a$, $v_a = -(R + R_K)i_a$ and $v_g = v_s - R_K i_a$. Since we are dealing with small changes, the Valve Equation can be used, giving $i_a = g_m v_g + v_a/r_a$.

$$i_a = g_m(v_s - R_K i_a) - (R + R_K)i_a/r_a,$$

i.e.,
$$i_a = g_m v_s/\{1 + g_m R_K + (R + R_K)/r_a\}.$$

$$A_f = -Ri_a/v_s = -g_m R/\{1 + g_m R_K + (R + R_K)/r_a\}.$$

With a pentode $r_a \gg R + R_K$, and then

$$A_f = -g_m R/(1 + g_m R_K).$$

There are several interesting points which may be made about this circuit. Firstly, this is a case of negative feedback. The voltage appearing between grid and cathode, when the signal is applied, is not v_s but v_s reduced by $R_K i_a$. The reduction is proportional to the output voltage $-Ri_a$. The fraction of the output voltage fed back to the input is R_K/R. There is also a reduction in the voltage amplification brought about by the presence of R_K. If the bias had been obtained from a battery, then the voltage amplification would have been $A = v_o/v_g = -g_mR$, as is shown in Section 7.3. We thus find that the voltage amplification with feedback is given by

$$A_f = \frac{A}{1 - \dfrac{AR_K}{R}}.$$

If the inherent gain A of the amplifier is sufficiently great for $AR_K/R \gg 1$ numerically, then the gain with feedback is $A_f = -R/R_K$. Thus the amplifier gain depends only on the ratio of two circuit components and is independent of the valve constants. The valve may change with age or be replaced, or the power supply voltage may vary, but as long as $|A|$ is sufficiently great, the amplifier gain is unaffected. This most desirable effect is obtained at the cost of considerable loss of gain. If, for example, $R = 20{,}000\ \Omega$ and $R_K = 2{,}000\ \Omega$, then $A_f = -10$. If $g_m = 10$ mA/V then $A = -200$; $AR_K/R = -20$, so that the condition for stability is satisfied fairly well, but the gain is reduced from 200 to 10.

10.4. Cathode Bias Condenser

When automatic bias is introduced in Section 7.11 in connection with a.c. amplifiers it is stated that the cathode resistor R_K is shunted by a condenser C_K whose reactance is much less than R_K at the operating

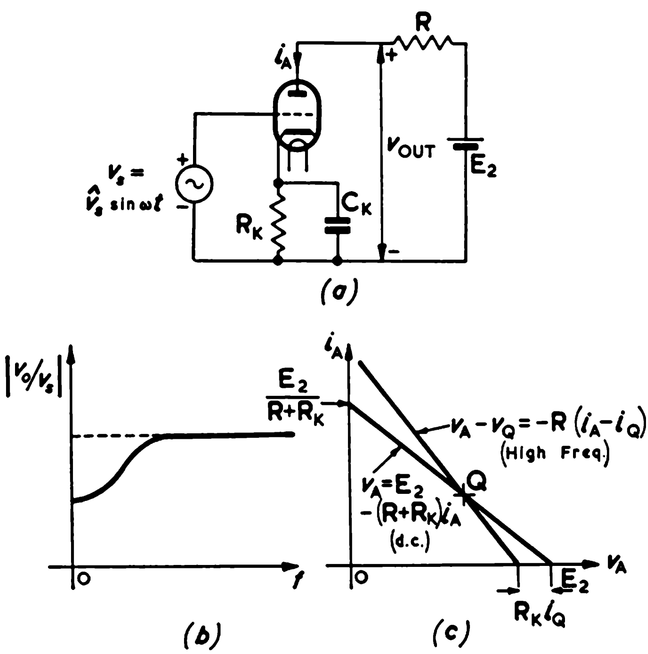


FIG. 10.5

frequency. Under these conditions there is no appreciable a.c. voltage drop across R_K (see Fig. 10.5.*a*). The amplifier behaves as though it had fixed bias equal to $-R_K i_Q$, and there are no negative feedback effects on the signal. However, the feedback still operates for d.c. changes, and the Q-point must be determined by the method described in Section 10.2. Also, there is feedback at low frequencies where $1/\omega C_K$ is not much less than R_K. The voltage amplification varies with frequency, as shown in Fig. 10.5.*b*. The anode load line varies from $v_A = E_2 - (R + R_K)i_A$ with d.c. to $v_A = E_2 - R_K i_Q - Ri_A$ at high frequencies (see Fig. 10.5.*c*). The Q-point is the same in both cases.

10.5. Feedback—General Considerations

In Fig. 10.6.*a* a conventional amplifier is shown diagrammatically with input terminals v_g and output terminals v_o. The amplifier may be of any type with one or more stages. A signal v_s is shown connected to the input terminals. In Fig. 10.6.*b* the amplifier has part of the output fed back to the input through a feedback network. The input to the amplifier now consists of the signal v_s and the feedback voltage βv_o in series, so that $v_g = v_s + \beta v_o$. The signs in this diagram define the polarities of the various voltages. In the analysis of the network it is convenient to assume a sinuosoidal a.c. signal and to use vector quantities. The phase of $\mathbf{V}_0$ depends on any phase changes introduced in the amplifier. The feedback network may also introduce phase change so that a phase angle must be assigned to $\boldsymbol{\beta}$. We define $\mathbf{A}_f$ and $\mathbf{A}$ to be the voltage amplifications for the amplifier with and without feedback where

$$\mathbf{A}_f = \mathbf{V}_0/\mathbf{V}_s \quad \text{and} \quad \mathbf{A} = \mathbf{V}_0/\mathbf{V}_g.$$

$\mathbf{A}_f$ and $\mathbf{A}$ are also vector quantities. Since $\mathbf{V}_g = \mathbf{V}_s + \boldsymbol{\beta}\mathbf{V}_0$,

$$\mathbf{A}_f = \mathbf{V}_0/(\mathbf{V}_g - \boldsymbol{\beta}\mathbf{V}_0) = \frac{\mathbf{V}_0/\mathbf{V}_g}{1 - \boldsymbol{\beta}\mathbf{V}_0/\mathbf{V}_g} = \mathbf{A}/(1 - \boldsymbol{\beta}\mathbf{A}).$$

Positive or negative feedback is then defined according as $|\mathbf{A}_f/\mathbf{A}|$ is greater or less than unity

i.e., $\quad |1 - \boldsymbol{\beta}\mathbf{A}| < 1$ gives positive feedback

and $\quad |1 - \boldsymbol{\beta}\mathbf{A}| > 1$ gives negative feedback.

These conditions are illustrated in Fig. 10.6.*c* and *d*. In these figures OB, OD and DB are vectors representing the grid voltage, the feedback voltage and the signal voltage respectively. Fig. 10.6.*c* represents negative feedback, since DB is greater than OB; on the other hand, Fig. 10.6.*d* represents positive feedback. It is obvious that positive or negative feedback occurs according as the point D lies inside or outside the circle with its centre at B and its radius equal to OB.

When all the reactances in the amplifier and feedback circuits are negligible in comparison with the resistances, then the voltages at various points in the circuits are either in phase or differ by 180°. A vector diagram showing negative feedback in this case is drawn in Fig. 10.6.*e*.

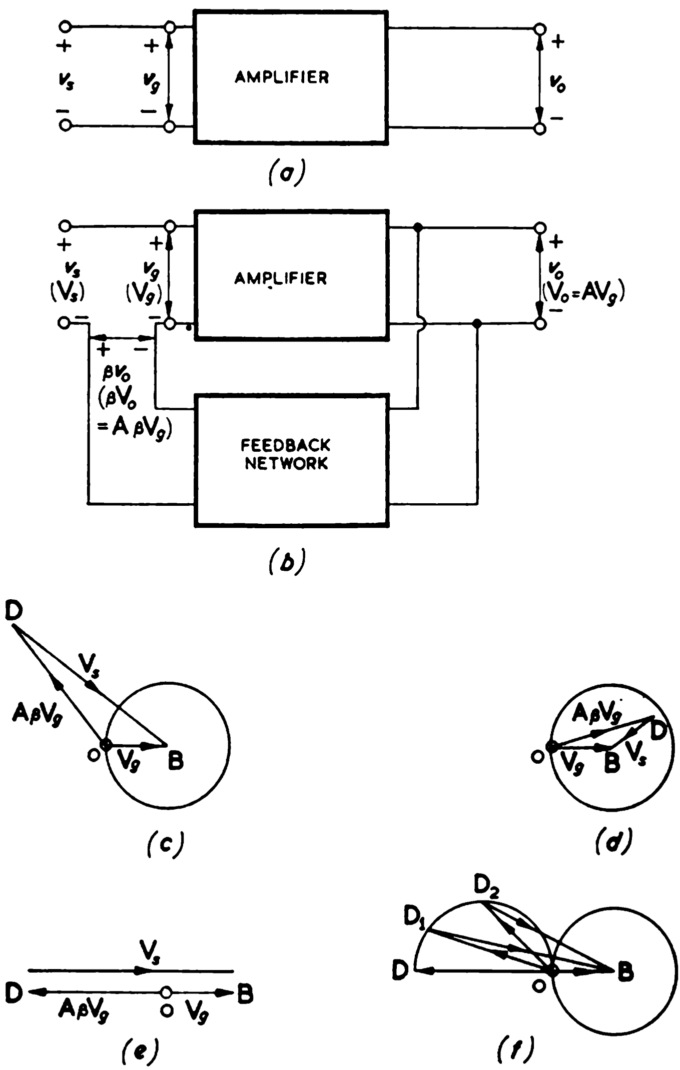


FIG. 10.6

In Section 10.3 we consider a simple amplifier of this type in which $\beta = R_K/R$. However, in any amplifier reactances become effective at some frequencies, and then the feedback conditions change. We have one example of this in the cathode bias circuit with a condenser in Section

10.4. At the higher frequencies the feedback is zero, and the amplifier gain is independent of frequency (see Fig. 10.5.*b*). With d.c. and at very low frequencies there is feedback due to R_K, and the gain is constant but at a lower level. In the intervening range the feedback and the gain vary with frequency. At the same time the phase angle of the voltage across R_K is no longer 180° relative to $\mathbf{V}_s$, but varies from 180° to 90° as the frequency rises. When the phase angle is 90°, the amplitude of $\boldsymbol{\beta}$ is zero and the feedback is negligible. The vector diagram in this case is shown in Fig. 10.6.*f*. As the frequency rises from zero the point D moves round the curve $D\ D_1\ D_2\ O$.

10.6. Effect of Feedback on Non-linear Distortion

When an amplifier with a curved dynamic characteristic is used with large values of v_g, non-linear distortion occurs. This distortion arises entirely from the nature of the valve characteristic and the size of v_g. For a given output the size of v_g is fixed and is given by v_o/A, where A is the inherent voltage amplification of the valve amplifier with resistance load. None of these factors is affected by feedback, and so feedback does not affect the amount of non-linear distortion produced by the valve. However, the amount of distortion in the output may be reduced by negative feedback provided sufficient signal amplitude is available. This may be explained as follows. If v_o is the fundamental output required from the valve, then $v_g = v_o/A$. If there is no feedback $v_g = v_s$. We assume that under these conditions a harmonic of voltage v_H is produced in the output. Now let us apply negative feedback to the amplifier and at the same time increase v_s to give the same fundamental output v_o as before. The valve has the same v_g as before and produces the same distortion. However, the harmonic voltage is also subject to feedback, so that the total harmonic content in the output is due to the part produced by the valve and the part fed back and re-amplified. If v_H' is the total harmonic voltage in the output with feedback, then $\beta v_H'$ is fed back and amplified A times. Thus

$$v_H' = v_H + A\beta v_H'.$$

$$v_H' = v_H/(1 - A\beta).$$

The harmonic distortion may therefore be reduced to a small value if the feedback is negative and the magnitude of $A\beta$ is sufficiently great. It is assumed above that the reactive components of the feedback amplifier have no effect and that the voltage amplification is the same for the fundamental and the harmonic. We have also ignored the effect of the non-linearity of the amplifier on the harmonic component which is fed back. This is obviously a second-order effect.

The reduction in non-linear distortion with feedback requires a greater signal v_s. However, this may be obtained from a voltage amplifier operating at a level at which little distortion occurs.

10.7. Effect of Feedback on Frequency Distortion and Noise

If an amplifier has an excessive gain at some particular frequency, then there is a greater voltage fed back at this frequency, and this offsets the gain to some extent. The quantitative effect of negative feedback on frequency distortion can be found readily from the feedback equation. If $\mathbf{A}_{f1}$ and $\mathbf{A}_1$ are the voltage amplification with and without feedback at one frequency, and $\mathbf{A}_{f2}$ and $\mathbf{A}_2$ the corresponding values at a second frequency, then

$$\mathbf{A}_{f1} = \mathbf{A}_1/(1 - \boldsymbol{\beta}_1\mathbf{A}_1) \text{ and } \mathbf{A}_{f2} = \mathbf{A}_2(1 - \boldsymbol{\beta}_2\mathbf{A}_2),$$

$$\therefore \quad \mathbf{A}_{f1}/\mathbf{A}_{f2} = \mathbf{A}_1(1 - \boldsymbol{\beta}_2\mathbf{A}_2)/\mathbf{A}_2(1 - \boldsymbol{\beta}_1\mathbf{A}_1).$$

If, with negative feedback, the magnitudes of $\boldsymbol{\beta}_1\mathbf{A}_1$ and $\boldsymbol{\beta}_2\mathbf{A}_2$ are sufficiently great compared with unity, then, provided $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$, $\mathbf{A}_{f1} = \mathbf{A}_{f2}$.

Negative feedback may improve the signal-to-noise ratio (S/N) of an amplifier if the noise is produced in the amplifier. For example, if hum is introduced in one of the stages of a multi-stage amplifier, then negative feedback may be used to reduce the amount of hum in the output. This feedback is applied between the output and a point just before the origin of the hum. The hum is then reduced in the same way as harmonic distortion. The gain of the earlier stages of the amplifier is now increased to restore the output voltage of the whole amplifier to the original level, using the same signal amplitude as before. The value of S/N at the output stage is improved. This effect on noise assumes that the extra gain in the early stages may be obtained without introducing comparable noise.

If feedback is introduced in the early stages of a high-gain amplifier, S/N may be reduced if more valves have to be used to obtain the required gain; this increases the proportion of noise due to the valves (see Chapter 20).

10.8. Current and Voltage Feedback

In the feedback circuit with a cathode resistor the voltage fed back to the input is proportional to the output current. This is an example

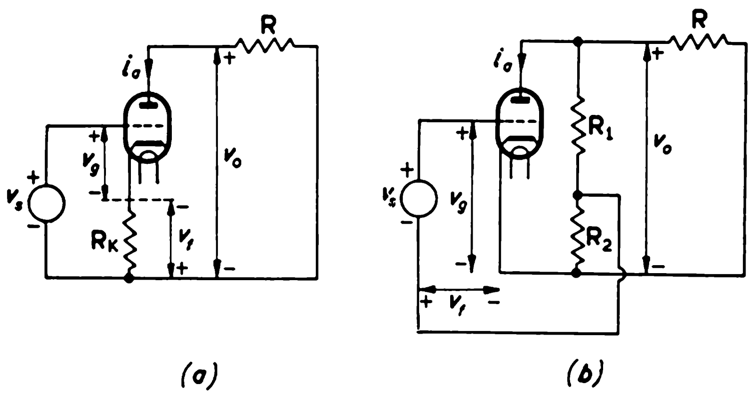


FIG. 10.7

of a single-valve amplifier with current negative feedback (Fig. 10.7.*a*). An alternative circuit is shown in Fig. 10.7.*b*. Here $\beta v_o = v_o R_2/(R_1 + R_2)$, and the fed-back voltage is proportional to the output voltage. This is a single-valve amplifier with voltage negative feedback. To check the

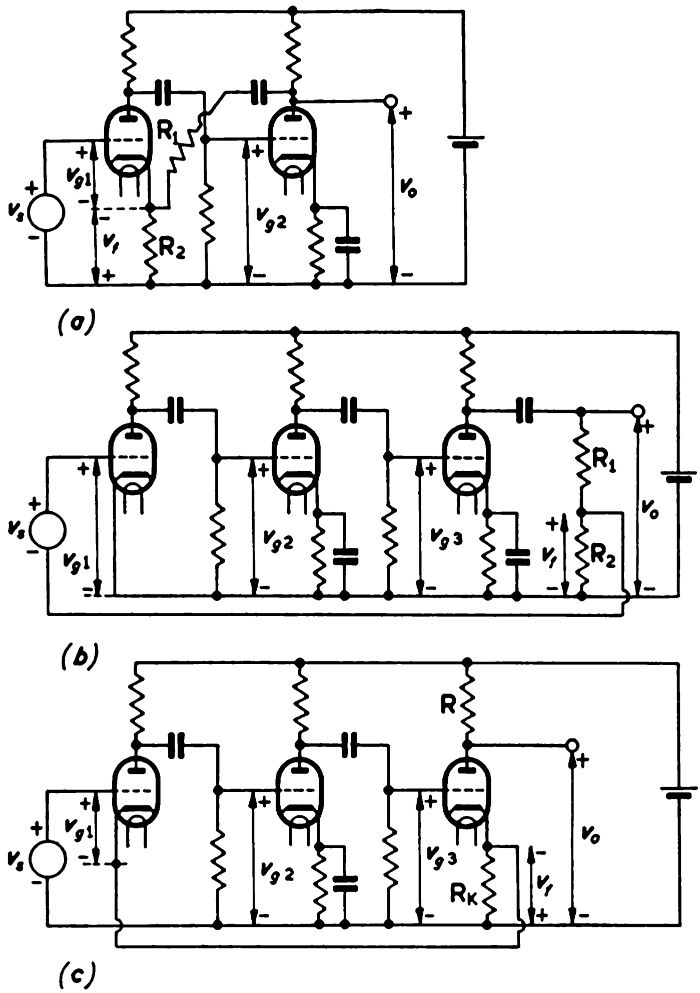


FIG. 10.8

nature of the feedback in resistance-loaded amplifiers it is convenient to assume that the effects of the reactances are negligible. Then the currents and voltages throughout the circuit are either in phase or in antiphase with the signal. It should be verified that all the circuits in Figs. 10.7 and 10.8 give negative feedback.

The common-emitter transistor amplifier in Fig. 9.14.*a* has current negative feedback when the capacitor C_E is omitted. There is also some feedback inherent in the transistor itself. This is measured in terms of $h_{12e} = \partial v_{BE}/\partial v_{CE}$.

10.9. Output and Input Impedance of Feedback Amplifiers

In any generator the output obtainable from it depends on the relative values of the generator internal resistance and the load resistance. For maximum power output the load resistance should equal the generator resistance. For maximum voltage output the load resistance should be much greater than the generator resistance. For a generator with low internal impedance the value of the load impedance may vary considerably, with frequency or otherwise, and the change in output voltage is small as long as the load impedance is high compared with the generator impedance. On the other hand, if the generator impedance is much greater than the load impedance the output current does not vary much when the load varies. The internal impedance of a generator is therefore an important property. A conventional valve amplifier acts as a voltage generator of internal resistance r_a feeding the anode load, and the

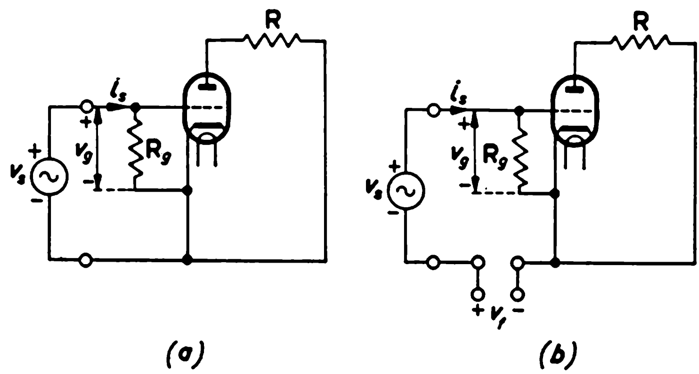


FIG. 10.9

amplifier is said to have an output impedance equal to r_a. We are now going to see the effect of negative feedback on the amplifier internal impedance or output impedance.

When the load impedance of an amplifier or any other generator changes, the output voltage and output current both change. For example, if the load impedance rises the output voltage increases and the output current decreases. If this amplifier has voltage negative feedback, then the voltage fed back to the input increases, the effective voltage on the grid of the first valve decreases and the output voltage of the amplifier drops. Thus the output voltage tends to remain constant.

This is equivalent to saying that an amplifier with voltage feedback has a low internal impedance. The cathode follower is an example (see Section 10.10).

If the load impedance rises in the case of an amplifier with current negative feedback, then the decrease in output current brings about a reduction in feedback which offsets the change in output current. Thus with current feedback the amplifier tends to give constant current output, i.e., it behaves as though it has a high internal impedance.

The input impedance of an amplifier, which is defined in Section 7.13, is another property of considerable importance, since it determines how much the amplifier acts as a load on the source of the signal. If a signal $\mathbf{V}_s$ is connected to the input terminals and a current $\mathbf{I}_s$ flows as shown in

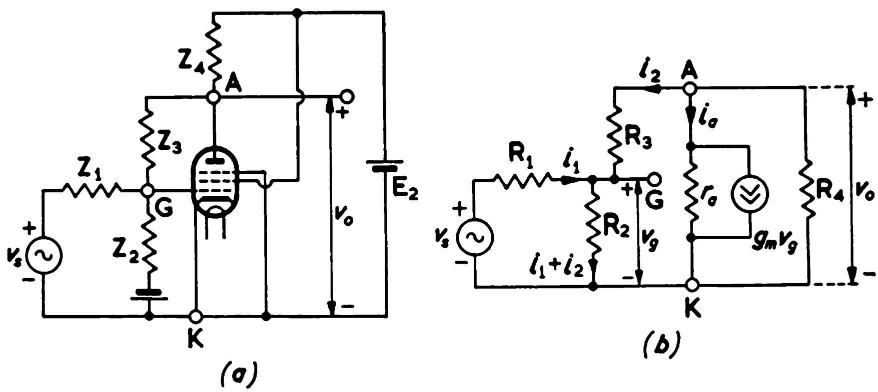


FIG. 10.10

Fig. 10.9.*a*, then the ratio $\mathbf{V}_s/\mathbf{I}_s$ is the input impedance. For a conventional amplifier of the type shown, the input impedance is R_g. When no resistance is connected in the circuit, R_g is the leakage resistance across the valve insulators. The effect of the inter-electrode capacitances is considered in Section 10.11.

In all the cases of negative feedback which we have considered so far the feedback voltage has been connected in series with the input voltage between the grid and cathode. This always results in an increase in the input impedance of the amplifier. In the circuit in Fig. 10.9.*b*, $|\mathbf{V}_g|$ is less than $|\mathbf{V}_s|$, since the feedback is negative; $\mathbf{I}_s$ is the current taken from the generator and $\mathbf{I}_s = \mathbf{V}_g/R_g$. This is less than the current $\mathbf{V}_s/R_g$ in the circuit without feedback, so that negative feedback increases the input impedance.

It is possible to have feedback connected in parallel with the signal voltage, and this may result in a decrease of input impedance. The Miller Effect, described in Section 10.11, arises from this type of feedback. Another example of parallel voltage feedback is shown in Fig. 10.10.

The impedances $\mathbf{Z}_1$ $\mathbf{Z}_2$, $\mathbf{Z}_3$ and $\mathbf{Z}_4$ are, in general, combinations of resistance and capacitance; $\mathbf{Z}_3$ is the feedback impedance. The grid voltage is derived from the signal through $\mathbf{Z}_1$ and also from the output through $\mathbf{Z}_3$, so that the feedback and the signal are essentially in parallel. In this case the value of $\boldsymbol{\beta}$ cannot be written down easily, and it is necessary to proceed from the basic circuit equations. For simplicity it is assumed that the impedances are purely resistive. If the inherent gain of the valve is large, then approximately

$$\mathbf{I}_2 = \mathbf{V}_0/R_3.$$

Also
$$\mathbf{I}_1 = (\mathbf{V}_s - \mathbf{V}_g)/R_1.$$

When there is appreciable negative feedback $\mathbf{V}_g \ll \mathbf{V}_s$ and then

$$\mathbf{I}_1 \simeq \mathbf{V}_s/R_1.$$

If R_2 is sufficiently large

$$\mathbf{I}_1 = -\mathbf{I}_2.$$

Then
$$\mathbf{A}_f = \frac{\mathbf{V}_0}{\mathbf{V}_s} \simeq -\frac{R_3}{R_1}.$$

The feedback is negative and, as in previous feedback amplifiers, the gain can be indépendent of the valve parameters. The input impedance of the amplifier is approximately equal to R_1, as compared with $R_1 + R_2$ when the feedback resistance R_3 is removed. The output impedance is low, as it is a case of voltage feedback. The circuit of Fig. 10.10.*a* has important applications in electronic computers (see Exx. XIX).

Another example of an amplifier with low input impedance is the common-grid amplifier, which is discussed in Section 7.12; it is shown in Fig. 10.11.*a* and its equivalent circuit in Fig. 10.11.*b*. This amplifier

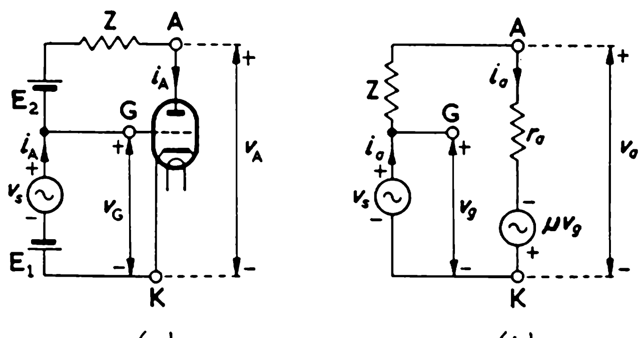


FIG. 10.11

is one with feedback in which the whole of the anode current flows through the signal circuit. Thus the input impedance is $\mathbf{V}_s/\mathbf{I}_s = (r_a + \mathbf{Z})/(\mu + 1)$. When the valve is a pentode and $r_a \gg |\mathbf{Z}|$, then the input impedance is $r_a/(\mu + 1)$, which is approximately equal to $1/g_m$. Since g_m usually lies between 1 and 10 mA/V, the input impedance of a common-grid amplifier is of the order of a few hundred ohms.

A further example of parallel negative feedback is the common-emitter transistor amplifier with automatic bias, which is given in Fig. 9.11.*a*. The feedback occurs through the resistance R_F.

10.10. The Cathode Follower

The cathode follower is discussed in Section 7.12 as a common-anode amplifier in which the signal is connected between grid and anode and the output is taken between cathode and anode. The circuit is shown in

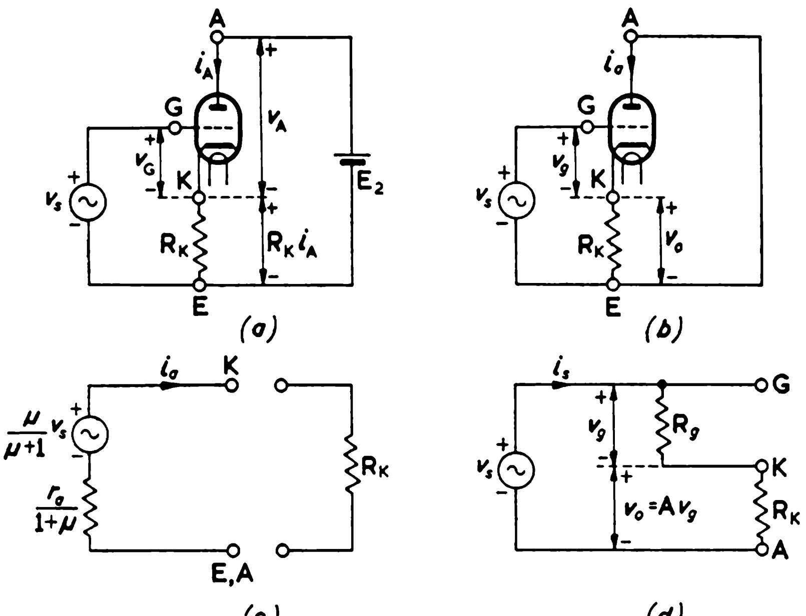


FIG. 10.12

Fig. 10.12.*a*, and in its form for small changes in Fig. 10.12.*b*. The cathode follower is a special case of a negative feedback amplifier in which the whole of the output voltage is fed back to the input. It is a voltage negative feedback amplifier in which $\beta = -1$. In Fig. 10.12.*b* the valve acts as a normal amplifier of its grid voltage v_g with R_K as the load resistance. As usual, $v_o = Av_g$, where A is large. Since $v_g = v_s - v_o$, it follows that v_s and v_o are nearly equal, and that v_s and v_o are

in phase relative to their common point E. Thus the cathode voltage is in phase with the signal and is practically equal to it; hence the name cathode follower. Since this is a case of voltage feedback, the cathode follower has a low output resistance.

Quantitative values for the voltage amplification and the output resistance may be established quite readily. The voltage amplification follows directly from the feedback equation by putting $\beta = -1$. Then $A_f = A/(1 + A)$. This is always less than unity when A is positive, as it is when R_K is purely resistive. The value of A may be found from the formula derived in Section 7.3 by putting R_K for R. Then

$$v_o/v_g = A = \mu R_K/(R_K + r_a).$$

The positive sign is taken in this formula, since the polarity of the output voltage is reversed when the load is in the cathode lead instead of its usual place in the anode lead. Also, since $v_g = v_s - v_o$, we get, after substitution and rearrangement,

$$\frac{\mu}{\mu + 1} v_s = R_K i_a + \frac{r_a}{\mu + 1} i_a.$$

In this equation the first term is an e.m.f., the second is the voltage drop due to the current i_a flowing in the load resistance R_K and the third is the voltage drop in a resistance $r_a/(\mu + 1)$ with the same current flowing through it. The term $R_K i_a$ gives the output voltage of the cathode follower. The same equation would have been obtained for a generator of e.m.f. $\mu v_s/(\mu + 1)$ with internal resistance $r_a/(\mu + 1)$ feeding a load R_K with current i_a. Thus the cathode follower acts as though it were a generator of internal resistance $r_a/(\mu + 1)$ and e.m.f. $\mu v_s/(\mu + 1)$, as shown in Fig. 10.12.*c*. Usually $\mu \gg 1$, particularly for pentodes, and hence the e.m.f. is approximately equal to v_s and the internal resistance is r_a/μ or $1/g_m$. Thus the output impedance of the cathode follower is of the order of 100 to 1,000 ohms. This low output impedance means that the cathode follower can give an output voltage which is independent of the load impedance for wide variations of the latter. When a reactive or low impedance load is to be coupled to a voltage amplifier a cathode follower is sometimes put between the output of the amplifier and the load.

It may easily be shown from Fig. 10.12.*d* that the input resistance of a cathode follower is approximately equal to $R_g(1 + A)$, where R_g is the resistance between the grid and the cathode.

Since v_s is nearly equal to v_o, and the input resistance of the cathode follower is very high, whilst the load resistance may be low, it follows that the current or the power in the load may be much greater than the current or the power of the signal. Thus, although it cannot give voltage amplification, the cathode follower may, like the conventional amplifier, be used for current or power amplification.

10.11. The Miller Effect

In a valve there is some capacitance between the anode and grid electrodes. In a conventional common-cathode amplifier this provides coupling or feedback between the output and input circuits. This is a case where the feedback is in parallel with the signal. The phase of the feedback depends on the phase difference between the input and output circuits. For an amplifier with resistance load of the type shown in Fig. 10.13.*a* the phase difference is very nearly 180°. When the current through C_{ag} is small compared to the current through the valve, the assumption of 180° phase difference is still justified. If $\mathbf{V}_s$ is the input voltage and $|\mathbf{A}|$ is the magnitude of the voltage amplification, then the output

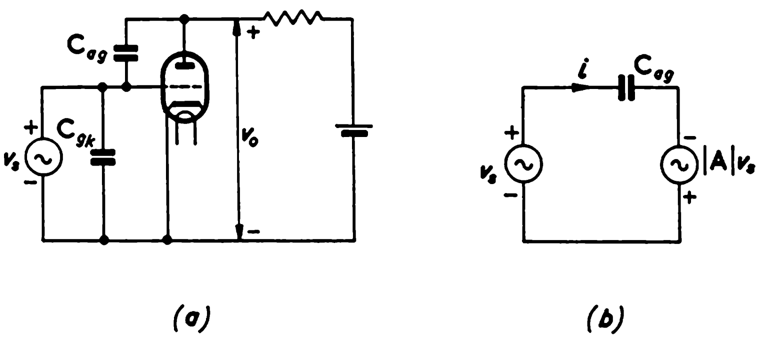


FIG. 10.13

voltage is $-|\mathbf{A}|\mathbf{V}_s$. These two voltages have one common terminal and the other terminals are separated by C_{ag}, the anode–grid capacitance, as shown in Fig. 10.13.*b*. There is, therefore, across C_{ag} a potential difference of $(|\mathbf{A}| + 1)\mathbf{V}_s$. The current $\mathbf{I}$ through C_{ag} is equal to $j\omega C_{ag}(|\mathbf{A}| + 1)\mathbf{V}_s$. Thus, when the signal $\mathbf{V}_s$ is applied to the input terminals, it has to supply a reactive current as though it were connected to a capacitance of value $(|\mathbf{A}| + 1)C_{ag}$. The signal is also connected directly across C_{gk}, the grid–cathode capacitance, so that the effective input capacitance of the amplifier is $C_{gk} + (|\mathbf{A}| + 1)C_{ag}$. In a triode C_{gk} and C_{ag} may be about 10 μμF each. If $|\mathbf{A}| = 20$, then the input capacitance is 220 μμF. This capacitance is across the load resistance of the previous valve, and may cause considerable frequency distortion. In pentodes C_{ag} is of the order of 0·01 μμF, and pentodes are therefore much better than triodes in amplifiers which have to be used at high frequencies. This marked effect of the feedback through the anode-grid capacitance in a valve amplifier is called the Miller effect.

The input capacitance of a cathode follower may be found by a method similar to that used for determining the input resistance. The value is approximately $C_{ag} + C_{gk}/(1 + |\mathbf{A}|)$. Thus a cathode follower amplifier has a much smaller input capacitance than a common-cathode amplifier.

10.12. Stability with Negative Feedback—Nyquist Diagram

The voltage gain of an amplifier with feedback is given by the equation,

$$\mathbf{A}_f = \mathbf{A}/(1 - \boldsymbol{\beta}\mathbf{A}).$$

In much of this chapter we have assumed that we are operating with resistive loads and resistive feedback networks, and that any reactance effects are negligible. It is quite easy to design an amplifier to satisfy these conditions and give negative feedback over a required frequency band, such as the audio-frequencies for sound reproduction. However, at very low frequencies, reactances such as those of the grid coupling condensers and the cathode and screen decoupling condensers have some

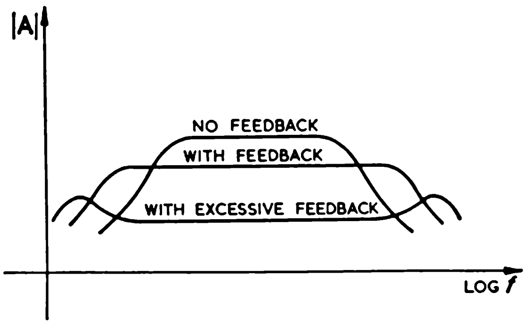


FIG. 10.14

effect. The magnitude of **A** is altered, and at the same time the phase angle of $\mathbf{A}\boldsymbol{\beta}$ is affected so that the feedback may become positive. Unless **A** has then become very small, self-oscillation may occur at a low frequency. The same possibility arises at high frequencies outside the wanted range, due to stray capacitances shunting the load resistances. Even though self-oscillation does not occur, there may be a rise in gain of the amplifier at low and high frequencies, as shown in Fig. 10.14.

The behaviour of a feedback circuit and a criterion for its stability may be determined by means of a Nyquist diagram. This is based on a vector diagram of the type shown in Fig. 10.6.*c* to *f*. The starting point is $\mathbf{V}_g$, which is drawn as a unit vector *OB* in Fig. 10.15. The output $\mathbf{V}_0$ is the vector **A**, and $\boldsymbol{\beta}\mathbf{A}$ is the feedback voltage. Since $\mathbf{V}_g = \mathbf{V}_s + \boldsymbol{\beta}\mathbf{V}_0$, the vector $1 - \boldsymbol{\beta}\mathbf{A}$ represents the signal voltage. In normal operation the magnitude of $\boldsymbol{\beta}\mathbf{A}$ is much greater than unity. Over the required operating range $\boldsymbol{\beta}\mathbf{A}$ is negative and in antiphase with *OB*, as shown by *OC*; *CB* then represents the signal. As the frequency departs from the operating range, $\boldsymbol{\beta}\mathbf{A}$ has a phase angle differing from *OC*; *OD* represents $\boldsymbol{\beta}\mathbf{A}$ at a lower frequency and *DB* is the vector difference $1 - \boldsymbol{\beta}\mathbf{A}$. If the magnitude of *DB* exceeds *OB*, then $\mathbf{A}_f$ is less than **A** and the feedback is negative; i.e., if *D*, the end of the $\boldsymbol{\beta}\mathbf{A}$ vector, lies outside the unit circle round *B* the feedback is negative. At the point or points where the locus of D cuts the OB line to the right of O, the feedback voltage is in phase with the grid voltage and with the signal voltage. For intersections between O and B there is an increase in gain but the amplifier

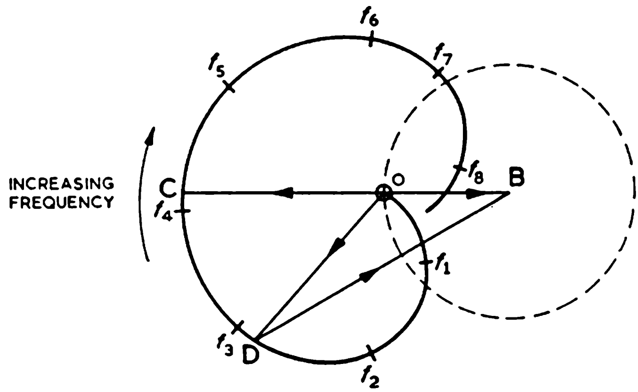


FIG. 10.15

remains stable. If the locus of D passes through B, then the feedback voltage equals the grid voltage, and no signal is necessary to maintain the output, i.e., the amplifier oscillates. Also, if the point B lies inside the locus of D the amplifier is unstable.

CHAPTER 11

TRANSIENTS IN AMPLIFIERS

11.1. Steady State and Transients

In considering the response of an electric circuit to an applied voltage it is necessary to distinguish two different states of the circuit. There is the steady state, i.e., the condition existing some considerable time after switching, and there is the transient state which occurs immediately after switching. In the electronic circuits which have been dealt with in this book so far, attention has been restricted to steady states. In this chapter we consider some of the features of the transient conditions in valve circuits. Before doing so, a brief review is made of transients in R, L and C circuits.

11.2. Transients in Passive Circuits

Some of the fundamental properties of R, L and C circuits may readily be established from energy considerations. When a current i flows in a coil with inductance L, then energy of amount $\frac{1}{2}Li^2$ is stored in the magnetic field surrounding the coil. Similarly, when a condenser of capacitance C has a potential difference v across it, energy of amount $\frac{1}{2}Cv^2$ is stored in the electric field of the condenser. In a resistor R there is no storage of energy. When a current i flows in the resistor, energy is dissipated as heat of amount Ri^2t, where t is the length of time the current flows. By the nature of energy it is impossible to take it at an infinite rate from a finite supply. Thus, when a battery is connected through a switch to an inductance the current cannot rise instantaneously. It must be zero immediately after closing the switch, and a definite time must elapse before it reaches its final or steady state with energy $\frac{1}{2}Li^2$. Similarly, it is not possible to change the voltage across a condenser instantaneously. There is no such limitation with a resistor, and the current rises immediately to its steady-state value E/R, where E is the e.m.f. of the battery. This involves no energy consumption until some time has elapsed. Inductance or capacitance is therefore responsible for the transient conditions in electric circuits. The impossibility of sudden change of current in an inductance implies that immediately after the sudden change in applied voltage the inductance acts as an open-circuit to the change. Similarly, the impossibility of sudden change of voltage across a condenser implies that immediately after the initiation of the transient a condenser acts as a short-circuit to the change. These considerations are useful in determining the initial conditions in transient changes.

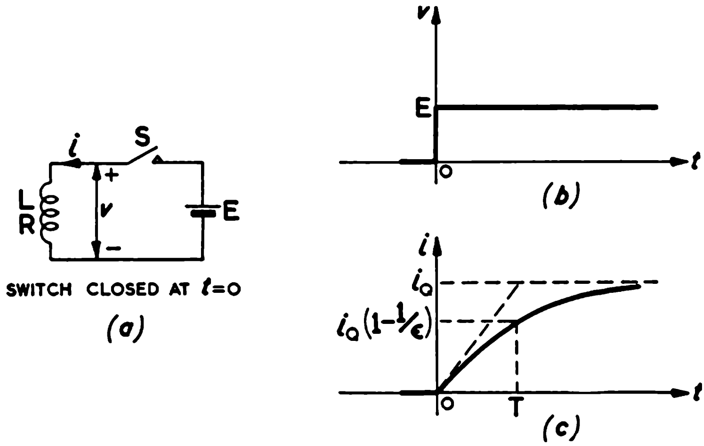


FIG. 11.1

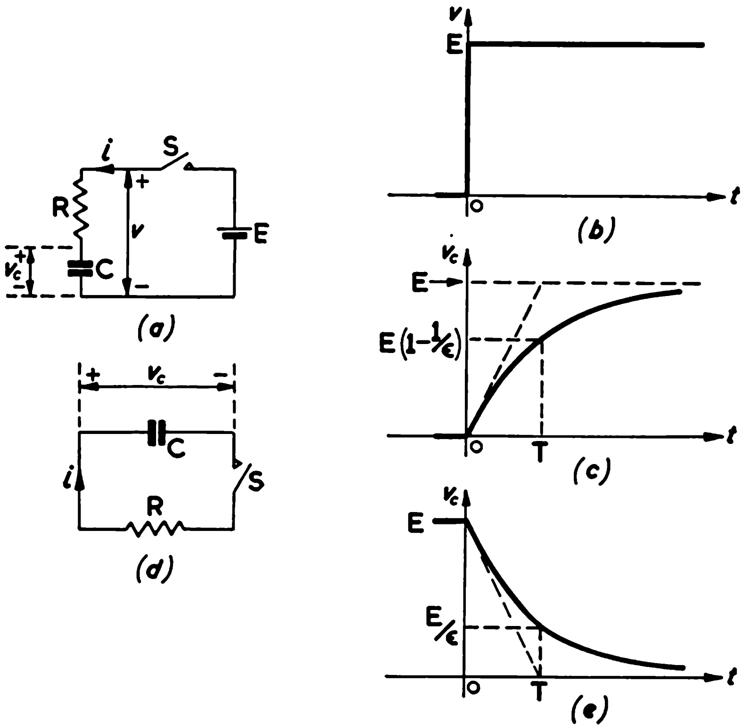


FIG. 11.2

The mathematical laws of transient currents and voltages are derived in numerous text-books, and it is found that they follow exponential laws. For example, in the case of a coil with inductance L and resistance R connected suddenly to a steady supply of e.m.f. E (see Fig. 11.1.*a*), we have the relation $L\frac{di}{dt} + Ri = E$. Solution of this equation shows that the current i after time t follows the law

$$i = i_Q(1 - \varepsilon^{-t/T}),$$

where i_Q is the steady-state current equal to E/R, and $T = L/R$. T is called the time constant; it is the time for the current to reach $(1 - 1/\varepsilon)$ of its final value (i.e., approximately $\frac{2}{3}$ of i_Q). The variation with time of v, the voltage across the inductance, and i, the current through it, are shown in Fig. 11.1.*b* and *c*.

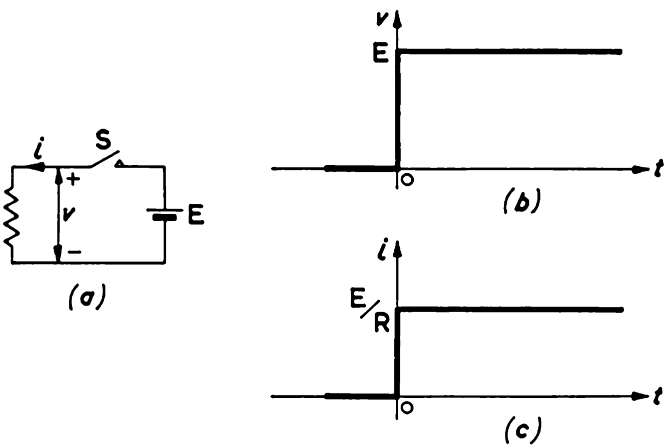


FIG. 11.3

When a battery of e.m.f. E is connected suddenly to a condenser C through a resistance R, as shown in Fig. 11.2.*a* and *b*, the voltage v_c across the condenser follows a similar law to the current in the inductance, and

$$v_c = v_Q(1 - \varepsilon^{-t/T}) \text{ and } i = C\frac{dv_c}{dt} \text{ (see Fig. 11.2.}c\text{)},$$

where $v_Q = E$, the final value of v_c. The time constant T equals RC in this case. If the same condenser after being charged were then discharged through R (Fig. 11.2.*d*), the condenser voltage would decay exponentially to zero according to the law

$$v_c = E\varepsilon^{-t/T}$$

which is shown in Fig 11.2.*e*. The corresponding case of sudden connection of a supply to a pure resistance is shown in Fig. 11.3.*a*, *b* and *c*. As mentioned already, there is no transient delay in this case.

In circuits with inductance and capacitance there are two separate energy stores, and it is possible, under certain conditions, for the energy to be transferred backwards and forwards from one store to the other. An example of a charged condenser suddenly connected to a coil is shown in Fig. 11.4.*a*. At the instant of switching, all the energy is stored in the condenser with no current in the inductance. The condenser then discharges through the inductance, and a current grows until the condenser

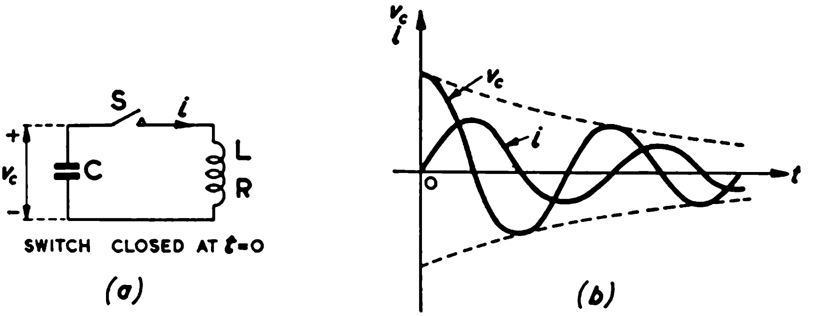


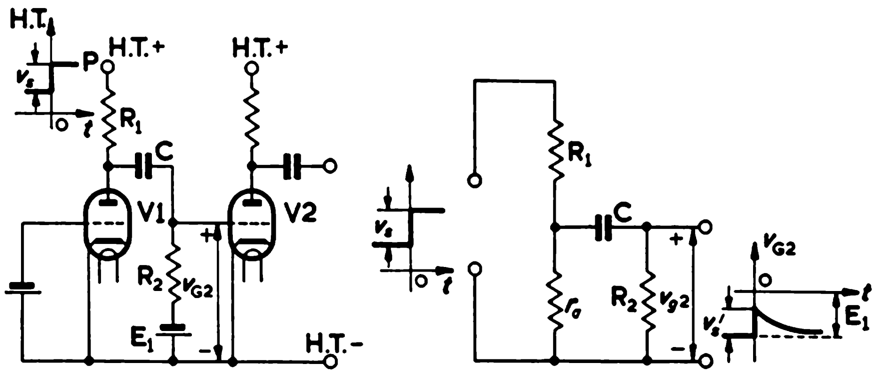
FIG. 11.4

is completely discharged. At this instant the current has its maximum value, and all the stored energy is in the inductance. The current then decreases and the condenser charges again but with reversed polarity. The charging continues until the current is zero and so on. Charge oscillates backwards and forwards round the circuit. While the current is flowing energy is dissipated in the circuit resistance, so that the amount of stored energy decreases and ultimately the oscillation dies away. The variation of current and voltage are shown in Fig. 11.4.*b*. The oscillatory current varies sinusoidally with steadily decreasing amplitude. The actual law is $i = A\varepsilon^{-\alpha t} \sin \omega t$, where $\alpha = R/2L$ and $\omega \simeq \dfrac{1}{\sqrt{LC}}$. If the circuit resistance is large the energy may be dissipated so rapidly that no oscillation takes place. The condition for oscillation is $R < 2\sqrt{\dfrac{L}{C}}$.

11.3. Transients in Valve Circuits

The anode–cathode path of a valve may be represented by a resistance r_a, at least for small changes of the currents and voltages. Also, there are capacitances between the electrodes. Thus the valve makes contributions to the resistance and capacitance of the circuits associated with it, and so must affect the transient behaviour for sudden changes either of the supply voltages or of a signal. The behaviour of the valve circuit depends to some extent on the place where the change occurs.

Firstly, we consider the circuit in Fig. 11.5.*a*, which shows a valve amplifier with *RC* coupling to another valve. Inter-electrode capacitances are ignored at present. It is assumed that a sudden change occurs in the supply voltage at the point *P* as shown by the step function in the diagram. This affects the anode voltage of the first valve *V*1, but does



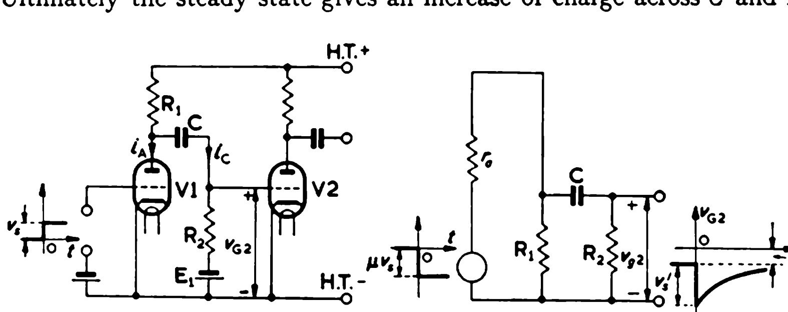
(a) (b)

$T = RC$ where $R = R_2 + \dfrac{r_a R_1}{r_a + R_1}$

$$v_{g2} = \frac{R_2 r_a v_s}{R_1 r_a + R_2 r_a + R_1 R_2} \epsilon^{-t/T} = v_s' \epsilon^{-t/T}$$

FIG. 11.5

not affect its grid voltage. If the changes are small, then any anode current or voltage changes are related by $i_a = v_a/r_a$. The valve is therefore replaced by r_a, as shown in Fig. 11.5.*b*. The effect of the sudden voltage change on the amplifier depends on the voltage v_{g2} appearing at the grid of *V*2. On account of the coupling condenser *C*, a transient occurs. Ultimately the steady state gives an increase of charge across *C* and no

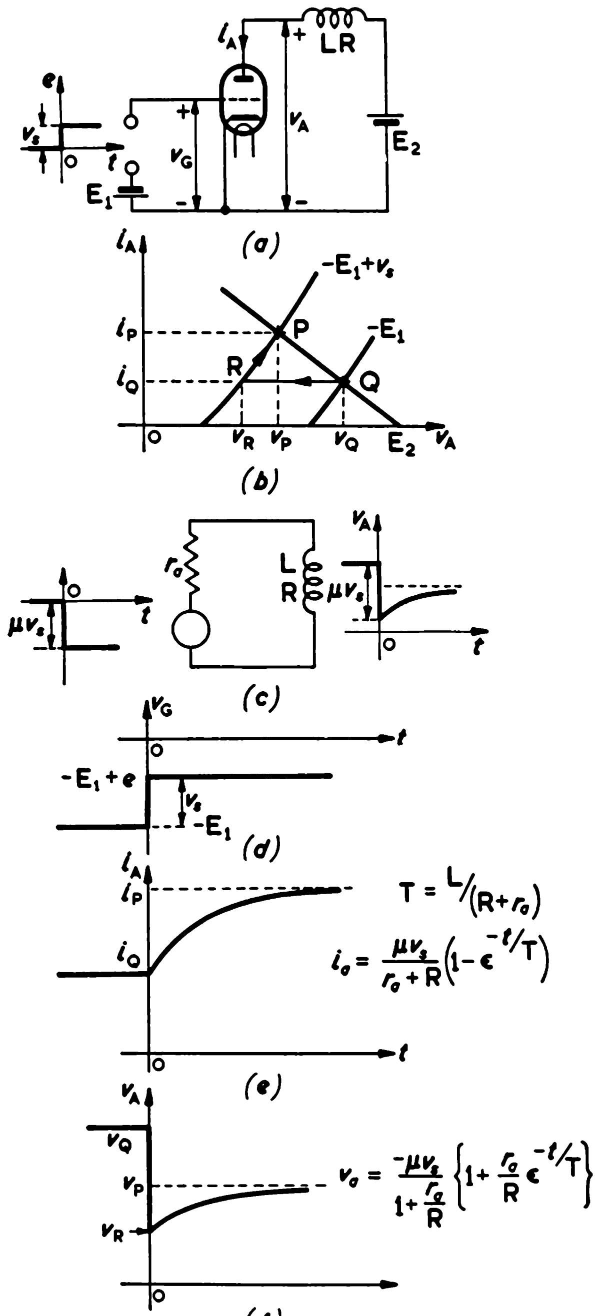


(a) (b)

$T = RC$ where $R = R_2 + \dfrac{r_a R_1}{r_a + R_1}$

$$v_{g2} = -\left(\frac{\mu R_2 R_1 v_s}{R_1 r_a + R_2 r_a + R_1 R_2}\right) \epsilon^{-t/T} = -v_s' \epsilon^{-t/T}$$

FIG. 11.6



e
v_s
t
O
E_1
v_G
i_A
LR
v_A
E_2
(a)
i_A
$-E_1+v_s$
i_P
P
$-E_1$
R
i_Q
Q
O
v_R
v_P
v_Q
E_2
v_A
(b)
μv_s
t
r_a
L
R
μv_s
v_A
(c)
v_G
t
$-E_1+e$
v_s
$-E_1$
(d)
i_A
i_P
$T = L/(R+r_a)$
i_Q
$i_a = \frac{\mu v_s}{r_a + R}\left(1 - \epsilon^{-t/T}\right)$
t
(e)
v_A
v_Q
v_P
$v_a = \frac{-\mu v_s}{1 + \frac{r_a}{R}}\left\{1 + \frac{r_a}{R}\epsilon^{-t/T}\right\}$
v_R
(f)

FIG. 11.7

permanent effect on v_{g2}. While C is charging there is a current through R_2, and v_{g2} varies as shown by the exponential waveform in the diagram.

As a second example of transients in valve circuits we consider the case of a sudden small change in grid voltage in the first valve of a two-stage RC-coupled amplifier as shown in Fig. 11.6.*a*. After its initial sudden change the grid voltage remains constant, and the transient conditions are determined by the rest of the circuit. As the changes are small, the equivalent circuit may be used as shown in Fig. 11.6.*b*, from which the transient change in v_{G2} may be determined. The actual shape and magnitude are indicated in the figure. The value of r_a is, of course, determined from the i_A, v_A characteristic at the operating point.

Finally, we consider an amplifier with an inductive load (Fig. 11.7.*a*), and we assume a sudden grid voltage rise as shown. Since the current through an inductance cannot be changed instantaneously, it follows that the anode voltage must drop suddenly in order to maintain constant anode current. Thereafter the anode current changes gradually to a new steady value. The variations may be determined as in Fig. 11.7.*b* from the valve characteristics and the load line corresponding to the resistance R. The initial quiescent point is Q. The sudden application of v_s results in an instantaneous anode voltage change from v_Q to v_R at constant i_A. The anode current then rises from i_Q to i_P, whilst the anode voltage rises from v_R to v_P along the $-E_1 + v_s$ grid characteristic. If the changes are small the transient may be analysed by using the equivalent circuit of Fig. 11.7.*c*. The nature and magnitude of the transient changes are shown in Fig. 11.7.*d*, *e* and *f*.

11.4. Amplification of Square Pulses

An RC-coupled amplifier is incapable of amplifying faithfully a single step signal. Frequently it happens that an amplifier is required for a flat-topped pulse of the type shown in Fig. 11.8.*a*. The circuit of Fig. 11.6.*a* could give at the grid of the second valve a signal of the type shown in Fig. 11.8.*b*, which has little resemblance to the original signal. The nature of the waveform produced by the amplifier varies with the time constant of the valve and circuit. For the case in Fig. 11.8.*b* the time constant is much less than the width of the pulse. For a time constant of the order of the pulse width the amplified signal would be as shown in Fig. 11.8.*c*. With a time constant much longer than the pulse width the amplified waveform would be as shown in Fig. 11.8.*d*, and would be almost the same shape as the original pulse. The conditions required for longer time constant are the same as those for extending the low-frequency range of the amplifier. Thus good low-frequency response is necessary for the faithful reproduction of the flat top of a pulse.

In dealing with transients in valve circuits so far, we have neglected inter-electrode capacitances or stray capacitances. In the case of an amplifier with a resistance load the stray capacitance from the anode, C_S, is effectively across the load resistance as shown in Fig. 11.9.*a*. When

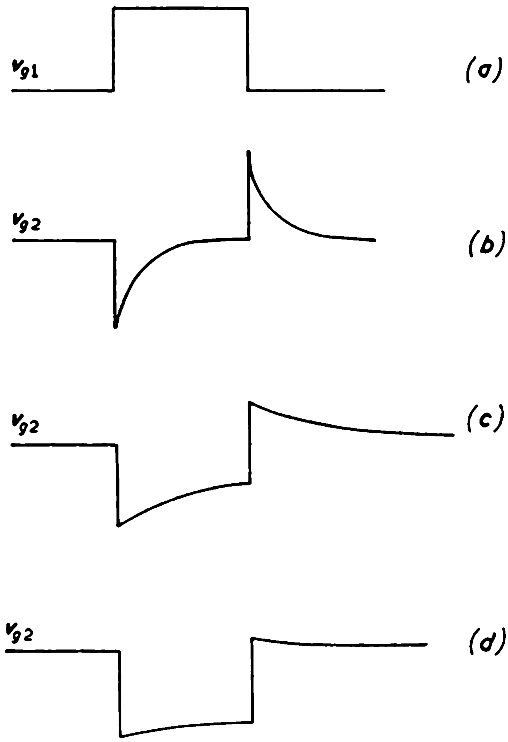


FIG. 11.8

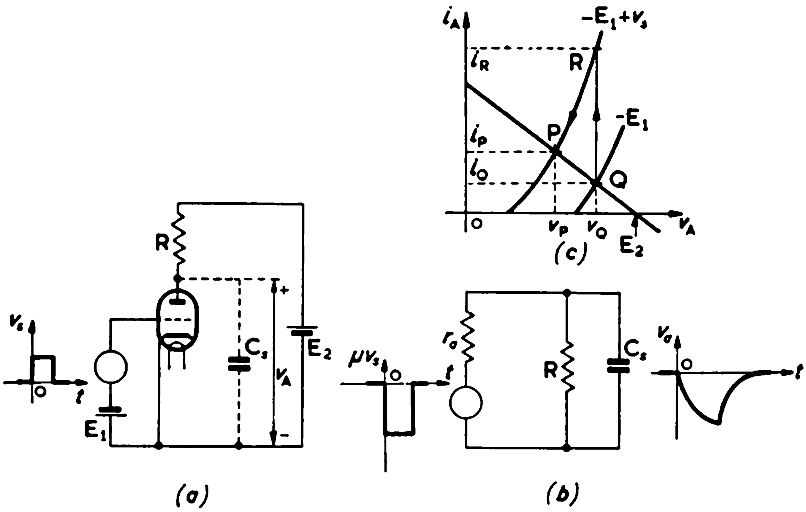


FIG. 11.9

a small square pulse is applied to the grid of such an amplifier the transient behaviour is found from Fig. 11.9.*b* or *c*. The effect of C_S is to prevent the sudden rise and fall at the beginning and end of the anode voltage pulse, as shown in the figure. The anode current varies along the path QRP in Fig. 11.9.*c*. For an amplifier to give rapid rise and fall to the sides of a pulse it must obviously have good high-frequency response.

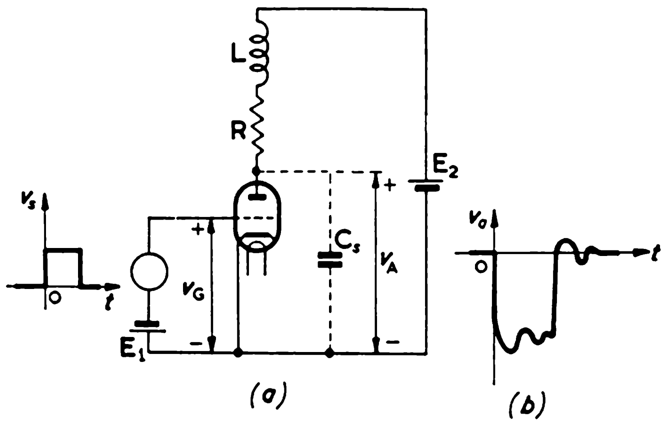


FIG. 11.10

In some amplifiers an inductance is used in series with the load resistance to extend the high-frequency range of the amplifier (Fig. 11.10.*a*). The load circuit now has L, C and R, and hence oscillation is possible. When a square pulse is applied to the input of such an amplifier the output voltage may have oscillations as shown in Fig. 11.10.*b*.

11.5. Large Transients in Valve Circuits

In most of this chapter we have assumed that the transients have been sufficiently small to justify the use of the valve equivalent circuit. Frequently large transients occur, and then mathematical analysis is extremely difficult, as the valve behaviour may be highly non-linear over the range of the transient. The performance may usually be determined qualitatively on the basis of the principles discussed in this chapter by using anode characteristics and a load line.

11.6. A.c. Transients

We have considered transients arising from a step change in a steady voltage, and we have seen that the resultant behaviour can be determined in terms of the initial conditions, the final steady state and the transient state. Similar effects are obtained on introducing any sudden change. The nature of the transient part is the same in all cases, and any differences arise from the different initial conditions and final steady state. In the case of a.c. supplies the initial conditions vary with the time of switching; the magnitude of the resulting transient is affected. However, the duration of the transient is the same as in switching d.c. supplies.

11.7. Class C Amplifier as a Switch

A sudden change in a circuit with both L and C may give an oscillatory current which gradually dies away as the energy is dissipated in the circuit resistance. In Fig. 11.11.*a* an oscillatory circuit is shown connected to a power supply E_2 through a switch S. Let us assume that the condenser

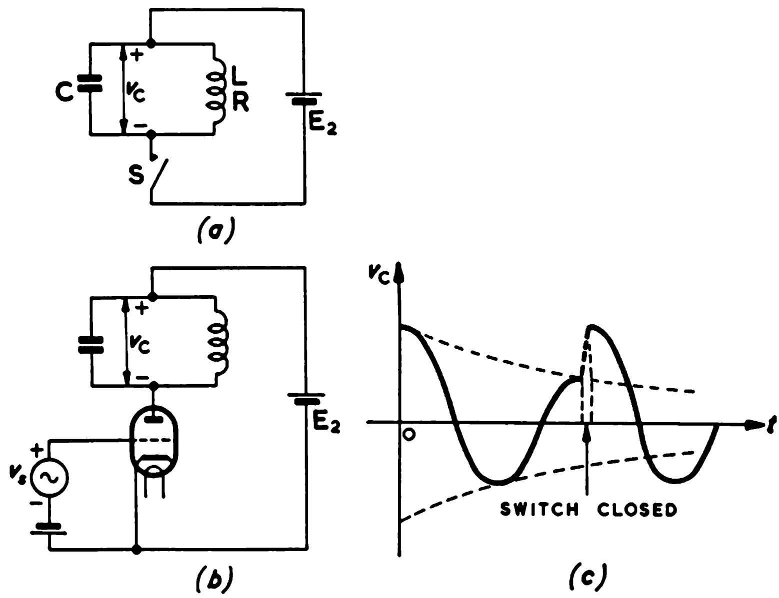


FIG. 11.11

is charged initially to a potential difference equal to E_2 and that the switch is open. The condenser discharges through the inductor, and after one cycle the voltage across C is slightly less than E_2, on account of the energy dissipated in R during the cycle. If now the switch S is closed for a short time the condenser voltage is restored to E_2, and the switch may be opened again (see Fig. 11.11.*c*). A second oscillatory cycle occurs and, provided S is always closed at the right time each cycle, the battery restores the energy dissipated; the oscillation therefore continues without loss of amplitude. Provided the energy lost each cycle is only a small fraction of the energy stored, i.e., provided the circuit has a high Q, the oscillation is practically sinusoidal. The rate at which a condenser charges depends on the resistance of the charging circuit. If the resistance of the condenser–battery–switch circuit is negligible, then the switch need be closed only momentarily.

The valve in a Class C amplifier behaves very much like the switch in Fig. 11.11.*a*. Once every cycle the valve passes a pulse of current to restore the initial charge amplitude to the condenser C (Fig. 11.11.*b*). Since the valve has some resistance, it must conduct for a definite part of the cycle in order to recharge the condenser, and also there is some voltage drop across the valve, so that the condenser voltage amplitude is less than E_2. In the switch circuit we postulated that S should be closed at the right time every cycle. With the valve this is achieved by connecting to the grid a signal of the same frequency as the oscillation. It may be noted in passing that the grid signal does not supply any power to the oscillatory circuit. It merely serves as a suitable timing device to ensure that the losses in R are made good from the anode supply. In practice, R arises mainly from a coupled load resistance. The resistive components of the circuits are made as small as possible. The operation of the Class C amplifier as a switch supplying power to the oscillatory circuit once every cycle is closely analogous to the operation of a clock pendulum, which is given an impulse once every cycle by means of the escapement. The energy is supplied from the potential energy stored in weights or in a coiled spring. Another analogy is found in the child's swing, which is kept going by giving it a push at the right time each cycle.

It may be seen that it is not essential to supply energy to these oscillators each cycle. Every other cycle, or one cycle in three, is enough, provided the Q is sufficiently high. When a valve is supplied with a grid signal of frequency f and the anode circuit oscillates at $2f$, $3f$ or higher, then we have a frequency-multiplying Class C amplifier.

11.8. Transients in Circuits with Feedback

In Section 11.3 we considered the response of valve amplifiers to sudden changes of grid voltage. It was assumed that after the initial change the grid voltage remained constant and that the transient behaviour depended on the rest of the valve and the circuit. In an amplifier with feedback this assumption is unjustified, as the grid voltage continues to change with the transient changes in the output circuit. The problem is then much more complicated. Some consideration is given to the case of an amplifier with positive feedback in Chapter 13.

11.9. Some General Comments on Transients

We consider above transient effects in circuits with R, L or C in various combinations. In particular, in circuits which are purely resistive, changes in current and voltage occur instantaneously. However, when a resistor carries current there must be a magnetic field around it, so that the resistor has some inductance as well as resistance. Also, when a potential difference exists across a resistor there is an electric field around it, so that it must also have capacitance. The inductance and capacitance, which are distributed along the resistor, are small. Similarly, inductors have distributed resistance and capacitance and condensers have inductance and resistance. The time constants arising from these distributed circuit components are small, usually much less than 10^{-6} sec. For longer time constants than this the circuits can be represented by simple combinations of R, L and C. For shorter time constants it may be necessary to consider each circuit component in terms of its distributed properties, rather like a transmission line.

In circuits with very short time constants the transit time of the electrons across the inter-electrode spaces may not be negligible. In some transistor circuits diffusion or reactive effects may be important at quite low frequencies.

CHAPTER 12

DIRECT-COUPLED AMPLIFIERS

12.1. The Amplification of d.c. Changes

A common requirement in electronics is the amplification of a steady current or voltage, or of the mean component of a changing current or voltage. When several stages of amplification are required, then the capacitance or mutual inductance couplings previously described cannot be used, since neither of these couplings is capable of transmitting a d.c. change. A direct connection must be made between the anode of one valve and the grid of the next valve. If a d.c. signal is applied to the input of a direct-coupled amplifier, then the output is also a d.c. change. However, changes in the grid or anode supplies or in the heater supplies also give d.c. changes in output, and hence are indistinguishable from signals. This effect, sometimes referred to as zero drift, constitutes a serious limitation in direct-coupled amplifiers, and special circuits are used to minimize the effects of supply changes. Similar changes in supply voltages are of much less account in a.c. amplifiers with capacitance or mutual inductance coupling, since they do not affect the output, except in so far as the valve constants, r_a and g_m, vary with the supply changes.

Direct-coupled amplifiers are also capable of amplifying a.c. signals. However, when an amplifier is required for a.c. signals only it is usual to avoid direct coupling, except perhaps when the frequency of the signals is very low.

12.2. Direct Coupling

A simple two-stage direct-coupled amplifier is shown in Fig. 12.1. The direct connection from the anode of the first valve to the grid of the second valve is made through a battery E_{12}. The purpose of this

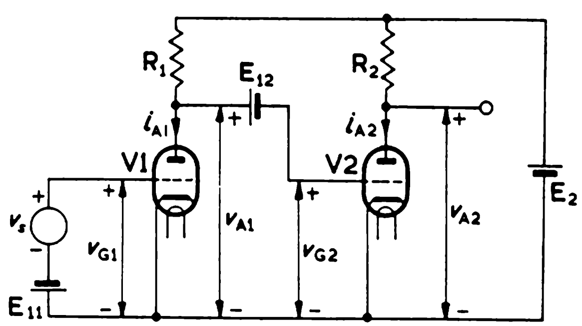


FIG. 12.1

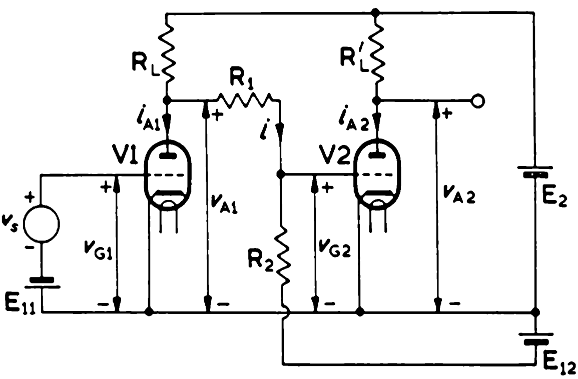
battery is to ensure that the second valve has the correct operating grid bias. From the circuit it may be seen that $v_{G2} = v_{A1} - E_{12}$ and $v_{A1} = E_2 - R_1 i_{A1}$. Since v_{G2} has a small negative value, E_{12} is approximately equal to v_{A1}. Normally, v_{A1} is an appreciable fraction of E_2, so that the battery E_{12} is fairly large. When a small signal v_s is applied to the input of the amplifier the amplified output of the first valve is passed directly to the grid of the second valve, and hence

$$v_{g2} = v_{a1} = -g_{m1}v_s \bigg/ \left(\frac{1}{R_1} + \frac{1}{r_{a1}}\right).$$

The signal output of the amplifier is v_{a2}, where

$$v_{a2} = g_{m1}g_{m2}v_s \bigg/ \left(\frac{1}{R_1} + \frac{1}{r_{a1}}\right)\left(\frac{1}{R_2} + \frac{1}{r_{a2}}\right).$$

If the valves are identical pentodes with equal load resistances R, then the voltage amplification is $A = (g_m R)^2$. With an a.c. signal the output is in



$$v_{G2} = v_{A1} - R_1 i$$
$$(R_1 + R_2)i = E_{12} + v_{A1}$$

FIG. 12.2

phase with the signal. If a small change e_{11} should occur in the grid battery supply E_{11}, this would be amplified and give an output change of $-(g_m R)^2 e_{11}$ (assuming pentodes). A variation in E_{12} of amount e_{12} would give a change in output of $g_m R e_{12}$. A change of amount e_2 in the main h.t. supply E_2 would give an output change with pentodes of $(1 - Rg_m)e_2$. Normally $Rg_m \gg 1$, and then the change in output is $-g_m R e_2$. On account of these variations in output it is essential to use highly stable supplies with this circuit. Also, the need for the fairly large battery E_{12} has certain disadvantages. Not only is it an extra cost, but its capacitance is in parallel with the load resistance R_1, and so limits the high-frequency and transient response of the amplifier.

An alternative method of obtaining direct coupling is shown in Fig. 12.2. Here the anode of the first valve is joined to the grid of the second valve through a resistor R_1. The appropriate quiescent grid voltage for $V2$ is obtained from an additional supply E_{12} through the resistor R_2. Here

$$v_{G2} = v_{A1} - R_1 i,$$

where

$$i(R_1 + R_2) = E_{12} + v_{A1}$$

(it is assumed that v_{G2} is negative so that $V2$ has no grid current). It follows that

$$v_{G2} = \frac{R_2 v_{A1}}{R_1 + R_2} - \frac{R_1 E_{12}}{R_1 + R_2}.$$

Since v_{G2} is a small negative voltage, then

$$E_{12} \simeq R_2 v_{A1}/R_1,$$

and this is one design equation for the circuit. When a signal is applied

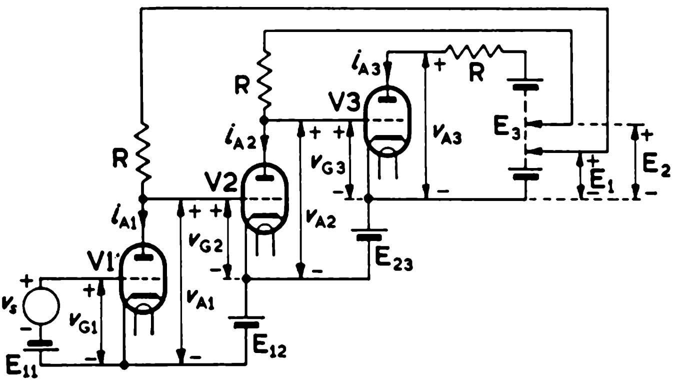


FIG. 12.3

to the grid of $V1$ only part of the resulting change in anode voltage, v_{a1}, is applied to the grid of $V2$, since R_1 and R_2 act as a potential divider. Thus

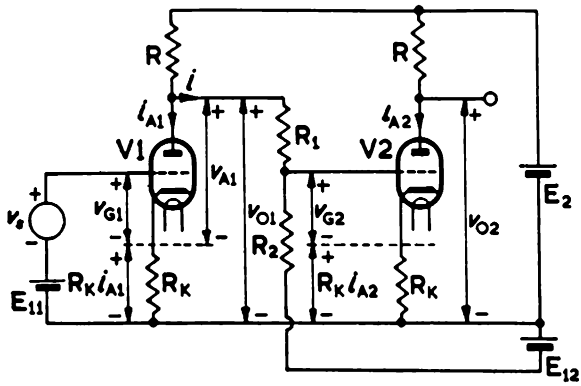
$$v_{g2} = R_2 v_{a1}/(R_1 + R_2).$$

(It is assumed that the internal resistance of E_{12} is negligible.) It is therefore desirable to make R_1/R_2 small to give high stage gain. From the first design equation this means large E_{12}, and obviously there must be some compromise. This circuit is sometimes called the " third rail circuit ". In most other amplifiers there are two h.t. " rails ", the h.t. positive and h.t. negative lines. In Fig. 12.2 there is an additional h.t. rail to the grid of $V2$.

A third type of direct coupling is shown in Fig. 12.3, where each anode is connected straight to the next grid. This circuit involves the cathodes of the valves being at different potentials, which may give rise to cathode-heater insulation problems unless separate heater supplies are used for each valve.

12.3. Use of Negative Feedback

In Chapter 10 it is shown that the use of negative feedback can render amplifiers insensitive to changes in supply voltage, and it is therefore worth considering the introduction of feedback to direct-coupled amplifiers. A two-stage amplifier with current feedback applied to each stage is shown in Fig. 12.4. The cathode resistors R_K provide the feedback and at the same time give automatic bias. This bias may be too



$$v_{G2} + R_K i_{A2} = v_{O1} - R_1 i$$
$$(R_1 + R_2) i = E_{12} + v_{O1}$$

FIG. 12.4

great, particularly in the first valve, and then an additional battery E_{11} is required. In each of the stages the feedback fraction β is equal to R_K/R. Provided each stage has sufficiently high inherent gain, the gain of each with feedback is $-1/\beta$. The overall gain of the amplifier is then $(R/R_K)^2$; this assumes that R_1/R_2 is small. The gain of the amplifier is thus independent of changes in supply voltage. However, this is far from the whole story, since changes in supply voltages still give changes in output. For example, a change in E_{11} is indistinguishable from a signal v_s. Also, it may be shown that a change e_{12} in E_{12} gives an output change of

$$\frac{R}{R_K} \cdot \frac{R_1}{R_1 + R_2} e_{12}$$

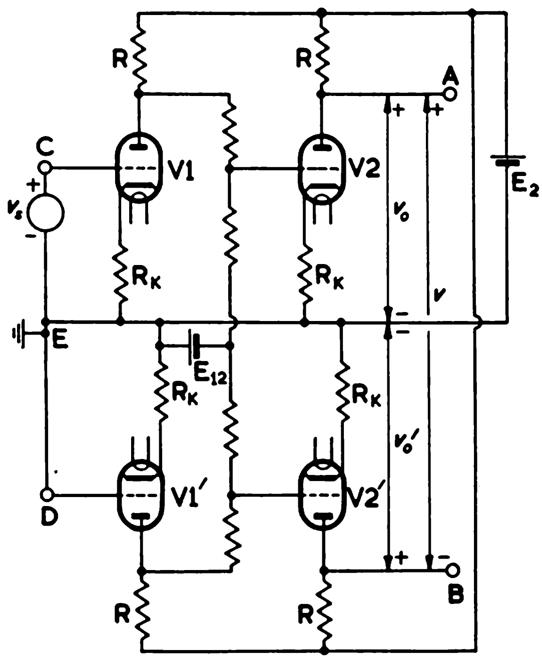
and a change in E_2 of amount e_2 gives an output voltage change of

$$-\frac{R}{R_K} \cdot \frac{R_2}{R_1 + R_2} e_2.$$

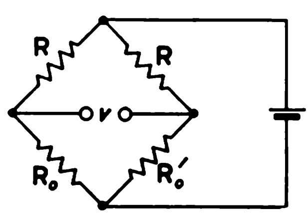
Thus, although feedback renders the amplifier gain independent of supply changes, it is still necessary to have highly stable supplies to prevent zero drift.

12.4. Balanced or Differential Amplifiers

The effects of power supply changes on zero drift can be considerably reduced by the use of certain balanced or differential circuits. These usually employ twice the number of valves that would be used in a normal amplifier, and the connections are arranged so that the changes produced in the output circuit by the extra valves cancel those in the



(a)



(b)

Fig. 12.5

normal amplifier. An example of such a circuit is shown in Fig. 12.5.*a*. $V1$ and $V2$ form a third rail, two-stage amplifier with feedback similar to the one considered in the last section. The output of this amplifier v_0 depends on the signal applied to the input of $V1$ and also on any changes in the supply voltages. $V1'$ and $V2'$ form a second and similar amplifier, whose output v_o' depends only on changes in the supply voltages. There is no signal applied to this amplifier. The output v of the whole system is taken between the points A and B, and so $v = v_o - v_o'$. The valves and components of the two amplifiers are identical, so that with no signal v_o and v_o' are equal and v is zero. Now any changes in the power supplies cause equal changes in v_o and v_o' so that v is still zero. A signal applied to $V1$ gives a change in v_o, but does not affect v_o', and hence v changes by the same amount as v_o. With the same assumptions as in the previous section, the overall gain of the amplifier is therefore $(R/R_K)^2$. It may be noted in passing that the valves $V2$ and $V2'$,

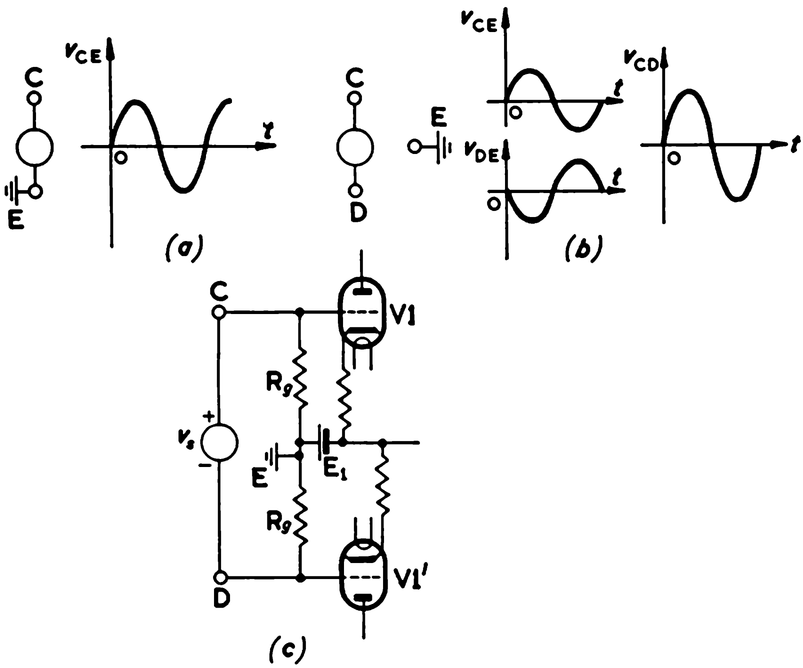


FIG. 12.6

together with their anode loads, form a Wheatstone bridge network, as shown in Fig. 12.5.*b*, where R_o and R_o' represent the equivalent d.c. resistances of $V2$ and $V2'$. Changes in supply voltages cause equal changes in R_o and R_o' so that the bridge remains balanced. A d.c. signal, however, alters the value of R_o only, and the balance is upset. The amount of unbalance is then a measure of the signal.

If the signal is omitted from Fig. 12.5.*a* the circuit is completely symmetrical, and by comparison with the circuits of Section 9.9 it may be seen that this is a push-pull circuit with resistive components. In push-pull circuits changes in the supply voltages do not appear in the output. In the use of the amplifier described above the signal is connected asymmetrically between one grid C and the earth point E. Such a signal is said to be unbalanced. One terminal of the signal is assumed to remain constant at earth potential, and the other terminal varies positively and negatively with respect to earth, as shown in Fig. 12.6.*a*. Sometimes the signal is balanced with respect to earth, i.e., the two terminals both vary, but by equal and opposite amounts measured from earth potential, as shown in Fig. 12.6.*b*. Such a signal would be connected to the amplifier symmetrically, as shown in Fig. 12.6.*c*, and this is the normal input connection for a push–pull amplifier.

12.5. High-gain Amplifier with High Stability

We see above that stability against power supply changes can be obtained by the use of a second amplifier giving a differential action. At the same time stage-gain stability is achieved by the use of current

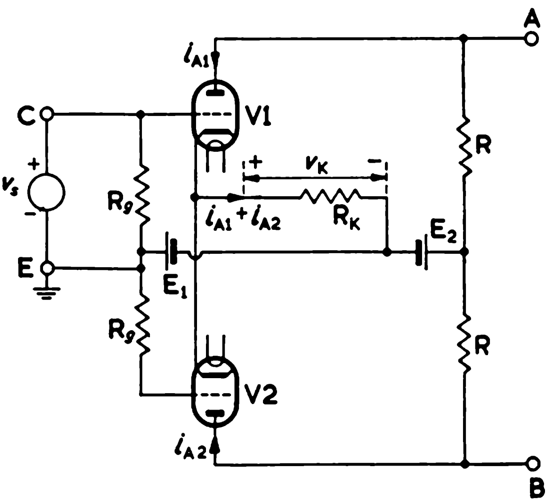


FIG. 12.7

feedback separately on each valve. Even greater stability against supply changes can be realized by using a common-bias resistor for $V1$ and $V1'$. As shown below, this gives feedback against supply changes but no feedback on the signal. Thus high signal gain with high stability to supply changes is obtained. The circuit is shown in Fig. 12.7. The anode currents of both valves flow through the common cathode resistor R_K, giving a bias voltage v_K. If need be, this may be offset by a supply E_1 in order to bring the valves to a suitable operating point. If there is an increase in supply voltage E_2 this causes an increase in the anode currents of both valves, and so an increase in v_K. This increase in bias reduces the anode current, and so opposes the original change. Thus, in addition to the differential action obtained by identical valves and components, there is also a feedback against supply changes, and this amplifier is extremely stable against such changes.

When a signal is applied which makes C positive with respect to E the anode current in $V1$ rises, and so increases v_K. This reduces to some extent the effect of the signal on the anode current to $V1$, and at the same time reduces the anode current to $V2$. The valve currents thus vary in opposite senses. The variations are very nearly equal and opposite, as is now shown. For pentodes and small signals

$$i_{a1} = g_m v_{g1} \text{ and } i_{a2} = g_m v_{g2},$$

where $\quad v_{g1} = v_s - R_K(i_{a1} + i_{a2}) \quad$ and $\quad v_{g2} = -R_K(i_{a1} + i_{a2}).$

Hence $\quad v_{g2} = -R_K g_m(v_{g1} + v_{g2}), \quad$ i.e., $\quad v_{g2} = \dfrac{-R_K g_m}{1 + R_K g_m} v_{g1}.$

If $R_K g_m \gg 1$ then $v_{g2} = -v_{g1} = -v_s/2$. The input signal is thus shared equally by the two valves, but in opposite senses. The output voltage

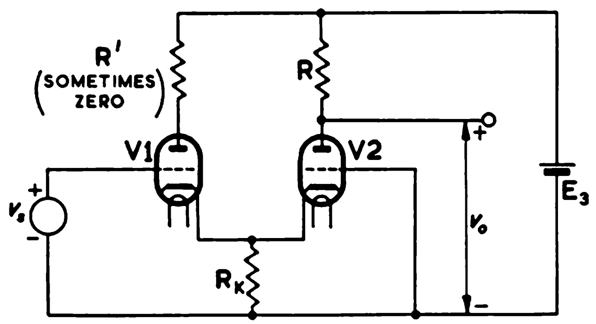


FIG. 12.8.

of each valve is $g_m R v_s/2$ and the total output across the terminals A and B is $g_m R v_s$. There is therefore no feedback on the signal.

The circuit of Fig. 12.7 has several forms and it is known by a variety of names, such as a cathode-coupled amplifier and a long-tailed pair. One variation of the circuit is shown in Fig. 12.8, in which the output is taken from $V2$. This output may be connected to the grid of another valve in a second cathode-coupled amplifier. It may be noted that the output voltage of a cathode-coupled amplifier is in phase with the signal voltage.

12.6. Other Methods of Amplifying Steady Signals

Throughout this chapter we have come across the problem of zero drift in direct-coupled amplifiers—a difficulty which does not occur to anything like the same extent in a.c. amplifiers. Other methods are sometimes used for the amplification of steady signals. For example, the steady signal may be connected to the terminals of an interrupter, i.e., a mechanical vibrator which interrupts the steady signal, thereby producing an alternating signal which may be amplified by a conventional a.c. amplifier.

The magnetic amplifier is another means of amplifying d.c. by using the d.c. to produce an a.c. The d.c. is passed through one winding of a saturable iron-cored inductor. The magnetization of the core varies with the magnitude of the d.c. A second winding is fed with a.c., and the inductance of this winding varies with the magnetization. Thus the d.c. change is converted into an a.c. change, which may be amplified further and finally rectified to restore a steady signal.

CHAPTER 13

OSCILLATORS

13.1. Introduction

Electronic circuits are used as generators of oscillations at frequencies from 1 to 10^{11} c/s. To cover this enormous range, valves or transistors and circuits of great variety are used. In considering the operating principles it is convenient to divide oscillators into two main classes, negative resistance and feedback. In the former an electronic device acts, between two terminals, as a negative resistance which cancels the normal ohmic resistance of a tuned circuit connected across the terminals. A feedback oscillator can be realized from any device capable of power amplification. Sufficient output power is fed back to the input so that the amplifier supplies its own input signal. This chapter deals with both classes of oscillator, but only with those producing sinusoidal oscillations and using negative grid valves and transistors. Special valves for very high frequencies are considered in Chapter 15, and non-sinusoidal oscillators in Chapter 18. In all oscillators the ultimate source of energy is a d.c. power supply. The valves and circuits serve to convert some of the d.c. energy into a.c. energy.

13.2. Negative Resistance Oscillators.

When a condenser is discharged through an inductance the current is oscillatory provided that the resistive damping of the circuit is not too great (see Section 11.2). The amplitude of the current decreases exponentially, and ultimately becomes negligible when all the original energy stored in the condenser has been dissipated in the resistance. In the circuit in Fig. 13.1 a capacitance is connected in parallel with an inductance L and a resistance R. When the switch is closed an oscillatory current flows and the rate of decay of the oscillation depends on R; the larger the value of R, the slower the rate of decay. At any instant the rate of dissipation of energy in R is v^2/R, where v is the instantaneous value of the voltage across C. When R is infinite there is no loss of energy, and once the current starts it continues indefinitely with constant amplitude. If R is negative the resistance acts as a source of energy and the amplitude of the oscillation steadily increases.

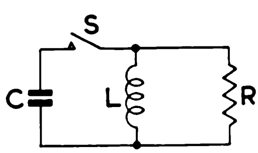


FIG. 13.1

Under certain conditions, a tetrode valve behaves as though there is a negative resistance between the anode and the cathode. An anode characteristic for a tetrode is shown in Fig. 13.2.*a*. Over the region *BC* the anode slope resistance r_a is negative. An arrangement for using this negative resistance to generate oscillations is shown in Fig. 13.2.*b*, where

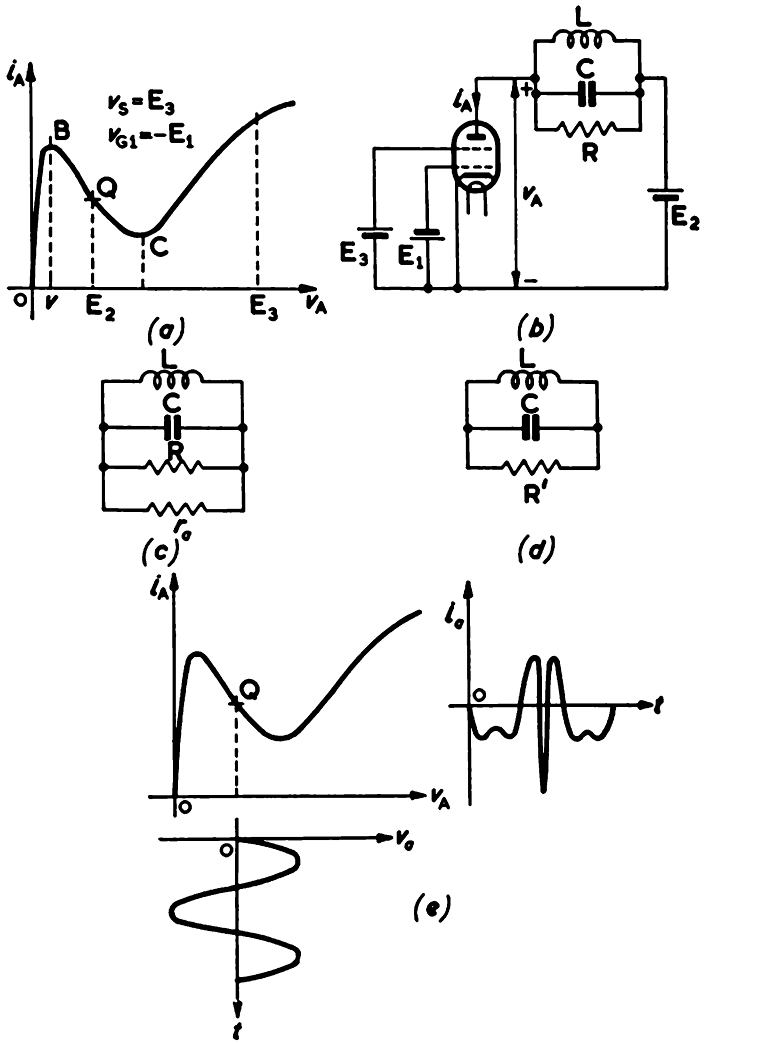


FIG. 13.2

a parallel tuned circuit is used as the anode load. As far as alternating currents are concerned, the valve is in parallel with the tuned circuit (Fig. 13.2.*c*), and the resultant resistance across *L* and *C* is $R' = Rr_a/(R + r_a)$ (Fig. 13.2.*d*). If the quiescent point is chosen near the middle of *BC*, then r_a is negative, and the sign of *R′* depends on the relative sizes of *R* and r_a.

If R and r_a are numerically equal R' is infinite, and an oscillatory current once started does not decay. In practice, it is extremely difficult to arrange that R and r_a are exactly equal. The circuit and voltages are adjusted so that r_a is negative and numerically less than R and then R' is negative. The amplitude of oscillation therefore increases. However, it may soon reach a value where the anode voltage goes beyond the range BC; r_a is then positive and no longer acts as a source of power. In operation the anode voltage variation is just enough to give r_a an average value over one cycle equal to R. The anode voltage and current variations are shown in Fig. 13.2.*e*. Although the anode current is not sinusoidal, it produces voltage across the load only at the fundamental frequency, since the load impedance is small at frequencies away from the resonant value, f_o, where $2\pi f_o = 1/\sqrt{LC}$. The initial adjustment of this oscillator is more than enough to produce oscillation, and the amplitude is limited by the valve characteristics. This is an example of a practice which is common in valve oscillators.

In Section 5.13 it is shown that a gas diode may produce an arc discharge which has a negative resistance. In the early days of radio, arc oscillators were frequently used in transmitters. Diodes and other vacuum tubes may have the property of negative resistance at certain very high frequencies, when the electron transit time is greater than a period of the oscillation. The tetrode negative resistance oscillator is usually called a dynatron oscillator.

13.3. Feedback Oscillators

The vast majority of electronic oscillators are amplifiers which provide their own input. In the general feedback circuit of Fig. 10.6.*b*, which is discussed in Sections 10.5 and 10.12, it is shown that any amplifier oscillates when $1 - \boldsymbol{\beta}\mathbf{A}$ is zero, where $\boldsymbol{\beta}\mathbf{A}$ is the feedback factor. This means that the feedback voltage is exactly equal in magnitude and phase to the grid voltage. Oscillation may also occur if the feedback voltage is greater than v_g, but again the phases must be identical. In general, there are, therefore, two separate conditions to be satisfied in feedback oscillators. The phase condition determines the frequency of oscillation. The other condition is concerned with the feedback magnitude.

The principle of the feedback oscillator may be illustrated with reference to a single-stage resistance-loaded amplifier of the type shown in Fig. 13.3.*a*. In this circuit the output voltage v_a is greater than the signal voltage, and differs from it in phase by 180°, provided the capacitances of the valve and circuit elements are negligible. If this amplifier is to provide its own input there must be a further change of phase of 180° in the feedback circuit. This can be achieved with a phase-shifting network of the form illustrated in Fig. 13.3.*b*. The phase difference between the voltages v_1 and v_2 varies from 90° to zero, as the frequency is raised from zero to a very high value. At the same time the phase difference between v_o and v_1 ranges from 270° to zero. If this network is put in parallel with the anode load R_1, then at some frequency, f_o, the voltage v_o is in phase with v_s. If $R \gg R_1$ there is approximately 180° phase shift through the amplifier and a further 180° in the R, C network. The magnitude of the voltage v_o is considerably less than v_a. Provided that the voltage gain of the amplifier at frequency f_o is greater than v_a/v_o,

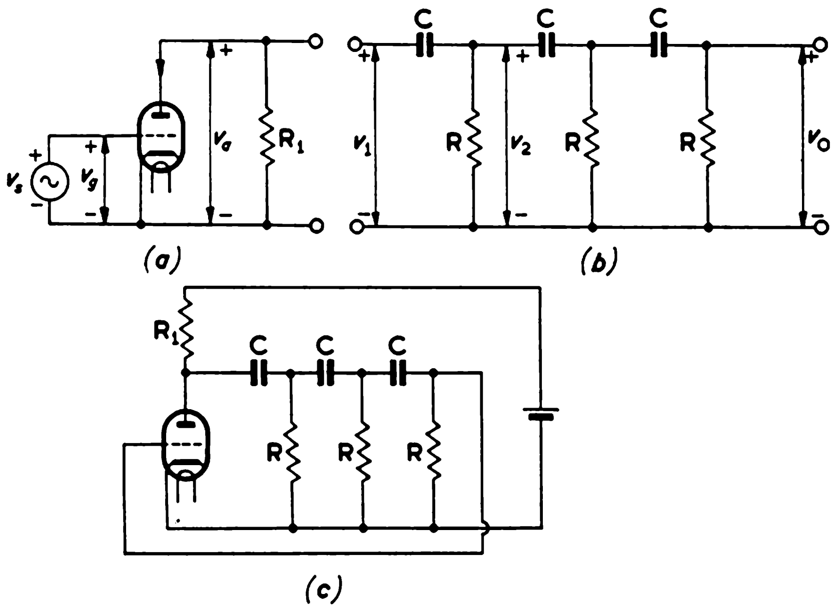


FIG. 13.3

then v_o is greater than and in phase with v_s. Hence, on joining the output of the phase-shifting network to the input of the amplifier, as shown in Fig. 13.3.*c*, the circuit oscillates at frequency f_o.

13.4. Tuned-anode Oscillator

A common form of feedback oscillator is shown in Fig. 13.4, where a tuned anode load is used, and the feedback to the grid is through the mutual inductance M. In order to determine the conditions for oscillation certain assumptions are made. Firstly, small a.c. changes are assumed so that the valve equivalent circuit may be used with vector currents and voltages. Also, it is assumed that no grid current flows so that the mutual inductance has no effect on the current in L. The equivalent circuit for the valve and load is given in Fig. 13.4.*b*. If this amplifier provides its own input, then the voltage across the grid winding of M must be $\mathbf{V}_g$ as shown. If $\mathbf{Z}$ is the vector impedance of the LC circuit then

$$\mathbf{Z} = (R + j\omega L)/(1 - \omega^2 LC + jR\omega C).$$

The following equations follow directly from the network

$$\mu\mathbf{V}_g = \mathbf{I}_a(r_a + \mathbf{Z}),\ (R + j\omega L)\mathbf{I}_l = \mathbf{I}_c/j\omega C,$$

$$\mathbf{I}_a = \mathbf{I}_l + \mathbf{I}_c \text{ and } \mathbf{V}_g = j\omega M\mathbf{I}_l.$$

Then $\qquad \mathbf{I}_a = \mathbf{I}_l(1 + j\omega CR - \omega^2 LC),$

i.e., $\qquad \mathbf{I}_a = (1 + j\omega CR - \omega^2 LC)\mathbf{V}_g/j\omega M$

and $\qquad \mu\mathbf{V}_g = (1 + j\omega CR - \omega^2 LC)\mathbf{V}_g(r_a + \mathbf{Z})/j\omega M.$

On substituting for $\mathbf{Z}$ it is found that

$$j\omega M\mu = (1 + j\omega CR - \omega^2 LC)r_a + R + j\omega L.$$

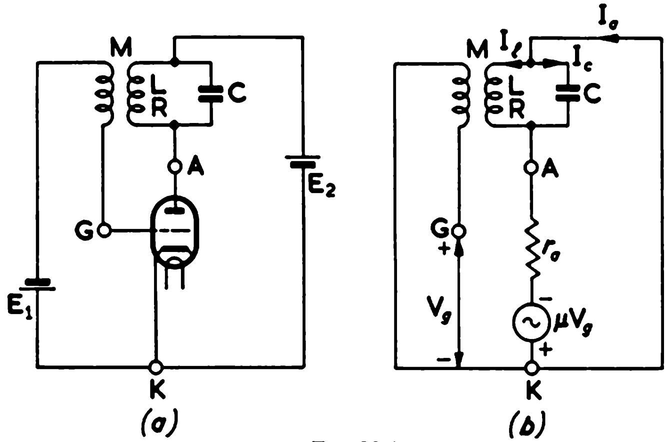


Fig. 13.4

This equation must be satisfied if the circuit is to provide its own input. The real and imaginary parts must be equal, and hence

$$M = L/\mu + CR/g_m \text{ (using } g_m = \mu/r_a)$$

and $\qquad \omega^2 LC = 1 + R/r_a.$

Both of these equations are necessary conditions for oscillation. The first gives the value of M for given values of L, C, R, μ and g_m. The second equation gives the angular frequency of oscillation

$$\omega^2 = \frac{1}{LC}(1 + R/r_a).$$

Normally $r_a \gg R$ and then $\omega^2 \simeq 1/LC$, the familiar value of the resonant frequency of the LC circuit when resistances are neglected. This analysis shows how self-oscillation can be obtained from a tuned-anode amplifier working under Class A conditions with small a.c. amplitudes and feeding back to the grid a voltage which just maintains these conditions.

13.5. Class C Oscillators and Amplitude Limitation

In practice, it is extremely difficult to operate stably under the conditions described in the last paragraph. The feedback is usually greater than the value required to maintain Class A operation. Then the amplitude of the oscillation increases until it is limited by the valve characteristics in much the same manner as with the negative-resistance oscillator described in Section 13.2. Large amplitudes of anode and grid voltages are built up, and the analysis based on small signals no longer applies, except as an indication of the starting conditions. The frequency of oscillation is still given approximately by $\omega^2 = 1/LC$, but the M condition becomes

$$M > L/\mu + CR/g_m.$$

However, this applies only in the early stages of the build-up of oscillation, since g_m, and to some extent μ, vary over the cycle as the amplitudes grow.

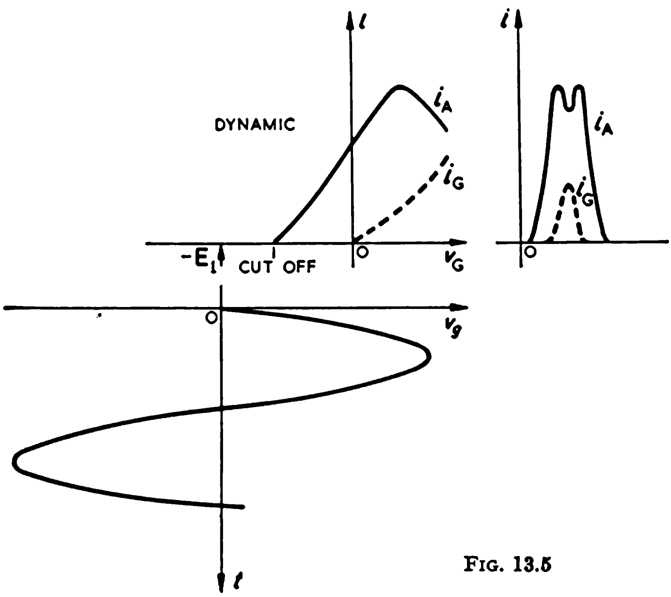
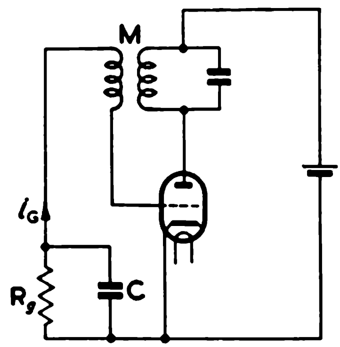


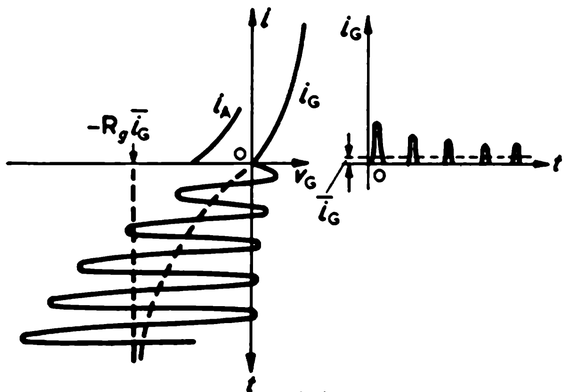
FIG. 13.5

When the grid voltage goes into the positive region, grid current flows. The oscillator is usually adjusted to operate under the conditions described for Class C amplifiers in Chapter 8. This means that the grid bias is beyond cut-off and the grid is driven well into the positive region. The behaviour of such an oscillator can be determined to some extent if the dynamic grid characteristic is known into the positive region. Such a characteristic is shown in Fig. 13.5. When the grid is positive the anode current continues to rise at first, but subsequently the curve may turn over. This is largely due to the increasing current taken by the grid (shown by broken line). In this region g_m and μ both drop. Thus at the peak of the grid voltage the condition for maintenance of oscillation is not satisfied. The amplitudes settle to equilibrium values in which the average values of g_m and μ over a cycle satisfy the M condition. It may be noted that g_m is zero over an appreciable part of the cycle.

A self oscillator cannot operate under Class C conditions with fixed grid bias. On switching on, the anode current would be cut-off, and no oscillation could build up. The bias is usually obtained automatically by utilizing the flow of grid current through a grid leak R_g, as shown in Fig. 13.6.*a*. There is a condenser C across R_g such that $1/\omega C \ll R_g$. Then, only the mean grid current flows through R_g, and the bias is equal



(*a*)



(*b*)

FIG. 13.6

to R_g times the mean current, i.e., $-R_g\bar{i}_G$. With this arrangement there is no bias on switching on. The bias gradually builds up until the equilibrium condition is reached, rather after the manner shown in Fig. 13.6.*b*.

The equilibrium condition for a Class C oscillator with grid leak bias adjusts itself automatically so that, during the portion of the cycle when the valve is conducting, it passes enough energy from the h.t. supply to the oscillatory circuit to make up the loss of energy during the whole cycle. If for any reason the amplitude of oscillation decreases slightly, then it is essential that the bias decreases also, otherwise the circuit losses are not made up and the amplitude of the oscillation decreases further and ultimately falls to zero. The rate at which the bias voltage may be changed depends on the time constant of R_g and C, and there is therefore a limit to the value of R_gC. The actual value is related to the rate at which oscillations decay in the resonant circuit. When oscillation ceases in this manner the valve remains cut-off until C is discharged sufficiently. Then the valve passes current and amplifies once more. The oscillation builds up again with steadily increasing amplitude and bias. If the bias lags behind too much, it increases beyond the equilibrium value and oscillation again decreases. Thus when R_gC is too great, intermittent oscillation may be obtained. The effect is known as squegging.

The operating conditions for a Class C oscillator are the same as those for a Class C amplifier. The power output is rather lower, since the oscillator has to supply its own grid power.

13.6. Other Tuned Oscillators

In the tuned-anode oscillator of Fig. 13.4 the grid and anode a.c. voltages are very nearly in antiphase. The grid voltage $\mathbf{V_g}$ differs in phase by 90° from $\mathbf{I_l}$, and $\mathbf{I_l}$ differs from $\mathbf{V_a}$ by 90°, except for the small effect of R. The vector diagram is shown in Fig. 13.7. In this diagram $\mathbf{V_g}$ is drawn first and then $\mathbf{I_l}$, 90° behind $\mathbf{V_g}$. The voltage vectors for $R\mathbf{I_l}$ and $j\omega L\mathbf{I_l}$ are drawn next, in phase with and 90° ahead of $\mathbf{I_l}$ respectively. The resultant gives the voltage across the load, which is $-\mathbf{V_a}$. $\mathbf{V_a}$ and $\mathbf{V_g}$ are very nearly in antiphase, but the small departure from 180° is essential to give the correct feedback. For many purposes it is convenient to ignore the effects of resistances of the oscillatory circuits, and then $\mathbf{V_a}$ and $\mathbf{V_g}$ are directly in antiphase. This is a useful approximate rule for applying to other tuned oscillators. Then the rough criterion of suitability of a circuit as an oscillator is that the grid and anode voltage should be in antiphase, and there should be ready means of adjusting the size of the grid voltage to ensure that it is sufficient.

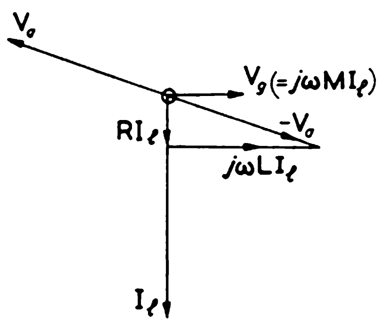


FIG. 13.7

A common triode oscillator, known as a Hartley circuit, is shown in Fig. 13.8.*a*. Only the essential a.c. connections are given. In this circuit $\mathbf{V_a} = j\omega L_1\mathbf{I_l}$ and $\mathbf{V_g} = -j\omega L_2\mathbf{I_l}$, ignoring mutual inductance between L_1 and L_2. (The oscillatory current $\mathbf{I_l}$ is large compared with any current in the other leads, and we may assume that the same current $\mathbf{I}_l$ flows in L_1 and L_2.) Thus $\mathbf{V}_g/\mathbf{V}_a = -L_2/L_1$ and the size of $\mathbf{V}_g$ may be varied by moving the tapping point on the inductance. This is a simple oscillator and easy to adjust. The frequency is determined by L and C, where L is equal to $L_1 + L_2$. A circuit diagram showing the d.c. con-

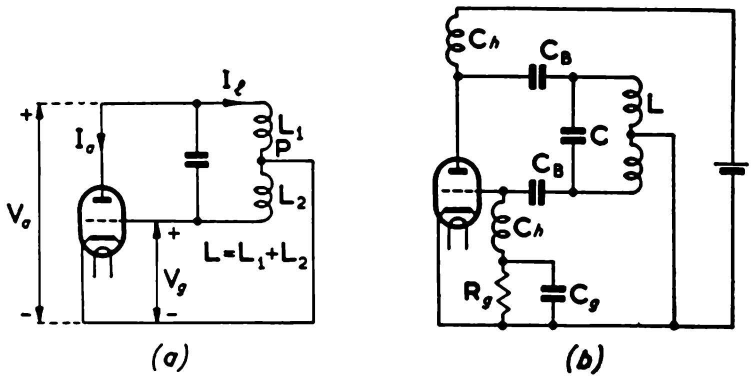


FIG. 13.8

nections as well is given in Fig. 13.8.*b*. The condensers C_B, of large capacitance, are for d.c. separation; the inductances Ch are of high value to act as chokes at the frequency of oscillation; R_g and C_g provide the grid bias.

Another commonly used oscillator, known as the Colpitts' circuit, is shown in Fig. 13.9. Here $\mathbf{V}_g/\mathbf{V}_a = -C_1/C_2$.

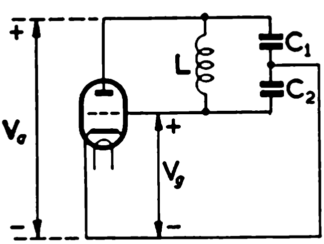


FIG. 13.9

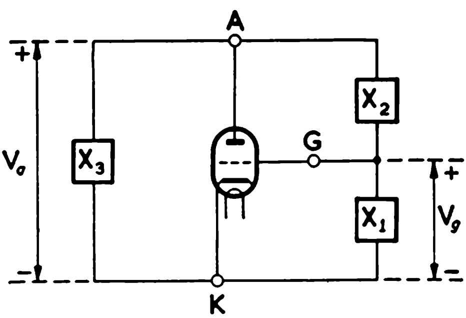


FIG. 13.10

In Fig. 13.10 a generalized triode oscillator circuit is drawn showing reactances between each pair of the electrodes. Circuit and valve resistances are neglected, and inter-electrode capacitances are included in X_1, X_2 and X_3. The three reactances form a closed oscillatory circuit. Obviously they cannot all be inductances or all capacitances, as there would then be no tuned circuit. The same current flows through all the reactances. Since $\mathbf{V}_g$ and $\mathbf{V}_a$ must be of opposite sign for oscillation, it follows that X_1 and X_3 must both be inductances or both capacitances. Thus the possible arrangements of the reactances are as shown in the table.

	X_1	X_2	X_3
I	L	C	L
II	C	L	C

It may be seen that the Hartley circuit conforms to type I and the Colpitts circuit to type II. This generalized circuit may be used to interpret any tuned oscillator in which there are no mutual couplings.

Another circuit, known as a tuned-anode, tuned-grid oscillator, is shown in Fig. 13.11. This case conforms to the generalized circuit if the anode–grid capacitance is included to form X_2. This is a type I oscillator with X_1 and X_3 both inductive. Thus the two tuned circuits, including the other inter-electrode capacitances, must both be tuned to a frequency above the frequency of oscillation.

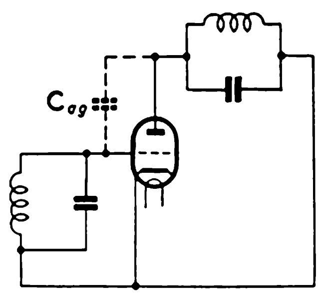


FIG. 13.11

In the generalized oscillator circuit of Fig. 13.10 the current flowing round the circuit depends on the values of the three reactances. If any one reactance is very high the current is small and the grid voltage may not be sufficient to give oscillation. Oscillation in the circuit of Fig. 13.11 occurs more readily at high frequencies when the reactance of C_{ag} is small. In pentodes C_{ag} is very small and there is much less likelihood of oscillation. Hence the circuit of Fig. 13.11 with a pentode may be used as a stable amplifier; sometimes it is necessary to place a screen between the input and output circuits.

13.7. Transistor Oscillators

Transistor amplifiers with suitable feedback may act as oscillators, and most of the triode circuits described in Sections 13.4 and 13.6 have corresponding transistor versions. The circuit in Fig. 13.12 shows a tuned-collector oscillator in which a common-emitter arrangement is used. The main oscillatory circuit is L_1C_1, and the feedback is obtained through the mutual inductance between L_1 and L_2; the resistance R provides the correct operating base voltage, and C is a d.c. blocking condenser. The similarity to the tuned-anode oscillator is obvious. In considering the latter in Section 13.4 we assumed that the resistance between grid and cathode was infinite. In the transistor case the input impedance with the common-emitter circuit is a few hundred ohms, and this factor constitutes a major difference between valve and transistor oscillators. In Section 13.6 we established some general properties of triode oscillators on the assumption that only reactances between the electrodes need be considered. Now the resistance between base and emitter may be comparable with the reactance, and the conditions for oscillation are rather more complicated. One result is that, in order to obtain the correct phase of the feedback, the frequency of oscillation may differ appreciably from the resonant frequency of the LC circuit. Also, the input impedance acts as a damping load on the oscillatory circuit.

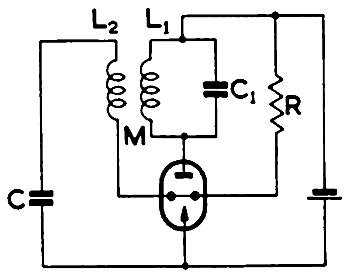


FIG. 13.12

Transistor versions of the Hartley and Colpitts oscillators are shown in Fig. 13.13.*a* and *b*. These circuits should be compared with Fig. 13.8.*a* and 13.9.

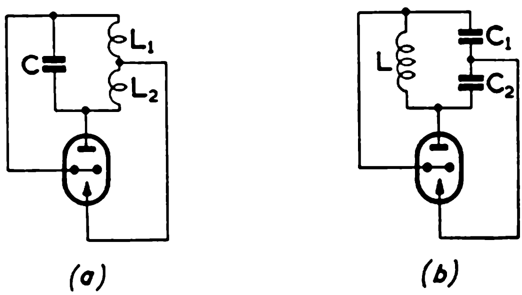


FIG. 13.13

13.8. Feedback and Negative Resistance Oscillators

So far we have distinguished between feedback and negative resistance oscillators. Actually, it is always possible to deal with a feedback oscillator in terms of negative resistance. For example, consider the two-stage, direct-coupled transistor amplifier of Fig. 13.14.*a*, in which the output is fed back to the input; R_1 and R_2 are the coupling resistances, and R is the load of the second stage. If the relation between v_1 and i_1 is measured it is found that there is a negative resistance portion of the curve, as shown in Fig. 13.14.*b*. Thus if a parallel tuned circuit is connected with a suitable biasing voltage, as shown in Fig. 13.14.*c*, oscillation can occur. A triode circuit with similar properties is shown in Fig. 13.15.

Another interesting oscillator which may be discussed in terms of feedback or negative resistance is the transitron, which uses a pentode valve in a rather unusual way. In Section 6.11 we discuss the variation of

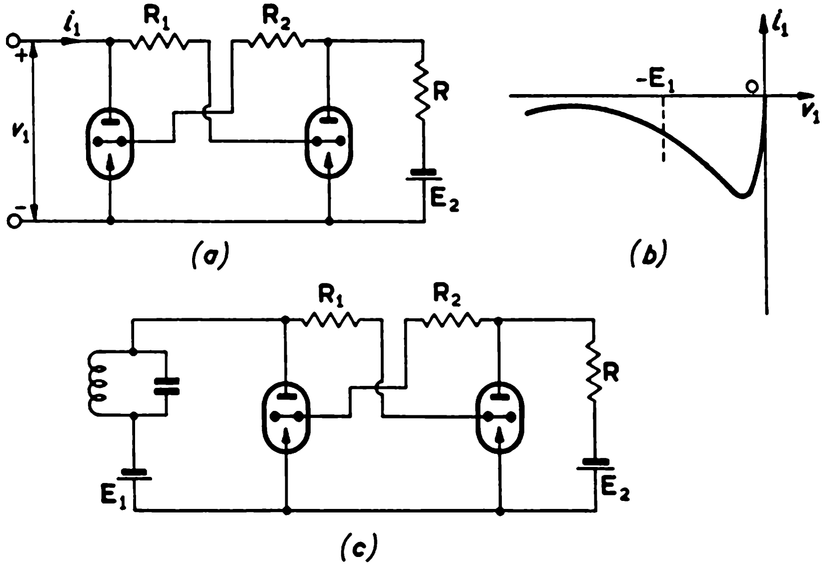


FIG. 13.14

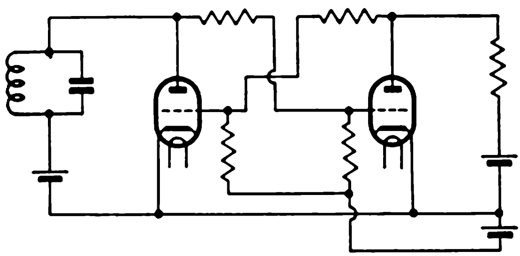


FIG. 13.15

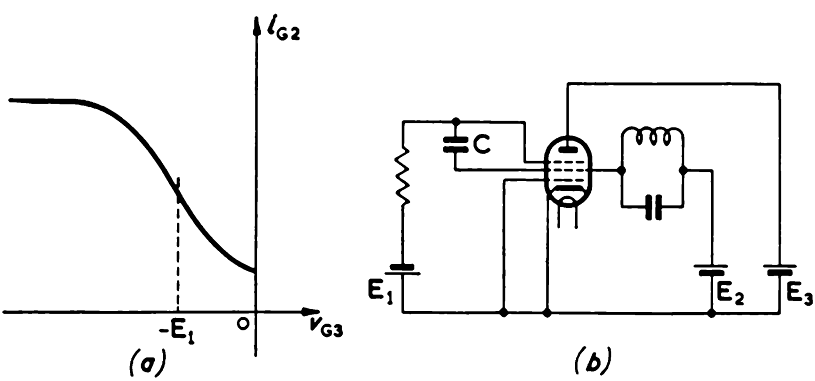


FIG. 13.16

screen current with suppressor voltage, when all the other voltages are constant. The characteristic takes the form of Fig. 13.16.*a*, and as v_{G3} rises, i_{G2} falls. If G_3 is used as a control grid and G_2 as an output electrode, then the output current is in antiphase with the signal voltage. Hence, with a resistance load, the output voltage is in phase with the signal. This means that direct feedback from the screen to the suppressor gives the necessary phase condition for oscillation. The transitron oscillator, based on this principle, is shown in Fig. 13.16.*b*; the condenser C provides the feedback. The oscillator can also be analysed in terms of negative resistance. For constant v_A and v_{G1},

$$i_{G2} = f(v_{G3}, v_{G2}),$$

and for small changes

$$i_{g2} = \frac{\partial i_{G2}}{\partial v_{G3}} v_{g3} + \frac{\partial i_{G2}}{\partial v_{G2}} v_{g2} = g_{23} v_{g3} + v_{g2}/r_2.$$

From Fig. 13.16.*a* it is seen that g_{23} can be negative in value; r_2 is positive. When G_2 and G_3 are joined together changes in v_{G3} equal changes in v_{G2}. Then

$$\mathbf{I}_{g2} = (g_{23} + 1/r_2)\mathbf{V}_{g2}$$

or

$$\frac{\mathbf{V}_{g2}}{\mathbf{I}_{g2}} = \frac{1}{g_{23} + 1/r_2} = r.$$

As long as g_{23} is negative and numerically greater than $1/r_2$, the dynamic resistance r is negative. Hence, with a suitable parallel tuned circuit between the screen grid and the cathode, oscillation can be obtained.

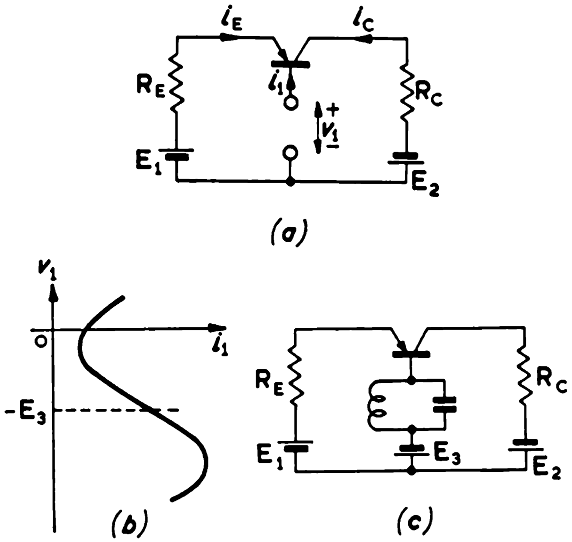


FIG. 13.17

The transitron oscillator is sometimes referred to as a negative mutual conductance (or transconductance) oscillator.

Transistors can be used in a variety of ways to give negative resistance oscillators. For example, if a transistor has α_{ce} greater than unity this can lead to a negative resistance characteristic. With a point-contact transistor in the circuit of Fig. 13.17.*a* the relation between v_1 and i_1 is as shown in Fig. 13.17.*b*. If a parallel tuned circuit is connected in the base lead with suitable bias, as in Fig. 13.17.*c*, oscillations occur when the magnitude of the negative resistance is less than the parallel resonant resistance of the tuned circuit. This circuit can also be considered to

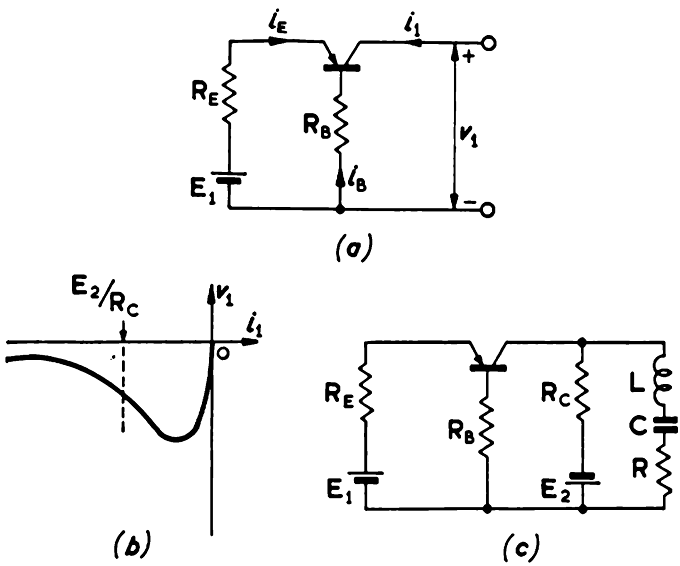


FIG. 13.18

give positive feedback through the resistance of the tuned circuit, since the changes in collector current are greater in magnitude than changes in emitter current.

All the negative resistance oscillators discussed above have used a parallel resonant circuit as the frequency-determining element. Continuous oscillations may be obtained in a suitable series tuned circuit, and an example is given in Fig. 13.18, using a point-contact transistor. Oscillation occurs in this case, provided the magnitude of the negative resistance exceeds R, the series resistance of the tuned circuit. The quiescent point is determined by R_C and E_2. The base resistance R_B can be considered to give positive feedback. It should be noted that, in this case, the negative resistance portion of the characteristic exists uniquely for a definite range of current, whereas in the previous cases of negative resistance the controlling factor is the voltage range.

13.9. Triode Oscillators for Ultra-high Frequencies

As the frequency of operation is increased the inductance of the electrode leads and the inter-electrode capacitances become increasingly important. With ordinary valves it is impossible to make a simple Hartley or Colpitts oscillator, since the connections from the valve to the tuned circuits have appreciable reactance. Some high-frequency oscil-

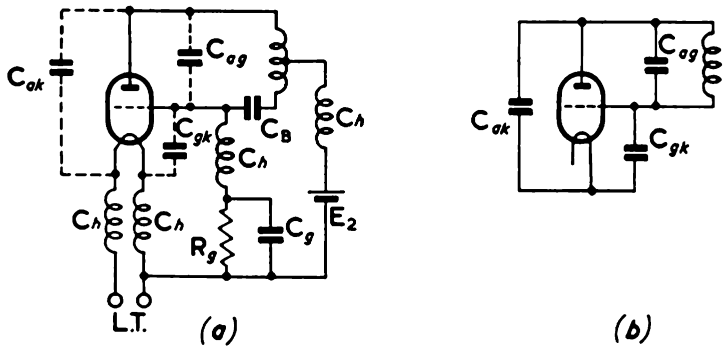


FIG. 13.19

lators use the anode–cathode and grid–cathode capacitances as C_1 and C_2 in a Colpitts circuit, and then there is no need for the connection from the common point to the cathode. The ratio of $\mathbf{V}_g$ to $\mathbf{V}_a$ equals—C_{ak}/C_{gk}. This modified Colpitts circuit is shown in Fig. 13.19. The anode and cathode supplies are fed through chokes. Circuits of this type may operate quite successfully at high frequencies, but they do not conform to any of the simple oscillators which have already been described, since there are always stray and indeterminate couplings between different parts of the circuits. Special triodes have been designed which do permit of fairly simple and controllable operation even at frequencies of 4,000 Mc/s and higher. Such a triode is shown diagrammatically in Fig. 13.20. This valve has a planar electrode system, and the leads to the anode and the grid are in the form of copper disks. The cathode lead is essentially a metal tube. This valve is constructed in such a way that it forms an integral part of the external circuit, which, instead of using lumped inductances and capacitances, employs short sections of high-frequency transmission lines.

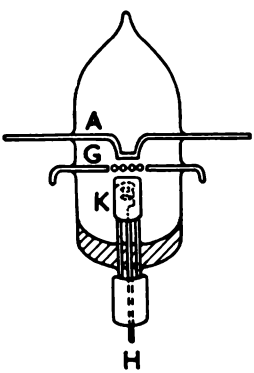


FIG. 13.20

The arrangement is illustrated in Fig. 13.21. The triode plugs into a circuit consisting of three co-axial tubes, each electrode being attached to one tube with the anode outermost and the cathode innermost. There are then two co-axial transmission lines attached to the electrodes. The grid tube is part of both lines. Its outside surface is the inner conductor of the anode–grid line, and its inside surface is the outer conductor of the grid–cathode line. Although one conductor is shared by these two circuits, the high-frequency currents are completely isolated from one another, as they flow only on the surfaces of the conductors. The two co-axial transmission lines are short-circuited by means of sliding plungers, P_1 and P_2, whose positions may be varied for tuning purposes. The lengths of short co-axial lines act as inductances or capacitances, depending on their length relative to the wavelength of the oscillation. When the length lies between 0 and $\lambda/4$ the reactance is inductive and

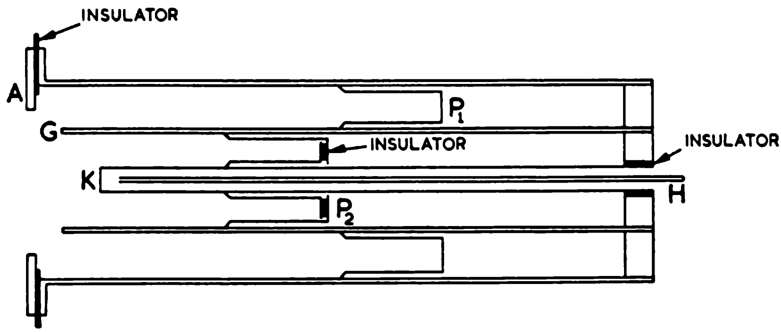


FIG. 13.21

when it lies between $\lambda/4$ and $\lambda/2$ the reactance is capacitive. Both of these relationships are unaffected by the addition of any integral number of half wavelengths to the length of the transmission line. It may be seen from Fig. 13.21 that the valve and its capacitances are attached to the two transmission lines in such a way that they conform to the simple generalized circuit of Fig. 13.10; X_1 is the cathode–grid transmission line with C_{gk} in parallel, X_2 is the anode–grid line with C_{ag} in parallel and X_3 is simply C_{ak}. Since X_3 is capacitive, then X_1 must be capacitive and X_2 inductive. The transmission lines are adjusted to give the required reactances for oscillation.

The reactances of X_1, X_2 and X_3 have all to be sufficiently small to give oscillation. In some cases C_{ak} has to be increased if oscillation is to occur in the circuit of Fig. 13.21. On the other hand, if C_{ak} is very small there is no oscillation, and this circuit may be used as a stable amplifier. The signal is connected to the cathode–grid line, and the output is taken from the anode–grid line. This is an example of a common-grid amplifier. It is frequently but inaptly called a grounded-grid amplifier.

Common-anode circuits have also been used with transmission lines as oscillators for very high frequencies. The common-cathode oscillator, which is another name for the tuned-anode, tuned-grid circuit of Fig. 13.11, is not suitable for operation at the highest frequencies.

CHAPTER 14

ELECTRONS AND FIELDS

14.1. Induced Currents due to Moving Charges

In this chapter we investigate further the interchange of energy between moving electrons and electric fields with a view to explaining the operation of some of the special valves which are used for the generation of alternating currents at ultra-high frequencies. Consider the movement of a negative

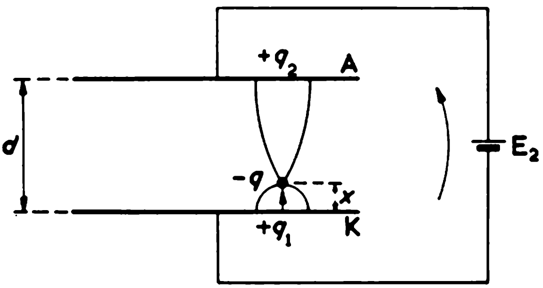


FIG. 14.1

charge between two large parallel planes (Fig. 14.1), under the influence of the steady field produced by the battery E_2. When the charge $-q$ is in the space between the two planes it induces positive charges $+q_1$ and $+q_2$ in the planes, such that

$$q_1 + q_2 = q.$$

When the charge leaves K,

$$q_1 = q \text{ and } q_2 = 0$$

and when it reaches A,

$$q_1 = 0 \text{ and } q_2 = q.$$

As the charge moves across the space the positive charge is continuously transferred from K to A. The transfer must take place through the battery and the total transfer is $+q$. When the charge moves a distance dx in the space we assume that charge dq_1 is transferred from K to A. The electric field strength in the space is $-E_2/d$, and the force on the charge is qE_2/d. The work done in moving the charge the distance dx is qE_2dx/d. This energy is obtained from the battery by the transfer of dq_1 through it in the direction of the arrow. Hence

$$qE_2dx/d = E_2dq_1 \quad \text{or} \quad dq_1 = qdx/d.$$

The current flowing in the external circuit is

$$\frac{dq_1}{dt} = qu/d, \text{ where } u = \frac{dx}{dt}.$$

Thus the movement of the charge induces a current in the external circuit, and the magnitude of the current depends on the velocity of the charge. The induced current starts when the charge leaves K and ends when it reaches A. We thus see that the motion of a charge to or from an electrode produces a current to that electrode in the external circuit. When there are many moving charges the external current is the sum of all the induced currents.

Similar considerations may be applied to the movement of a charge

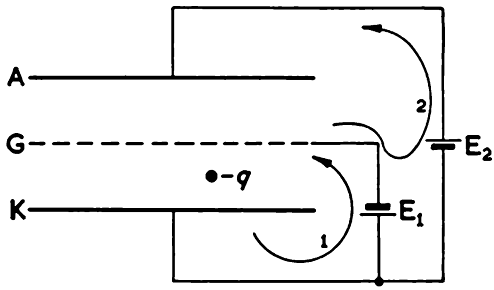


FIG. 14.2

from the cathode to the anode of a triode as illustrated in Fig. 14.2. As the charge moves from the cathode to the grid an induced current flows through the grid battery in the direction of arrow 1. After the charge passes through the grid the induced current flows through the grid and anode batteries as shown by arrow 2. If the triode has a high amplification factor μ there is very little charge induced on the anode when the moving charge is in the cathode–grid space, and similarly there is negligible induction on the cathode after the charge passes the grid. Of course, the values of E_2 and E_1 have to be such that $-E_1 + E_2/\mu$ is positive in order to make the charge move from the cathode. The values of the induced currents again depend on the charge velocity. The directions of current flow in Fig. 14.1 and 14.2 show that in both cases the current in the E_2 battery is in a direction to take energy from the battery. The motion of the charge between K and G in Fig. 14.2 causes a charging current in the grid battery, i.e., energy is given to the battery. The subsequent movement from G to A gives a discharging current, and an equal quantity of energy is taken from the battery. We have here an example of how current may flow to an electrode even when it collects no charge from the space. Further consideration of the energy conditions in the batteries shows that energy is given to a battery when the negative charge moves towards a negative electrode. Energy is taken from the batteries when the negative charge moves away from a negative electrode or towards a positive electrode.

14.2. Energy Considerations

In the triode in the previous paragraph there are two separate sources of the steady field, namely the grid and anode batteries. The energy relationships between the moving charge and the sources may be stated in an alternative form. When charge moves against, or is retarded by, the field of force produced by a source it gives energy to that source. When a charge moves with, or is accelerated by, the force produced by a source, it takes energy from the source. In the triode example both sources are steady. However, the same energy relationship applies when one or more of the sources is varying. This may be confirmed by reference to Class A, B and C amplifiers in Section 8.11. There it is shown how more energy can be transferred to the a.c. circuit by arranging for a greater proportion of the electrons to move towards the anode during the negative half cycle of the a.c. anode voltage. We may consider the case of the Class A amplifier in more detail in terms of the ideas of the present chapter. The relative waveforms of anode voltage, grid voltage and anode current are shown in Fig. 14.3 for a Class A amplifier with a resistive load. The electrons which flow from the cathode to the grid during the first half cycle move against the alternating grid force (the alternating grid voltage is negative). These electrons therefore give energy to the source of the grid a.c. field. As these same electrons move from the grid to the anode, they are moving with the alternating grid force (they are being accelerated by the negative a.c. grid voltage), and so take back the energy which they have already given to the grid. At the same time they are moving towards the anode whilst the alternating anode force is accelerating, and so they take energy from the anode a.c. source. During the second half cycle the electrons first take energy from the a.c. grid source and then return it. They now move against the a.c. anode field of force, and so give energy to it. More electrons flow during the second half cycle, and on balance the electrons give more energy to the a.c. anode source than they take from it. The resultant energy exchange with the grid source is zero.

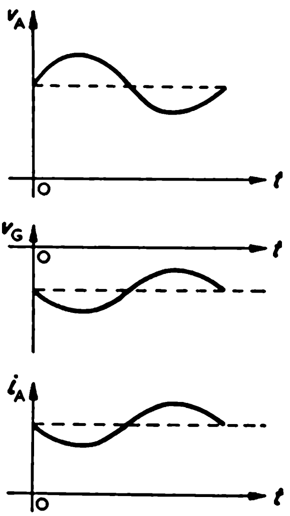


FIG. 14.3

We have considered above each alternating field of force separately and ignored the d.c. fields whilst considering the exchange between the electrons and each separate field. There is, of course, energy exchange going on continuously between the electrons and all the sources of field in each region. The d.c. sources supply the energy, some of which is transferred to the a.c. circuits, and the whole process goes on simultaneously. In Chapter 15 there are examples where the electrons move first in the d.c. field to acquire energy, some of which they then give up in an a.c. field. For energy conversion from d.c. to a.c. the electrons should be given high velocities by the d.c. source and then slowed down in the retarding part of the a.c. field. The slower the electrons are finally, the greater the efficiency of energy conversion.

When an amplifier has a resistive load the anode voltage is in antiphase with the anode current, and the concentration of the electrons at the negative half cycle of the anode voltage is an essential feature of the

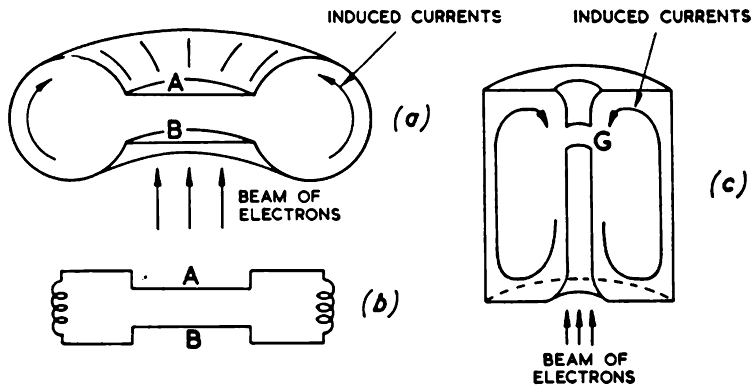


FIG. 14.4

operation. In most amplifiers and oscillators the load is resistive (frequently parallel-resonant) and the conditions for energy conversion arise automatically in the adjustment of the circuits. A common example at high frequencies of this type of operation is found in space resonators, in which the electrons move between two electrodes which are integral parts of the circuit. Examples are shown in Fig. 14.4.*a* and *b*. In the former there is a section of a space resonator in which two parallel plates A and B are joined by a doughnut-shaped conductor. This resonator, which is sometimes called a rhumbatron, is an extension of the " lumped " circuit in Fig. 14.4.*b*, where two plates act as a capacitor and are tuned to resonance by a number of inductors in parallel. In such a space resonator the plates A and B are usually grids permitting the passage of a beam of electrons normally. The beam is modulated in density so that one complete cycle of modulation passes through the resonator during the natural resonant period. Induced currents flow round the doughnut and build up the a.c. field across AB with the retarding half cycles coinciding with the maximum density of electrons. A second type of space resonator is shown in section in Fig. 14.4.*c*. In this case the beam of electrons builds up the a.c. field across the gap G.

14.3. The Energy Equation

When an electron moves in a steady field between two points at potentials v_1 and v_2 the gain in kinetic energy equals the loss of potential energy. This is expressed in the Energy Equation

$$e(v_2 - v_1) = \tfrac{1}{2}m(u_2^2 - u_1^2),$$

u_1 and u_2 being the velocities at the two points. In the case of a diode with steady potential difference E_2 between anode and cathode, the energy equation becomes $eE_2 = \frac{1}{2}mu_A^2$, where u_A is the electron velocity at the anode and it is assumed that the electron leaves the cathode with zero velocity. If the diode has an alternating potential difference as well, then the total voltage between anode and cathode is $v_A = E_2 + \hat{v}\sin\omega t$. Provided the electron transit time is negligible compared with the duration of one cycle of the field, then each electron traverses the diode in a steady field, and the kinetic energy on arrival at the anode is given by

$$\tfrac{1}{2}mu_A^2 = e(E_2 + \hat{v}\sin\omega t).$$

14.4. Transit Time Loading

When the electron transit time is not negligible the electric field changes during the electron's flight, and the kinetic energy at the anode is not given by the simple energy equation. The curve in Fig. 14.5 shows the variation in anode voltage with time in a parallel-plane diode. At any instant the force on an electron is ev_A/d if the effects of space charge are neglected. An electron which leaves the cathode at time t_1 and arrives at t_2 experiences

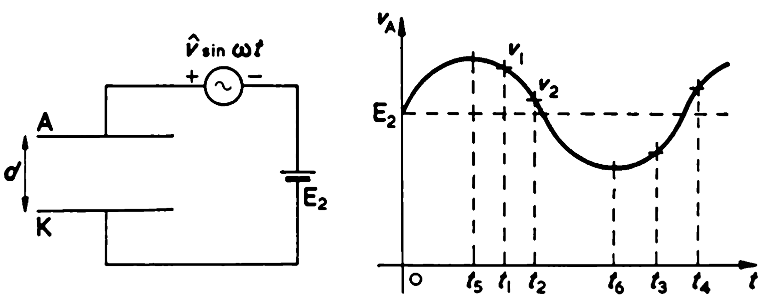


FIG. 14.5

a decreasing force during its transit. At the anode its kinetic energy exceeds what it would have been if it had moved throughout in a steady force given by ev_2/d, the value when it reaches the anode. Such an electron arrives at the anode with kinetic energy greater than the loss of potential energy. Similarly, an electron crossing between t_3 and t_4 while the force is increasing arrives at the anode with kinetic energy less than the loss of potential energy. The excess or deficiency of the kinetic energy over the potential energy is obtained from the alternating field. To find the net energy exchange between the electrons and the field an average must be taken over a complete cycle. These conclusions apply whether the current flowing is temperature-limited or space-charge-limited. In the latter case the greatest number of electrons leave the cathode when v_A is maximum at t_5, i.e., when the force on electrons is beginning to decrease. Also, the smallest number leave at t_6 when the force is starting to increase. Thus, on balance, there is an excess of kinetic energy gained at the expense of the a.c. field, and the effect of the finite electron transit time is to put an additional load on the a.c. generator. When the electron transit time is very large the energy exchange becomes complicated.

Similar considerations may be applied to the grid circuit of a tetrode or pentode amplifier. Usually the transit time between the grid and screen is much less than that between cathode and grid, on account of the high screen voltage. Transit-time effects may therefore be neglected between the grid and the screen when they begin to be appreciable in the cathode–grid space. The energy exchanges in the latter space are practically the same as for the diode considered above. Thus in the pentode at high frequencies the electron transit time causes a load on the grid signal, even though no electrons are collected by the grid. This effect is frequently called input damping due to electron transit time. Transit-time damping also occurs in a triode. However, the conditions are complicated by the varying anode voltage. When the anode voltage is low transit-time effects in the grid–anode space must also be taken into account.

CHAPTER 15

SPECIAL VALVES FOR VERY HIGH FREQUENCIES

15.1. The Klystron

In Chapter 14 it is shown that electrons can be used for transferring energy from a d.c. source to an a.c. circuit. The electrons are given high velocity by a steady field, and are slowed down in an alternating field by making them traverse the latter whilst it is in its retarding phase. Any electrons which move through the alternating field during its accelerating phase decrease the field and reduce the efficiency of power conversion from

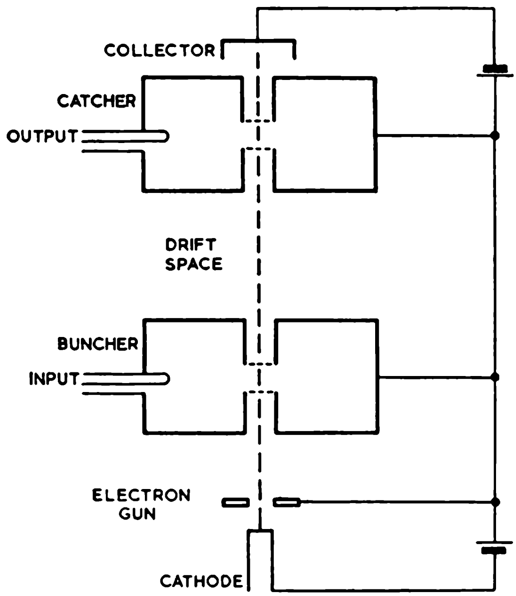


FIG. 15.1

the d.c. supply to the a.c. circuit. In some electronic devices, such as diodes, triodes and pentodes, the electrons move simultaneously in the steady and alternating fields. In others, the two fields are separated. The klystron amplifier is a good example of the latter. The essential parts of one type of klystron are an electron gun to produce a beam of high-velocity electrons and two space resonators of the type described in Section 14.2 and having the same resonant frequency. The high-frequency fields are confined to the spaces inside the two resonators. The arrangement is shown in Fig. 15.1. The final anode of the gun and the two resonators are at the same steady potential. The electrons emerge from the gun with high velocity and enter the first resonator. This resonator is energized at its resonant frequency from some signal source, so that there is a small alternating field across the resonator gap. The electrons which traverse the gap during the half cycle that the field is accelerating emerge with slightly increased velocity. During the other half cycle the electrons leave the resonator with reduced velocity. The electron beam is said to be velocity modulated by the resonator. In the drift space between the two resonators the faster electrons overtake the slower ones and there is electron concentration into bunches by the time the second resonator is reached. The bunches are repeated each alternat-

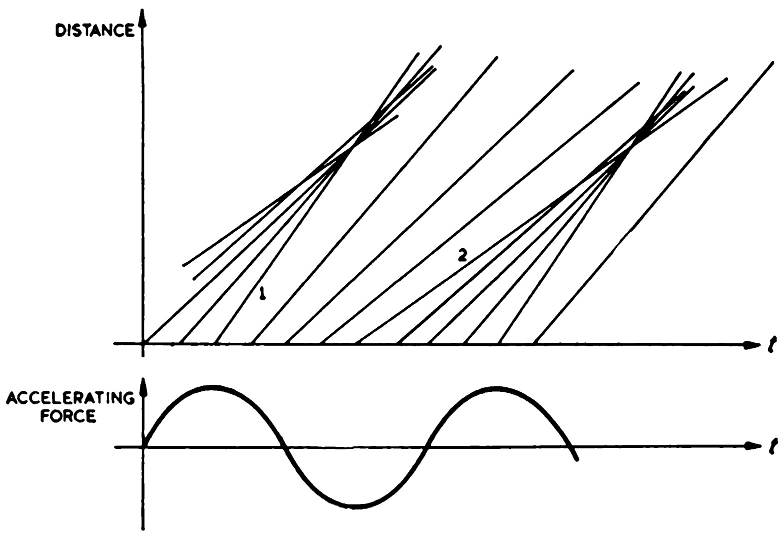


FIG. 15.2

ing cycle. In passing through the second resonator the bunches of charge induce pulses of current round the resonator and a voltage builds up across the gap at the resonant frequency. The phase of this voltage automatically gives a retarding field while the bunches are passing through, and the electrons emerge from the second resonator with reduced velocity, having given up their energy to the resonator. An equilibrium condition is reached when the rate of energy removal by the output coupling equals the rate at which energy is extracted from the beam. The electrons are finally collected by a separate electrode, whose voltage is just sufficient to collect all the electrons.

In Section 14.4 it is shown that energy losses occur in the grid circuit of a conventional amplifier when the electron transit time is comparable with the high-frequency period. Similar losses occur in klystron resonators. However, as the electrons traverse these resonators with high velocity, the electron transit time is much less than in the grid circuit of a pentode where the voltages are low. Hence in klystrons transit-time limitations occur at considerably higher frequencies.

If space-charge effects are neglected the formation of electron bunches in the drift space of a klystron may be illustrated graphically as in Fig. 15.2. The horizontal axis represents time, and the vertical axis represents distance along the drift tube from the first resonator, or buncher, as it is frequently called. On the time axis is shown one cycle of the buncher voltage. The electrons emerge from the buncher with modulated velocities and then move in the field free drift space. The subsequent distance travelled by an electron along the drift space may be found from a straight line whose slope represents the velocity. The greatest slope occurs when the resonator voltage is at its positive maximum (line 1)

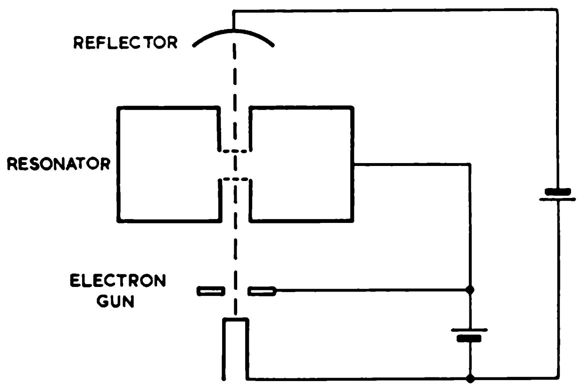


FIG. 15.3

and the least at the negative maximum (line 2). The lines starting from the buncher are equally spaced in time, representing the uniform beam density leaving the buncher. It may be seen from this diagram that at a distance along the drift space there are regions of high and low charge density. Also, the separation in time of the bunches is one period. The second resonator, or catcher, is placed where appreciable bunching has occurred.

Considerable voltage and power amplification can be obtained from a resonator klystron. If coupling is introduced between the output and input resonators the amplifier may provide its own input and so become a self-oscillator.

In the reflex klystron which is shown diagrammatically in Fig. 15.3 only one resonator is used and the electron beam is reflected back so that it traverses the resonator twice. In this way the one resonator acts as both buncher and catcher. Provided the phase of the bunches is correct, self-oscillation may be obtained. The phase may be controlled by adjustment of the d.c. voltages of the resonator and the reflector.

15.2. Travelling-wave Tubes

For the transfer of energy from moving electrons to an a.c. circuit the electrons must be bunched or density-modulated, and the bunches must move in an alternating field during the retarding half cycle. Alternating fields can exist in the form of travelling electromagnetic waves as in a transmission line or a waveguide. If a stream of electrons could be

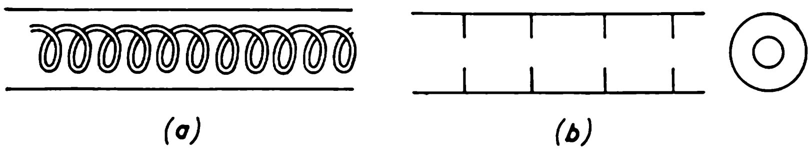


FIG. 15.4

made to move with the same velocity as the travelling wave, then it would be possible for continuous exchange of energy between the stream and the wave. For a net gain of energy to the wave it would be necessary to bunch the stream and for the bunches to move in the retarding regions of the wave. This is the principle of operation of certain types of high-frequency valves, including travelling-wave tubes, space-charge-wave tubes and cavity magnetrons.

Waves on a transmission line or in a normal wave-guide travel with velocities of the order of the velocity of light. Electron beams with such

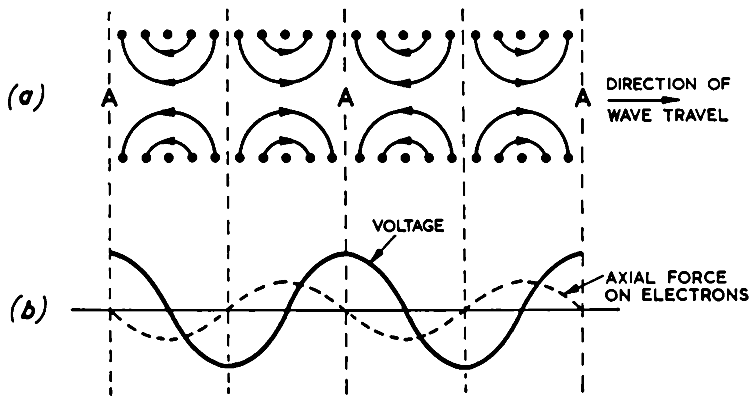


FIG. 15.5

velocities are not practicable. If the waves are to travel in step with the electrons the wave velocity must be reduced. This can be done in several ways, two of which are shown in Fig. 15.4.*a* and *b*. In the former a helix of wire is enclosed inside a conducting tube. This is like a co-axial transmission line with a coiled inner conductor. The velocity of a wave on a transmission line is equal to $1/\sqrt{(LC)}$, where L and C are the inductance and capacitance per unit length. For the helical line,

C is much the same as it would be for a line with a rod inner conductor whose diameter equalled that of the helix. However, L is very much greater and the wave velocity along the axis is considerably reduced. The wave is guided round the wire rather than along the axis, and the velocity reduction is about equal to the axial length divided by the wire length. Fig. 15.4.*b* shows a cylindrical wave-guide with diaphragms at intervals along the length. The axial wave velocity in such a guide is much less than the free space velocity.

In Fig. 15.5.*a* the instantaneous field is shown along the helix. The lines and arrows show the direction of the force on electrons, i.e., in the opposite direction to the electric field strength. The voltage and force

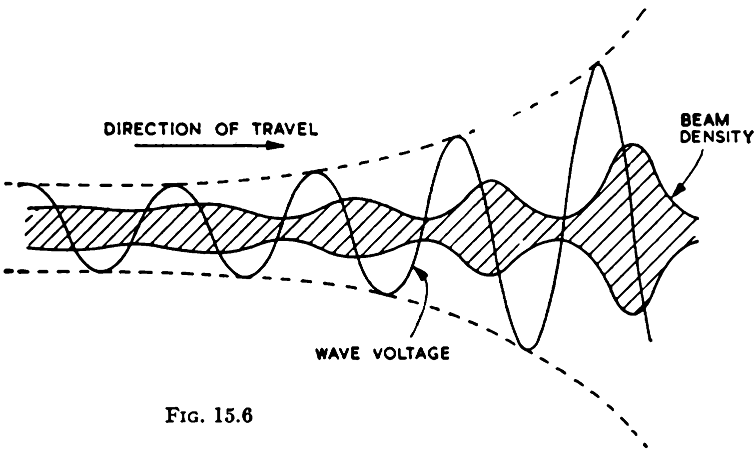


FIG. 15.6

distribution at the same instant are shown in Fig. 15.5.*b*. If now a beam of electrons travels along with a velocity equal to the wave velocity (i.e., the relative velocity of the electrons and the wave is zero) some are accelerated and others retarded, and bunches gradually form at regions A. At A the axial force is zero, and the bunched electrons then travel along in step with the wave, but there is no further exchange of energy, since the bunches are in regions of zero field. If the beam velocity is slightly greater than the wave velocity, bunches still form, but they move to regions to the right of A, where the force is retarding. As long as these conditions are maintained the electrons give up some of their energy to the wave, whose amplitude increases. The relationship between the beam and the wave along the tube is illustrated in Fig. 15.6. Initially the beam is uniform, but it gradually becomes more bunched, and the wave amplitude increases at the same time. Since energy is transferred to the wave at the expense of the kinetic energy of the electrons, the electron velocities ultimately approach the wave velocity and the bunches move to the positions of zero field. No further increase in amplitude is then obtained.

The relation between the electrons and the field in a travelling-wave tube may be illustrated by the movement of traffic along an undulating road, where the undulations correspond to the variation in field along the helix. The excess velocity of the electrons over the wave velocity corresponds to the mean speed of the traffic relative to the road. Because of the undulations there are also fluctuations in the traffic speed. Vehicles slow down when going up hill and speed up on the down gradients. Thus concentrations of traffic or bunches occur on the rises and there is relatively little traffic on the descents. Similar behaviour is obtained with the electrons and the bunches, though it must be remembered that, as the electrons have negative charge, they concentrate in the regions where the potential is decreasing. Also the size of the " hills " increases steadily as the electrons give up energy to the wave. The traffic analogy may also be used to illustrate the behaviour of the electrons when their

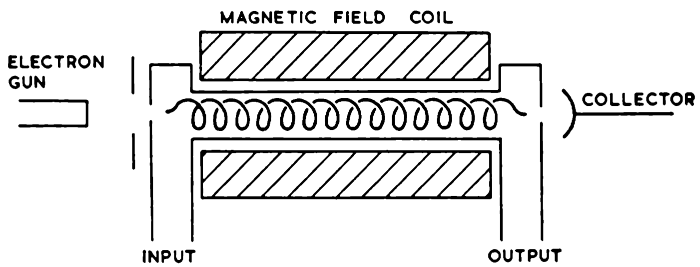


FIG. 15.7

velocity equals the wave velocity. The vehicles would have no mean forward velocity relative to the road, and they would all concentrate in the valleys, which correspond to the regions A in Fig. 15.5.*a*.

In a travelling-wave amplifier the wave is injected into the helix with a wave-guide coupling as shown in Fig. 15.7. The output is taken from the helix with a similar coupling arrangement. This circuit may be used for self-oscillation by feeding back some of the output to the input circuit. There may be some reflected wave travelling backwards from the output end of the helix to the input. If the amplitude of the reflected wave is sufficiently great this may cause self-oscillation. Usually it is necessary to introduce some attenuation in the helix to prevent oscillation in this manner. The length of the helix in a travelling-wave tube may be 20 cm or more. An axial magnetic focusing field is used to keep the beam inside the helix.

One feature which distinguishes travelling-wave amplifiers from other high-frequency amplifiers is the absence of resonant circuits. A wide range of frequencies may be covered without tuning. This feature is valuable in any system which requires a wide frequency bandwidth. Klystron amplifiers, which use sharply tuned resonators, are limited to narrow bandwidths.

15.3. Linear Accelerators

When a travelling-wave tube is used with the electron velocity slightly less than the wave velocity, the bunches form in the accelerating parts of the wave and the electron energies increase at the expense of the wave. This principle is used in the linear accelerator for giving extremely high velocities to electrons. If the wave velocity is gradually increased, say by opening out the helix, the electron velocity may be increased until it approaches close to the velocity of light.

15.4. Space-charge-wave Tubes

If a beam of electrons is density modulated as, say, in a long drift tube of a klystron amplifier, there is a series of bunches along the tube with low density regions between them. The axial electric forces set up by this charge distribution are similar to those shown in Fig. 15.5.*a*. Such a beam is equivalent to a slow travelling wave, and can be used to exchange energy with a second beam of electrons flowing along with the first and with a slightly different velocity. Double-beam arrangements of this type are called space-charge-wave tubes, or double-stream amplifiers.

15.5. Cavity Magnetrons

In Fig. 15.4 two arrangements are shown for producing slow travelling electromagnetic waves. Another possible arrangement is shown in Fig. 15.8, in which the wave travels round the slot and hole cavities and the

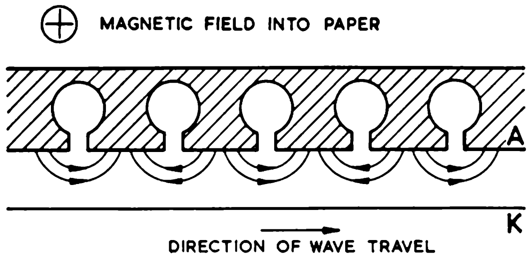


FIG. 15.8

horizontal velocity is considerably less than the free space velocity. Each slot and hole acts rather like a resonant circuit. The electric field is greatest near the slot and the magnetic field within the hole. The lines with the arrows show a possible instantaneous distribution of the electric force on electrons, and the correspondence between this and Fig. 15.5 is obvious. If a horizontal beam of electrons travels in the space between A and K with the correct velocity, this arrangement produces bunching and energy exchange between electrons and field.

A horizontal flow of electrons may be maintained in this structure if it is used as a planar magnetron of the type described in Section 2.8, with K as the cathode, A as the anode and a magnetic field into the plane of the paper. With a steady electric field, the electrons describe cycloidal paths provided space-charge effects are neglected. The nature of the path depends on the relative sizes of the electric and magnetic fields. In the absence of a high-frequency field and with the magnetic field much greater than the cut-off value, an electron leaving the cathode with zero velocity describes a path as shown in the curve in Fig. 15.9.*a* and just returns to the cathode with zero velocity. There are electrons describing similar paths all along the cathode, and there is therefore a horizontal

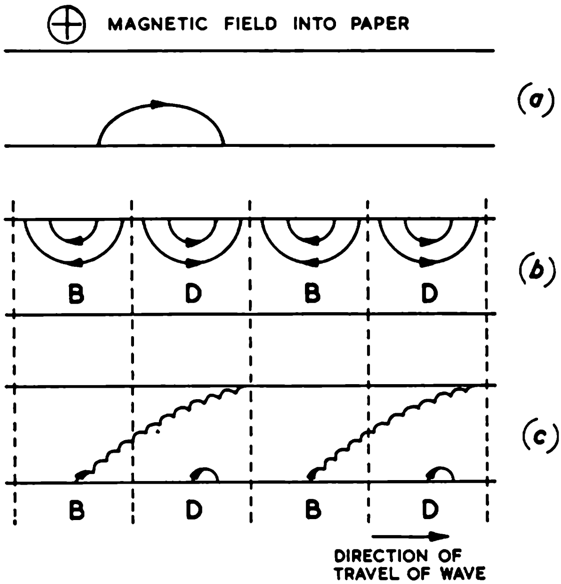


FIG. 15.9

flow of charge. If now an electromagnetic wave is introduced at the left and travels along the system with a velocity equal to the mean horizontal velocity of the electrons, energy interchange takes place between the electrons and the wave. An instantaneous distribution of electric field is shown in Fig. 15.9.*b*, where the arrows again show the direction of the force on electrons. Electrons in regions *D* have the horizontal components of their velocities increased, whilst those at *B* lose velocity. The effect of the increase in velocity is to increase the force (*Beu*) due to the magnetic field, and these electrons are driven back to the cathode with appreciable velocity and are removed from the field. Electrons with reduced velocity have a reduced force due to the magnetic field, and are therefore brought to rest before they reach the cathode. They then move on again under the influence of electric field (steady and varying). As the mean horizontal velocity of the electrons equals the wave velocity, the electrons continue in the retarding high-frequency field and describe a path as shown, ultimately reaching the anode. From the nature of the path it is seen that the mean velocity of the electrons remains constant during the passage from cathode to anode. At first sight it might appear that they have given no energy to the wave. However, although they have maintained their mean kinetic energy, they have lost potential energy in moving from the zero potential cathode to the positive anode, and this energy has been given to the wave. Thus all the unfavourable electrons which leave the cathode at D are quickly removed, and those from B continue to move along with the travelling wave, giving up energy to it as they go.

This planar magnetron acts as a travelling-wave amplifier. Most cavity magnetrons are used as self-oscillators, and this is achieved by bending the planar structure into cylindrical form, as shown in Fig. 15.10, the output end being joined to the input. For this arrangement to work it is essential for there to be an exact number of waves round the circuit.

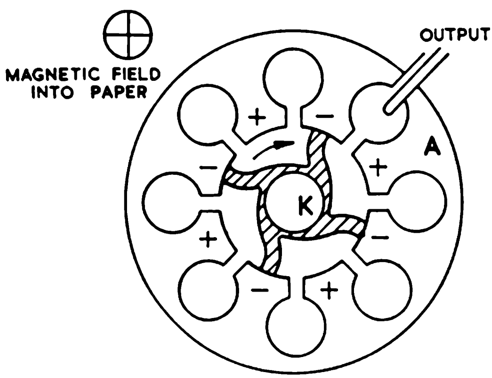


FIG. 15.10

With the eight-cavity magnetron it is possible to have one, two, three or four complete waves. With four waves there is a phase difference of π radians across each cavity, as shown by the signs in the figure; this is known as the π-mode of operation. There are four regions for favourable electrons, and the electron distribution forms an axle-and-spoke arrangement as shown shaded in the figure. The field and the electrons travel together round the system in the direction of the arrow. The favourable electrons in a magnetron can give up most of their potential energy (equal to ev_A) to the high-frequency circuit. The unfavourable electrons are in the field for only a very short time. The magnetron may therefore operate as a generator of high efficiency. Conversion efficiency from d.c. to a.c. of 70 per cent can be obtained.

The main field of use of the special valves described in this chapter is for frequencies from about 1,000 to 100,000 Mc/s.

CHAPTER 16

RECTIFICATION

16.1. Simplified Diode Characteristics

The applications of diodes, whether vacuum, gas or semi-conductor, are based on the asymmetrical features of the characteristics. It may be seen from Fig. 16.1 that these characteristics all show similar general trends. When v_A is positive the resistance is very much lower than when

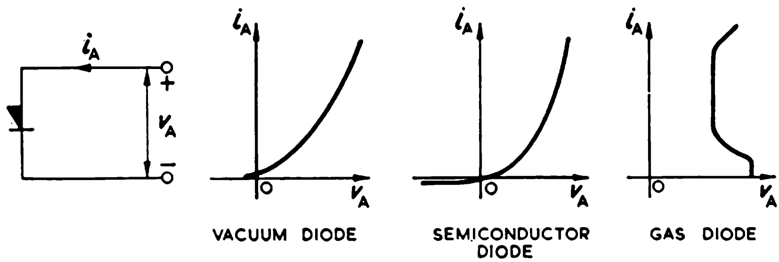


FIG. 16.1

v_A is negative. When v_A is zero the current is either zero or small. In order to simplify the analysis of circuits using diodes, the characteristics can be assumed to have the ideal form shown in Fig. 16.2, in which $i_A = 0$ when v_A is negative, and $i_A = v_A/R_0$ when v_A is positive.

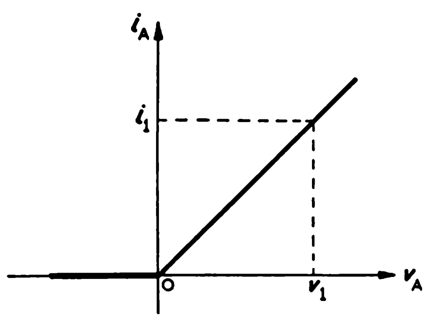


FIG. 16.2

The first application which is to be considered is that of the rectification of an alternating current to produce a direct current. The frequency of the a.c. supply may be 50 c/s, but the application is not limited to this frequency.

16.2. A.c. Supply, Diode and Resistance in Series—Half-wave Rectifier

The circuit arrangement is shown in Fig. 16.3, where a transformer is used to provide a voltage, $v_{12} = \hat{v} \sin \omega t$, and also to act as a low resistance return path for the diode current. The resistance R represents the load

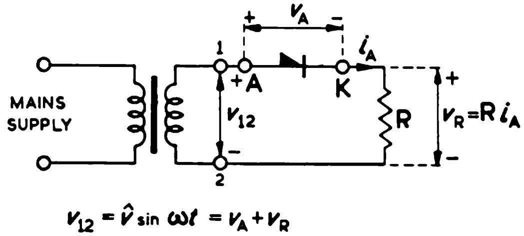


FIG. 16.3

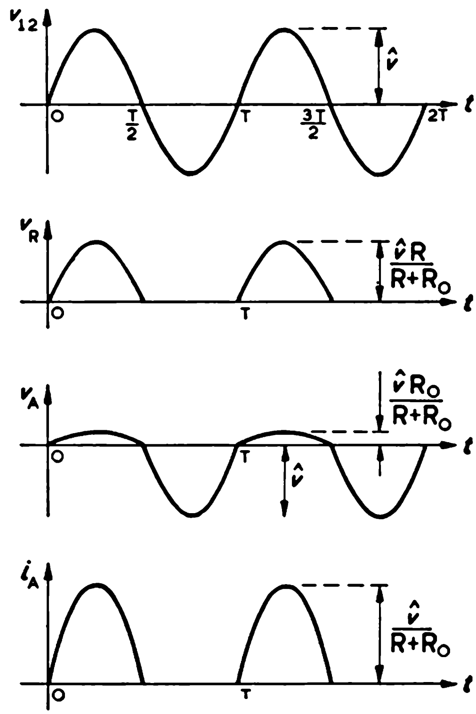


FIG. 16.4

which has to be supplied with d.c. When terminal 1 is positive with respect to terminal 2 current flows through the diode and the resistance. On assuming the ideal diode characteristic, it is found that the current is given by

$$i_A = \frac{\hat{v}}{R + R_0} \sin \omega t \text{ from } t = 0 \text{ to } T/2,\ T \text{ to } 3T/2, \text{ etc.}$$

At the same time the voltages across the load and the diode are

$$v_R = \frac{R\hat{v}}{R + R_0} \sin \omega t$$

and $$v_A = \frac{R_0\hat{v}}{R + R_0} \sin \omega t.$$

When terminal 1 is negative with respect to terminal 2 no current flows through the diode, so that

$$i_A = 0 \text{ from } t = T/2 \text{ to } T,\ 3T/2 \text{ to } 2T, \text{ etc.,}$$

whilst $$v_R = 0 \text{ and } v_A = \hat{v} \sin \omega t.$$

The variations of current and voltage are shown graphically in Fig. 16.4.

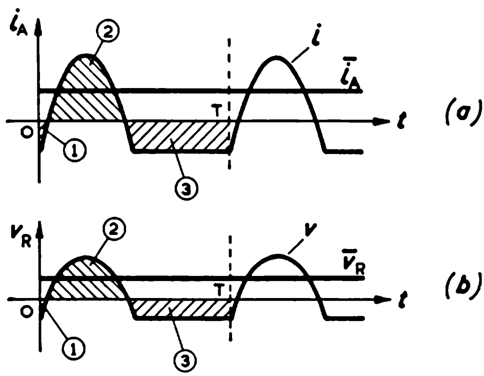


FIG. 16.5

The current through the resistance and the voltage v_R across it are both cyclic but unidirectional. Each may be separated into a steady component and a cyclic component. The steady component of the current is the average value of i_A, and the cyclic component has zero average value over one cycle. The two added together give the actual i_A. The average value of the current is

$$\bar{i}_A = \frac{1}{T}\int_0^T i_A dt = \hat{v}/\pi(R + R_0).$$

The cyclic component i is shown in Fig. 16.5.*a* along with $\bar{i}_A$. The areas 1 and 3 together equal area 2, since the cyclic component has zero average value. The voltages are shown similarly in Fig. 16.5.*b*; the steady component of the voltage is

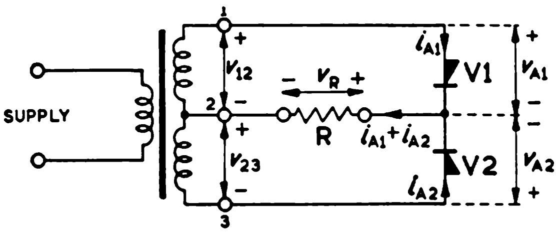
$$\bar{v}_R = R\bar{i}_A = R\hat{v}/\pi(R + R_0).$$

From Fig. 16.4 it is seen that the maximum voltage across the diode occurs when it is not passing current and the anode is negative. This voltage is equal to the peak value of the supply voltage, and is called the peak inverse voltage.

The above circuit has produced a d.c. supply with a superimposed cyclic component, which is usually called a " ripple ". For some purposes this ripple is undesirable, and various methods of reducing it are described later in this chapter. Since current is drawn from the supply only during alternate half cycles, this circuit is known as a half-wave rectifier. A more efficient arrangement, which is described in the next section, gives full-wave rectification by using a second diode during those half cycles when the first diode is inoperative.

16.3. Full-wave Rectification

The circuit of the full-wave rectifier is shown in Fig. 16.6. Two identical



$v_{12} = v_{A1} + v_R = \hat{v} \sin \omega t \,; -v_{23} = v_{A2} + v_R = -\hat{v} \sin \omega t$

$v_R = R(i_{A1} + i_{A2})$

FIG. 16.6

valves are used and $v_{12} = v_{23} = \hat{v} \sin \omega t$. From the circuit it is found that

$$v_{12} = v_{A1} + v_R = \hat{v} \sin \omega t \text{ and } -v_{23} = v_{A2} + v_R = -\hat{v} \sin \omega t.$$

From $t = 0$ to $T/2$, terminal 1 is positive with respect to terminal 2 and diode 1 passes current. At the same time terminal 3 is negative with respect to terminal 2 and diode 2 is non-conducting. Thus for this period of time

$$i_{A1} = \frac{\hat{v}}{R + R_0} \sin \omega t, \quad i_{A2} = 0,$$

$$v_R = \frac{R\hat{v}}{R + R_0} \sin \omega t,$$

$$v_{A1} = \frac{R_0\hat{v}}{R + R_0} \sin \omega t$$

and

$$v_{A2} = -\hat{v}\left(1 + \frac{R}{R + R_0}\right) \sin \omega t.$$

From $t = T/2$ to T, diode 2 conducts and diode 1 is non-conducting, and hence

$$i_{A2} = -\frac{\hat{v}}{R + R_0}\sin\omega t, \qquad i_{A1} = 0,$$

$$v_R = -\frac{R\hat{v}}{R + R_0}\sin\omega t,$$

$$v_{A1} = \hat{v}\left(1 + \frac{R}{R + R_0}\right)\sin\omega t$$

and
$$v_{A2} = -\frac{R_0\hat{v}}{R + R_0}\sin\omega t.$$

During this time interval $\sin\omega t$ is negative, and hence i_{A2} and v_R are positive. This may be verified readily from Fig. 16.7. Now v_R consists

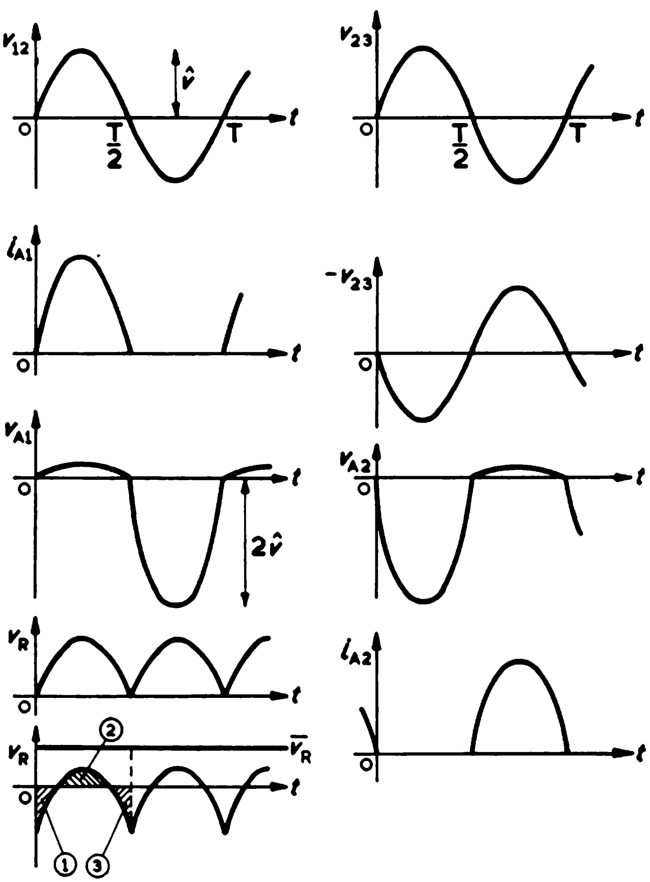


FIG. 16.7

of a unidirectional voltage in half sine wave pulses, and the average value is

$$\bar{v}_R = 2R\hat{v}/\pi(R + R_0).$$

This mean voltage is double the value obtained from the half-wave rectifier, but, for the full-wave ciruit, the total transformer secondary voltage is also doubled. It may be seen that the frequency of the ripple is twice the supply frequency. The maximum voltage across each diode again occurs during the non-conducting half cycles. Usually $R_0 \ll R$, and hence the peak inverse voltage is nearly equal to $2\hat{v}$. With the same approximation $\bar{v}_R = 2\hat{v}/\pi$.

16.4. Choke-input Full-wave Rectifier

When an inductance is used in series with the load the rectifier behaviour is modified in certain respects. The circuit, known as a choke-

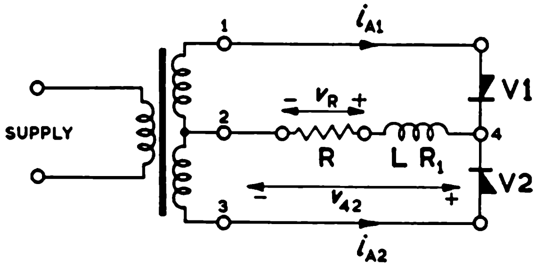


FIG. 16.8

input rectifier, is shown in Fig. 16.8. Provided R_0 is sufficiently low, the voltage between terminals 4 and 2 is a series of half sine waves of amplitude $\hat{v}$ as shown in Fig. 16.9.*a*. As before, v_{42} may be separated into its steady component $\bar{v}$ and a ripple component v (Fig. 16.9.*b*); again

$$\bar{v} \simeq 2\hat{v}/\pi.$$

The steady voltage causes a steady current $\bar{i}$ through the load R, where

$$\bar{i} = \bar{v}/(R + R_1)$$

and R_1 is the resistance of the choke. The d.c. voltage across the load is therefore $\bar{v}R/(R + R_1)$, and it is seen that R_1 should be small compared with R. The cyclic voltage v gives rise to a cyclic current i, where

$$v = (R + R_1)i + L\frac{di}{dt}.$$

If L is sufficiently large

$$v = L\frac{di}{dt}$$

and hence

$$i = \frac{1}{L}\int_0^t v dt,$$

i.e., i may be determined from the area under the curve of v against t. This is done in Fig. 16.9.*c* using Fig. 16.9.*b*. At the point B, i is zero and then increases in the negative direction to point C. Here v goes positive and i becomes less negative, until at D, i is zero and area 1 equals area 2. The current now goes positive, passes through a maximum value at E and

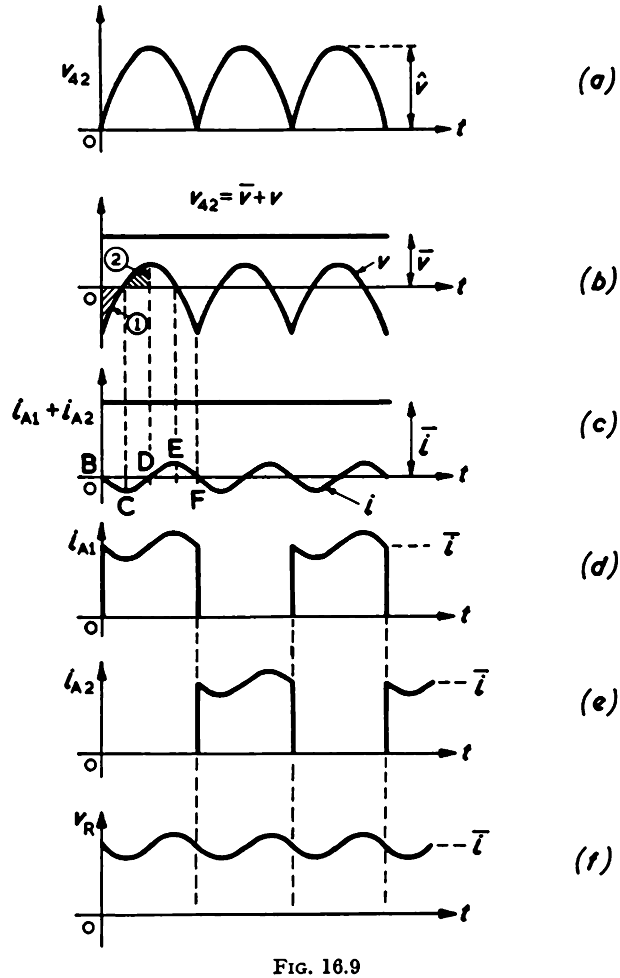


FIG. 16.9

drops to zero again at F. The cyclic current adds to the steady current $\bar{i}$, giving the total current, which is supplied in alternate half cycles by the two diodes as indicated in Fig. 16.9.*d* and *e*. The ripple voltage across the load is iR, and this may be made very small with large L. Thus one advantage of the choke-input rectifier is considerable reduction of ripple. The individual diode currents consist essentially of rectangular pulses of maximum height little in excess of the mean load current $\bar{i}$. This is another feature in favour of the choke-input circuit in comparison with some other rectifiers. For satisfactory operation of this circuit the choke must have high inductance but low resistance. Also it must have high inductance with some d.c. flowing in it. To meet these requirements the choke usually has an iron core but with an appreciable air gap.

16.5. A.c. Supply, Diode and Condenser in Series

The circuit for this case is shown in Fig. 16.10. It is assumed that $v_{12} = \hat{v} \sin \omega t$ and that the condenser C is uncharged at the time of switching on. When terminal 1 is positive with respect to terminal 2, current flows through the diode and charges up the condenser to voltage

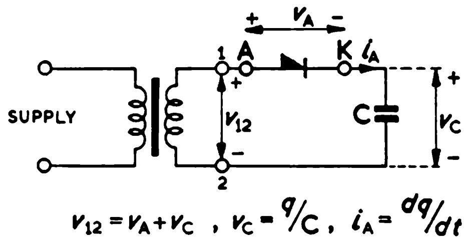


FIG. 16.10

v_C. During the negative half cycle no current flows and the condenser retains its charge, provided its insulation resistance is infinite. When the next positive half cycle occurs the diode anode is positive with respect to the cathode for only part of the time on account of the voltage v_C. The charging process increases v_C during each positive half cycle until the condenser is charged to a voltage equal to $\hat{v}$, when no further current flows. This process may be analysed as follows using Fig. 16.11. At all times

$$v_{12} = v_A + v_C = \hat{v} \sin \omega t.$$

At $t = 0$, v_A is just about to become positive and current i_A flows through the diode of amount

$$i_A = (\hat{v} \sin \omega t - v_C)/R_0 \text{ as long as } \hat{v} \sin \omega t > v_C.$$

While the current flows v_C increases, since the condenser charge

$$q = \int_0^t i_A dt \text{ and } v_C = q/C.$$

At time t_1, v_A is zero and $v_C = v_{C1}$, where $v_{C1} = \hat{v} \sin \omega t_1$. At t_1, i_A also is zero, and it remains zero as long as v_A is negative, i.e., until t_2 where $\omega t_2 = 3\pi - \omega t_1$. The condenser voltage remains constant at v_{C1} during the time t_1 to t_2. After t_2, v_A is again positive, diode current flows and the condenser charges to a voltage v_{C3} at time t_3, where $\hat{v} \sin \omega t_3 = v_{C3}$.

Current again ceases from t_3 to t_4, flows from t_4 to t_5 and so on. An equilibrium condition is reached in which:

(*a*) the condenser is charged to a steady voltage equal to $\hat{v}$ (this is the rectified or d.c. voltage, and C is called the reservoir condenser),

(*b*) the potential difference between anode and cathode of the diode is cyclic and equal to $-\hat{v}(1 - \sin \omega t)$, i.e., it varies from zero to $-2\hat{v}$, thus giving a peak inverse voltage of $2\hat{v}$, and

(*c*) there is no diode current at any time.

The equilibrium condition is approached more quickly with smaller C or smaller R_0. Maximum diode current occurs during the first cycle,

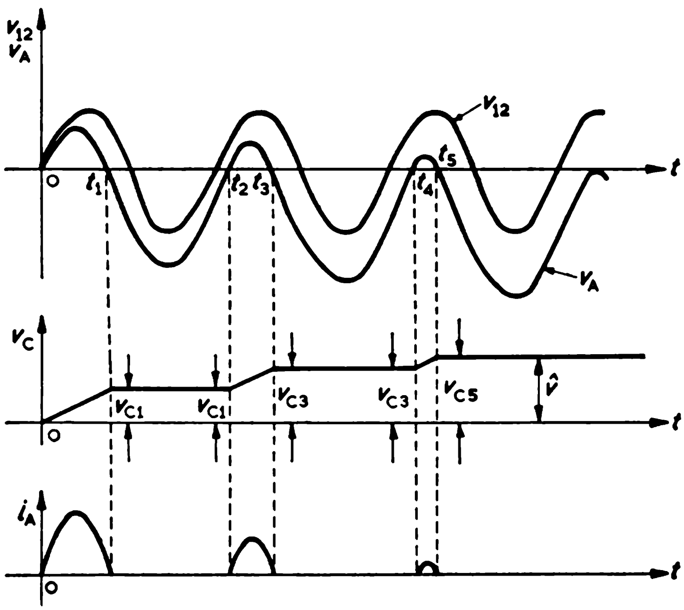


FIG. 16.11

and its value increases with the capacitance of the reservoir condenser. This circuit has applications for small rectifiers, voltage doubling and peak voltage measurement.

16.6. Condenser-input Full-wave Rectifier

The steady voltage across the reservoir condenser may be used for supplying d.c. to a load resistance by connecting the latter across C. Usually a full-wave circuit is used, and this involves a centre-tapped transformer and two diodes, as shown in Fig. 16.12. With this arrangement a current i_R flows in the load where $i_R = v_C/R$. The condenser is being discharged continuously by R, and it is being charged only during those periods when the anode–cathode voltage of either diode is positive.

In analysing this system it is assumed that the equilibrium condition has been reached. At time t_1 (Fig. 16.13) when v_{A1} becomes positive, $V1$ passes current and charges C. During the interval t_1 to t_2 the diode supplies current to both C and R. Then

$$i_{A1} = i_C + i_R, \text{ where } i_R = v_C/R \text{ and } i_C = C\frac{dv_C}{dt}.$$

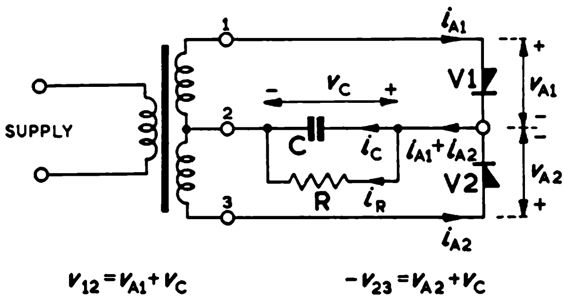


FIG. 16.12

From t_2 to t_3, $i_{A1} = i_{A2} = 0$ and $i_C = -i_R$. The condenser is now being discharged. If at time t_2, $v_C = v_2$, then at time τ after t_2,

$$v_C = v_2\varepsilon^{-\tau/RC}.$$

If τ/RC is small

$$v_C \simeq v_2(1 - \tau/RC).$$

At time t_3,
$$v_C = v_3 \simeq v_2\left(1 - \frac{t_3 - t_2}{RC}\right).$$

The greatest possible value of $t_3 - t_2$ is $1/2f$, where f is the supply frequency, hence the above approximation is justified if $RC \gg 1/2f$. The waveforms of the various currents and voltages are shown in Fig. 16.13. It is seen that there is a ripple current which can be reduced by increasing the capacitance of the reservoir condenser. The average voltage across the load is $\bar{v}_C$ and the average load current is $\bar{v}_C/R$. The recharging of the condenser by the two diodes takes place during the intervals of time equal to $t_2 - t_1$. The shorter the value of $t_2 - t_1$, the nearer $\bar{v}_C$ is to $\hat{v}$, but the larger is the maximum current taken by each diode to produce the average load current. The value of $t_2 - t_1$ may be shortened by increasing CR. Thus for a given load resistance the greater the value of the reservoir condenser, the lower the ripple, but the greater the maximum diode current. The condenser-input rectifier may require a maximum diode current several times the mean load current. In this respect the choke-input rectifier has a distinct advantage, since its maximum diode current is approximately equal to the load current. On the other hand, the condenser-input circuit gives a rectified voltage equal to $\hat{v}$ from a given transformer, which is about 50 per cent more than the value $2\hat{v}/\pi$ obtained from the choke-input circuit.

The average rectified voltage with the condenser-input rectifier is given by

$$\bar{v}_C \simeq (v_2 + v_3)/2 = v_2\left(1 - \frac{t_3 - t_2}{2RC}\right).$$

As rough approximations $v_2 = \hat{v}$ and $t_3 - t_2 = \dfrac{1}{2f}$.

Hence $$\bar{v}_C \simeq \hat{v}\left(1 - \frac{1}{4fRC}\right).$$

Also $\hat{v}/R \simeq \bar{i}_R = \bar{i}_O$, the rectified current, and so

$$\bar{v}_O = \bar{v}_C \simeq \hat{v} - \frac{\bar{i}_O}{4fC}.$$

This shows that the rectified voltage decreases with the current to the load. This variation of the d.c. voltage with the load current is called the

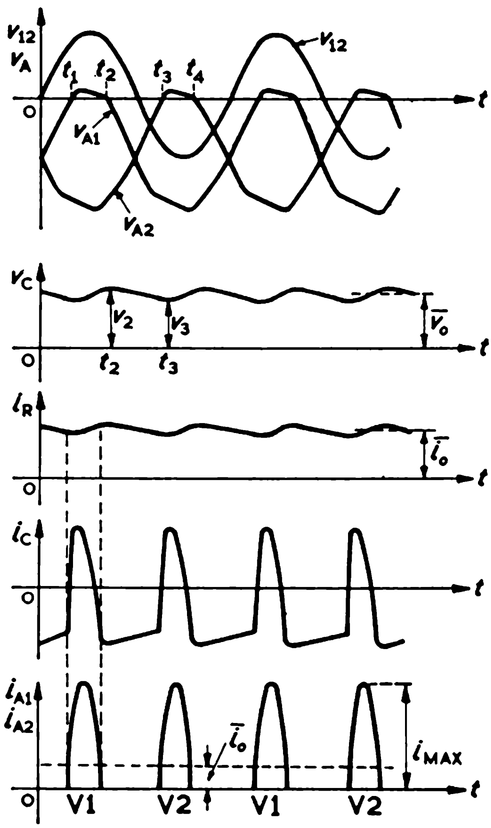


FIG. 16.13

regulation of the rectifier. In this case the regulation may be improved by increasing C. In practice, the regulation is also affected by voltage drop in the resistance of the transformer winding and in the diodes. The regulation of the choke-input filter depends on similar resistances, including the resistance of the choke. However, in that case there is no drop in voltage comparable to that due to the discharging of the reservoir condenser. As a result, the choke-input rectifier has better regulation. The condenser-input full-wave rectifier is used mainly for small power supplies of a few hundred volts and 200 mA or less. Larger supplies use the choke-input circuit.

16.7. Voltage-doubling Circuits

It is shown in Section 16.5 that the voltage across the diode in a diode-condenser rectifying circuit varies from 0 to $2\hat{v}$, with the greatest value

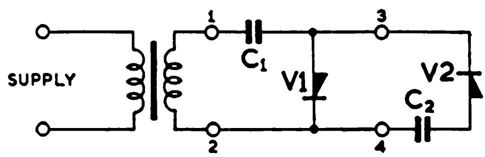


FIG. 16.14

occurring when the anode is negative with respect to the cathode. If a second diode and condenser are connected across the first diode with the correct polarity, as shown in Fig. 16.14, then this diode passes current until the condenser C_2 is charged to a steady voltage equal to $2\hat{v}$, thus giving double the rectified voltage of the single circuit. The peak inverse

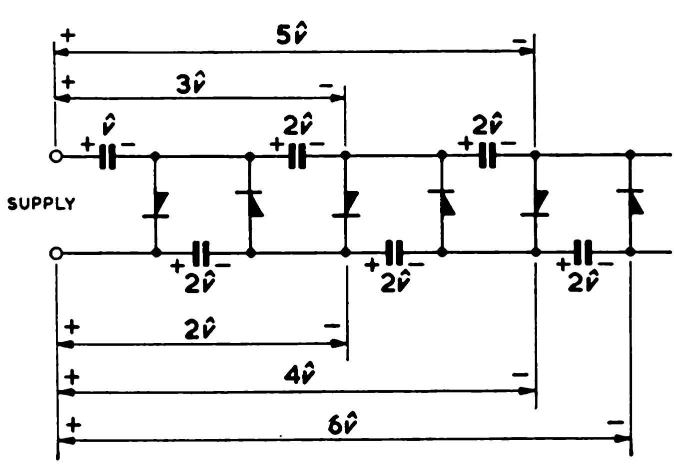


FIG. 16.15

voltage across the second diode is $2\hat{v}$, and occurs when the anode is negative with respect to the cathode. A third diode and condenser may be connected across the second diode, and this condenser is also charged to a voltage $2\hat{v}$. The process may be continued as often as required, as illus-

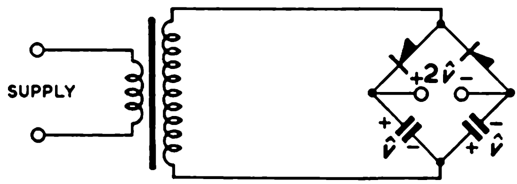


FIG. 16.16

trated in Fig. 16.15, where voltages up to $6\hat{v}$ can be obtained from six diodes and condensers.

An alternative voltage-doubling circuit is shown in Fig. 16.16. Here two half-wave rectifiers use a common transformer, and the voltage across the two reservoir condensers in series is $2\hat{v}$.

16.8. Filter Circuits

Rectifier circuits produce a current in the load which consists of a steady part and a superimposed ripple. The amount of ripple is usually expressed in terms of the ratio of the r.m.s. ripple current through the load

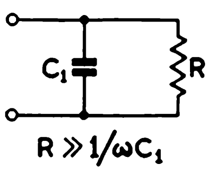


FIG. 16.17

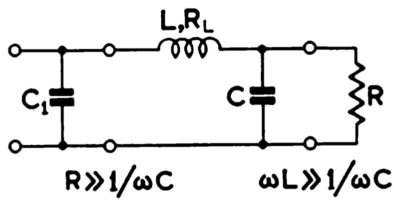


FIG. 16.18

to the mean current (or the r.m.s. ripple voltage across the load to the mean voltage).

In the choke-input rectifier the ripple current through the load may be reduced by connecting a condenser C_1 in parallel with it. The cyclic ripple current passes mainly through C_1, provided $\frac{1}{\omega C_1} \ll R$, where the ripple is assumed to be sinusoidal and of frequency $f = \omega/2\pi$. The ripple current through the load is reduced in the ratio $1/\omega C_1 R$ (see Fig. 16.17).

Further reduction in ripple may be achieved by the connection of a "filter circuit" between the rectifier output terminals and the load. The filter circuit consists of one or more inductors and capacitors. A single L, C filter is shown in Fig. 16.18, where $\omega L \gg 1/\omega C$ and $\frac{1}{\omega C} \ll R$.

In this circuit the condenser C_1 is either the parallel condenser used above in the choke-input circuit or else the reservoir condenser in the condenser-input circuit. Using the above conditions, the magnitude of the ripple voltage across the load is reduced to $1/\omega^2 LC$ of its value across C_1. Any resistance R_L associated with the inductor causes a reduction in steady voltage across the load. The higher the supply frequency, the smaller the values of L and C needed for filtering. Also, the filtering or smoothing of a full-wave rectifier is easier than a half-wave rectifier, since the ripple has twice the frequency in the former case.

16.9. Diode Peak Voltmeter

The series diode–condenser circuit is used widely, and particularly at high frequencies, as an a.c. voltmeter for reading the peak value of any applied voltage. The circuit in Fig. 16.19.*a* has an equilibrium state when the condenser is charged to a steady voltage v_C equal to $\hat{v}$, and the

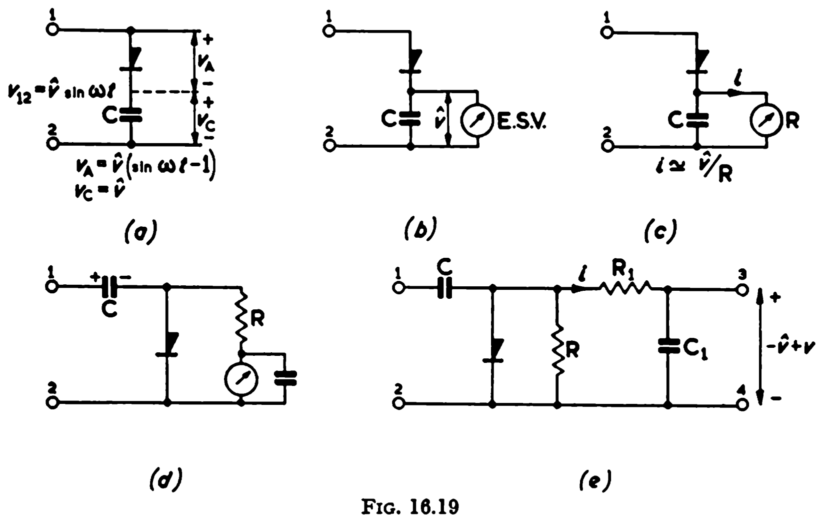


FIG. 16.19

measurement of the peak voltage reduces to the measurement of a steady voltage. Thus, if an electrostatic voltmeter is placed in parallel with C, as shown in Fig. 16.19.*b*, it reads the peak value of the applied voltage. As an electrostatic voltmeter has itself a capacitance, it can act as C. An alternative method of measuring the voltage across C is shown in Fig. 16.19.*c*, which uses a voltmeter of high resistance R. The current taken by this voltmeter must be negligible, and this means a large value of CR, as shown in Section 16.6. For this circuit to work the source of the a.c. voltage must have a d.c. path, otherwise there is no circuit for the voltmeter current. The d.c. path must also have a resistance much less than R to avoid reduction in the voltage indicated by the voltmeter. A circuit which avoids the need for the supply to have a d.c. path is shown in Fig. 16.19.*d*. It is based on the equilibrium voltage v_A across the diode. For a sinusoidal applied voltage this is shown in Section 16.5 to be given by $v_A = -\hat{v} + \hat{v}\sin\omega t$. The d.c. voltmeter, consisting of a high resistance and a galvanometer, is connected across the diode, and it measures the steady value of the voltage, irrespective of whether or not there is a d.c. path between terminals 1 and 2. To protect the galvanometer from the current arising from the cyclic component of v_A, it may be shunted by a condenser as shown. The form of this circuit which is usually used in practice is drawn in Fig. 16.19.*e*, where a filter R_1C_1 is inserted between the diode and the voltmeter. Across terminals 3 and 4 there is a steady voltage $\hat{v}$ and a cyclic voltage of magnitude $\hat{v}/\{1 + R_1{}^2\omega^2C_1{}^2\}^{1/2}$, which can be made very small. It might appear at first sight that the resistance R is not necessary in this circuit. However, even with terminals 1 and 2 shorted there is some diode current due to initial velocities of the electrons. This current requires a path back to the cathode, and it produces a voltage drop across the resistance of the path, thereby giving the diode a negative anode voltage for its quiescent condition. Without R the resistance of the path depends on leakages and is indefinite. The purpose of R is therefore to stabilize the quiescent voltage across the diode. The value of R is normally many megohms. The steady voltage across C_1 is usually measured with a d.c. amplifier; the actual peak value of the a.c. voltage is the difference between this steady voltage and the quiescent voltage.

Any good voltmeter should take a negligible current from the supply being measured. The impedance at high frequencies of the diode peak voltmeter of Fig. 16.19.*e* is determined almost entirely by the effective shunt capacitance C_i between terminals 1 and 2, and this should be as small as possible. It depends mainly on the capacitance of the diode and stray capacitance from the condenser C to terminal 2. When the voltmeter is used at very high frequencies the leads from the circuit under test to terminals 1 and 2 should be as short as possible. Any inductance in series with the input capacitance C_i across terminals 1 and 2 effectively changes the input impedance to $j(-1/\omega C_i + \omega L)$; the voltage across the capacitance C_i is greater than the voltage to be measured, as LC_i is a series circuit. At the resonant frequency of this circuit the voltmeter reading may be many times the applied voltage.

In using a diode voltmeter care must be taken when the magnitude of the applied voltage varies with time. The rate at which equilibrium is reached depends largely on the product CR. In order to measure rapid changes in amplitude CR must be small in comparison with the duration of the changes. However, CR must be large in comparison with the alternating period if the measured voltages are to approximate to the peak values. Thus there must be some compromise in the choice of C when a high-frequency voltage of varying amplitude is to be measured.

16.10. Some Practical Considerations in Rectifier Design

In the design of rectifiers there are several important practical considerations, some of which have already been mentioned. In the choice of a suitable diode there are two main factors: (*a*) the maximum current in the forward direction, and (*b*) the maximum peak inverse voltage. The maximum permissible values are specified by the valve maker. These factors become increasingly important in larger rectifiers. When large currents are required it is obviously desirable to have the mean rectified current nearly equal to the maximum allowed current. The choke-input rectifier is therefore preferred in such cases. The peak current demand on switching on also favours the choke-input circuit. When the current taken from a rectifier varies, then an important factor is the regulation, i.e., the change in the output voltage when the load current varies between its minimum and maximum values. The ratio of the change in output voltage to the corresponding change in load current is called the internal resistance of the supply. For good regulation the internal resistance should be low. In any rectifier the resistance of chokes, transformer windings and R_0, the diode forward resistance, all contribute to the internal resistance. Thermionic gas diodes, mercury-arc diodes and crystal diodes have very low values of R_0, and are therefore used in rectifiers when high currents are required. In addition to having stability of the output voltage against load changes, it is sometimes important to have stability of output voltage against changes in the mains voltage. Both types of stability may be achieved with special circuit arrangements, examples of which are described in the next two sections. When thermionic diodes are used in rectifiers care must be taken with the insulation of the windings of the heater transformers. This is particularly necessary in voltage-doubling circuits.

16.11. Voltage Stabilization—Gas Diode

A cold-cathode gas diode can maintain a discharge in which the voltage across the diode remains almost constant while the current varies over considerable limits. A typical characteristic is shown in Fig. 16.20.*a*. Over the current range from $i_{\min}$ to $i_{\max}$, the voltage across the diode varies from $v_{\min}$ to $v_{\max}$. The average value of v_A over the range is the maintenance voltage v_M; v_S is the striking voltage. This diode may be used to provide a stable voltage equal to v_M across a varying load R_L with the circuit shown in Fig. 16.20.*b*, in which E is the supply voltage. The arrangement provides stability against changes in the load over quite an appreciable range. The value of R is chosen so that i_A equals its maximum value $i_{\max}$, when R_L is infinite, i.e., zero load current. Then

$$v_A = E - Ri_{\max}.$$

But $$v_A \simeq v_M$$

and hence $$R = (E - v_M)/i_{\max}.$$

When there is some load current i_L, the diode current adjusts itself to give

$$i_A + i_L = i_{\max},$$

and the voltage across the load is still v_M approximately. Over the current range $i_{\min}$ to $i_{\max}$ the diode acts as a current reservoir; as the load current increases, the diode current decreases by a corresponding amount.

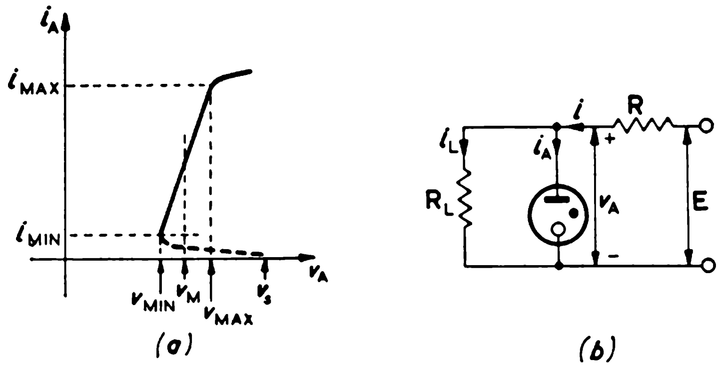


FIG. 16.20

This circuit also gives a stable voltage v_M across a fixed load R_L for an appreciable range of variation of supply voltage E.

If $$i_L = v_M/R_L,$$

then $i(= i_L + i_A)$ may vary from

$$i_L + i_{\min} \text{ to } i_L + i_{\max}.$$

The supply voltage may therefore vary from

$$v_M + R(i_L + i_{\min}) \text{ to } v_M + R(i_L + i_{\max}),$$

and the load voltage remains nearly constant at v_M.

In all uses of the circuit shown in Fig. 16.20.*b* it is essential to ensure that the discharge strikes initially. This means that $\dfrac{R_L E}{R + R_L}$ must be greater than v_S when the circuit is switched on.

Gas diodes with values of v_M of 50 V and upwards are available, but one of the disadvantages of this type of stabilizer is that the voltage cannot be varied once a particular diode is chosen. Another limitation is that the maximum load current is determined by the diode. An alternative stabilizer, which gives greater freedom in these respects and at the same time gives greater stability, is described in Section 16.12. Crystal diodes are available with characteristics which make them suitable as stabilizers for low voltages of the order of 5 to 10 V.

16.12. Voltage Stabilization—Feedback

By using the principle of negative feedback a power supply can be constructed with low internal resistance. One suitable circuit is shown in Fig. 16.21. The load R_L is fed from the supply E through a triode $V1$ in a cathode follower circuit. Part of the output voltage, across the resistance R_2, is connected to the grid of a pentode amplifier. The actual grid–cathode voltage of the pentode $V2$ is the difference between the voltage across R_2 and the voltage across a gas diode $V3$, which acts as a diode stabilizer. The output of the pentode amplifier is direct-coupled to the grid of the triode.

The operation of this circuit is somewhat as follows. If, for any reason,

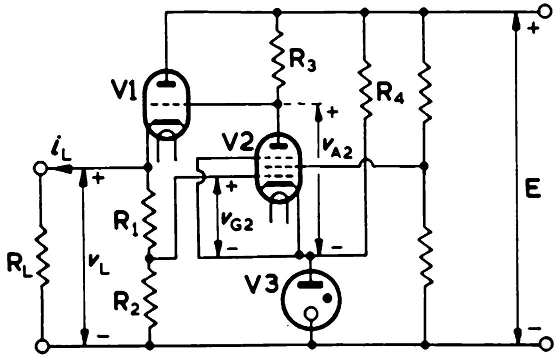


FIG. 16.21

the value of v_L tends to rise, then the voltage across R_2 rises. The reference voltage across the diode $V3$ remains constant over a range of variation of current through it. Hence the grid–cathode voltage of $V2$ becomes less negative, its anode current increases and its anode voltage decreases. This means that the grid voltage of $V1$ becomes more negative and the triode passes less current, thus offsetting to some extent the original increase in v_L. In effect, the pentode provides amplified negative feedback between the input and output circuits of the triode cathode follower, thus keeping the output voltage nearly constant. Suppose that the supply voltage changes by amount e and gives rise to changes v_{g2} and v_{a2} in the grid and anode voltages of $V2$. Then

$$v_{a2} = e - g_{m2}v_{g2}R_3.$$

This is the input voltage to $V1$ and,

provided $$R_L g_{m1} \gg 1,$$

the output voltage of the cathode follower equals the input voltage, and

hence $$v_l = e - g_{m2}v_{g2}R_3.$$

But $$v_{g2} = R_2 v_l/(R_1 + R_2) \text{ and}$$

so $$v_l = e\bigg/\left(1 + g_{m2}R_3 \cdot \frac{R_2}{R_1 + R_2}\right).$$

The factor $g_{m2}R_3$ is the gain of the pentode amplifier, and the fraction $R_2/(R_1 + R_2)$ need not differ greatly from unity, so that

$$v_l = e\bigg/\left(g_{m2}R_3 \cdot \frac{R_2}{R_1 + R_2}\right)$$

gives a measure of the stability of this circuit against changes in the supply voltage. It has been assumed above that the voltage across the gas diode remained constant. Actually, the voltage does change by a small amount in the same direction as the change across the load, so that the control is reduced somewhat. Where very high stability is required the gas diode is sometimes replaced by an h.t. battery. Alternatively, some improvement may be obtained if the resistance R_4 to the diode is fed from the stabilized side of the supply. The magnitude of the output voltage may be varied by adjusting the value of $R_2/(R_1 + R_2)$. When this fraction equals unity the output voltage approximately equals the maintenance voltage of the diode. When the fraction is decreased the output voltage rises, but for small values of the fraction the stability is reduced.

MODULATION AND DETECTION

17.1. Modulation

If the output of an oscillator is connected to an aerial, some of the output is radiated into space as an electromagnetic wave. A small part of the radiation may be intercepted by a second aerial which is connected to a receiver. In order to convey information over this communication system some characteristic of the original oscillation must be varied in time in accordance with the information. The sinusoidal output of the oscillator may be varied in amplitude, giving amplitude modulation; alternatively, the phase angle may be varied, in which case there is either frequency or phase modulation. There are also several ways of using pulses, which are varied by the information, and which themselves control the output of a sinusoidal oscillator, thus producing pulse modulation.

Modulation techniques are also used in other applications of electronics, such as instrumentation.

17.2. Amplitude Modulation

The main sinusoidal oscillation is known as the carrier. In amplitude modulation the amplitude of the carrier is determined by the instantaneous value of the information, or modulation.

If

$$e_c = \hat{e}_c \sin \omega_c t$$

represents the carrier voltage wave and the modulating signal is given by

$$e_m = \hat{e}_m f(t),$$

then the modulated carrier is given by

$$e = \hat{e}_c \left\{ 1 + \frac{\hat{e}_m}{\hat{e}_c} f(t) \right\} \sin \omega_c t.$$

In the case when the modulating signal is a single sine wave, $e_m = \hat{e}_m \cos \omega_m t$, the modulated carrier wave is

$$e = \hat{e}_c\{1 + m \cos \omega_m t\} \sin \omega_c t,$$

where $m = \hat{e}_m/\hat{e}_c$. This quantity m is called the modulation factor or the depth of modulation, and it should not exceed unity, otherwise distortion is introduced into the system (see Fig. 17.1). Both experiment and analysis show that the modulated wave may be represented by three separate sinusoidal components, the carrier and two side bands of angular frequency $\omega_c - \omega_m$ and $\omega_c + \omega_m$. Since

$$e = \hat{e}_c\{1 + m \cos \omega_m t\} \sin \omega_c t$$

then
$$e = \hat{e}_c \sin \omega_c t + \frac{m\hat{e}_c}{2} \sin (\omega_c + \omega_m)t + \frac{m\hat{e}_c}{2} \sin (\omega_c - \omega_m)t.$$

If this modulated voltage is produced across a resistance whose value is the same at the carrier and the side-band frequencies, the power in each

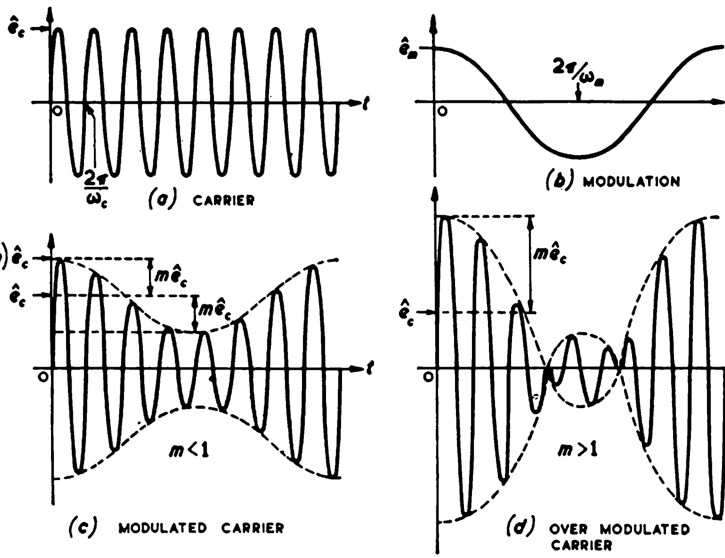


FIG. 17.1

side band is $m^2/4$ of the carrier power, with a maximum value of 1/4 when m is unity.

For a more complicated modulating signal the modulated wave is

$$e = \hat{e}_c\{1 + mf(t)\} \sin \omega_c t$$

and $f(t)$ is defined so that it has a maximum value of unity and m can have any value up to unity. For distortionless amplitude modulation the envelope of the modulated wave has the same wave shape as the modulating signal as shown in Fig. 17.1.b and c.

17.3. Circuits for Amplitude Modulation

It is not usual to attempt to modulate the amplitude of a self-oscillator, as this produces some unwanted frequency modulation as well. Amplitude modulation is effected in the power amplifier stages following the oscillator. At low power levels amplitude modulation may be carried out using a device with a non-linear current–voltage characteristic such as a diode. The carrier and modulating voltages are connected in series with the diode as shown in Fig. 17.2.*a*. The dynamic characteristic of the diode can be expressed as a power series

$$i_a = av + bv^2 + cv^3 + .$$

where the applied voltage is given by

$$v = \hat{e}_c \sin \omega_c t + \hat{e}_m \cos \omega_m t.$$

These equations are similar to those used in Section 8.7, where intermodulation in a triode is discussed. If we consider only the first two terms in the power series we find that the current contains components of angular

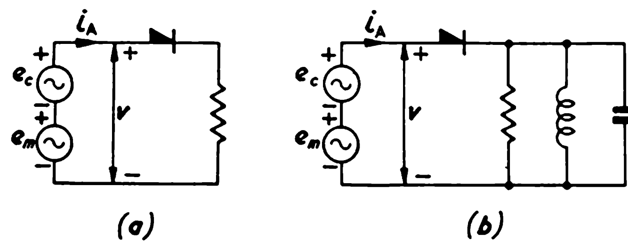


FIG. 17.2

frequency ω_c, ω_m, $2\omega_c$, $2\omega_m$, $\omega_c - \omega_m$ and$\omega_c + \omega_m$. The components in ω_c, $\omega_c - \omega_m$ and $\omega_c + \omega_m$ represent the carrier and side bands. The output voltage of these components is proportional to

$$a\hat{e}_c \sin \omega_c t + b\hat{e}_c\hat{e}_m \sin (\omega_c + \omega_m)t + b\hat{e}_c\hat{e}_m \sin (\omega_c - \omega_m)t.$$

The depth of modulation is given by

$$\frac{m}{2} = \frac{b\hat{e}_c\hat{e}_m}{a\hat{e}_c}, \text{ i.e., } m = \frac{2b\hat{e}_m}{a}.$$

If the cubic term of the power series is included the output current also includes components of angular frequency, $\omega_c + 2\omega_m$ and $\omega_c - 2\omega_m$. These components introduce unwanted side bands whose amplitudes are proportional to $\hat{e}_m^2$. Since m is proportional to $\hat{e}_m$, it is important to keep the modulation depth small to avoid distortion. In addition to the carrier and side bands the diode current contains components at frequencies ω_m, $2\omega_m$, $3\omega_m$, $2\omega_c$, $3\omega_c$, etc. In practice the diode load is arranged to have appreciable impedance only over the range covering the carrier and side bands. One way of achieving this is shown in Fig. 17.2.*b*, where the parallel resonant circuit is tuned to the carrier frequency.

Amplitude modulation may also be produced by connecting the carrier and modulation voltages in series with the grid and cathode of a triode or other amplifier and operating over a curved region of the dynamic grid characteristic.

In a pentode, amplitude modulation can be achieved by applying the carrier voltage to the control grid and the modulating voltage to the suppressor grid, as shown in Fig. 17.3. In this case the dynamic characteristic may be represented approximately by the expression

$$i_a = a(v_{g1} + kv_{g3}) + b(v_{g1} + kv_{g3})^2$$

and the square term gives rise to side bands as before.

In high-power transmitters it is more usual to modulate the anode of the final stage of a power amplifier operating under Class C conditions.

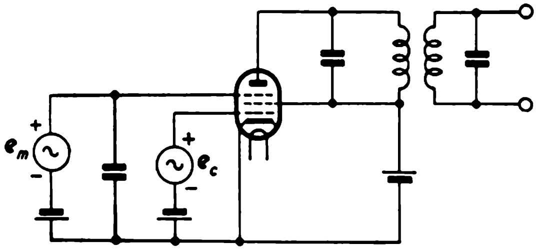


FIG. 17.3

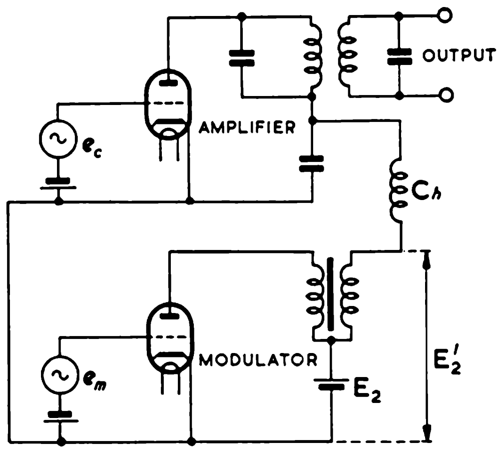


FIG. 17.4

The circuit is shown in Fig. 17.4. In a Class C amplifier the amplitude of the sinusoidal component of the anode voltage is approximately proportional to the quiescent value of the anode voltage, i.e., to E_2'. Thus, if the latter is varied at the modulation frequency so that

$$E_2' = E_2(1 + m \cos \omega_m t),$$

the alternating component of the anode voltage is

$$e_c = \hat{e}_c(1 + m \cos \omega_m t) \sin \omega_c t.$$

Negative feedback can be used to reduce distortion in the modulation process. Some of the modulated carrier is demodulated in a linear detector (see Section 17.7) and the output is fed to the input of the modulating amplifier. The action is similar to that described in Section 10.6 for the reduction of harmonic distortion.

17.4. Frequency Modulation

When a carrier wave is frequency modulated the instantaneous deviation of the frequency from the carrier is proportional to the instantaneous value of the modulating signal, so that

$$\omega_i = \omega_c + A\hat{e}_m \cos \omega_m t.$$

The high-frequency wave completes one cycle of frequency variation during one cycle of the modulation, as shown in Fig. 17.5.*a* and *b*.

In the unmodulated carrier, $\hat{e}_c \sin \omega_c t$, the total phase angle is θ, where

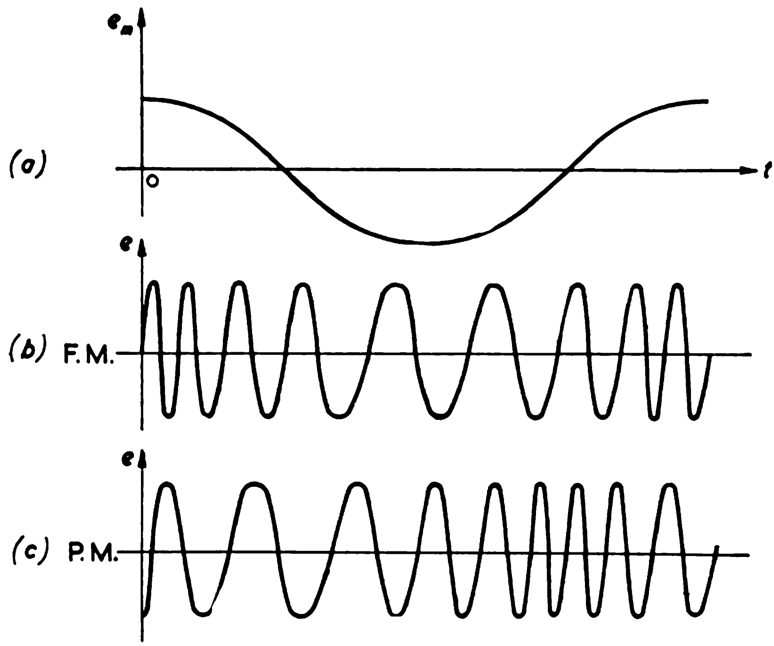


FIG. 17.5

$\theta = \omega_c t$. The angular frequency ω_c is seen to be the time rate of change of θ. Similarly, the instantaneous frequency in the frequency-modulated wave is defined as the rate of change of the instantaneous phase angle θ,

where $$\theta = \omega_c t + \frac{A\hat{e}_m}{\omega_m} \sin \omega_m t.$$

Thus the frequency-modulated voltage wave is given by

$$e = \hat{e}_c \sin \left\{ \omega_c t + \frac{A\hat{e}_m}{\omega_m} \sin \omega_m t \right\}.$$

When a carrier wave is phase modulated the phase angle varies according to the modulation so that

$$e = \hat{e}_c \sin \{\omega_c t + B\hat{e}_m \cos \omega_m t\}.$$

Thus, in addition to the normal linear increase of the total phase angle with time, there is a varying component of the angle which is proportional to the modulating signal (see Fig. 17.5.*c*). The instantaneous angular frequency in this case is given by

$$\omega_i = \omega_c - B\omega_m \hat{e}_m \sin \omega_m t,$$

and the frequency deviation is proportional to the frequency of the modulating signal. This means that, in a phase-modulated wave, the carrier is frequency modulated by a signal whose amplitude varies proportionately with the frequency of the signal. Obviously, frequency and phase modulation are closely related.

17.5. Circuits for Frequency Modulation

An oscillator can be frequency modulated if it has, in parallel with its resonant LC circuit, a small susceptance whose magnitude varies with the modulating signal. If the susceptance is capacitive, $C_m(t)$, then the resonant frequency is

$$f_i = \frac{1}{2\pi\sqrt{L(C + C_m(t))}} = f_0 \Big/ \sqrt{1 + \frac{C_m(t)}{C}},$$

where $$f_0 = \frac{1}{2\pi\sqrt{LC}}.$$

When $$C_m(t) \ll C,$$

then $$f_i = f_0 \left\{1 - \frac{C_m(t)}{2C}\right\}.$$

Thus if $C_m(t)$ is proportional to the instantaneous value of the modulating signal the oscillator is frequency modulated. As long as the output of the oscillator remains constant over the range, then there is no amplitude modulation. A device giving a susceptance with the required properties for $C_m(t)$ is the reactance valve, shown in Fig. 17.6.*a*. The essential features of this circuit are a pentode valve with a capacitor C connected between its anode and control grid, and a resistor R between the control grid and cathode, as shown in Fig. 17.6.*b*. In the former diagram all the other capacitors have negligible reactance at high frequency. The choke Ch has a very high reactance and merely serves the purpose of feeding the d.c. supply to the anode of the pentode. The pentode is connected across the tuned circuit of the oscillator so that there is an alternating voltage v_a between the anode and the cathode. Provided $1/\omega C \gg R$, the voltage across R, i.e. v_g, leads the anode voltage by nearly 90° and its magnitude is $\hat{v}_a \omega CR$. The anode current of a pentode is nearly independent of the anode voltage and $i_a \simeq g_m v_g$. Thus the anode current has a magnitude $g_m \omega C R \hat{v}_a$ and it also leads the anode voltage by nearly 90°. The total current taken from the source of v_a is, therefore, of magnitude $(g_m \omega C R + \omega C)\hat{v}_a$ and the current leads v_a by 90°. The effective susceptance across v_a is $(g_m R + 1)\omega C$ and is equivalent to a capacitance $(g_m R + 1)C$. If $g_m R \gg 1$, then the reactance valve behaves as a capacitance of $g_m RC$ across the tuned circuit. In some pentodes there is almost a linear relationship between mutual conductance and grid voltage. Then if the

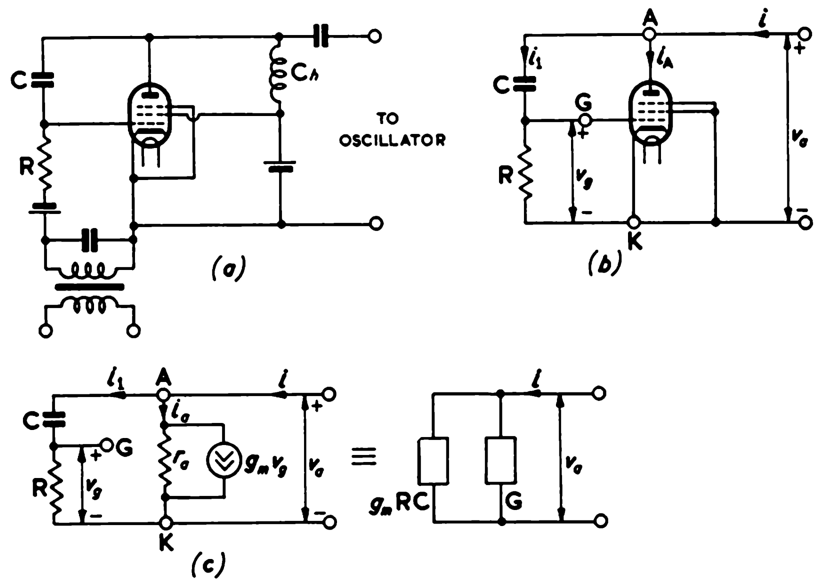


FIG. 17.6

grid voltage is varied at modulation frequency through a transformer, as shown in Fig. 17.6.*a*, the capacitance $g_m RC$ is proportional to the modulating signal, and this is the required condition for frequency modulation. Using the equivalent circuit of Fig. 17.6.*c* and the conditions $\omega CR \ll 1$ and $g_m R \gg 1$, it can be confirmed that there is an effective capacitance $g_m RC$ across v_a. It can also be shown that there is a parallel conductance G_m given by

$$G_m \simeq \omega^2 C^2 R^2 g_m + \frac{1}{r_a}.$$

The conductance is small and is usually neglected. However, as it varies with g_m and frequency, it may cause some amplitude modulation.

Some alternative reactance-valve circuits are given in Exx. XVII.

17.6. Detection or Demodulation

The useful information in a modulated carrier wave is contained in the modulation. The extraction of the information is achieved in the process known as detection or demodulation. In the case of amplitude modulation the envelope of the modulated wave has to be extracted, whereas with frequency modulation a voltage has to be produced proportional to the instantaneous frequency deviation from the carrier frequency. Before detection, the modulated wave is usually amplified by means of a tuned high-frequency amplifier. In the detection process use is made of the non-linear characteristics of diodes, triodes or other electronic devices. For small signals detection may be dealt with in terms of the curvature of the characteristics. If the received signal is amplified sufficiently before detection the electron device is best represented as a switch. In this case the principle of detection is very similar to that of rectification, which is discussed in Chapter 16.

17.7. Detection of Amplitude-modulated Waves

A thermionic or a crystal diode is frequently employed as the non-linear device. The modulated carrier voltage

$$e = \hat{e}_c(1 + m \cos \omega_m t) \sin \omega_c t$$

is applied in series with the diode and a load resistance R, as shown in

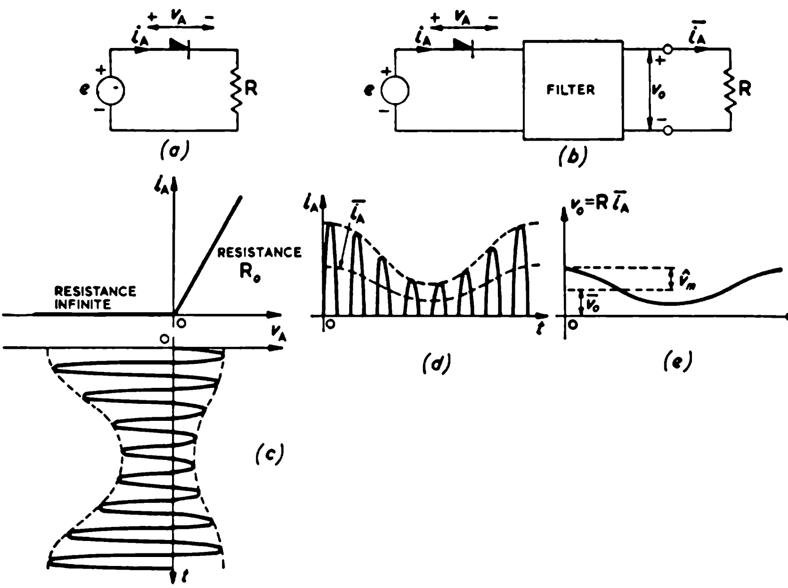


FIG. 17.7

Fig. 17.7.*a*. The diode characteristic is assumed to be ideal, and then the average value of the diode current over one cycle at the carrier frequency is found by the method used in Section 16.2. It is found that

$$\bar{i}_A = \frac{\hat{e}_c(1 + m \cos \omega_m t)}{\pi(R + R_0)}.$$

It is assumed here that $\omega_m \ll \omega_c$, so that $\cos \omega_m t$ does not change appreciably during one cycle of the carrier frequency. In addition to the mean anode current there is also the carrier component (see Fig. 17.7.*c* and *d*). This component can be removed by means of a suitable filter circuit, shown in Fig. 17.7.*b*. Then the output voltage consists of a

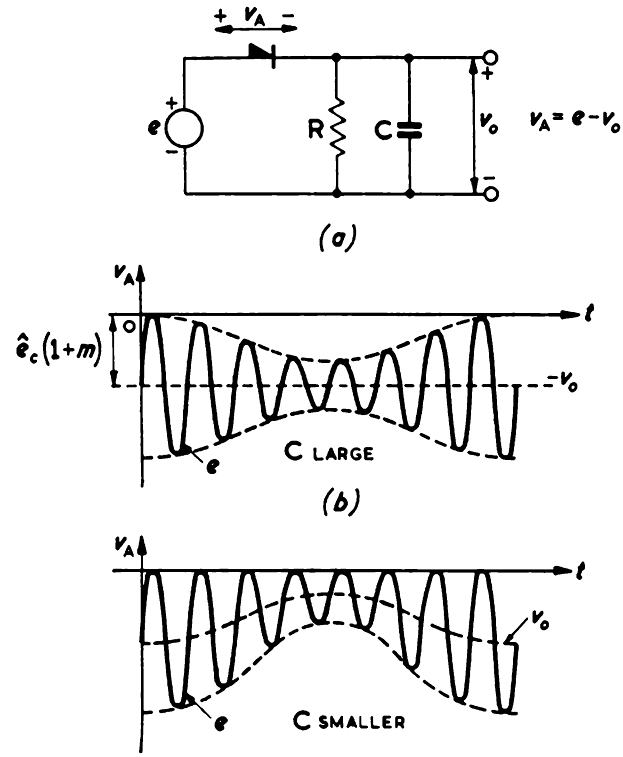


FIG. 17.8

steady part proportional to the carrier amplitude and a part, $\hat{v}_m \cos \omega_m t$, proportional to the original modulating signal, as shown in Fig. 17.7.*e*. Since the amplitude $\hat{v}_m$ varies as m, the modulation depth, this is called linear detection. The linearity, of course, depends essentially on the assumed ideal characteristics of the diode.

The detection efficiency η is defined as $\hat{v}_m/m\hat{e}_c$ and is given by

$$\eta = R/\pi(R + R_0).$$

When $R \gg R_0$ the efficiency has a maximum value of $1/\pi$. The maximum efficiency can be increased by connecting a capacitor C across the load resistance, as illustrated in Fig. 17.8.*a*. If this condenser is made very large the output of the detector is a steady voltage and, as shown in Section 16.5, it is equal to the maximum value reached by the applied voltage, i.e., $(1 + m)\hat{e}_c$ (see Fig. 17.8.*b*). However, if the capacitance is reduced in value the detector output is able to follow changes at modulation frequencies (Fig. 17.8.*c*.). In order to give improved efficiency the time constant CR must be large compared with the period of the carrier,

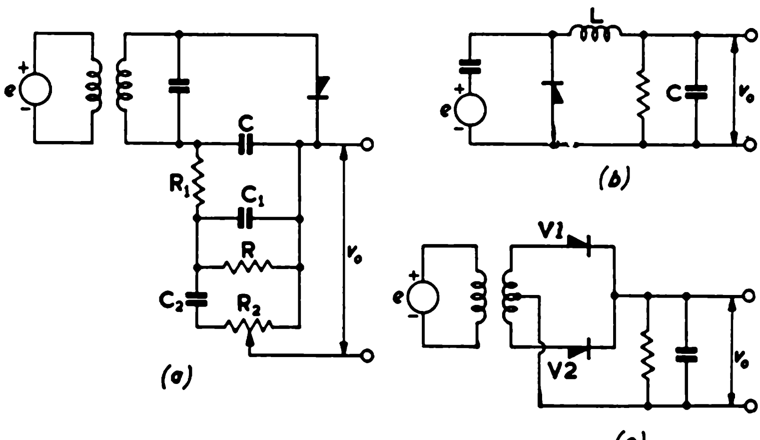


FIG. 17.9

but to follow the modulation envelope CR must be small compared with the period of the highest modulation frequency.

Several linear detector circuits are shown in Fig. 17.9. The first, Fig. 17.9.*a*, has a filter R_1, C_1, and a variable resistor R_2 to control the size of the output; C_2 provides d.c. separation between R_2 and the rectifier circuit. Fig. 17.9.*b* shows parallel connection of a diode detector with L, C filtering. Finally, Fig. 17.9.*c* is an example of a push-pull detector.

The assumption of the ideal characteristic for the diode detector is justified only if the signal is large. For small signals the dynamic diode characteristic is represented by a power series

$$i_a = av + bv^2 +$$

where

$$v = \hat{e}_c(1 + m \cos \omega_m t) \sin \omega_c t.$$

The output voltage across the load resistance includes high-frequency components which may be filtered out as before. The square term in the power series gives rise to low-frequency components in the output voltage of value

$$v_m = bm\hat{e}_c^{\,2}R \cos \omega_m t + \frac{bm^2\hat{e}_c^{\,2}R}{4} \cos 2\omega_m t.$$

The first term is the required modulating signal, and the second represents second-harmonic distortion. If other modulation frequencies are present there are also inter-modulation terms. Thus small signal detection is always accompanied by appreciable distortion, and normally this method is not used.

Detection can be achieved with triode or other valves by using the non-linear variation of anode current with grid voltage when the valve is biased near to cut-off. Alternatively, the non-linear grid current–grid voltage relation near to zero grid bias may be used.

17.8. Receivers

For satisfactory radio reception the signal applied to a linear detector must be of the order of a few volts. The received radio signal is not likely to be more than a few millivolts, and may be much less. Considerable

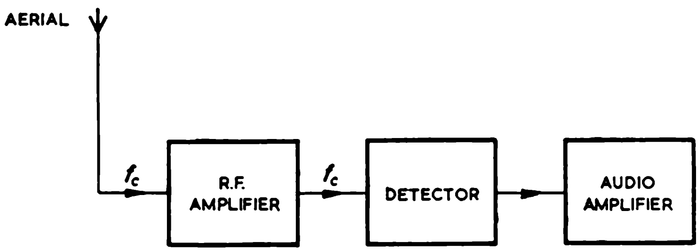


FIG. 17.10

amplification is therefore necessary. The signal covers a band of frequencies which includes the carrier and the side bands. To avoid frequency distortion, the whole of this bandwidth should be amplified uniformly. At the same time it is necessary to discriminate against signals arising from other transmitters whose carrier frequencies are adjacent to the wanted signal. For example, in a sound broadcasting system with amplitude modulation, the side bands may cover a range on either side of the carrier of about 5,000 c/s. The carrier frequencies of the transmitters are separated by 9,000 c/s. Thus the receiver must satisfy fairly stringent requirements of bandwidth and frequency selection. Also it must have considerable voltage amplification prior to the detector. Added to all this is the requirement that the receiver must be capable of selecting any one of many signals of widely differing radio-frequencies. One simple type of receiver is shown schematically in Fig. 17.10. This includes a high-frequency amplifier, a detector and an audio amplifier. With this kind of receiver it is virtually impossible to satisfy simply the above requirements of high-frequency amplification, bandwidth and selectivity over a wide tuning range.

17.9. Superheterodyne Reception

The difficulty of maintaining adequate and constant performance with wide tuning is overcome to a large extent in the superheterodyne receiver, whose essential features are represented in the diagram in Fig. 17.11. The main tuning element of the receiver is in the local oscillator, whose frequency is adjusted so that it differs from the signal frequency by an amount equal to the intermediate frequency. The signal voltage and local oscillator voltage are both applied to a "mixer" or frequency changer, which is a non-linear device producing output current components whose frequencies are the sum and difference of the signal and local oscillator frequencies. The output circuit of the mixer selects the

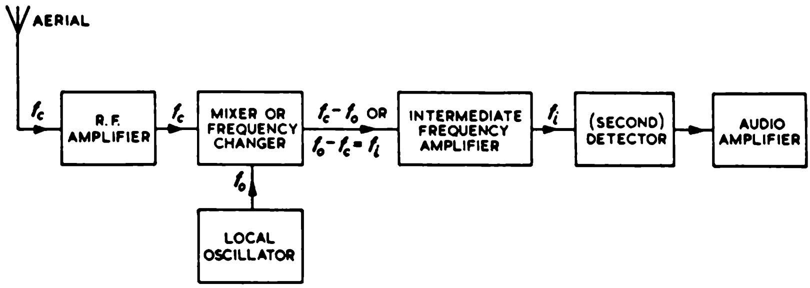


FIG. 17.11

difference frequency, which is also the intermediate frequency. The selected output contains all the components of the modulated wave, with the intermediate frequency replacing the carrier frequency. The signal is now amplified at intermediate frequency, and the modulation is extracted by using a linear detector, as before. When a different carrier frequency is to be received the local oscillator is tuned so that the difference between its frequency and the wanted carrier is again the intermediate frequency. Thus gain, selectivity and bandwidth are all achieved almost entirely at the fixed intermediate frequency. In some super-

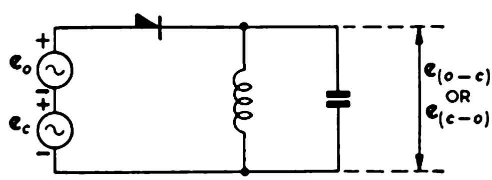


FIG. 17.12

heterodyne receivers there is a single-stage high-frequency amplifier before the mixer to give additional selectivity.

Crystal or thermionic diodes are frequently used as the non-linear devices in the mixer stage. The signal voltage and local oscillator voltage are connected in series with the diode and the load circuit, which is tuned to an angular frequency of either $\omega_0 - \omega_c$ or $\omega_c - \omega_0$, as shown in Fig. 17.12. The oscillator voltage is usually much greater than the signal voltage. The operation of this circuit is similar to the circuit of Fig. 17.2. The diode current contains components at frequencies ω_0, $2\omega_0$, ω_c, $2\omega_c$, $\omega_c - \omega_0$, $\omega_c + \omega_0$, etc. The impedance of the load is negligible at all angular frequencies except $\omega_c - \omega_0$, and hence the output voltage of the mixer is at this frequency. Only the carrier frequency has been considered in the above analysis. However, the side-band components are treated similarly, and the tuned load circuit has a bandwidth sufficient to accept the side bands. Thus the modulation of the carrier wave is transferred unchanged to the intermediate-frequency wave.

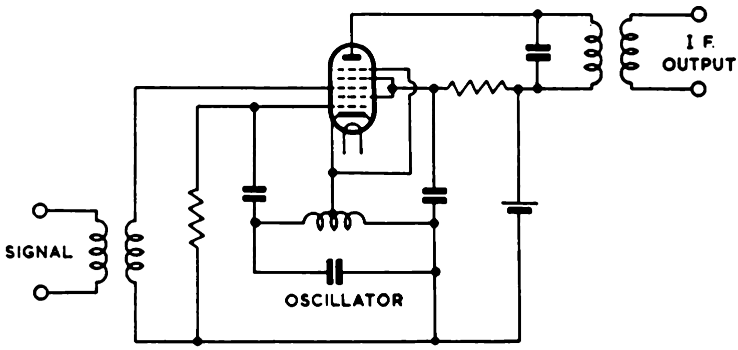


FIG. 17.13

Mixing may also be achieved with specially designed multi-grid valves, much in the same way as the pentode is used for producing amplitude modulation, as described in Section 17.3. One form of multi-grid mixer is shown in Fig. 17.13. The valve has five grids, and it serves the triple function of mixer, local oscillator and first intermediate-frequency amplifier. The cathode and first two grids act as a triode Hartley oscillator. The signal is connected to the third grid. The fourth and fifth grids act like the screen and suppressor grids of a pentode amplifier, and the anode circuit is tuned to the intermediate frequency. The oscillator operates in the Class C condition so that the anode current consists of a series of pulses at angular frequency ω_0. When the small signal e_c is applied to the third grid the anode current varies linearly according to the relation $i = g_m e_c$. However, owing to the action of the oscillator grid, g_m varies at oscillator frequency and may be represented by the series

$$g_m = g_0 + g_1 \sin \omega_0 t + g_2 \sin 2\omega_0 t +$$

The conditions are similar to those for the diode mixer and output is obtained at the intermediate frequency.

17.10. Automatic Volume Control

In a communication system it is sometimes found that the received signals vary slowly in magnitude due to changes in the propagation properties of the medium between the transmitter and the receiver. These slow variations are known as fading, and they are detected in the receiver along with the wanted modulations. The fading usually occurs

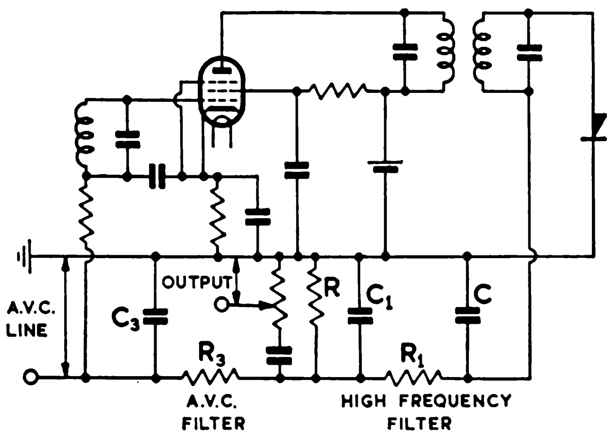


FIG. 17.14

at a rate much slower than the lowest modulation frequencies, and it is possible to introduce suitable compensation with a circuit arrangement of the type shown in Fig. 17.14. The detector circuit is the same as the one shown in Fig. 17.9.*a*, but it has an additional filter circuit C_3, R_3 which has a time constant of about 0·1 sec. Thus across C_3 there is a voltage which remains steady except for the slow variations due to fading. This voltage can be used as a grid-bias voltage for the preceding amplifying valves, varying the gain in such a way that the effect of fading is reduced considerably. Special amplifying valves are used in the automatic volume control circuit. By varying the pitch of the control-grid wires along the length of the system a valve is produced whose mutual conductance varies over a wider range of grid bias than in a normal pentode. These valves are known as variable-μ pentodes.

17.11. Detection of Frequency-modulated Waves

The detection of a frequency-modulated signal requires a circuit whose output is proportional to the instantaneous frequency difference between the modulated and unmodulated carrier. This can be achieved simply by adjusting a resonant circuit so that the carrier frequency is in the sloping part of the resonance curve, as shown in Fig. 17.15. Satisfactory operation requires the range of frequency deviation to be within the straight portion of the resonance curve. This is possible only for very small deviations, and this circuit is not often used.

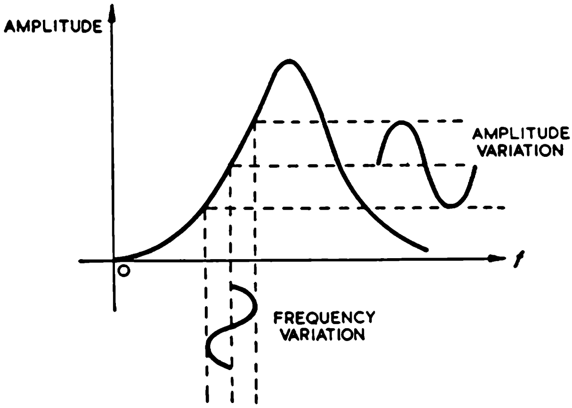


Fig. 17.15

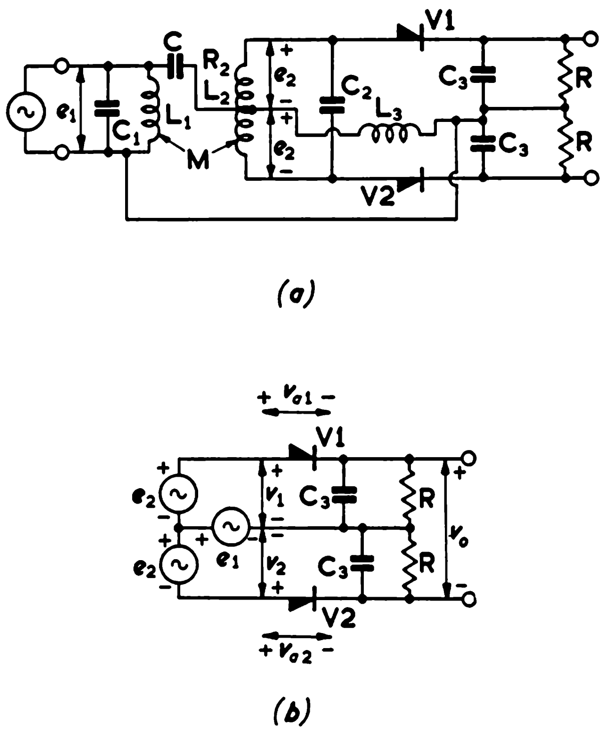


Fig. 17.16

A more useful form of detector is shown in Fig. 17.16.*a*, in which e_1, the frequency-modulated signal, is connected across the primary section of a tuned high-frequency transformer. The windings are tuned to resonate at the carrier frequency. The secondary inductance is centre-tapped so that equal voltages are developed across the two halves. The high potential end of the primary circuit is joined to the centre tap through a condenser C, whose reactance is small at high frequencies. The two diodes and their load circuits C_3, R are identical. The common point of the load circuits is joined to the centre tap on L_2 through an inductance L_3. The reactance of L_3 is large compared with the reactance of L_1; L_1 and L_3 are effectively in parallel so that the voltage across L_3 is e_1. For sinusoidal voltages it can be shown that

$$\frac{2\mathbf{E}_2}{\mathbf{E}_1} \simeq -\frac{j\omega M}{R_2 + j\left(\omega L_2 - \dfrac{1}{\omega C_2}\right)} \cdot \frac{C_1}{C_2},$$

where M is the mutual inductance between L_1 and L_2, and R_2 is the series resistance of the secondary winding. This assumes that the circulating current in the primary is large in comparison with the current taken from the generator e_1. At the carrier frequency $\omega_c L_2 = \dfrac{1}{\omega_c C_2}$ and $\mathbf{E}_2$ and $\mathbf{E}_1$ differ in phase by $\pi/2$. At a frequency $\omega_c + \delta\omega$ near to resonance it can be verified that the phase difference between $\mathbf{E}_2$ and $\mathbf{E}_1$ is $\pi/2 + \phi$, where $\tan\phi \simeq 2\delta\omega L_2/R_2$. When $\delta\omega$ is small ϕ is therefore proportional to the difference between the actual frequency of the signal and the carrier frequency.

For determining the voltages applied to the diodes the circuit can be redrawn as shown in Fig. 17.16.*b*. Then

$$\mathbf{V}_1 = \mathbf{E}_2 + \mathbf{E}_1 \quad \text{and} \quad \mathbf{V}_2 = -\mathbf{E}_2 + \mathbf{E}_1.$$

These voltages are rectified separately by the diodes operating as linear detectors. The voltage v_0 across the load is proportional to the difference between the outputs of the diodes, i.e.,

$$v_0 = K\{|\mathbf{E}_2 + \mathbf{E}_1| - |-\mathbf{E}_2 + \mathbf{E}_1|\}.$$

If $$e_1 = \hat{e}_1 \sin \omega t \text{ then } e_2 = \hat{e}_2 \sin\left(\omega t + \frac{\pi}{2} + \phi\right).$$

The numbers of turns on the windings are chosen so that $\hat{e}_1 = \hat{e}_2$ and then

$$|\mathbf{E}_2 + \mathbf{E}_1| = 2\hat{e}_1 \cos\left(\frac{\pi}{4} + \frac{\phi}{2}\right)$$

and $$|-\mathbf{E}_2 + \mathbf{E}_1| = 2\hat{e}_1 \sin\left(\frac{\pi}{4} + \frac{\phi}{2}\right).$$

Hence we find that

$$v_0 = \frac{4K\hat{e}}{\sqrt{2}} \sin\frac{\phi}{2} \simeq K\hat{e}\phi\sqrt{2}.$$

As ϕ is proportional to $\delta\omega$, then the output voltage is proportional to the instantaneous difference between the modulated frequency and the carrier frequency, which is the required condition. Since the output of this

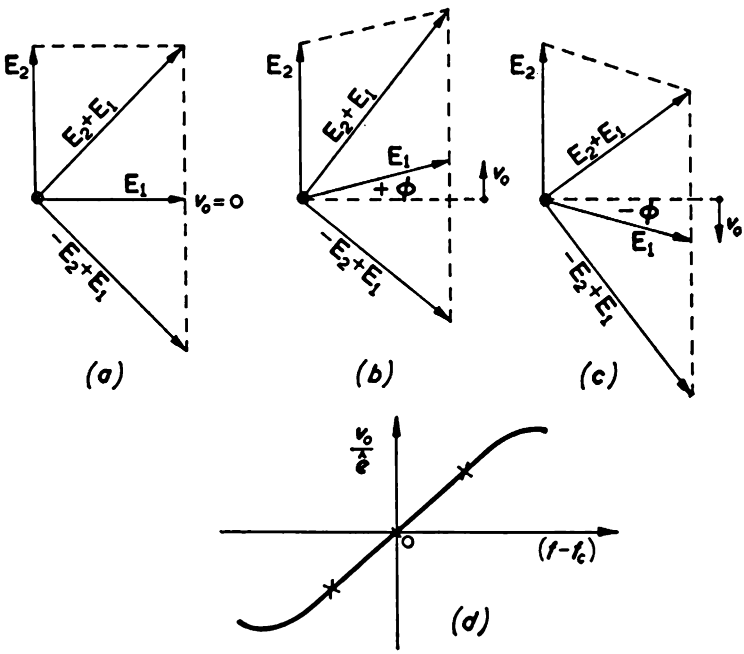


FIG. 17.17

circuit is proportional to the phase angle ϕ, it is called a phase discriminator. As it gives an output proportional to the frequency difference $\delta\omega$, it is also a frequency discriminator. The conditions in the circuit are shown diagrammatically in Fig. 17.17.

The output of the discriminator is proportional to the signal amplitude as well as to the frequency deviation. Thus amplitude variations due to interference or noise may produce unwanted output. It is usual to eliminate these amplitude variations before the discriminator by means of an amplitude limiter, which consists of a cut-off triode clipping circuit, similar to the one described in Section 19.2. This ensures constant amplitude without affecting frequency variations. Thus any amplitude modulated interference is removed.

17.12. Automatic Frequency Control

A frequency discriminator produces an output voltage proportional to the difference between the frequency of a signal and some fixed frequency. If the output voltage is used to control the bias of a reactance valve the frequency of an oscillator may be kept automatically at a value that gives, on the average, zero output at the detector.

CHAPTER 18

RELAXATION OSCILLATORS AND SWITCHES

18.1. Relaxation Oscillators

Various types of valve oscillator for producing sinusoidal voltages are discussed in Chapter 13. We now consider relaxation oscillators which produce non-sinusoidal waves. There are two distinctive features of these oscillators. Firstly, they depend on valves acting as voltage-sensitive switches, and secondly, the frequency of oscillation is determined by the time constants of the circuits. The switching action may cause a sudden change or relaxation of the conditions and is responsible for the non-sinusoidal waveform. The principle may be explained with reference to the circuit of Fig. 18.1.*a*, which shows a resistance and condenser in series with a battery. Across the condenser is a rather special type of switch S. This switch has the property of closing automatically when the condenser has charged up to some specific voltage v_S. The condenser then discharges rapidly down to some lower voltage v_M, when the switch automatically opens. The condenser charges again and the process is repeated, giving an alternating voltage across the condenser of the shape shown in Fig. 18.1.*b*. The period of the relaxation oscillation is T_0, which is determined primarily by the time T_1 taken for the condenser C to charge from a voltage v_M to v_S. This is given by the equation

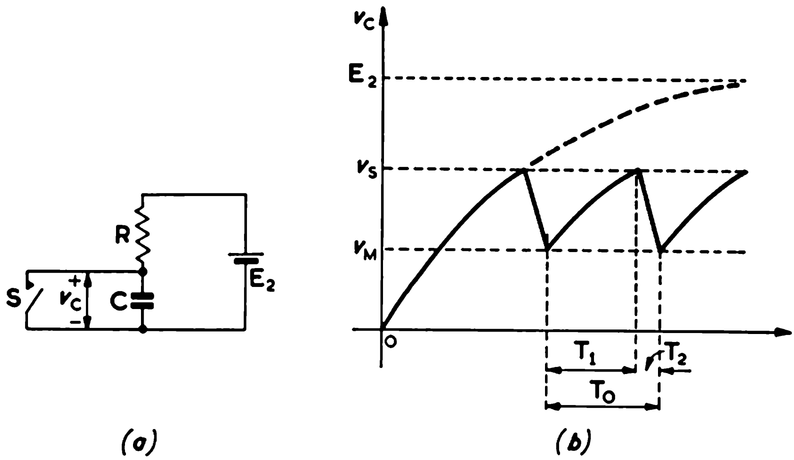


Fig. 18.1

$$\varepsilon^{-T_1/RC} = (E_2 - v_S)/(E_2 - v_M).$$

The time T_2 of discharge of the condenser from voltage v_S down to v_M is determined by the series resistance of the condenser plus the resistance of the switch. Usually $T_2 \ll T_1$, so that the period of oscillation is nearly equal to T_1, which depends on the time constant RC. It is shown in Section 18.2 how certain types of valve may act as the switch S. In Fig. 18.1.*b* the condenser is assumed to be uncharged at time $t = 0$. If the switch is not present, the condenser ultimately charges up to the battery voltage E_2. The initial part of the rise of voltage is nearly linear, but, as the condenser voltage approaches E_2, the rate of rise of voltage decreases. For linear voltage rise from v_M to v_S it is necessary for E_2 to be considerably greater than v_S. The waveform, usually called a " saw-tooth ", is used for the time base of a cathode-ray tube. The period of oscillation can be altered by changing R or C, and also by changing either v_S or v_M or both.

18.2. Gas Diode or Triode Oscillator

A cold-cathode gas diode may be used as the switch S, as shown in Fig. 18.2. The voltages v_S and v_M correspond to the striking and maintenance voltages respectively (see Section 5.12). This is a simple type of oscillator whose frequency is altered by changing the time constant RC. The amplitude of the oscillation is constant and equal to $v_S - v_M$. As long as this is small compared to E_2, the condenser voltage rises reasonably linearly with time. A small resistance R_1 is included to limit the diode current during the discharge of the condenser. The actual value of R_1 is determined from the maximum permissible diode current.

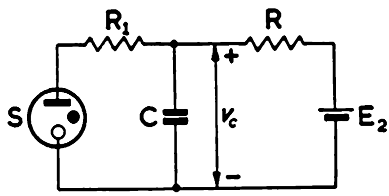


FIG. 18.2

The hot-cathode gas triode or thyratron may also be used as the switch

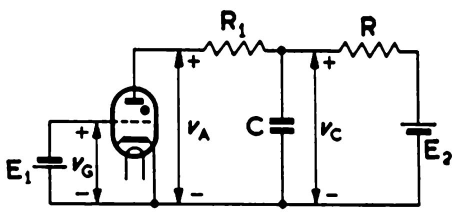


FIG. 18.3

in a saw-tooth generator with the circuit shown in Fig. 18.3. The thyratron has the advantage that the striking voltage v_S can be changed by altering the value of the grid bias E_1. The greater the magnitude of the negative bias, the higher the striking voltage. However, as shown in Section 6.12, the maintenance voltage v_M is approximately equal to the ionization potential of the gas and is almost independent of the grid bias. Thus it is possible to control the amplitude of oscillation by varying the bias, but this is accompanied by a change of frequency. The frequency can also be changed by varying the time constant, and this does not alter the amplitude.

A refinement of this circuit is the use of a pentode in place of the charging resistance, as shown in Fig. 18.4. The pentode has the property of

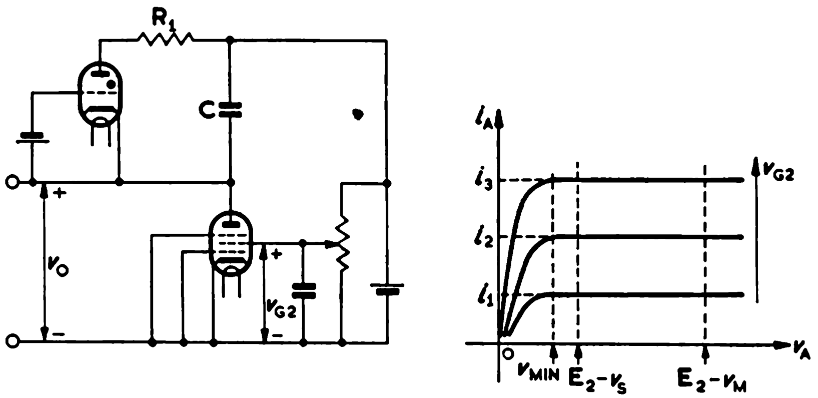


FIG. 18.4

passing an anode current which is almost independent of the applied anode voltage, as long as the voltages on all the grids are kept constant, and the anode voltage is greater than a certain minimum, v_{min}. The constant charging current produces a linear rise of voltage across C. The anode current, and hence the frequency of oscillation, can be altered by changing the screen voltage.

18.3. Feedback Relaxation Oscillators—The Multivibrator

In dealing with feedback oscillators in Chapter 13 it is shown that two conditions are necessary for oscillation. The phase of the feedback voltage must be exactly correct, and its magnitude must be sufficiently great. If the output of a two-stage resistance–capacitance coupled amplifier is connected back to the input, as shown in Fig. 18.5.*a*, the magnitude of the feedback may be much more than enough for oscillation. On account of the coupling condensers the phase of the feedback is exactly correct when $\frac{1}{\omega C_1} = \frac{1}{\omega C_2} = 0$, i.e., at infinite frequency. In practice, because of lead inductances and valve capacitances, this would probably mean some fairly high frequency. However, it is found that this circuit gives rise to a relaxation oscillation in which each valve conducts in turn while the other is cut off; the frequency is determined by the time constants of the circuits. The mechanism of the oscillation may be explained roughly as follows. Because of the positive feedback, any small change of voltage produces large and sudden changes throughout the circuit. This results in one of the valves being driven beyond cut-off and amplification ceases. The grid bias of this valve now increases at a rate de-

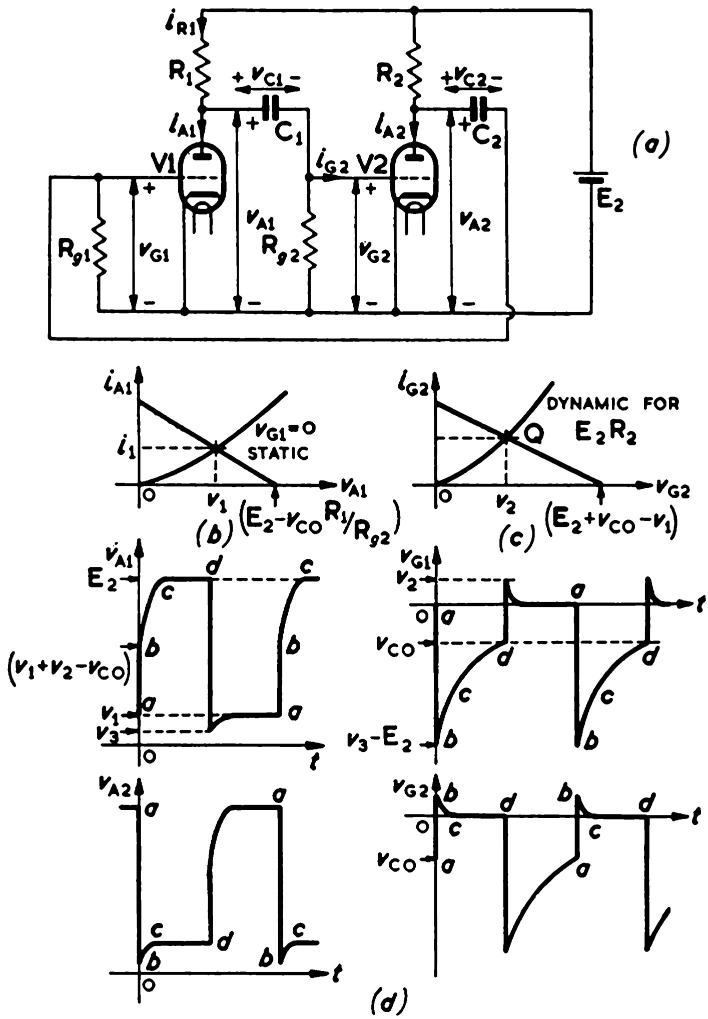


FIG. 18.5

termined by the circuit time constant. As the bias passes through the cut-off value amplification occurs again, resulting in large and sudden changes in the opposite direction, and ending in the other valve being cut off. The whole process is repetitive and the anode voltage waveforms are nearly rectangular (see Fig. 18.5.*d*). We now attempt detailed explanation of these waveforms. Let it be assumed that at a certain instant $V1$ is conducting with $v_{G1} = 0$ and $i_{A1} = i_1$, and $V2$ is cut off with $v_{G2} = v_{C0}$ and $i_{A2} = 0$. Under these conditions

$$v_{A1} = E_2 - v_{C0}R_1/R_{G2} - Ri_1 = v_1 \text{ (see Fig. 18.5.}b\text{)}$$

and $$v_{A2} = E_2.$$

The corresponding points are marked a in Fig. 18.5.d, which shows the variations with time of the anode and grid voltages of the two valves. At the same instant the voltages across the condensers C_1 and C_2 are

$$v_{C1} = v_{A1} - v_{G2} = v_1 - v_{C0}$$

and $$v_{C2} = v_{A2} - v_{G1} = E_2.$$

Now let v_{G2} increase slightly for any reason so that anode current just starts to flow in $V2$. Then v_{A2} becomes less than E_2. As the voltage across C_2 cannot change instantaneously, then v_{G1} goes negative by the same amount. This reduces i_{A1} so that v_{A1} becomes greater than v_1. The voltage across C_1 cannot change abruptly so that v_{G2} increases further. This means that i_{A2} increases, v_{A2} falls and v_{G1} becomes even more negative. The whole process is cumulative giving a sudden avalanche of change in which $V1$ becomes cut off, whilst $V2$ conducts and its grid is driven positive. Then

$$i_{A1} = 0 \text{ and } v_{A1} = E_2 - R_1 i_{R1}.$$

The end point of the avalanche may be determined by considering the grid current taken by $V2$. If R_{g2} is large so that v_{G2}/R_{g2} is negligible compared with i_{G2}, then, as $i_{A1} = 0$,

$$i_{G2} = i_{R1} \text{ and } v_{A1} = E_2 - R_1 i_{G2}.$$

The avalanche has taken place instantaneously without change of condenser voltages so that v_{A1} must have changed from its initial value of v_1 by an amount equal to the change in v_{G2}. Thus after the avalanche, i.e., at b in the diagrams

$$v_{A1} = v_1 + v_{G2} - v_{C0}$$

so that $v_{G2} = (E_2 + v_{C0} - v_1) - R_1 i_{G2}$. This equation may be represented by a load line in a v_{G2}, i_{G2} plot as shown in Fig. 18.5.c. In this diagram is shown also the dynamic characteristic of i_{G2} against v_{G2} for $V2$ operating with battery voltage E_2 and anode load R_2. The point Q in this figure gives the end of the avalanche where

$$v_{G2} = v_2 \text{ and } v_{A1} = v_1 + v_2 - v_{C0}.$$

If i_{A2} is the anode current of $V2$ at b then

$$v_{A2} = E_2 - R_2 i_{A2} = v_3 \text{ (see Fig. 18.5.}d\text{)}$$

and $$v_{G1} = v_3 - E_2.$$

The condenser C_2 now begins to discharge from E_2 and condenser C_1 begins to charge from $v_1 - v_{C0}$. As $V2$ is taking grid current, the charging of C_1 is determined mainly by the time constant R_1C_1, as long as $v_{G2}/i_{G2} \ll R_1$ and R_{g2}. Thus C_1 charges up rapidly to E_2 whilst v_{G2} falls to zero. The anode current of $V2$ falls somewhat and v_{A2} rises to v_1. There is a consequent change of v_{G1} of amount $v_1 - v_3$, although v_{G1} still remains beyond cut-off. The corresponding positions on the waveforms are marked *c*. The next part of the waveform arises from condenser C_2 discharging through R_2 and R_{g1} in series, with a time constant $C_2(R_2 + R_{g1})$. During this discharge everything else is quiescent until

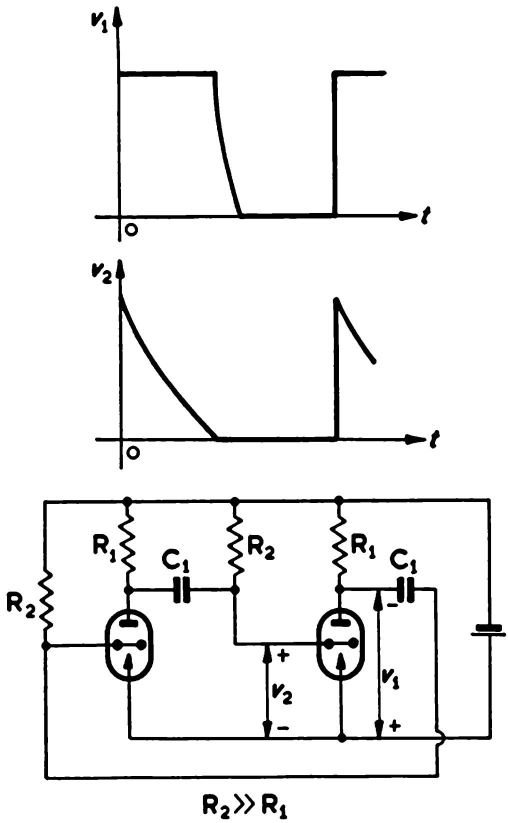


FIG. 18.6

v_{G1} reaches v_{C0} at points *d*. $V1$ now begins to take current and a second avalanche occurs but in the reverse direction, giving an abrupt transfer of current from $V2$ to $V1$. The second avalanche is limited as before, and the whole process is repetitive with waveforms shown in Fig. 18.5.*d*. In drawing these diagrams it has been assumed that the circuit is symmetrical with identical valves and $R_1 = R_2$, $C_1 = C_2$ and $R_{g1} = R_{g2}$. This relaxation oscillator is known as a free-running multivibrator.

A circuit for a transistor multivibrator is shown in Fig. 18.6. The principle of operation is similar to the one described above, and the waveforms are summarized in the diagram.

To some extent the multivibrator may be likened to a limiting case of a squegging oscillator of the type described in Section 13.5. The valves attempt to build up an oscillation at a very high frequency in agreement with the phase relation mentioned at the beginning of this section. However, the oscillation builds up only part of one quarter cycle before the bias condition of one valve causes the oscillation to cease. The bias now changes slowly until conditions for oscillation are again correct, when a further quarter cycle occurs before the oscillation stops once more.

18.4. The Transitron Relaxation Oscillator

Another relaxation oscillator using feedback may be based on the transitron circuit which is described in Section 13.8. When such a circuit has a resistive load (R in Fig. 18.7) the output voltage v_{g2} is in

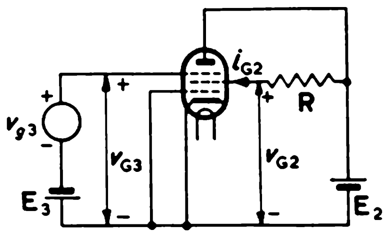


FIG. 18.7

phase with the input voltage v_{g3}. Thus direct feedback between screen and suppressor may give the required phase conditions, and the magnitude of the feedback may be more than is necessary for oscillation. In order to separate the d.c. supplies, the feedback is introduced by means of a capacitance C as shown in Fig. 18.8.*a*. The presence of C gives the

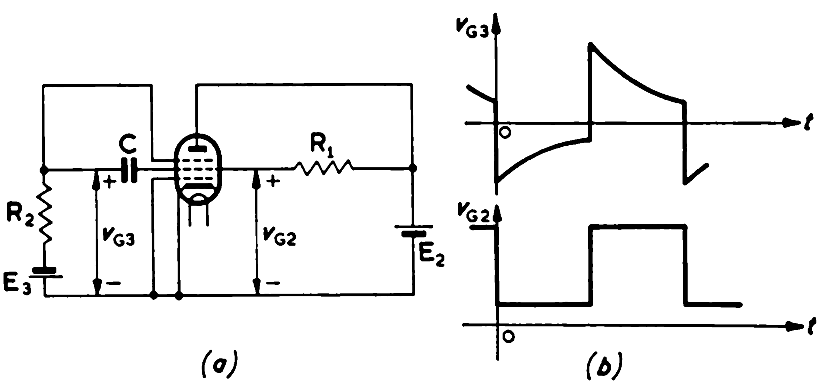


FIG. 18.8

correct phase condition for oscillation only at very high frequency. It may be seen that the conditions are similar to those in the multivibrator, and again relaxation oscillations are obtained. Sudden large changes occur in v_{G3} and v_{G2}, with resulting waveforms of the type shown in Fig. 18.8.*b*. The limits to the changes are set by the regions of the characteristics where the suppressor voltage no longer affects the screen current and amplification ceases. The recovery time after each sudden change depends on the time constant $C(R_1 + R_2)$.

18.5. The Blocking Oscillator

Frequently it is required to produce voltage pulses whose width is much less than the separation between the pulses. It is possible to do this with a multivibrator by proper choice of the circuit components. However, a more suitable arrangement is the blocking oscillator, which is

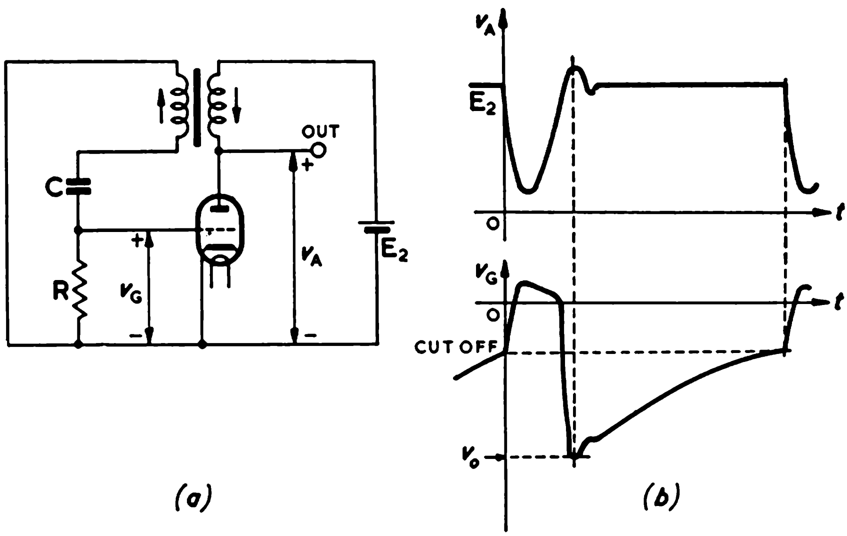


FIG. 18.9

shown in Fig. 18.9.*a*. In Section 18.3 the multivibrator is likened to a very high-frequency squegging oscillator which has excessive feedback and which builds up only a fraction of one cycle of the high-frequency oscillation before the valve is cut off. The blocking oscillator is similar to some extent, but it uses a single valve with mutual inductance feedback, usually through an iron-cored transformer. The detailed action of the oscillator is quite complicated, as it depends not only on the non-linear characteristics of the triode but also to some extent on the magnetic saturation of the transformer. The sequence of operations may be followed by assuming that the grid of the triode is initially beyond cut-off but is becoming less negative. When the grid rises above the cut-off value anode current begins to flow and an induced voltage, proportional to the rate of change of anode current, makes the grid less negative, thus increasing the anode current even further. The action of the feedback is to give a sharply decreasing anode voltage and rising grid voltage, as shown in Fig. 18.9.*b*. Two factors bring this feedback action to a stop. Firstly, when the grid becomes positive there is effectively across the transformer a low resistance damping the oscillation, and secondly, the mutual conductance of the triode is decreasing. The grid voltage then falls slowly for a short time until it again becomes slightly negative, whilst the anode goes more positive. The effect of removing the load from the transformer together with the falling anode current is to produce a rapid feedback action in the opposite direction, which drives the grid beyond cut-off and charges the condenser C to a voltage v_0. The condenser now discharges through the resistance R. When the grid voltage again reaches the cut-off value the whole process is repeated. The length of the pulse depends upon the characteristics of the transformer and the value of C, and is usually much shorter than the time interval of repetition. The latter is determined mainly by the time constant of the RC circuit, so that the frequency varies approximately linearly with the reciprocal of R.

18.6. Monostable Circuits

The circuits described in this section have one stable and one unstable state. When a suitable signal is applied the circuit goes from one state to the other, and then returns to the stable condition after a time determined by the constants of the circuit. The circuit produces an output

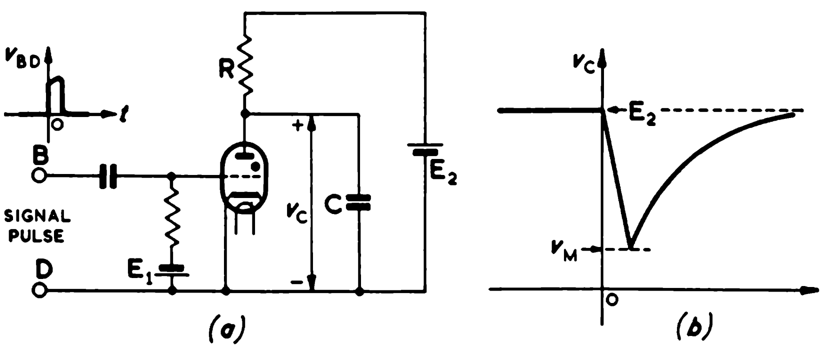


FIG. 18.10

voltage of fixed magnitude and waveform for input signals which may vary in amplitude and shape over wide limits.

A monostable thyratron circuit is shown in Fig. 18.10. The stable state occurs when the anode voltage is below the striking value, and the condenser C is charged to a voltage E_2. When a positive pulse is applied to the grid the thyratron strikes and rapidly discharges the condenser down to the maintenance voltage of the thyratron when the discharge is extinguished. The condenser then recharges to a voltage E_2 through resistance R, returning the circuit to its stable condition, in which it remains until it is triggered again by another signal.

In the multivibrator described in Section 18.3 the operation is closely connected with the charging and discharging of the coupling capacitances between the stages of the two-stage resistance-loaded amplifier. We know that such amplifiers may have direct coupling between the stages, and it is interesting to consider the effect of this on the behaviour of the multivibrator. In the circuit shown in Fig. 18.11 one of the coupling condensers is replaced by a battery E_1. We assume that the circuit has initially both valves operating with zero grid voltage, and we find the

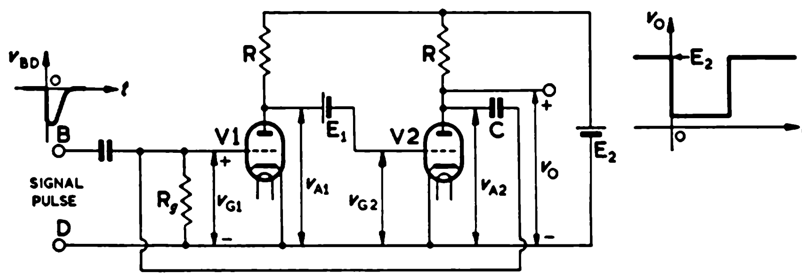
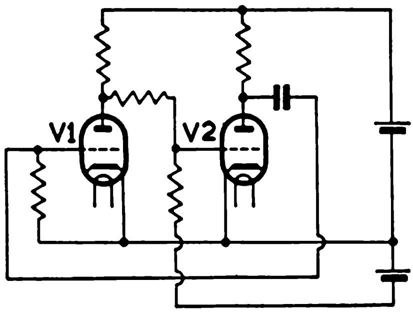


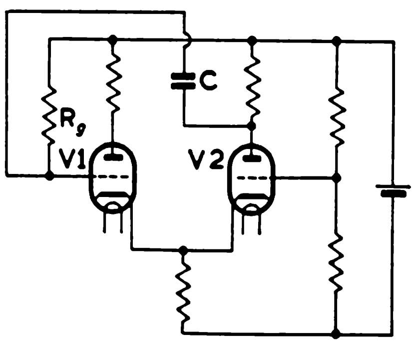
FIG. 18.11

effect of a small change in one of them. For example, let the first grid voltage become slightly negative. This change is amplified by the two stages and is fed back to the first valve, making its grid much more negative. An avalanche occurs and the valve is driven beyond cut-off. At the same time the second valve passes a large current and v_{A2} drops to a low value. The condenser C then discharges at a rate depending on $C(R + R_g)$, just as in the free-running multivibrator, and the negative grid voltage on the first valve decreases until current starts to flow. This initiates another avalanche which makes the first grid positive and the second grid negative. At the same time the first anode voltage drops to a low value. Now the actual value of the second grid voltage is given by $v_{G2} = v_{A1} - E_1$. It is possible therefore, by suitable choice of E_1, for v_{G2} to cut off the current in $V2$ after the second avalanche. There is now no mechanism in the circuit to change this condition. Thus, this circuit has one unstable condition with $V1$ cut-off and one stable condition with $V2$ cut-off; hence the name monostable multivibrator, or univibrator. When in its stable state, the second valve may be rendered conducting by the application of a positive signal to its grid. This circuit has many applications as a switch or relay, whose operation is controlled by an external signal or trigger of the correct polarity. After operation the relay is automatically reset. When required to operate on receiving a negative signal the latter is connected to the first grid.

Practical univibrator circuits do not have battery coupling. One of the alternative methods of direct coupling is used, as described in Chapter 12. Two circuits using a third rail and cathode coupling are shown in Fig. 18.12.*a* and *b* respectively. An interesting point in the circuit of Fig. 18.12.*b* is that the grid resistance R_g is returned to the h.t. positive



(a)



(b)

FIG. 18.12

rail instead of h.t. negative. Following an unstable avalanche making v_{G1} very negative, the grid voltage gradually becomes less negative until the stable avalanche takes place. When R_g is returned to h.t. negative the cut-off value of v_{G1} occurs near the end of the discharge of C when the voltage is varying slowly, and the time of occurrence is rather uncertain. When R_g is returned to h.t. positive the cut-off voltage occurs whilst the discharge rate is still changing rapidly and the time of occurrence is more certain. This arrangement may also be used with advantage in the free-running multivibrator, particularly when R_g is large.

The blocking oscillator circuit can be adapted for monostable working by using a grid battery which cuts off the anode current in the stable condition, as shown in Fig. 18.13. When a positive pulse is applied to the grid, then one cycle of oscillation is induced just as in the free-running

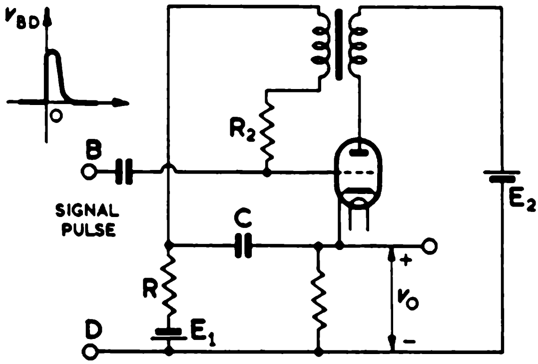


Fig. 18.13

circuit. At the end of the pulse of anode current the condenser C is charged to a voltage somewhat greater than E_1 and then discharges to its original state. The short pulse of output voltage can be made nearly rectangular in shape by adjusting the value of resistance R_2.

18.7. Bistable Circuits

Some circuits possess two separate stable states and can be transferred from one to the other by suitable input signals. However, they can exist indefinitely in any one of the two states when required. The circuit of Fig. 18.14 shows a bistable thyratron circuit. In one stable condition the

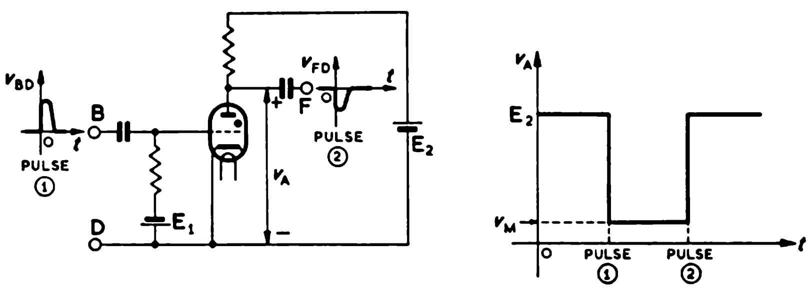
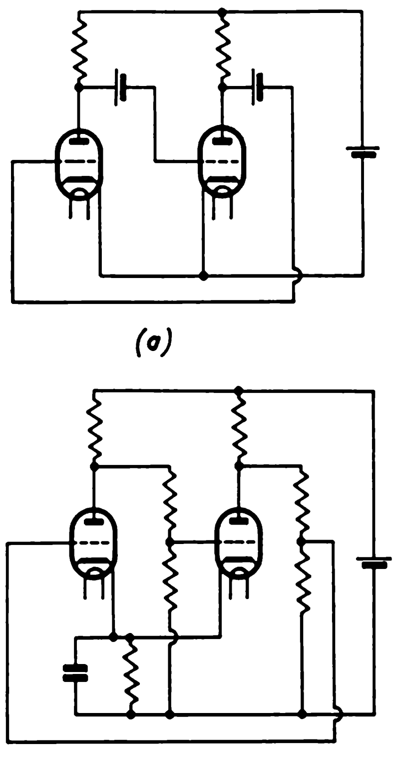


Fig. 18.14

anode voltage is less than the striking voltage. The application of a positive pulse to the grid causes the thyratron to strike, giving the second stable condition, which is unaffected by any further positive pulses to the grid. The new condition remains until a negative pulse is applied to the anode, extinguishing the discharge and returning the circuit to its original stable state, which in turn is unaffected by further negative pulses to the anode.

The substitution of direct coupling for one of the coupling condensers in the free-running multivibrator gives a circuit with one stable and one unstable state. When the second coupling condenser is also replaced by direct coupling, as in Fig. 18.15.*a* or *b*, there are two stable states. When



(*a*)

(*b*)

FIG. 18.15

this circuit receives a suitable triggering signal at one grid it changes from one stable state to another. It returns to its original state when a signal of reversed polarity is applied to the same grid. This circuit may be used as a switch or relay when it is desired that the action be controlled by two separate signals.

18.8. Cathode-coupled Trigger Circuit

A direct-coupled multivibrator can have two stable states. If a common-cathode resistor is used for coupling, then bistable operation

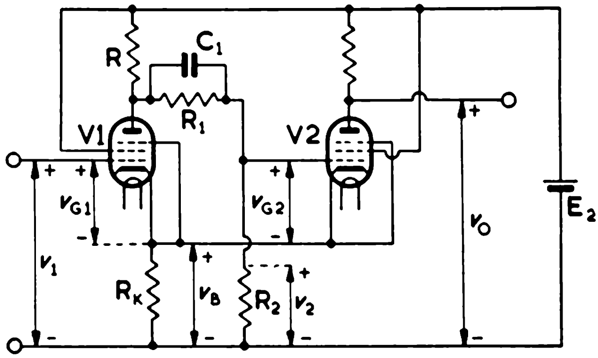


FIG. 18.16

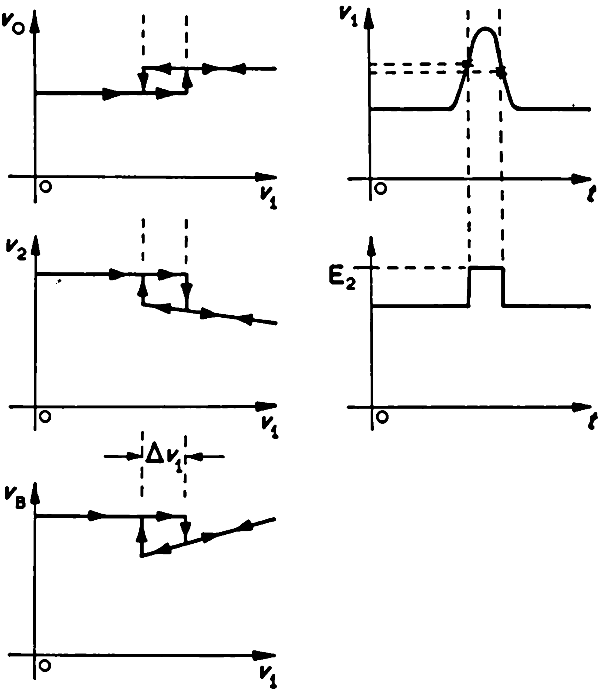


FIG. 18.17

is still obtained, but the existence of either state depends on whether the signal voltage is above or below a threshold value. The circuit is therefore able to discriminate between the amplitudes of signals. It is shown in Fig. 18.16, and it is sometimes called the Schmitt Trigger circuit. One stable state occurs with zero input voltage and with $V1$ cut-off. Its grid–cathode voltage is equal to v_B, which is due to the anode current of $V2$ flowing through the cathode resistor R_K. Under this condition

v_{G2} is slightly negative. This voltage is determined by v_B and the potential divider R, R_1, R_2. If v_1 is now increased to approach v_B, $V1$ begins to pass current and causes a fall in v_{A1}. As a result v_{G2} becomes more negative and v_B decreases, thus increasing the current in $V1$. There is a sudden transfer of current from $V2$ to $V1$ until the second valve is cut off, giving the second stable state, which is unaffected by further increase of v_1, though v_B now varies with v_1. When v_1 is reduced the circuit remains in its second stable state until $V2$ starts to take current, when there is a sudden transfer back to the original stable state. The diagrams in Fig. 18.17 show the voltage variations. The threshold voltage depends to some extent on whether v_1 is increasing or decreasing, but the "backlash" Δv_1 can be made small compared with v_1. The capacitance C_1 is included to avoid the delay in transfer of rapid changes in v_{A1} to v_{G2} arising from the stray capacitance across R_2.

18.9. Counting and Scaling

Some of the circuits described in this chapter are used for counting electronically, as in digital computing machines or in counting the pulses produced by radiation particles in ionization counters. In any counting process the principle of scaling is used. In ordinary arithmetical count-

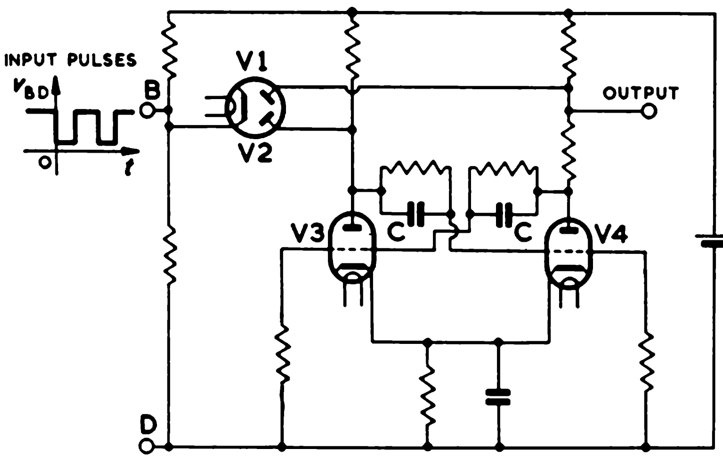


FIG. 18.18

ing we use the decimal scale of 10. Any particular number is expressed in powers of 10, e.g., 971 is equal to $9 \times 10^2 + 7 \times 10^1 + 1 \times 10^0$. Electronically it is convenient to use the binary scale of 2, and this can be done by means of a number of bistable multivibrator circuits. In Section 18.7 it is shown that a bistable circuit requires two signals to take the circuit through one complete cycle of operation and that two signal pulses produce only one output pulse from the anode of one of the valves. If this output pulse is passed on to a second bistable circuit, then four signal pulses are required to produce one output pulse from an anode of the second pair of valves. Each pair of valves gives a division by 2 of the number of signal pulses and so a counting system in the scale of 2 can be produced with a series of bistable circuits in tandem. If there are n circuits, pulses can be counted up to 2^n. One basic bistable circuit for binary scaling is shown in Fig. 18.18. The bistable multivibrator of valves $V3$ and $V4$ is the same as that of Fig. 18.15.*b* except that the condensers C are added to speed up the response to rapid changes. The diodes $V1$ and $V2$ are used to feed the signal pulses to $V3$ and $V4$. Sometimes a bistable circuit has the signals connected directly to both valves through isolating condensers. This may work all right, but the application of a large negative pulse to both valves simultaneously may give rise to undesirable effects. The purpose of the diodes in Fig. 18.18 is to ensure that the signal pulse affects only one of the valves $V3$ or $V4$ at a time. If the circuit is quiescent with $V3$ passing current and $V4$ cut off, then the anode of $V3$, is at a lower potential than the anode of $V4$. Thus the negative voltage across $V1$ is less than the negative voltage across $V2$. When the negative signal pulse is applied to the common cathode of $V1$ and $V2$, $V1$ conducts but not $V2$, and the triggering signal is applied to $V4$ but not to $V3$. In the other stable state the reverse situation exists. This is sometimes referred to as a steering circuit.

18.10. Decade Scaling

An electronic decade counter can be produced using four bistable multivibrators suitably connected (see Exx. XVIII). An interesting cold-cathode gas valve, called a Dekatron, has been specially designed for decade counting. Fig. 18.19 shows diagrammatically one type of Dekatron, in which there is a cylindrical anode, surrounded by nine cathodes " *a* ", ten cathodes " *b* ", ten cathodes " *c* " and one cathode " *d* ". All the cathodes of one type are connected together, and the " *d* " cathode is normally at

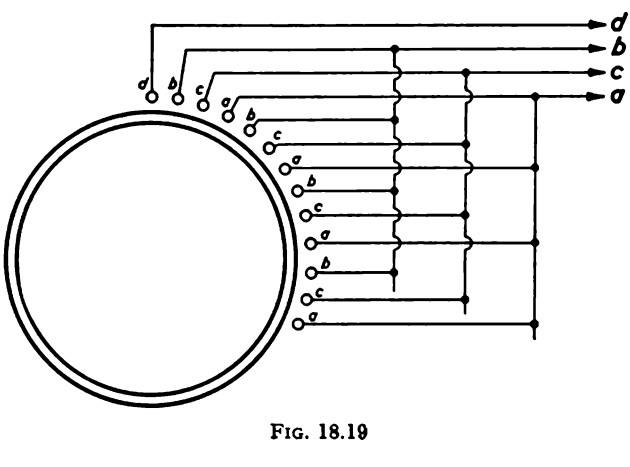


FIG. 18.19

the same potential as the " *a* " cathodes. Each " *a* " cathode has a " *b* " and a " *c* " electrode on either side of it. Under quiescent conditions the " *b* " and " *c* " cathodes are at the same negative potential with respect to the anode and the " *a* " cathodes are more negative still. Thus a local glow discharge is set up between one " *a* " cathode and the anode. A signal pulse now makes the " *b* " cathodes more negative than the " *a* " cathodes and the discharge moves to the adjacent " *b* " electrode. A short time later the same pulse is arranged to make the " *c* " cathodes even more negative so that the discharge moves on one place to a " *c* " electrode. The " *b* " and " *c* " potentials now return to their quiescent values and the discharge moves to the next " *a* " cathode. The signal pulse has thus advanced the discharge in a clockwise direction by one " *a* " cathode. The single " *d* " cathode functions similarly to the " *a* " cathodes, but, although it is at the same potential as the " *a* " electrodes, it is connected separately through a resistance. When the discharge reaches the " *d* " electrode there is a pulse of current through the resist-

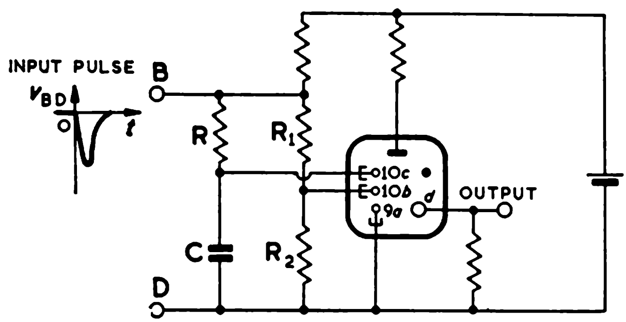


FIG. 18.20

ance. Thus for every ten signal pulses there is one output pulse, which may, after suitable amplification, be passed on to another similar Dekatron. One output pulse is obtained from the second Dekatron for every 100 signal pulses. Provision is usually made for zero setting by switching a large negative pulse to the " *d* " electrode so that the discharge moves to that electrode. Normally the gas discharge can be seen through the end of the Dekatron, and a scale marked 0 to 9 gives a visible indication of the count.

A typical circuit for use with a Dekatron is shown in Fig. 18.20. The potential divider R_1, R_2 ensures that about half of the input pulse is applied to the " *b* " cathodes. At a short time later, determined by the *RC* circuit, the full pulse is applied to the " *c* " cathodes.

18.11. Amplitude Control and Discrimination

For satisfactory operation of the counting circuits described in Sections 18.9 and 18.10 it is desirable to have pulses of constant amplitude. This can be achieved by using the signal pulse to operate a monostable multivibrator (see Section 18.6). As long as the signal pulse is sufficient to operate the multivibrator the latter produces an output pulse of constant amplitude and shape.

Sometimes it is desired to count only pulses which are greater than a certain amplitude. This can be done readily with a cathode-coupled trigger circuit (see Section 18.8). Frequently, counting circuits embrace a combination of trigger, switching, shaping and scaling circuits.

CHAPTER 19

WAVE SHAPING

19.1. Wave-shaping Circuits

In the operation of certain electronic circuits it is essential for the signal to have some particular waveform. For example, pulses may be required with flat top, steep sides, suitable duration or suitable magnitudes. To achieve these requirements the signal may have to pass through a wave-shaping circuit. There are two main classes of such circuits. In the first, use is made of the non-linear characteristics of valves for changing the shape of a sinusoidal or other type of signal. In the second class, the transient properties of linear circuits are exploited.

19.2. Non-linear Wave-shaping Circuits

The asymmetric conductivity of thermionic or crystal diodes may be used for wave shaping. For example, if a sinusoidal voltage is applied to a resistance and a diode in series, as shown in Fig. 19.1, the voltage across the resistance consists of a series of half sine waves. Provided the series resistance is much greater than the diode forward resistance, the voltage

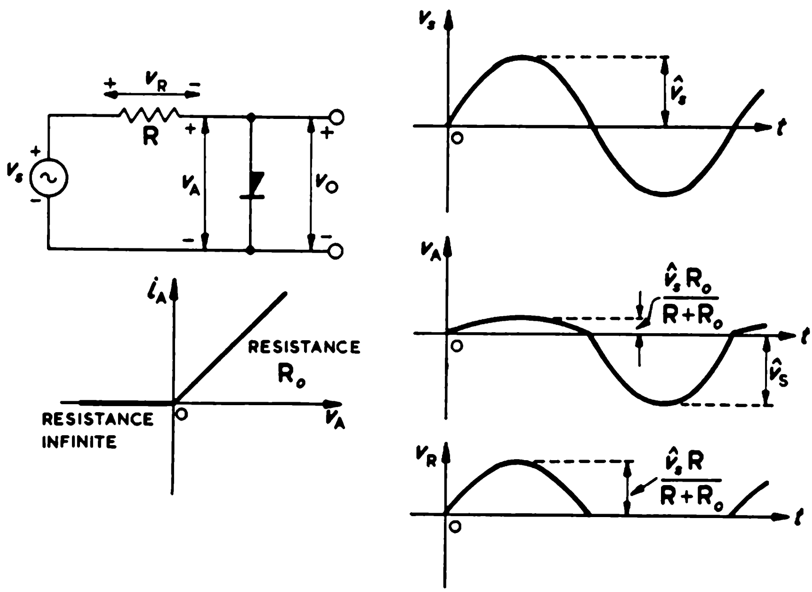


Fig. 19.1

across the diode also consists very nearly of half sines. The diode has clipped or limited the waveform. When, as in Fig. 19.2, there is a battery in the series circuit another clipped waveform is obtained. By

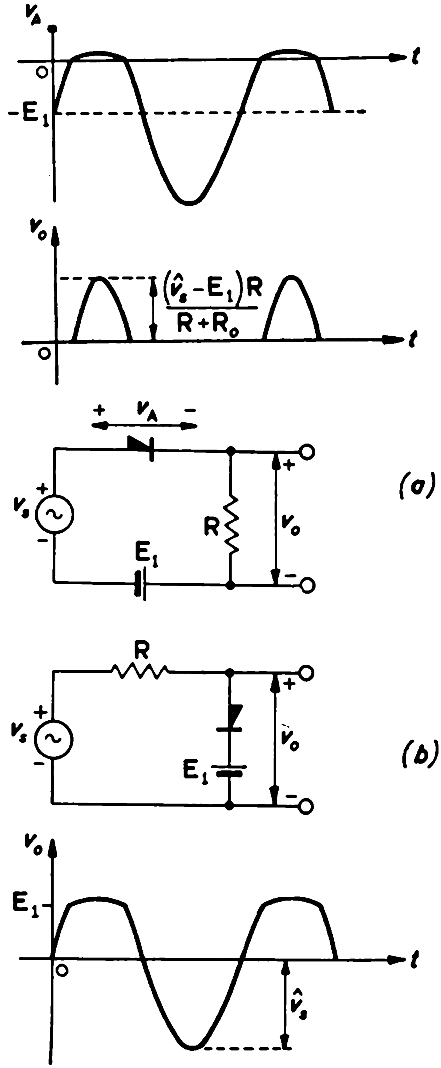


FIG. 19.2

using two diodes in the circuit of Fig. 19.3 both halves of the sine wave are limited.

Wave clipping is frequently achieved by means of the non-linear grid characteristic of a triode. With a high resistance in series with the input

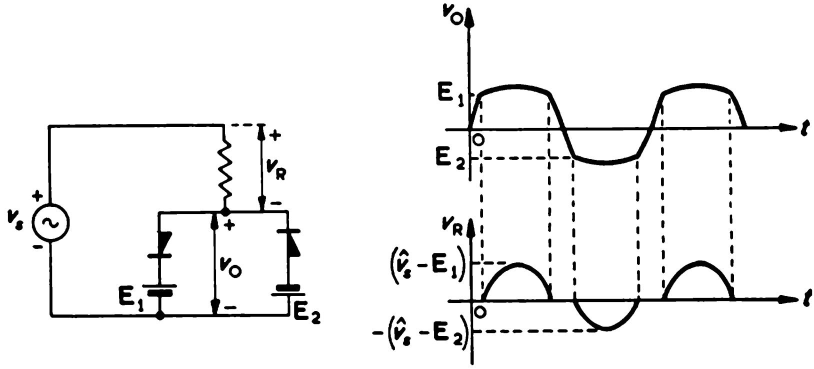


FIG. 19.3

signal the grid and cathode can act together like the diode in the previous paragraph. When the grid is driven positive the grid-cathode voltage is limited, and the resulting output voltage takes the form shown in Fig. 19.4. When the grid voltage is positive the cathode current is shared between the grid and the anode. As the voltage rises an increasing share of the current goes to the grid, and the anode current passes through

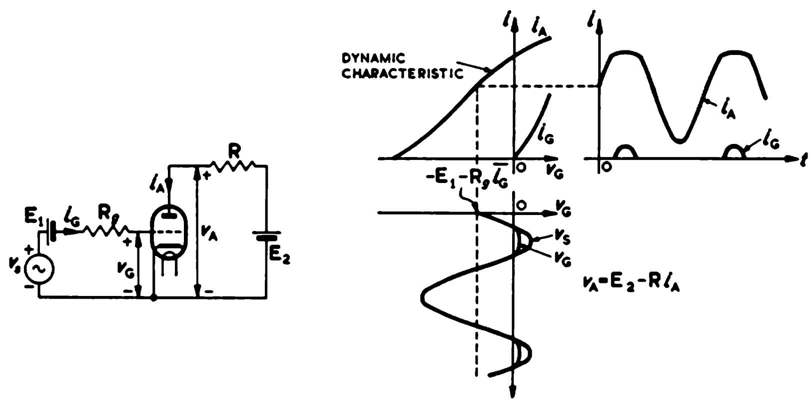


FIG. 19.4

a maximum value as seen in Fig. 19.5. Thus wave clipping can arise even without a high resistance in series with the grid signal. If, at the same time, the signal is sufficiently large to give cut-off over part of the cycle the other half of the waveform is also limited, as shown in the diagram. By suitable adjustment of the grid bias cut-off clipping is easily achieved. A two-stage amplifier can give cut-off clipping in each stage. Because of the phase reversal in a resistance-loaded amplifier both peaks are limited as shown in Fig. 19.6.

A junction transistor amplifier of the form illustrated in Fig. 19.7 can be used for shaping a signal waveform. The transistor amplifies linearly only for a limited range of signal. Cut-off occurs when the base is positive

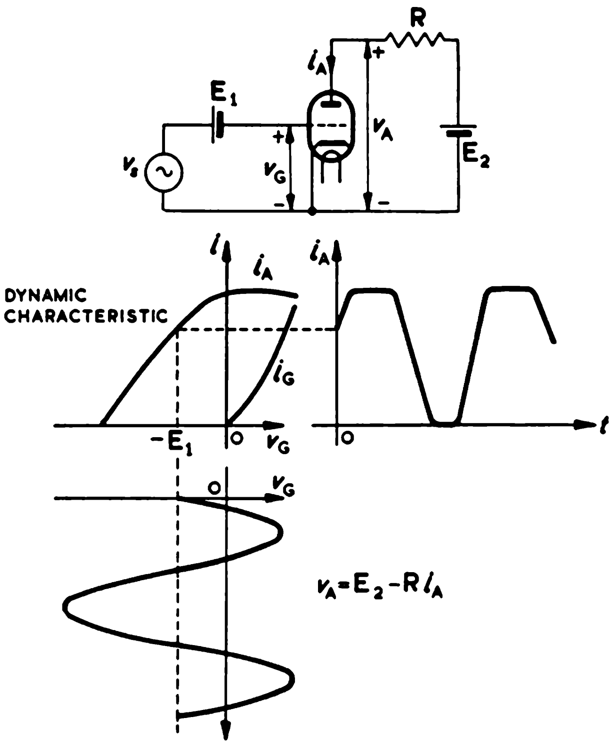


FIG. 19.5

with respect to the emitter. Also, with a large enough signal, particularly when the load resistance is high, the collector voltage is almost zero over part of the negative peak. Thus both halves of the signal are limited in the manner shown.

19.3. Clamping Circuits and d.c. Restoration

In the diode clipping circuit of Fig. 19.1 the output voltage across the diode does not rise appreciably above zero. The output voltage is said to be clamped at zero. Similarly, v_A in Fig. 19.2.*a* and v_0 in Fig. 19.2.*b* are clamped at zero and E_1 respectively. In the triode circuit of Fig. 19.4 the grid voltage is clamped at zero.

A clamping circuit is frequently used to restore the d.c. component of a pulse which has been passed through an a.c. amplifier or through any other circuit which does not pass d.c. In Fig. 19.8.*a* is shown a

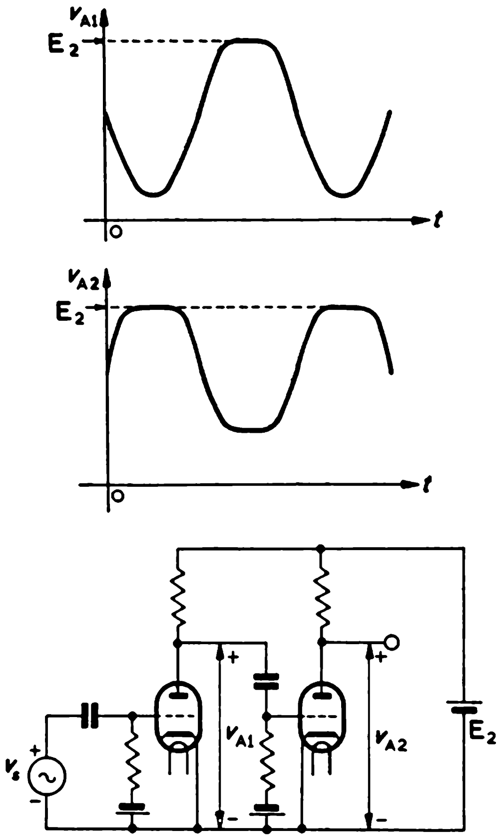


FIG. 19.6

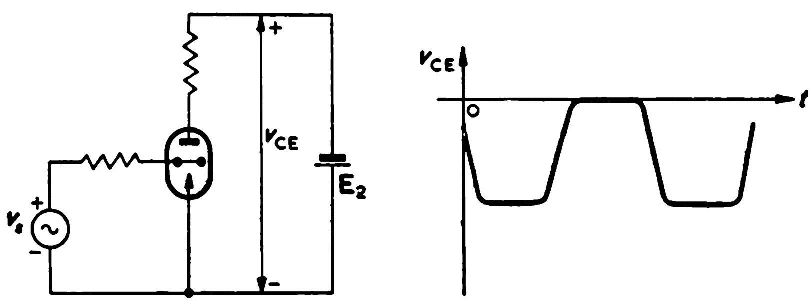


FIG. 19.7

signal pulse which has a d.c. component. If this is passed through a transformer or an a.c. amplifier it is possible to reproduce the shape of the pulse without distortion, but the d.c. component disappears and the average

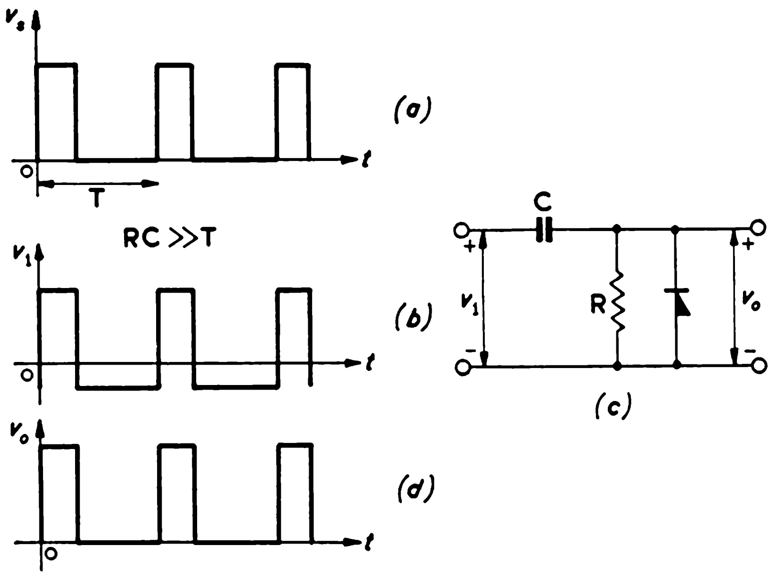


FIG. 19.8

value of the output is zero as shown in Fig. 19.8.*b*. The d.c. component of the pulse can be restored by using the circuit of Fig. 19.8.*c*, in which the time constant RC is large compared with one complete cycle of the pulse. During the negative portion of the pulse the diode conducts and

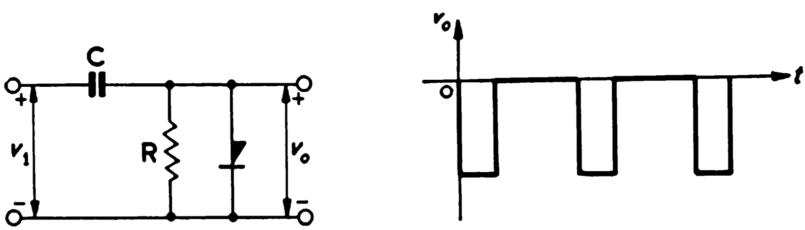


FIG. 19.9

charges condenser C. This continues until at no point of the cycle does the total voltage v_0 fall below zero, as shown in Fig. 19.8.*d*. Whatever the mean level v_1 of the pulse, this circuit automatically produces positive output pulses with zero level at the base. The circuit in Fig. 19.9 gives a negative output pulse which just rises to zero.

19.4. Linear Wave Shaping—Differentiating and Integrating

In Chapter 11 we consider the transient response of various circuits possessing resistance, capacitance and inductance. When the time constants of the circuits are very small or very large the response may be used to modify certain wave-forms. Consider the circuit of Fig. 19.10.*a*,

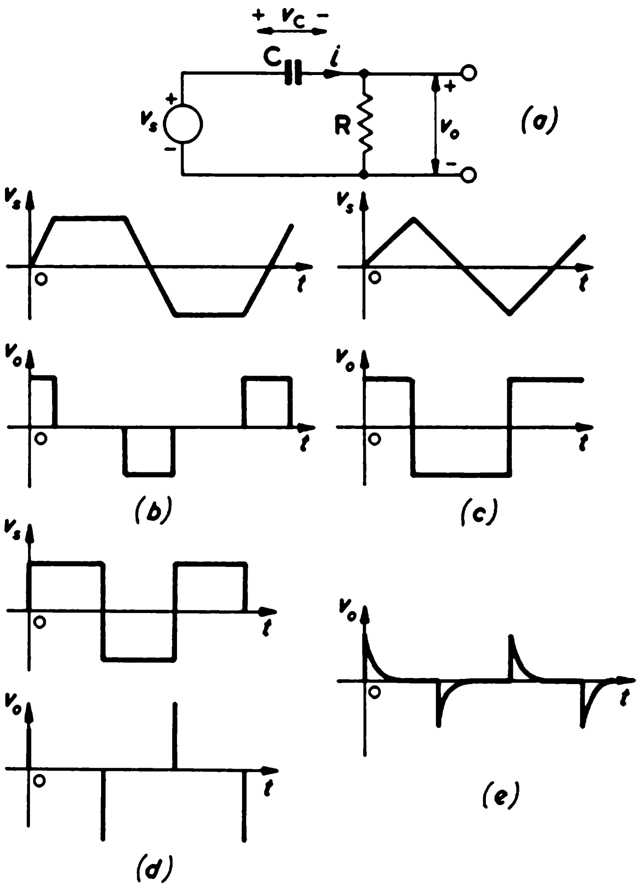


FIG. 19.10

in which a voltage v_s varying in time is connected in series with a resistor R and a condenser C. The circuit equations are

$$i = C\frac{dv_c}{dt},$$

$$v_o = RC\frac{dv_c}{dt}$$

and

$$v_s = v_c + v_o.$$

If RC is very small in comparison with the time required for any signal change, then $v_s \simeq v_c$ and

$$v_o \simeq RC\frac{dv_s}{dt}.$$

Thus the output voltage is approximately equal to the differential coefficient of the signal voltage. The response of this circuit to three different signals is shown in Fig. 19.10.*b*, *c* and *d*. The sharp spikes in Fig. 19.10.*d* in response to the step signals are really limiting values of the exponential waveforms of Fig. 19.10.*e*. The smaller the value of RC, the more nearly does the actual response correspond to Fig. 19.10.*d*. Another

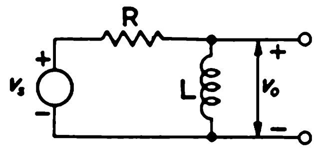


FIG. 19.11

differentiating circuit may be achieved with resistance and inductance connected as shown in Fig. 19.11. Provided that the time constant L/R is sufficiently small it follows that

$$v_o \simeq \frac{L}{R}\frac{dv_s}{dt}.$$

When the positions of the components in Fig. 19.10.*a* or Fig. 19.11 are interchanged then the circuits give a response corresponding to the mathematical process of integration, but now it is necessary for the time

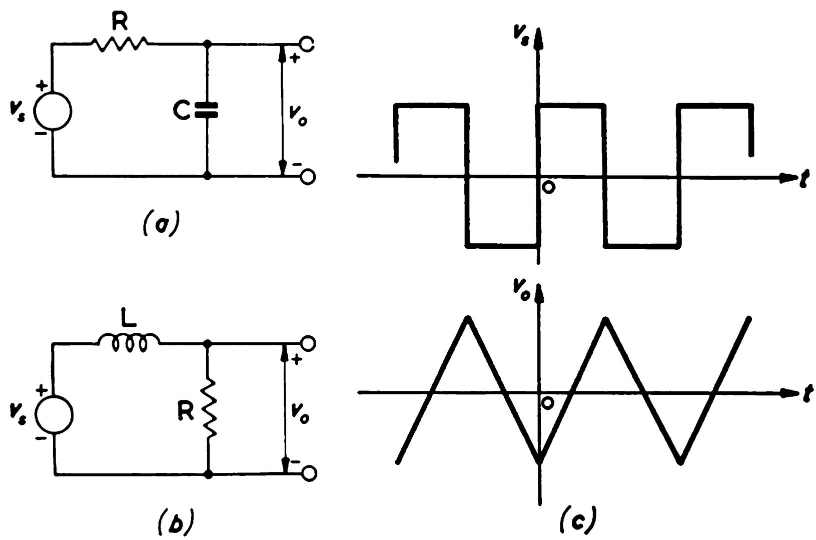


FIG. 19.12

constants to be large in comparison with the time of duration of signal changes. In Fig. 19.12.*a*

$$v_o = \frac{1}{C}\int i dt = \frac{1}{RC}\int v_R dt$$

and

$$v_s = v_R + \frac{1}{RC}\int v_R dt.$$

Then, provided RC is sufficiently great, $v_s \simeq v_R$ and

$$v_o \simeq \frac{1}{RC}\int v_s dt.$$

In the LR circuit of Fig. 19.12.*b* it may be shown similarly that

$$v_o \simeq \frac{R}{L}\int v_s dt$$

provided that L/R is sufficiently large. These integrating circuits give an output of triangular waveform in response to a square wave signal, as shown in Fig. 19.12.*c*.

19.5. Electronic Integrating Circuits

An interesting electronic integrator can be realized with a single pentode valve and by exploiting the Miller Effect. The circuit is shown in Fig. 19.13.*a*, in which there is a capacitance C between the anode and control grid of a pentode amplifier. It is shown in Section 10.11 that the effective capacitance between grid and cathode in this circuit is

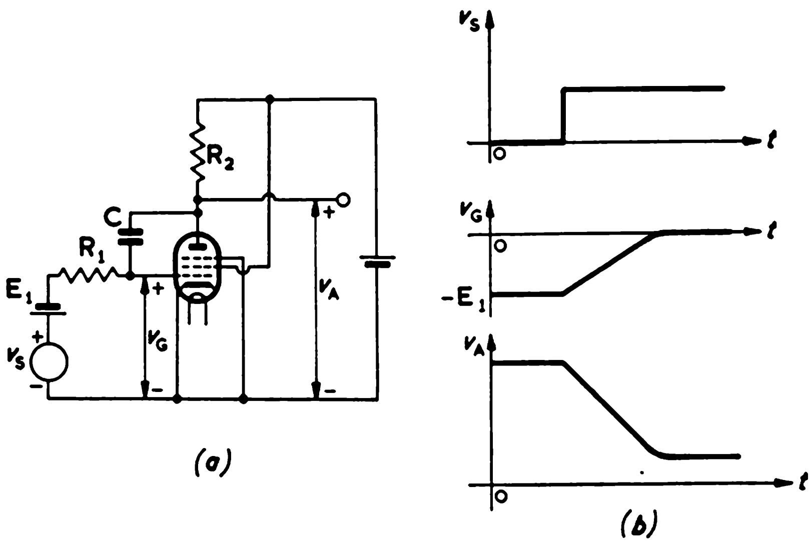


FIG. 19.13

$(A + 1)C$, where A is the voltage gain of the amplifier and, in this case, $A \simeq g_m R_2$. The time constant of the grid circuit is $R_1C(1 + g_m R_2)$, which is much greater than R_1C. When a step signal is applied to the grid circuit it is effectively integrated, the grid voltage rises linearly and the amplified output voltage drops linearly. If R_1 is large the input

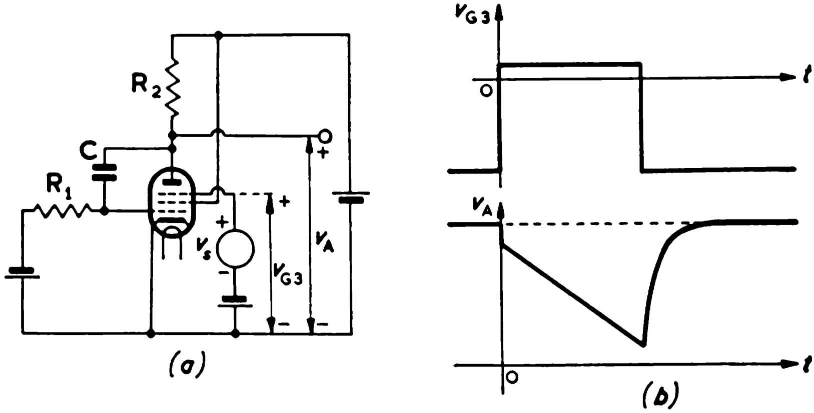


FIG. 19.14

voltage is clamped at $v_G = 0$, so that the linear variations cease abruptly. The waveforms are shown in Fig. 19.13.*b*. The Miller integrator is usually controlled on the suppressor grid. The suppressor is biased negatively so that the anode current is zero and the cathode current flows to the

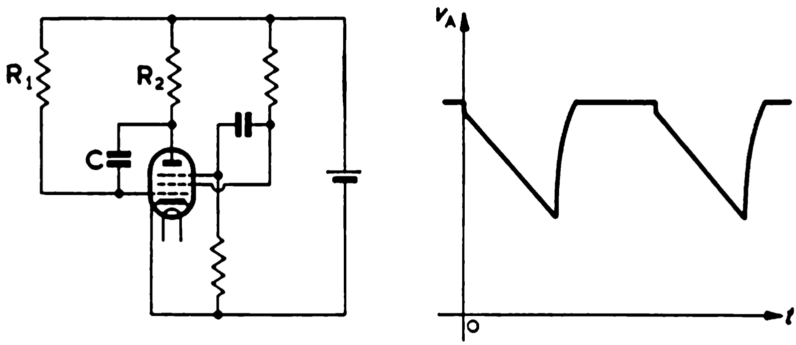


FIG. 19.15

screen. A positive pulse is applied to $G3$, so that anode current starts to flow, and there is a sudden drop in anode voltage which is passed on to $G1$, through C, and the integrating action occurs as the grid voltage returns to its quiescent value. The circuit and waveforms are shown in Fig. 19.14.*a* and *b*. This circuit may be used for producing a single-stroke time-base for a cathode-ray tube. A continuous time base can be achieved by combining this integrator with the transitron relaxation oscillator described in Section 18.4. Successive sweeps are initiated by each transitron avalanche, but the flow of anode current is controlled by

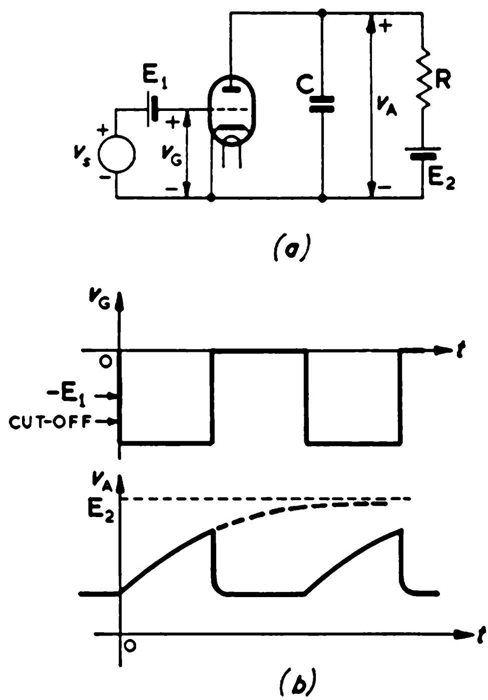


FIG. 19.16

the integrator. The circuit and waveforms are shown in Fig. 19.15; the integrating resistor R_1 is returned to h.t.+, but this does not affect the principle of operation. The circuit gives, from a single valve, a waveform which is highly linear and which is of considerable amplitude.

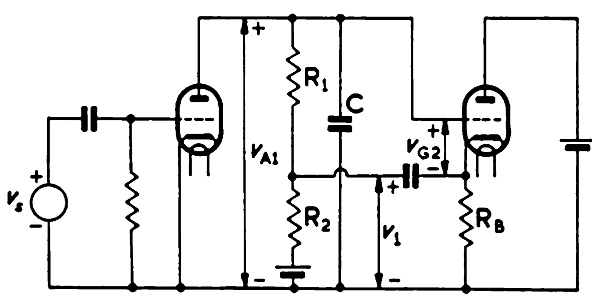


FIG. 19.17

Another electronic integrating circuit is illustrated in Fig. 19.16.*a*, where a triode has an anode load consisting of resistance R and capacitance C in parallel. A square wave voltage is applied to the grid, and the bias is chosen so that the grid voltage varies between zero and beyond cut-off, as shown in Fig. 19.16.*b*. Whilst the valve current is cut off the condenser C charges through R from E_2. When the grid voltage changes to zero the anode slope resistance is effectively in parallel with C and, provided $r_a \ll R$, the condenser discharges rapidly to a voltage $v_O = E_2 - Ri_O$, where i_O is the anode current at zero grid voltage. The total range of anode voltage variation is limited to a small fraction of the possible range, and then the output voltage varies linearly with time. A common adaptation of this circuit is shown in Fig. 19.17, where the output is

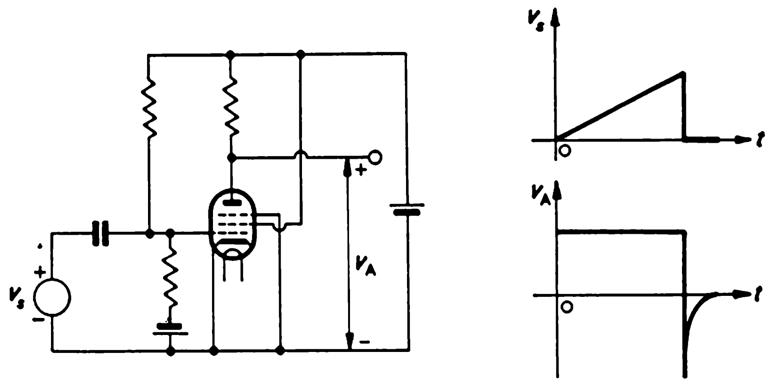


FIG. 19.18

applied to a cathode follower, which in turn feeds back to a tapping point on the load resistor of the triode amplifier. The cause of non-linearity in a condenser charging circuit is the decrease in charging current as the condenser voltage opposes the supply voltage. If the supply voltage is increased appropriately, then the charging current may be kept constant. This is what is attempted in the circuit of Fig. 19.17. As the condenser voltage rises so does the output voltage of the cathode follower, and this contributes to the charging of the condenser C. Thus the charging current is maintained very nearly constant and the circuit gives a high degree of linearity of output voltage against time. This circuit, usually called a Bootstrap circuit, is used as a single-stroke or repetitive time-base.

An electronic differentiating circuit and its waveforms are shown in Fig. 19.18. This circuit and also the Miller integrator can be considered as special cases of the parallel voltage feedback amplifier analysed in Section 10.9. (See Exx. XIX.)

CHAPTER 20

NOISE

20.1. Noise

Before the information in a small signal can be used the signal frequently has to be amplified considerably. Any output from the amplifier other than that due to the signal introduces extraneous information and is classed as noise. Noise may arise from many sources, and some of these have been mentioned in other chapters. Variations in the output may arise from the a.c. sources used to heat the valve cathodes or to produce the electrode supply voltages; this type of noise is usually referred to as mains hum. Other noises come from sources outside the amplifier and its associated equipment. A common example is radiation from sparks in motor car ignition systems or in the commutators of electric motors. Interference of this type may be reduced by taking suitable precautions in the offending equipment. There is one source of noise which is inevitable in all electrical apparatus, including amplifiers. This noise arises from the finite size and random movements of electrons. As a result of these random movements an electric current does not have a perfectly steady value but varies randomly about an average value. The variations are very small, but they set a limit to the size of the signal which may be usefully amplified. When the output voltage due to the signal is less than that due to the noise no more information can be obtained from the signal by further amplification.

Noise due to random electronic fluctuations is sometimes called fundamental noise. It may arise from resistances or valves, when it is referred to as Johnson or shot noise respectively.

20.2. Johnson Noise

In a conductor the electrons move through the crystal lattice and they continually exchange energy with the thermal vibrations of the ions in the lattice. As a result, there are random variations in the charge density of electrons, and hence variations in potential, through the conductor. When the conductor is part of an amplifier circuit these potential variations are amplified just as any other voltage across the conductor. It has been shown experimentally by Johnson that the mean-square value of the electronic voltage fluctuations across an open-circuited resistance R is given by

$$\overline{e^2} = 4kTRB$$

where k is Boltzmann's constant, T is the absolute temperature and B is the measured frequency bandwidth of the equipment which is used to observe the effect. It is noted that the mean-square noise voltage per unit bandwidth is independent of frequency. The formula is applicable to the effective resistance of any circuit, however complicated.

20.3. Shot Noise

Consider the flow of electrons from the cathode to the anode of a diode. As shown in Section 14.1, each electron during its flight induces on the anode a charge which varies from zero up to $+e$. The change of charge with time constitutes a pulse of current in the external circuit. The total current is the sum of the pulses due to all the electrons in transit at any instant. When the current is temperature-limited all the electrons reach the anode. Due to the random motion of the electrons in the cathode material, the number and velocity of the electrons emitted in equal time intervals fluctuate in a random manner about average values. Thus the anode current fluctuates also, and produces a noise voltage across the load in the diode circuit. This fluctuation is called shot noise. It has been shown experimentally that the mean-square noise current in a temperature-limited diode passing a current i_A is given by

$$\overline{i^2} = 2ei_A B,$$

where B is again the bandwidth of the equipment which is used to observe the effect. The mean-square current per unit bandwidth is independent of frequency.

In the temperature-limited diode each electron is assumed to move independently of the others. In the space-charge-limited diode this is no longer the case, and it is found that the effect of space charge is to "smooth out" the fluctuations, giving a considerable reduction in the noise current. The formula is now

$$\overline{i^2} = 2ei_A BF^2,$$

where F^2 is the space-charge reduction factor which is determined experimentally; it may be as low as 0·03, and is about 0·1 on the average.

The noise current in a space-charge-limited triode is given by the same formula, as long as the grid is at a negative potential and collects no electrons. In valves where the cathode current is collected by more than one electrode there is an increase in shot noise. This arises since the fraction in which the current divides between the electrodes fluctuates in accordance with the random transverse components of the electron velocities. In the case of a pentode with anode and screen currents the fluctuating anode current is given by

$$\overline{i^2} = 2ei_A B\left\{\frac{F^2 i_A + i_{G2}}{i_A + i_{G2}}\right\}.$$

The value of F^2 for a pentode is about 0·1 as in a triode. However, i_{G2} may be as much as 0·2 of the total current, and a pentode thus has more noise than a triode with the same anode current. When the cathode current is temperature-limited it is seen that the noise current is unaffected by the partition of current between screen and anode.

20.4. Addition of Noise Voltages

Noise arises from unrelated random events, and noise from one source can be treated independently of noise from a different source. The resultant effect is then obtained by the direct addition of the appropriate

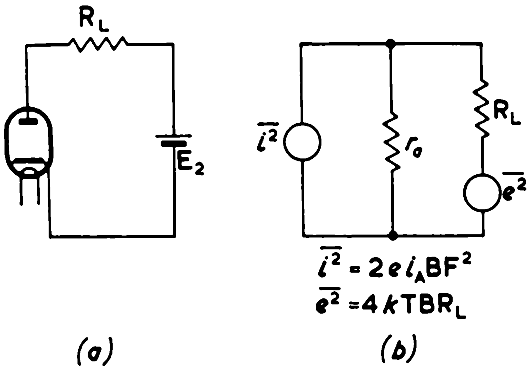


FIG. 20.1

mean-square values. As an example we may consider the diode with load resistance shown in Fig. 20.1.*a*. The diode shot noise may be represented by a current generator supplying its current in parallel with the diode slope resistance r_a. The noise in the load resistance R_L may be represented by a voltage generator in series with R_L. The resultant circuit is shown in Fig. 20.1.*b*. Due to shot noise the mean-square noise voltage between anode and cathode is

$$\overline{v_1^2} = \overline{i^2}\left(\frac{R_L r_a}{R_L + r_a}\right)^2 = \overline{i^2}\, R^2,$$

where R is equivalent to R_L and r_a in parallel. Due to resistance noise the corresponding mean-square noise voltage is

$$\overline{v_2^2} = \frac{\overline{e^2} r_a^2}{(R_L + r_a)^2} = \frac{\overline{e^2} R^2}{R_L^2}.$$

The resultant noise voltage is therefore given by

$$\overline{v^2} = \overline{v_1^2} + \overline{v_2^2} = R^2(\overline{i^2} + \overline{e^2}/R_L^2) = R^2 B(2ei_A F^2 + 4kT/R_L).$$

20.5. Equivalent Noise Resistance

In Fig. 20.2 a triode is shown with load resistance R_L. It may be verified, in most practical cases, that the noise voltage across R_L due to its own resistance noise is much less than the noise voltage across R_L due to the shot noise current of the valve. The latter may be evaluated from

Fig. 20.2.*b* exactly as in the previous section for the diode, and it is found that the output voltage is

$$\overline{v^2} = \overline{i^2}R^2 = 2ei_A BF^2R^2.$$

For some purposes it is convenient to express this noise voltage in terms of an equivalent noise voltage from a resistance between grid and cathode

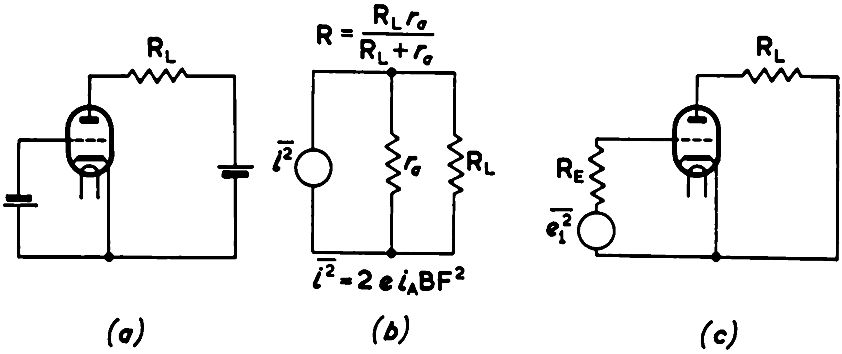


FIG. 20.2

and to assume a noiseless valve. If R_E is this resistance, then it produces an input noise voltage of

$$\overline{e_1^2} = 4kTBR_E.$$

After amplification this gives an output voltage of

$$\overline{v^2} = \frac{g_m^2\overline{e_1^2}}{\left(\dfrac{1}{R_L} + \dfrac{1}{r_a}\right)^2} = 4g_m^2 kTBR_E R^2.$$

The noise voltages are equivalent if

$$R_E = \frac{ei_A F^2}{2g_m^2 kT}.$$

Thus the shot noise of a valve may be represented by an equivalent noise resistance between its grid and cathode, as shown in Fig. 20.2.*c*. This is convenient for many purposes, since R_E can be determined by measurement, and then the noise performance of the valve may be compared with noise from any resistance in the actual input circuit. The formula for R_E shows that a low noise valve has a high g_m at low i_A.

20.6. Noise Factor

Any signal is always associated with a source of some internal resistance. For example, the aerial of a receiver has a radiation resistance. In the case of a signal generator there is the internal resistance of the generator. The resistance is a source of noise, and from the source there is therefore a definite signal-to-noise ratio. This may be expressed as a signal-to-noise power ratio $\overline{v_{is}^2}/\overline{v_{in}^2}$. When the signal is applied to the input of an amplifier as shown in Fig. 20.3 it is amplified giving an output voltage v_{os}. However, in the amplifier there is some noise arising from the effective parallel resistance of the input circuit, and there is also some shot noise from the first valve. If this valve has appreciable gain the noise of subsequent circuits and valves is negligible. In any case, it is obvious

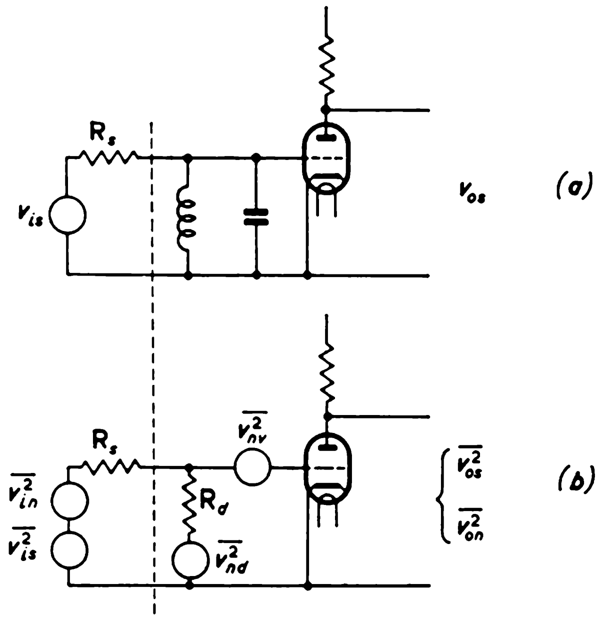


FIG. 20.3

that the amplifier has reduced the signal-to-noise ratio. If v_{on} is the noise voltage in the output of the amplifier, then the quantity N defined by

$$N = \frac{\overline{v_{is}^2}/\overline{v_{in}^2}}{\overline{v_{os}^2}/\overline{v_{on}^2}}$$

is called the noise factor of the receiver. It is usually expressed in decibels. An ideal receiver producing no noise would have a noise factor of 0 dB. The various signal and noise voltages in the circuit are shown in Fig. 20.3.*b*. In this figure R_s is the resistance of the signal circuit, R_d is the resistance of the input circuit of the amplifier, v_{nd} is the noise voltage associated with R_d, and v_{nv} is the noise voltage equivalent to the valve shot noise. In using the concept of equivalent noise resistance it must be realized that R_E has no physical existence, and it must not be included in the circuit in determining the actual voltages applied between grid and cathode.

20.7. Other Sources of Noise in Valve Amplifiers

The presence of small quantities of gas in valves may increase valve noise appreciably. Electrons ionize the gas atoms, producing more electrons and positive ions. The former flow to the anode and produce additional fluctuations, whilst the ions flow to the grid and cause fluctuations in grid current.

In some valves it is found that the mean-square noise current per unit bandwidth increases as the mean frequency is reduced below about 4 kc/s. This flicker effect, as it is called, is most notable with oxide-coated cathodes and especially under temperature-limited conditions, though it is still present when the current is space-charge-limited.

Finally, mention may be made of noise arising from mechanical vibration of the valve system, particularly in the early stages of a high-gain amplifier. This may cause variation in the valve currents which may be amplified and appear in the output. This type of noise can be minimized by using valves specially designed to reduce " microphony ", as it is called.

20.8. Transistor Noise

Transistors and crystal diodes produce Johnson noise like any other resistance. When they are passing steady currents additional noise is produced. Unlike shot noise in valves, transistor noise is not constant over the frequency band. It is greater at lower frequencies. Although, in some cases, transistors are inferior to vacuum valves as far as noise is concerned, inherently solid state valves have a distinct advantage over the thermionic valve. The latter requires a high temperature cathode as the electronic source. There is no such temperature requirement in crystal devices; indeed they may be operated at low temperatures where electron random movements are much reduced.

Examples II

1. Find the percentage increase in mass of an electron accelerated through a p.d. of 5,000 V.

(1 per cent.)

2. Calculate the transit time of an electron in the deflecting plates of a cathode-ray tube if the plates are 2 cm long and the final anode voltage is 1,000.

($1{\cdot}1 \times 10^{-9}$ sec.)

3. A cathode-ray tube has deflecting plates 2 cm long, 0·5 cm apart and 20 cm distant from the screen. If a deflecting p.d. of 40 V produces a spot deflection of 2 cm, calculate the approximate value of the final anode voltage. Point out any assumptions which you make.

(800 V.)

4. What is the shortest time it would take for a proton starting from rest to move between two points differing in potential by 1 V and separated by a distance of 1 cm? What factors might increase this time?

(1·4 μs.)

5. How would you show experimentally that electrons acquire energy in moving from a cathode to an anode under the influence of a p.d.?

6. The potential distribution between co-axial cylinders is given by the formulae $v = k \log(r/r_1)$ and $k = v_A/\log(r_2/r_1)$, where r_1 and r_2 are the radii of the cylinders and v_A is the potential difference between the cylinders. Draw a graph showing the variation of v in a cylindrical magnetron with anode radius of 1 cm and cathode radius of 0·025 cm. Hence show that with a magnetic field parallel to the axis the path of an electron in a magnetron with a thin wire cathode is very nearly circular.

7. A cylindrical magnetron has a filamentary cathode 0·2 mm in diameter and the anode diameter is 2 cm. If $v_A = 1{,}000$ V, calculate the approximate value of the magnetic flux density for an electron just to reach the anode. Explain any assumptions that you make.

(0·022 Wb/m^2.)

8. If 10^{16} electrons per second pass steadily along a 100V electron beam, find the beam current and the power dissipated at the collector.

(1·6 mA, 0·16 W.)

9. An electron with energy 400 eV moves in a uniform magnetic field of flux density 0·001 Wb/m^2, the field and the velocity being mutually perpendicular. Calculate the radius of the electron path.

(6·8 cm.)

10. A 1,000V electron moves in a uniform magnetic field of flux density

0·01 Wb/m^2, the electron velocity making an angle of 5° with the field. Calculate the path of the electron.

(Helix of radius 0·93 mm and pitch 6·7 cm.)

11. Discuss critically some of the differences in construction and use of cathode-ray tubes with electrostatic and magnetic deflection.

12. Determine the path of an electron which enters a uniform magnetic field of flux density B with a velocity v at right angles to the field. How is the path modified when the velocity and the field are inclined at an angle θ? Mention briefly one important practical example of each of these cases. [I.E.E., II, 1954.]

13. With the aid of a sketch show the essential parts of a cathode-ray tube with electrostatic deflection. Derive an approximate expression for the deflection sensitivity in terms of the deflecting voltage and the final anode voltage.

When there is no deflecting voltage how does the electron velocity vary between the final anode and the screen? (Give reasons.)

[I.E.E., II, October 1956.]

14. The deflecting plates of a cathode-ray tube are 3 cm long and 0·5 cm apart. The distance from the centre of the plates to the fluorescent screen is 20 cm. A deflecting potential difference of 100 V produces a spot deflection of 4 cm. Calculate the final anode voltage. Derive any formula that you use, and explain any assumptions you make in the derivation. What is the velocity of the electrons leaving the final anode? (1500 V, 2·3 × 10^7 m/s.) [I.E.E., II, October 1957.]

15. An electron of charge e and mass m is projected with velocity v into a uniform magnetic field of flux density B. If the direction of projection is normal to the field, determine the path of the electron.

Describe with the appropriate theory any experiment for the determination of e/m for an electron. [I.E.E., II, April 1956.]

16. A cathode-ray tube has deflecting coils which produce a uniform field of 6 × 10^{-3} Wb/m^2 when the coil current is 1 A (d.c.). This field extends an axial distance of 2·5 cm and its centre is 25 cm from the screen. If an alternating current of 0·25 A (r.m.s.) produces a trace 15 cm long on the screen, what is the final anode voltage of the tube? Prove any formulae used relating to electron ballistics. (2750 V.) [I.E.E., III, 1954.]

17. A high vacuum diode has a cylindrical anode of diameter 1 cm. The cathode, of very small diameter, is on the axis of the cylinder. The anode is maintained at a positive potential of 800 V relative to the cathode. What value of uniform axial magnetic field is required just to cause the anode current to be zero? Derive any necessary formulae and state clearly any assumptions made.

(1·35 × 10^{-2} Wb/m^2.) [I.E.E., III, October 1956.]

18. A cathode-ray oscillograph has a final-anode voltage of +2·0 kV with respect to the cathode. Calculate the beam velocity.

Parallel deflecting plates are provided, 1·5 cm long and 0·5 cm apart, their centre being 50 cm from the screen: (*a*) find the deflection sensitivity in volts applied to the deflecting plates per millimetre deflection at the screen; (*b*) find the density of a magnetic cross field, extending over 5 cm of the beam path and distant 40 cm from the screen, that will give a deflection at the screen of 1 cm. Prove all the formulae used.

($2{\cdot}7 \times 10^7$ m/s, 2·8V, $7{\cdot}6 \times 10^{-5}$ Wb/m^2.) [I.E.E., III, April 1956.]

19. Derive an expression for the electric field strength in the annular space bounded by two concentric cylinders when there is a potential difference between them.

An electron is injected with a certain velocity and at a certain radius into the evacuated space between the cylinders in a tangential direction. Determine the relation that must exist between electron velocity, cylinder radii and potential difference if the electron is to follow a concentric circular orbit. Calculate the potential difference required to give a circular orbit if the electron velocity is 10^7 m/sec and the relevant cylinder radii are 2 cm and 6 cm, respectively.

(The ratio of charge to mass of the electron is $e/m = 1{\cdot}76 \times 10^{11}$ coulomb/kg.) (630 V.) [I.E.E., III, April 1957.]

20. An electron moves with velocity 2×10^7 m/sec mid-way between and in a plane parallel to the electrodes of a planar magnetron. Calculate the p.d. between the electrodes. If the distance between cathode and anode is 0·5 cm, calculate the magnetic flux density. Indicate clearly any assumptions that you make.

(1,100 V, $1{\cdot}1 \times 10^{-4}$ Wb/m^2.)

21. Explain how an electron beam can be focused by: (*a*) a magnetic field; (*b*) an electric field. Describe briefly one application of each method, indicating the way in which the field is provided and its approximate magnitude.

The anode and cathode of a vacuum diode are parallel plates 1 cm apart. The cathode is at zero potential and the potential of the anode is given by $V = \sin 2\pi ft$ volt, where $f = 50$ Mc/s. At time $t = 0$ an electron is at rest near the cathode. Describe its subsequent motion and find: (i) its velocity at time $t = 2 \,.\, 10^{-8}$ sec; (ii) its maximum velocity.

(0, $1{\cdot}1 \times 10^5$ m/s.) [I. of P., 1957.]

EXAMPLES III

1. Compare the mechanism of electrical conduction in gases and semi-conductors.

2. What is the evidence for the existence of electron energy levels in matter?

3. Explain the difference between *n*-type and *p*-type semi-conductors and discuss the conditions at the junction between *p*- and *n*-type germanium.

4. Discuss the important differences and similarities between diamond and silicon.

5. Discuss the equilibrium potential, charge and energy conditions at the junction of: (*a*) two different metals; (*b*) two different semi-conductors.

6. Give an account of the electron theory of electrical conduction in solids. Explain the differences in the variation of conductivity with temperature in conductors, insulators and semi-conductors.

[I.E.E., II, April 1955.]

7. Write an account of the conduction of electricity in solids, referring particularly to the factors determining the electrical resistivity of metals, semi-conductors and insulators.

Explain what is meant by the work function of a surface, and describe briefly its importance in (*a*) thermionic emission, (*b*) metal rectifiers.

[I. of P., 1957.]

8. Write an account of the mechanism of the conduction of electricity in solids. [I. of P., 1954.]

9. If the conductivity of a semi-conductor is given by

$$\sigma = \sigma_0 \varepsilon^{-A/T}$$

where T is the absolute temperature, show that the change in conductivity caused by a small change in temperature is $A\sigma/T^2$ times the change in temperature.

(*Note.* If we write $A = ev/k$, then ev is related to the value of the energy gap between the top of the valency band and the bottom of the conduction band, i.e., E_G.)

Examples IV

1. Define the work function of a metal and show how it is related to the Fermi level.

2. Discuss the relative advantages and disadvantages of the more commonly used thermionic emitters.

3. How could you demonstrate experimentally that electrons are emitted from a hot cathode with a distribution of velocities?

4. Describe briefly the phenomenon of photo-emission.

The work function of the cathode of a photo-cell is 3·5 electron-volts. What is the maximum velocity of the emitted electrons when the cell is irradiated with light of frequency 4×10^{15} c/s? How could the maximum velocity of emission be determined experimentally?

($2{\cdot}1 \times 10^6$ m/s) [I.E.E., II, October 1955.]

5. Explain the meaning of the various symbols in Richardson's emission equation $I = AT^2\varepsilon^{-\phi/kT}$.

Describe and compare the main features of the various types of thermionic cathode which are in general use. [I.E.E., II, April 1956.]

6. In what way does the current from a vacuum photocell vary with the intensity of the incident radiation? How is the variation affected by the presence of gas in the cell?

The work function of barium is 2·5 eV. Would barium be suitable as a cathode in a photocell for violet light of wavelength 4,300 Å? (Give reasons.) [I.E.E., II, April 1957.]

7. How does the thermionic emission from a valve cathode depend upon: (*a*) the nature of the cathode surface; (*b*) the heater power?

Describe how the cathodes of modern valves are designed so as to minimize the heater power required for a given emission.

By what percentage will the emission from a tungsten filament at 2,400° C be changed by a change in temperature of 10° C? (8·1.) [I. of P., 1953.]

8. Explain what is meant by thermionic emission and describe how you would investigate its variation with temperature for a particular surface.

Explain briefly why, although the three common types of thermionic cathode have widely different emission efficiencies, all three are nevertheless in general commercial use.

By how many electron-volts must the work function of a surface change in order to reduce the emission from that surface at 2,400° C by 10 per cent? (+ 0·025.) [I. of P., 1955.]

9. Explain what is meant by "secondary emission" and describe how you would measure the secondary emission properties of a surface.

Discuss the importance of secondary emission in: (*a*) triodes and pentodes; (*b*) cathode-ray tubes; (*c*) photomultiplier tubes.

[I. of P., 1955.]

10. Describe a suitable model by means of which the emission of electrons from a metal surface may be described. State the condition under which thermionic emission, field emission and photo-electric emission will take place, and draw attention to common features and to differences in the three processes. [I. of P., 1956.]

11. Explain what is meant by: (*a*) secondary emission; (*b*) photo-electric emission.

Describe briefly the principles of operation of an electron multiplier photocell. Indicate suitable materials for the various component parts of the device and discuss its advantages and disadvantages compared with a vacuum photocell followed by a high-gain amplifier.

[I. of P., 1956.]

12. When monochromatic radiation of wavelength 2,000 Å falls upon a nickel plate the latter acquires a positive charge. The wavelength is increased, and at a wavelength of 3,400 Å the effect ceases, however intense the beam may be. Explain this, calculate the maximum velocity of the electrons emitted in the first case and describe, with a diagram and a circuit diagram, the construction and use of a practical photocell based on this effect. ($9{\cdot}5 \times 10^5$ m/s.) [I. of P., 1952.]

Examples V

1. A parallel-plane diode is operated at an anode voltage of 10 V. Calculate the velocity of an electron half-way between the cathode and anode when: (*a*) the current is space-charge-limited, and (*b*) temperature-limited. (Ignore initial velocities of the electrons.)

($1{\cdot}2 \times 10^6$ m/sec, $1{\cdot}3 \times 10^6$ m/sec.)

2. A p–n junction and a junction between two dissimilar metals both give a contact potential difference, but only the p–n junction can act as a rectifier. Explain this.

3. With the aid of potential distribution diagrams distinguish between temperature-limited and space-charge-limited current in a planar diode. State the Child–Langmuir formula for the space-charge-limited current density and explain how the formula is modified when the initial velocities of the electrons are taken into account.

4. The distance between the cathode and anode of a planar diode is d and the anode potential is v_A relative to the cathode. At what distance from the cathode is the potential equal to $v_A/2$ when a space-charge-limited current flows?

($0{\cdot}6\,d$.)

5. The anode current of a particular thermionic diode is given by

$$i_A = i_o \varepsilon^{Kv_A} \text{ when } v_A \text{ is negative.}$$

A resistance R is connected directly between the anode and the cathode. Calculate the voltage across the diode when $R = 1{,}000$ MΩ, $k = 11$ V^{-1} and $i_o = 60\mu$A.

($v_A = -1{\cdot}0$ V.)

(The load-line relation gives $i_A = -v_A/R$.)

6. The relation between current and voltage for a junction diode is given approximately by

$$i = i_o(\varepsilon^{Av} - 1).$$

Draw the characteristic when $i_o = 1\ \mu$A and $A = e/kT$ between $v = -1$ V and $v = 0{\cdot}5$ V. (e and k are the electronic charge and Boltzmann's constant, and T is the absolute temperature.)

7. When the value of v is sufficiently positive the characteristic of the junction diode is approximately

$$i = i_o \varepsilon^{Av}.$$

Show that under these conditions the a.c. conductance of the diode is proportional to the current.

8. The voltage–current characteristic of a diode valve is given for positive values of v_A by

$$i_A = 2 \times v_A^{3/2} \times 10^{-3}\text{ A}$$

and for negative values of v_A by $i_A = 0$.

A voltage, $v_s = 4 \cos \omega t$ is applied to the diode. Plot the variation of current through the diode for values of ωt from 0 to 2π.

A d.c. voltage of 3 V is applied to the diode in series with a small alternating voltage. Show that the ratio of the amplitude of the alternating current through the diode to the amplitude of the alternating voltage is about $5{\cdot}2 \times 10^{-3}$ A/V.

9. A parallel-plane diode has an anode–cathode clearance of 1 cm and a cathode area of 16 cm^2. Assuming space-charge-limitation and an anode voltage of 100 V, find:

(*a*) the current density;
(*b*) the total current;
(*c*) the number of electrons in the anode–cathode space;
(*d*) the average density of electrons.

((*a*) 2·3 mA/cm^2; (*b*) 37 mA; (*c*) $1{\cdot}2 \times 10^9$; (*d*) $7{\cdot}4 \times 10^7$ cm^{-3}.)
(*Note.* The density of free electrons in metals is of the order of 10^{22} cm^{-3}.)

10. Describe the formation of a space charge in a thermionic valve and explain: (*a*) why this modifies the distribution of electric field between the electrodes, and (*b*) its effect on the anode current.

11. Explain how electric current can be carried between cold electrodes in a gas at low pressure and describe how and why the nature of the discharge varies with the pressure in the gas. Why is such a discharge often accompanied by the emission of light and what factors determine the intensity and colour of the light emitted?

Mention two applications of gas discharge devices and indicate briefly, with circuit diagrams, how they are used. [I. of P., 1953.]

12. Discuss briefly any two of the following: (*a*) grid current; (*b*) space charge; (*c*) electron temperature. [I. of P., 1953.]

13. Explain, with reference to a typical diode valve: (*a*) cathode emission; (*b*) space charge; (*c*) saturation.

Calculate the space charge density at (i) the anode, and (ii) the cathode, of a plane parallel diode the plates of which are 5 mm apart, when the potential difference between them is 300 volts. Assume the current through the diode to be limited by space charge. What is the significance of the calculated value of the space charge density at the cathode?

($4{\cdot}7 \times 10^{-5}$ c/m^2, infinite.) [I. of P., 1954.]

14. Give a simple explanation in terms of energy-levels of the difference between conductors, intrinsic semi-conductors and impurity semi-conductors of the *p*-type and *n*-type. Using either *p*-type or *n*-type as your example, account for the rectifying action exhibited by an impurity semi-conductor in contact with a metal. What are the respective fields of application of point-contact and junction rectifiers?

[I.E.E., III, April 1956.]

15. A plane diode has an anode–cathode spacing of s metres. A sinusoidal voltage of maximum value V_m at a frequency $\omega/2\pi$ is applied between anode and cathode. Show that the distance travelled by an

electron released with zero initial velocity from the cathode at time t_o is, in the absence of space-charge:

$$x = \frac{qV_m}{\omega^2 ms}\,[(\omega t - \omega t_o)\cos\omega t_o + \sin\omega t_o - \sin\omega t].$$

If the anode–cathode spacing of such a diode is 0·5 mm, and the applied voltage has maximum value of 100 V at a frequency of 1,000 Mc/s, how long will it take an electron to reach the anode, assuming it to be released from the cathode with zero initial velocity at the commencement of a cycle? ($2{\cdot}5 \times 10^{-10}$ sec.) [I.E.E., III, April 1956.]

16. Sketch a graph showing how the current between two cold electrodes in an inert gas at low pressure varies as the voltage between the electrodes is increased.

Explain (giving skeleton circuit diagrams) how certain parts of the curve may be applied in the following ways:

(i) Detection of charged particles.
(ii) Voltage stabilization.
(iii) High-power rectification.

[I.E.E. III, October 1956.]

Examples VI

1. The anode current of a triode is given by the equation $i_A = f(v_G, v_A)$ When v_G and v_A change to $v_G + \delta v_G$ and $v_A + \delta v_A$ then i_A becomes $i_A + \delta i_A$. Taylor's Theorem for functions of two variables shows that

$$\delta i_A = \frac{\partial i_A}{\partial v_G}\,\delta v_G + \frac{\partial i_A}{\partial v_A}\,\delta v_A + \tfrac{1}{2}\left\{\frac{\partial^2 i_A}{\partial v_G^2}\,(\delta v_G)^2 + \frac{2\,\partial^2 i_A}{\partial v_G \partial v_A}\,\delta v_A \delta v_G + \frac{\partial^2 i_A}{\partial v_A{}^2}\,(\delta v_A)^2\right\} + .$$

Use this to show that

$$\delta i_A = g_m \delta v_G + \delta v_A / r_a$$

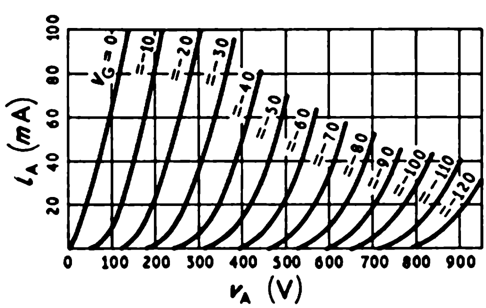


Fig. VI.i

provided that *either* δv_G and δv_A are small *or* the characteristics are free from curvature.

2. Use the previous example to prove that $\mu = -\left(\frac{\partial v_A}{\partial v_G}\right)_{i_A} = g_m r_a$.

3. Using the i_A, v_A characteristics of Fig VI. i, draw i_A, v_G and v_A, v_G characteristics for values of v_A from 0 to 500 V and v_G from 0 to $-$ 50 V. From the three sets of characteristics determine μ, g_m and r_a for (i) $v_G = -10$ V, $i_A = 60$ mA; (ii) $v_G = -30$ V, $i_A = 20$ mA; and (iii) $v_G = -50$ V, $i_A = 10$ mA. Verify in each case that $\mu \simeq g_m r_a$.

4. Draw typical static characteristics for a screen-grid tetrode and a pentode and explain in detail the reasons for their shapes.

5. Explain why the control grid current of a thyratron varies from negative to positive values as the grid voltage is made less negative.

6. Draw common-base hybrid characteristics for a transistor from the following data:

i_E(mA)	1·00	1·00	1·00	0·75	0·75	0·75
v_{CB}(V)	− 5·0	− 2·0	0·0	− 5·0	− 2·0	0·0
v_{EB}(mV)	135	139	142	127	129	131
i_C(mA)	− 1·00	− 0·99	− 0·97	− 0·75	− 0·74	− 0·74
i_E(mA)	0·50	0·50	0·50	0·25	0·25	0·25
v_{CB}(V)	− 5·0	− 2·0	0·0	− 5·0	− 2·0	0·0
v_{EB}(mV)	115	117	119	95	96	98
i_C(mA)	− 0·5	− 0·48	− 0·46	− 0·25	− 0·25	− 0·24

Find the values of the hybrid parameters at $i_E = 1$ mA, $v_{CB} = -4$ V and $i_E = 0{\cdot}5$ mA, $v_{CB} = -2$ V.

7. From the common-base transistor equation, changes in electrode currents and voltages are related by $\delta v_{EB} = h_{11b}\delta i_E + h_{12b}\delta v_{CB}$ and $\delta i_C = h_{21b}\delta i_E + h_{22b}\delta v_{CB}$. Using the conditions $i_E + i_B + i_C = 0$ and $v_{CB} = v_{CE} + v_{EB}$, show that

$$-\delta v_{BE}(1 - h_{12b}) = -h_{11b}(\delta i_B + \delta i_C) + h_{12b}\delta v_{CE}$$

and
$$\delta i_C(1 + h_{21b}) = -h_{21b}\delta i_B + h_{22b}(\delta v_{CE} - \delta v_{BE}).$$

Hence, by keeping δv_{CE} constant, i.e., $\delta v_{CE} = 0$, show that

$$h_{21e} = \frac{-h_{21b} + h_{21b}h_{12b} - h_{22b}h_{11b}}{1 + h_{21b} - h_{12b} - h_{21b}h_{12b} + h_{22b}h_{11b}}$$

$$\equiv \frac{-h_{21b} + h_{21b}h_{12b} - h_{22b}h_{11b}}{\Sigma} \quad \text{and} \quad h_{11e} = \frac{h_{11b}}{\Sigma}$$

i.e.,
$$\alpha_{cb} \simeq \frac{\alpha_{ce}}{1 - \alpha_{ce}} \quad \text{and} \quad h_{11e} \simeq \frac{h_{11b}}{1 - \alpha_{ce}}$$

since h_{12b} and $h_{22b}h_{11b}$ are both much less than 1 (see Table I, Chapter 6).

8. By setting $\delta i_B = 0$ in the previous example show that

$$h_{12e} = \frac{-h_{12b} - h_{21b}h_{12b} + h_{11b}h_{22b}}{\Sigma} \quad \text{and} \quad h_{22e} = \frac{h_{22b}}{\Sigma}$$

i.e.,
$$h_{12e} \simeq -h_{12b} + \frac{h_{11b}h_{22b}}{1 - \alpha_{ce}} \quad \text{and} \quad h_{22e} = \frac{h_{22b}}{1 - \alpha_{ce}}.$$

9. Discuss the physical reasons why the parameters h_{11}, h_{21} and h_{22} for the common-emitter connections are approximately equal to the corresponding common-base values multiplied by the factor $1/(1 - \alpha_{ce})$.

10. Define h-parameters for the common-collector connection using i_B and v_{EC} as the independent variables. Show that

$$h_{11c} \simeq \frac{h_{11b}}{1 - \alpha_{ce}}, \quad h_{12c} \simeq 1,$$

$$h_{21c} \simeq \frac{-1}{1 - \alpha_{ce}} \quad \text{and} \quad h_{22c} \simeq \frac{h_{22b}}{1 - \alpha_{ce}}.$$

11. To a first approximation the emitter current of a p–n–p transistor is given by

$$i_E = i_0 \varepsilon^{Av_{EB}} \text{ when } v_{EB} \text{ is positive.}$$

Show that the rate of change of v_{EB} with i_E is inversely proportional to i_E.

12. Explain how and why the grid voltage control action in a gas-filled relay valve differs from that in a high-vacuum triode. What is meant by the control ratio of a gas-filled relay and how does it depend upon: (*a*) the nature and pressure of the gas filling; (*b*) the anode voltage?

Why is it usual to include a resistor in series with the grid of a gas-filled relay? [I. of P., 1954.]

13. Describe the action of a triode valve and explain how the properties of the materials with which it is made, and the size and spacing of the electrodes, determine: (*a*) the maximum current and voltage which can be handled; (*b*) the highest frequency signal which can usefully be handled; (*c*) the maximum voltage and power amplifications; (*d*) the useful life of the valve.

Describe briefly how you would determine the anode characteristics of a pentode valve. [I. of P., 1955.]

14. Explain how the potential applied to the suppressor grid of a pentode valve determines the division of the cathode current between the anode and the screen-grid.

What are the advantages of the pentode over: (*a*) a triode; (*b*) a tetrode? Discuss the types of applications for which each of these three types of valve is most suitable. [I. of P., 1956.]

15. Describe the construction of a low-power pentode valve. Sketch the following static characteristic curves and explain their shapes with the aid of potential distribution diagrams:

(i) anode current/anode voltage;
(ii) anode current/suppressor-grid voltage;
(iii) screen-grid current/suppressor-grid voltage.

Sketch also characteristic (i) for a beam tetrode, and explain the differences. [I.E.E., III, April 1957.]

EXAMPLES VII

1. A single-stage amplifier has an anode load of 20 kΩ. Calculate the amplifier gain when the valve is: (*a*) a triode with $g_m = 2$ mA/V and $r_a = 10$ kΩ, and (*b*) a pentode with $g_m = 2$ mA/V.
(13, 40.)

2. The amplifier of Example 1 is coupled to another stage through a capacitor of 0·005 μF and a grid leak of 1 MΩ. If the total effective capacitance across the load is 50 μμF, calculate the frequencies at which the gain has dropped by 3 db with the triode and the pentode.
(32 c/s, 480 kc/s; 32 c/s, 160 kc/s)

3. A pentode amplifier has a load consisting of an inductor of 250 μH, a capacitor of 0·0001 μF and a resistor of 20 kΩ all in parallel. If the pentode has $g_m = 5$ mA/V, calculate the frequencies at which the amplifier has: (*a*) maximum gain, and (*b*) gain of 3 db below the maximum. What is the maximum gain?
(1·01 Mc/s, 0·97 Mc/s, 1·05 Mc/s, 100.)

4. For freedom from phase distortion in an amplifier show that the

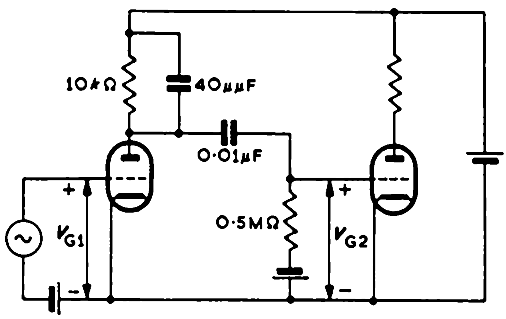


FIG. VII.i

phase angle delay produced by the amplifier should be proportional to the frequency.

5. For the circuit shown in Fig. VII.i calculate the " middle frequency " gain (i.e., the value of v_{g2}/v_{g1} when the reactances are negligible) when the first valve is a triode with $g_m = 5$ mA/V and $r_a = 10$ kΩ. Also calculate the two frequencies at which the gain has dropped to half the middle frequency value and the corresponding phase shifts.
(25, 18 c/s, 1·4 Mc/s, 240°, 120°.)

6. A pentode amplifier has an anode load consisting of a resistance R in parallel with a capacitance C. Show that the complex gain $\mathbf{A}$ at an angular frequency ω is given by

$$\mathbf{A} \simeq -\frac{g_m R}{1 + j\dfrac{\omega}{\omega_0}},$$

where $\omega_0 = 1/RC$, i.e., ω_0 is the " half-power angular frequency ".

7. If an inductance L is put in series with the load resistance R of Example 6, show that

$$\frac{\mathbf{A}}{g_m R} = -\frac{1 + jQ\left(\frac{\omega}{\omega_o}\right)}{1 - Q\left(\frac{\omega}{\omega_o}\right)^2 + j\left(\frac{\omega}{\omega_o}\right)}$$

where $Q = \omega_0 L/R$.

8. Use Example 7 to draw curves showing how the magnitude of the gain varies with frequency when $Q = 0$, 0·25, 0·5 and 1·0.

(*Note.* Amplifiers for wide frequency bands often use inductance compensation to extend the frequency range. A value of about 0·5 is usually chosen for Q.)

9. Show that the voltage gain of a cathode follower is given by the formula

$$A = g_m R/(1 + g_m R + R/r_a)$$

where R is the cathode load resistance.

10. A pentode with $g_m = 5$ mA/V has a signal of 0·1 V connected between its grid and earth. The output is taken across a resistance of 1,000 Ω connected between the cathode and earth. There is an h.t. supply between earth and the anode to give the required operating conditions. Calculate the output voltage.

(0·083 V.)

(*Note.* When a pentode is used in a cathode follower, then, in order to keep r_a large, the total voltage between screen grid and cathode must remain unchanged on application of the signal.)

11. An automatic bias circuit uses cathode resistance and capacitance of 400 Ω and 20 μF. If the valve current is $(10 + 5 \sin 2\pi ft)$ mA calculate the magnitudes of the direct and alternating voltages across the bias circuit when $f = 10$, 50 and 1,000 c/s.

(4 V, 1·8 V; 4 V, 0·7 V; 4 V, 0·04 V.)

12. Show that the input impedance of a common grid amplifier is

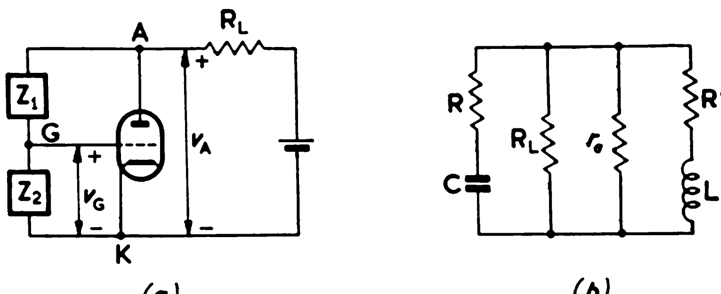


FIG. VII.ii

equal to $(\mathbf{Z} + r_a)/(\mu + 1)$ where $\mathbf{Z}$ is the load impedance between anode and grid.

13. In circuit of Fig. VII.ii.*a* show that $\mathbf{V}_g = \mathbf{V}_a\, \mathbf{Z}_2/(\mathbf{Z}_1 + \mathbf{Z}_2)$.

Hence, if $\mathbf{Z}_1$ is a resistance R and $\mathbf{Z}_2$ a capacitance C, show that the impedance between anode and cathode is equivalent to that shown in Fig VII.ii.*b*, where $R' = 1/g_m$ and $L = CR/g_m$.

14. Two identical triode valves are connected in series, the anode of one being connected to the cathode of the other. The two in series are connected across a steady h.t. voltage supply. The valve which is on the negative side of the h.t. supply is provided with an alternating voltage $\mathbf{V}_1$ connected between its grid and its cathode. The other valve has its grid connected to h.t. negative through a very large capacitance and to its own cathode through a very large resistance. Prove that the alternating component of anode current is given by

$$g_m(1 + \mu)\mathbf{V}_1/(2 + \mu).$$

Also show that the alternating voltage across the second valve, is

$$\mu\mathbf{V}_1/(2 + \mu).$$

15. We may write the characteristics of a triode valve when used in a common-cathode circuit as

$$i_G = f(v_G, v_A) \quad \text{and} \quad i_A = F(v_G, v_A),$$

so that for small changes

$$i_g = g_{11k}v_g + g_{12k}v_a \quad \text{and} \quad i_a = g_{21k}v_g + g_{22k}v_a.$$

Show that for the conventional negative grid triode

$$g_{11k} = 0, \quad g_{12k} = 0$$
$$g_{21k} = g_m \quad \text{and} \quad g_{22k} = 1/r_a.$$

16. We may write the characteristics of a triode valve when used in a common-grid circuit as

$$i_K = f(v_{KG}, v_{AG}) \quad \text{and} \quad i_A = F(v_{KG}, v_{AG}),$$

so that for small changes

$$i_k = g_{11g}v_{kg} + g_{12g}v_{ag} \quad \text{and} \quad i_a = g_{21g}v_{kg} + g_{22g}v_{ag}.$$

Show that for the negative grid triode

$$g_{21g} = -g_{11g} \quad \text{and} \quad g_{22g} = -g_{12g}$$

whilst $\quad g_{11g} = g_m + 1/r_a \quad \text{and} \quad g_{12g} = -1/r_a.$

The input resistance of the system is given by $r_i = v_{kg}/i_k$. Show that when the output is short-circuited so that $v_{ag} = 0$ then

$$r_i = \frac{1}{g_{11g}} = \frac{1}{(g_m + 1/r_a)}.$$

17. We may write the characteristics of a triode valve when used in a common-anode circuit as

$$i_G = f(v_{GA}, v_{KA}) \quad \text{and} \quad i_K = F(v_{GA}, v_{KA})$$

so that for small changes

$$i_g = g_{11a}v_{ga} + g_{12a}v_{ka} \quad \text{and} \quad i_k = g_{21a}v_{ga} + g_{22a}v_{ka}.$$

Show that for the conventional negative grid triode

$$g_{11a} = 0, \; g_{12a} = 0, \; g_{21a} = -g_m, \text{ and } g_{22a} = g_m + 1/r_a.$$

The effective internal resistance of the system is given by $r_0 = v_{ka}/i_k$ when the input terminals are shorted. Show that $r_0 = \dfrac{1}{g_{22a}} = 1/(g_m + 1/r_a)$.

18. A common-cathode pentode amplifier has an anode load consisting of resistance R_2 in series with resistance R_1. Across R_1 is a condenser C_1. Show that the gain at low frequencies is $\dfrac{R_1 + R_2}{R_2}$ times greater than at high frequencies.

(This can be used to give bass boost in a valve amplifier.)

19. Discuss in a qualitative manner the variation in stage gain of a triode common cathode amplifier as the load resistance is altered, keeping the high tension constant.

20. Explain why, over a large range of anode voltage, the anode current in a pentode valve is almost independent of the anode voltage. Explain also how it is possible to increase the anode slope resistance of a pentode without affecting its mutual conductance, and how this enables a large voltage amplification to be obtained with a pentode valve.

An amplifying valve has a grid leak of 1 MΩ. Neglecting all other forms of low-frequency attenuation calculate the minimum capacitance of the coupling condenser in order that the amplification at a frequency of 30 c/s shall be within 3 dB of that at mid-frequencies. Assume the load resistance of the previous valve to be small compared with that of the grid leak. (0·0053 μF.) [I. of P., 1954.]

21. Determine the expressions for the frequency response and phase characteristics of a conventional *RC*-coupled amplifier stage. The effects of anode and screen-grid decoupling and of the cathode by-pass capacitor are to be neglected.

Explain how the high-frequency response of such a stage may be extended. [I.E.E., III, October 1956.]

(See Examples 7 and 8 above.)

EXAMPLES VIII

1. A valve used as an audio-frequency power amplifier takes a quiescent current of 30 mA from an anode supply of 200 V. When a sinusoidal signal is applied to the grid the anode voltage varies from 40 to 360 V and the anode current from 50 to 10 mA. Calculate: (*a*) the power output;

(*b*) the amplifier efficiency; and (*c*) the turns ratio of the output transformer if the valve is to feed maximum power to a load of 20 Ω.

(1·6 W, 27 per cent, 20.)

2. Explain the difference between the d.c. and a.c. load lines in a transformer-coupled power amplifier.

A triode, whose characteristics are given in Fig VI.i, is to be used to supply power to a resistance load of 5 Ω using a supply voltage of 440 V. The anode dissipation must not exceed 20 W. Choose a suitable operating point and load line and determine: (*a*) maximum power output; (*b*) grid driving voltage; (*c*) transformer ratio; (*d*) percentage distortion.

3. Explain why a moving-coil meter is suitable but a moving-iron or a thermal meter is unsuitable for indicating distortion in a triode amplifier.

4. By considering anode characteristics, or otherwise, explain why a pentode gives higher gain and greater output than a triode of comparable size.

5. For the ideal transformer-coupled amplifier represented by Fig. 8.9 show that the amplitude of the grid signal for maximum output power is

$$\frac{E_2}{\mu}\left\{\frac{R + r_a}{R + 2r_a}\right\}.$$

6. With the notation of Section 8.6 show that the fundamental and second harmonic output powers are given respectively by the expressions $(P + N)^2R/8$ and $(P - N)^2R/32$, where R is the load resistance.

7. A push-pull amplifier has a sinusoidal signal of value $\hat{e}\sin \omega t$ supplied to each grid. The output transformer has n_1 turns on each half of the primary and n_2 secondary turns. Show that the power output is

$$P_o = \frac{2n^2R_L(\mu\hat{e})^2}{(r_a + 2n^2R_L)^2}$$

where R_L is the load resistance, $n = n_1/n_2$ and the operation is assumed to be linear.

8. Describe, with circuit diagrams, the principles of operation of choke-capacity-coupled, transformer-coupled and resistance-capacity-coupled amplifiers. How do the characteristics of each type vary with frequency?

In a low-frequency amplifier the input voltage is applied across a 600 Ω resistor in parallel with the grid and cathode of a triode: the valve has an amplification factor of 20 and an anode slope-resistance of 12,000 Ω. In the anode circuit there is a 15 : 1 transformer supplying a resistive load of 60 Ω. Calculate the overall gain in decibels. (7dB.) [I. of P., 1952.]

9. Describe briefly the push-pull method of amplification and list its advantages over the single-valve method.

The anode current/anode voltage characteristics of a triode are given in the table below. Two such valves are to be used in push-pull in the output stage of an audio-frequency amplifier feeding a non-reactive load of 500 Ω via a transformer. Each half of the primary winding of the transformer has 332 turns and the secondary winding has 100 turns. If the available h.t. supply is 400 V and the peak a.c. input signal is 18 V determine: (*a*) the grid bias required; (*b*) the maximum power output; (*c*) the overall voltage gain.

Grid volts	0	− 6	− 12	− 18	− 24
Anode volts					
120	17·5				
180	42	8			
240	80	20			
300	128	42	10		
360	180	80	24		
420		129	50	9	
480		178	88	28	2
540			138	57	12
600			180	100	32

(−18 V, 4·7 W, 12·2.) [I. of P., 1953.]

10. Explain carefully the purposes and properties of the second and third grids in a pentode valve.

How would you use a pentode valve or valves: (*a*) in an amplifier stage; (*b*) in a push–pull output stage?

In each case draw a typical circuit showing the use of the valve, and indicate the characteristics required of the valve and the circuit.

[I. of P., 1955.]

11. The anode current/anode voltage characteristics of a triode valve are given below. The valve is transformer-coupled to a resistive load of 13·3 Ω, the ratio of primary to secondary turns being 20, and is used with an h.t. supply of 400 V. The grid bias voltage is — 36 V, and a sinusoidal signal voltage of amplitude 30 V is applied to the grid. Calculate: (*a*) the percentage of second harmonic in the output voltage; (*b*) the fundamental output power; and (*c*) the efficiency. The resistance of the windings of the transformer may be neglected.

	Anode current, mA				
Grid voltage	− 6	− 18	− 36	− 54	− 66
Anode voltage					
50	5				
100	35				
150	90	5			
200	150	25			
250		68			
300		125	7		
350			25		
400			55		
450			100	10	
500				25	
550				50	11
600					25
650					50

(7·5, 4·2 W, 19 per cent.) [I. of P., 1956.]

12. Explain the principle and the advantages of push–pull operation in a valve amplifier. Each of the triodes used in a push–pull Class A power amplifier has a mutual conductance of 6 mA/V and an anode slope resistance of 1,500 Ω. The combined anode current of 120 mA is fed from a 350 V d.c. supply to the centre tap of a primary winding of an iron-cored transformer, the secondary winding of which is connected to a non-reactive resistor of 2·5 Ω resistance. The ratio of the number of primary turns to secondary turns is 40. Estimate the anode power efficiency of the stage when a sinusoidal signal of 20 V peak is applied to each valve. (Assume ideal linear characteristics for the valves.)

(12·6 per cent.) [I. of P., 1957.]

13. Show that, if two frequencies f_1 and f_2 are applied simultaneously to a non-linear impedance, component voltages having frequencies $f_1 \pm f_2$ are present in the output.

Discuss the importance of this result in telecommunication practice. [I.E.E., III, April 1956.]

14. A non-linear device has the current/voltage relation

$$i = a + bv + cv^2 + dv^3.$$

Show that the third-order term leads to cross-modulation between two amplitude-modulated sinusoidal signal voltages.

A carbon microphone, when subjected to a sinusoidal sound wave of frequency ω, has a resistance given by

$$100(1 + 0{\cdot}2 \sin \omega t)\ \Omega.$$

Find the percentage second-harmonic distortion current in its output.

(10.) [I.E.E., III, October 1956.]

15. Explain what is meant by the term non-linear distortion when applied to an amplifier.

The dynamic i_A/v_G characteristic of a triode valve with a resistive load may be represented by

$$i_A = A + Bv_G + Cv_G^2.$$

Deduce the ratio of the second-harmonic component in the output to the fundamental.

If the valve has a steady no-signal anode current of 60 mA and the application of a sinusoidal grid voltage causes the anode current to vary between 105 and 25 mA, calculate the percentage second harmonic in the output current. (6·3.) [I.E.E., III, April 1957.]

EXAMPLES IX

1. Using the transistor of Example VI.6, find the operating point for a common-base circuit with $v_{EB} = 110$ mV and a load resistance and a battery of -5 V between the collector and the base, when the load resistance is: (*a*) 2·5 kΩ; (*b*) 5 kΩ; and (*c*) 10 kΩ. Find the output voltage for signal changes of 10 mV, 20 mV and 30 mV, when the load resistance is 5 kΩ. Explain how the output could be made to follow the input more linearly.

2. Determine the approximate value of the power gain in Example 1.

3. The current gain of an amplifier A_i is defined as

$$A_i = \frac{\text{Current change in load}}{\text{Corresponding current change in the input circuit}}.$$

Show that for small signals the following formulae give the current gain for common-base, common-emitter and common-collector amplifiers respectively

$$A_i = \frac{h_{21b}}{1 + h_{22b}R_L} \simeq -\alpha_{ce},$$

$$A_i = \frac{h_{21e}}{1 + h_{22e}R_L} \simeq \alpha_{cb}$$

and

$$A_i = \frac{h_{21c}}{1 + h_{22c}R_L} \simeq \frac{-1}{1 - \alpha_{ce}}.$$

4. The output impedance of a common-base amplifier is defined as $r_o = v_{cb}/i_c$ when no signal is applied to the input. If the input circuit consists of a resistance R between the emitter and base show that

$$r_o = \frac{R + h_{11b}}{h_{11b}h_{22b} - h_{21b}h_{12b} + Rh_{22b}}.$$

5. Write down an expression for the output impedance of a common-emitter amplifier.

6. Show that the voltage gain of the common-collector amplifier shown in Fig. IX.i is given by

$$A = \frac{-h_{21c}R_L}{h_{11c} + R_L(h_{11c}h_{22c} - h_{12c}h_{21c})} \simeq 1.$$

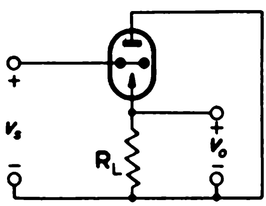


FIG. IX.i

7. Draw an equivalent circuit for a transistor based on the equations

$$v_{be} = h_{11e}i_b + h_{12e}v_{ce} \quad \text{and} \quad i_c = h_{21e}i_b + h_{22e}v_{ce}.$$

8. Common-base transistor characteristics may be expressed in the form

$$v_{EB} = f(i_E, i_C) \quad \text{and} \quad v_{CB} = F(i_E, i_C).$$

Use these expressions to show that, for small changes, v_{eb} and v_{cb} may be put in the forms

$$v_{eb} = r_{11b}i_e + r_{12b}i_c \quad \text{and} \quad v_{cb} = r_{21b}i_e + r_{22b}i_c,$$

where r_{11b}, r_{12b}, r_{21b} and r_{22b} are resistances, whose values may be found from the characteristics.

9. Show that the resistance parameters r_{11b}, r_{12b}, r_{21b} and r_{22b} are related to the h-parameters by the following equations:

$$r_{11b} = \frac{h_{11b}h_{22b} - h_{12b}h_{21b}}{h_{22b}},\ r_{12b} = \frac{h_{12b}}{h_{22b}},\ r_{21b} = -\frac{h_{21b}}{h_{22b}} \text{ and } r_{22b} = \frac{1}{h_{22b}}.$$

(Start with the transistor equations using h-parameters and use the definitions of the r-parameters. See Examples VI, 1 and 2.)

10. Use Example 8 to show that Fig. IX.ii, iii and iv are suitable equivalent circuits for a transistor,

where
$$r_e = r_{11b} - r_{12b} = h_{11b} - \frac{(1 + h_{21b})h_{12b}}{h_{22b}},$$

$$r_b = r_{12b} = \frac{h_{12b}}{h_{22b}},\ r_c = r_{22b} - r_{12b} = \frac{1 - h_{12b}}{h_{22b}} \simeq \frac{1}{h_{22b}},$$

$$\alpha = \frac{r_{21b} - r_{12b}}{r_{22b} - r_{12b}} = -\frac{h_{21b} + h_{12b}}{1 - h_{12b}} \simeq \alpha_{ce}.$$

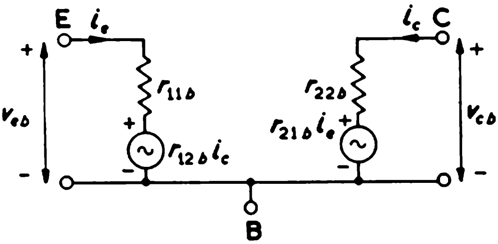


FIG. IX.ii

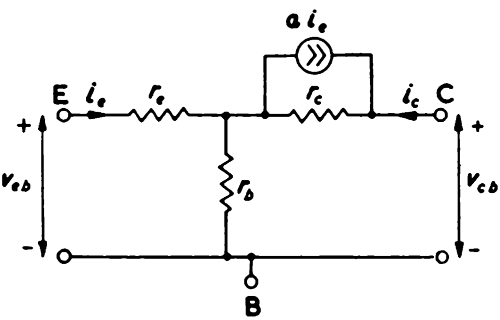


FIG. IX.iii

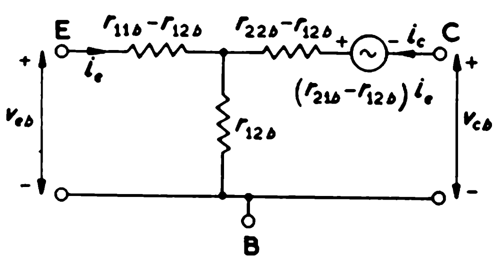


FIG. IX.iv

The circuits of Fig. IX.iii and iv are known respectively as the current generator and voltage generator T-equivalent circuits.

11. Find the values of r_e, r_b, r_c and α for the transistor of Example VI. 6.

12. Using the T-equivalent circuit, show that for a common-base transistor amplifier with load resistance R_L and resistance R between E and B

$$A = \frac{v_{cb}}{v_{eb}} = \frac{R_L(\alpha r_c + r_b)}{R_L(r_e + r_b) + r_e r_c + r_e r_b + r_b r_c(1 - \alpha)}$$

$$\simeq \frac{\alpha R_L}{r_e + r_b(1 - \alpha)},$$

$$r_i = v_{eb}/i_e = r_e + r_b \cdot \frac{R_L + r_c(1 - \alpha)}{R_L + r_b + r_c} \simeq r_e + r_b(1 - \alpha)$$

and

$$r_o = v_{cb}/i_c = r_b + r_c - \frac{r_b(r_b + \alpha r_c)}{R + r_e + r_b} \simeq r_c - \frac{\alpha r_b r_c}{R + r_e + r_b}.$$

13. The output of a common-emitter amplifier is connected to a parallel combination of resistance R_2 and condenser C_2 which is in series with the input of a second amplifier of input resistance R_1. The collector load of the first amplifier is a resistance R_3. Assuming that

$$v_{be} = h_{11e} i_b \quad \text{and} \quad i_c = \alpha_{cb} i_b,$$

find the ratio of voltage across the input of the second amplifier to that at the first amplifier.

(Such a circuit can be used to give treble boost in a low-frequency amplifier.)

$$\left[- \frac{\alpha_{cb}}{h_{11e}} \left\{\frac{R_1 R_3(1 + j\omega C_2 R_2)}{jR_2(R_1 + R_3)\omega C_2 + (R_1 + R_2 + R_3)}\right\}\right].$$

14. A common-emitter amplifier of input resistance R_1 is fed from a signal generator of resistance R_2. Across the input of the amplifier is connected a series circuit of resistance R_3 in series with a condenser C_3. Show that the gain of the amplifier varies with frequency according to

$$\mathbf{A} \propto \frac{R_1(R_3 + 1/j\omega C_3)}{(R_1 R_2 + R_2 R_3 + R_1 R_3) + (R_1 + R_2)/j\omega C_3}.$$

(This circuit can be used to give a bass boost in a low-frequency amplifier.)

15. Let i_{CO} be the collector current for a given transistor when the emitter current is zero. If α is assumed constant, then

$$i_C = - \alpha i_E + i_{CO}.$$

Hence show that the collector current is also given by

$$i_C = \frac{\alpha}{1 - \alpha} i_B + \frac{1}{1 - \alpha} i_{CO}.$$

(This shows how much more serious can be the effects of temperature in a common-emitter than in a common-base circuit, particularly where

$\alpha \simeq 1$ and the base current is determined mainly by the external circuit: i_{CO} varies considerably with temperature.)

16. A p–n–p transistor is connected in a common-emitter circuit. A by-passed resistance R_3 is included in series with the emitter lead to the positive terminal of the battery E_2. The base bias is obtained from a potentiometer chain R_1 to the negative and R_2 to the positive terminal of the battery. The base is connected to the centre point. If we assume that the resistance R_3 is such that the d.c. voltage across it is large compared with v_{BE}, show that when there is zero resistance in the collector lead

$$i_C = \frac{i_{CO}(R_1R_2 + R_2R_3 + R_3R_1) - \alpha R_2E_2}{R_1R_2(1-\alpha) + R_2R_3 + R_3R_1}.$$

17. In the previous example with $R_1 = 22$ kΩ, $R_2 = 4{\cdot}7$ kΩ and $R_3 = 1$ kΩ, find the rate at which the collector current changes with changes in i_{CO}.

With R_1 unchanged but R_3 equal to zero and R_2 removed so that i_B is constant, find the rate at which the collector current now changes with i_{CO} for $\alpha = 0{\cdot}98$.

(5, 50.)

18. Explain how a metal semi-conductor contact can act as a rectifier.

In a thermionic triode currents at a low power level in the grid circuit can be used to control larger currents in the lower impedance anode circuit. Describe briefly how the action of a transistor may be explained in similar terms, and point out the principal points of similarity and difference between the two devices. [I. of P., 1956.]

19. Assume that a transistor has ideal characteristics such that i_C is independent of v_{CE} and is proportional to i_B. If two such transistors operate in a Class A push–pull amplifier using a 12 V supply and giving 10 W output in a load resistance of 15 Ω, find the collector current, the collector dissipation and the transformer turns ratio.

(0·83 A, 10 W, 0·7 + 0·7 : 1.)

20. If, in the previous example, the operation is in Class B find the peak collector current.

(1·7 A.)

Examples X

1. Explain briefly in words why negative feedback makes the gain of an amplifier independent of variations in supply voltage.

A single-stage amplifier without feedback has a voltage gain of 10. A second amplifier, operated from the same power supply, has two stages each with a gain of 10, but there is negative feedback reducing the overall gain of the amplifier to 10. Calculate the percentage feedback. As a result of supply voltage variations the gain of the first amplifier drops to 9. What is now the gain of the second amplifier?

(9 per cent, 9·8.)

2. An amplifier, in the absence of feedback, has a gain which is liable

to fall by 40 per cent of its rated value as a result of uncontrollable variations of supply voltages. If, by the application of negative feedback, an amplifier is to be produced with a rated gain of 100 and with the requirement that the gain shall never fall below 99, determine the required initial gain of the amplifier in the absence of feedback.

(6,600.)

3. Show that the output impedance of a common-grid amplifier is $r_a + (\mu + 1)\mathbf{Z}_g$, where $\mathbf{Z}_g$ is the impedance between grid and cathode.

4. Show that the output conductance of the circuit in Fig. 10.10 is approximately equal to $g_m/2$ (assuming R_2 is very large and $R_1 = R_3$).

5. At sufficiently high frequencies the inductance of the electrode leads

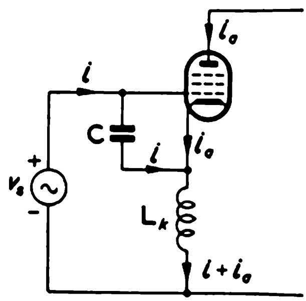


FIG. X.i

may be appreciable. In Fig. X.i L_k represents the inductance of the cathode lead and C is the grid-cathode capacitance of a pentode. Show that the input admittance is given by the expression

$$\frac{\mathbf{I}}{\mathbf{V}_s} = \frac{g_m\omega^2 L_k C - j\omega C(\omega^2 L_k C - 1)}{(g_m \omega L_k)^2 + (\omega^2 L_k C - 1)^2}.$$

If the operating frequency is not too high, show that the effect of the lead inductance is to introduce a conductance between grid and cathode of amount $g_m\omega^2 L_k C$.

6. With the aid of circuit diagrams explain current negative feedback. Discuss qualitatively its effect on the input and output impedance of an amplifier.

Calculate the voltage gain of a triode cathode follower stage in which $g_m = 2{\cdot}5$ mA/V, $r_a = 10{,}000\ \Omega$ and the cathode load resistor is $5{,}000\ \Omega$. Derive any formula that you use.

(0·9.)

7. Discuss in detail the purpose, operation, and design of a cathode-follower stage.

8. Discuss the effects of current and voltage negative feedback on: (*a*) the gain, input impedance and output impedance of an amplifier, and (*b*) the distortion produced by an amplifier.

If the gain of an amplifier without feedback is 90 dB, what must be the attenuation in the feedback loop if, with feedback, the gain is reduced to 60 dB? (60 dB.) [I. of P., 1952.]

9. Explain, with reference to the equivalent circuit, the inherent disadvantages of a triode valve having a large amplification factor, when used as an amplifier at high frequencies. Why is a tetrode superior for this purpose?

Calculate the mutual conductance of a pentode valve used in a single stage I.F. amplifier operating at 465 kc/s given that the voltage gain between grid and anode is 58·6 dB and that the anode tuned circuit consists of a 200 $\mu\mu$F condenser in parallel with an inductor having a Q of 100. (5 mA/V.) [I. of P., 1953.]

10. Describe what is meant by feedback and explain its effect with reference to: (*a*) a cathode-follower circuit; (*b*) the Miller effect.

A triode having an a.c. impedance of 10,000 Ω and an amplification factor of 30 is used in a cathode-follower circuit with a cathode load resistance of 200 Ω. Calculate the voltage gain of the circuit and the effective internal impedance. (0·37, 320 Ω.) [I. of P., 1953.]

11. Explain in physical terms the effects of current and voltage negative feedback in any system for the amplification of electrical signals. Consider the effects of both types of feedback on: (*a*) the amplification; (*b*) the effective internal resistance; (*c*) distortion.

Two identical triode valves are connected in parallel. The resistor in the anode circuit of the combination has a value of 25,000 Ω, and the value of the cathode bias resistor is 1,000 Ω. The grid bias is arranged so that the anode slope resistance of each valve is 10,000 Ω and the amplification factor is 30. Calculate how much of the cathode resistor must be by-passed to alternating current in order to increase the output impedance of the combination to 25,000 Ω. (355 Ω.) [I. of P., 1954.]

12. Explain in words the effects of current and voltage negative feedback in any system for the amplification of electrical signals. Refer particularly to: (*a*) input and output impedance; (*b*) frequency response; (*c*) distortion.

In a single-valve pentode amplifier the load resistance is 65,000 Ω, $g_m = 1{\cdot}75$ mA/V, the impedance of the valve is 0·95 MΩ, and the anode current is 3·15 mA. Draw the equivalent a.c. circuit and from it deduce the gain of the amplifier. Calculate the values of: (*a*) the cathode resistor necessary to reduce the gain to 25; (*b*) the bias resistor required for a grid bias of $-2{\cdot}2$ V. Sketch the complete circuit and indicate suitable values for components in the grid and cathode portion of the circuit. (2000 Ω, 700 Ω.) [I. of P., 1955.]

13. The anode resistor of a triode valve circuit has a resistance of 50,000 Ω, and so also has the resistor connecting the cathode to the negative terminal of the h.t. supply and earth. The grid is so biased that the anode slope resistance of the valve is 10,000 Ω and the amplification factor is 25. Calculate the approximate value of the internal resistance of (*a*) the anode circuit, (*b*) the cathode circuit, when each is used to drive a

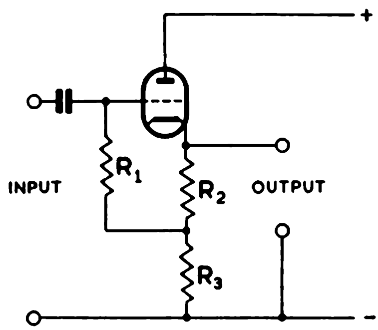
subsequent stage of amplification. The input voltage of the triode is applied between grid and earth. (1·3 MΩ, 2·3 KΩ.) [I. of P., 1956.]

14. Draw the equivalent circuit, including inter-electrode capacitances, for a simple triode amplifier feeding a resistive load, and derive an expression for the equivalent input impedance. Hence explain why such a circuit is unsuitable for the amplification of high-frequency signals. Why is a pentode more suitable than a triode for this purpose?

What is the effect, on the properties of the amplifier, of increasing the effective grid–anode capacitance by connecting a condenser between grid and anode? Indicate two applications of this type of circuit.

[I. of P., 1957.]

15. In a single-valve amplifier a fraction β of the output voltage is fed back to the input as negative feedback, and negative current feedback is also applied through a resistance R_f. The anode slope resistance of the valve is R_a and its amplification factor is μ. Calculate from first principles the output impedance of the amplifier.



R_1 = 2MΩ, R_2=400Ω, R_3 =15kΩ.

Fig. X.ii

Explain briefly why the non-linear distortion of the amplifier is reduced by the use of negative feedback.

Calculate the input impedance, output impedance and voltage gain of the cathode–follower stage shown in Fig. X.ii.

(*Note.* Neglect the effect of condenser.) [I. of P., 1957.]

16. Derive an expression for the shunt resistive component of the input impedance of an earthed-cathode amplifier valve due to the inductance of its cathode connection. Calculate the input resistance for a pentode valve having the following properties:

Capacitance between control grid and cathode = 5 μμF
Inductance of cathode connection = 0·03 μH
Mutual conductance = 8 mA/V

The frequency is 100 Mc/s.

(See Example 5 above.) (2000 Ω.) [I.E.E., III, April 1957.]

17. A triode amplifier has a resistance R_K between the cathode and earth. The anode load is a resistance R. The input is applied between the grid and earth, whilst the output is taken between the anode and earth. A by-pass condenser C is in parallel with the resistance R_K. Show that the ratio of the voltage amplification at high frequencies to that for d.c. signals is given by $1 + R_K g_m$, if $1/r_a$ is assumed negligible.

If $R = 20\ k\Omega$, $R_K = 100\ \Omega$, $g_m = 10$ mA/V, and $r_a = \infty$, find the value of C such that the stage gain at 5 kc/s is 1·5 times the d.c. stage gain.

(0·54 μF.)

(*Note.* This type of circuit can be used to give a treble boost in an audio amplifier.)

18. A pentode amplifier has an anode load resistance R_1, and the signal is applied in series with a resistance R_2 between the grid and cathode. The output voltage is taken between anode and cathode. A feedback path $\mathbf{Z}_3$, consisting of resistance R_3 in series with a condenser C, is connected between anode and grid. Show that at any frequency

$$\mathbf{A} = \frac{-\left(g_m - \dfrac{1}{\mathbf{Z}_3}\right) R_1}{1 + \dfrac{1}{\mathbf{Z}_3}(\mathrm{R}_1 + R_2 + R_1 R_2 g_m)}$$

if $1/r_a$ is assumed negligible.

Find the ratio of high-frequency gain to d.c. gain, when $R_1 = 20$ kΩ, $R_2 = 1{,}000\ \Omega$, $R_3 = 100$ kΩ and $g_m = 10$ mA/V.

(0·3.)

(This type of circuit can be used to give a bass boost in an audio amplifier.)

19. A transistor amplifier has a resistance R_1 between collector and earth, and a parallel combination of resistance R_2 and capacitance C_2 between emitter and earth. The input is applied between base and earth, and the output taken between collector and earth. Show that the voltage gain is given by

$$-\alpha_{cb} R_1 \Bigg/ \left\{ h_{11e} + \alpha_{cb}\left(\frac{R_2}{1 + j\omega C_2 R_2}\right)\right\}$$

if we assume that $v_{be} = h_{11e} i_b$ and $i_c = \alpha_{cb} i_b$.

(This feedback circuit can be used to give treble boost.)

Examples XI

1. Use static characteristics and load lines to show the transient variation of v_A and i_A in the first valve of the amplifier in Fig. 11.6.*a*.

(Through the Q-point draw a load line corresponding to the load resistance and the grid leak in parallel.)

2. Show that the initial value and the time constant of v_{g2} in Fig. 11.5.*b* are respectively $v_s R_g r_a/(R_L r_a + R_g r_a + R_L R_g)$ and RC, where R_L = load resistance, R_g = grid leak and $R = R_g + r_a R_L/(r_a + R_L)$.

3. For the circuit of Fig. 11.9.*a* show that the output voltage due to a sudden change e in grid voltage is $A(1-\varepsilon^{-t/T})$, where $A = -\mu eR/(R+r_a)$ and $T = Rr_aC_s/(R+r_a)$.

4. The anode load of a triode amplifier consists of resistance R and inductance L in parallel. If the grid voltage is changed suddenly by a small amount e, show that the anode current is $i_Q + g_m e(1 - A\varepsilon^{-t/T})$, where $A = R/(R+r_a)$ and $T = L(R+r_a)/Rr_a$. Show on anode characteristics how the anode current changes.

5. In the previous example show that the variation in anode voltage is given by $v_a = g_m eR'\varepsilon^{-R't/L}$, where $R' = Rr_a/(R+r_a)$.

6. A diode is switched in series with a battery E and an uncharged condenser C. The condenser has a parallel leakage resistance R. Using the characteristic of the diode, indicate qualitatively the pulse of current through the diode, and the pulse of voltage across the condenser.

(The steady-state condition is given by the cross-over of the diode characteristic and the load line $E = v_A + Ri_A$; the initial diode current is given by $v_A = E$.)

Show that the steady-state condenser voltage is equal to E if the diode current for $v_A = 0$ is E/R.

7. A diode is switched in series with a battery E and an inductance L of negligible resistance. Using the characteristic of the diode, indicate qualitatively the pulse of current through the diode.

8. A thermionic diode has a resistance R between anode and cathode. This parallel combination is suddenly switched in series with an uncharged condenser and a battery. Indicate qualitatively the change with time of the voltage across the diode, and show that the voltage across the condenser can finally exceed the battery voltage.

9. A common-cathode triode amplifier has an anode load of resistance R in series with inductance L. The d.c. grid voltage is suddenly changed by a small amount v_s. Show that

$$v_a = \frac{-g_m v_s}{\left(\frac{1}{R} + \frac{1}{r_a}\right)}\left\{1 + \frac{r_a}{R}\varepsilon^{-t/T}\right\},$$

where $T = L/(r_a + R)$.

10. A pentode amplifier ($r_a = \infty$) has a resistance R in series with an h.t. battery E_2 to earth. From the cathode to earth is a parallel combination of resistance R_K and condenser C. The voltage between the grid and earth is changed suddenly by a small amount v_s. Show that

$$i_a = \frac{g_m v_s}{(1 + g_m R_K)}\{1 + g_m R_k \varepsilon^{-t/T}\},$$

where
$$T = \frac{CR_K}{(1 + g_m R_K)}.$$

11. A common-cathode triode amplifier has a resistance load of R_1. It is coupled to a second amplifier through a condenser C in series with a

resistance R. Show that when the voltage on the grid is suddenly changed by a small amount v_s that the voltage across the resistance R is

$$-\frac{g_m v_s}{\left(\frac{1}{R_1}+\frac{1}{r_a}+\frac{1}{R}\right)}\varepsilon^{-t/T},$$

where $$T = C\left\{R + \frac{R_1 r_a}{R_1 + r_a}\right\}.$$

12. A common-cathode triode amplifier has the primary of a transformer as the anode load (inductance L_1, negligible resistance). If M is the mutual inductance between primary and secondary, show that, when the grid voltage changes by a small amount v_s, the secondary voltage is

$$\frac{\mu M v_s}{L_1}\varepsilon^{-t/T},$$

where $T = L_1/r_a$.

(*Note.* In Examples 9 to 12 find qualitatively, using the valve characteristics, the changes in currents and voltages in the cases where these changes are not small.)

13. A common-cathode pentode amplifier ($r_a = \infty$) has an anode load composed of C, R and L in parallel. Show that if the grid voltage changes by a small amount v_s and $T = 2CR \gg \sqrt{LC}$,

then $$v_a = -g_m v_s\sqrt{\frac{L}{C}}\,\varepsilon^{-t/T}\sin\frac{1}{\sqrt{LC}}t.$$

14. A common-cathode triode amplifier has an anode load resistance R_1 and a capacitance C between anode and grid. Show that, if the grid voltage changes by a small amount v_s,

$$v_a = -g_m R_1 v_s\left\{1 - \left(1 + \frac{1}{g_m R_1}\right)\varepsilon^{-t/T}\right\},$$

where $T = CR_1$.

15. A common-cathode pentode amplifier ($r_a = \infty$) has an anode load composed of resistance R in parallel with a capacitance C and with the primary L of a transformer. The secondary of the transformer (mutual inductance M) is in series with the d.c. grid supply between the grid and cathode. If the grid voltage changes by a small amount v_s, show that

$$v_a = -g_m v_s\sqrt{\frac{L}{C}}\,\varepsilon^{-t/T}\sin\frac{1}{\sqrt{LC}}t,$$

where $$\frac{1}{T} = \frac{1}{2C}\left\{\frac{1}{R} - \frac{g_m M}{L}\right\} \ll 1/\sqrt{LC}.$$

16. The table on next page gives some data on the static characteristics of a photo-electric cell:

Anode-cathode potential difference, V	0	10	25	250	*Amount of light falling on cathode, lumens*
Anode current, μA	0	0·75	0·8	0·9	0·02
	0	1·5	1·6	1·8	0·04
	0	3·0	3·2	3·6	0·08

(*a*) What type of cell is this, and for what applications is it particularly suited?

(*b*) The cell is used in the circuit shown in Fig. XI.i. The light falling on the cathode is mechanically chopped into " square " pulses such that

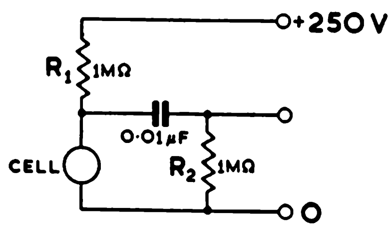


FIG. XI.i

it is constant at 0·08 lumen for 0·02 sec and is zero for 0·02 sec. Assuming that steady conditions have been reached, what are the maximum and minimum values of pulse-height developed across R_2?

(1·31 V, −1·31 V.) [I.E.E., III, April 1956.]

17. For applications involving transient phenomena it is often necessary to use an amplifier whose gain is constant over a wide frequency range, and whose phase-shift changes in proportion to the frequency. Explain simply why these characteristics are necessary, and show why it is difficult in practice to obtain them simultaneously with high gain.

[I.E.E., III, 1954.]

18. A single-stage amplifier is to use a pentode having an anode slope resistance of 1 MΩ and a mutual conductance of 2 mA/V. The amplifier is to feed a circuit of resistance 0·5 MΩ and shunt capacitance 12 $\mu\mu$F through a coupling capacitor. The gain is to be uniform, within 3 dB, from 20 c/s to 100 kc/s. If the anode–earth capacitance plus stray capacitance is 8 $\mu\mu$F, what load resistance is required, what middle-frequency gain can be achieved and what is the minimum value of coupling capacitor required?

If the input to this amplifier consists of rectangular pulses, estimate:

(i) The rise-time (to 90 per cent) of the output pulses.

(ii) The greatest pulse length that can be handled if the decay of the output pulse is not to exceed 10 per cent.

(104,000 Ω, 160, 0·016 μF, 3·7 μs, 22 ms.) [I.E.E., III, 1956.]

19. A valve having an amplification factor of 5 and a mutual conductance of 5 mA/V feeds a load circuit of inductance 25 henry and resistance 1,000 Ω. Negative feedback is applied by a non-bypassed resistor of 1,000 Ω in the cathode circuit. By what value will the anode current increase if a positive step signal of 35 V is applied to the grid circuit, and how long will it take to reach 90 per cent of its final value?

(21·9 mA, 7·2 ms.) [I.E.E., III, 1954.]

EXAMPLES XII

1. In the direct-coupled amplifier of Fig. 12.1 show that a small change of amount e_2 in the battery voltage E_2 would give a change in output voltage of

$$\left\{\frac{\dfrac{1}{R} - \dfrac{g_m}{1 + R/r_a}}{\dfrac{1}{R} + \dfrac{1}{r_a}}\right\} e_2,$$

when both valves are identical and have the same load resistance R.

2. For the amplifier in Fig. 12.4 show that the change in output arising from a change e_{12} in the supply E_{12} is approximately equal to

$$\frac{RR_1e_{12}}{R_K(R_1 + R_2)}.$$

For a change e_2 in E_2 show that the output changes by

$$-\frac{RR_2e_2}{R_K(R_1 + R_2)}.$$

3. Explain why the performance of the circuit in Fig. 12.7 may be improved by replacing R_K by a suitably biased pentode.

4. Draw the circuit diagram of a cathode-coupled amplifier stage and explain its operation. Show how such a circuit can be used as a difference-amplifier and derive an expression for the gain in this case.

5. What is meant by " drift " in connection with direct-coupled amplifiers, and why is it so serious in high-gain d.c. amplifiers, whereas it is not usually troublesome in equally high-gain systems using capacitive or transformer inter-stage coupling? Mention the main causes of drift, and explain how their effects can be minimized.

[I.E.E., III, October 1956.]

6. The diagram of Fig. XII.i shows the circuit of a d.c. valve voltmeter. Describe the manner of operation of the circuit and explain the function of each component. What are the disadvantages of this type of circuit for this purpose?

E — 100 V, R_1 — 6,500 Ω, R_2 — 3,500 Ω, R_3 — 500 Ω, R_4 — 9,500 Ω.

Meter resistance 300 Ω. Valve μ — 19, g_m — 2·1 mA/V.

Find: (*a*) the anode current when the meter reads zero; (*b*) the d.c. input required to produce a current of 1 mA through the meter.

(3·5 mA, 0·85 V.) [I. of P., 1955.]

7. Two identical triodes, with identical anode resistors but with a common-cathode resistor, are connected across an h.t. supply. Analyse

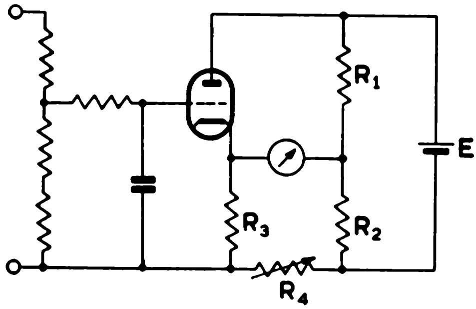


FIG. XII.i

the action of this circuit and explain how the anode voltages vary when a sinusoidal voltage is applied to one grid, the potential of the other grid remaining constant. Why is it desirable that the resistance in the cathode circuit should be as great as possible?

Indicate possible uses for this type of circuit. [I. of P., 1956.]

8. Discuss the difficulties which are encountered in the construction and operation of a direct current amplifier, and explain how these may be minimized.

Give an example of a situation in which it is possible to modify apparatus which would normally yield a direct current response so that an alternating current output can be obtained. [I. of P., 1957.]

EXAMPLES XIII

1. Sketch a typical anode current/anode voltage curve for a tetrode when the grid and screen voltages are fixed. Explain fully the nature of the curve. Show how a tetrode may be used as an oscillator.

2. In the vector diagram of Fig. 13.7 add vectors representing $\mathbf{I}_a$ and $\mu\mathbf{V}_g$, and explain the conditions to be satisfied for oscillation.

3. In the phase shift oscillator of Fig. 13.3 show that the frequency of oscillation is given by the formula

$$= \frac{1}{2\pi RC\sqrt{6}}. \quad \text{(Assume } R_1 \ll R.\text{)}$$

4. In the phase shift oscillator of Fig. 13.3 show that, for oscillation, the amplifier voltage gain must be at least 29.

5. Draw a phase shift transistor oscillator corresponding to the triode oscillator in Fig. 13.3.

6. It is shown in Section 13.6 that both circuits of a tuned-anode–

tuned-grid oscillator must be tuned to resonant frequencies above the frequency of oscillation. Use this result to explain why this type of oscillator is not suitable for operation at very high frequencies.

7. When resistive components are taken into account the generalized triode circuit of Fig. 13.10 takes the form shown in Fig. XIII.i, where the

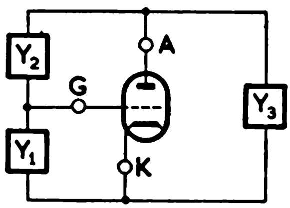


FIG. XIII.

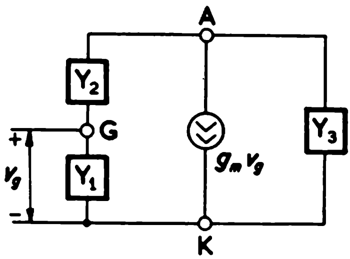


FIG. XIII.ii

$\mathbf{Y}$'s are admittances which may be expressed in the form $\mathbf{Y} = G + jB$, where G is a conductance and B is a susceptance. Using the equivalent circuit of Fig. XIII.ii and assuming that an oscillator is an amplifier which provides its own input, show that the condition for oscillation is

$$\mathbf{Y}_1\mathbf{Y}_2 + \mathbf{Y}_2\mathbf{Y}_3 + \mathbf{Y}_3\mathbf{Y}_1 + \mathbf{Y}_2 g_m = 0.$$

(*Note.* The anode slope conductance of the valve g_a is included in $\mathbf{Y}^3$.)

8. By separating the $\mathbf{Y}$'s into their real and imaginary parts show that the condition for oscillation of the last example reduces to two separate and necessary conditions

$$(G_1G_2 + G_2G_3 + G_3G_1) - (B_1B_2 + B_2B_3 + B_3B_1) + g_mG_2 = 0$$

and $$G_1(B_2 + B_3) + G_2(B_3 + B_1) + G_3(B_1 + B_2) + g_mB_2 = 0.$$

Use these two equations and the knowledge that all the conductances are positive to show that B_1 and B_3 must have the same sign and B_2 the opposite sign.

9. If the circuit resistances in Example 8 are negligible (i.e., all the conductances are zero except g_a and g_m) show that the condition for oscillation becomes $B_3/B_1 \geqslant 1/\mu$.

10. Derive the condition for sustained oscillation, and the exact value of the oscillation frequency, in a circuit consisting of a coil (L, R), a capacitance (C), and a negative resistance ($-r$), all connected in parallel.

11. Compare briefly, with reference to the electric field in the valves, the action of: (*a*) a tetrode; (*b*) a pentode; (*c*) a beam tetrode.

Show that sustained oscillations may be obtained with any device for which the voltage–current characteristic has a negative slope and describe briefly an oscillator based on this principle. [I. of P., 1952.]

12. Describe two possible circuit arrangements for the maintenance of continuous oscillations in an oscillatory circuit by a triode valve, and

derive the necessary initial conditions for the continuous increase of oscillatory amplitude in one of them.

How is the amplitude of oscillation limited by the characteristics of the valve? [I. of P., 1954.]

13. It can be shown that all single-valve feedback oscillators can be represented by the basic circuit of Fig. XIII.iii (a.c. conditions only).

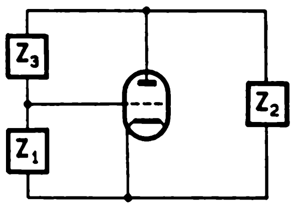


FIG. XIII.iii

For oscillation to be possible either of the following sets of conditions must be fulfilled:

$\mathbf{Z}_1$ and $\mathbf{Z}_2$ inductive; $\mathbf{Z}_3$ capacitive.

$\mathbf{Z}_1$ and $\mathbf{Z}_2$ capacitive; $\mathbf{Z}_3$ inductive.

Illustrate each of these criteria with a circuit diagram (showing supply connections) of a practical oscillator arrangement. What properties of the valve limit the highest frequency at which such circuits will operate, and which of the two basic arrangements is preferable when the highest possible frequency is required? [I.E.E., III, April 1956.]

14. State the conditions for self-oscillation of a valve-amplifier circuit, and explain the operation of the oscillator shown in Fig. XIII.iv.

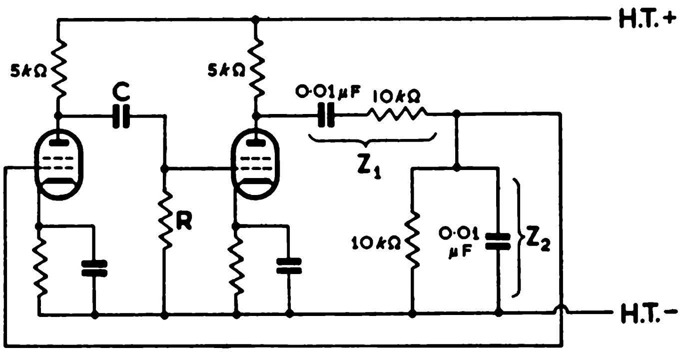


FIG. XIII.iv

Assuming valves of identical characteristics and high slope resistance, determine the frequency of oscillation and the least value of mutual conductance for the valves. The effects of C and R are to be neglected.

(1·30 kc/s, 0·32 mA/V.) [I.E.E., III, October 1956.]

15. A direct-coupled amplifier consists of three identical resistance-

loaded stages in cascade, and has an overall gain–frequency characteristic (relative to zero frequency) as follows:

Frequency, kc/s	0	1	10	50	100	200	400	800	1,600
Change in gain, dB	0	0	3·6	9	21	37	54	72	90

Voltage gain at zero frequency = 8,000. A fraction of the output voltage is tapped off the load resistance of the last stage and returned without modification to the input circuit of the first stage, so that the feedback at low frequencies is negative. It is found that oscillation occurs if the voltage feedback fraction exceeds a certain value. Determine this value, and the frequency of oscillation.

(10^{-3}, 88 k/cs.) [I.E.E., III, April 1956.]

16. In terms of the quantities shown in the skeleton circuit diagram (Fig. XIII.v), deduce expressions for: (*a*) the minimum mutual inductance

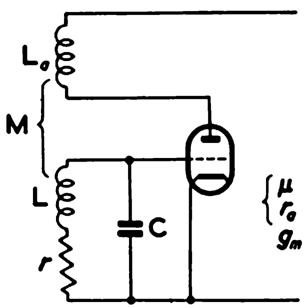


FIG. XIII.v

required for oscillation to commence, and (*b*) the frequency of oscillation. What method might be used, in a practical oscillator, to ensure reasonable amplitude stability of the output? [I.E.E., III, October 1956.]

17. Describe with the aid of a circuit diagram a valve oscillator suitable for generating frequencies in the audio-frequency range. Give a physical explanation of the generating process and derive an expression for the conditions determining the inception of oscillations.

[I.E.E., III, October 1956.]

18. Draw a circuit diagram for a tuned-anode valve oscillator and explain its action. Sketch a vector diagram for the equivalent circuit and derive an expression in terms of the circuit constants for the least value of the mutual inductance necessary for oscillations to commence.

[I.E.E., III, April 1957.]

19. The a.c. circuit of a transistor is equivalent to a resistance R_E between emitter and earth, and a resistance R_C between collector and earth. If we assume that for small signals

$$v_{eb} = h_{11b} i_e \quad \text{and} \quad i_c = -\alpha_{ce} i_e,$$

show that the a.c. resistance between the base and earth is

$$\frac{h_{11b} + R_E}{1 - \alpha_{ce}}.$$

(If the transistor has a value of α_{ce} greater than unity, then the a.c. resistance is negative.)

20. The same transistor is in an a.c. circuit which can be represented by a resistance R_B between base and earth, whilst there is a resistance R_C between collector and earth. Show that if

$$\alpha_{ce} > \left(1 + \frac{h_{11b}}{R_B}\right)$$

the a.c. resistance between emitter and earth is negative.

21. Two identical triodes have their cathodes connected together, and at the frequency under consideration the anode of each triode is effectively connected to the grid of the other triode. Show that the value of the a.c. resistance between the two anodes is

$$\frac{2}{(1/r_a - g_m)}$$

if grid current is negligible.

(Note that $i_{a1} = - i_{a2}$.)

Examples XIV

1. Show that the movement of a charge between two planar electrodes under the influence of a uniform field produces a saw-tooth pulse of current in the external circuit.

2. Indicate approximately the effect of space charge on the shape of the current pulse in the external circuit in Example 1.

3. The constant current leaving the cathode of a temperature-limited planar diode is i_0 and the anode voltage is $v_0 + \hat{v} \sin \omega t$. If d is the anode–cathode distance show that at time t the electron velocity and displacement are given by

$$\frac{dx}{dt} = \frac{ev_0}{md}(t - t_0) - \frac{e\hat{v}}{\omega md}(\cos \omega t - \cos \omega t_0)$$

and $x = \dfrac{ev_0}{2md}(t - t_0)^2 + \dfrac{e\hat{v}}{\omega md}(t - t_0) \cos \omega t_0 - \dfrac{e\hat{v}}{\omega^2 md}(\sin \omega t - \sin \omega t_0)$,

where it is assumed that the electron leaves the cathode at time t_0 with zero velocity, and space-charge effects are neglected.

Using the expression $\frac{1}{2}m\left(\frac{dx}{dt}\right)^2$ for the kinetic energy and $\frac{evx}{d}$ for the potential energy, find, in terms of t_0 and T, the difference between the K.E. and P.E. at the anode of an electron which reaches the anode at time $t = t_0 + T$, where T is the transit time.

Since $\frac{i_0 dt_0}{e}$ is the number of electrons leaving the cathode in time dt_0, the integral

$$P = \frac{\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}} \frac{i_0}{e}\,(\text{K.E.} - \text{P.E.})\,dt_0$$

gives the additional power consumed by the electrons at the expense of the h.f. field. If $v_0 \gg \hat{v}$, so that T is the same for all electrons, show that

$$P = \frac{i_0\hat{v}^2}{v_0}\left\{\frac{2 - 2\cos\omega T - \omega T\sin\omega T}{(\omega T)^2}\right\}.$$

4. If the additional source of power consumed by the electrons in the previous example is represented by a conductance g in parallel with the diode show that

$$\frac{g}{g_0} = 2(2 - 2\cos\omega T - \omega T\sin\omega T)/(\omega T)^2,$$

where $g_0 = i_0/v_0$.

Draw a graph of g/g_0 against ωT and hence explain why a diode can be used as a negative resistance oscillator at very high frequencies.

Examples XV

1. Discuss the interchange of energy between electric fields and moving charges.

Explain qualitatively the growth of the amplitude of the wave along the helix of a travelling-wave tube.

2. The mean voltage of a klystron resonator is v_0 and the bunching voltage is $\hat{v}_1 \sin\omega t$. If $\hat{v}_1 \ll v_0$, show that the velocities of the electrons leaving the buncher are given approximately by the equation

$$u = u_0\left(1 + \frac{\hat{v}_1}{2v_0}\sin\omega t\right),$$

where u_0 is the velocity corresponding to v_0.

3. Show that the distance–time diagram for an electron in the resonator–reflector space of a planar reflex klystron is parabolic, and hence show how bunching occurs.

4. Explain why the collector in a double resonator klystron is sometimes operated at a voltage below the resonator voltage.

5. On a diagram of a planar magnetron similar to Fig. 15.8 show the electron distribution.

6. Explain briefly how oscillations are maintained in: (*a*) a Hartley oscillator; (*b*) a magnetron; (*c*) a multivibrator; (*d*) a dynatron oscillator; and (*e*) a resistance–capacity oscillator.

A triode with a mutual conductance of 1·5 mA/V is to be used as an oscillator. The parallel-tuned grid circuit has an equivalent series

resistance of 25 Ω and a total capacitance of 200 μμF. Calculate the minimum value of the mutual inductance between the untuned anode coil and the grid coil in order that oscillations may just be maintained.
(3·3 μH.) [I. of P., 1957.]

7. A diode has a cylindrical anode of radius a metres and a cathode of radius b metres ($a > b$). A constant potential difference of V volts is maintained between them, and the axis of the diode lies along a uniform magnetic field of flux density B webers per square centimetre.

Calculate the value of B which just allows electrons leaving the cathode radially with zero initial velocity to reach the anode.

Explain briefly the basic principles of the operation of the cavity magnetron oscillator. [I.E.E., III, October 1956.]

8. Show how velocity modulation of an electron beam can produce bunching in a drift space. Show the application of this effect to: (i) an amplifier, and (ii) an oscillator suitable for very high frequencies.

Upon what factors does the operating frequency of these devices depend?
[I.E.E., III, April 1957.]

Examples XVI

1. A diode has the following values of v_A and i_A:

v_A, V	0	5	10	15	20	25	30	35	40	60
i_A, mA	0	4	10	20	32	50	72	96	120	240

Find graphically the anode current when a sinusoidal alternating voltage of peak value 50 V is applied to the diode.

2. In the rectifier circuit of Fig. 16.3 show that the d.c. power output in the load is $\hat{v}^2R/\pi^2(R_0 + R)^2$, the power dissipated in the diode is $\hat{v}^2R_0/4(R_0 + R)^2$ and the efficiency (i.e., d.c. power in load/power taken from source) is

$$\frac{4}{\pi^2(1 + R_0/R)}.$$

3. Show that in a condenser-input half-wave rectifier the voltage across the load is given approximately by the expression $\bar{v}_C = \hat{v} - \bar{i}/2fC$, where the symbols have the same meanings as in Section 16.6.

4. Compare the relative merits of condenser-input and choke-input rectifiers.

5. Draw a diagram similar to Fig. 16.11 to show how the voltage v_{C2} builds up across the condenser C_2 in the voltage-doubling circuit of Fig. 16.14.

6. Explain how the thyratron circuit shown in Fig. XVI.i may be used to control a unidirectional current through the load R. Draw waveforms of the anode voltage and anode current.

7. A certain gas diode has a striking voltage of 105 V and over the

voltage range 72 to 78 V the diode current varies from 5 to 45 mA. Show that this diode may be used to supply a load current from 0 to 40 mA whilst the voltage remains constant at 75 V $\pm$ 4 per cent.

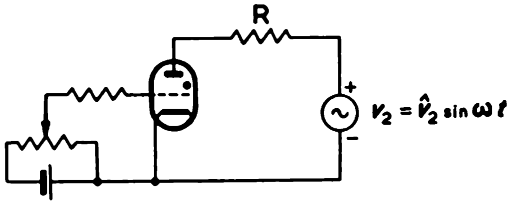


FIG. XVI.i

A diode with similar characteristics is used to maintain a stable voltage of 75 V across a resistance of 5,000 Ω using a nominal 200 V d.c. supply, with the circuit shown in Fig. 16.20.*b*. Calculate the value of R and show that the d.c. supply voltage may vary by 30 per cent for a 4 per cent voltage variation across the load. What is the minimum value of the supply voltage for switching on?

(3,120 Ω, 170 V.)

8. A half-wave diode rectifier feeds a load of resistance 10,000 Ω, across which is connected a 1 μF capacitor. The supply is sinusoidal, of frequency 50 c/s, and the diode has zero resistance when conducting and infinite resistance when non-conducting. Determine the instants in the cycle at which diode conduction commences and ceases.

(280° and 380° from peak) [I.E.E., III, October 1956.]

9. Explain the action of the two diode voltmeter circuits in

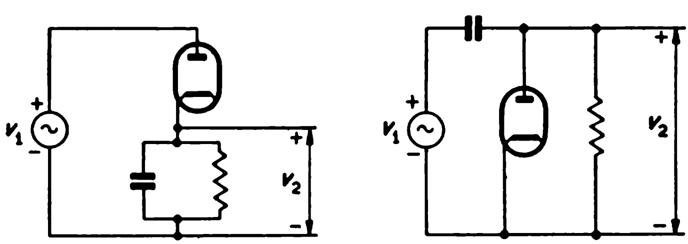


FIG. XVI.ii

Fig. XVI.ii and indicate the type of application for which each is suitable. If the output voltage from these circuits is to be amplified before measurement, sketch suitable direct current amplifiers for use with each type of circuit. [I. of P., 1952.]

10. Compare the actions of a thyratron and a discharge-tube voltage stabilizer.

Sketch: (*a*) a circuit using a thyratron to control the opening and closing of a relay in response to an external signal, and (*b*) a circuit using a discharge-tube voltage stabilizer. Comment on the performance of the voltage stabilizer and indicate briefly how any residual voltage fluctuations might be eliminated. [I. of P., 1952.]

11. Explain, with circuit diagrams, the action of two types of full-wave rectifier circuit suitable for supplying an X-ray tube at 200 kV from a 50 c/s supply. Describe the essential features of the components used and explain briefly how the voltage across the X-ray tube could be measured.

If the X-ray tube current is 10 mA, what must be the capacity of the smoothing condensers if the ripple is not to exceed 0·5 per cent?

(0·1 μF.) [I. of P., 1953.]

12. The circuit diagram of a series–parallel voltage-stabilizer circuit is given in Fig. XVI.iii. $V1$ and $V2$ are identical triodes for which

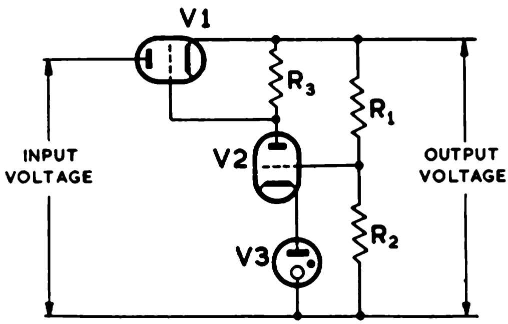


FIG. XVI.iii

$R_a = 40{,}000\ \Omega$ and $\mu = 20$. $V3$ is a gas-discharge tube. $R_1 = R_2$ and $R_3 = 1\ \text{M}\Omega$. Calculate, approximately, the change in the output voltage if the input voltage changes by 10 V. (0·04 V.) [I. of P., 1956.]

EXAMPLES XVII

1. Explain what is meant by an amplitude-modulated wave and show that it may be represented by a carrier and side-bands. What is the significance of this for communication purposes?

A transmitter radiates a power of 1,000 W when fed with a carrier amplitude modulated to a depth of 50 per cent. Calculate the power in each side-band.

(56 W.)

2. Describe and explain the main features of a superheterodyne receiver for the reception of amplitude-modulated broadcast signals. Describe in more detail the action of the second detector.

3. A superheterodyne receiver has a calibrated r.f. dial which indicates the frequency of the received signal. A certain station is received strongly at 100 Mc/s on the dial and weakly at 76 Mc/s. Explain this and determine

the actual frequency of the station and the intermediate frequency of the receiver.

(100 Mc/s, 12 Mc/s.)

4. The receiver in Example 3 receives the same station very weakly when the dial indicates 82, 84, 92 and 94 Mc/s. Show how these results arise from the harmonics of the station and the local oscillator.

5. A second station is received on the same receiver when the dial is set to 106 or 94 Mc/s. Explain this and find the actual frequency of the station.

(224 Mc/s.)

6. Explain how a pentode valve may be used as a variable reactance.

7. Draw the circuit of a variable-reactance valve in which the grid bias may be varied at audio frequency. The phase-shift potential divider consists of a capacitor of 30 pF connected between grid and negative h.t. and a 50 kΩ resistor.

Show that the impedance between anode and cathode of the valve is equivalent to an inductance shunted by a resistance, and find the values of these components for mean values of $g_m = 2$ mA/V and $f = 1{,}000$ c/s.

Explain how the circuit may be used for frequency-modulation of an oscillator.

What steps could be taken to avoid amplitude modulation?

(8·5 H, 500 Ω.) [I.E.E., III, April 1956.]

8. Show that the processes of modulation, demodulation and frequency-changing are essentially the same.

The dynamic characteristic of a triode with a 10,000-ohm resistive load is represented by

$$I_a = 2{\cdot}5\,(V_{gk} + 5) + 0{\cdot}2(V_{gk} + 5)^2 \text{ milliamperes,}$$

where V_{gk} is the potential difference between grid and cathode in volts. The valve is operated with a fixed bias of -3 V, and sinusoidal signals of amplitudes 1 V and 0·5 V at frequencies of 2 kc/s and 5 kc/s, respectively, are applied simultaneously to the grid circuit. Determine the amplitudes and frequencies of the various components of voltage across the load.

(59·3, 33, 16·5, 1, 0·25, 1, 1 V; 0, 2, 5, 4, 10, 7, 3 k/cs.) [I.E.E., III, 1954.]

9. Explain the operation of a diode rectifier used for the demodulation of a modulated radio signal.

Show that for an unmodulated signal the load presented to the preceding valve by the diode circuit is a resistance of half the value of the series load resistor in the diode circuit. Assume the diode to act as a perfect rectifier and to have a large condenser in parallel with the load resistor.

[I. of P., 1957.]

EXAMPLES XVIII

1. Describe in detail the operation of a free-running multivibrator. What general considerations govern whether a trigger circuit will be free-running, mono-stable or bi-stable? Give examples of each type of circuit.

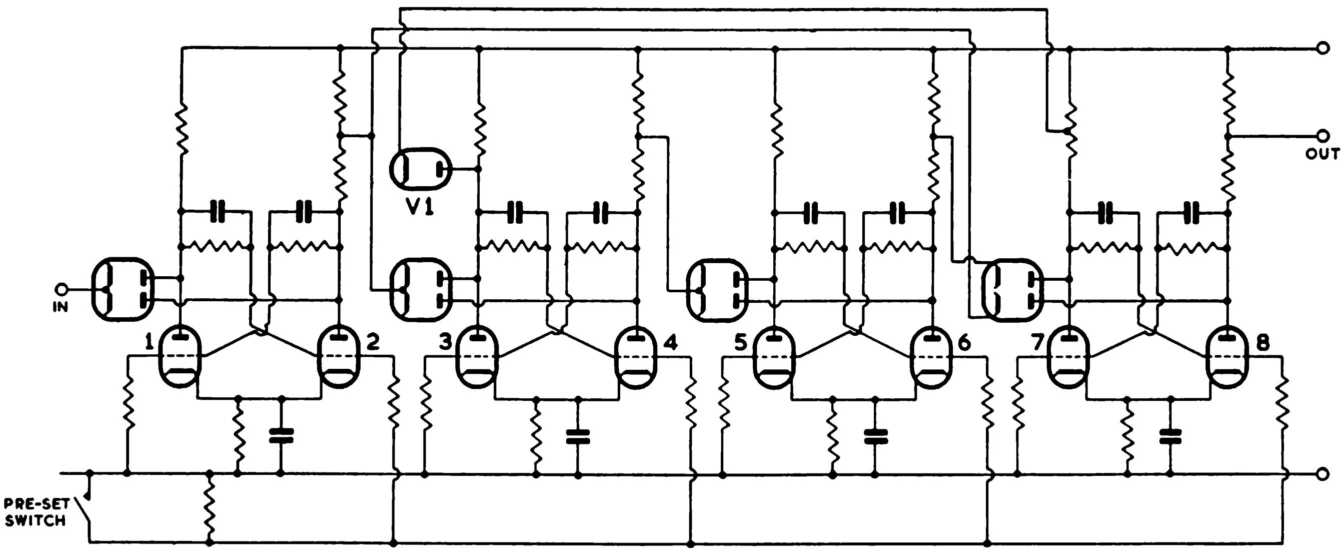


FIG. XVIII.ii

2. Use Sections 6.9, 18.1 and 18.4 to show how a transitron relaxation oscillator may be used as a switch across a charging condenser for the production of a saw-tooth waveform with the circuit shown in Fig. XVIII.i.

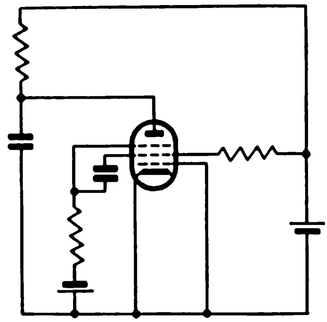


FIG. XVIII.i

3. Valves 1 to 6 in the circuit shown in Fig. XVIII.ii form three bistable binaries which count up to seven in the normal manner. The output of the first binary is also arranged to switch one stage of the output bistable circuit. Show that the whole circuit gives one negative output pulse for every ten input pulses.

4. Analyse the transistor multivibrator circuit of Fig. 18.6, assuming that the collector current is a linear function of the base current but is almost independent of the collector–emitter voltage.

EXAMPLES XIX

1. Show that a pentode valve used with a high-resistance load acts as a clamping device at approximately zero voltage. Hence show that the

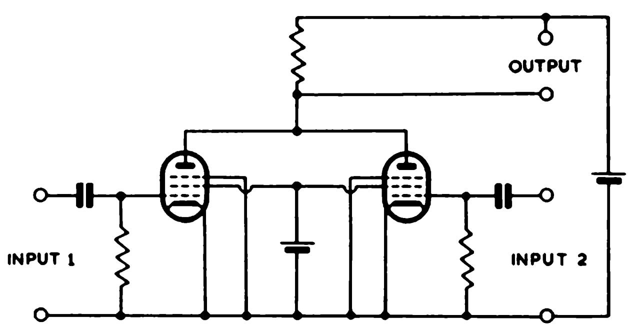


FIG. XIX.i

circuit of Fig. XIX.i acts as a coincidence counter, i.e., a device which operates only when two negative signal pulses occur simultaneously.

2. Explain how the circuit of Fig. XIX.ii can also act as a coincidence counter for negative pulses.

3. Explain how the circuits of Examples 1 or 2 may be adapted to act as " gates ", i.e., circuits which will count only for specified time intervals.

4. *The Operational Amplifier.* The circuit in Fig. XIX.iii shows an

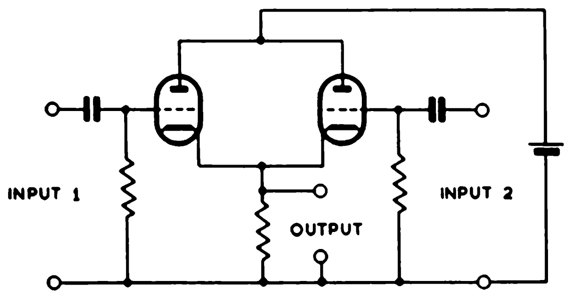


FIG. XIX.ii

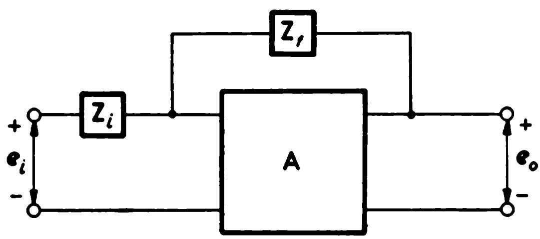


FIG. XIX.iii

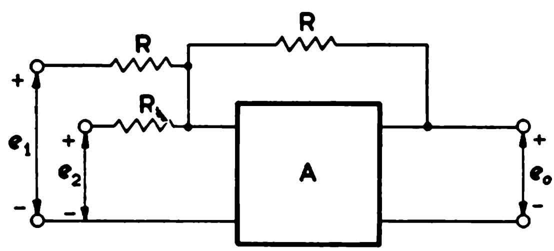


FIG. XIX.iv

amplifier with a network z_i in series with the input signal and a feedback impedance network z_f. Show that for a sinusoidal signal

$$\mathbf{E}_0 \simeq -\frac{\mathbf{z}_f}{\mathbf{z}_i}\mathbf{E}_i$$

when the inherent gain of the amplifier is very large.

When z_i is a resistance R_i and z_f is due to a capacitance C_f show that for any signal

$$e_0 \simeq -\frac{1}{R_i C_f}\int_0 e_i dt.$$

Also, when z_i is due to a capacitance C_i and $z_f = R_f$ show that

$$e_0 \simeq -R_f C_i \frac{de_i}{dt}.$$

(*Note.* These amplifiers are able to carry out respectively the mathematical operations of integration and differentiation. They are used in differential analysers and analogue computers.)

5. Show that the operational amplifier in Fig. XIX.iv can be used to carry out the mathematical operation of addition.

Examples XX

1. Write an account of the causes and effects of amplifier noise with particular reference to resistance noise, shot noise, equivalent noise resistance and noise factor.

2. A signal generator has an open-circuit output e.m.f. e, and its output impedance is resistive and of value R. Show that the ratio of signal to noise across the generator terminals on open-circuit is $e^2/4kTBR$. When a resistance R_1 is connected across the output terminals show that the signal-to-noise ratio is $e^2R_1/4kTBR(R + R_1)$ and the noise factor is $R_1/(R + R_1)$.

Hence show that in a receiver with a first stage of zero noise resistance and high gain the noise factor is 3 dB when the receiver input resistance, assumed not to be due to feedback, is matched to the generator.

3. Give a full account of how noise may be generated in thermionic amplifiers, and indicate why noise considerations are important in: (*a*) carrier-telephony repeaters; (*b*) television amplifiers.

(I.E.E., III, April 1957.]

4. Available power is defined as the maximum power which may be obtained in a load from a source, and it is achieved by matching the load to the source. Show that the maximum available noise power from any resistance is kTB.

5. Show that in a receiver with power gain G and an available noise power output of P_N per unit band width, the noise factor is P_N/kTG when the receiver is matched to a generator at temperature T.

6. A receiver has its input matched by a resistance R, across which is connected a temperature-limited diode. The available output noise power of the receiver is doubled when the diode current is increased from zero to i_A. Show that the noise factor of the receiver is $ei_AR/2kT$.

(*Note.* This method is used for the experimental determination of receiver noise factor.)

7. A wide-band amplifier is to consist of a number of identical stages, each giving a voltage gain of 10 times, and is to respond uniformly from direct current to 10 Mc/s. The signal-source resistance is 3,000 Ω, the input resistance is 2,000 Ω and the equivalent noise resistance of the valves used is 1,500 Ω, referred to the grid. The final stage can develop 30 V (r.m.s.) output. How many stages are required if, at full gain in the absence of signal, full output is to be provided by noise? The temperature of the equipment is 27° C.

The mean-square noise voltage developed by a resistance R Ω is:

$$\overline{e^2} = 4\ kTBR,$$

where $k = 1{\cdot}37 \times 10^{-23}$ joule per degree; T = absolute temperature; B = effective band width, c/s.

(7.) [I.E.E., III, April 1957.]

8. A 75 Ω resistive signal source is coupled to the first stage of an amplifier by a network giving an impedance step-up of 100 times and an effective band width of 5 Mc/s. Valve and circuit damping amount to 5,000 Ω, and the noise resistance of the valve, referred to its grid, is 2,000 Ω. Assuming a uniform temperature of 17° C, what is the r.m.s. noise voltage referred to the grid of the valve?

You may assume that the "available power" from a source of noise is kTB watts, where $k = 1{\cdot}37 \times 10^{-23}$ joule per degree absolute; T = temperature in degrees absolute; B = effective band width in cycles per second. (20 μV.) [I.E.E., III, 1954.]

9. A common-cathode amplifier is the first stage of a receiver. If a resistance R is across the input terminals the noise output of the receiver is found to be twice the noise output when the input terminals are short-circuited. Show that the noise resistance of the amplifier is equal to R. (Assume that there is no change in band-width.)

APPENDIX I

LIST OF SYMBOLS

THE list below contains certain symbols which are used frequently in the book. Some symbols are used with more than one meaning, but where this occurs the context makes it clear which meaning is intended. Symbols in heavy type denote complex or vector quantities. The rationalized M.K.S. unit in which each quantity is measured is also given.

$\mathbf{A}$ or A = voltage amplification.
$\mathbf{A}_f$ or A_f = voltage amplification with feedback.
A_p = power amplification.
$|\mathbf{A}|$ = stage gain.
$\alpha_{cb} = \dfrac{\partial i_C}{\partial i_B}$ at constant v_{CE} in a transistor $= h_{21e}$.
$\alpha_{ce} = -\dfrac{\partial i_C}{\partial i_E}$ at constant v_{CB} in a transistor $= -h_{21b}$.
B = magnetic flux density, Wb/m^2.
B = frequency bandwidth, c/s.
$\boldsymbol{\beta}$ or β = fraction of output voltage fed back in series with the input.
C = capacitance, F.
C_{ag}, C_{gk}, C_{ak} . . = inter-electrode capacitances, F.
d = distance between parallel planes, m.
d_{kg}, d_{ga} = interelectrode clearances, m.
δ = secondary emission coefficient.
E = energy, W.
E_1, E_2, etc. = e.m.f. of battery supplies, V.
E = electric field strength, V/m.
η = efficiency $= P_O/P_A$.
F = force, N.
f = frequency, c/s.
$g_m = \dfrac{\partial i_A}{\partial v_G}$ at constant v_A = mutual conductance or trans-conductance of a triode, A/V.
$g_m = \dfrac{\partial i_A}{\partial v_{G1}}$ at constant v_A, v_{G2}, v_{G3} = mutual conductance or transconductance of a pentode, A/V.
$G1$ refers to control grid in multi-electrode valves.
$G2$ refers to screen grid in multi-electrode valves.
$G3$ refers to suppressor grid in multi-electrode valves.

h_{11} or $h_{11b} = \dfrac{\partial v_{EB}}{\partial i_E}$ at constant v_{CB} in common base transistor, Ω.

h_{12} or $h_{12b} = \dfrac{\partial v_{EB}}{\partial v_{CB}}$ at constant i_E in common base transistor.

h_{21} or $h_{21b} = \dfrac{\partial i_C}{\partial i_E}$ at constant v_{CB} in common base transistor $= -\alpha_{ce}$.

h_{22} or $h_{22b} = \dfrac{\partial i_C}{\partial v_{CB}}$ at constant i_E in common base transistor, mho.

$h_{11e} = \dfrac{\partial v_{BE}}{\partial i_B}$ at constant v_{CE}, in common emitter transistor, Ω.

$h_{12e} = \dfrac{\partial v_{BE}}{\partial v_{CE}}$ at constant i_B, in common emitter transistor.

$h_{21e} = \dfrac{\partial i_C}{\partial i_B}$ at constant v_{CE}, in common emitter transistor $= \alpha_{cb}$.

$h_{22e} = \dfrac{\partial i_C}{\partial v_{CE}}$ at constant i_B, in common emitter transistor, mho.

i = current, A.

i_A = total anode current, A.

i_a = varying component of anode current, A.

$\mathbf{I}_a$ = vector value of sinusoidal component of anode current, A.

i_Q = steady value of anode current in absence of a signal, A.

$\bar{i}_A$ = mean value of total anode current, A.

$\hat{i}_a$ = peak value of varying component of anode current, A

i_G = total grid current, A.

i_g = varying component of grid current, A.

$\mathbf{I}_g$ = vector value of sinusoidal component of grid current, A.

$\bar{i}_G$ = mean value of total grid current, A.

$\hat{i}_g$ = peak value of varying component of grid current, A.

i_K = total cathode current, A.

i_E, i_B, i_C = total emitter, base, collector current, respectively, A.

i_e, i_b, i_c = varying component of emitter, base, collector current respectively, A.

$\mathbf{I}_e, \mathbf{I}_b, \mathbf{I}_c$ = vector value of sinusoidal component of emitter, base, collector current respectively, A.

J = current density, A/m^2.

J_S = total emission current density, A/m^2.

l = effective length of deflecting system of a cathode-ray tube, m.

L = inductance, H.
L = length from deflecting system to screen of a cathode-ray tube, m.
λ = wavelength, m.
m = mass, kg.
m or m_e = mass of electron, kg.
m_i = mass of ion, kg.
μ = charge mobility = ratio of drift speed to applied electric field, $m^2\,s^{-1}\,V^{-1}$.
$\mu = -\dfrac{\partial v_A}{\partial v_G}$ at constant i_A = triode amplification factor.
$\mu_S = -\dfrac{\partial v_{G2}}{\partial v_{G1}}$ at constant i_A and v_A = tetrode amplification factor.
P_O = power output, W.
P_A = anode power dissipation, W.
q = charge, C.
R = resistance, Ω.
r = radius, m.
r_a = anode slope resistance where $1/r_a = \dfrac{\partial i_A}{\partial v_A}$.
ρ = charge density, C/m^3.
σ = conductivity, $\Omega^{-1}\,m^{-1}$.
t = time, sec.
T = alternating period, sec.
T = absolute temperature, ° K.
τ = electron transit time, sec.
ϕ = phase angle, rad.
ϕ = work function, eV.
u = velocity, m/s.
v = potential, V.
v_A = total anode voltage measured from cathode, V.
v_a = varying component of anode voltage, V.
$\mathbf{V}_a$ = vector value of sinusoidal component of anode voltage.
v_{AQ} (sometimes v_Q) = steady or quiescent value of anode voltage in absence of signal, V.
$\bar{v}_A$ = mean value of total anode voltage, V.
$\hat{v}_a$ = peak value of varying component of anode voltage, V.
v_G = total grid voltage, V.
v_g = varying component of grid voltage, V.
$\mathbf{V}_g$ = vector value of sinusoidal component of grid voltage, V.
v_{GQ} = steady or quiescent value of grid voltage in absence of signal, V.
$\bar{v}_G$ = mean value of total grid voltage, V.

$\hat{v}_g$ = peak value of varying component of grid voltage, V.
$v_{G1}, v_{G2}, v_{G3}, \ldots$ = total voltages of $G1, G2, G3 \ldots$, V.
v_{G1}, v_{G2} (sometimes) = total voltages of the control grids of two different valves, V.
$v_{g1}, v_{g2}, v_{g3}, \ldots$ = varying components of voltages of $G1, G2, G3 \ldots$, V.
v_{g1}, v_{g2} (sometimes) = varying components of two different voltages applied to control grid, or, varying components of voltages of the control grids of two different valves, V.
v_s = signal voltage, V.
v_O = total output voltage, V.
v_o = varying component of output voltage, V.
v_M = value of potential minimum in space charge, V.
v_M = maintenance or operating or burning voltage in gas discharge.
v_S = breakdown or striking voltage in gas discharge.

All the above voltages are measured from the cathode.

v_{EB}, v_{BC}, v_{CE} = total transistor voltages between electrodes indicated, V.
v_{eb}, v_{bc}, v_{ce} = varying components of transistor voltages between electrodes indicated, V.
$\mathbf{V}_{eb}, \mathbf{V}_{bc}, \mathbf{V}_{ce}$ = vector value of sinusoidal component of transistor voltages between electrodes indicated, V.
$\hat{W}_A$ = maximum permissible anode dissipation, W.
W_F = Fermi energy level, eV.
ω = angular frequency = $2\pi f$, rad/s.
X = reactance, Ω.
Z = impedance, Ω.
$\mathbf{Z}$ = vector impedance, Ω.

APPENDIX II

USEFUL CONSTANTS

c = velocity of light in free space = $3{\cdot}00 \times 10^8$ m/s.
e = electronic charge = $1{\cdot}60 \times 10^{-19}$ C.
ε_0 = primary electric constant = $8{\cdot}86 \times 10^{-12}$ F/m.
$\varepsilon = 2{\cdot}72$.
h = Planck's constant = $6{\cdot}62 \times 10^{-34}$ J–s.
k = Boltzmann's constant = $1{\cdot}38 \times 10^{-23}$ J/° K.
m = electronic rest mass = $9{\cdot}11 \times 10^{-31}$ kg.

APPENDIX III

BIBLIOGRAPHY

THE following list gives a selection of books for further reading:

General

Gray, T. S. *Applied Electronics*, 2nd edition (Wiley, 1954).
Parker, P. *Electronics* (Arnold, 1950).
Terman, F. E. *Electronic and Radio Engineering*, 4th edition (McGraw Hill, 1955).
Williams, E. *Thermionic Valve Circuits*, 3rd edition (Pitman, 1955).

Valves

Beck, A. H. W. *Thermionic Valves* (Cambridge University Press, 1953).
Harman, W. W. *Fundamentals of Electronic Motion* (McGraw Hill, 1953).
Spangenberg, K. *Vacuum Tubes* (McGraw Hill, 1948).

Transistors

Evans, J. *Fundamental Principles of Transistors* (Heywood, 1957).
Lo, A. W., Endres, R. O., Zawels, J., Waldhauer, F. D., and Cheng, C. *Transistor Electronics* (Prentice-Hall, 1955).

INDEX