ELECTRONIC CIRCUIT THEORY

Devices, Models, and Circuits

H. J. Zimmermann
S. J. Mason

Emphasizes the use of linear
and piecewise-linear circuit models

JOHN WILEY & SONS, Inc.
Publishers

About the book . . .

This book deals primarily with methods of analysis of electronic circuits. The model concept is stressed. Resistive models for electronic devices are synthesized. Special attention is given to piecewise-linear models suitable for large-signal operation. The authors devise models to approximate the characteristics of diodes, triodes, pentodes, transistors, and other control valves. It contains extensive graphical and geometrical interpretations of analyses. The effect of circuit and signal on device operation is shown by means of locus plots. Basic circuit functions are classified as follows: rectification and detection, waveshaping and amplification, and waveform generation.

HG-JF-3-59

One of the publications
in the M.I.T.
"core-curriculum" program
in electrical engineering
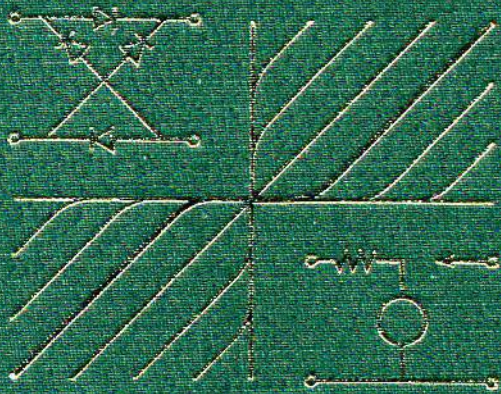
About the authors . . .

Henry J. Zimmermann is Professor of Electrical Engineering and Associate Director of the Research Laboratory of Electronics at the Massachusetts Institute of Technology.

Samuel J. Mason is Associate Professor of Electrical Engineering and a staff member of the Research Laboratory of Electronics at the Massachusetts Institute of Technology.

ELECTRONIC

CIRCUIT

THEORY

ELECTRONIC

CIRCUIT

THEORY

Devices, Models, and Circuits

HENRY J. ZIMMERMANN

Professor of Electrical Engineering

SAMUEL J. MASON

Professor of Electrical Engineering

**Department of Electrical Engineering
and Research Laboratory of Electronics**

Massachusetts Institute of Technology

TO

PRISCILLA AND JEANIE

F O R E W O R D

This book is one of several resulting from a recent revision of the Electrical Engineering Course at The Massachusetts Institute of Technology. The books have the general format of texts and are being used as such. However, they might well be described as reports on a research program aimed at the evolution of an undergraduate core curriculum in Electrical Engineering that will form a basis for a continuing career in a field that is ever-changing.

The development of an educational program in Electrical Engineering to keep pace with the changes in technology is not a new endeavor at The Massachusetts Institute of Technology. In the early 1930's, the Faculty of the Department undertook a major review and reassessment of its program. By 1940, a series of new courses had been evolved, and resulted in the publication of four related books.

The new technology that appeared during World War II brought great change to the field of Electrical Engineering. In recognition of this fact, the Faculty of the Department undertook another reassessment of its program. By about 1952, a pattern for a curriculum had been evolved and its implementation was initiated with a high degree of enthusiasm and vigor.

The new curriculum subordinates option structures built around areas of industrial practice in favor of a common core that provides a broad base for the engineering applications of the sciences. This core structure includes a newly developed laboratory program which stresses the role of experimentation and its relation to theoretical model-making in the solution of engineering problems. Faced with the time limitation of a four-year program for the Bachelor's degree, the entire core curriculum gives priority to basic principles and methods of analysis rather than to the presentation of current technology.

J. A. STRATTON

P R E F A C E

The importance of electronic devices, circuits, and systems in modern technology is apparent to most electrical engineering students; many have acquired practical experience with circuits before they begin formal study of the subject. Electronic circuit theory can be introduced to the student in a variety of ways. In view of the many devices available and the numerous applications of electronic circuits, it is important in any plan of presentation to seek unifying principles. Such principles permit the student to extend his knowledge in a rapidly advancing field.

We have organized our approach to electronics around circuit models and methods of circuit analysis in order to reduce the number of separate ideas and concepts. The many functions performed by electronic systems can be understood in terms of a few fundamental circuits if similarities are sought.

This book deals with electronic devices, models, basic circuits and circuit functions. Many of the interesting properties of electronic devices are a consequence of nonlinearity accompanied by regional linearity. As a result, piecewise-linear circuit models can be used to convert a nonlinear circuit problem to a number of related linear problems. Thus, the mathematics of linear circuit theory can be applied to a broad class of physical circuits and systems operating in a nonlinear manner.

The model concept emphasizes the need for making approximations as part of the process of analyzing a physical problem. The student is thus encouraged to exercise judgment in order to arrive at the simplest circuit models that will give an adequate result. Extremely

simple models can be used to explain the general mode of operation
of electronic circuits. Simple resistive models together with one major
energy-storage element suffice to explain the behavior of most basic
circuits. Refinements in the resistive model and one or two additional
energy-storage elements provide adequate accuracy for nearly all
design purposes. Sinusoidal or rectangular-wave circuit response illus-
trates basic operations such as waveform generation, wave shaping,
amplification and modulation. A companion volume presents methods
of linear and quasi-linear analysis pertinent to more complicated elec-
tronic circuits, signals and systems.

The emphasis on general methods illustrated by specific examples
is in keeping with the present trend in engineering education. The
emphasis on fundamentals is the inevitable consequence of rapid
development, particularly in such fields as electronics, communications,
and computation, which have literally exploded in the past decade.
In today's technology, a specialized education becomes obsolete too
soon after graduation. We do not mean to say that real problems
and applications should be avoided. However, too much specialization,
either in fact or in attitude, deprives the student of the background
and the confidence that will enable him to enter new fields. More-
over, technical problems often span several disciplines so that breadth
of understanding becomes more important in the long run than detailed
knowledge.

Some of the material presented in this book evolved from a
graduate subject on pulse circuits, and some was developed during
the revision of the introductory undergraduate subject on electronic
circuits (part of the core curriculum for all electrical engineering
students at M.I.T.). With minor variations, portions of the material
in this book have been used for five years in this core subject.

The development of this presentation of electronic-circuit theory
was influenced by the early work of Godfrey T. Coate. Contributions
have also been made by other colleagues on the teaching staff; in
particular, many of the problems were prepared by section instructors.
Ideas have come from staff members of the Research Laboratory of
Electronics or have resulted from the stimulation of the research
environment. The inspiring leadership of Professor Ernst A. Guillemin
in circuit theory research and teaching has had both tangible and
intangible effects on the project. Many worth-while suggestions have
been made by our students.

During the final stages of the book, we had the invaluable aid of
Professor Campbell L. Searle, whose critical technical editing con-
tributed greatly to the improvement of the manuscript. In addition,

he gave his time and effort unsparingly to galley reading in order to help us meet publication deadlines. As in any other book, a number of errors inevitably remain. The number would have been greater had it not been for the perceptive checking and page proofing done by Professor Richard D. Thornton, who also took major responsibility for organizing the index.

We are most grateful to Professor Gordon S. Brown for creating a departmental environment in which academic experiments are the rule rather than the exception. His constant encouragement has provided a real stimulus throughout the subject revision, note writing, manuscript, and production stages of this project.

Our acknowledgements would be incomplete without an expression of thanks to the secretaries who typed rough draft, notes, and manuscript. They are Bertha Hornby, Rosemarie Connell, Dorothea Scanlon, Marjorie D'Amato, and Louise Juliano.

<div align="right">

HENRY J. ZIMMERMANN
SAMUEL J. MASON

</div>

Cambridge, Massachusetts
January, 1959

C O N T E N T S

9.1 Introduction. 9.2 Properties of Oscillators.
9.3 Negative Resistance in the Common-Base,
Point-Contact Transistor Circuit. 9.4 Negative
Resistance in a Common-Emitter, Point-Contact
Transistor Circuit. 9.5 Negative Resistance in a

Series Triode Circuit. 9.6 Some General Prop-
erties of Negative Resistance. 9.7 A Simple
Relaxation Oscillator. 9.8 Relaxation Oscillator
Transition Time. 9.9 Relaxation Oscillator
Locus and Waveforms with Nonzero Transition
Time. 9.10 Shifting from Astable to Monostable
Operation. 9.11 Triggering the Monostable Cir-
cuit. 9.12 Bistable Operation. 9.13 Limitations
of Negative-Resistance Methods. 9.14 Plate-
Coupled Multivibrator. 9.15 Plate-Coupled
Bistable Circuit. 9.16 Blocking Oscillator.
9.17 Sinusoidal Oscillations.

Introduction

1.1 Electronic Circuit Theory

Electronic devices such as diodes, triodes, and transistors operate as switches or valves to control or modulate the flow of electric current. Electronic switches and valves are useful because of their sensitivity and speed of operation, which exceed the sensitivity and speed of mechanical or electromechanical devices.

Electronic circuit theory is the mathematical study of circuits containing electronic devices. The ultimate purpose of electronic circuit theory is to provide a basis for the design of electronic systems. Such systems are combinations of electric and electronic devices assembled and connected to perform some desired operation on an electrical signal. These systems may also be associated with nonelectrical devices as in servomechanisms or industrial process control.

Like any other body of theory, electronic circuit theory deals with *models*. A model is a simple, idealized abstraction which approximates the behavior of a physical system. *A model is always a compromise between simplicity and reality.* The three phases of electronic circuit theory are, therefore, as follows:

1. Development of suitable *circuit models for electronic devices.* Model making is facilitated by a knowledge of the theoretical "models" for the

underlying *physical processes* which determine the externally observed electrical behavior of a given device.

2. Development of *analysis methods* applicable to electronic circuits. A natural and useful approach is the extension of elementary *electric circuit theory* so that familiar techniques can be brought to bear upon *electronic circuits*.

3. Accumulation of a store of ideas from working with the models and analyzing the circuits. A knowledge of device and circuit capabilities and limitations provides specific criteria and general technical intuition for the *design of electronic systems*.

This book emphasizes the first two phases of electronic circuit theory, namely, model making and methods of circuit analysis. Design information accumulates, as we proceed, from illustrative examples, problems, and interpretations of basic results.

1.2   Ideal Circuit Elements

In general, electronic circuits are not linear and do not obey reciprocity. This means that we cannot make satisfactory circuit models from the familiar R, L, and C building blocks of linear circuit theory. However, the difficulty can be overcome with only two additional circuit elements. These are the *ideal diode* and the *controlled source*.
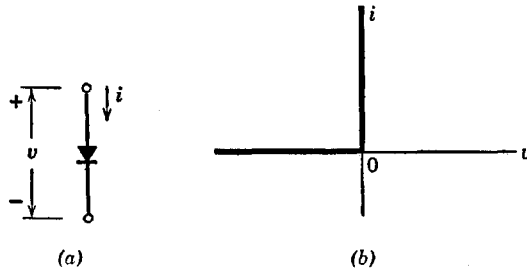
Fig. 1.1.   The ideal diode.   (*a*) Circuit symbol; (*b*) Current versus voltage curve.

As indicated in Fig. 1.1, the ideal diode is a *self-operated switch* that *opens* when the terminal voltage is negative and *closes* when the current is positive. The ideal diode, sometimes called an ideal rectifier, is a *non-linear* circuit element. It is an idealized approximation of the class of electronic devices known as diodes or rectifiers.

A controlled source is shown in Fig. 1.2. The control variable and the

source variable may each be either a current or a voltage, but only the current-controlled current source is shown here. The controlled source is a linear three-terminal circuit element. It is described by the linear two-terminal-pair relations between its input variables e_1 and i_1 and its output variables e_2 and i_2. The *controlled source* is a *basic unilateral circuit element.* "Unilateral" means that a signal at the input has an
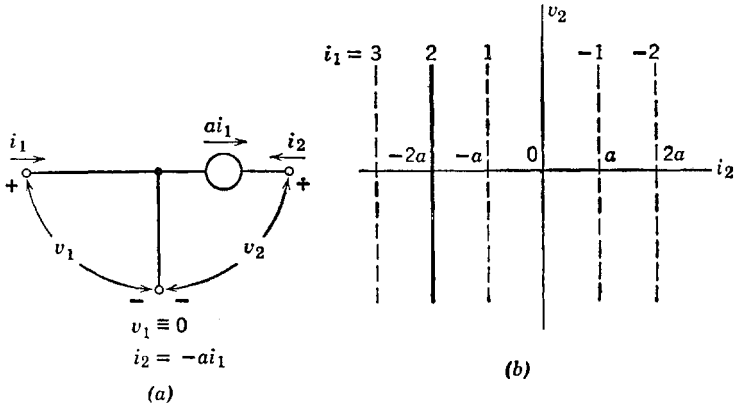
Fig. 1.2. The controlled source. (*a*) Circuit symbol for a current-controlled current source; (*b*) Output curve (v_2 vs. i_2) showing the effect of the control signal.

effect on the output but a signal applied at the output has no influence on the input. Thus in Fig. 1.2, the input current i_1 controls the position of the line relating v_2 and i_2. Varying i_1 from $+2$ to -2 sweeps the v_2 vs. i_2 line across the plane from left to right. Conversely, v_1 and i_1 are independent of applied v_2. Such performance is impossible to achieve with resistance, capacitance, or inductance (including mutual inductance). Coupling between two terminal-pairs due to any of these elements is "bilateral" or "reciprocal" (two-way coupling).

1.3 Circuit Models

The ideal diode allows us to make piecewise-linear models of nonlinear devices. From the current-versus-voltage curve for a typical vacuum diode [Fig. 1.3(*a*)], we see the general character of diode curves that suggests the ideal-rectifier approximation. By combining just a single resistance with an ideal diode, as in Fig. 1.3(*b*), we have a piecewise-linear model that matches the actual curve very closely.

Inclusion of a controlled (or dependent) source allows us to devise models that approximate the output curve of a control valve (such as a

(a)   (b)
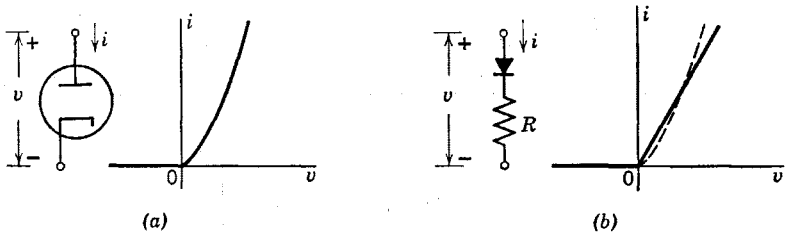
Fig. 1.3. Diode curve and piecewise-linear approximation.

(a)

(b)
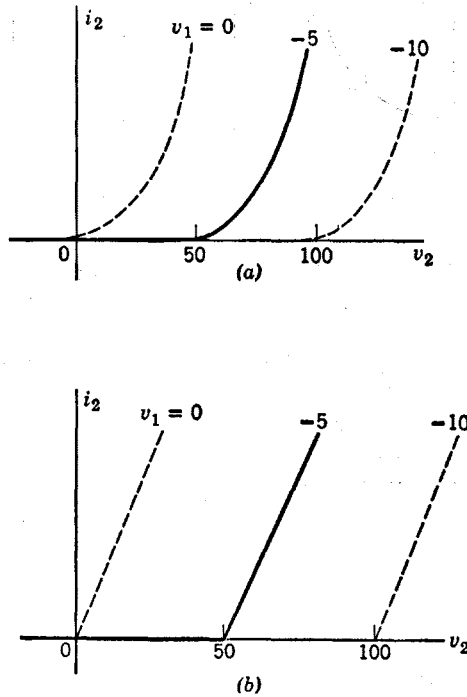
Fig. 1.4. Control valve curve and piecewise-linear approximation. (a) i_2 vs. v_2 for $v_1 = -5$; (b) Approximation.

vacuum triode or transistor). A typical triode curve is shown in Fig. 1.4(a) and a simple piecewise-linear approximation is indicated in (b).

Many nonlinear electronic devices are fairly linear in certain regions of operation, with rather abrupt transitions connecting the different re-

gions. In other words, a piecewise-linear model is often a very realistic representation.

1.4 Analysis Methods

From the analysis standpoint, piecewise-linear models lead to linear equations with restricted ranges of validity. A piecewise-linear analysis problem consists of a number of linear problems, each one pertinent to a separate "piece" or "range" of the variables involved. The convenience of the piecewise-linear approach lies in the ease of solution of linear equations.

$$v_1 = Ri_1(1 - a)$$

$$R_{\text{input}} = \frac{v_1}{i_1} = R(1 - a)$$
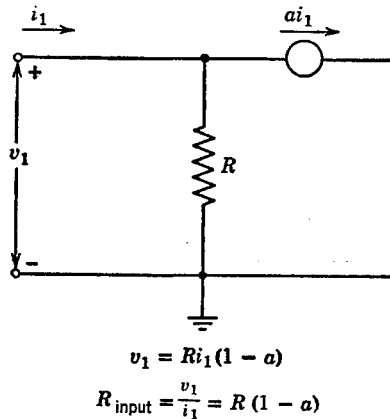
Fig. 1.5. Effect of controlled source on apparent resistance.

Background gained in the study of linear RLC circuits is useful since it applies directly to nonlinear electronic circuit problems when the devices are represented by piecewise-linear models. The very fact that piecewise-linear approximations provide a facility for handling nonlinear problems provides a strong stimulus for further study of linear circuit theory. The more we know about linear circuit theory, the better our preparation for handling electronic circuit problems.

The controlled source, though linear, introduces effects not treated in elementary circuit theory. Figure 1.5 offers a specific example. The controlled source *dictates* the current through R, thereby establishing v_1 and determining the apparent input resistance. If we apply Thevenin's theorem and calculate the effective resistance between terminals by "short-circuiting the internal voltage sources and open-circuiting the

current sources," we obtain the value R, which is incorrect. The error arises from an incorrect use of Thevenin's theorem. Only *independent* sources may be open-circuited or short-circuited to find effective resistance, since only independent sources contribute to the constant A (open-circuit voltage) in the general linear terminal relation

$$v = A + Bi \qquad (1.1)$$

Controlled sources are *dependent* on the control signal. Since they change when the control signal changes, they contribute to the constant B (effective resistance). Thus, controlled sources do not violate the laws of elementary linear circuit theory but they do require us to *understand* the laws, rules, and theorems in order to apply them properly.

In this book we shall describe certain conduction processes, from these deduce the terminal behavior of various electronic devices, make circuit models for the devices, and use the models to study the operation of basic electronic circuits. Attention is given to the important functions performed by each class of circuits. As we proceed, it will become clear that a relatively small number of fundamental ideas, properly interpreted, permit one to understand the operation of a very large variety of electronic devices, circuits, functions, and systems. An integration of such ideas is the basis for circuit and system design.

PROBLEMS

1.1. Which of the following terms apply to the ideal diode? (*a*) linear, (*b*) capacitive, (*c*) resistive, (*d*) inductive, (*e*) nonlinear, (*f*) lossless, (*g*) bidirectional, (*h*) reactive, (*i*) passive, (*j*) active.

1.2. Do the following relations between instantaneous current and voltage specify the behavior of the ideal diode completely?

$$i \leqslant 0$$
$$v \geqslant 0$$
$$vi = 0$$

1.3. Compare the following relations with those of Problem 1.2.

$$v = 0 \quad \text{for } i > 0$$
$$i = 0 \quad \text{for } v < 0$$

1.4. What is the circuit model M for a driving-point curve like that shown in Fig. P1.1?

1.5. Sketch the output curve i_2 vs. v_2 for the voltage-controlled voltage source shown in Fig. P1.2. What are the dimensions of μ?
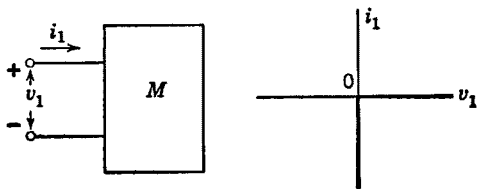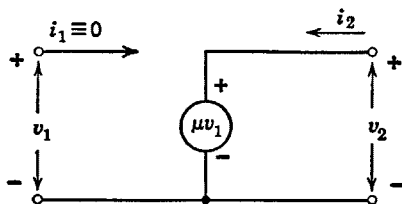
Fig. P1.1

Fig. P1.2

1.6. Sketch the output curve i_2 vs. v_2 for the voltage-controlled current source shown in Fig. P1.3. What are the dimensions of g?
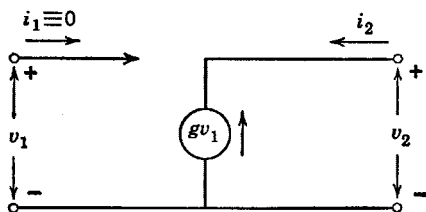
Fig. P1.3

1.7. Sketch the output curve v_2 vs. i_2 for the current-controlled voltage source shown in Fig. P1.4. What are the dimensions of r?
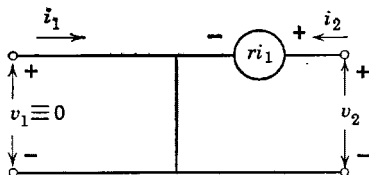
Fig. P1.4

1.8. Plot the terminal relation v vs. i for a linear resistance.

1.9. Show by graphical addition of voltages that the curve in Fig. 1.3(b) is composed of the curves for a resistance R and an ideal diode.

1.10. Plot the terminal relation v vs. i for a linear resistance R in parallel with an ideal diode.

1.11. Plot i vs. v for the circuit shown in Fig. P1.5.
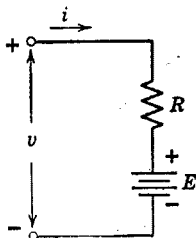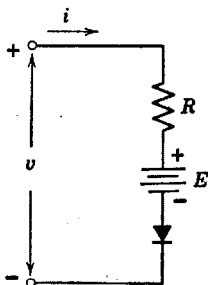
Fig. P1.5   Fig. P1.6

1.12. Plot i vs. v for the circuit shown in Fig. P1.6.

1.13. How can the circuit of Fig. P1.6 be modified to represent the curve of Fig. 1.4(b) in the text?
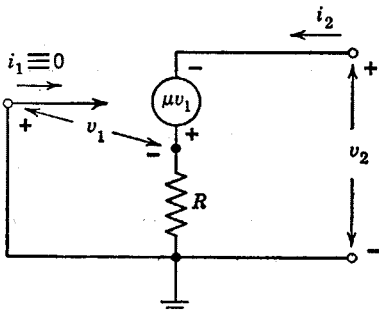
Fig. P1.7

1.14. Determine the ratio of v_2 to i_2 for the circuit shown in Fig. P1.7.

Electrical Conduction and Diodes

2.1   Introduction

Electrical conduction processes govern the basic properties of many electronic devices. Vacuum, gas, or semiconductor diodes are important examples because they are in common use and because they form the basis for describing the behavior of transistors, triodes, pentodes, thyratrons, and other devices.

Circuit models for electronic devices can be based entirely on measured electrical properties as typified by the average characteristic curves provided by manufacturers. However, without some knowledge of the theory of operation one might conclude that drastic differences exist between different types of devices. In terms of conduction processes, we can recognize similarities that unify and simplify the subject. At best, measurement of the external terminal behavior of a device can tell us only "what" happens, whereas the conduction processes help to explain "why" it happens.

The following discussion of conduction processes is aimed primarily at establishing the plausibility of the theoretical forms for the current-versus-voltage curves of vacuum, gas, and semiconductor diodes. The operation of a few other devices is also described. The results obtained for the diodes are extrapolated to transistors, triodes, pentodes and thyratrons in Chapters 5, 6, and 7.

2.2   Atomic Particles

For our purposes it will suffice to consider the nucleus of an atom as a single entity with a net positive electrical charge and a particular atomic weight. The remainder of the atom consists of electrons, each bearing a negative charge and orbiting around the nucleus.

The orbital electrons can be considered to be arranged systematically in successive layers or shells. The ones nearest the nucleus are more tightly bound to the nucleus by the force of attraction between oppositely charged particles than are the more remote electrons. The total number of orbital electrons is just sufficient to neutralize the positive charge on the nucleus, so that atoms normally display no net charge.

The similarity of the chemical properties of various elements, which results in the familiar periodic table, is related to the electron arrangement around the nucleus. The shell nearest the nucleus can contain no more than two electrons, whereas successive layers may contain eight or more. The outermost shell contains the valence electrons which govern the nature of chemical reactions. These valence electrons also play a large part in determining electrical behavior and crystalline structure.

The metallic elements tend to have one, two, or three valence electrons, and nonmetals have five, six, or seven. Filled valence shells correspond to the inert gases, such as helium, argon, neon, and xenon. Half-filled valence shells characterize a class of elements known as semiconductors (e.g. carbon, silicon, germanium). Although mechanically brittle rather than ductile, these elements have a metallic appearance. As we shall see in the discussion of conductivity, the semiconductors are centered somewhere between the metals that are good electrical conductors and the nonmetals which tend to be good insulators.

2.3   Electrical Conduction in Metals

An electric field applied to a conducting material produces a slow, steady drift of electrons which accounts for the conduction current. Conduction current is defined as the rate at which charge passes through a given cross section of a conductor (amperes equal coulombs per second). Although this current results from the motion of many electrons, we shall neglect statistical fluctuations in the numbers of electrons and in their velocities so that current is like a smooth fluid flow.*

* The fluid-flow analogy is not valid for very small currents, because in this case the fluctuations due to the particle nature of the current cannot be neglected.

Copper, the most commonly used conducting material, has two valence electrons per atom. The atoms of copper are so densely packed that valence electrons are readily interchanged by the atoms and are free to
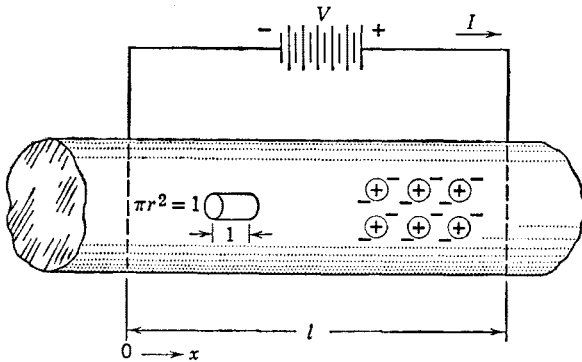
$-e$ = charge on electron = 1.6×10^{-19} coulombs

$E = \dfrac{-V}{l}$ = electric field (volts/meter)

v = drift velocity (meters/sec)

$M = v/E$ = mobility (meter^2/volt sec)

n = no. of electrons/meter^3

$J = -nev$ = current density (amperes/meter^2)

$\sigma = J/E = Mne$ = conductivity (mhos/meter)

$I = JA$ = current (amperes)

$I = \sigma AE = \dfrac{\sigma A V}{l}$

$I = GV$ where $G = \dfrac{\sigma A}{l}$ = conductance (mhos)

$V = RI$, where $R = \dfrac{1}{G}$ = resistance (ohms)

$R = \dfrac{\rho l}{A}$, where $\rho = \dfrac{1}{\sigma}$ = resistivity

$\rho = 1.72 \times 10^{-6}$ ohm-cm for copper

Fig. 2.1. Conduction in a metal.

move when a force is applied. Current is proportional to the number of electrons involved, their mobility or freedom of motion, and the magnitude of the applied force. The mobility is defined as the drift velocity per unit of applied electric field. (See Fig. 2.1.)

The fluctuations of current due to thermal agitation of the charged particles can be heard as a hissing sound in the output of a high-gain audio amplifier or radio receiver. As a result, such thermal fluctuations of current (or voltage) are called "noise."

The drift velocity of the electrons in a metal represents an equilibrium condition analogous to the terminal velocity of a body falling in the gravitational field. In vacuum, the force of gravity produces constant acceleration; in air, a falling body attains a terminal velocity determined by the force of gravity and by friction forces. Similarly, in vacuum, the force on an electron due to an electric field produces constant acceleration since there are no restraining forces. In a metal, the moving electrons collide with the atoms and lose some of the kinetic energy gained from the electric field. The collisions increase the thermal energy of the atoms and thus produce the familiar heating effect due to an electric current. The drift velocity represents an equilibrium between the kinetic energy gained from the applied field and that lost in collisions.

Mobility corresponds roughly to the reciprocal of viscous friction in a mechanical system. The force on an electron is the product of the electronic charge e and the electric field strength E. Thus, the force equals $(e/M)v$, whereas in mechanics the corresponding term is $f\,dx/dt$ or fv, where f is the coefficient of viscous friction.

2.4   Conduction in a Semiconductor

The elements called semiconductors have four valence electrons. Silicon or germanium are good examples. They are normally in crystalline form with each valence electron shared by two atoms. As indicated in Fig. 2.2, the atoms are arranged along the diagonals and centers of cubes. Each bond in this crystal lattice consists of two valence electrons shared by two atoms. The planar diagram in Fig. 2.2 is convenient for visualization of valence bonds. In Fig. 2.3(a) each link represents a single valence electron shared by two atoms. The circles each represent the nucleus of a germanium atom plus the orbital electrons other than valence electrons. A particular valence electron may be assumed to spend half time with each of two atoms. On the average each atom sees eight half-time electrons, which satisfies the valence requirements for four electrons. Compared with the situation that exists in a metal, the valence electrons are relatively tightly bound.

At room temperatures (say 300°K), the thermal vibrations of the atoms shake loose a few electrons. The number depends upon the purity of the crystal and other factors as well as the temperature. In a pure material, such as germanium (called an intrinsic semiconductor), the number of valence electrons free of the atomic bonds at any given time is apt to be one in every 10^8 to one in 10^{10} of the total number of valence electrons. These free electrons, except for the fact that they

are less numerous, behave somewhat like the free electrons in a metal. They have a definite mobility, and under the influence of an applied field they produce a conduction current. The conductivity of an intrinsic semiconductor can be expected to be much less than that of a metal because of the small number of free electrons available.
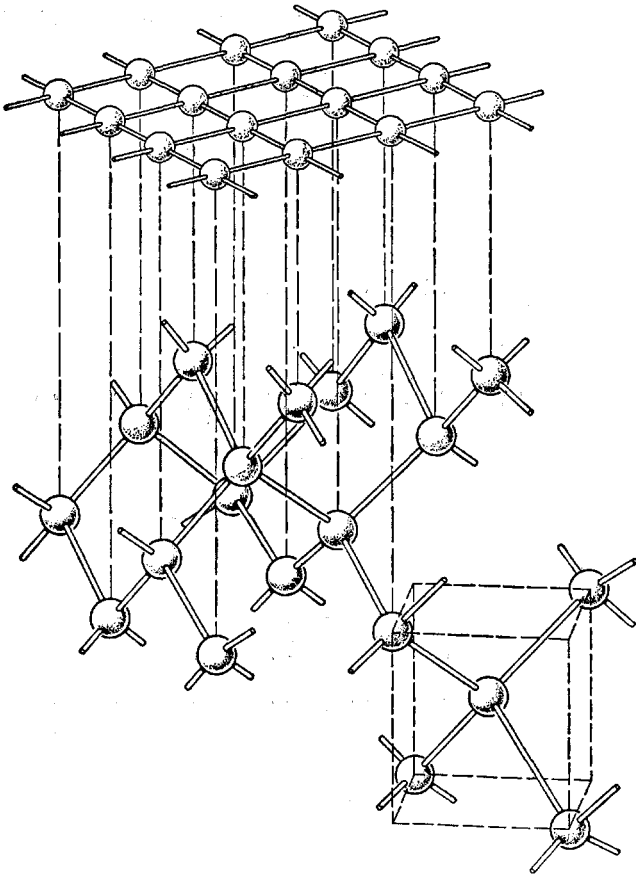
Fig. 2.2. Semiconductor crystal structure.

A free electron wandering through the semiconductor crystal is very likely to be recaptured by any other atom which happens to have a missing valence electron. In fact, when an electron is jarred loose from its atomic bond, it leaves a vacancy in the crystal structure. This vacancy in the crystal lattice is very aptly called a "hole." The atom from which the electron has departed is effectively ionized (net positive

charge equal to $+e$), but since the total number of such ionized atoms equals the total number of free electrons, the material displays no

$$\sigma = ne(M_h + M_e)$$

(a)

Electric field →

(b)

Electric field →

(c)
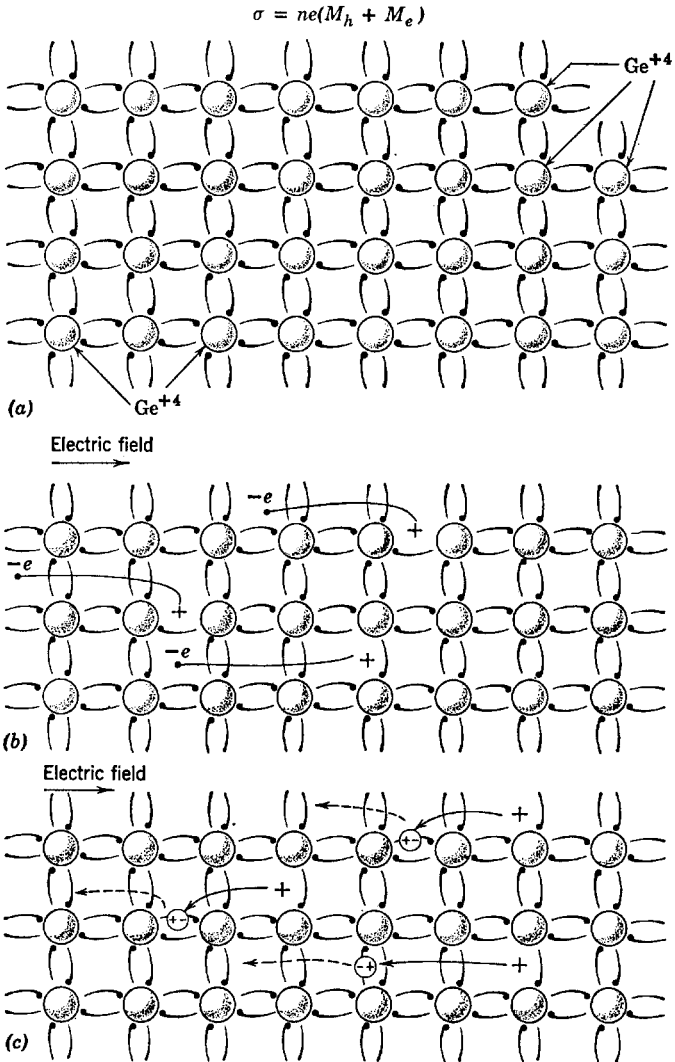
Fig. 2.3. Conduction in an intrinsic semiconductor.

charge in the macroscopic sense. Nevertheless, in the course of the endless thermal jostlings, the ionized atom seeks to acquire from a

neighbor an electron to replace the missing one. Whenever such a transfer takes place, the electron vacancy or hole has effectively moved from one atom to another. Thus, the positive charge associated with an ionized atom may be transferred from one atom to another in the lattice. Since the atoms themselves do not migrate through the material, it is convenient to associate a positive charge with the hole.

Designating the hole as the equivalent of a positively charged particle provides a physical model that agrees quite well with experimental observations. Holes display a mobility differing from that of free electrons, but like electrons, they contribute a component to the conduction current. In an intrinsic semiconductor, the number of holes and their total charge equal the number and total charge of the free electrons; hence, the fractions of the conduction current contributed by the electrons and by the holes are proportional to the mobilities in an intrinsic semiconductor. The sketches in Fig. 2.3(b) and (c) indicate the formation, migration, and recombination of electrons and holes. An applied electric field E is assumed to be directed from left to right. Under the influence of the applied field, the electrons move to the left and the holes to the right. Since they are oppositely charged, the electron current and hole current add to produce the total conduction current.

The recapture of an electron by a vacancy in the crystal lattice amounts to the recombination of an electron with a hole. The breaking of a bond represents the creation of a free electron ($-e$) and a hole ($+e$). The normal life span of the free charges is of the order of a millisecond in germanium. This lifetime is sufficient to make semiconductors interesting and important materials.

From a practical standpoint the conductivity of an intrinsic semiconductor like germanium is too low to be very useful. Therefore, in normal use controlled amounts of impurities are injected to provide additional conduction electrons (n-type) or additional holes (p-type).

2.5   n-Type and p-Type Semiconductors

Since free electrons and holes exist in equal numbers in an intrinsic semiconductor, the charge carriers added by an impurity cause either the holes or the electrons to predominate. Thus, in n-type semiconductors, electrons are the majority carriers and holes are in the minority; in p-type the reverse is true.

The n-type material is produced by adding a carefully controlled minute quantity (e.g., one part in 10^8) of a "donor" impurity such as arsenic to the intrinsic semiconductor. The arsenic atoms merely

replace occasional germanium atoms in the crystal lattice. As indicated in Fig. 2.4(a) they produce an excess of free electrons, since arsenic has five valence electrons. The number of free electrons "donated" by the

(a)

(b)
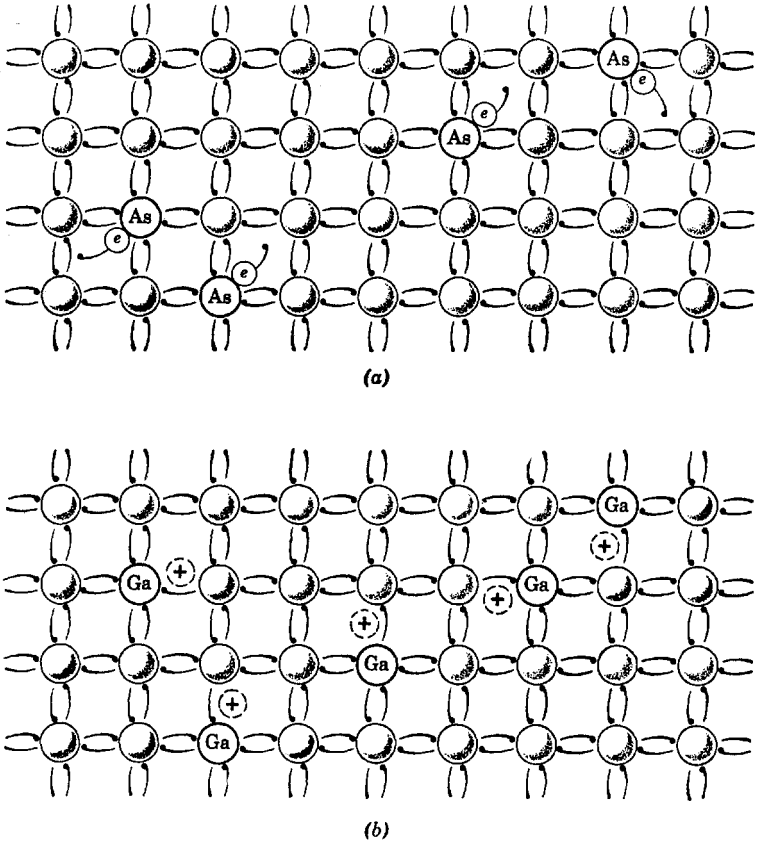
Fig. 2.4. Conduction in (a) n-type and (b) p-type semiconductors.

arsenic far exceeds the number of electron-hole pairs formed; hence, the conductivity due to the majority carrier (electrons in this case) far exceeds that of the intrinsic material.

The p-type semiconductor is formed by adding an "acceptor" impurity like gallium to an intrinsic semiconductor. The acceptor has only three valence electrons, and when it replaces the germanium in the crystal lattice, it leaves a hole as indicated in Fig. 2.4(b). Thus, the holes predominate and the conductivity increase is mainly due to this charge carrier which in actuality is a shortage of electrons.

2.6   Semiconductor Junction

Various methods are used to bring about the close proximity of p-type and n-type materials necessary to create a p-n junction. The so-called grown junction is actually a single crystal with a donor impurity on one side of the junction and an acceptor impurity on the other. The transition occurs in a very small distance (of the order of 0.001 millimeters).
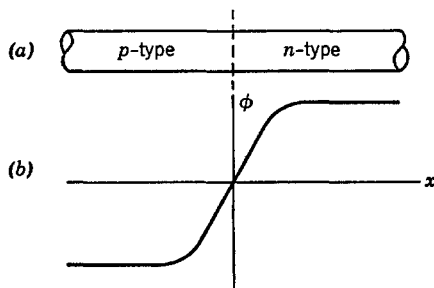
Fig. 2.5.   Potential diagram of a semiconductor p-n junction.

The sketch in Fig. 2.5(a) shows a p-n junction. Each type of semiconductor material is electrically neutral, but the p-type contains an excess of holes, and the n-type contains free electrons. Consequently, by thinking of the excess holes or electrons as a gas or cloud of particles free to move, it is natural to expect some holes to diffuse from the p-type to the n-type and some electrons to diffuse the other way.

The holes moving to the right represent positive charge and leave the left side of the junction slightly negative. Similarly, electrons diffusing to the left add negative charge there. The loss of electrons from the right half further raises the potential of the right side of the junction. Thus, the diffusion of both holes and electrons results in an additive effect illustrated by the potential diagram in Fig. 2.5(b). The potential difference developed across the junction by the diffusion of particles tends to oppose the diffusion. The higher the potential difference, the smaller the number of particles that have sufficient energy to migrate across the potential barrier. An equilibrium value is reached when the force due to the potential barrier equals the diffusion force. Now note what happens to the few hole-electron pairs formed by thermal agitation. If a pair is formed in the n-type material just to the right of the junction, the potential difference provides a force of attraction for the hole; hence, it moves to the left. Similarly, for a pair formed on the left,

the electron is pulled to the right. These directions of motion are opposite to those involved in the diffusion process. When equilibrium occurs, the net value of the electron migration and the net value of the hole migration are zero across the junction. The disappearance of holes and free electrons through recombination can be considered qualitatively as a reduction in the number of pairs formed.

The value of the equilibrium potential difference for a germanium diode is a fraction of a volt. This is a measure of the energy required to move a hole from left to right or an electron from right to left.

2.7   Junction Diode

The form of the current-versus-voltage curve for a semiconductor junction diode can be deduced from the principles just discussed. Suppose the voltage source in Fig. 2.6 is adjustable over a suitable range of values of either polarity. The desired diode curve is obtained by measuring the value of current i for specific values of applied voltage v.

For positive values the source voltage opposes the p-n junction potential. Lowering the potential barrier by this means permits more holes to migrate from left to right and more electrons from right to left. The additive effects of the hole and electron motions constitute the current i. For small values of v the effective potential barrier is lowered only slightly; hence, only a few carriers get through. These are the ones having maximum thermal energy. For larger values of the source voltage more carriers can overcome the junction potential. Since the statistical distribution of the energies of holes or electrons is exponential,* more and more charges migrate across the junction as the barrier is lowered and we may therefore expect current i to increase exponentially with an increase in the positive value of v.† The exponential relation should hold over a fair range of values of current and voltage. For large values of these variables the heating effect of the current will modify the behavior. Also, the ohmic resistance of the semiconductor material in the diode body (away from the junction) will modify the curve of i vs. v toward a more nearly linear resistive curve.

For slightly negative values of v the source effectively aids the junction potential in retarding diffusion of the majority carriers provided by the impurity content. However, this reverse voltage aids minority carriers

---

* At any particular temperature the number of particles that have a specific energy decreases exponentially with increasing values of energy.

† The constant $V_0 = kT/e$ is of the order of 0.025 volts for germanium at room temperature.

created by the formation of hole-electron pairs. For a pair created in the p-type material the electron migrates to the right. The particles that remain in the p-type or n-type material are the same as those provided by the impurities. The polarity of the applied potential pulls majority
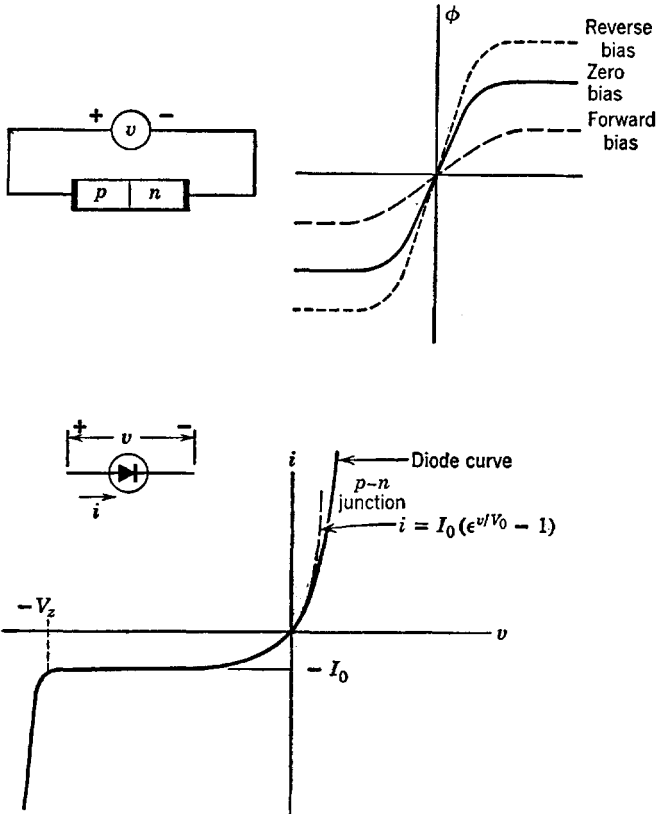
Fig. 2.6. Current versus voltage curve for a p-n junction diode.

carriers away from the junction and tends to prevent their diffusing across the junction. Thus, the current carried across the junction depends entirely on the thermally-created pairs of carriers and not on the applied voltage. After the voltage is a few tenths of a volt negative, the current approaches a small constant value $(-I_0)$. The magnitude of I_0 is likely to be only a few microamperes and increases with temperature, but for a fixed temperature the current is virtually independent of voltage from a few tenths of a volt negative to a few volts or a few tens of volts negative.

As the voltage is made more negative, the electric field acting on the charged particles increases correspondingly. For some value of v, say $-V_z$, a large increase in current occurs. This increase results from the fact that the high value of electric field accelerates the minority carriers proportionately more. The kinetic energy thus acquired by the holes or electrons can be imparted to atoms in the crystal lattice by collision. Increasing the energy of an atom in this fashion may cause it to break an electron bond and thus free another hole-electron pair. At the voltage $-V_z$, this process, which is called the avalanche effect, becomes cumulative, and a very large change in current occurs for very little increase in negative voltage. The breakdown voltage $-V_z$ is an important parameter in many applications of junction diodes.
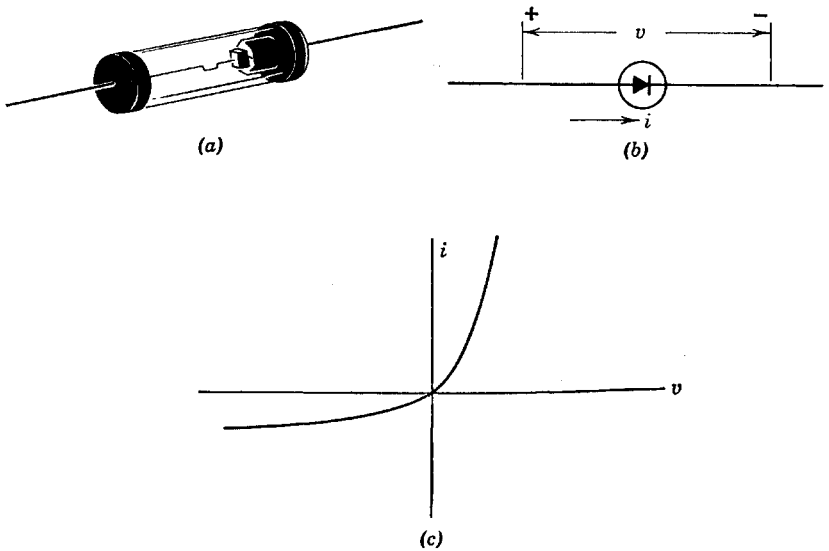
Fig. 2.7. Point-contact diode. (a) Sketch; (b) Symbol; (c) i vs. v curve.

2.8 Point-Contact Diode

The point-contact rectifier consists of a semiconductor on which the tip of a fine wire rests as shown in Fig. 2.7(a). The curve of current versus voltage shown in Fig. 2.7(c) is qualitatively similar to that of the junction diode. However, for a given positive voltage, the point-contact conducts less current than the junction diode. Also, for a given reverse voltage the point-contact conducts somewhat more current. Furthermore, as negative voltage increases, the reverse current tends to

increase rather than remaining nearly constant. The sharp break in the junction diode curve at $-V_z$ does not occur in point-contact diodes, since heating of the sharp point occurs at much lower voltages and causes a gradual increase in the conductance in the negative direction.

(a)

(b)
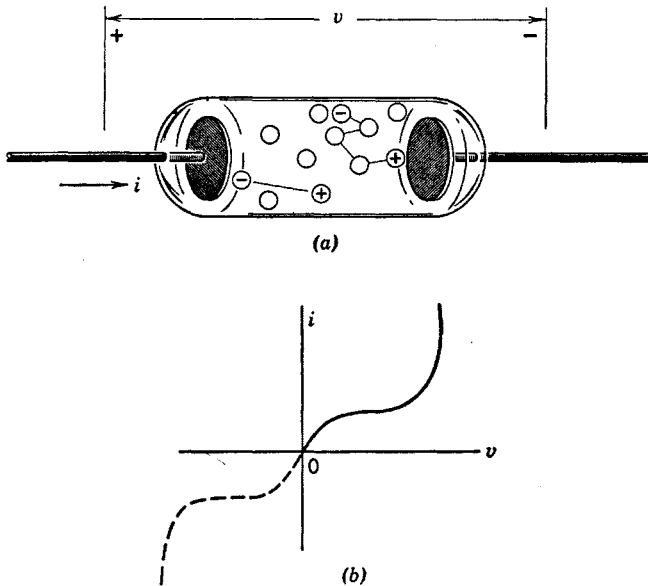
Fig. 2.8. Properties of a gas discharge. (a) Gas-filled diode; (b) Current versus voltage curve.

2.9 Conduction in Gases

Whereas in metals the atoms are so closely packed that the orbits of valence electrons overlap, the atoms in a gas can usually be considered as isolated particles. When pressures and temperatures begin to approach those for which the gas becomes a liquid, this simplifying assumption no longer applies. However, gas diodes contain gas or vapor at relatively low pressures for which the assumption applies fairly well.

The sketch shown in Fig. 2.8 illustrates a simple electron tube. Assume the sealed glass tube contains a gas at low pressure. For example, let the gas be one of the inert elements, such as argon or neon. The electric field, $-v/d$, owing to an applied voltage v, should have no particular effect on the gas, since the atoms are electrically neutral. The atoms

move in random fashion because they have thermal energy. As they move about, the atoms occasionally collide with each other. Some of these collisions may cause one or both of the colliding atoms to ionize; that is, one or more valence electrons receives sufficient energy from the collision to escape from the atom. Thus, ionization results in the creation of one or more free electrons and a positively charged ion.

In the absence of an applied field, the free electrons and ions move about in the gas and eventually recombine. The total number of charged carriers in the tube will depend on the gas temperature, the pressure, and the amount of high-energy radiation present. Cosmic rays or gamma rays passing through the tube may leave a trail of ionized atoms and free electrons. With a voltage v applied to the electrodes, the electrons tend to move toward the positive electrode. Since the mass of the ion is many times greater than that of the electron, the force imparts a high acceleration to the electron but a relatively small one to the ion.

The electron accelerates until it collides with an atom or ion (or the left electrode). The probability that the electron will cause an ionizing collision depends mainly on the pressure. If the pressure is extremely low, there are few atoms in the tube, and the chances of hitting any of them are small. Nevertheless, for a given applied voltage the electron accelerates freely for a longer time (on the average); hence, if a hit occurs the probability of ionization is high. If the pressure is too great, the electron collides with atoms so frequently that the probability of ionizing one is negligibly small. Since the electron is so light compared with an atom, it is deflected in random manner by each collision. Thus, the acceleration due to the applied field is lost at each collision (on the average). Between extremes of pressure a value exists for which the probability of ionization by collision is a maximum.

Now let us consider the form of the current versus voltage curve likely to be obtained for a simple tube similar to the one shown in Fig. 2.8. The rate of formation and of recombination of the free charges in the tube establishes the equilibrium number present. For voltage equal to zero the current is zero. As voltage is increased from zero, some electrons are accelerated through the tube and a small current is observed. A further increase of voltage begins to provide sufficient electron energy to ionize a few atoms. This increases the number of free charges present and permits a somewhat greater number of electrons to cross the tube. Average velocity also increases. The current is the product of the number of electrons, the electronic charge, and the velocity. Thus, the current should increase more rapidly than linearly with increasing voltage, because both the number and the velocity of electrons increase with the applied voltage.

A continued increase of the applied voltage increases the kinetic energy of the free electrons. A certain minimum of energy must always be inparted to a bound electron to free it and ionize the atom. However, with sufficient energy available, the ejected electron may have sufficient energy to ionize another atom by collision without first gaining energy from the applied field. Under these conditions the process is regenerative, and the result is glow discharge (a form of electrical arc). The available current increases tremendously with no increase in voltage. In fact, once the discharge has been started the voltage drop across the tube diminishes below the value required to initiate the discharge.

From the symmetry of the electrodes in the gas-filled tube sketched in Fig. 2.8(a), it is obvious that reversal of the polarity of v will result in a similar curve with the direction of current reversed. Thus, although this tube has a highly nonlinear i vs. v curve, the curve is bidirectional.

2.10   Electron Emission from Metals

The large numbers of free electrons in a metal give the metal its high conductivity. Although free to migrate within the metal, these electrons are restrained from leaving the metal by the potential barrier known as the work function. For most metals the work function has a value of a few electron volts. An electron volt is 1.6×10^{-19} joules. The electronic charge is 1.6×10^{-19} coulombs, hence the work function W in electron volts corresponds numerically to the voltage v. The several methods for supplying the necessary energy for electron emission constitute an important aspect of electronic-device behavior. For emission to take place, it is always necessary to supply the difference between the required escape energy (work function) and the energy already possessed by the electron.

Figure 2.9(a) shows the binding force holding an electron to a metal as a function of the distance from the surface. The surface is assumed to be at $x = 0$ with the metal to the left and free space to the right of the origin. When the electron is at point A, it is embedded in the metal and (on the average) feels no net force. At point B the metallic atoms are all to the left of the electron in question; hence, there is a net binding force tending to hold the electron to the metal. Right at the surface, the force on the electron is due primarily to positive charges on atoms nearest the electron. In going outward from B to C, the force increases somewhat. Although the increased distance diminishes the force due to immediately adjacent charges, the influence of positive charges over a greater part of the surface tends to increase the total

force. The distances from A to B and B to C are small multiples of atomic dimensions.
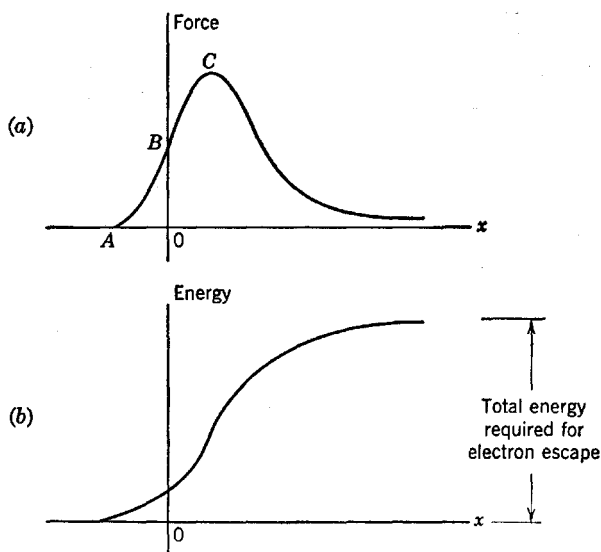
Fig. 2.9. Force and energy involved in electron emission.

The sketch of Fig. 2.9(a) shows the applied force required to hold an electron in equilibrium at any value of x. The total energy required to remove an electron from the metal is therefore the area under the curve. The integral of the force curve is sketched in Fig. 2.9(b).

More precisely, the work function is defined as the difference of two energies $(W_x - W_0)$ where W_x is the kinetic energy associated with the x component of electron motion (normal to the surface). The energy W_0 is the maximum value of energy (associated with the x component of electron motion) at a temperature of $0°K$.

2.11   Field Emission

Since free electrons are bound to a metal by electrical forces, let us consider electrical force as a means of releasing them from the metal. Emission of electrons as a result of an electrical force is called field emission, since the force on an electron is the product of charge and electric field intensity. The field intensity required is of the order of 10^{10}

volts per meter. In order to obtain such large field intensities with nominal voltages between electrodes, special geometry is required.

As an example, consider the cylindrical electrode structure shown in Fig. 2.10(a). With a voltage applied between the two cylinders, the electric field intensity is given by the expression in Fig. 2.10(b). Since there is circular symmetry, the lines of force are radial and the intensity varies inversely with r for $r_1 < r < r_2$. This expression applies only if
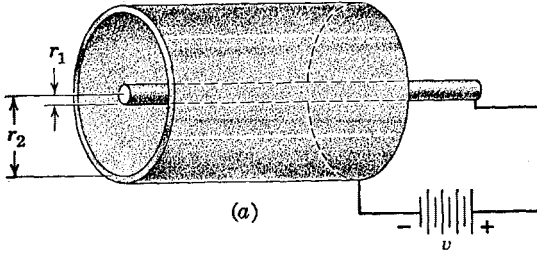
$$(b) \quad E_r = \frac{v}{r \ln \frac{r_2}{r_1}} \text{ for } r_1 < r < r_2$$

Fig. 2.10. Field emission from a cold cathode.

the space between the electrode is evacuated or contains no charged particles. For $v = 100$ volts, $r_1 = 0.01$ cm and $r_2 = 1$ cm, $r_2/r_1 = 100$, $\ln r_2/r_1 = 4.6$, and the maximum field intensity (which occurs for $r = r_1$) is approximately 2×10^5 volts per meter. This value is still far from that required for field emission. The inclusion of a fine, sharp-pointed wire or annular disc extending radially inward from the outer cylinder (cathode) can bring the field intensity up high enough for field emission. The so-called cold-cathode, gas-discharge voltage regulators use such an electrode structure. The high field intensity at the sharp point or edge causes electrons to be emitted to initiate the glow discharge. When the interelectrode space is filled with ionized atoms and free electrons, the field is radically modified, and the simple expression for electric field intensity no longer applies.

2.12   Thermionic Emission

The type of electron source most commonly used in vacuum tubes or gas-filled tubes is the heated cathode. The temperature of the emitting material is raised until a suitable number of the free electrons have thermal energies comparable to or exceeding the work function. Under

these conditions an electron leaves the cathode with a kinetic energy that depends on the amount by which its thermal energy exceeds the work function. Most of the electrons emerge from the hot cathode with a velocity corresponding to energies from zero to a few electron volts. Electrons are thus released whether or not an external field is applied. This means of releasing electrons is called thermionic emission.

The cathode may be a filament of tungsten or thoriated tungsten wire heated to a temperature of 1500°K or more by passing an electric current through it. Indirectly heated cathodes are more efficient in terms of thermionic current available per watt of heating power. This type of cathode consists of a filament or heater surrounded by a small cylindrical sleeve of nickel or similar metal. The outer surface of the sleeve is coated with a material like barium oxide, strontium oxide, calcium oxide, or a mixture of oxides, which are better emitters than tungsten or thoriated tungsten because they have lower work functions.

The efficiency of various cathode materials is a critical function of the detailed processing of the material. As a result, considerable variations occur even with the same type of material. The following empirical equation gives the form of the temperature dependence of thermionic current density.

$$J = AT^2 \epsilon^{-b/T} \text{ amperes/unit area} \tag{2.1}$$

In this equation b is the work function in temperature units,* T is temperature in degrees Kelvin, and A is an empirical constant. The thermionic emission equation indicates that a saturation value of emission current exists for a specific value of cathode temperature.

2.13 Temperature-Limited Emission in a Vacuum Diode

The effect of temperature-limited emission on the behavior of a vacuum diode is illustrated by Fig. 2.11 where (a) shows the complete symbol† for a vacuum diode, and (b) shows the i vs. v curves for various values of T. These are established by adjusting the voltage E applied to the heater. The currents I_1, I_2, and I_3 are the product of cathode area and the corresponding current densities given by Eq. 2.1. Normal

---

* $b = e\phi/k$, where e is electronic charge (1.6×10^{-19} coulombs), ϕ is voltage equivalent of work function, and k is Boltzmann's constant (1.38×10^{-23} joules per degree Kelvin).

† The more commonly used symbol [Fig. 2.11(c)] shows only the anode and cathode. The heater and its power source are usually omitted, since it is assumed that the heater will be operated at rated voltage. The term diode is obviously a misnomer unless the heater and cathode are considered as a single electrode.

diode operation does not include temperature-limited operation, since other design ratings, such as anode power dissipation, are usually exceeded long before saturation is reached.

(a)   (b)

(c)
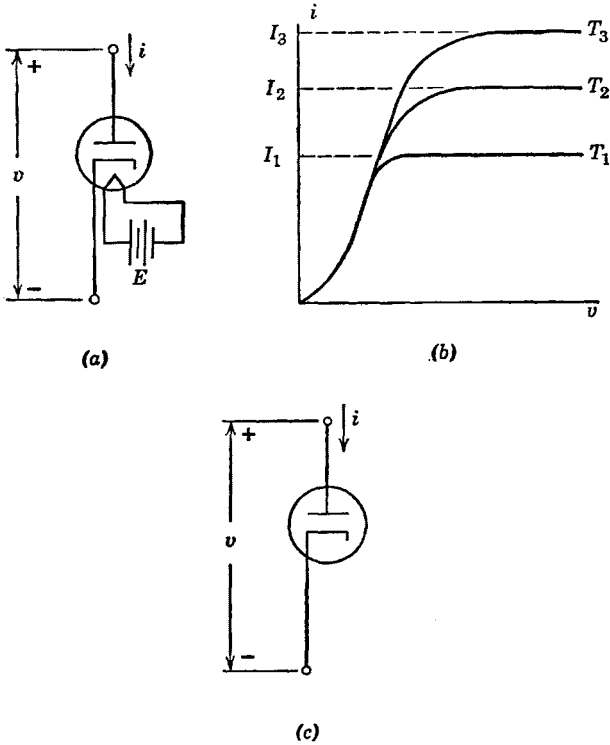
Fig. 2.11. Temperature-limited emission in a vacuum diode.

A vacuum diode designed to operate in the saturation region may be considered as a control valve with heater temperature as the control parameter. Since thermal time constants are large, the response to changes in T (effected by adjusting heater power source E) for such a device is too sluggish to compete with other available control valves, such as triodes or transistors.

2.14 Vacuum Diode

The electrode structure of a typical vacuum diode is shown in Fig. 2.12. With the cathode cold, the electrodes constitute a small capaci-

tance (a few micromicrofarads). The capacitive properties of the diode need be considered only for high-frequency operation. With the cathode heated, the diode is primarily a nonlinear conductance even for anode currents considerably less than the saturation value. This nonlinearity is the result of a cloud of electrons (called space charge) that accumulates near the cathode. The presence of this cloud of negative charges depresses the potential near the cathode to the same value as the cathode potential or even slightly below.

The potential near the cathode depends to a great extent on initial velocities of electrons. For example, suppose both the cathode and the anode are connected to the same potential (say, zero). If we postulate an ideal cathode, namely one that brings all electrons just to the energy required to leave the cathode, electrons will emerge with zero energy or zero velocity, hence none will actually leave. The potential will then remain zero throughout the interelectrode space. If this ideal cathode is heated further, all electrons will emerge with a finite initial velocity and some will reach the anode. However, the presence of electrons between the electrodes depresses the potential below zero and tends to retard electrons. From the region of negative potential, some electrons will be accelerated toward the anode while others will be returned toward the cathode.

Fig. 2.12. Typical vacuum diode.

An actual cathode produces a distribution of initial velocities, since the electrons in the metal have an exponential distribution of energies. Again the space charge will depress the potential below cathode potential, but a small anode current will exist even with anode voltage zero. If we make the anode voltage negative with respect to cathode, this small current will decrease exponentially with the negative voltage. This small exponential tail in the negative-voltage region can usually be neglected, so that for most engineering purposes we can assume the anode current to be zero for zero anode voltage.

As the anode voltage is raised to positive values the potential minimum is raised slightly. Thus more electrons tend to reach the anode and fewer electrons remain in the space-charge cloud.
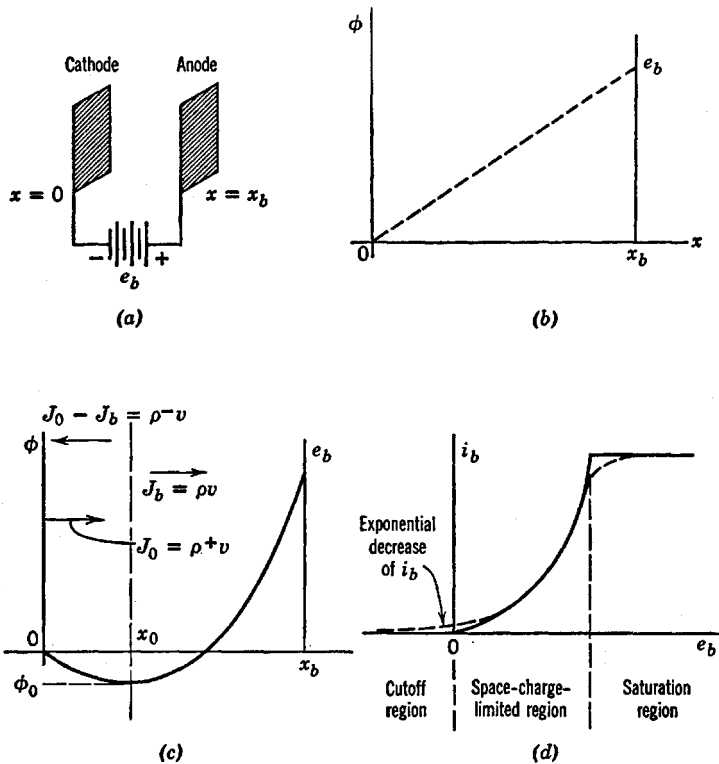
Fig. 2.13. Parallel-plane diode.

2.15 The Parallel-Plane Diode

In normal operation the current versus voltage curve for a vacuum diode is quite closely approximated by an expression of the form $i_b = K e_b{}^{3/2}$. The quantity K, called the perveance, depends upon electrode geometry, but the functional relation between e_b and i_b may be shown to be independent of geometry. Thus, the i_b vs. e_b relation for current limited by space charge in a vacuum tube is usually called the three-halves power law.

A quantitative derivation of the behavior of a vacuum diode is simplified by choosing a parallel-plane electrode structure. As indicated by Fig. 2.13(a), we shall consider the conduction per unit area between parallel plates of infinite extent. This implies no fringing of fields; that is, no variations except as a function of x. For a given anode voltage e_b, the potential distribution between electrodes is therefore

linear in the absence of electrons (cold cathode) and the electric field intensity (slope of the potential curve) is a constant. We shall also assume that for any given temperature, the cathode emits electrons at a constant rate in a direction normal to the cathode plane and that all electrons have the same initial kinetic energy.

With the cathode emitting copious quantities of electrons, the potential distribution between the cathode ($x = 0$) and the anode ($x = x_b$) takes on the form shown in Fig. 2.13(c). The electron cloud depresses the curve below the straight line that exists under cold-cathode conditions. Assuming the general form of the curve to be reasonable, we shall calculate the mathematical expression for the curve from electron behavior in the interelectrode space.

Electrons leave the cathode with an initial velocity v_0. Assuming each electron acts independently, we can neglect the loss of energy by collision with other electrons. A change in kinetic energy must therefore appear as a change in the potential energy $e\phi$, where e is the electronic charge and ϕ is the electric potential. Thus, at any distance x from the cathode:

$$e\phi = \frac{m}{2}(v^2 - v_0{}^2) \tag{2.2}$$

where m is electron mass and v is velocity. The initial kinetic energy is related to the initial velocity v_0 by the expression:

$$e\phi_0 = \frac{mv_0{}^2}{2} \tag{2.3}$$

so that Eq. 2.2 can be expressed as follows:

$$\frac{mv^2}{2} = e(\phi + \phi_0) \tag{2.4}$$

Differentiating Eq. 2.2 or 2.4 with respect to x and remembering that $v = dx/dt$ yields:

$$e\frac{d\phi}{dx} = mv\frac{dv}{dx} = m\frac{dv}{dt} \tag{2.5}$$

The first term is the force acting on the electron (product of charge and electric field). The last term is the product of mass and acceleration.

As an electron leaves the cathode, it encounters a retarding electric field. If the depth of the potential minimum is just equal to ϕ_0, as shown in Fig. 2.13(c), the electron comes to rest at x_0. This is a point of unstable equilibrium, and the electron either falls back to the cathode or begins the longer fall to the anode. (Remember that electrons "fall" *up*

a potential hill because of the negative charge.) Our problem is to calculate the fractional number of emitted electrons that reach the anode.

The total electron current passing through unit area of any plane for which x is constant is measured by the charge crossing the plane in unit time. To the right of x_0 the current density J_b is ρv, where ρ is the local space-charge density per unit volume. Both ρ and v vary with x, but continuity of current demands that their product be independent of x when the flow is in equilibrium.

To the left of x_0, current can be considered to consist of two components, as indicated by Fig. 2.13(c). The emission current density J_0 is $\rho^+ v$, and the returned current density $(J_0 - J_b)$ is $\rho^- v$. Equation 2.4 shows that for a given x, the velocities are the same for electrons traveling in either direction. This is physically reasonable, since the acceleration from x_0 back to the cathode is due to the same potential curve as the deceleration from cathode to x_0. Thus, the total space-charge density at any point on the left of x_0 is related to the current densities by the equation:

$$2J_0 - J_b = (\rho^+ + \rho^-)v \qquad (2.6)$$

The relation between the potential curve and the space charge that generates it is provided by Poisson's equation, which in one dimension reduces to:

$$\frac{d^2\phi}{dx^2} = -\frac{\rho}{\epsilon_0} \qquad (2.7)$$

since both ϕ and ρ are independent of y and z. In this equation, ρ is the total space-charge density at any point, and ϵ_0 is the permittivity of free space.

Relation $J_b = \rho v$ (independent of x) together with Eqs. 2.4 and 2.7, after elimination of ρ and v, yields

$$\frac{d^2\phi}{dx^2} = \frac{J_b}{\epsilon_0}\sqrt{\frac{m}{2e}}\,(\phi + \phi_0)^{-\frac{1}{2}} \qquad (2.8)$$

Solution of this equation subject to the boundary conditions:

$$\phi = -\phi_0, \quad \frac{d\phi}{dx} = 0 \quad \text{at} \quad x = x_0$$

$$\phi = e_b \quad \text{at} \quad x = x_b \qquad (2.9)$$

yields the result:

$$J_b = A\,\frac{(e_b + \phi_0)^{\frac{3}{2}}}{(x_b - x_0)^2} \qquad (2.10)$$

where

$$A = \frac{4\epsilon_0}{9} \sqrt{2e/m} = 2.34 \times 10^{-6} \text{ in MKS units} \qquad (2.11)$$

The result given by Eq. 2.10, which applies to the right of x_0, shows a three-halves power dependence on voltage and an inverse square dependence on distance. This is consistent with the dimensions of Eq. 2.8. In fact, by dimensional analysis it can be shown that a three-halves power law holds for any electrode geometry as we shall see in Art. 2.16.

For the region to the left of x_0, the same method is used. The total space charge $(\rho^+ + \rho^-)$ is used in Poisson's equation together with the boundary condition $\phi = 0$ at $x = 0$. The conditions $\phi = -\phi_0$ and $d\phi/dx = 0$ at $x = x_0$ also apply. The latter insure continuity of potential and electric field at x_0. Using these factors to modify Eq. 2.10 yields the following result:

$$2J_0 - J_b = A \frac{\phi_0^{3/2}}{x_0^2} \qquad (2.12)$$

Equations 2.10 and 2.12 frame the problem. Given emission current J_0 and emission energy ϕ_0, plate voltage e_b and electrode spacing x_b, find current density J_b and the position of the minimum x_0. Practical considerations lead to a few simplifying assumptions. At normal operating temperatures ϕ_0 for an oxide coated cathode is usually a few electron volts. The plate voltage required to saturate a typical diode may be several hundred volts. At saturation $J_b = J_0$; hence, Eqs. 2.10 and 2.12 become

$$J_0 = A \frac{(e_b + \phi_0)^{3/2}}{(x_b - x_0)^2} = \frac{A\phi_0^{3/2}}{x_0^2} \qquad (2.13)$$

Since $e_b \gg \phi_0$ it follows that $x_0 \ll x_b$. For example, if $e_b = 100\phi_0$, then x_0 is approximately $x_b/30$. When e_b is slightly negative, $J_b = 0$; hence, Eq. 2.12 yields:

$$2J_0 = \frac{A\phi_0^{3/2}}{x_0^2} \qquad (2.14)$$

Comparison with Eq. 2.13 shows that x_0 is roughly one-third smaller at cutoff $(J_b = 0)$ than at saturation $(J_b = J_0)$. If the maximum value of x_0 is $x_b/30$, then $(x_b - x_0)^2$ varies only about two per cent over the entire range of operation from cutoff to saturation. Ignoring this small variation in x_0, the space-charge-limited diode law becomes:

$$i_b = k(e_b + \phi_0)^{3/2}, \text{ for } e_b \geq 0 \qquad (2.15)$$

where k and ϕ_0 are constants. The curve is sketched in Fig. 2.13(d). This relation, called the Child–Langmuir law or three-halves power law, agrees closely with experimental tube curves over a wide range of positive voltages. Empirically determined values of ϕ_0 run from less than a volt to a few volts. Considered as an empirical constant, ϕ_0 accounts for contact-potential difference as well as initial velocities.

Actually, because of the exponential distribution of emission velocities, the plate current does not cut off abruptly as shown in Fig. 2.13(d). Instead, i_b remains slightly positive and decreases exponentially as e_b is made negative; hence, the cutoff region should more realistically be called the exponential region.

An experimental curve also shows a gradual transition from space-charge limit to saturation limit as indicated by the dotted line in Fig. 2.13(d). This is due to the distribution of initial velocity, and to non-uniformities in the cathode temperature, the cathode potential, and the interelectrode spacing.

2.16   Generalized Three-Halves Power Law

The parallel-plane electrode structure considered in the previous article is not as common in vacuum diodes as is the concentric cylinder structure. Anodes with rectangular or elliptical cross sections are also used with cylindrical cathodes. Nevertheless, except for the value of the constant multiplier, the three-halves power law applies for any electrode geometry. For concentric cylinders the law may be derived with but slightly more difficulty than for the parallel-plane case. For other geometries the derivation is analytically very difficult, if not impossible. The assumptions that are helpful in demonstrating plausibility of the three-halves power law for a diode with arbitrary electrode cross sections are no more restrictive than those made for the parallel-plane structure.

Two physical effects govern the behavior of the charged particles moving in an electric field, once the potential distribution and the current have reached equilibrium. These are as follows:

(a) Charge distribution determines the potential. This means that if the position of all charges in space is known, the potential at any point can be calculated by superposition of the potential due to each charge.

(b) Potential distribution determines charge motion. This means that if the potential is known everywhere, the motion of charges can be calculated, assuming zero velocity at the point of origin. This effect may be expressed as $\phi e = \frac{1}{2}mv^2$, where ϕ is potential, e is electronic charge, m is electron mass, and v is electron velocity.

Now consider a diode with arbitrary electrode structure. For convenience, take the cathode potential as reference and let it be zero. Let a voltage e_b be applied to the anode relative to the cathode and allow current to reach an equilibrium value. In accordance with (b), velocity is proportional to the square root of potential $(v \sim \sqrt{\phi})$ at any point in the interelectrode space. In accordance with (a), doubling charge density ρ everywhere doubles ϕ everywhere $(\phi \sim \rho)$. Under equilibrium conditions there must be continuity of current across any equipotential surface. In particular, the anode is an equipotential surface, and the anode current must be proportional to the charge density (ρ) times the velocity (v) at the anode. Since $\rho v \sim \phi \sqrt{\phi}$, we have $i_b \sim e_b^{3/2}$. The generality of this result is of considerable importance, since it establishes the idealized behavior of any vacuum diode* over a fairly wide range of current and voltage values, namely from the cutoff or exponential region to the saturation region. Thus, the three-halves power law unifies the interpretation of measured vacuum tube curves.

2.17 Gas-Filled Thermionic Diode

The discussion of conduction in a gas indicated that a discharge could be established between two electrodes as shown in Fig. 2.8. A few stray electrons, producing ionization by collision, initiate a discharge that can sustain itself. If one of the electrodes is a hot cathode, the large number of thermionic electrons thus made available will modify the properties of the discharge tube. With parallel-plane electrodes (both cold) the curve of current versus voltage is symmetrical for positive or negative v and i. With a hot cathode, a preferred direction of conduction exists, and the discharge tube becomes a rectifier somewhat analogous to the vacuum diode. Gas-filled diodes usually contain one of the inert gases, such as argon or neon. The gas pressure within the tube is made low compared with atmospheric pressure in order to make the breakdown voltage low. Sometimes a drop of liquid mercury is included. In this case the mercury vapor provides the atoms which are to be ionized.

Suppose electrons are emitted from the cathode and attracted toward the anode of a parallel-plane structure by a positive value of e_b. As the electrons accelerate, they ionize gas molecules by collision, releasing additional electrons and positive ions. The electrons, having high mobility, are swept rapidly toward the anode. The positive ions, being

* It can also be extended to multi-electrode tubes.

much larger and heavier, tend to move slowly toward the cathode. As indicated by Fig. 2.14, the presence of positive ions in the interelectrode space tends to raise the potential curve.
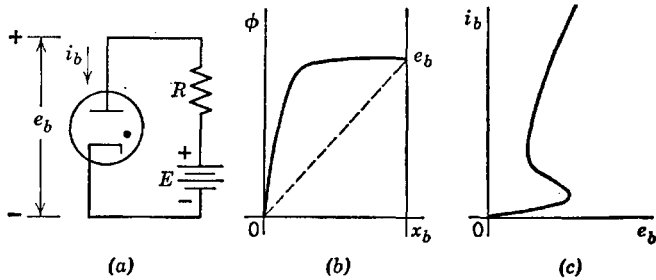

Fig. 2.14. Thermionic gas diode.

The symbol for a gas diode is shown in Fig. 2.14(a), and the expected form of the potential curve is shown in (b). The current versus voltage curve is shown in (c). The form of this curve differs from that of the cold-cathode discharge. Since there is a large quantity of electrons available from the thermionic cathode, the low-current conduction region shown in Fig. 2.8(b) does not exist here. Note, however, from Fig. 2.14(c) that once the discharge is initiated, the current will tend to become excessively large if a fixed voltage is applied. However, if the current is controlled, the voltage adjusts itself to the appropriate value. The simple circuit of Fig. 2.14(a) might be used to determine the i_b vs. e_b curve experimentally, provided that E and R are adjustable. The resistance R limits the current to values safe for the diode.

The form of the potential curve shown in Fig. 2.14(b) shows a high field intensity near the cathode. This field accelerates the electrons almost to final velocity (corresponding to e_b); hence, ionization of gas molecules starts near the cathode. Over the majority of the space between electrodes, the potential variation is small. This means small field intensity and, hence, electrons tend to proceed at almost constant velocity. Meanwhile, positive ions in this region, being subject to low field intensity and having low mobility, drift slowly toward the cathode. This portion of the interelectrode space is called the plasma region. Here, under equilibrium conditions, ions and electrons are formed by collision and some may be lost by recombination. Near the cathode, the high field intensity accelerates emission electrons toward the plasma and then to the anode; concurrently, positive ions are accelerated toward the cathode, where they recombine with electrons.

The plasma is highly conducting because of both the positive ions and the electrons, which tend to cancel each other insofar as space charge is concerned, however the ion and electron currents are additive. Once the plasma has been established, the potential curve varies only slightly with the magnitude of the current.

The large conduction current possible in a gas diode with very low voltage drop makes the device a very efficient rectifier in comparison with a vacuum diode. The highly conducting plasma effectively places the anode very close to the cathode. This helps to account for the low anode voltage required to sustain a large current.

2.18   Excitation by Collision

The conduction in a gas-filled tube is called a glow discharge because of the characteristic light given off. For neon the color is a deep red, for argon violet, and for mercury vapor a greenish blue. The light is emitted by the conducting gas following excitation of the atoms by collision with fast-moving electrons. Excitation involves an increase in the energy of an electron. The electron then loses that energy by emitting it in the form of electromagnetic radiation.

According to quantum theory, the electrons associated with an atom can assume only a finite number of distinct values of energy. Only one electron at a time can assume a specific state. The lower energy states belong to the inner shells of electrons, and, in general, electrons seek the lowest energy states. Valence electrons have the highest energies and, hence, require the least additional energy for escape. Ionization normally occurs by removal of a valence electron. However, suppose a high-speed electron gives up enough energy in colliding with an atom to raise the energy of one of the electrons in an inner shell sufficiently to eject it from the atom. This leaves a vacancy in one of the low-energy states, and the atom or molecule is said to be excited as well as ionized. One of the high-energy electrons will then fall to the low-energy state, and the decrease in electron energy will result in the emission of a quantum of light (photon). The low-energy electron need not be completely ejected from the atom for excitation to occur. Its energy could be raised from the normal level to a higher level within the atom. Upon losing that energy, a photon is emitted. The frequency of the light emitted is determined by the product of the energy difference between the two states involved and a natural constant known as Planck's constant, $h = 6.63 \times 10^{-34}$ joule sec.

2.19   Photoelectric Emission

Photoelectric emission is the converse of excitation by collision.   When an electron collides with an atom, giving up some or all of its energy, the result may be ionization, or excitation which then results in photon radiation.   The converse applies to the situation in which photons of light impinge on the atoms of a material such as cesium, calcium, or aluminum.   If the quantum of energy $h\nu$ for light having a frequency ν imparts sufficient energy to an electron to overcome the work function of the material, photoelectric emission occurs.   For most materials the photoelectric work function is nearly equal (but not identical) to the thermionic work function.

The materials mentioned above are good photoelectric emitters and are used as the cathode structure in a vacuum or gas-filled tube together with an anode to form a diode.   The anode is connected to a positive potential relative to the cathode and thus attracts the emitted electrons. Photoelectric diodes (vacuum or gas-filled) may be compared in a general way to thermionic diodes.   With light of constant intensity and frequency distribution impinging on a photoelectric diode, the current versus voltage curve is roughly similar to that of a thermionic diode with constant cathode temperature.   However, constant cathode heating power (and therefore temperature) represents the normal mode of operation for thermionic diodes, whereas photoelectric diodes are more often used under conditions of varying light intensity and frequency.

The sketch shown in Fig. 2.15(a) indicates the electrode geometry used in a common type of vacuum phototube.   The cathode is part of a cylinder and the anode a wire placed before it.   Light impinging on the cathode releases electrons which are attracted to the anode if the latter has a positive potential with respect to the cathode.   Photoelectric emission from typical cathode materials yields electron currents of the order of microamperes (thermionic emission from an oxide-coated cathode may yield currents of many milliamperes or even amperes). The successively higher values of saturation current in Fig. 2.15(d) correspond to increasing levels of light intensity.   Specific values of current for a given light intensity depend upon the light frequency or wavelength, since the metals used as photoelectric cathodes show a definite peak in sensitivity.   The peak for cesium occurs at a light wavelength of about 5400 angstroms, whereas sodium and potassium show peaks at about 4200 and 4400 angstroms, respectively.

The symbol for a vacuum phototube is shown in Fig. 2.15(b).   If a dot

is placed within the circle, the resulting symbol represents a gas-filled tube [Fig. 2.15(c)]. Anode curves for a gas-filled tube are much like those for a vacuum phototube until gas breakdown occurs. Then the available current increases rapidly with voltage [Fig. 2.15(e)]. In fact,
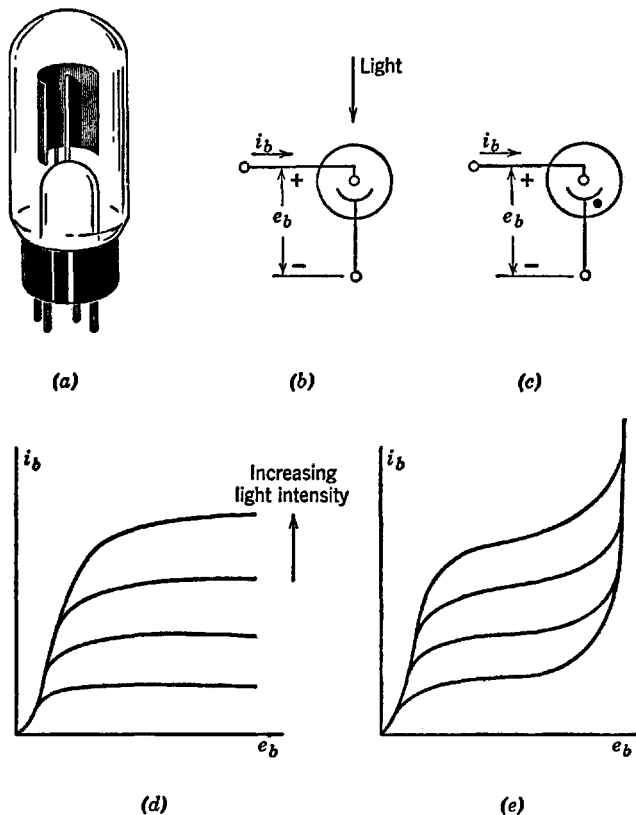

Fig. 2.15. Properties of phototubes. (a) Phototube; (b) Symbol for phototube; (c) Symbol for gas-filled phototube; (d) Anode curves for vacuum phototube; (e) Anode curves for gas-filled phototube.

as in the case of gas-filled thermionic diode, a series resistance is necessary to limit the current to reasonable values. The gas-filled tubes are useful in operating relays which then execute a function, such as door-opening, under control of a light signal.

When it is desirable or necessary to increase the current output from a vacuum phototube, a device called a secondary emission multiplier is sometimes used.

2.20   Secondary Emission

Electrons can be ejected from a metal, glass, or other material under bombardment by electrons. Thus, an electron impinging on a metal may impart sufficient energy to atoms near the surface to eject one or more electrons. This process, called secondary emission, often plays an important part in vacuum tube operation. For example, a secondary
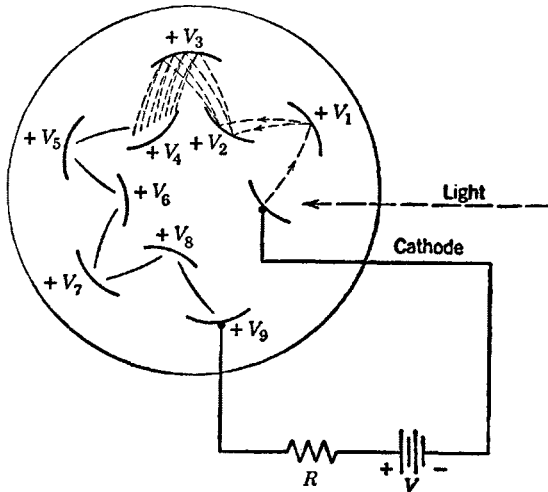

Fig. 2.16.   Secondary-emission multiplier.

emission multiplier associated with a vacuum phototube is shown in Fig. 2.16. The cathode and the anode form a diode and the other electrodes, going counterclockwise from cathode to anode, are connected to successively higher voltages (usually by means of a resistive voltage-divider). An electron emitted from the cathode is attracted to the first electrode where it may release one or more secondary electrons. The higher potential of the next electrode attracts these secondaries. If each electron releases r secondaries and there are n electrodes on the way from cathode to anode, the anode current will be r^n times the photoelectrically-emitted current. The secondary emission ratio r usually falls in the range of two or three secondary electrons per primary electron.

In many vacuum tube applications secondary emission is undesirable rather than useful. If bombardment of an anode or other electrode in a tube results in the release of secondary electrons, the characteristics of

the tube may be seriously affected. This point will be discussed further in Chapter 7 in connection with multi-electrode tubes.

SUPPLEMENTARY READING

L. B. Arguimbau, R. B. Adler, *Vacuum Tube Circuits and Transistors*, John Wiley and Sons, New York, 1956.

H. Bruining, *Physics and Applications of Secondary Electron Emission*, McGraw-Hill, New York, 1954.

W. G. Dow, *Fundamentals of Engineering Electronics*, John Wiley and Sons, New York, 1952.

D. V. Geppert, *Basic Electron Tubes*, McGraw-Hill, New York, 1951.

T. S. Gray, *Applied Electronics*, 2nd edition, John Wiley and Sons, New York, 1954.

Willis W. Harman, *Fundamentals of Electronic Motion*, McGraw-Hill, New York, 1953.

L. P. Hunter, editor, *Handbook of Semiconductor Electronics*, McGraw-Hill, New York, 1956.

R. D. Middlebrook, *An Introduction to Junction Transistor Theory*, John Wiley and Sons, New York, 1957.

Karl R. Spangenberg, *Fundamentals of Electron Devices*, McGraw-Hill, New York, 1957.

Robert L. Sproull, Modern Physics: A Textbook for Engineers, John Wiley and Sons, New York, 1956.

Aldert van der Ziel, Solid State Physical Electronics, Prentice-Hall, New Jersey, 1957.

V. K. Zworykin and G. A. Morton, *Television: The Electronics of Image Transmission in Color and Monochrome*, 2nd edition, John Wiley and Sons, New York, 1954.

PROBLEMS

2.1. Give a brief qualitative description of the mechanism of electrical conduction in a metal.

2.2. What basic processes govern electrical conduction in an intrinsic semiconductor like silicon or germanium?

2.3. What is the effect of adding a small amount of a donor impurity like arsenic to an intrinsic semiconductor? Of an acceptor impurity like indium or gallium?

2.4. What physical processes influence the equilibrium value of the potential barrier at a p-n junction with no external fields applied?

2.5. Make a qualitative sketch of charge distribution, electric field and potential variations in the vicinity of a p-n junction.

2.6. (a) Sketch the curve of current versus applied voltage for a p-n junction and give a qualitative physical explanation of the form of the curve.

(b) Sketch another curve of current versus voltage for the same p-n junction at a higher equilibrium temperature.

2.7. (a) How much d-c current flows in a short-circuited semiconductor junction diode when the entire circuit is in thermal equilibrium at room temperature? Explain.

(b) With reverse voltage applied, a p-n junction diode conducts a small, nearly constant reverse current. Explain.

(c) With forward voltage applied, a p-n junction diode conducts a relatively large forward current which increases exponentially with applied voltage. Explain.

2.8. Sketch the general form of the current versus voltage curve for a cold-cathode gas diode.

2.9. Describe the operation of a hot-cathode gas-filled diode.

2.10. For a vacuum diode, sketch curves of current versus anode voltage for various values of cathode temperature.

2.11. For a vacuum diode, sketch curves of current versus temperature for various values of anode voltage.

2.12. Explain the effects of space-charge limitation and temperature limitation of current in a vacuum diode.

2.13. State the basic principles involved in establishing the plausibility of the generalized three-halves power law for vacuum diodes.

2.14. For a parallel-plane vacuum diode with cathode grounded and anode at $+E_b$, sketch the cold-cathode potential distribution, electric field, and charge distribution. Assume the anode and cathode each have an area A and are separated a distance d (neglect fringing).

2.15. Modify the curves of Problem 2.14 to include the effects of a sheet of charge midway between cathode and anode.

$$i_b = I_0 \epsilon^{-e_b/E}$$
$$\text{for } e_b < 0$$
$$\text{where } I_0 = k\phi_0^{3/2}$$

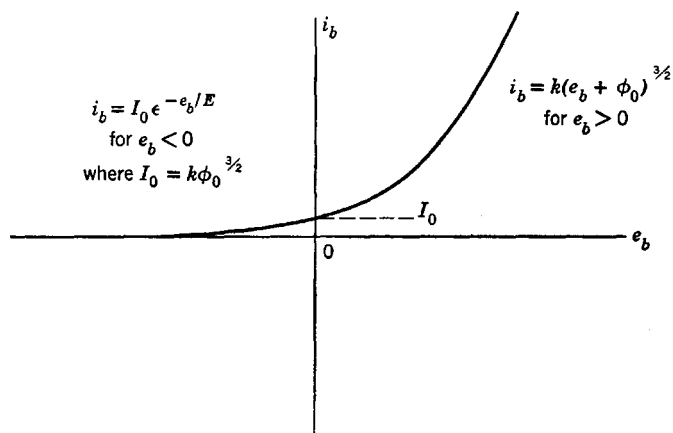$$i_b = k(e_b + \phi_0)^{3/2}$$
$$\text{for } e_b > 0$$

Fig. P2.1

2.16. Measurements on the plate-to-cathode circuit of a vacuum diode yield results that are closely approximated by the curve in Fig. P2.1.

(a) What are possible reasons for $I_0 \neq 0$?

(*b*) What is the significance of the product $e_b i_b$ in the first quadrant?

(*c*) What is the significance of the product $-e_b i_b$ in the second quadrant?

2.17. The voltage-current relation for a particular diode is given by

$$i = \frac{e^{3/2}}{1000} \text{ amperes}$$

In the vicinity of the operating point (E_0, I_0) this relation may be expanded in the power series

$$i = c_0 E_0 + c_1 (e - E_0) + c_2 (e - E_0)^2 + \cdots c_n (e - E_0)^n$$

Evaluate the coefficients c_0, c_1, and c_2 at the operating point $I_0 = 1$ ampere.

2.18. (*a*) Show that for a three-halves power law curve ($i_b = K e_b^{3/2}$), which closely approximates the i_b versus e_b curve of a high-vacuum diode, the slope of the curve at any point is just three-halves times the slope of a chord drawn from the origin to that point.

(*b*) If $i_b = 10$ ma when $e_b = 100$ v, what is the incremental resistance de_b/di_b of the diode at $i_b = 10$ ma?

(*c*) Show that the incremental resistance along a three-halves power i_b versus e_b curve varies inversely as the cube root of the current.

2.19. Two parallel plates, spaced 1 cm apart, are connected in series with a current-indicating device and a battery of voltage $V = 200$ volts, as shown in Fig. P2.2. The area of each plate is A cm². At $t = 0$, an electron is introduced at a small hole in the left-hand plate with zero initial velocity. Plot, as a function of time, the current which flows through the indicating device. Show that the energy *acquired* by the electron in its flight between plates is equal to the energy *delivered by* the battery.

The polarity of the battery is reversed so that the left-hand plate is made
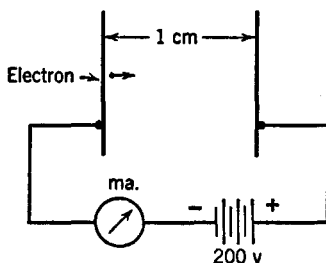

Fig. P2.2

positive. At $t = 0$ an electron is injected through the hole with an x-directed initial velocity corresponding to an energy of 200 electron volts. Plot the current-time relation. Show that the energy *lost* by the electron is equal to the energy *delivered to* the battery.

Field fringing effects at the plate edges and in the vicinity of the hole may be neglected.

2.20. (*a*) Assuming that the cathode of a high-vacuum diode is a limitless source of zero-velocity electrons, what happens to the electric field at the cathode, to the electron velocity at a given point, and to the space-charge density at a given point when the plate-to-cathode voltage is doubled?

(b) Sketch log i_b vs. log e_b for a high-vacuum diode. What portion or portions of the curve are straight, and why?

(c) Repeat part (b) with log e_b replaced by e_b.

2.21. An idealized parallel-plane diode consists of a pair of semi-infinite plates spaced 8 mm apart. With the plate 100 volts positive with respect to the cathode, the current flow is space-charge limited. Under this condition, and neglecting initial velocities and contact potential, find the velocity attained by an electron which has reached a point 1 mm distant from the cathode.

2.22. A particular germanium junction diode conducts 0.1 ma of current in the reverse direction when a few volts of reverse voltage are applied. The incremental resistance of the diode, measured at the origin of the i vs. v diode curve, is found to be 200 ohms. What are the values of the constants I_0 and V_0 in the theoretical junction law $i = I_0(\epsilon^{v/V_0} - 1)$?
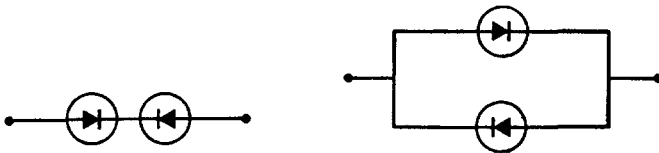

Fig. P2.3

2.23. Sketch the current versus voltage characteristic for two identical junction diodes, each characteristic given by $i = I_0(\epsilon^{v/V_0} - 1)$, connected in (a) series and (b) parallel as in Fig. P2.3.

2.24. The voltage-ampere relationship of a germanium diode may be described by the following equation:

$$i = I_0(\epsilon^{qv/kT} - 1)$$

At room temperature $kT/q = 0.025$ volt. Typical values for the constant I_0 and the reverse voltage at which avalanche breakdown takes place are 50×10^{-6} amperes and -80 volts, respectively. The maximum average power which can be dissipated by such a device without internal damage due to heating is typically 0.1 watt.

(a) Determine the maximum average forward current rating of the device.

(b) Determine the voltage v at which the forward current rating is reached.

Resistive Diode Circuits

3.1   Introduction

Resistive diode circuits include diodes, linear resistances, and sources, but no energy-storage elements. Although actual devices include small capacitances and inductances (for example the capacitance between electrodes in a vacuum diode), we shall represent diodes as nonlinear *resistances* in this chapter. By setting aside the consideration of the small energy-storage elements for the time being, we are in effect confining our attention to situations in which the currents and voltages are either in static equilibrium or they are varying so slowly that their time derivatives can be neglected. Circuit equilibrium equations are therefore algebraic rather then differential.

Since interelectrode and other stray capacitances, as well as wiring inductances, are very small in actual circuits, the resistive approximation fits many practical situations quite well. Both the circuit functions and the methods of analysis examined here will apply to circuits involving electronic devices other than diodes and to situations in which energy-storage elements must be included to describe circuit behavior satisfactorily. In fact, resistive circuit analysis forms the base upon which to build the analysis of circuits containing energy-storage elements.

3.2 Graphical Analysis of Linear Resistance Networks

Graphical analysis is seldom used for linear resistive circuits because quantitative results are more readily obtained by algebraic methods. However, a graphical solution may be the most convenient when the terminal relations for a device are available in graphical form and not
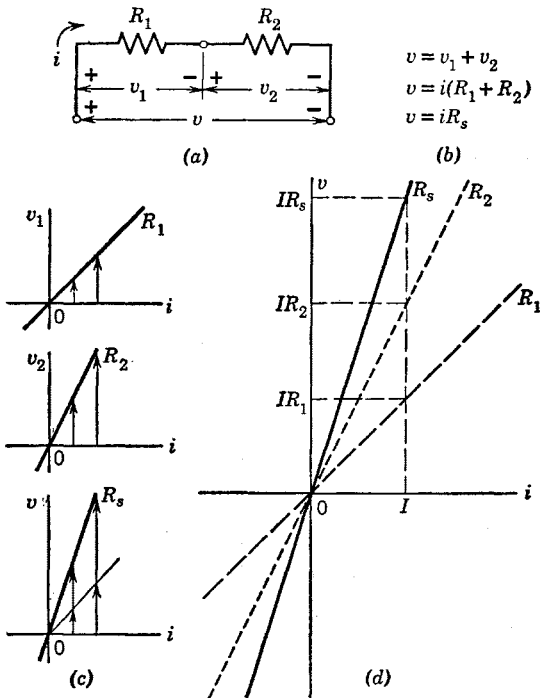

$$v = v_1 + v_2$$
$$v = i(R_1 + R_2)$$
$$v = iR_s$$

(a)   (b)

(c)   (d)

Fig. 3.1. Resistances in series.

readily expressible as a simple function. If nonlinear curves are approximated to permit an analytic solution, a qualitative graphical solution may be helpful in making the approximations and interpreting the results. A graphical sketch is an aid to understanding the operation of a nonlinear device. A few graphical solutions of linear circuits will serve to introduce the techniques that also apply to nonlinear circuits.

Figure 3.1 shows the graphical method of combining resistances in series. The circuit is shown in (a), and the pertinent equations are given in (b). The three sketches in (c) represent resistances R_1 and R_2

as well as the sum R_s. Voltage drops across R_1 and R_2 for any particular value of current are added to obtain the voltage drop across R_s for that current, as shown in (d). The same procedure can be extended to the summation of any number of resistances in series.
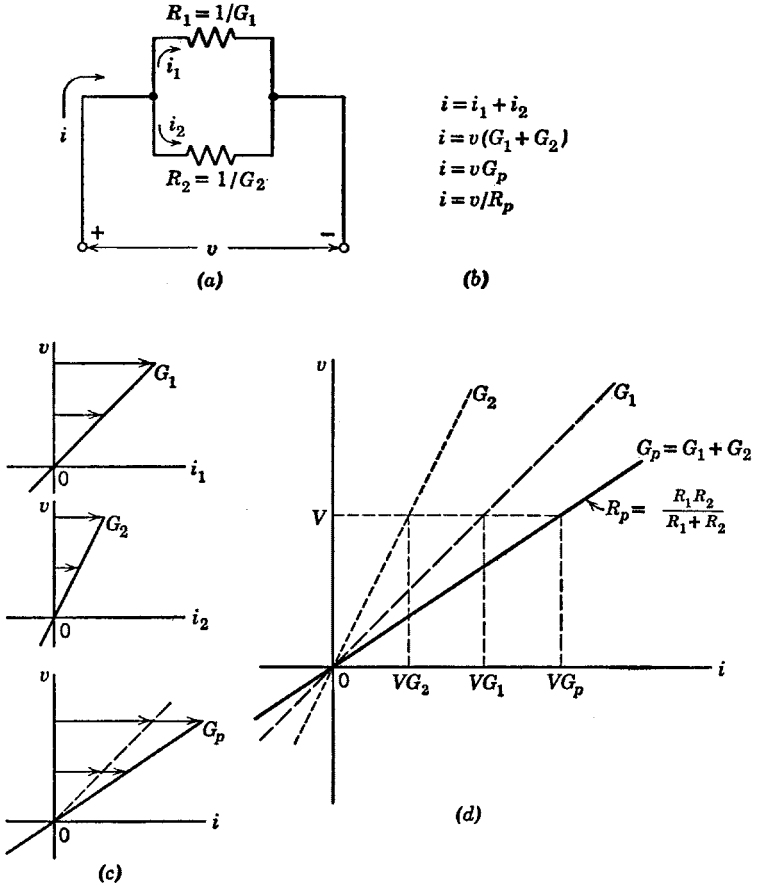

$$i = i_1 + i_2$$
$$i = v(G_1 + G_2)$$
$$i = vG_p$$
$$i = v/R_p$$

(a)   (b)

$$G_p = G_1 + G_2$$

$$R_p = \frac{R_1 R_2}{R_1 + R_2}$$

(c)   (d)

Fig. 3.2. Resistances in parallel.

Figure 3.2 presents the dual problem of combining resistances in parallel. Here the voltage across the resistances is common, and the equivalent resistance of the parallel combination is obtained graphically by summing currents. Expressing the resistances R_1 and R_2 as the conductances $G_1 = 1/R_1$ and $G_2 = 1/R_2$ puts the equations in the same form as those for the series circuit. This is a consequence of the duality

principle. In fact, by applying this principle to the series-resistance problem given in Fig. 3.1, the solution for the parallel-resistance problem of Fig. 3.2 is immediately known. Using these methods, the total
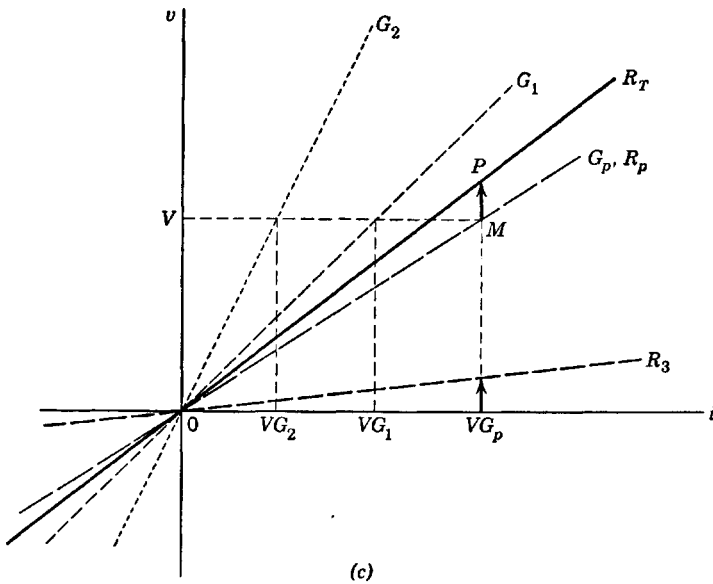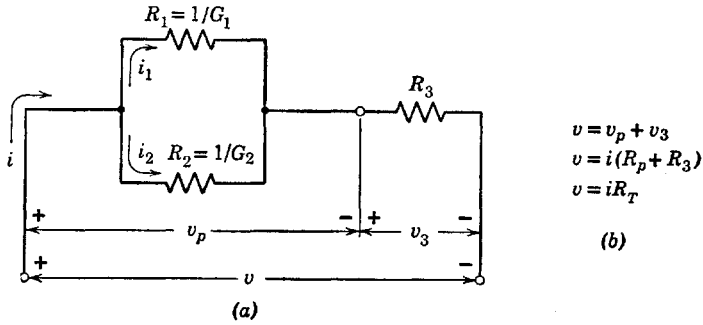

$$v = v_p + v_3$$
$$v = i(R_p + R_3)$$
$$v = iR_T$$

(b)

(a)

(c)

Fig. 3.3. Series-parallel combination.

resistance R_T for a series-parallel combination of resistances is determined graphically in Fig. 3.3.

The foregoing examples illustrate several important points. A straight line through the origin of a voltage versus current plane represents a

linear resistance R or a conductance $G = 1/R$. Thus, data given either as resistance or conductance may be plotted equally well, and no conversion from one to the other is necessary. Any number of resistances or conductances can be combined in series and parallel, and the resultant
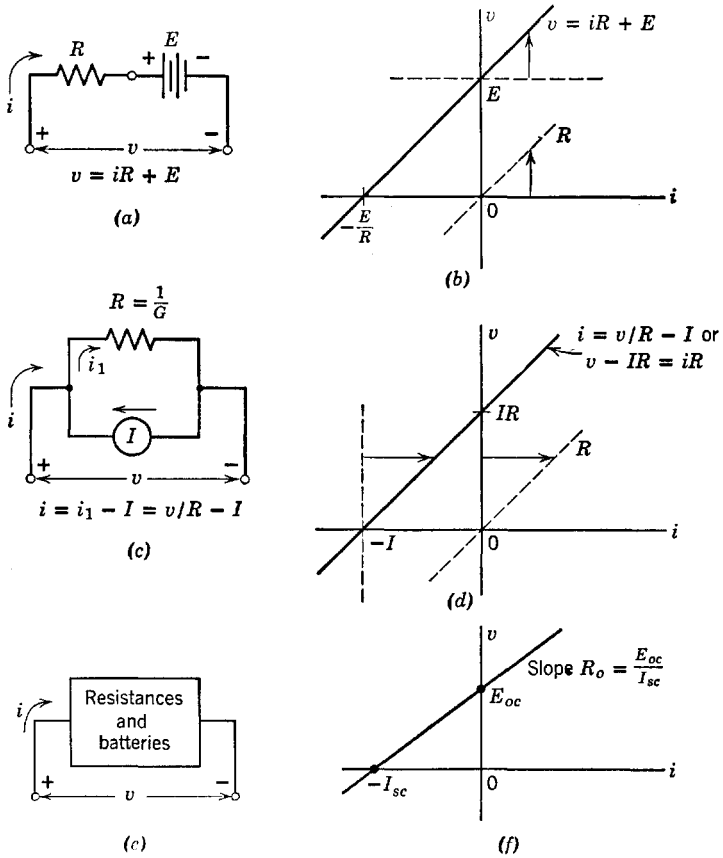

Fig. 3.4. Resistance and voltage or current source. Circuits (a) and (c) are equivalent if the resistances are equal and if $I = E/R$.

plot will still be a line through the origin with a slope equal to the total resistance.

The graphical effect of including a source as part of a resistive circuit amounts to a translation of coordinates or, alternatively, a translation of the resistive line so that it does not pass through the origin. Since an ideal voltage source is independent of current, it appears as a line of zero slope, thus placing in evidence the fact that it has no internal resistance.

Similarly, a current source appears as a line of constant current for all voltages.

Consider the series combination of a resistance and battery shown in Fig. 3.4(a). The resistance R might, of course, represent the composite effect of any number of resistances, in accordance with the results above. In the graphical representation of this circuit shown in Fig. 3.4(b), the line has a slope R, a voltage intercept E, and a current intercept $-E/R$.

Referring to Fig. 3.4(c) and (d), we see that a current source in parallel with a resistance has the same graphical appearance as a voltage source in series with a resistance. If the source values are related by $E = IR$, and a common value of resistance is chosen, the circuits (a) and (c) are equivalent, and the plots in (b) and (d) are identical. Thus, we have pictorial evidence of the equivalence of voltage source with series resistance and current source with parallel resistance. Conversely, for an arbitrary straight line in a voltage versus current plane, the slope and intercepts determine the values for the parameters in the voltage-source or current-source representation of the line. Thus, the circuits in Fig. 3.4(a) and (c) are general circuit models for any straight line in a v vs. i plane. If the intercept is zero, the circuit model contains no source, and the line represents only a resistance.

Note specifically that the voltage intercept corresponds to zero current and is therefore the open-circuit voltage, whereas the current intercept corresponds to zero voltage and hence is short-circuit current. Thus we see that *resistance between a pair of terminals* in a network is the *ratio of the open-circuit voltage to the short-circuit current (with proper sign)*. Furthermore, it is evident from these facts that the behavior of an arbitrary resistive network is completely described at a pair of terminals by this resistance and either the voltage or the current source.

Note that the v vs. i plot of a positive resistance will have a positive slope only if current and voltage reference directions are appropriately chosen. Negative slope does not necessarily imply negative resistance.*

As a more general example, consider the circuit shown in Fig. 3.5. To relate the graphical interpretation to the algebraic relations, the lines in the v vs. i plane are numbered to correspond to the algebraic equations. The solution may be summarized as follows:

(a) Plot lines (1) and (2) with appropriate slopes and intercepts.

(b) Choose convenient points (like A and B) and sum currents of (1) and (2) to obtain line (3).

* Negative resistance is a useful concept in electronic circuit analysis. In contrast to positive resistance, which dissipates power, negative resistance is a convenient way to represent certain types of power sources. Further discussion is deferred until Chapter 9.
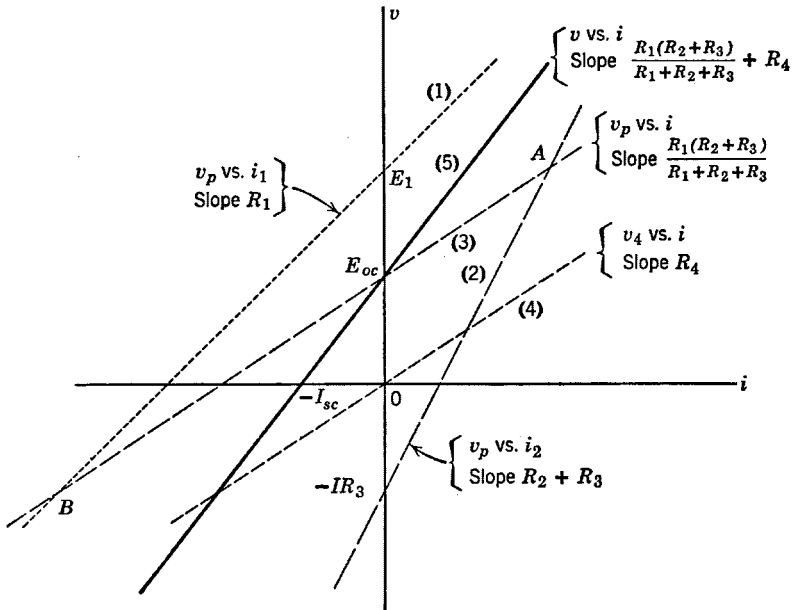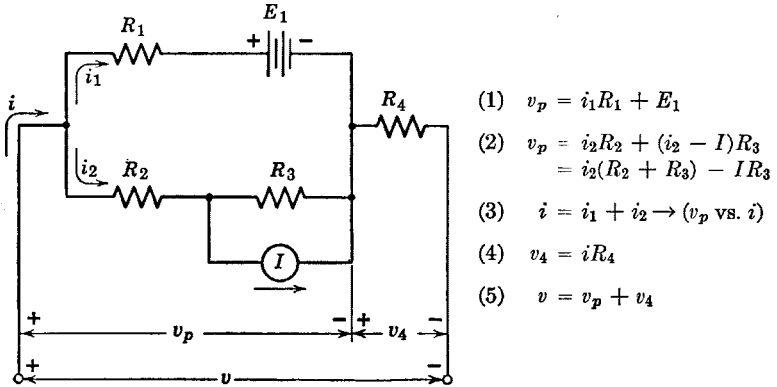
(1) $v_p = i_1 R_1 + E_1$

(2) $v_p = i_2 R_2 + (i_2 - I)R_3$
$\quad\quad = i_2(R_2 + R_3) - IR_3$

(3) $i = i_1 + i_2 \rightarrow (v_p \text{ vs. } i)$

(4) $v_4 = iR_4$

(5) $v = v_p + v_4$

Fig. 3.5. Multiple resistances and sources.

(c) Plot line (4).

(d) Choose convenient points (like E_{oc} and $-I_{sc}$) and sum voltages of (3) and (4) to obtain line (5).

Comparison of Figs. 3.4(b) and (d) and 3.5 shows that the circuit of Fig. 3.5 is equivalent to a voltage source E_{oc} in series with a resistance R_o, or a current source I_{sc} in parallel with R_o, where R_o equals E_{oc}/I_{sc}.

The generalization of this concept is illustrated in Fig. 3.4(e) and (f). Regardless of circuit complexity, we can always reduce any number of resistances, batteries, and current sources to the same simple form of voltage versus current relation at a pair of terminals. Figures 3.4(b), (d), and (f) are identical in form, and hence they illustrate the Thevenin equivalent circuit concept. The graphical approach to resistive circuits has enabled us to deduce this result without stating or attempting to prove Thevenin's theorem. Furthermore, the graphical approach emphasizes the fact that this simple terminal representation for any resistive circuit is a direct consequence of linearity. The addition of currents or voltages specified by nonlinear relations does not lead to such generality.
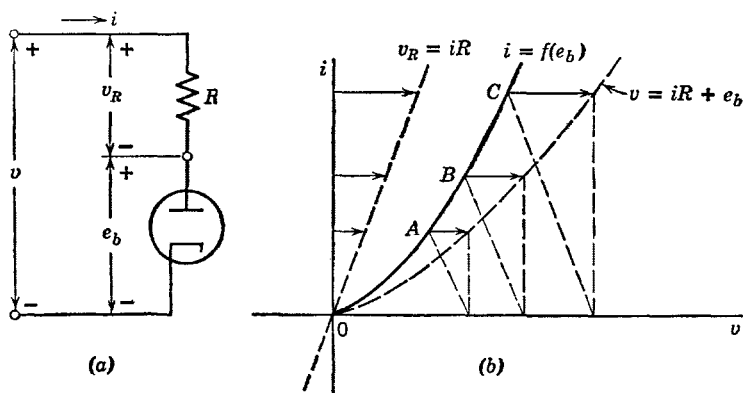
Fig. 3.6. Diode and series resistance.

3.3 Graphical Solution of Nonlinear Circuits

Having solved the general linear resistive circuit problem, let us analyze some elementary nonlinear circuits. A diode in series with resistance R, as shown in Fig. 3.6(a), is a nonlinear circuit analogous to the combination of linear resistances in series. The construction in Fig. 3.6(b) illustrates the summation of voltage drops for specific values of current. A number of points must be plotted to obtain the new curve, since the result is not a straight line. This construction is only useful when a particular resistance R is to be associated with a particular diode; another curve must be plotted when either the diode or the resistor is changed. Since such plots may require a fair number of points, it is desirable to seek a simpler procedure. Restricting a problem to finding a specific solution, i.e., the equilibrium current and voltage

at one operating point, removes the need for replotting. A succession of such operating points would form a functional relation between two variables (like i and v at a pair of terminals).
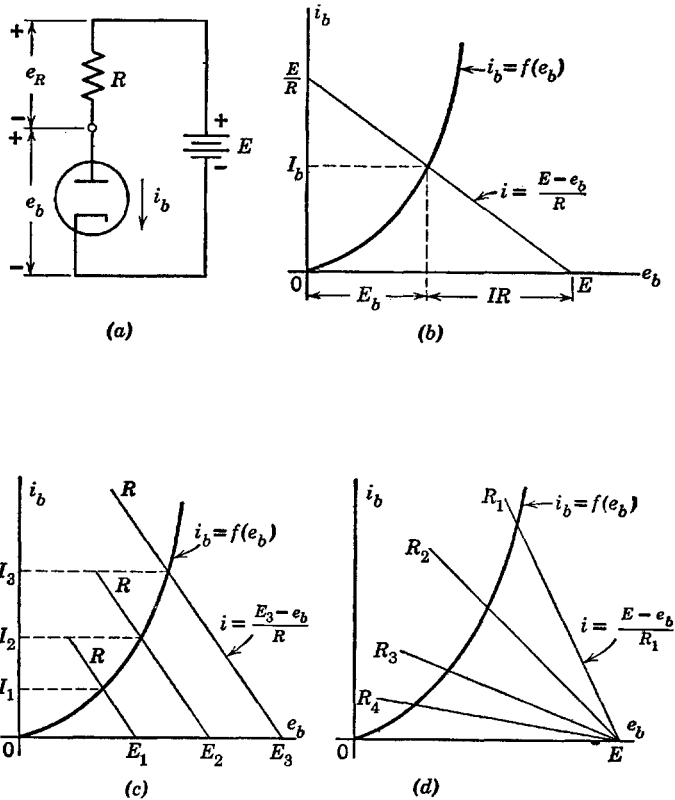

Fig. 3.7. Diode operating point.

The diode circuit of Fig. 3.7 (a) has a specific voltage E applied, supplying power to the circuit. The source E and the resistance R, often called the *load resistance*, provide a linear constraint that aids in solving the problem. This constraint or "load line," as plotted in Fig. 3.7 (b), represents the relation $i = (E - e_b)/R$ and amounts to plotting the resistive line as a voltage drop subtracted from supply voltage E. The effects of varying the supply voltage for a fixed value of load resistance and varying the load resistance for a fixed value of supply voltage are illustrated by Fig. 3.7 (c) and (d), respectively.

Having solved the simplest one-diode problem, the previous results

obtained for linear resistances and sources tell us that we have also solved the most general one-diode problem. As indicated by Fig. 3.8, the load line may actually be representative of an arbitrary network of resistances and batteries. Graphical or algebraic reduction of the network to a single resistance and source permits solution of the operating-point problem. Once the current and voltage are specified at the diode terminals, any other current or voltage within the network can be determined.
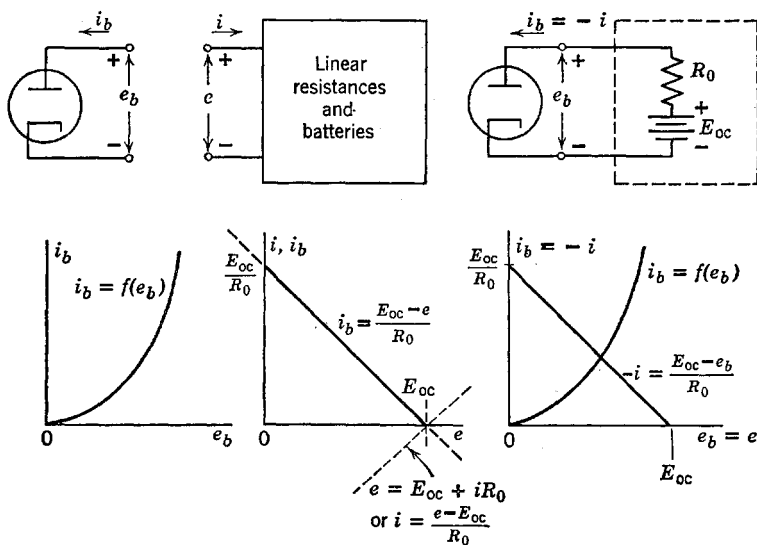
Fig. 3.8. General one-diode circuit.

Only a few relatively simple circuits with two diodes can be analyzed graphically. Obviously, any number of identical diodes in series or in parallel poses no greater problem than a single diode, since the voltage scale or the current scale of the diode curve can be multiplied by the number of diodes.

As shown in Fig. 3.9(a), two different diodes in series present a more difficult problem, since the composite curve for the two (b) must be plotted point-by-point. The same circuit connected to a resistive load R and supply voltage E is shown in (c). The operating-point problem for this case could be solved by using the composite curve $i = f(v)$ from Fig. 3.9(b) and plotting the load line $i = (E - v)/R$, just as was done for a single diode.

Graphical solution of the general two-diode circuit requires a number

of construction lines to carry out a double cut-and-try process in order to establish a pair of operating points for the two diodes. The solution is seldom necessary, since most two-diode circuits use identical diodes and usually exhibit symmetry, which can be exploited to simplify the
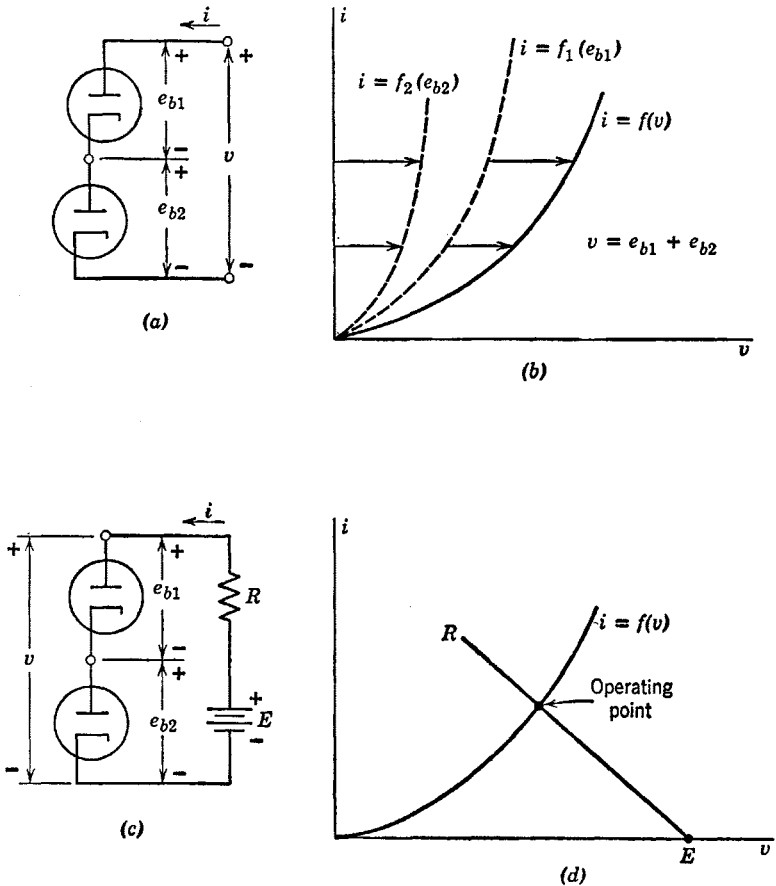

(a)

(b)

(c)

(d)

Fig. 3.9. Diodes in series.

analysis. Circuits containing more than two diodes (or any other type of nonlinear device) are so cumbersome to handle that simplifying approximations become essential.

The essence of the general two-diode problem is shown in Fig. 3.10. Note the simplicity with which any three-terminal linear circuit can be represented. The circuit shown in Fig. 3.10(b) is only one of many

possible models for a three-terminal network. With three terminals there are two separate conditions that can be specified, say i_1 and i_2. Each of these currents should be a linear function of the voltages e_1 and e_2. The model and the equations given in Fig. 3.10 satisfy these conditions.

(a)

(b)
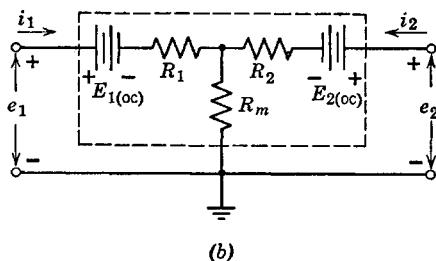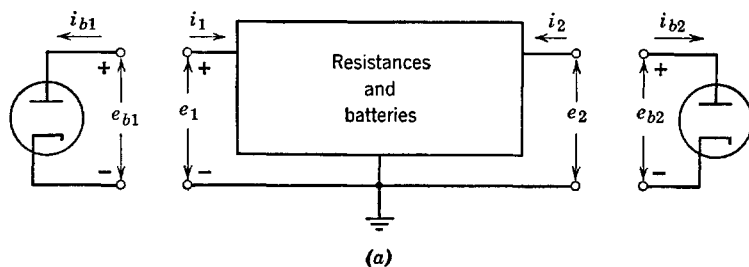
$$e_1 - E_{1(OC)} = i_1(R_1 + R_m) + i_2 R_m$$

$$e_2 - E_{2(OC)} = i_1 R_m + i_2(R_2 + R_m)$$

(b)

Fig. 3.10. General two-diode circuit.

If values of i_1 and i_2 are assumed, e_1 and e_2 are determined; but with the diodes connected, $i_1 = -i_{b1}$ and $i_2 = -i_{b2}$, and these values determine e_{b1} and e_{b2}. There is little likelihood of finding the right values of the currents to make $e_1 = e_{b1}$ and $e_2 = e_{b2}$ without repeatedly assuming values for the currents. The cut-and-try process can be organized by means of graphical constructions.

Continued pursuit of such procedures is unwarranted, particularly since one need only contemplate a circuit with three or more diodes to stimulate a search for a simpler method. We can conclude that graphical analysis is most useful for the solution of resistive circuits containing

a single nonlinear element and an arbitrary number of resistances and batteries.

3.4   Algebraic Solution of Diode Circuit Problems

An algebraic expression for the i_b vs. e_b relationship of a diode can be obtained either by deriving it from physical laws as outlined in Chapter 2, or by approximating either experimental data or average characteristics with an appropriate function.

Algebraic determination of an operating point may be carried out by analogy to the graphical solution shown in Fig. 3.8(a) and (b). If the diode characteristic is specified by an equation such as $i = Ke^{3/2}$, this expression for current may be equated to the expression $i = (E - e_b)/R$ given by the load line. The solution of this equation yields E_b, the value of e_b at the operating point.

For exponential expressions, such as those describing semiconductor diodes, the solution of problems involving two or more diodes may result in transcendental equations that require graphical or cut-and-try solutions. However, in many cases the algebraic solution is direct and less time-consuming than a graphical construction.

The problem becomes more difficult if the circuit includes energy-storage elements since differential equations are then necessary to express circuit equilibrium conditions. Except for a few special cases, nonlinear differential equations cannot be solved by analytic methods.

3.5   The Ideal Rectifier as a Diode Model

The desirability of applying the powerful analytical tools of linear circuit theory to nonlinear problems provides the motivation for approximating nonlinear curves by straight-line segments; that is, by piecewise linear approximations. One additional type of circuit element, the ideal rectifier, enables us to represent a nonlinear curve in this manner with any desired degree of accuracy. Surprisingly enough, we shall see that a two-segment approximation often provides results of acceptable accuracy. By making piecewise linear approximations to nonlinear curves, a nonlinear problem is reduced to two or more connected linear problems for which solutions can be obtained by means of linear circuit theory.

The ideal rectifier is the basic piecewise linear circuit element. It is a resistive element in the sense that there is no energy-storage associated with it. As shown in Fig. 3.11(a), this element develops no voltage drop

for current in the forward direction (zero resistance) and permits no current for reverse voltages (infinite resistance). In contrast with the linear circuit elements R, L, C, E, and I, the ideal rectifier has no numerical value associated with it, only a preferred direction of con-
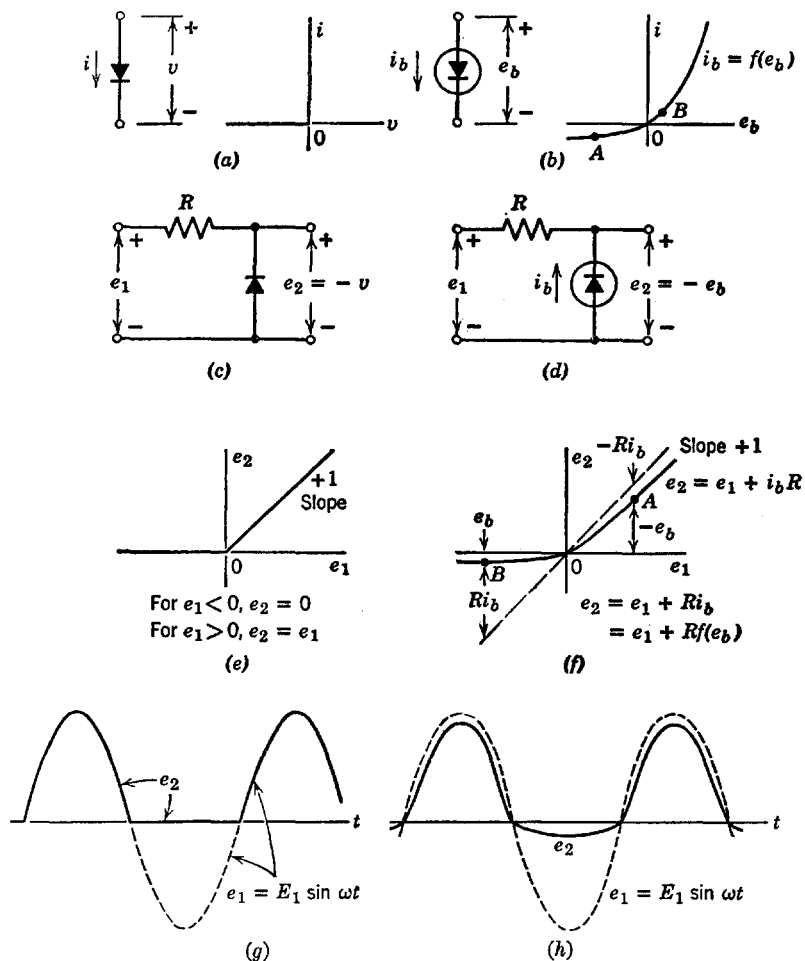

Fig. 3.11. Comparison of ideal rectifier and typical diode in a simple clipping circuit.

duction. Note also that the ideal rectifier is lossless, since the product of v and i is zero for all possible values of v and i.

Since the ideal rectifier is an idealization of a physical diode, it is interesting to see how well it approximates such a diode in a typical

circuit. An arbitrary diode curve $i_b = f(e_b)$ is shown in Fig. 3.11(b). Two simple rectifier circuits are shown, using in (c) an ideal rectifier and in (d) an actual diode. The circuits reproduce the positive portions of the input voltage waveform but do not pass the negative portions to the output so that the output is unidirectional or "rectified." This circuit is also the basic prototype of amplitude wave-shaping circuits. It is called a diode clipper, since it removes or clips a portion of the input waveform. The plots of e_2 vs. e_1 shown in Fig. 3.11(e) and (f), called transfer curves, present graphically the instantaneous relation between output and input voltages. The slope at any point $(\Delta e_2/\Delta e_1)$ is called the transfer ratio or gain, in this case a voltage ratio that cannot exceed unity. The sketches exaggerate the difference between the transfer curve for the circuit with the ideal rectifier and that of the circuit with an actual diode. In the high-conduction region $(e_1 < 0)$, if the voltage drop across R is very large compared with the drop across the diode, the percentage error is small. For e_1 greater than 0, the error is again small if $e_b \gg Ri_b$. These are contradictory conditions on R, but an intermediate value of R can usually be found to satisfy the conditions reasonably well over a fair range of values for e_1 and e_2. The close correspondence between these transfer curves is quite surprising in view of the fact that the diode curve and the ideal rectifier curve differ considerably. The rectifying or wave-shaping action of the two circuits, Fig. 3.11(g) and (h), also shows great similarity.
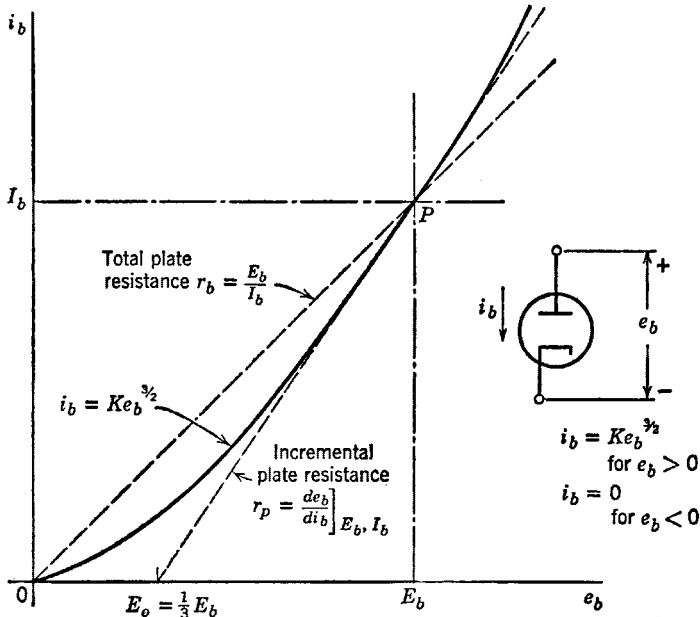
In some circuits the ideal rectifier may be an inadequate approximation to the actual diode. The remedy adds very little complexity in return for a considerably better approximation as discussed in the following articles.

3.6   Piecewise-Linear Models for Vacuum Diodes

Curves for vacuum diodes closely follow the three-halves power law (see Fig. 3.12), except at low voltages, where contact potential difference and the initial velocity of electrons become important. The geometry of this curve illustrates the considerations involved in approximating a nonlinear curve. Suppose the range of voltage over which the diode is to be operated includes negative values and extends to point P for positive values. For a particular point, such as P, let the voltage and current be designated as E_b and I_b, respectively.

The line OP through the origin determines a quantity r_b, called the *total* plate resistance. A tangent drawn through P defines the quantity r_p, called the *incremental* plate resistance. As indicated in Fig. 3.12,

the value of r_p determined at any point on a three-halves power curve is $2r_b/3$. The tangent-line intercept E_0 occurs at the value $E_b/3$. Note that the curve lies between the r_b line and the r_p line for $0 < e_b < E_b$. The tangent line r_p always lies below the curve, whereas the chord r_b crosses the curve at point P.

Total plate resistance $r_b = \dfrac{E_b}{I_b}$

$i_b = K e_b^{3/2}$

Incremental plate resistance

$r_p = \dfrac{de_b}{di_b}\bigg]_{E_b, I_b}$

$i_b = K e_b^{3/2}$ for $e_b > 0$

$i_b = 0$ for $e_b < 0$

$E_o = \tfrac{1}{3} E_b$

$$(1) \quad r_b = \frac{E_b}{I_b} = \frac{1}{K E_b^{1/2}}$$

$$(2) \quad r_p = \frac{de_b}{di_b} = \frac{2}{3K}\left(\frac{i_b}{K}\right)^{-1/3} = \frac{2}{3K E_b^{1/2}} = \frac{2}{3} r_b$$

$$(3) \quad E_o = E_b - I_b r_p = \tfrac{1}{3} E_b$$

Fig. 3.12. Geometrical properties of three-halves power curve for vacuum diode.

The equation for the tangent line represents the first two terms of the Taylor series expansion of the diode curve about point P. Expanding $i_b = K e_b^{3/2}$ at point P yields a constant term, and a linear term. Expressed in terms of the coordinates of point P, this becomes $i_b = -I_b/2 + 3I_b e_b/2E_b$. This tangent line may be considered to be a linear approximation to the diode curve. The corresponding circuit model is either a battery of value $E_0 = E_b/3$ in series with a resistance $r_p = 2E_b/3I_b$ or a current source of magnitude $I_b/2$ in parallel with the

resistance. The tangent line approximates i_b reasonably well near P. In fact, it yields current values within $0.015I_b$ of the actual diode curve
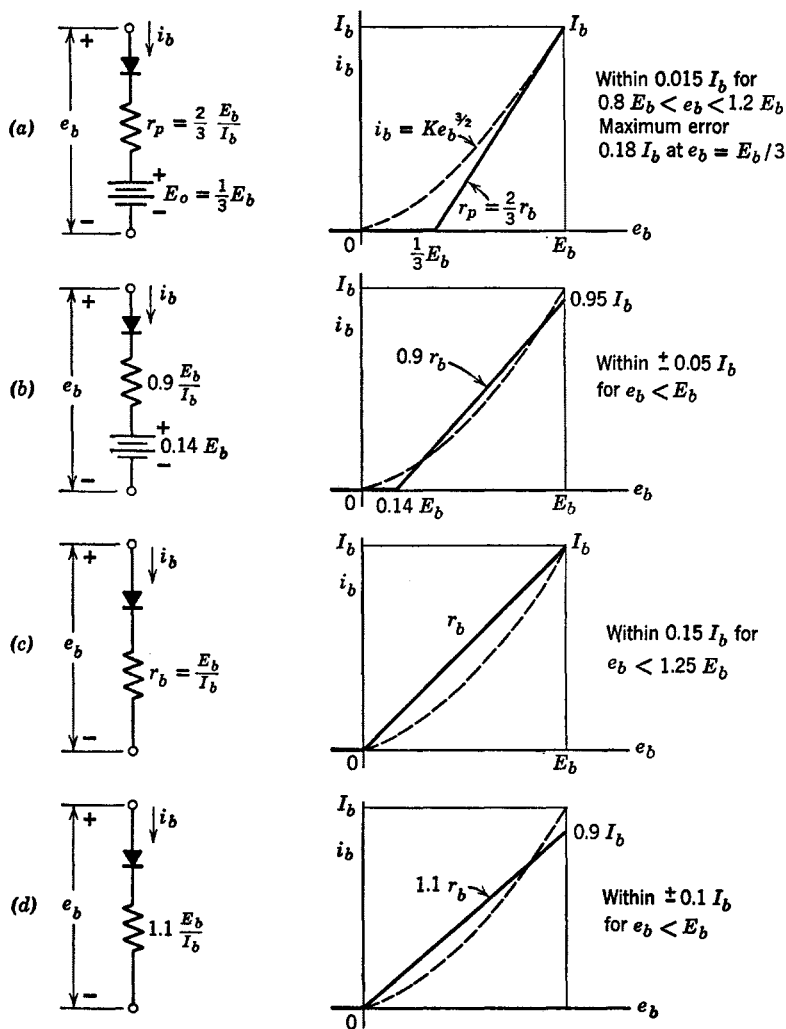

Fig. 3.13. Piecewise-linear models for vacuum diodes.

for $0.8E_b < e_b < 1.2E_b$. The addition of an ideal rectifier to this model yields the piecewise-linear model shown in Fig. 3.13(a).

A better fit over the specified range $e_b < E_b$ is obtained by changing circuit parameters as shown in Fig. 3.13(b). Obtaining such a straight-

line fit to a nonlinear curve can best be done by eye. Minimizing by
calculus the difference between a nonlinear curve and the equation of a

(a)

(b)
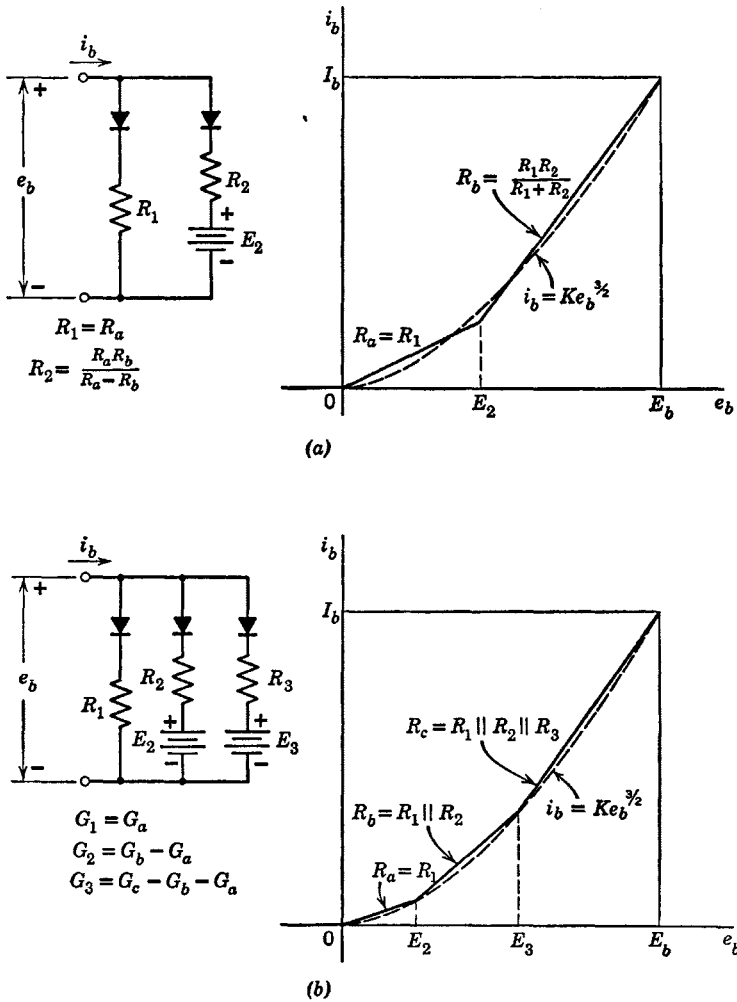
Fig. 3.14. Multiple-segment approximations.

straight line, with floating values for the intercept and slope, can be
time-consuming.

A simpler circuit model results from the use of r_b in conjunction with
an ideal rectifier. As shown in Fig. 3.13(c), the errors for this model are
somewhat larger than those of the previous two. Again, a better fit

may be obtained for the same circuit configuration by modifying the parameters to the values shown in Fig. 3.13(d).

The approximation to the vacuum diode obtained with any of these models is much better than that obtained with only an ideal rectifier. When one of these models is used to analyze the clipper circuit described in Article 3.5, the results obtained can hardly be distinguished from those obtained with the actual nonlinear curve. In fact, for almost any application of a vacuum diode, these models will yield adequately accurate results for the resistive-circuit problem. Nevertheless, if the need should arise, the concept can be extended to yield even better approximations. The circuit models and corresponding curves shown in Fig. 3.14 illustrate this point. Any number of segments can be produced by paralleling additional circuit branches made up of battery, resistance, and ideal rectifier in series. The relations between the resistive slopes of the piecewise-linear curve and the resistances or conductances used in the circuit model are indicated in the figure.

Each parallel branch in the circuit models shown in Fig. 3.14 includes a "back-biased" ideal rectifier. The applied voltage must exceed the reverse bias voltage in a branch in order for conduction to take place in that branch. Thus, as the applied voltage e_b is increased, successive branches of the circuit conduct.

The models used thus far approximate i vs. v curves that are concave upward; that is, conductance increases as applied voltage increases. However, some i_b vs. e_b curves have concave downward regions, for example, the more complete vacuum-diode curve shown in Fig. 3.15(a). Since the conductance decreases with increasing voltage in the saturation region, a new form of circuit is required to add a resistance in series when the voltage or current has increased to a given level. This effect can be achieved by using a "forward-biased" ideal rectifier. As indicated in Fig. 3.15(b), the rectifier must be so directed that the applied current eventually overcomes the current due to the forward bias. When the ideal rectifier is conducting, the voltage drop across its terminals is zero. When the applied current equals the bias current, the total current through the ideal rectifier is zero. This is the break point at which the rectifier becomes an open circuit, and current then passes through the resistance. An alternate form of this model is shown in Fig. 3.15(c). The complete vacuum-diode model is shown in (d).

3.7   Piecewise-Linear Models for Semiconductor Diodes

The semiconductor diode has a current versus voltage curve that differs in functional form from that of the vacuum diode. However,
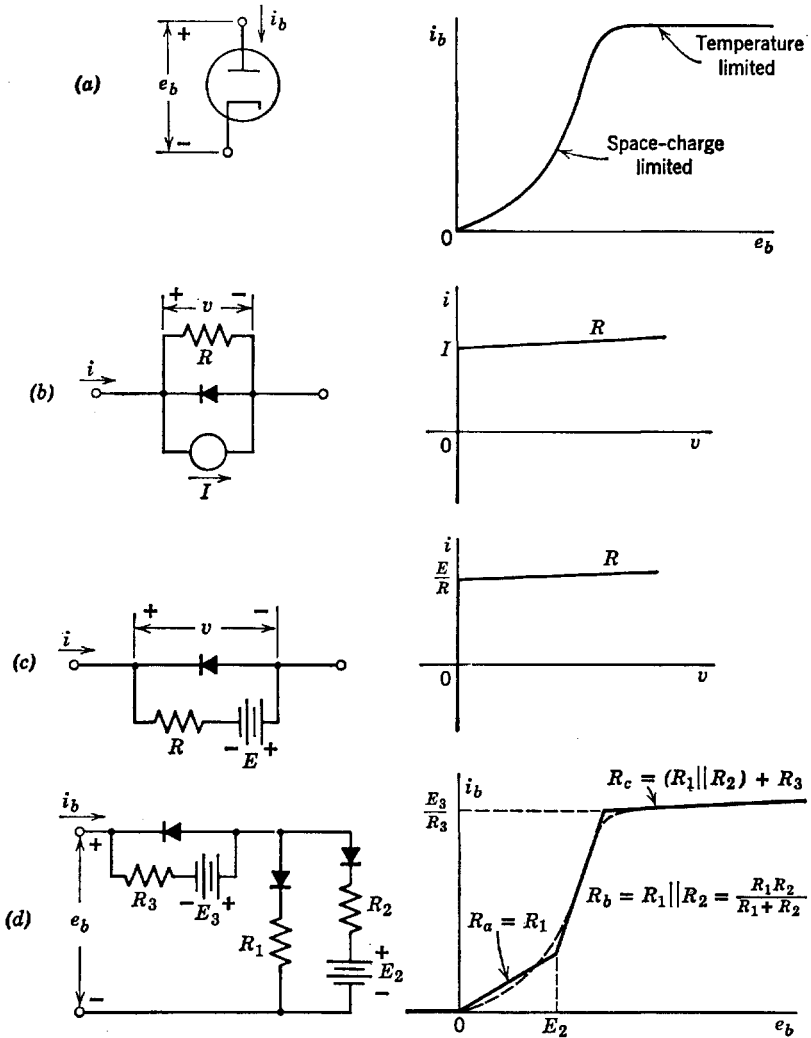
Fig. 3.15. Representing temperature saturation in a vacuum diode.

aside from the fact that reverse current may have to be considered, the problem of devising a circuit model is similar to that already solved for the vacuum diode.

The curve of current versus voltage for a semiconductor junction alone is exponential. The effective forward resistance may be as low as a few ohms and is ordinarily not more than a few hundred ohms, whereas the reverse resistance is likely to be hundreds of kilohms or even a few

megohms for reverse voltages up to the value required to produce reverse-voltage breakdown.

The sketch in Fig. 3.16(a) is a typical semiconductor-diode curve. The magnitude of the reverse current has been exaggerated in order to
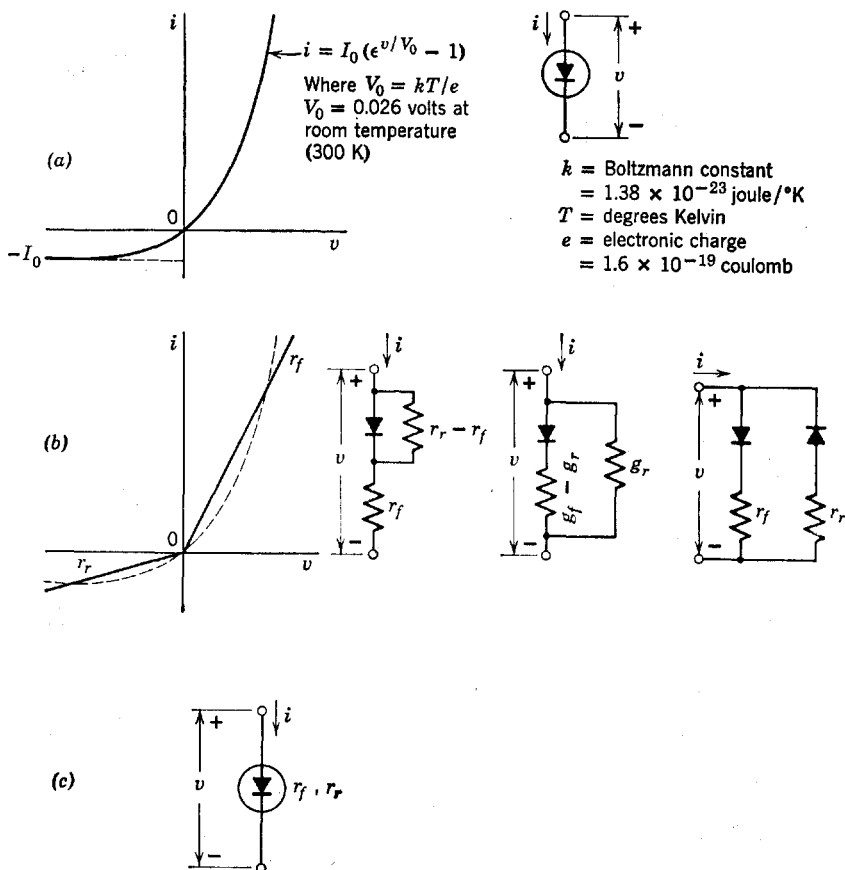
Fig. 3.16. Approximating semiconductor diodes with forward and reverse resistances.

clarify the drawings. An approximation to the curve based on two resistive lines is shown in Fig. 3.16(b). A number of alternative forms are possible for this circuit model. The models shown with a single ideal rectifier are equally convenient. Since reverse resistance r_r is much larger than forward resistance r_f, the two resistances (or conductances) in the model are very nearly those corresponding to the graphical slopes.

If the ratio of r_r to r_f is not sufficient to justify the simplifying approximation $r_r - r_f \approx r_r$ or $g_f - g_r \approx g_f$, the graphical slopes can be placed in evidence by introducing another ideal rectifier as shown. Since no sources are included in this two-rectifier model, the break point for
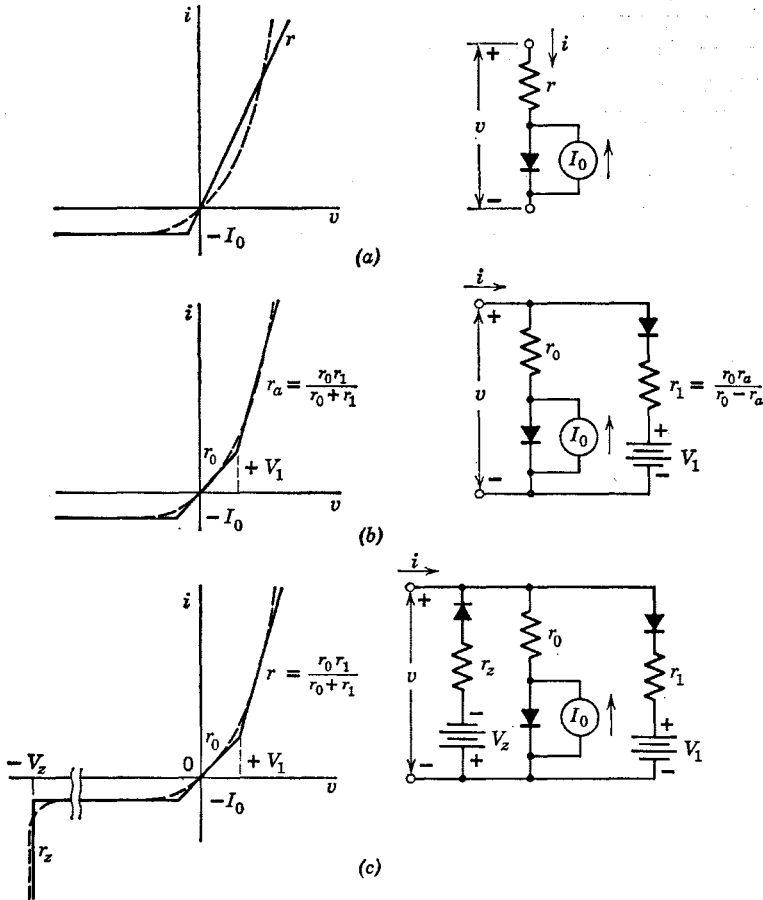
Fig. 3.17. Additional semiconductor diode models.

each of the diodes occurs at $i = 0$, $v = 0$ (one opening and the other closing).

The model with two ideal rectifiers is more complicated, but if it clarifies the resulting circuit problem, its use is justified. When circuit complexity increases, caution should be exercised in choosing models containing extraneous ideal rectifiers, as each rectifier may create additional circuit states.

It will usually be more convenient to designate a two-segment resistive approximation by the simple symbol shown in Fig. 3.16(c). Circuit equations for this approximation can be written as though the element were a resistance r with specific values r_r or r_f, depending upon the polarity of voltage v or current i.

If the reverse current of a semiconductor diode is critical to the behavior of a particular circuit, a better representation is obtained by the model shown in Fig. 3.17(a). For negative voltages, $v < (-I_0 r)$, the ideal rectifier is open, and the model represents the constant-current line $(-I_0)$. For $v > (-I_0 r)$, the diode closes, and the diode is represented by the small resistance r. An additional segment can be added to improve the approximation over a wider range of v and i by a parallel branch, as shown in Fig. 3.17(b). The effect of reverse-voltage breakdown can be included by a circuit that holds the voltage nearly constant at the appropriate value (say -25 volts) for a particular diode. A circuit model that suffices is shown in Fig. 3.17(c). There are no extraneous ideal rectifiers in this model, since each one corresponds to a different break point in the i vs. v curve. Nevertheless, it is apparent that this is an elaborate model to represent only the resistive behavior of a single circuit element, the semiconductor diode.

As familiarity with the subject develops, it is seldom necessary to draw the complete circuit model. For any particular value of v or i, we are concerned only with the relation between v and i, and in piecewise linear models this is always a straight line. Thus, if we keep track of the segment of the curve on which the diode is operating at any given time, the maximum complexity of the relation is $i = av + b$. Correspondingly, the circuit model for any one segment reduces to (at most) a resistance and a source (current or voltage). A series of single-segment models is shown in Fig. 3.18 for the same circuit model shown in Fig. 3.17(c). The slopes and intercepts for the linear segments correspond to those on the complete i vs. v curve, but do not necessarily correspond to the sources or resistances in the complete model. Note, for example, V_1 and V_1'.

3.8   Piecewise-Linear Models for Gas-Filled Diodes

The curve of i_b vs. e_b for a gas-filled thermionic diode is shown qualitatively in Fig. 3.19(a), together with the usual symbol for a gas-filled tube. The dot within the circle distinguishes the gas-filled tube from the corresponding type of vacuum tube (diode, in this case). Conduction in the tube is very small until e_b reaches the value E_m. At this
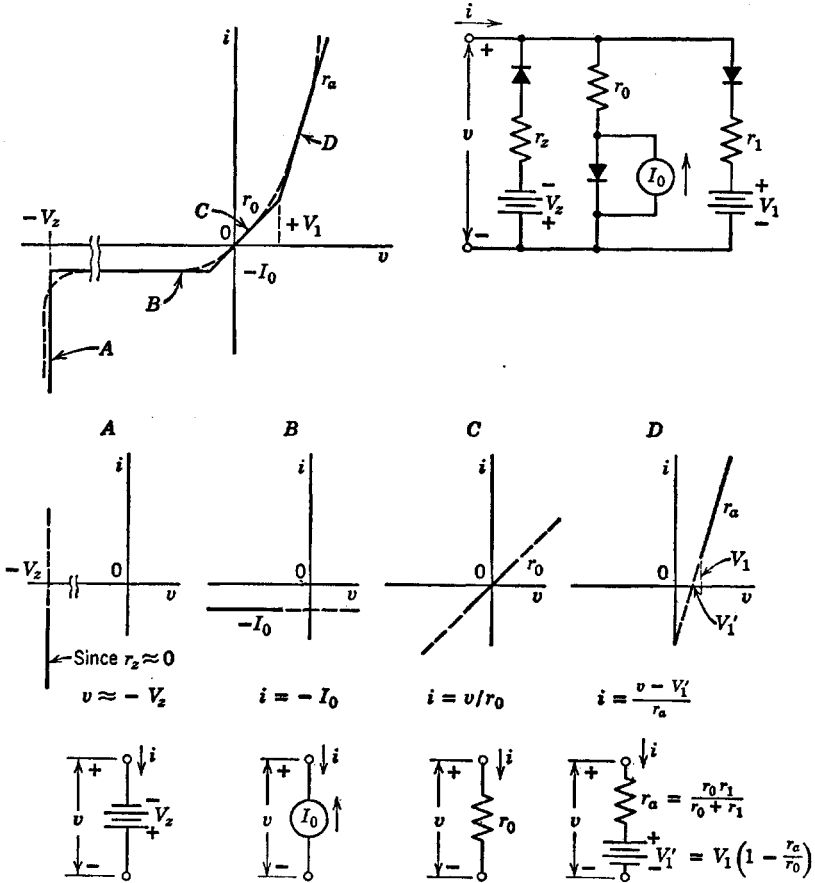
Fig. 3.18. Linear representation of segments of piecewise-linear curve.

point the molecules of the gas become completely ionized and form a plasma. The plasma consists of a nearly field-free region in which the charges of the electrons are essentially neutralized by positive-ion charge. Conduction is sustained for a wide range of currents with an almost constant voltage drop E across the tube. The voltage E may be substantially less than E_m. In the conduction state the curve is fairly well approximated by the simple model shown in Fig. 3.19(b). The onset of conduction may be represented by a resistance r in series with an ideal rectifier. This model is good up to voltage E_m, as indicated by the line of slope r in Fig. 3.19(c). When the two simple models are combined, as indicated in Fig. 3.19(c), the curve is as shown by

the solid line. It is therefore preferable to use each of these parallel branches as a separate model. The resistance applies prior to ionization, and the voltage source E applies after ionization.

The transition from one model to the other is left indeterminate by this use of two separate models. It is apparent that a piecewise-linear
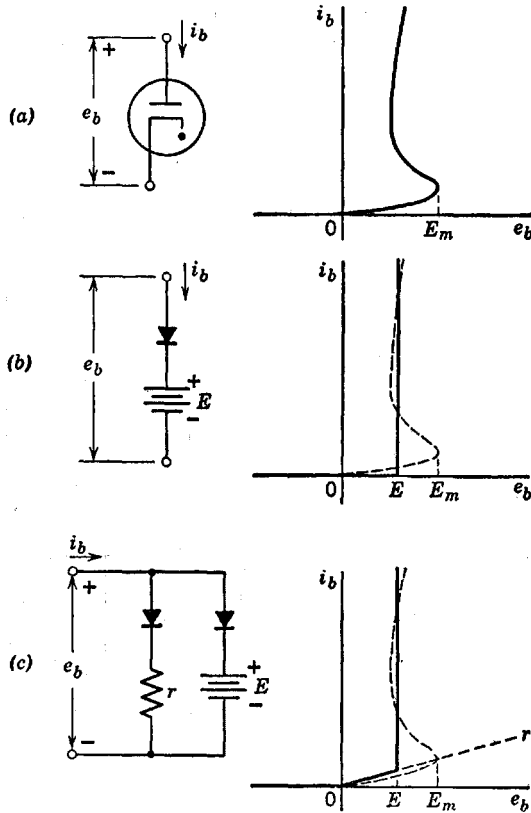
Fig. 3.19. Piecewise-linear models for gas-filled diodes.

approximation made to follow the curve more closely must consist of resistance r for $0 < e_b < E_m$. Then a line of negative slope is required in order to return to the vertical line at $e_b = E$. This negative slope implies the necessity for an element corresponding to negative resistance. A discussion of negative resistance and its properties is deferred until Chapter 9.

(a)   (b)

(c)   (d)

(1) $G_0 = 1/R_0$ (4) $R_3 = R_c - R_b$

(2) $G_1 = 1/R_1 = G_a - G_0$ (5) $R_4 = R_d - R_c$

(3) $G_2 = 1/R_2 = G_b - G_a$ (6) $G_5 = 1/R_5 = G_e - G_d$

(e)
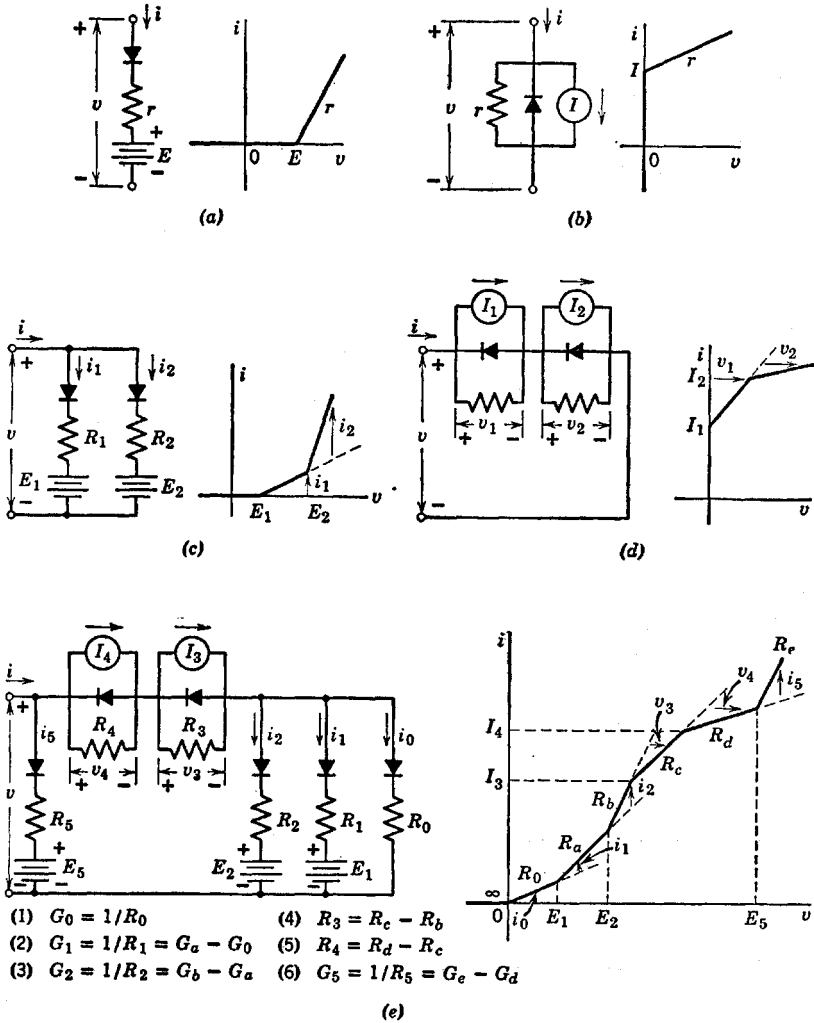
Fig. 3.20. Piecewise-linear model for an arbitrary nonlinear resistance.

3.9 Piecewise-Linear Model for Arbitrary Nonlinear Resistance

The basic ideas presented thus far can be extended to the representation of an arbitrary nonlinear resistance or conductance, provided we restrict ourselves to positive slopes or allow the use of negative resistance in the models.

The simple circuits shown in Fig. 3.20(a) and (b) yield concave-up

and concave-down breaks, respectively, on a plot of current versus voltage. A combination of such sections can be used for a succession of increasing or decreasing conductance segments, as shown in Fig. 3.20(c) and (d).

Representation of an "arbitrary" curve is shown in Fig. 3.20(e). The procedure followed here simply amounts to grouping together those segments that form concave-up regions of the curve. Similarly, concave-down regions are grouped together. Models of the successive "up" and "down" regions are then put together in series and parallel as required. The form of the resulting model is the familiar ladder network with "up" and "down" regions represented by shunt and series ladder network elements, respectively.

Suppose the sequence of segments in Fig. 3.20(e) represents the piecewise-linear approximation to some general resistive curve $i = f(v)$. As v increases from zero, the first segment of slope $G_0 = 1/R_0$ is represented directly by the shunt branch on the right end of the model. The second segment has conductance G_a; hence conductance $G_1 = G_a - G_0$ must be added in parallel, beginning at $v = E_1$. The same considerations apply to the next segment, where $G_2 = G_b - G_a$ must be added, beginning at $v = E_2$.

At current I_3, the slope decreases; hence series resistance must be added. It is apparent from the graph that $R_c > R_b$; hence R_3 is positive. The relations between circuit elements and slopes on the graph are tabulated in Fig. 3.20(e).

The sequence of diode operation for increasing v proceeds from right to left in the circuit model. Designating the diodes by the subscripts of the associated resistances, D_0 closes as v goes from negative to positive values. D_1 closes for $v = E_1$, and D_2 closes for $v = E_2$. Then, D_3 opens when v reaches the value corresponding to $i = I_3$, and D_4 opens when $i = I_4$. Finally, D_5 closes when $v = E_5$.

A few pertinent comments are in order about the basic shunt and series elements used in this ladder-network development of the arbitrary $i = f(v)$ curve. The form of the shunt element shown in Fig. 3.20(a) is convenient because it places the voltage coordinate of the break point in evidence. Also, since currents are added for shunt elements, it is convenient to have the break point at $i = 0$ so that no contribution of current from this element exists to the left of the break point. Similarly, the series element shown in (b) of the same figure places the current coordinate of the break point in evidence, and contributes no voltage drop to the left of the break point.

The variations of the basic forms shown in Fig. 3.21 serve to emphasize the important point that reference directions for v and i are part of the
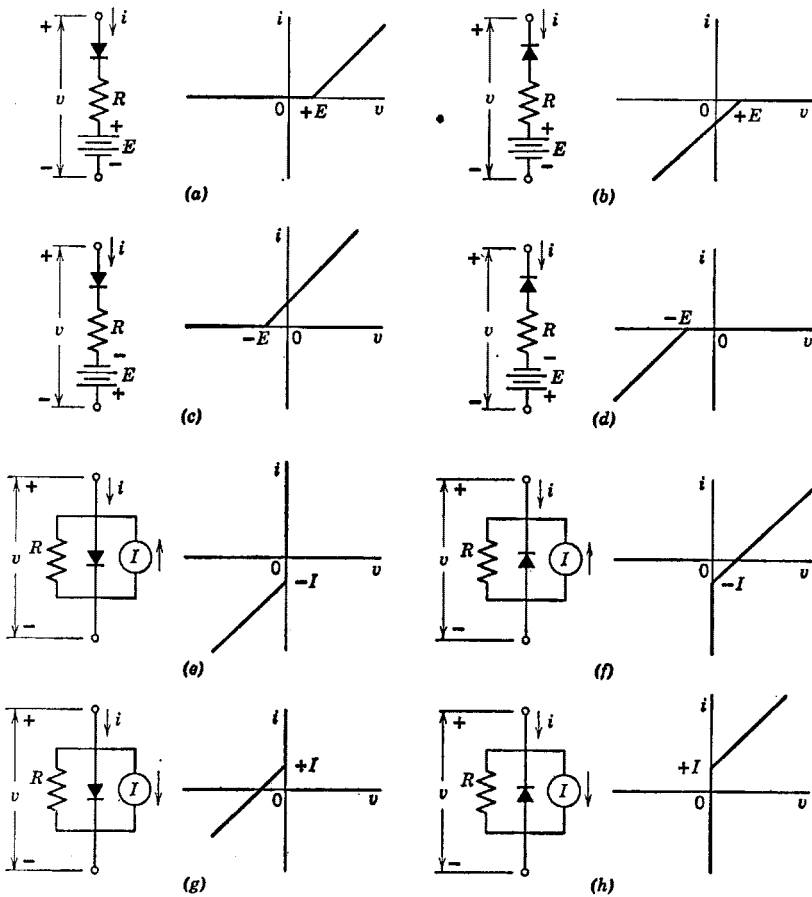
Fig. 3.21. Variations of basic resistive diode circuit forms.

circuit model. The series E and r forms shown in (a), (b), (c), and (d) yield a break point on the $i = 0$ axis. The shunt I and r forms shown in (e), (f), (g), and (h) yield a break point on the $v = 0$ axis. The position of the nonzero coordinate of the break point in each case depends only on the magnitude and polarization of E or I. The fact that the break joins two segments that are concave up or concave down depends only on the direction in which the diode is oriented relative to the assigned reference directions.

Circuits of the basic forms can be combined to yield a break point with both v and i nonzero. Examples of such forms are shown in Fig. 3.22. In the circuits shown in (a) through (f), the sources are
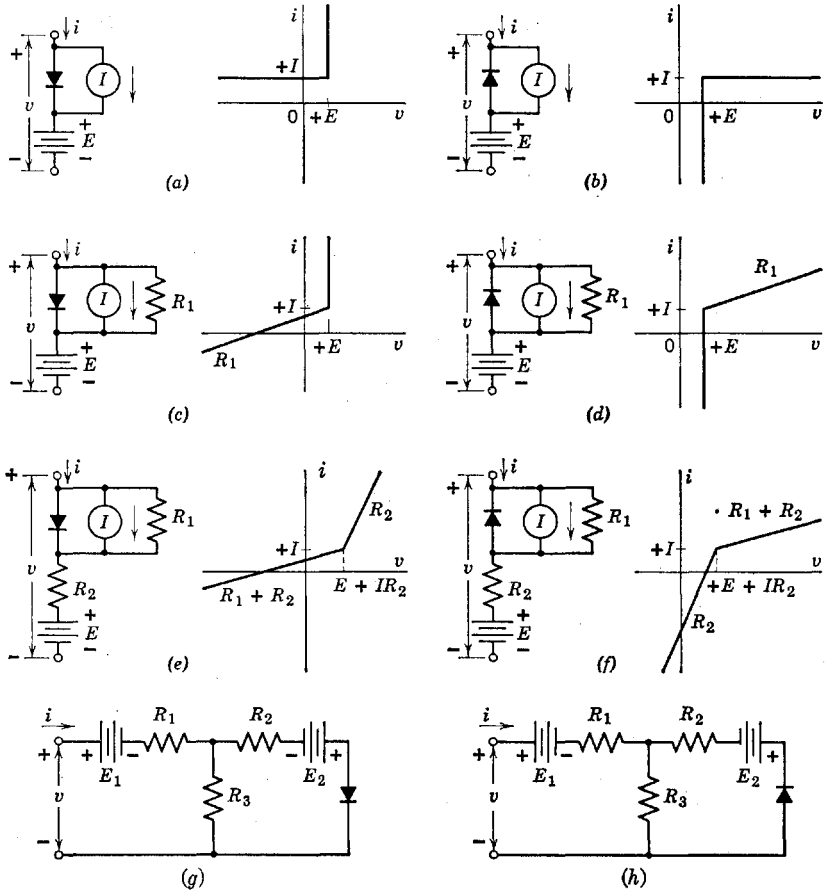
Fig. 3.22. Circuits with nonzero break-point coordinates.

polarized to position the break point in the first quadrant. By appropriate choice of polarities for E and I the breaks can be placed in the second, third, or fourth quadrants. The circuits shown in (g) and (h) are general forms that permit positioning of break coordinates and control of slopes.

More formal synthesis procedures can be used to realize a general nonlinear curve with resistive-diode circuits.* However, a remarkable degree of generality can be achieved using only the few simple ideas discussed thus far.

---

* T. E. Stern, Piecewise-Linear Network Theory, Technical Report No. 315, Research Laboratory of Electronics, M.I.T., June 15, 1956.

3.10   Analysis of Piecewise-Linear Circuits

Devising piecewise-linear circuits to approximate specified curves of voltage versus current, as illustrated by the several preceding articles, is an elementary form of synthesis or design. The analysis problem applies to a circuit that already exists and may require determination of voltage and current at a pair of terminals (driving-point problem), or the voltage or current response at one pair of terminals due to a voltage or current applied at another pair of terminals (transfer problem). For example, the response at the output of a circuit to a sinusoidal or square-wave input may be desired.

In the following articles, the principal methods of analyzing nonlinear resistive diode circuits will be used to derive piecewise-linear curves for a few simple circuits. These methods are based either on the determination of slopes and intercepts for each state or of coordinates for the break points between states.

3.11   The Method of Assumed Diode States

The method of assumed diode states is based on the fact that over any linear segment of a piecewise-linear curve, each diode in the corresponding circuit remains in a specific state, either on or off. Circuits are analyzed by assuming all possible combinations of states of the ideal rectifiers and finding the corresponding linear relation of i vs. v for each combination. In Fig. 3.23(a) the assumed states and the constraints on variables imposed by these assumptions are listed next to the circuit diagram. The linear relation for i vs. v in each assumed state is calculated from the simplified circuits shown in (b) and sketched as a dotted line in (c). The desired curve must consist of portions of these linear relations. Determination of the appropriate portions of the lines requires consideration of what is happening in the circuit. This is best accomplished by deducing circuit behavior while mentally varying v (or i) from large negative to large positive values.

To begin, let us consider values of v much more negative than $-E$. This condition would make D_1 nonconducting ($i = 0$); therefore, it corresponds to state I or II. When D_1 is *off*, the input is disconnected from the remainder of the circuit, and D_2 is held *on* by the current E/R_2 supplied from source E. Thus, state I can never actually exist and the circuit must be in state II when v is less than zero. Diode D_1 will switch from *off* to *on* when $v = 0$ (since $v_1 = v - v_2$). At that

point, current $i = 0$; hence, there is no possible effect on D_2. Thus, D_2 remains closed, and the circuit proceeds to state IV as v crosses zero. Now as v increases, the current i will increase; and since i passes through
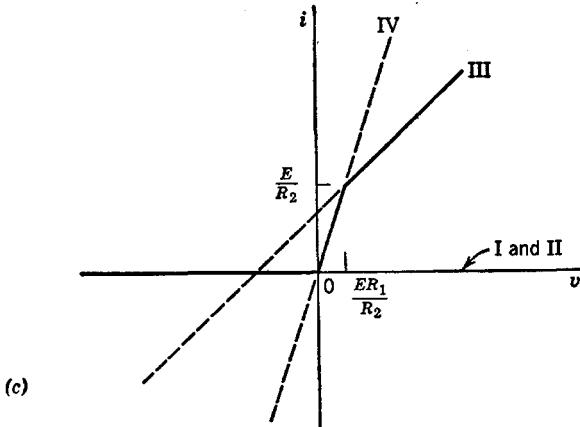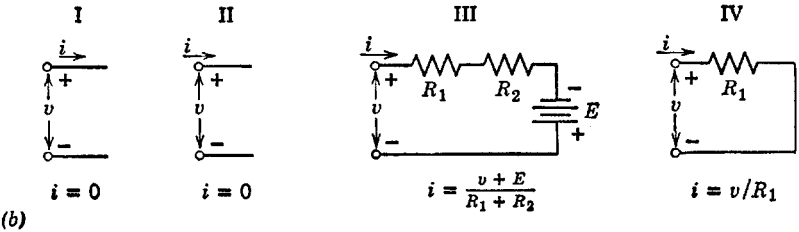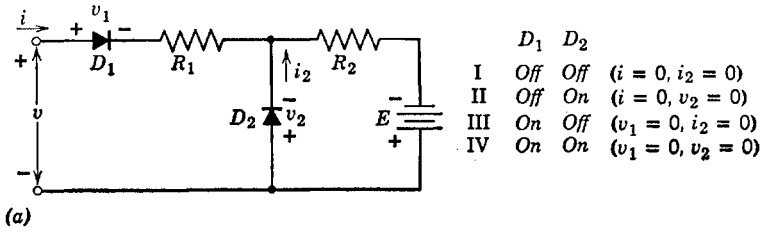

	D_1	D_2	
I	*Off*	*Off*	$(i = 0, i_2 = 0)$
II	*Off*	*On*	$(i = 0, v_2 = 0)$
III	*On*	*Off*	$(v_1 = 0, i_2 = 0)$
IV	*On*	*On*	$(v_1 = 0, v_2 = 0)$

(a)

(b)

(c)

Fig. 3.23.   Linear analysis based on assumed diode states.

D_2 in a direction opposite to that of the bias current, the net current i_2 is diminished. When $i = E/R_2$, the current i_2 is zero, and D_2 opens, putting the circuit into state III. Thus, as we vary v from large negative

values to large positive values, we proceed from state II to state IV, and then to state III.

If several diodes are involved in a circuit, the order of occurrence of the various states may be indefinite unless numerical values are specified for the resistances and batteries. With numerical values known, let the input variable (v or i) be increased from negative values to positive ones. Now suppose a given circuit is in a particular state with several diodes *on* and several *off*. To determine which state occurs next as the input variable (v or i) is increased, consider the negative voltage across each diode that is *off* and the positive current through each diode that is *on*. The first one of these quantities to reach zero determines the next state. Note when a diode is *on* that the positive current tells us how much it is conducting while the fact that the voltage is zero tells us only that the diode is *on*. Similarly, when a diode is *off*, the negative voltage tells us how far the diode is biased away from the break point, whereas the fact that the current is zero tells us only that the diode is *off*.

The method just outlined has a distinct shortcoming. It grows cumbersome very rapidly as circuit complexity increases. With three ideal diodes, there are eight combinations of *on* and *off*, and four diodes lead to sixteen combinations. In general, n diodes have 2^n combinations of *on* and *off*. Thus, even though preliminary reasoning eliminates some of the possible states, the problem mushrooms rapidly.

3.12   The Break-Point Method

The second approach is called the break-point method. It is based on the fact that a circuit containing only linear or piecewise-linear elements—such as batteries, resistances, and ideal rectifiers—must yield driving-point and transfer curves that are piecewise linear. The coordinates of the break points specify the character of the entire curve except for each of the end segments, which must be determined by the outermost break point and a resistive slope. Two adjacent break points (in terms of the independent variable) specify the state line joining the break points. Thus each break point joins two states, so that a circuit with n states has $n - 1$ break points.

The usefulness of the method stems from the fact that both the current and the voltage for a particular ideal rectifier are constrained to zero at the break point of that rectifier. These two constraints when applied to the circuit reduce an n-diode circuit to one containing only $(n - 1)$ diodes.

Let us illustrate the break-point method with the same circuit used

in the previous section. At the break point of D_1, $v_1 = i = 0$. Since R_1 carries no current, the right-hand loop of the circuit behaves as though R_1 were removed. Thus D_2 conducts a current $i_2 = E/R_2$ with voltage
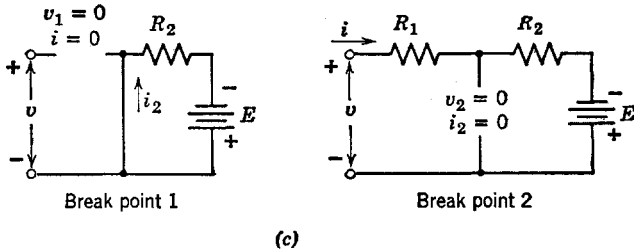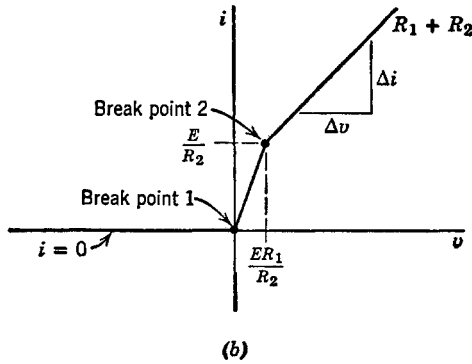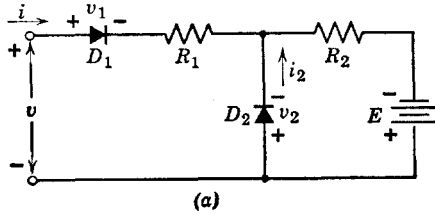

(a)

(b)

(c)

Fig. 3.24. Break-point method of analysis.

$v_2 = 0$. Now, since both v_1 and $v_2 = 0$, the input voltage v is zero, and the first break point is thus established at the origin where $v = i = 0$.

The break point of D_2 corresponds to $v_2 = i_2 = 0$. The current in R_2 must then be E/R_2; and since $i_2 = 0$, current i entering the circuit must also be E/R_2, and the first diode must be closed. Now, since $v_1 = v_2 = 0$, we have $v = iR_1 = ER_1/R_2$ for the voltage coordinate
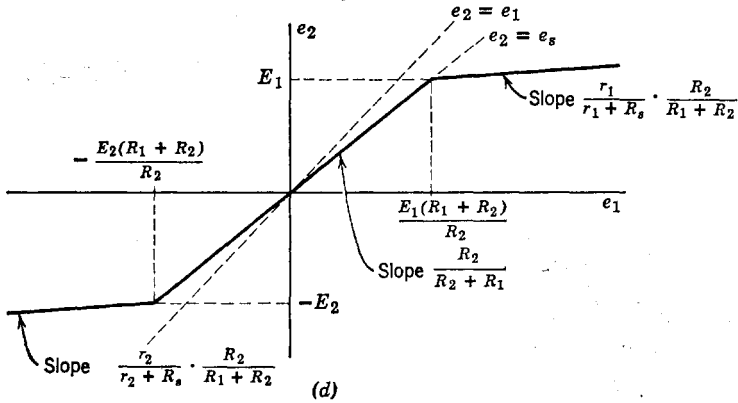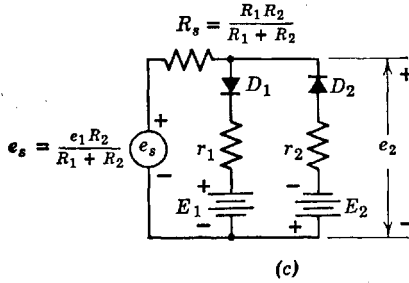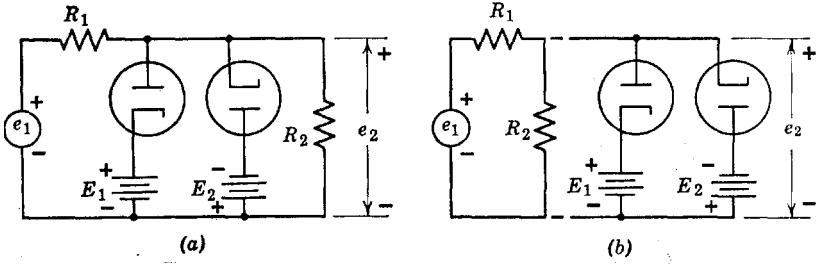
Fig. 3.25. Transfer curve for limiter.

of the break point, as shown in Fig. 3.24(b). The circuits that determine the break-point coordinates are shown in (c). To the left of break point 1 the first diode is open, and $i = 0$. The curve therefore proceeds to the left along the horizontal axis. To the right of break point 2 the first diode D_1 is closed, and the second diode D_2 is open, so that $\Delta v / \Delta i = R_1 + R_2$. Between the two break points the circuit is linear, so the curve must be a straight line joining points 1 and 2.

Let us now illustrate the determination of a transfer curve by the break-point method. The circuit shown in Fig. 3.25(a) is called a double-ended clipper or limiter. Before proceeding with the analysis, let us simplify the linear portion of the circuit, redrawn in (b) to emphasize the separation of the linear and nonlinear parts. From the terminals at which e_2 is to be obtained, the source e_1, its internal resistance R_1, and the load resistance R_2 may be replaced by a new source voltage e_s with a new internal resistance R_s, as shown in (c). The voltage e_s is the open-circuit effect of e_1, considering the two shunt-diode circuits as loads external to the linear circuit. The resistance R_s is the apparent resistance between the output terminals.

For the circuit of Fig. 3.25(c) each vacuum diode has been approximated by an ideal rectifier in series with a small resistance. The break point for ideal diode D_1 occurs at $e_2 = +E_1$. At this point, D_2 must obviously be *off*; hence, $e_2 = e_s = E_1$ and $e_1 = E_1(R_1 + R_2)/R_2$. For e_2 greater than $+E_1$, D_1 is *on*, D_2 is *off*, and the slope $\Delta e_2/\Delta e_s = r_1/(R_s + r_1)$. The break point for D_2 occurs at $-E_2$, and at this point D_1 is *off*; hence, $e_2 = e_s = -E_2$ and $e_1 = -E_2(R_1 + R_2)/R_2$. For e_2 more negative than $-E_2$, diode D_2 is *on*, and slope $\Delta e_2/\Delta e_s = r_2/(r_2 + R_s)$. The slope of the line joining the two break points can always be calculated from the break-point coordinates, but in this case the circuit is so simple (both diodes are *off*) that the value can be found by inspection.

3.13   Simple Multidiode Circuits

When a diode circuit contains no internal batteries or current sources, all break points occur at the origin. The analysis problem then becomes relatively simple, even when more than two ideal diodes are present. For illustration, consider the three full-wave rectifier circuits shown in Fig. 3.26. These circuits are so named because both positive and negative values of the input voltage e_1 produce a positive output voltage e_2. If a sinusoidal waveform of voltage is applied at e_1, the output e_2 consists of adjacent half-sinusoids of the same polarity.

Inspection of Fig. 3.26 indicates that the transfer curve for all three circuits should have even symmetry, that is, $f(-e_1) = f(+e_1)$. The circuits shown in (a), (b), and (c) have successively greater transfer slopes, as indicated by the corresponding curves of e_2 vs. e_1. However, the slope of the transfer curve is not the only basis for judging the relative merits of the three circuits. The input resistance $R_{in} = e_1/i_1$ lies between R_1 and $2R_1$ for circuit (a) and between $R_1/2$ and R_1 for

circuit (b). Hence more power must be supplied to circuit (b) for the same input voltage e_1. Circuit (c) provides maximum transfer slope
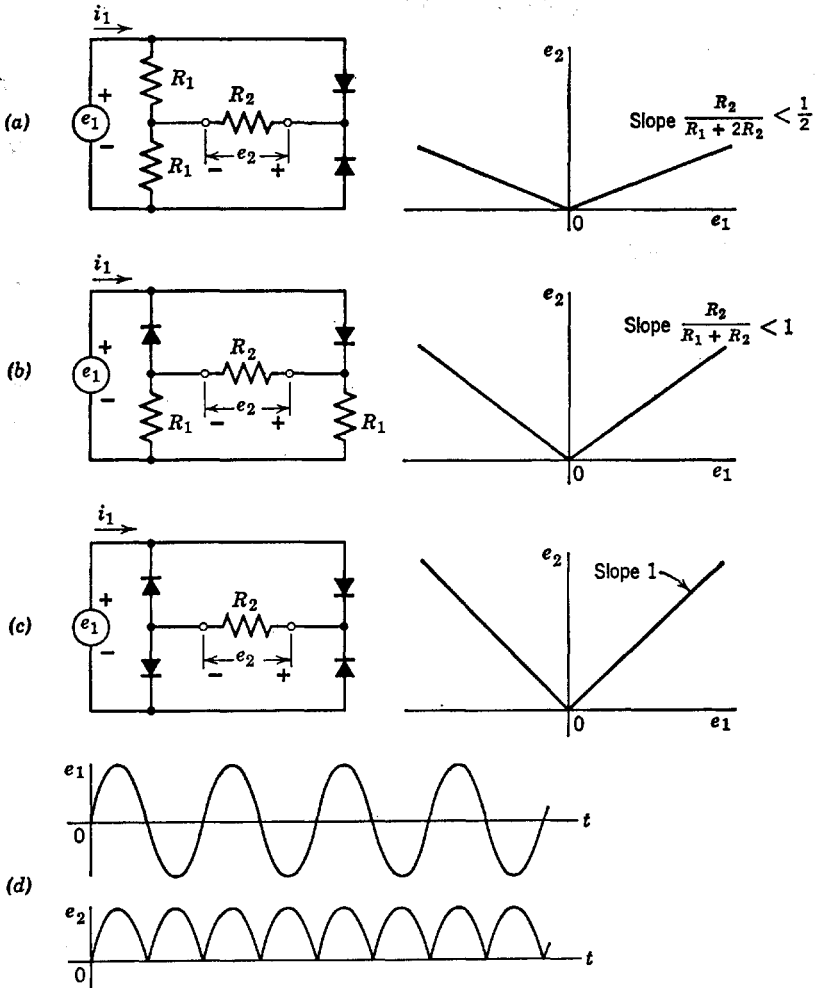

Fig. 3.26. Full-wave rectifier circuits.

(unity) and an arbitrary input resistance $R_{in} = R_2$, at the cost of two additional diodes.

Waveforms of input and corresponding output are shown in Fig. 3.26(d) for the circuit of (c). The output for circuit (a) or (b) would be similar in form but with smaller amplitude. The full-wave rectifier
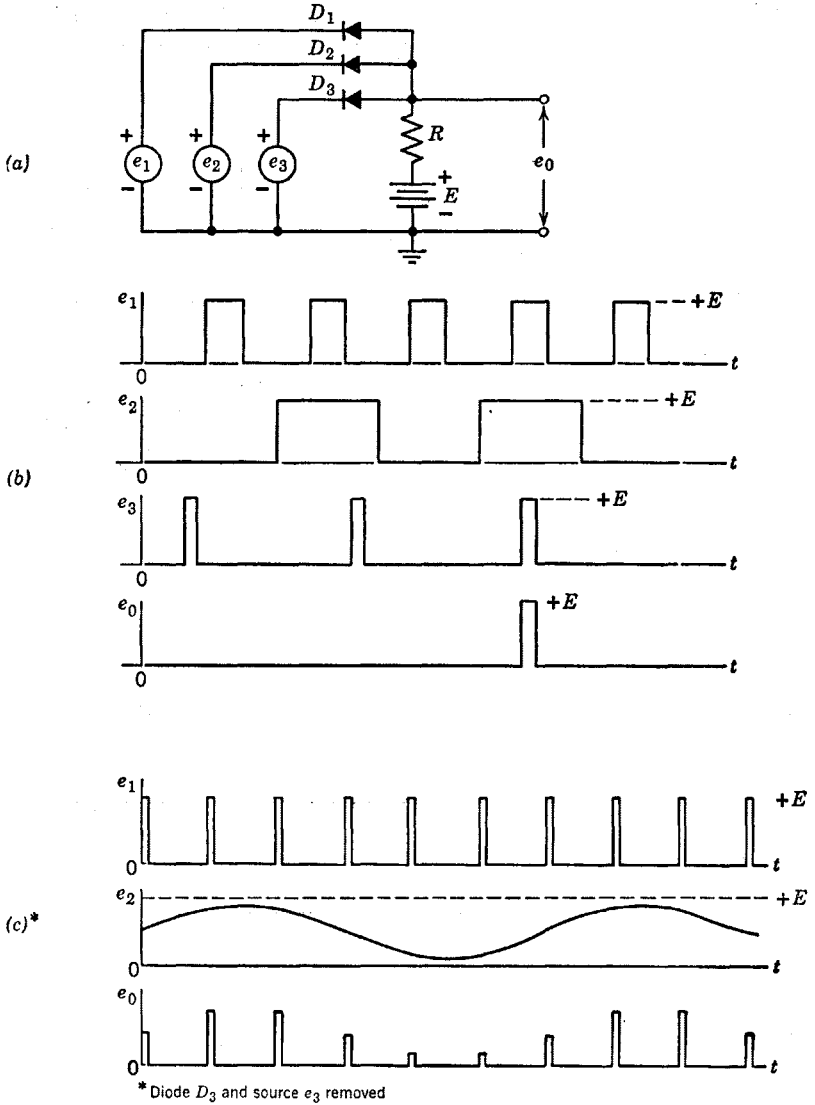
Fig. 3.27. Diode gate circuit (AND circuit).

circuits described here, or variations thereof, find many applications in electronic systems.

Another class of resistive diode circuits, called gate circuits, are used to permit or prevent the passage of electrical signals. Switching circuits

of all kinds and digital computers, in particular, use gate circuits to perform a variety of operations.

The gate circuit shown in Fig. 3.27(a) is called an AND circuit because it requires coincidence of signals at e_1, e_2, and e_3 to produce an output. Examination of the circuit shows that in the absence of inputs all diodes conduct (and the output voltage is zero) if the diodes

(a)

(b)
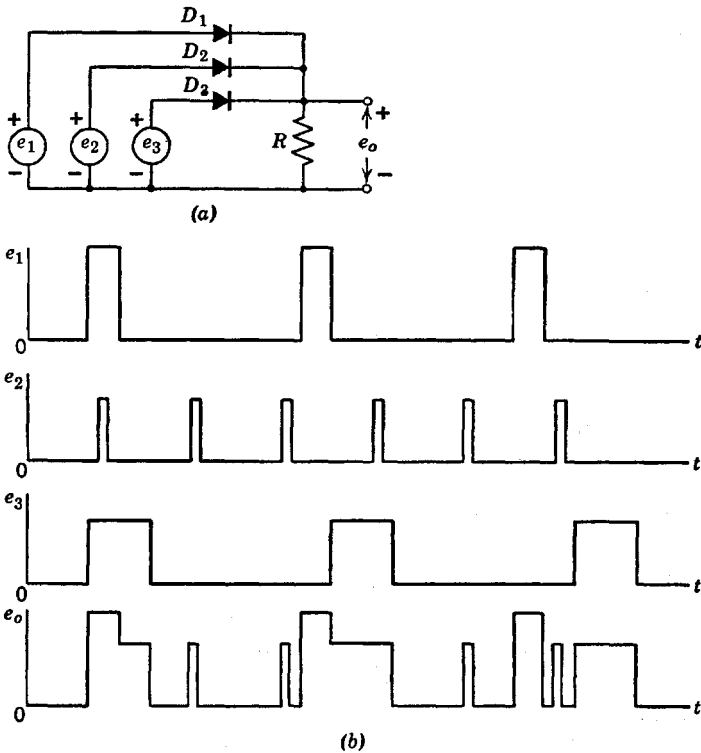
Fig. 3.28. Diode gate circuit (OR circuit).

and sources are assumed to be ideal (zero resistance). A positive pulse from any one or two of the sources simply turns off the corresponding diodes, but e_0 remains zero. Thus, e_1 and e_2 raise the cathodes of their diodes to a positive value, while the anodes are held at zero by the conduction of D_3. If a positive voltage is present at all three inputs, an output voltage occurs. This action of the circuit is illustrated in Fig. 3.27(b) for input voltages consisting of rectangular pulses. The output-pulse duration corresponds to the "overlap" of the input pulses.

If the input signals are unequal but less than E, the output voltage will have an amplitude very nearly equal to the amplitude of the smallest input signal. Thus the circuit is a basic form of pulse amplitude modulator, as indicated in (c). Alternatively, it may be called a sampling circuit, since each rectangular pulse, in effect, samples the instantaneous value of the sine wave. Only two inputs are necessary in this case; hence we may assume that source e_3 and diode D_3 have been disconnected from resistor R. The maximum positive value of the voltage e_2 should be slightly less than the peak pulse height; otherwise the pulse height will partially determine the output amplitude. Either of the circuit functions illustrated in Fig. 3.27(b) and (c) can be carried out with negative voltages by reversing the polarity of source E and reversing the diodes.

Another type of gate circuit, often called an OR circuit, is shown in Fig. 3.28(a). In this case, an input voltage at e_1, e_2, or e_3 will produce a corresponding output voltage. The circuit therefore performs the function of combining signals from various sources. As indicated in Fig. 3.28(b), the output amplitude will be dictated by the maximum input amplitude, in the event that several inputs occur simultaneously, because the largest input amplitude biases the remaining diodes in the reverse direction.

3.14   Stepwise Approximation

The circuit models devised for piecewise-linear approximations of nonlinear curves use resistances, ideal rectifiers, voltage sources, and current sources. If we eliminate resistances from this list of elements, the circuit models that can be devised result in stepwise approximations.

The ideal rectifier is actually a stepwise approximation to an actual diode rather than a piecewise-linear approximation, for neither segment of its i vs. v curve is really resistive. In the conducting direction, the ideal rectifier is a short circuit, which can be interpreted as a special case of voltage source $(E = 0)$. In the reverse direction we have infinite resistance, which can be interpreted as a special case of current source $(I = 0)$. Thus, the ideal rectifier can be considered to have a bit of the character of ideal sources. In fact, connected in series with a voltage source or in parallel with a current source it merely makes the ideal linear sources unidirectional.

The stepwise representation of a smooth curve, as shown in Fig. 3.29(a), may seem to be a rather gross approximation. Actually it can provide a very good average fit to the nonlinear curve. Since the area

under a stepwise approximation increases in a series of linear ramps, as shown in Fig. 3.29(b), we observe that the integral of a stepwise approximation leads to a piecewise-linear curve. A circuit model for the stepwise approximation is shown in (c).
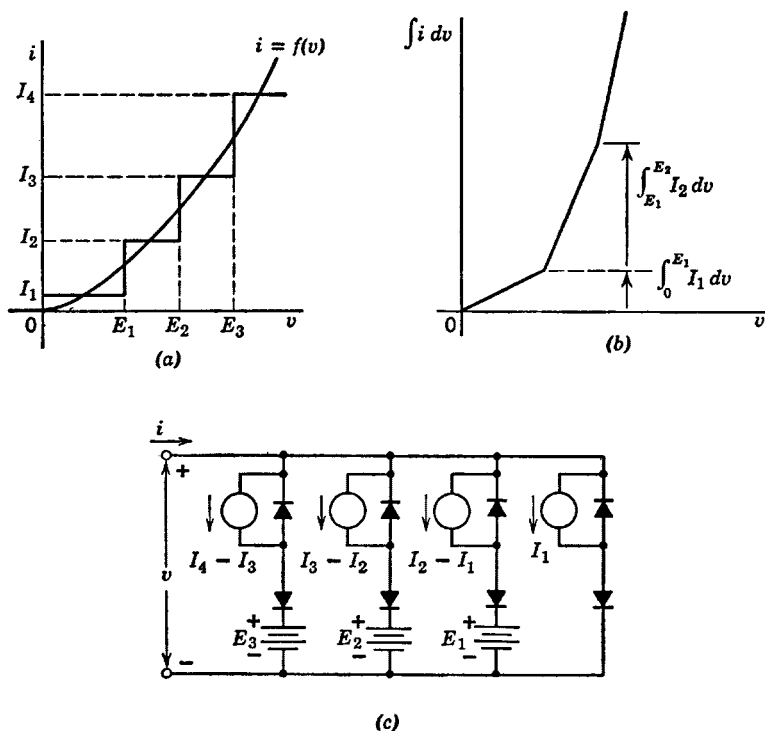

(a)

(b)

(c)

Fig. 3.29. Stepwise approximation.

3.15 Parabolic Approximation

A square-law curve can obviously be used to approximate some portion of a general nonlinear curve. As shown in Fig. 3.30(a), the parabola $i = K_2 e^2$ fits the vacuum-diode curve $i_b = K_1 e_b^{3/2}$ about as well as the linear resistance r_b over a given range of values. A somewhat better fit is obtained over the range of voltage and current indicated, by shifting the curve upward, as in (b), to place some of the parabola above and some below the vacuum-diode curve. By inspection of this curve we see that a reduction in the constant, accompanied by a shift to the

$$r_b = E_b/I_b$$

$$K_1 = I_b/E_b^{3/2}$$

$$K_2 = I_b/E_b^2$$
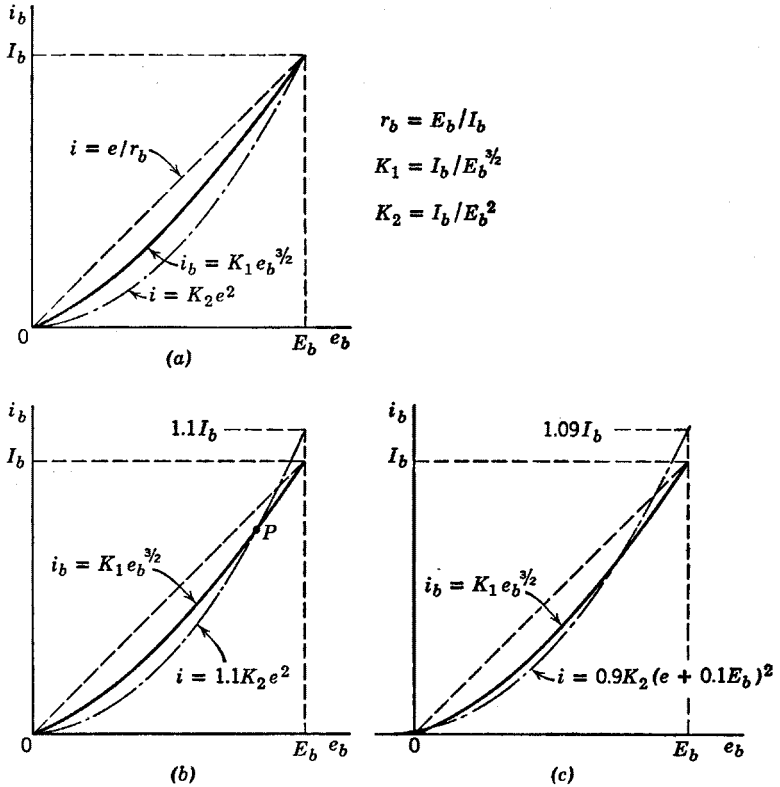
Fig. 3.30. Parabolic approximations to $\frac{3}{2}$-power curve.

left, should provide an even better fit, as shown in (c). This curve matches the diode curve very well over the range of values indicated.

If a nonlinear curve of i vs. e were approximated by several parabolic segments (concave up and down), the derivative curve di/de vs. e would be piecewise-linear, provided the slopes of the parabolic segments match at their junctions. If the slopes did not match, the curve of di/de vs. e would contain steps.

The square-law relation $i = Ke^2$ is often used in analyzing detector circuits, mixer circuits, and the like. See Art. 4.14.

3.16 A Simple Electrical Transient

Thus far in this chapter we have examined some of the properties of resistive-diode circuits. We have considered methods of approximating

the current versus voltage curves of various types of diodes or of an arbitrary nonlinear resistance. In this article we shall consider a circuit with one energy-storage element to show that functions of time or frequency also can be approximated to good advantage. Usually the resulting simplification of the analysis problem justifies the use of such approximations.

Let us consider the most elementary example possible. Suppose a capacitor C has been charged to a voltage $E_b = Q/C$. Let us determine the waveforms $e(t)$ and $i(t)$ resulting when the capacitor is discharged through various resistive circuits, both linear and nonlinear.

The circuit, equations, and waveforms shown in Fig. 3.31 (a) illustrate the familiar linear circuit problem. The solution of the linear differential equation is a decaying exponential characterized by the time constant $\tau = RC$.

Now suppose the capacitor C is to be discharged through a vacuum diode described by a three-halves power curve, as indicated in Fig. 3.31 (b). The simple nonlinear differential equation, when solved, shows that the voltage and current no longer vary exponentially with time, but rather, in an inverse square manner. The solution for $v(t)$ is given analytically and graphically in (b). If now the diode is approximated by r_b, the solution has the same form as in (a). This exponential waveform is shown as a dashed line in the waveform plot of Fig. 3.31 (b). Since the r_b approximation lies entirely above the three-halves power curve, and since $i = -C\,de/dt$, the discharge by the vacuum diode actually proceeds more slowly than the exponential determined by the r_b approximation. It is apparent that the exponential waveform can be matched to the inverse square expression more closely by using a closer approximation to the diode curve than that given by the r_b line.

In Fig. 3.31 (c) the waveform $v(t)$ is plotted for a square-law curve of i vs. v. Again the expression for $v(t)$ is readily obtained by separation of variables and integration of the simple expression. The exponential discharge produced by an r_b approximation is shown for comparison. Since the square-law curve of i vs. v lies farther below the r_b line than does the three-halves power law, the discharge proceeds more slowly in this case than in (b).

For each of the examples given in (a), (b), and (c), the waveform of current $i(t)$ can be obtained from $v(t)$ by substitution in the appropriate i vs. v equation, or by differentiation, since $i = -C\,dv/dt$.

Let us now consider the waveforms that result from a two-segment piecewise linear approximation for a diode curve, as shown in Fig. 3.32 (a) (a three-halves power curve is assumed for convenience). Each segment of this piecewise-linear approximation yields an exponen-

$$i = v/R = -C\,dv/dt \quad \begin{cases} \text{at } t = 0, \\ v = E_b \end{cases}$$

$$v = E_b \epsilon^{-t/RC}$$

(a)

$$i = -C\,dv/dt = Kv^{3/2} \quad \begin{cases} \text{at } t = 0, \\ v = E_b \end{cases}$$

$$r_b = E_b/I_b$$

$$i = Kv^{3/2}$$

Where $K = I_b/E_b^{3/2}$

$$\frac{v}{E_b} = \frac{4}{(2 + t/r_b C)^2}$$

$$\epsilon^{-t/r_b C}$$

(b)

$$i = Kv^2 = -C\,dv/dt$$

$$i = Kv^2$$

Where $K = I_b/E_b^2$

$$\frac{v}{E_b} = \frac{1}{1 + t/r_b C}$$

$$\epsilon^{-t/r_b C}$$

(c)
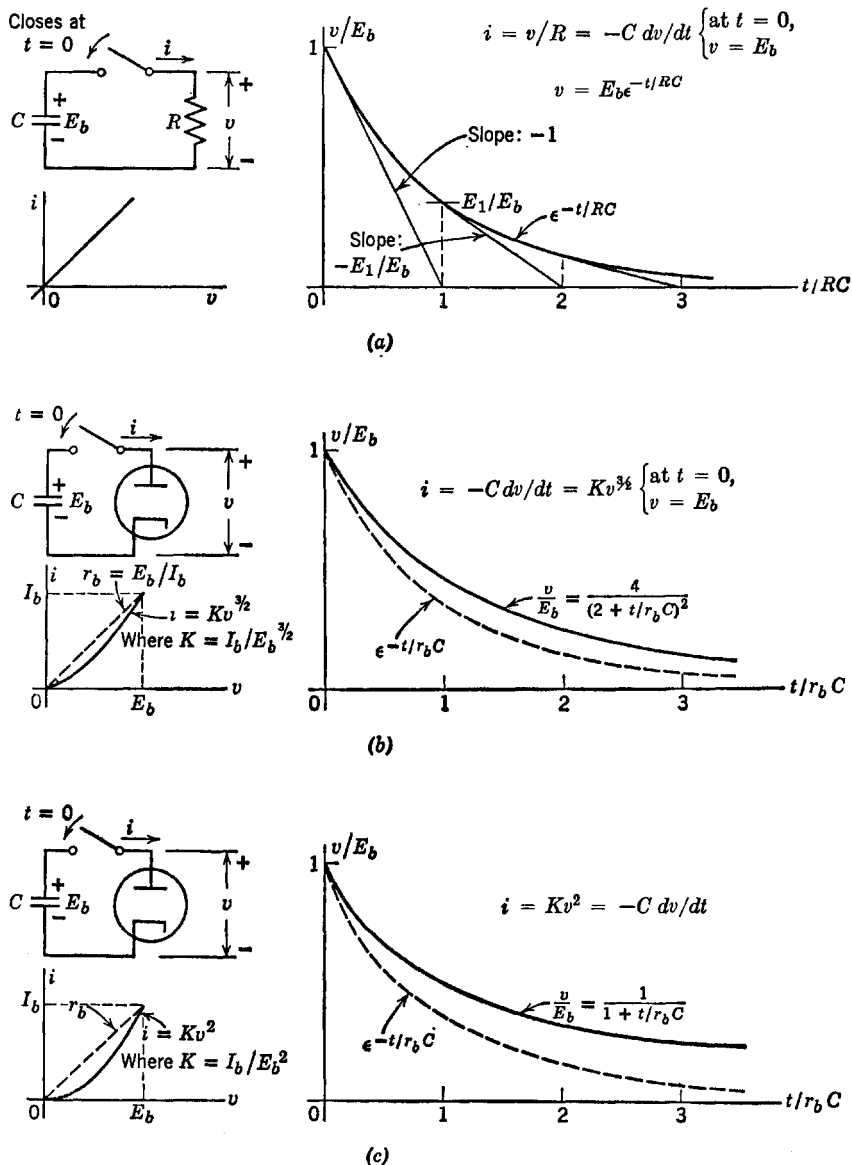
Fig. 3.31. Discharge of capacitor by linear and nonlinear resistances.

tial curve. The time constant and final value are different for each exponential; but since i is continuous at the break point, the slope of the voltage waveform (dv/dt) is also continuous at the transition from

one exponential to the other.   The slope of the approximation wave-
form of current, see Fig. 3.32(b), is not continuous at the break point.
   Note that the approximations match the true waveforms quite closely.
Since most of the applications of electronic circuits involve time-varying
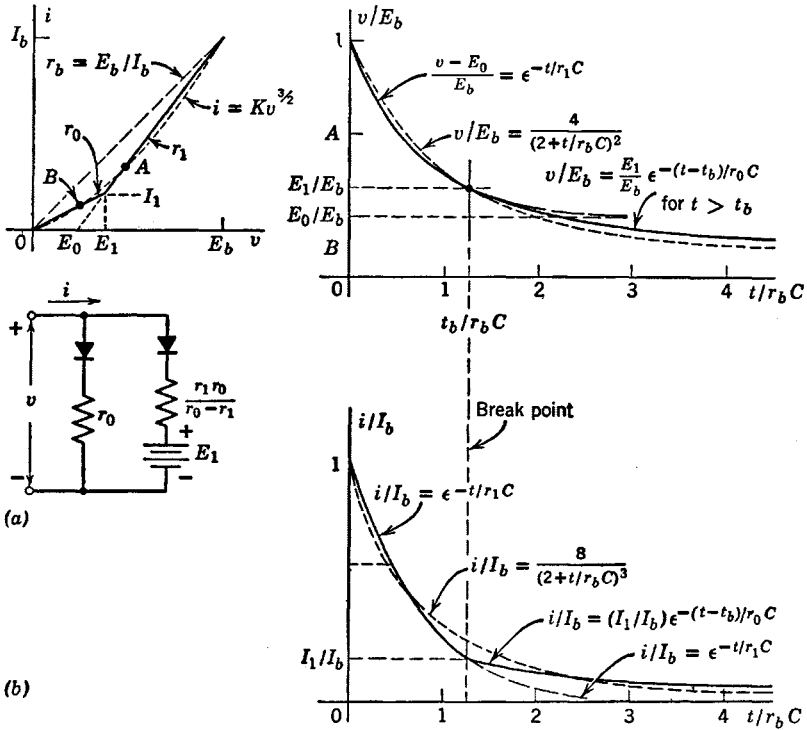
Fig. 3.32.   Use of a piecewise-linear approximation for calculating capacitor discharge.

voltages and currents in circuits containing inductance or capacitance
or both, it is encouraging to find that simple representations of
nonlinear i vs. v curves yield adequate approximations to time-vary-
ing behavior.

3.17   Solution of Transient Problem Using Stepwise Diode Approximation

If a nonlinear element exerts a major influence on the current or
voltage, a simple one- or two-segment piecewise-linear approximation
may be judged inadequate.   However, the piecewise-linear model

loses some of its convenience if more than one or two break points occur within the range of expected current or voltage variation.   In such cases the stepwise approximation for i vs. v is often useful.

Let the diode curve in Fig. 3.33 be approximated by a stepwise curve. The diode curve could be any function $i = f(v)$; but in order to have the true solution available for comparison, assume $i = Kv^{3/2}$.   The true
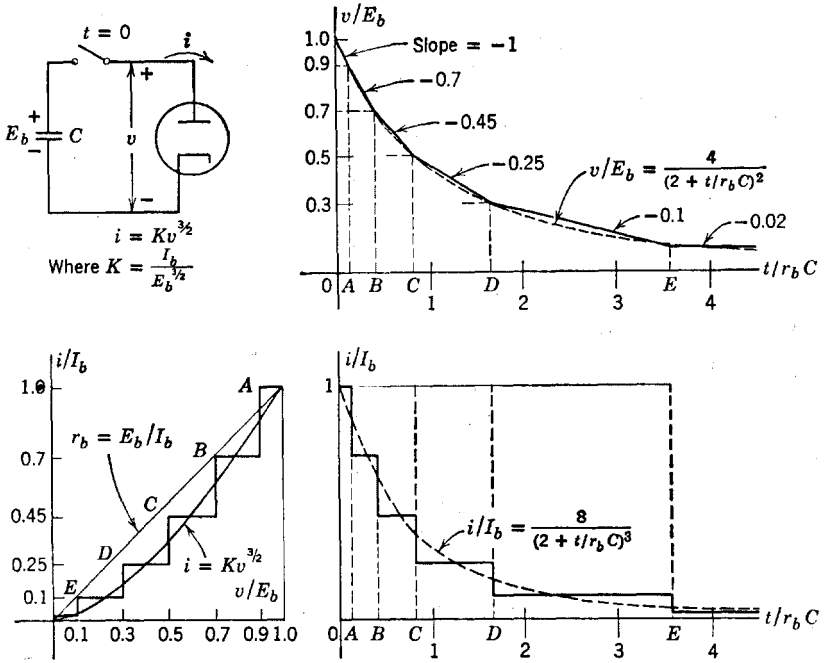
Fig. 3.33. Stepwise diode approximations to obtain piecewise-linear waveform approximations.

waveforms $v(t)$ and $i(t)$ for the capacitor discharge are shown by the dashed lines. The waveforms resulting from the stepwise approximation are shown by the solid line. Since $i = -C\,dv/dt$, the waveform $v(t)$ is piecewise-linear, whereas $i(t)$ is a stepwise curve.

In the example shown in Fig. 3.33, the stepwise curve was literally drawn to "look right." Despite this seemingly haphazard approach to the approximation problem, the resulting piecewise-linear waveform of $v(t)$ is almost everywhere within 10 per cent of the true curve derived from the three-halves power law. The waveform of $v(t)$ may be drawn very rapidly by beginning at $t = 0$, and drawing a line from the initial value of voltage to the value at break point A, with slope determined by the current.

The example shown in Fig. 3.33 includes many more segments than are required for making a reasonably good approximation. The sketches in Fig. 3.34 show the results obtained with smaller numbers of segments.
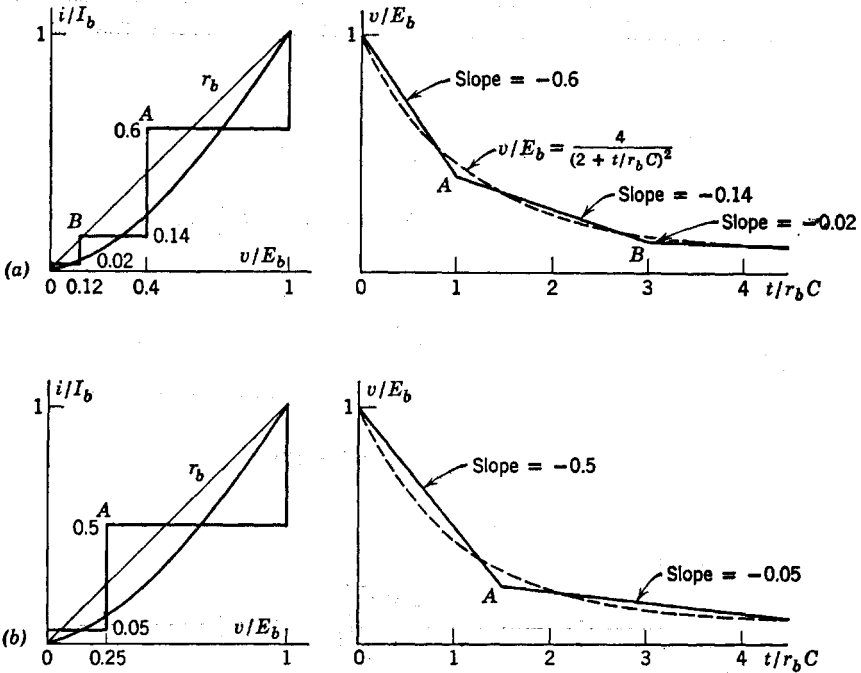
Fig. 3.34. Simple piecewise-linear waveform approximations.

3.18   Approximations for the Exponential Curve

The current and voltage transients of linear or piecewise-linear circuits are exponential in character. Since such transients will occur repeatedly in the analysis of electronic circuits, the properties of exponential waveforms, and a few simple approximations to them are important.

The function with which we shall most often be concerned is the form ϵ^{-x} as shown in Fig. 3.35(a). In this expression, $x = t/\tau$. The constant τ is merely a normalizing factor for the time scale. An exponential curve is completely determined when the initial and final values are known, together with the time constant. At each point on the curve, the slope is proportional to the magnitude (relative to the final value); that is, the slope at any specific value E_1 is $-E_1/\tau$. The area under the curve is finite and equal to the product of the initial value and the time constant. Mathematically, the value of the exponential approaches zero as time approaches infinity. For engineering purposes, however, the

exponential has reached completion after an elapsed time corresponding to several time constants. For example, when $t/\tau = 5$, the value of the exponential is 0.6 per cent of the initial value at $t = 0$.

The first term of the series expansion of the exponential makes a convenient approximation to the initial part of the exponential curve.
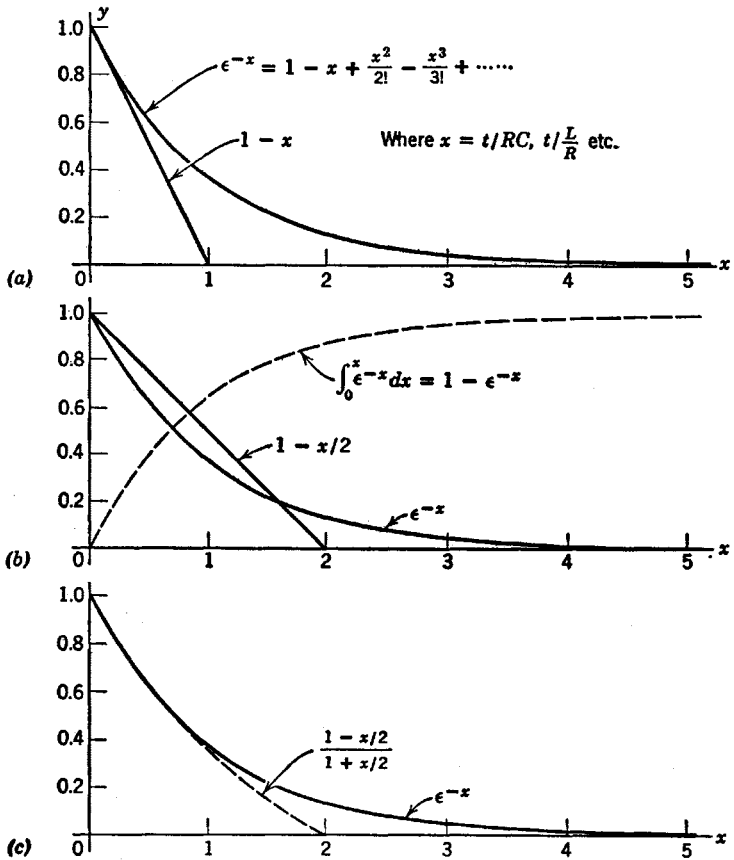
Fig. 3.35. Approximating exponential curves.

When x is small, the higher-order terms x^2, x^3, etc. may be neglected. This approximation is shown in Fig. 3.35(a).

The area under the exponential of (a) is unity, whereas the area under the triangle representing the approximation in (a) is only $\frac{1}{2}$. Hence, if a linear approximation is desired to match total area instead of initial slope, the line must reach zero at $x = 2$, as shown in Fig. 3.35(b).

The approximation given in Fig. 3.35(c) matches initial slope but not total area. It follows the true curve more closely than either of the other approximations. Higher-order approximations may be justified in some special cases, but since the exponential itself is fairly convenient to manipulate, one either uses simple approximations or the function itself.

SUPPLEMENTARY READING

B. Chance, *Waveforms*, McGraw-Hill, New York, 1948.
T. S. Gray, *Applied Electronics*, 2nd edition, John Wiley and Sons, New York, 1954.
E. A. Guillemin, *Introductory Circuit Theory*, John Wiley and Sons, New York, 1953.
J. Millman and H. Taub, *Pulse and Digital Circuits*, McGraw-Hill, New York, 1956.

PROBLEMS

3.1. Find the Thevenin equivalent of each of the circuits in Figs. P3.1–P3.4. Note that the two-terminal circuits can be simplified to a single resistor and single source, whereas the three-terminal circuits require two sources and three resistors.
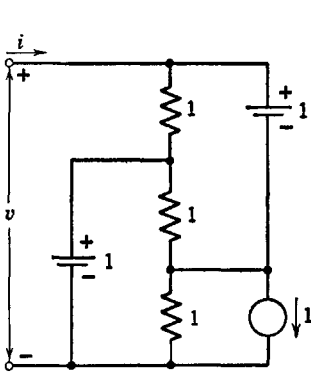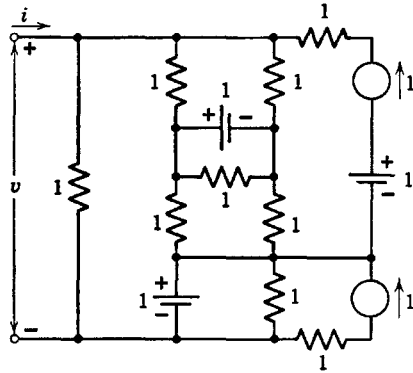
Fig. P3.1   Fig. P3.2

3.2. Reciprocity for a two-terminal-pair network can be expressed as a relation between two sets of measurements (A and B) performed on the same network (Fig. P3.5). If the networks are reciprocal, the following determinants must be equal.

$$\begin{vmatrix} V_1{}^A & I_1{}^A \\ V_1{}^B & I_1{}^B \end{vmatrix} = \begin{vmatrix} V_2{}^A & I_2{}^A \\ V_2{}^B & I_2{}^B \end{vmatrix}$$
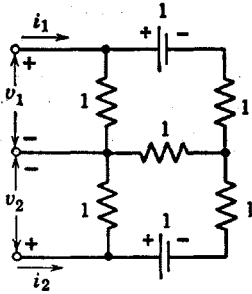
Fig. P3.3
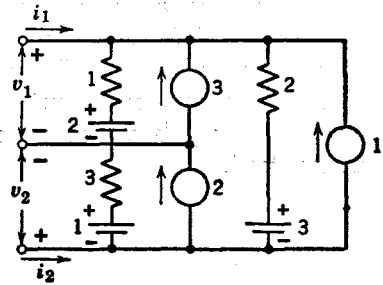
Fig. P3.4
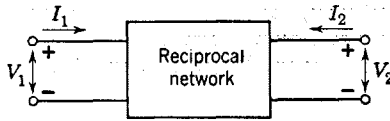
Fig. P3.5

This equation implies that

$$\left.\frac{I_2}{V_1}\right|_{V_2=0} = \left.\frac{I_1}{V_2}\right|_{V_1=0} \quad \text{and} \quad \left.\frac{V_2}{I_1}\right|_{I_2=0} = \left.\frac{V_1}{I_2}\right|_{I_1=0}$$

Thus any purely resistive, reciprocal two-terminal-pair network can be simplified to three resistors in a "T" or "π" configuration, as in Figs. P3.6 and P3.7.

Fig. P3.6

or
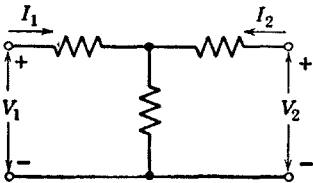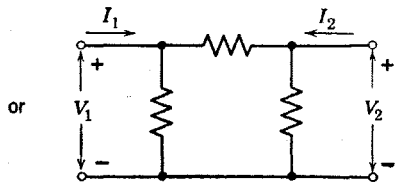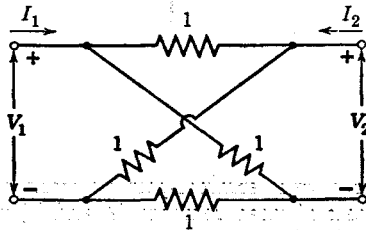
Fig. P3.7

Fig. P3.8

Using reciprocity, simplify the networks in Figs. P3.8, P3.9, and P3.10 to a T and π equivalent. Can you suggest how the above determinant expression
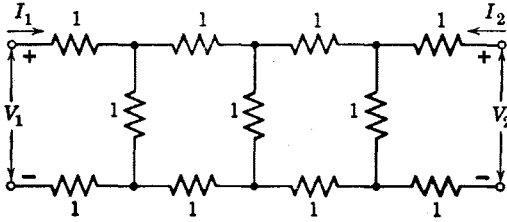
Fig. P3.9
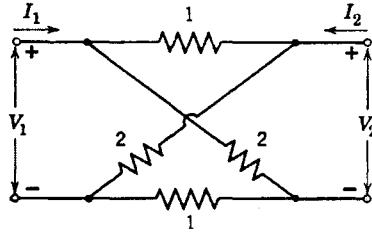
Fig. P3.10

of reciprocity might be proved for a purely resistive circuit using basic considerations (i.e., Ohm's law, conservation of energy, Kirchhoff's law, etc.)?

3.3. How many basically different circuits can be designed using a single battery, resistor, and ideal diode? Draw the circuits and sketch their associated $e-i$ characteristics. Some of the possible circuits are redundant, indeterminant, or merely special cases of another circuit.

3.4. (a) What is meant by the dual of a two-terminal linear circuit element? (b) What is the dual of an ideal rectifier?
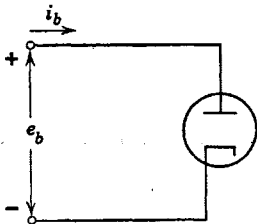
Fig. P3.11
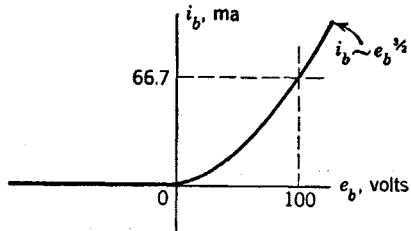
Fig. P3.12

3.5. A vacuum diode, Fig. P3.11, has the characteristic shown in Fig. P3.12. Determine suitable values of R and E to represent the diode by the circuit of Fig. P3.13, so that the approximation to the actual characteristic is very good in the neighborhood of $e_b = 100$ volts. (The ideal rectifier in Fig. P3.13 has the characteristic shown by Fig. P3.14.)

3.6. (a) Plot e_2 against e_1 for the circuit of Fig. P3.15, in which the diode characteristic is that of Fig. P3.12. Cover the range $-100 < e_1 < +200$. Also plot e_2 versus e_1 for Fig. P3.16.
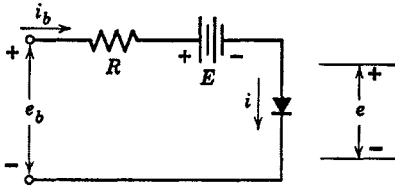
Fig. P3.13
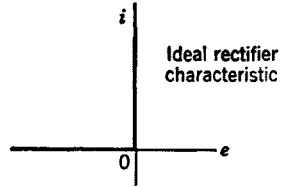
Fig. P3.14

(b) How serious is the error in the e_2 vs. e_1 characteristic due to replacement of the diode by an ideal rectifier?   How would the error be affected by a change in the load resistance from 10kΩ to some smaller value? Some larger value?
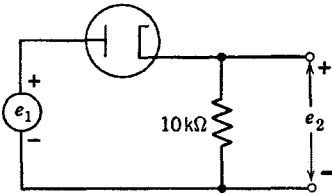
Fig. P3.15
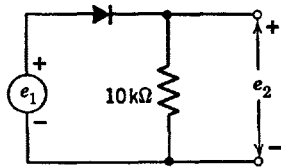
Fig. P3.16

3.7. A high-vacuum diode curve is closely approximated by the function $i_b = Ke_b^{3/2}$. It is known that $i_b = 100$ ma (milliamperes) when $e_b = 20$ volts.

(a) Determine K and specify its dimensions.

(b) Plot i_b vs. e_b over the range from 0 to 20 volts.

(c) It is desired to approximate the actual diode by a model consisting of an ideal rectifier in series with a resistance R. By adjusting a straightedge on the plot of part (b), find the value of R which makes the maximum magnitude of error in i_b a minimum in the region of interest lying between 0 and 20 volts. The error referred to is i_b (diode) minus i_b (model) at any given e_b.

(d) What is the maximum magnitude of error, expressed as a percentage of the maximum diode current in the region of interest?

3.8. The diode of Problem 3.7 is connected in series with a resistance R_L, and a voltage V_1 is applied across the combination. Let V_2 be the resulting voltage across R_L.
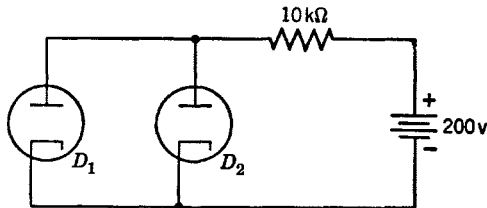
Fig. P3.17

(a) Plot V_2 vs. V_1 for $R_L = 4800$ ohms over the range of V_1 from 0 to 500 volts.

(b) Repeat part (a) with the diode replaced by its approximate model as in Problem 3.7(c).

(c) Does the approximation introduce an appreciable error? Explain.

3.9. Find the operating point for each diode in the circuit shown in Fig. P3.17. The characteristics for D_1 and D_2 are plotted in Figs. P3.18 and P3.19.
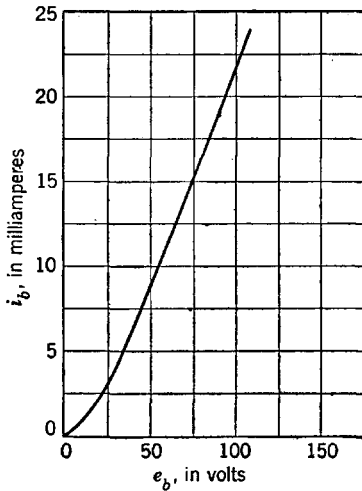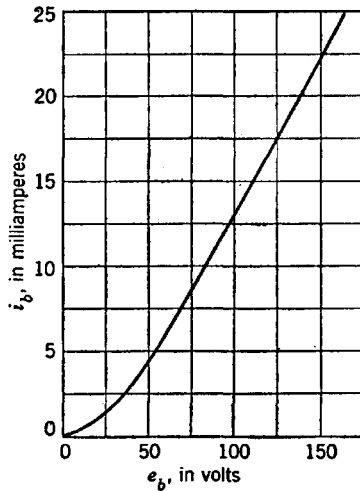

Fig. P3.18   Fig. P3.19

3.10. If $y = x^{3/2}$, the linear approximation $y = 0.9x$ is correct at $x = 0$, roughly 0.1 too high at $x = \frac{1}{3}$, correct at about $x = \frac{2}{3}$, and 0.1 too low at $x = 1$. If a space-charge-limited (i.e., operating somewhere between cut-off and saturation) high-vacuum diode draws 10 ma at 100 volts, what size linear resistor might you use as an approximate model of the diode over the region between 0 and 10 ma?

3.11. A voltage $e_1 = E_1 \cos \omega t$ is applied to the circuit of Fig. P3.20, and a voltage e_2 is measured at the output. The load resistance R is adjustable.
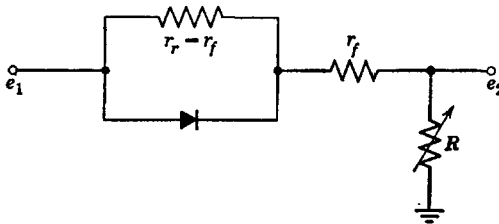
Fig. P3.20

In terms of the forward and reverse diode resistances, r_f and r_r, find the value of the load resistance R which gives the largest d-c component of output

voltage. What is the d-c value in terms of E_1, r_r, and r_f? (e_1 and e_2 are voltages measured with respect to ground.)
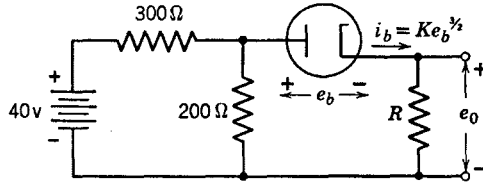

Fig. P3.21

3.12. The high-vacuum diode shown in Fig. P3.21 is assumed to obey the three-halves power law ($i_b = Ke_b^{1.5}$) and $i_b = 50$ ma when $e_b = 20$ volts. Compute e_0 as a function of R.

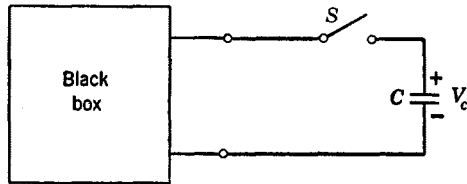**3.13.** A one-farad capacitor charged to 2 volts is suddenly connected to the terminals of a black box (Fig. P3.22). The capacitor voltage decreases lin-

Fig. P3.22

early (not exponentially) to 1 volt in 1 second, after which it continues exponentially toward zero, with no discontinuity in either the voltage or its time derivative. Using sources, resistances, and ideal diodes, synthesize a possible circuit for the black box.
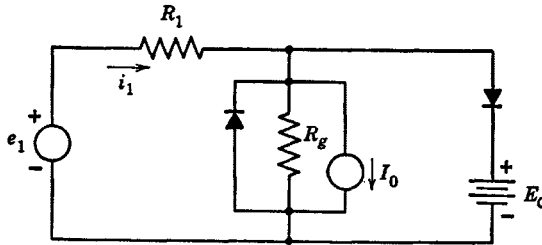
Fig. P3.23

3.14. Plot the driving-point curve (e_1 vs. i_1) for the circuit of Fig. P3.23. Indicate slopes and break points.

3.15. Draw the transfer relations between the variables indicated in Figs. P3.24 through P3.28.

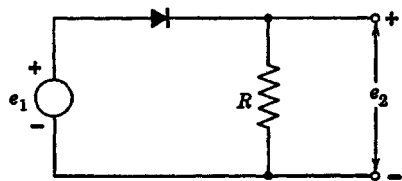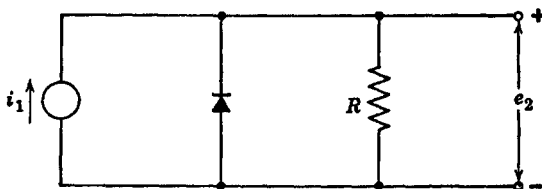**3.16.** Sketch the responses of the ideal circuits shown in Figs. P3.29 and P3.30.
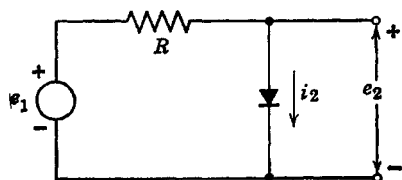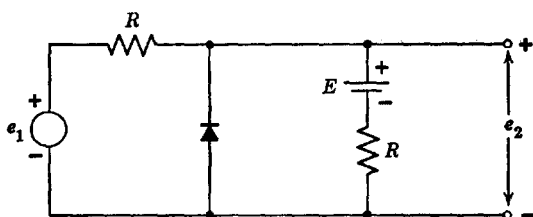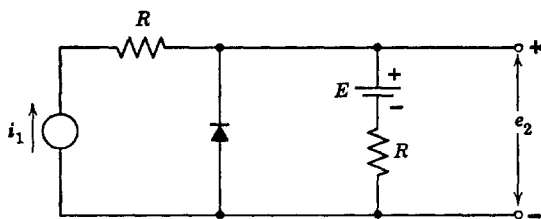
Fig. P3.24

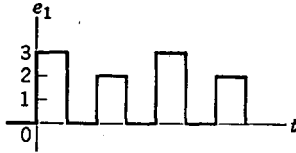Fig. P3.25

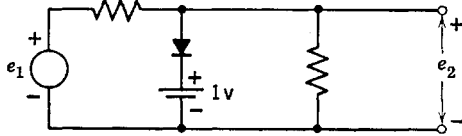Fig. P3.26

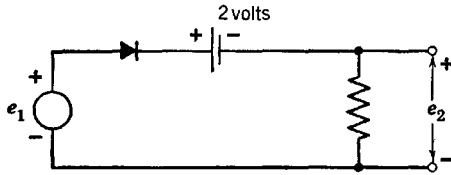

Fig. P3.27

Fig. P3.28

Fig. P3.29

Fig. P3.30

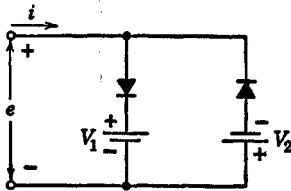**3.17.** Sketch the e vs. i characteristic of the circuits shown in Figs. P3.31–P3.34.
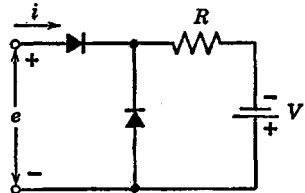

Fig. P3.31

Fig. P3.32
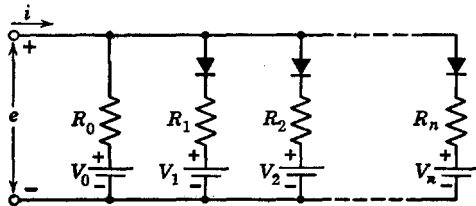
Fig. P3.33
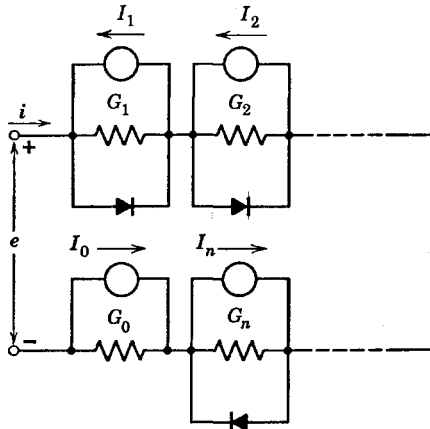
Fig. P3.34

3.18. Sketch v_2 vs. v_1 for the circuit shown in Fig. P3.35.
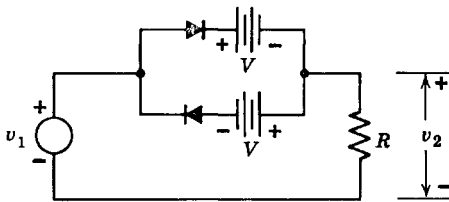
Fig. P3.35

3.19. Sketch and dimension the driving-point curve (v vs. i) for each of the circuits shown in Figs. P3.36–P3.38.
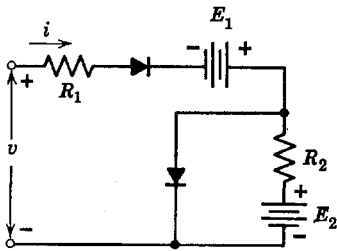
Fig. P3.36
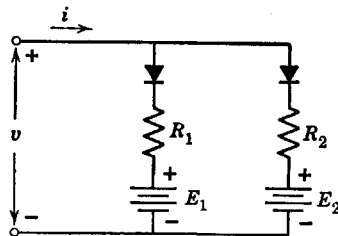
Fig. P3.37

3.20. Plot v vs. i for the circuit shown in Fig. P3.39, showing slopes and the dimensions of break points.

3.21. Sketch the graphs of e vs. i for the circuits shown in Figs. P3.40–P3.43. All components are ideal.

3.22. Plot the e vs. i graph for the circuit shown in Fig. P3.44.
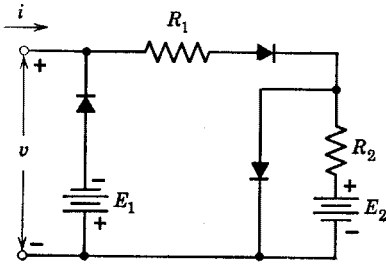
Fig. P3.38
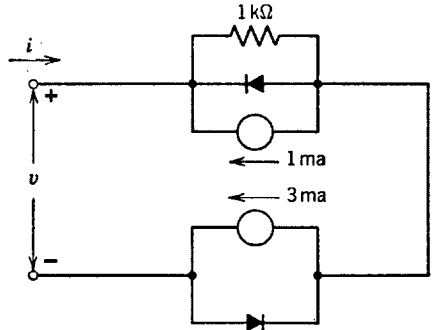
Fig. P3.39
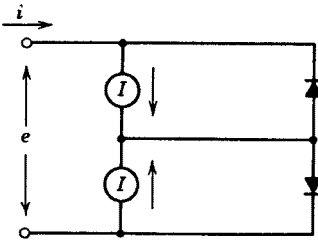
Fig. P3.40
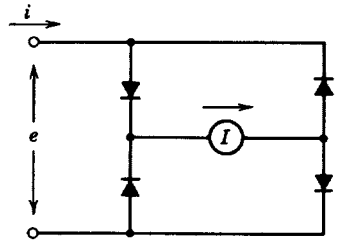
Fig. P3.41
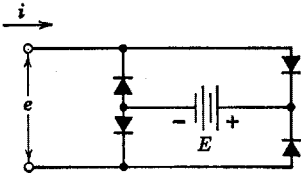
Fig. P3.42
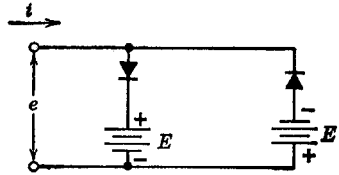
Fig. P3.43
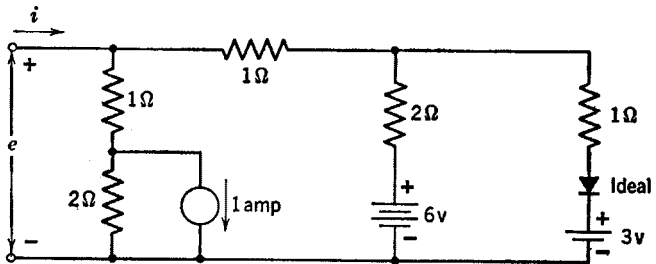
Fig. P3.44

3.23. In the circuit shown in Fig. P3.45, C is initially charged to the voltage E as shown. At time $t = 0$ the switch S is closed. Sketch $e(t)$ for the three cases: $R = 0$; $0 < R < E/I$; and $R > E/I$.
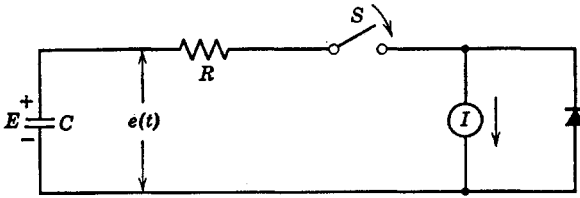
Fig. P3.45

**3.24.** (a) Synthesize a piecewise-linear circuit for the vacuum diode with the current-voltage characteristic (Fig. P3.46): $i = 10^{-3}e^{3/2}$ with the approximation tangent to the curve at the points $e = 0$; 1v; 2v (for $e < 0$, $i = 0$). (Two break points.)
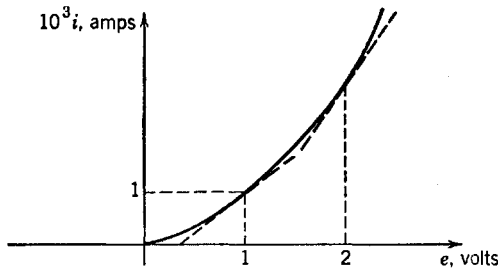
Fig. P3.46

(b) Using the circuit model found in part (a) for the diode in the circuit shown in Fig. P3.47, find the voltage transfer characteristic, e_2 vs. e_1.
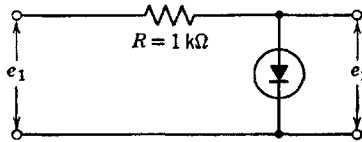
Fig. P3.47

3.25. The resistive-diode circuit shown in Fig. P3.48 has three ideal diodes, two of which have no influence on the driving point characteristic. Redraw the circuit with only one diode, and sketch i vs. e.
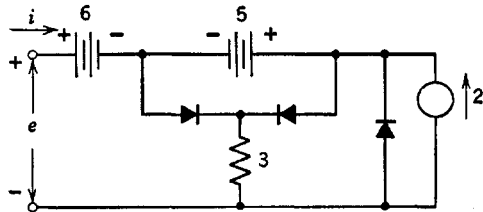
Fig. P3.48

3.26. Sketch and dimension e_2 vs. e_1 curves for the circuits shown in Figs. P3.49–P3.52. In each case e_1 is the applied voltage, e_2 is the measured output voltage, and R_0 is several times larger than R.
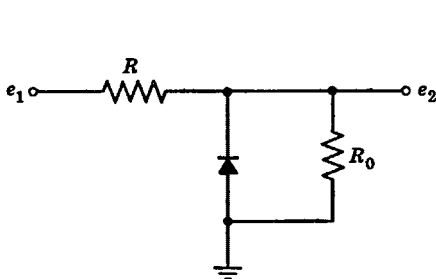
Fig. P3.49
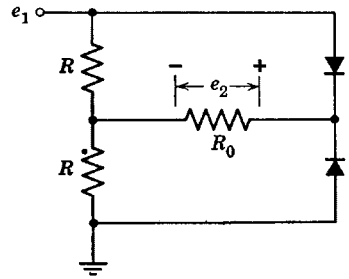
Fig. P3.50
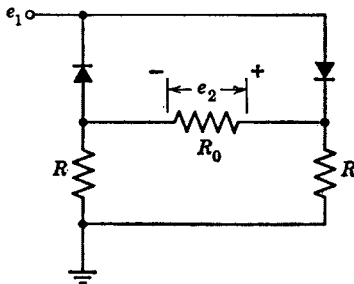
Fig. P3.51
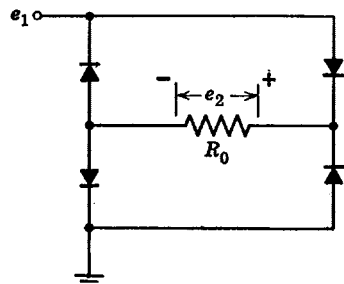
Fig. P3.52

3.27. Find the piecewise-linear e_2 vs. e_1 characteristic of the circuit shown in Fig. P3.53.
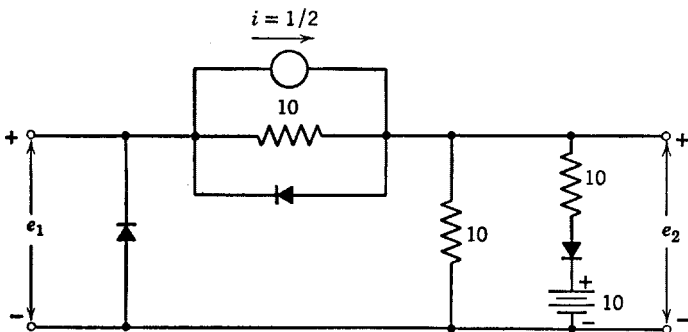
Fig. P3.53

3.28. (a) Find i vs. e for the circuit in Fig. P3.54. Give values of all slopes and break points.

(b) Assuming $|V_2| < |V_1|$ repeat (a) with V_2 reversed.
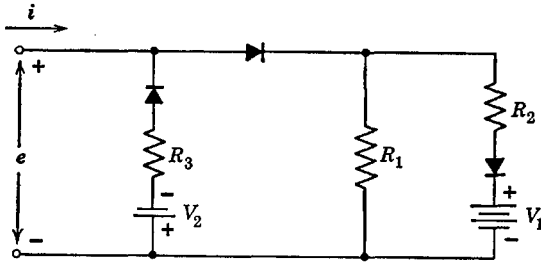
Fig. P3.54

3.29. Both rectifiers in the circuit shown in Fig. P3.55 may be considered ideal. Find i as a function of v, and sketch the result.
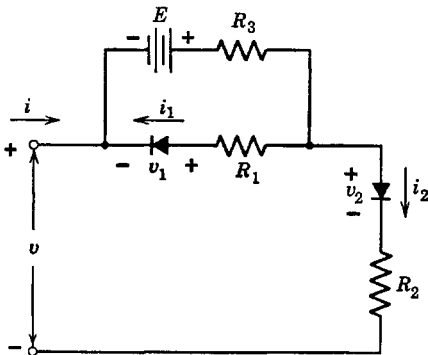
$E = 20$ volts
$R_1 = 10\text{k}\Omega$
$R_2 = 30\text{k}\Omega$
$R_3 = 40\text{k}\Omega$

Fig. P3.55

3.30. Sketch and dimension the current versus voltage curve of the circuit shown in Fig. P3.56.
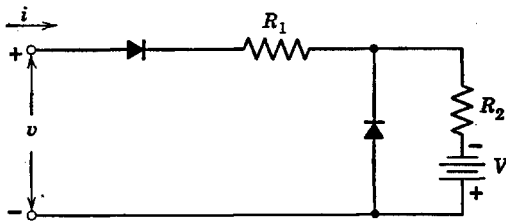
Fig. P3.56

3.31. The general form of the current-voltage characteristics of semiconductor diodes may be represented by a variety of piecewise-linear curves. A number of such curves are shown with numerical values chosen for convenience (Figs. P3.57–P3.59). Draw the piecewise-linear circuits corresponding to each of the above characteristics. Give numerical values for all resistances, and for any voltage or current sources which may be used.
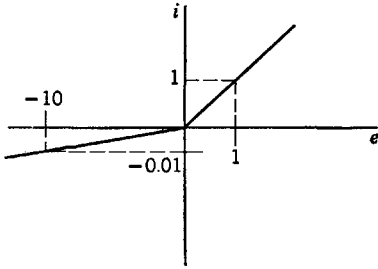
Fig. P3.57
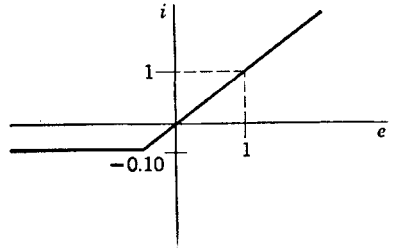
Fig. P3.58
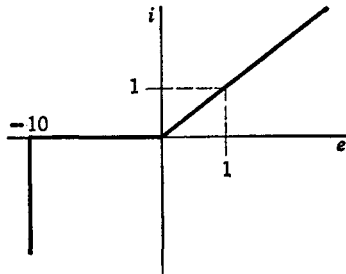
Fig. P3.59

3.32. Determine the e vs. i characteristic of the circuit shown in Fig. P3.60.
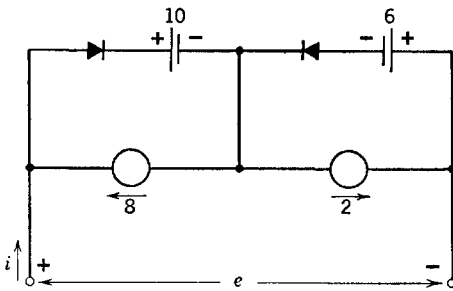
Fig. P3.60
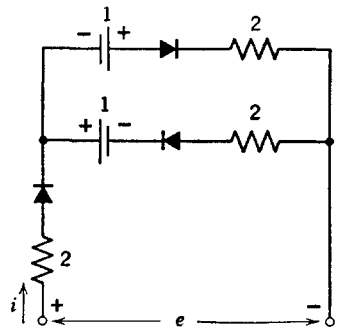
Fig. P3.61

3.33. Determine the e vs. i characteristic of the circuit shown in Fig. P3.61.

3.34. Plot the v vs. i characteristic of the circuit shown in Fig. P3.62 (there are three break points).

3.35. The scheme of Fig. P3.63 is used in a function generator to obtain low frequency, low distortion sine waves. Design a four-diode, four-break point, piecewise-linear network which will give a reasonable approximation to a sine
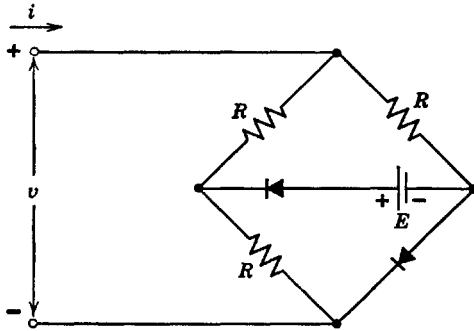
Fig. P3.62
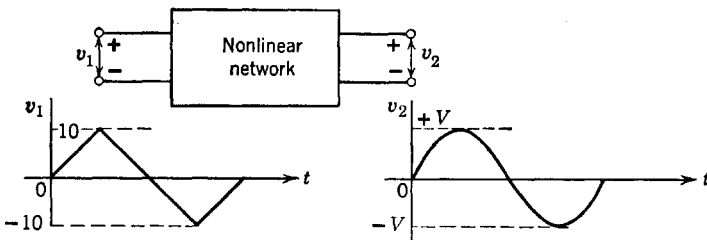
Fig. P3.63

wave for v_2. V should be on the order of 5 volts. The characteristic of the network should look somewhat like Fig. P3.64.
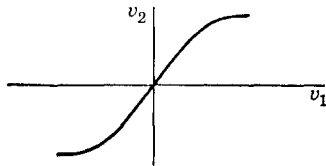
Fig. P3.64

3.36. Two nonlinear functions $y_1 = f_1(x_1)$ and $y_2 = f_2(x_2)$ are tabulated below:

x_1	0	1	2	3	4	6	8	10
y_1	0	2	3.5	4.5	5.25	6	6.25	6.5

x_2	0	2	4	6	7	8	10
y_2	1	2	3.5	5.5	8	10	11

Make a graphical simultaneous solution of these equations for the conditions that

$$y_1 = y_2; \qquad x_1 + x_2 = 10$$

Find the values of x_1 and x_2.

3.37. If in Problem 3.36 x_1 and x_2 are currents in two circuit elements, y_1 and y_2 are the voltages across them, and $f_1(x_1)$ and $f_2(x_2)$ are their volt-ampere characteristics, drawn an equivalent electric circuit.

3.38. If in Problem 3.36 x_1 and x_2 are the elongations of two springs, y_1 and y_2 are the spring forces, and $f_1(x_1)$ and $f_2(x_2)$ are the nonlinear force characteristics of the springs, draw an equivalent mechanical system.

3.39. A type of nonlinear resistor has a volt-ampere characteristic given approximately by $i = Kv^4$. For a certain resistor

$$i = 1.0 \text{ amp for } v = 100 \text{ volts.}$$

This resistor is to be used in a circuit where the operating range will be about 1.0 ± 0.5 amp.

(a) Find a linear incremental model at the 1.0-amp operating point.

(b) Find the percent error in voltage across the resistor at the two extremes of the current range if the linear model of (a) is used.

(c) Find a linear model which would have zero error at the two extremes. Also find the corresponding per cent error at $i = 1.00$ amp.

(d) The nonlinear resistor is now connected in series with a 100-ohm linear resistor and a 200-volt battery. Find the current.

(e) A sinusoidal voltage $e = 25 \sin \omega t$ is now applied in series with the circuit of (d). Use an incremental model to find the a-c component of current.

3.40. Plot i vs. v for the circuit shown in Fig. P3.65.
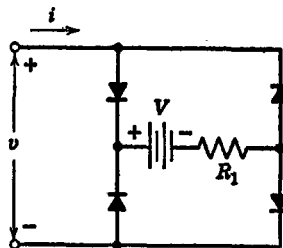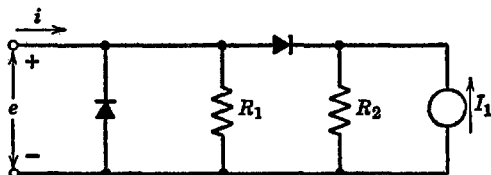
Fig. P3.65   **Fig. P3.66**

3.41. Find the i vs. e curve for the circuit shown in Fig. P3.66.

(a) by assumed state method;

(b) by break-point method.

3.42. It is desired to construct a diode limiter circuit which clips any input waveform in such a way that the magnitude of the voltage of the output waveform never exceeds ± 2 volts. The idealized transfer curve for this device is shown in Fig. P3.67.

Synthesize this curve using ideal diodes and linear elements. What modification of the transfer curve would result if vacuum diodes were used in place of the ideal diodes?

3.43. Devise one or more circuit models to represent the piecewise-linear driving-point curves shown in Figs. P3.68–P3.73. Use ideal rectifiers, resistances, batteries, and (if necessary) current sources.
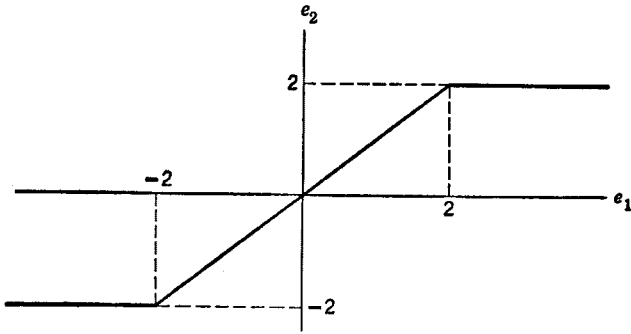
Fig. P3.67
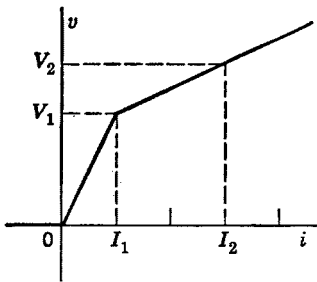
Fig. P3.68
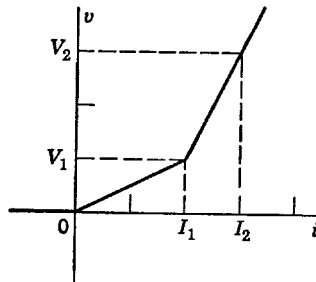
Fig. P3.69
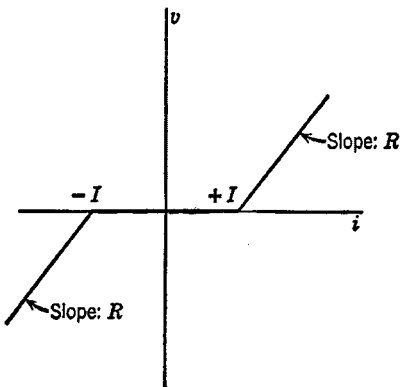
Fig. P3.70
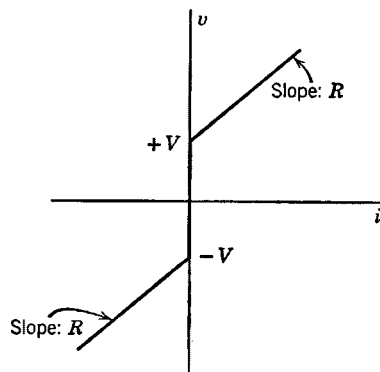
Fig. P3.71
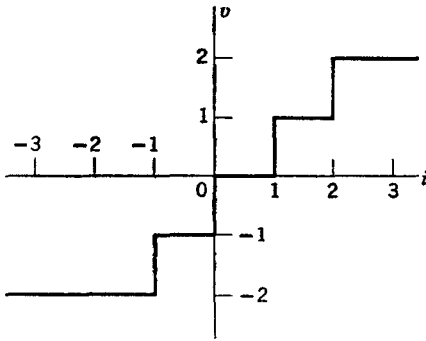
Fig. P3.72
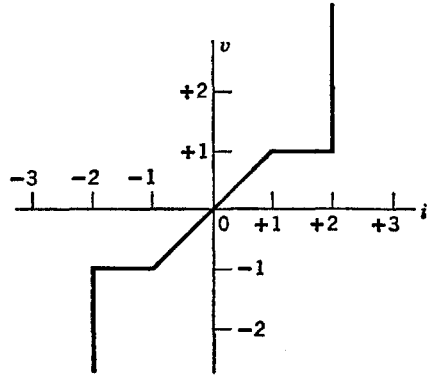
Fig. P3.73

3.44. Using ideal rectifiers, resistors, voltage sources, and (if needed) current sources, design a circuit whose input i vs. v is as shown in Fig. P3.74.
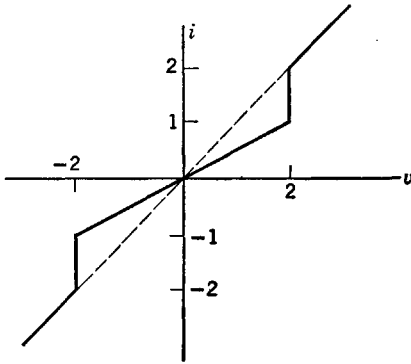
Fig. P3.74
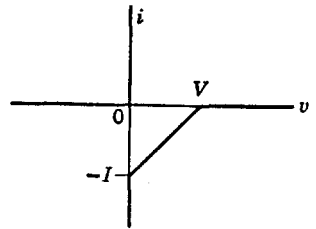
Fig. P3.75

3.45. Using ideal rectifiers, positive resistances, and batteries, design circuits to give the i vs. v curves shown in Figs. P3.75–P3.77. Curve of Fig. P3.77
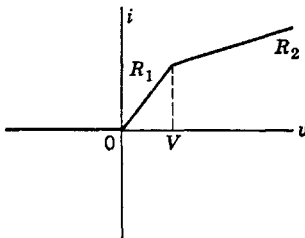
Fig. P3.76
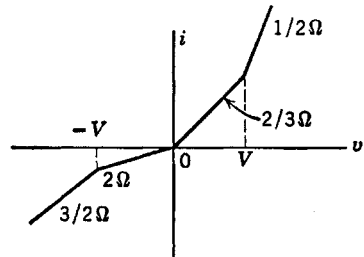
Fig. P3.77

can be realized with only two ideal rectifiers in a bridge-type circuit, or with three ideal rectifiers.

3.46. Write a set of linear equations for the curves in Fig. P3.77 and specify the range of validity of each equation.

3.47. It is desired to design a piecewise-linear circuit that will approximate the characteristic $i = e^2$ over the range $0 < e < 10$ with a maximum current error of ± 1 percent full scale, e.g., $i = (e^2 \pm 1)$ for $0 < e < 10$. What is the smallest number of diodes needed and what circuit do you propose?
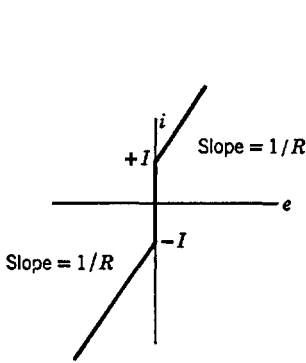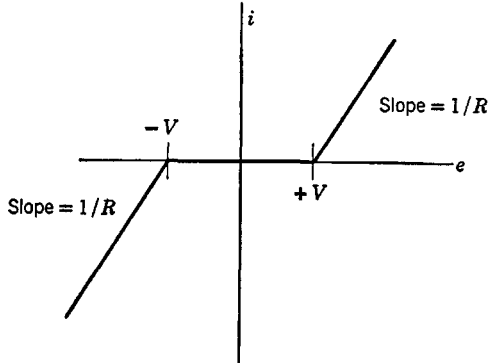
Fig. P3.78   Fig. P3.79

3.48. (a) Using linear circuit elements and ideal diodes, design circuits that will have the terminal volt-ampere relationships shown in Figs. P3.78 and P3.79.

(b) Sketch the terminal volt-ampere relationship that will result if the two circuits devised in (a) are connected in series.

(c) Repeat (b) for the circuits connected in parallel.

3.49. A gas diode is placed in series with a 100 kilohm resistance and the combination is connected to a variable voltage source. The voltage is slowly raised from zero. The circuit draws negligible current until the source voltage reaches 100 volts, at which time the current jumps to nearly 1 ma. The voltage is now slowly reduced, whence the current decreases slowly to the value 0.05 ma and then suddenly jumps back to zero. What information does this experiment give you about the i vs. e curve of the gas diode?
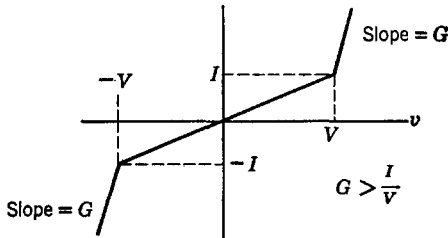
Fig. P3.80

3.50. Using batteries, positive resistances, and ideal rectifiers, design a circuit to give the volt-ampere curve shown in Fig. P3.80.

3.51. Sketch and dimension i_1 vs. e_1 and e_2 vs. e_1 for the circuit shown in Fig. P3.81.
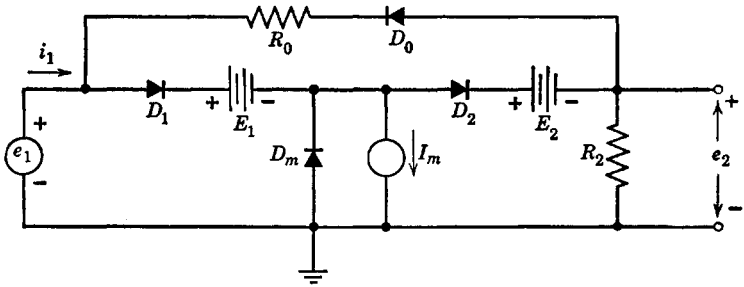
Fig. P3.81

Rectification and Detection

4.1 Introduction

Rectification can be defined rather broadly as the process that modifies the direct component of a voltage or current. Since rectification can occur in any circuit that contains a nonlinear element, this process plays a very important role in electronic circuits. In some cases the effect is desired, as in converting an alternating current to a direct current. In others the effect is undesirable: for example, the distortion of voltage or current waveforms in circuits that are supposedly linear but contain unavoidable nonlinearities. Since nonlinearities will be represented in circuit models by diodes or diode approximations, the circuit concepts and methods of analysis introduced in this chapter (as well as those of the preceding chapter) apply to circuits using triodes, transistors, and other devices.

Rectifier circuits usually include energy-storage elements to smooth the ripples in the output voltage, thus producing a more effective conversion of alternating current to direct current. Such rectifier-filter circuits are used in most electronic systems. They serve to convert a-c voltages from the power distribution system to d-c voltages required for the operation of most electronic devices.

In rectifier circuits used for power conversion, the input a-c voltage amplitude is fixed, but the output power varies with the changing de-

mands of the d-c load. Important design considerations for this case are: the ratio of d-c load power to a-c input power (efficiency), variation of output voltage with load current (regulation), and residual a-c fluctuations present in the output (ripple).

Where the speed of response of the rectifier is important (as it may be in a recording voltmeter), we are concerned with the detection problem. Detection (often called demodulation) is closely related to the rectification process. The amplitude of an a-c voltage may be time-varying, as is the case with the amplitude-modulated carrier of a broad-casting station. The amplitude varies in accordance with the sound level of the speech or music being broadcast. The demodulator or detector in an a-m receiver must produce an output that follows the audio-frequency variations faithfully.
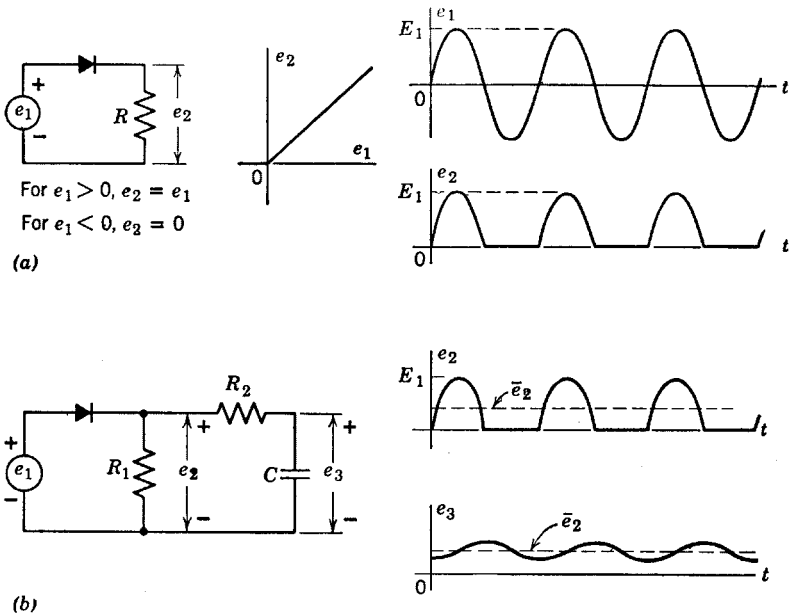
Fig. 4.1. Basic rectifier circuit.

4.2 The Basic Rectifier Circuit

The resistive portion of the basic rectifier circuit, consisting of an ideal diode and a load resistance, is shown in Fig. 4.1(a), together with the voltage transfer curve e_2 vs. e_1 and output waveform e_2 for a sinus-oidal input waveform.

Although the circuit converts a sinusoidal input voltage to a unidirectional voltage consisting of half-sinusoids, the efficiency of conversion from a-c to d-c is poor (the ratio of peak a-c to average d-c is $1/\pi$) and e_2 would hardly be considered a usable d-c voltage. The waveform passes through the average value instantaneously, twice during each cycle, and obviously has a large time-varying component. However, there *is* a direct component or average value. We need only pass the waveform through a low-pass filter [R_2 and C in Fig. 4.1(b)] to enhance the unidirectional character of the output voltage and suppress the time-varying part. This smoothing or filtering may be considered to be a selective separation of alternating and direct components of voltage and current. With $R_2 \gg R_1$ the filter does not appreciably load the rectifier circuit, hence the waveform of voltage e_2 remains nearly unchanged. This condition permits us to analyze the circuit as a cascade of two separate units, the rectifier circuit and the filter circuit.

4.3   Capacitor Smoothing

A single capacitor alone may be connected as in Fig. 4.2 to smooth the output of the basic rectifier circuit, but in this case the capacitor greatly modifies the current and voltage waveforms of the basic circuit, hence the circuit must be analyzed as a unit. The output of the circuit is very nearly a smooth d-c voltage, and the conversion from a-c to d-c has been materially improved by the addition of the capacitor. The residual ripple or fluctuation of the output voltage, which is a function of R, C, and the period T, will be calculated quantitatively, together with the average value or d-c component of the output voltage.

The waveform of output voltage e_2 with C disconnected is shown in the figure as a dotted line. This waveform, determined by the resistive circuit, forms the skeleton or framework for the complete solution. The input voltage is zero for $t < 0$.

The output voltage and capacitor current waveforms, with the capacitor connected to the circuit, are shown as solid lines. The current i_R is nearly constant, since e_2 has very little ripple. The ideal diode conducts during the first quarter-cycle to charge capacitor C to a voltage E_1. As the input voltage falls below E_1, the ideal diode opens. Capacitor C then discharges slowly through the large resistance R to the value designated as E_2. This value is reached at an increment of time δ before the peak of the next cycle. Since for small values of ripple the value of δ is much less than T, we can approximate $(T - \delta)$ by T to determine E_2, as indicated in Fig. 4.2. This approximation

For $RC \gg T$

$E_1\epsilon^{-t/RC} \approx E_1[1 - t/RC]$

And $\delta \ll T$

Hence

$\quad E_2 \approx E_1[1 - T/RC]$

Ripple (peak-to-peak)

$\quad \Delta e_2 = E_1 - E_2$

$\qquad = \dfrac{E_1 T}{RC}$

D-C output $= \overline{e_2} = E_1 - \dfrac{\Delta e_2}{2} = \tfrac{1}{2}(E_1 + E_2) = E_1\left[1 - \dfrac{T}{2RC}\right]$

Rectification efficiency $= \eta = \dfrac{\text{d-c output}}{\text{peak a-c input}} = 1 - \dfrac{T}{2RC}$

Per-unit ripple $= \dfrac{E_1 - E_2}{\overline{e_2}} \approx \dfrac{T}{RC}$

Fundamental ripple frequency $= \omega$
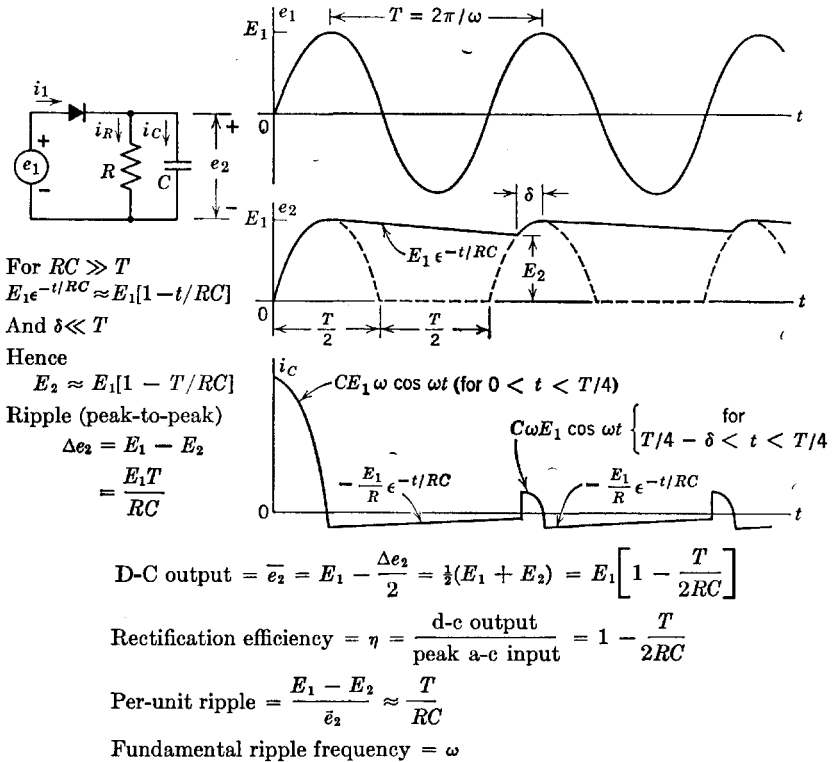
Fig. 4.2. Basic rectifier circuit with smoothing capacitor.

eliminates the need for solving a transcendental equation to find δ, and yields adequate accuracy for most purposes.

If the values of R and C are chosen so that $RC \gg T$, the exponential capacitor-discharge curve can be approximated by a linear relation that matches the initial slope. Using these two approximations, we find that

$$E_2 = E_1\left(1 - \frac{T}{RC}\right) \tag{4.1}$$

The ripple is thus a sawtooth voltage with a peak-to-peak amplitude

$$\Delta e_2 = E_1 \frac{T}{RC} \tag{4.2}$$

The average value of e_2 or the d-c voltage is

$$\overline{e_2} = E_1 - \frac{\Delta e_2}{2} = E_1\left(1 - \frac{T}{2RC}\right) \tag{4.3}$$

and the per-unit ripple is

$$\frac{E_1 - E_2}{e_2} \approx \frac{T}{RC} \tag{4.4}$$

The rectification efficiency (η) may be defined as the d-c output divided by the peak value of the a-c input (E_1). In terms of the input angular frequency ($\omega = 2\pi f$), it is apparent that the condition for small values of ripple and maximum efficiency will be $\omega RC \gg 1$. In the limit the d-c output becomes equal to the peak value of the a-c input, and the circuit becomes a perfect *peak detector*.

An alternative approach to the calculation of the ripple voltage is based on the fact that the output voltage is very nearly constant and approximately equal to E_1, the peak value of the input voltage. The current through R, very nearly constant at E_1/R, removes a total charge

$$\Delta Q \approx \frac{E_1(T\text{-}\delta)}{R} \approx \frac{E_1 T}{R} \tag{4.5}$$

from the capacitor when the diode is off. When the capacitor voltage has reached a steady-state value, the charge added to the capacitor during the interval δ in each cycle must equal the charge lost through R during the remainder of each cycle. In order to add a given amount of charge ΔQ to the capacitor, the voltage must change in accordance with the relation $\Delta Q = C\Delta e_2$. Substituting for ΔQ in Eq. 4.5 leads to the same value of peak-to-peak ripple voltage obtained by the previous method (Eq. 4.2).

The results of Fig. 4.2 are substantially unchanged if a diode with nonzero forward and finite back resistance is substituted for the ideal diode, provided that the forward resistance is small, and the back resistance large, compared with the load resistance R.

Rectifiers intended for power applications cannot always be operated with a very large value of R. With lower values of load resistance, the d-c output voltage drops further from the peak a-c input voltage and ripple voltage increases. Then a more elaborate smoothing circuit must be used in order to maintain low output ripple and good voltage regulation.

4.4   Rectifier Circuit with d-c and a-c Input Voltages

The analysis of nonlinear circuits with both a d-c source and an a-c source simultaneously applied is a problem of general interest. Whereas

the *individual sources* can be considered separately in linear circuit analysis, and the solution obtained by superposition, the *total applied voltage* must be considered (at least initially) in nonlinear or piecewise-linear circuits. The rectifier circuit under consideration will illustrate this fact.
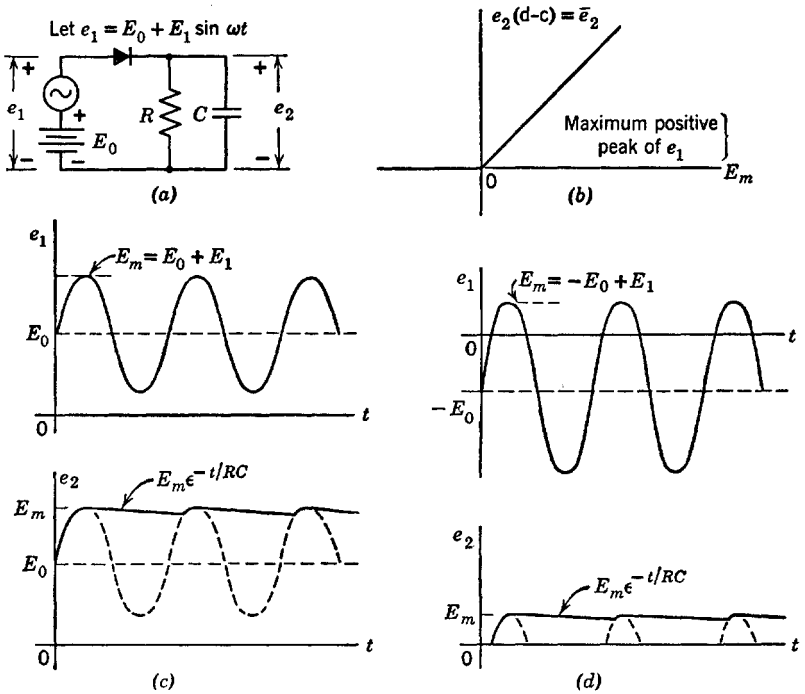
Fig. 4.3. Rectifier circuit with d-c and a-c input.

The circuit shown in Fig. 4.3(a) is identical to the one in Fig. **4.2**, except that the input voltage now has both a d-c and an a-c component. The a-c voltage waveform is again assumed to begin at $t = 0$. From the considerations of the preceding article, the transfer curve of Fig. 4.3(b) can be deduced. The unity slope shown for $\overline{e_2}$ (the d-c component of e_2) vs. the positive peak of e_1 is an idealization which assumes negligibly small ripple, source resistance, and diode forward resistance. The output waveforms for particular input combinations of a-c and d-c components follow directly from this transfer curve. Two examples are shown in Fig. 4.3(c) and (d).

In Fig. 4.3(c), the d-c source E_0 exceeds the magnitude E_1 and is so

polarized that e_1 is never negative. Thus the resistive circuit operates entirely on the unity slope portion of its transfer curve. In the absence of the capacitor, the input sine wave would be reproduced at the output, as shown by the dotted curve in Fig. 4.3(c). With the capacitor connected, the output voltage is nearly constant at E_m. Here E_0 may be assumed to apply a static charge to the capacitor, after which the circuit behaves like the one in Fig. 4.2. The reference level for the sine wave is $+E_0$ instead of zero.

With the polarity of E_0 reversed, the waveforms are as shown in Fig. 4.3(d). The source E_0 biases the sine wave negatively so that only the positive peaks cause diode conduction in the resistive case. Obviously, if E_0 exceeds E_1 in magnitude, the diode will never conduct and the output will be zero. The capacitor smooths the waveform and produces a d-c voltage near the maximum positive peak E_m.

4.5 The Clamping Circuit

Suppose the voltage across the diode of the basic rectifier circuit is considered to be the output of the circuit. The circuit is rearranged as shown in Fig. 4.4(a), to call attention to this fact, but it is still the same circuit. Results obtained previously can therefore be used directly. The waveform of diode voltage was not plotted in the previous figures, but it is readily deduced as the difference between the input voltage and the voltage across load resistance and capacitor. The voltage waveforms for the circuit are shown in Fig. 4.4(a), assuming that the sinusoidal input voltage begins at $t = 0$, and that the capacitor has no initial charge.

The output waveform consists of the input waveform pushed up by the d-c voltage across the capacitor. Since this voltage is approximately equal to the peak value of the input voltage (E_1), the *negative* peak of the output waveform is held or clamped at reference level (zero potential). With an ideal rectifier in the circuit, e_2 cannot be negative by even a small amount. Conduction of the diode immediately charges the capacitor to the value required to maintain the output voltage at zero as e_1 goes to its negative peaks. With an actual diode in the circuit, the output voltage is very slightly negative when the capacitor is being recharged.

Reversing the direction of the diode reverses the polarity of the charge on the capacitor and depresses the output waveform by an amount approximately equal to E_1. In this case, Fig. 4.4(b), the *positive* peak of the output waveform is clamped to the reference potential.

$$e_0 = e_1 + e_2$$

Assume $RC \gg T = 2\pi/\omega$

$e_1 = E_1 \sin \omega t$

(a)

$$e_0 = e_1 + e_2$$

Assume $RC \gg T = 2\pi/\omega$

$e_1 = E_1 \sin \omega t$

(b)
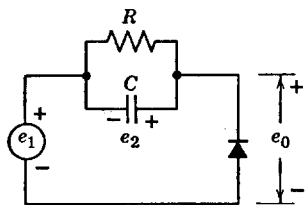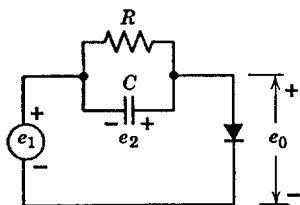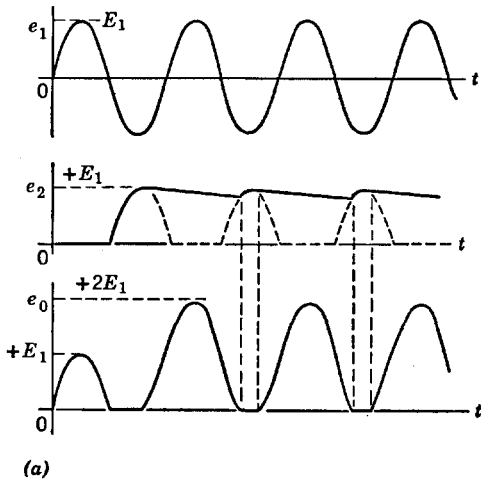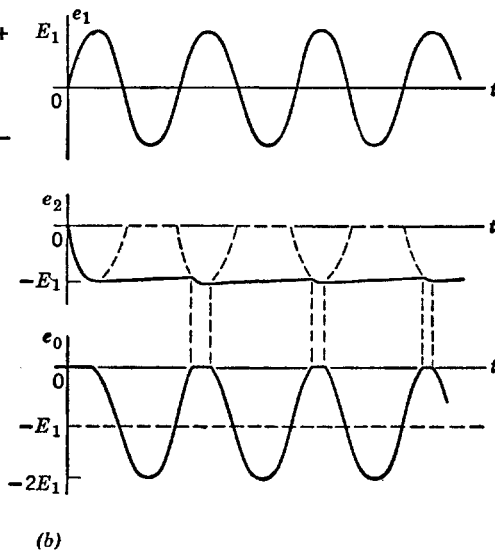
Fig. 4.4. Clamping circuit.

The reference potential can be shifted up or down by putting a d-c source in series with the diode, as shown in Fig. 4.5(a) and (b). The voltage level at which the output waveform is held depends on the

$$e_0 = e_1 + e_2$$
$$e_1 = E_1 \sin \omega t$$
$$RC \gg T = 2\pi/\omega$$

(a)

$$e_0 = e_1 + e_2$$
$$e_1 = E_1 \sin \omega t$$
$$RC \gg T = 2\pi/\omega$$

(b)
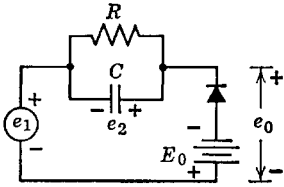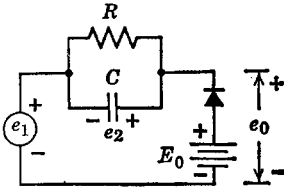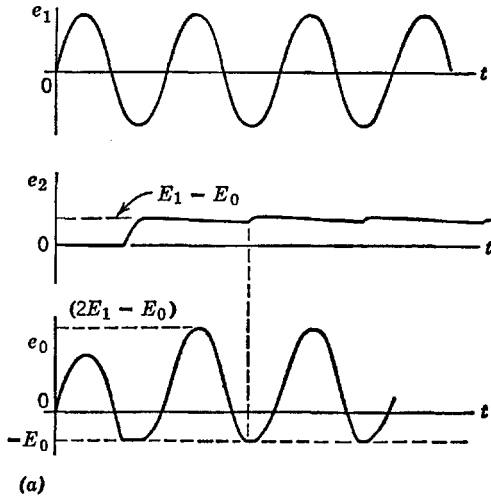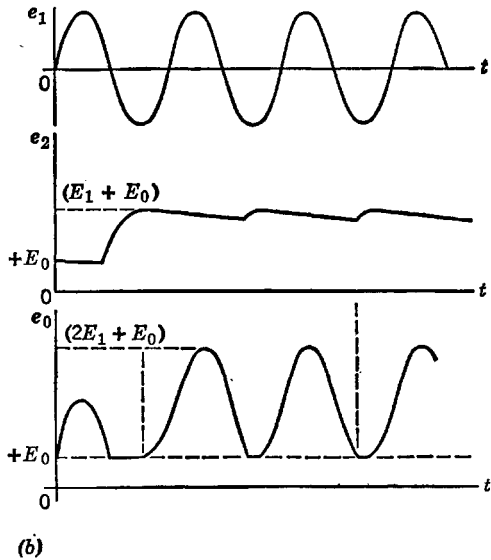
Fig. 4.5. Clamper with d-c source included.

magnitude and polarity of the d-c. However, once equilibrium has been established, the shape of the output waveform is the same as before.

The circuit under discussion is commonly called a clamping circuit or clamper, since it holds one peak of a waveform at a fixed potential.

Such descriptive names are frequently used because they are helpful in discussing the functional behavior of complex electronic systems. However, descriptive names may obscure basic similarities. Recog-
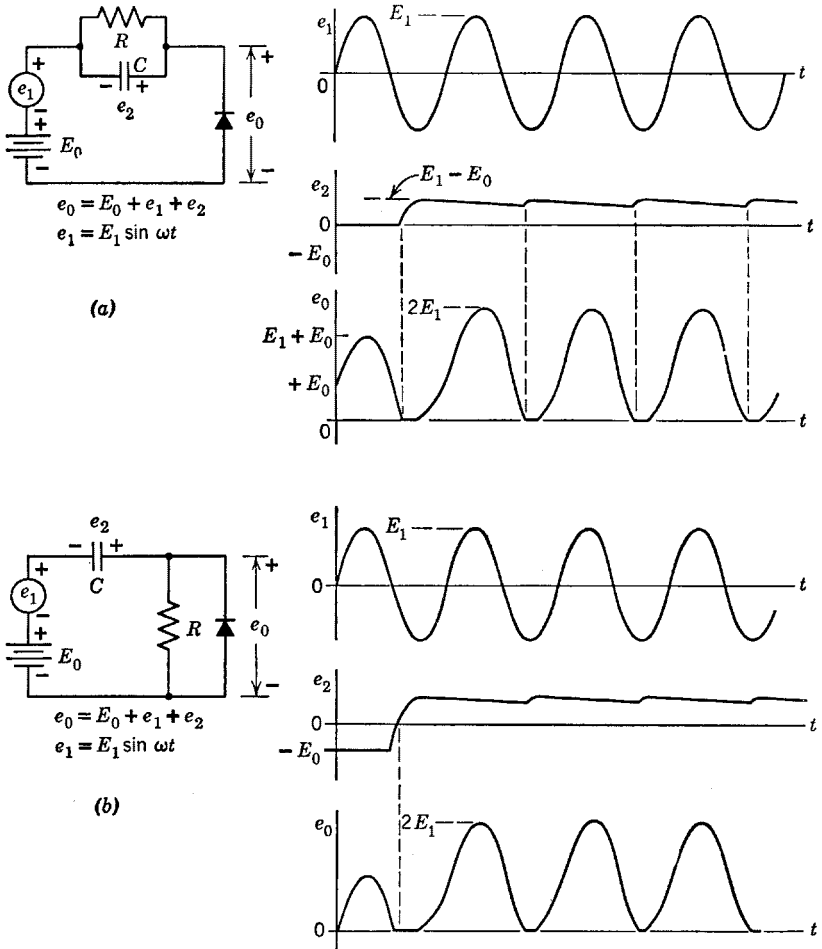
Fig. 4.6. Variations of clamping circuit.

nition of basic similarities between circuits, as in the case of the clamper and the basic rectifier, leads to a better understanding of both circuits and also saves analysis time. To emphasize this point, two other variations of the clamper circuit are shown in Fig. 4.6. The circuit shown in Fig. 4.6(a) has a d-c source in series with the input sinusoid, a

trivial variation of Fig. 4.5(a). Obviously the output is clamped at zero, and the charge on the capacitor takes up the effect of the d-c source. The circuit shown in Fig. 4.6(b) is a-c coupled. Since the capacitor blocks direct current, the d-c source cannot influence the output after a charge has been established.

4.6   The Voltage Doubler

The clamper of Fig. 4.6(b) yields an a-c output voltage that reaches a maximum value very nearly equal to the peak-to-peak value of the a-c input. The basic rectifier circuit, Fig. 4.3, produces a d-c output nearly equal to the maximum peak of the input voltage (a-c plus d-c). If the rectifier circuit is connected to the output of the clamper, a d-c output voltage is obtained which is nearly equal to the peak-to-peak value of the a-c input. Since this is nearly twice the d-c voltage output from a single-diode rectifier and smoothing capacitor, the circuit is called a voltage doubler.

The voltage-doubler circuit shown in Fig. 4.7 is seen to consist of a clamper followed by a basic rectifier circuit with a smoothing capacitor C_2. The resistance R_1 associated with the clamper is usually omitted from the complete circuit, since it tends to reduce the voltage across C_1. The waveforms shown in Fig. 4.7 depict the steady-state operation of the circuit. It is assumed that the rectifier circuit and smoothing capacitor C_2 constitute insufficient load to affect the operation of the clamper circuit. Thus voltage e_{01} is assumed to be a sine wave of peak amplitude E_1 added to a d-c component of value E_1, as in Fig. 4.6(b). The rectifier circuit, with e_{01} as an input, gives a nearly constant d-c voltage e_2, with a per-unit ripple equal to T/R_2C_2.

If the load resistance R_2 in Fig. 4.7 corresponds to the high resistance of an ordinary d-c voltmeter,* the combination constitutes an a-c voltmeter which reads the peak-to-peak value of the a-c input wave directly and linearly in terms of the d-c output voltage.

One terminal of the input source e_1 in Fig. 4.7 is common to one terminal of the output voltage e_0. Such a common terminal or node is usually very desirable in a circuit, since it permits the establishment of a common reference potential.

By way of contrast, a variation of the voltage-doubler circuit is shown in Fig. 4.8. In this case there is no terminal common to the a-c input

* For example, a d-c milliammeter (0 to 1 ma) with a series resistance of 1000 ohms per volt of d-c voltage to be measured; or a 0 to 50 d-c microammeter in series with 20,000 ohms per volt.
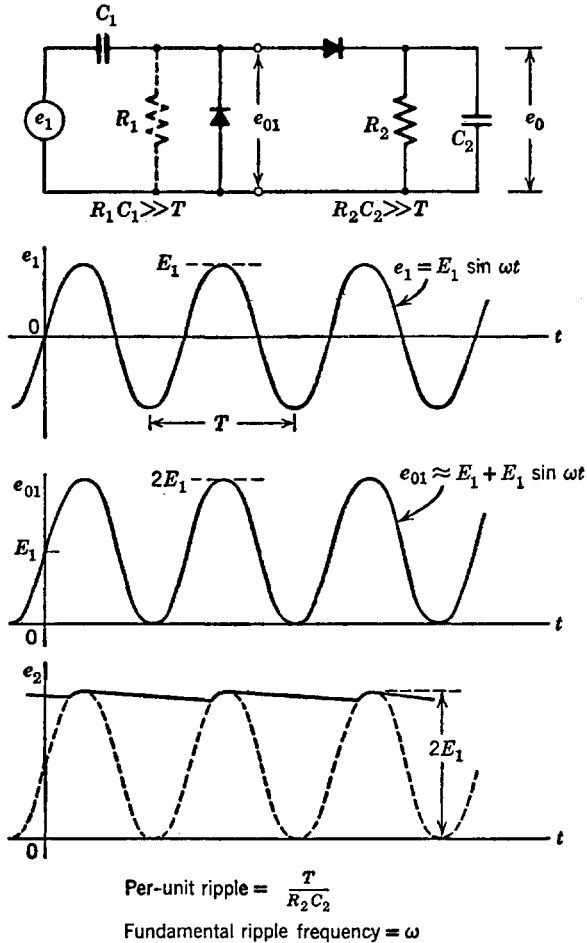
Fig. 4.7. Voltage doubler.

source and the d-c output voltage. The circuit of Fig. 4.8 has a symmetrical appearance, and if $C_a = C_b$, the usual case, the circuit is also electrically symmetrical. The d-c potential difference is the same between input terminal (B) and either output terminal. Since one diode conducts only on the positive peak of the a-c voltage, and the other conducts on the negative peak, input terminal (A) is alternately "connected" to the positive and negative output terminals. Thus e_0 has a magnitude approaching the peak-to-peak value of the input sine

wave, but the output voltage relative to the input terminal (B) extends from $+E_1$ to $-E_1$.

Let $C_a = C_b = C$

$$R_0 \frac{C_a C_b}{C_a + C_b} = \frac{R_0 C}{2} \gg T$$

$$e_0 = e_b + e_a$$
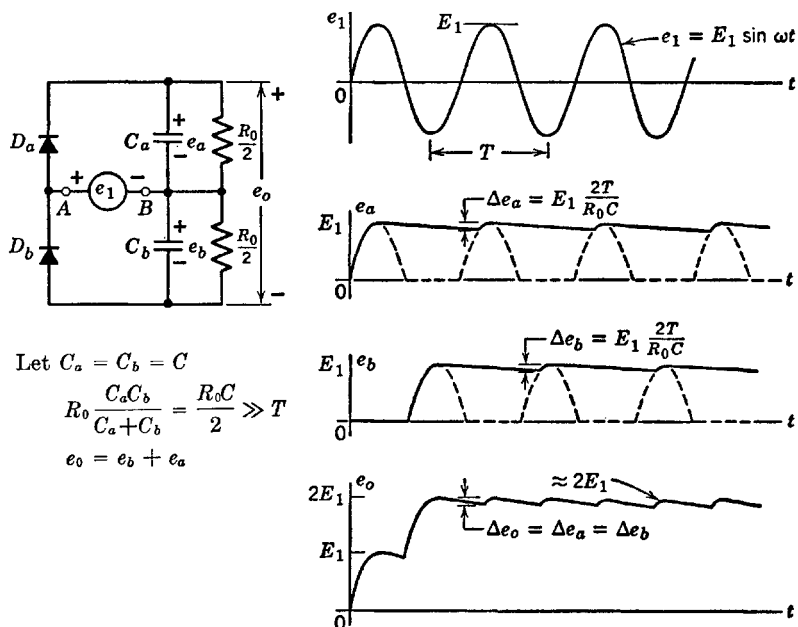
Per-unit ripple $= \dfrac{2E_1 \dfrac{T}{R_0 C}}{2E_1} = \dfrac{T}{R_0 C}$

Fundamental ripple frequency $= 2\omega$

Fig. 4.8. Balanced voltage doubler.

The circuit consists of two basic rectifiers connected back to back. The sum of waveforms e_a and e_b yields the output waveform e_0. Because e_a and e_b are "out of step" by half a cycle, the output ripple is the same as the ripple on either e_a or e_b. The per-unit ripple is

$$\frac{2E_1 T/R_0 C}{2E_1} = \frac{T}{R_0 C} \tag{4.6}$$

identical to that obtained for the unbalanced doubler. However, the fundamental ripple frequency is twice the input frequency in this case, whereas in the circuit of Fig. 4.7 the ripple frequency is equal to the input frequency.
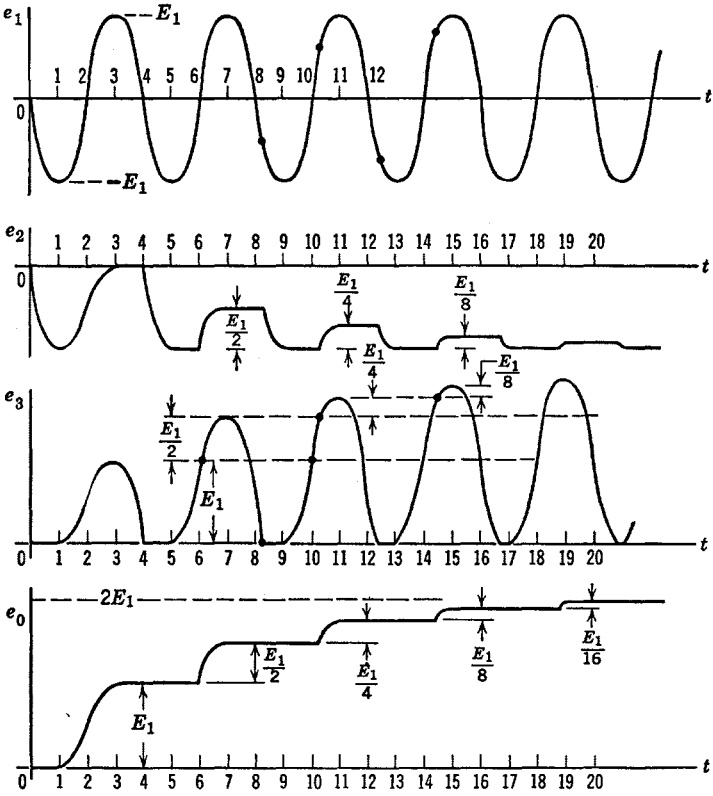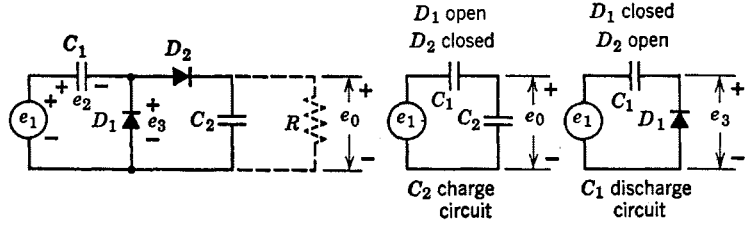
Fig. 4.9. Voltage doubler build-up.

4.7 Voltage-Doubler Build-Up

As indicated by the waveforms of Fig. 4.8, the transient build-up of voltage at the output of the balanced voltage doubler is very simple. The first positive half-cycle of e_1 applied to the circuit charges capacitor C_a to a voltage of about E_1 volts. The first negative half-cycle does the

same for C_b; after that the circuit is in a steady-state condition. Only the small charge lost through the load resistances must be supplied by successive cycles.

For the unbalanced voltage-doubler circuit of Fig. 4.7, the transient build-up proceeds more or less stepwise toward the final value, as indicated in Fig. 4.9. Initially, let us assume the load resistance to be so large that the current through it can be neglected. Two of the four possible diode states are shown in the figure. A third state, that with both diodes non-conducting, occurs twice every cycle. The state with both diodes closed does not occur. With the input waveform shown, the circuit alternates between the two states for which the simplified circuits apply. The waveforms are drawn for the special case of $C_1 = C_2$. As is often the case for nonlinear circuits, the build-up transient of e_0 is simpler to analyze if a square wave is applied to the circuit instead of a sine wave.

4.8   Step-Charging Circuit with Rectangular-Wave Input

Suppose the input voltage e_1 of the voltage doubler consists of a d-c component E_a and a rectangular waveform e_b having a peak-to-peak value E_b and an average value $\bar{e}_b$ equal to $E_{bb}\delta_1/(\delta_1 + \delta_2)$ as shown in Fig. 4.10. The total direct component of input voltage e_1, the sum of the d-c source voltage e_a and $\bar{e}_b$, influences the d-c component of waveform e_2 but has no effect on e_3 or e_0. For the waveforms shown, e_b is assumed to start from a positive peak, with a voltage $E_a + E_b$ on C_1. The waveforms are again drawn for $C_1 = C_2$, but are labeled for the general case.

The sequence of events (diodes opening and closing) proceeds in the same manner as with the sine wave input. On the negative-going steps of the input waveform, diode D_1 closes and capacitance C_1 is charged to the voltage E_a. On the positive-going steps of e_1, diode D_2 closes, D_1 opens, and a portion of the charge on C_1 is placed on C_2. With ideal rectangular waves, the events on each rise or fall occur at the same time. Since the transitions are instantaneous, the capacitors are charged or discharged instantaneously in the absence of resistance. This requires an impulse of current (an infinite amplitude, zero duration pulse). The effective area of the impulse corresponds to the charge $\Delta Q = C\Delta E$ added to or subtracted from the effective capacitance C by a voltage change ΔE.

The waveforms of Fig. 4.10 show considerable similarity to those of Fig. 4.9, aside from the effects of the d-c component. The time of open-

ing and closing of diodes, in this case, will coincide with the switching instants of the rectangular wave. As in the case here, the square-wave response of a circuit can often be determined and sketched more easily

Let $C_1/(C_1 + C_2) = k$

Then for the nth pulse

$(\Delta e_0)_n = E_b(1 - k)^{n-1}k.$

After the nth pulse

$e_0 = \Sigma (\Delta e_0)_n = E_b \left[1 - (1 - k)^n \right]$
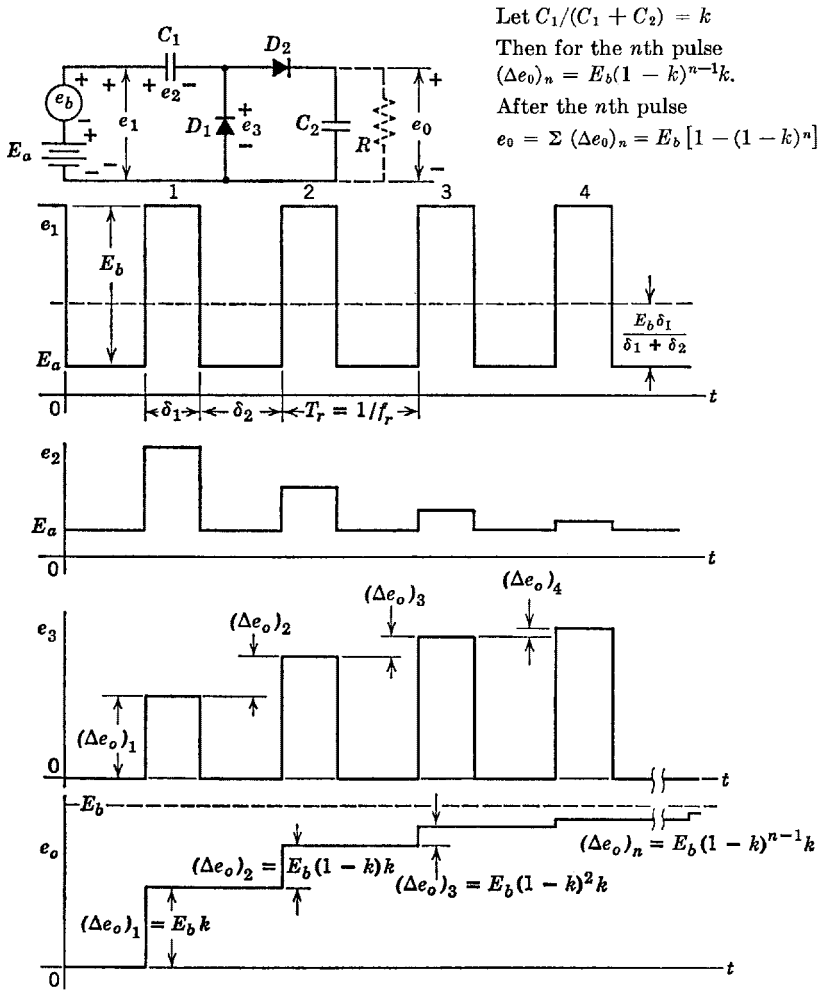
Fig. 4.10. Step-charging circuit with rectangular input waveform.

than the sine-wave response. In such cases it is not amiss to consider the square wave as a rough approximation to a sine wave. If a better straight-line or piecewise-linear approximation to a sine wave is desired, a trapezoidal waveform can be used. The trapezoidal waveform sepa-

rates diode events in time, and also results in constant charge or discharge current, since de_1/dt is constant during the rise or fall.

4.9   An Electronic Frequency Meter

In discussing the voltage doubler, the use of the circuit as a voltmeter to read peak-to-peak a-c voltages was suggested. This application requires a very high value of load resistance to approach the true reading. The capacitors C_1 and C_2 were logically, but not necessarily, equal. In the step-charging application, a ratio C_1/C_2 less than one yields smaller step increments and therefore a slower rise. Ratios of the order of 0.1 to 0.5 are common in practical circuits.

Let us now consider the same circuit, with C_2 very large (C_1/C_2 small) and with the output load resistance small. Let this resistance include the resistance of a 0–1 ma d-c milliammeter, as shown in Fig. 4.11. Since there can be no average current through C_2, the milliammeter registers the average value of charging current supplied to C_2 through diode D_2 and capacitor C_1.

If the discharge of C_1 goes to completion during the time the input voltage is at a minimum, then virtually the entire peak-to-peak change of voltage e_1 occurs across C_1. Actually, the change Δe_2 equals $E_b C_1/(C_1 + C_2)$ which is somewhat less than one per cent of E_b if $C_2 \geqslant 100 C_1$. Thus each positive-going change increases the charge on C_1 and C_2 by a fixed amount

$$\Delta Q = E_b[C_1 C_2/(C_1 + C_2)] \approx E_b C_1 \qquad (4.7)$$

This amount of charge is added once during each cycle of the input wave. Thus the average current through the meter (assuming $e_0 \ll E_b$) is $I = \Delta Q/\tau_r = \Delta Q f_r$; hence $f_r \approx I/E_b C_1$. So long as ΔQ remains constant, I is proportional to f_r, and the instrument can be calibrated to read f_r directly. The charge ΔQ per cycle is constant so long as peak-to-peak voltage is constant; hence the circuit is often preceded by an amplitude limiter. The reading is independent of waveform, provided the charge or discharge of C_1 has time to reach completion each time the voltage swings positive or negative. It is also assumed that the meter reading is not affected by the pulsations of the current. The current could be further smoothed by another large capacitor connected directly across the milliammeter.

Deviation from linearity in the relation f_r vs. I may be expected for very low values of I. Contact potentials, increase in actual diode resistances, or a variety of stray effects influence the behavior near zero

current. At the upper end of the scale, the curve will deviate from linearity if output voltage e_0 becomes comparable to Δe_1.
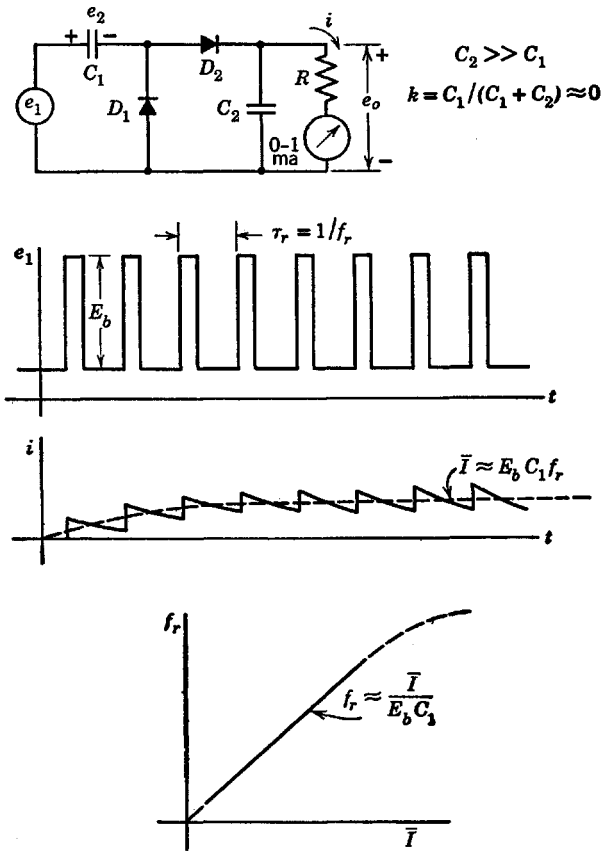
Fig. 4.11. An electronic frequency meter.

4.10 Full-Wave Rectifier with Smoothing Capacitor

The analysis of the full-wave rectifier circuit poses no new problems, but this class of circuits is sufficiently important to justify consideration. The transfer curves for several variations of the full-wave rectifier circuit were presented in Chapter 3. The addition of a smoothing capacitor to the circuit yields results similar to those obtained for the basic rectifier circuit. With a given capacitor, load resistance, and input frequency, the ripple is smaller for the full-wave rectifier, since each half-cycle

recharges the smoothing capacitor. The capacitor discharge involves an interval approximating $T/2$, compared with T for the half-wave rectifier.

For each circuit: $e_2 = E_2 \sin \omega t, \ \omega R_0 C \gg 1$
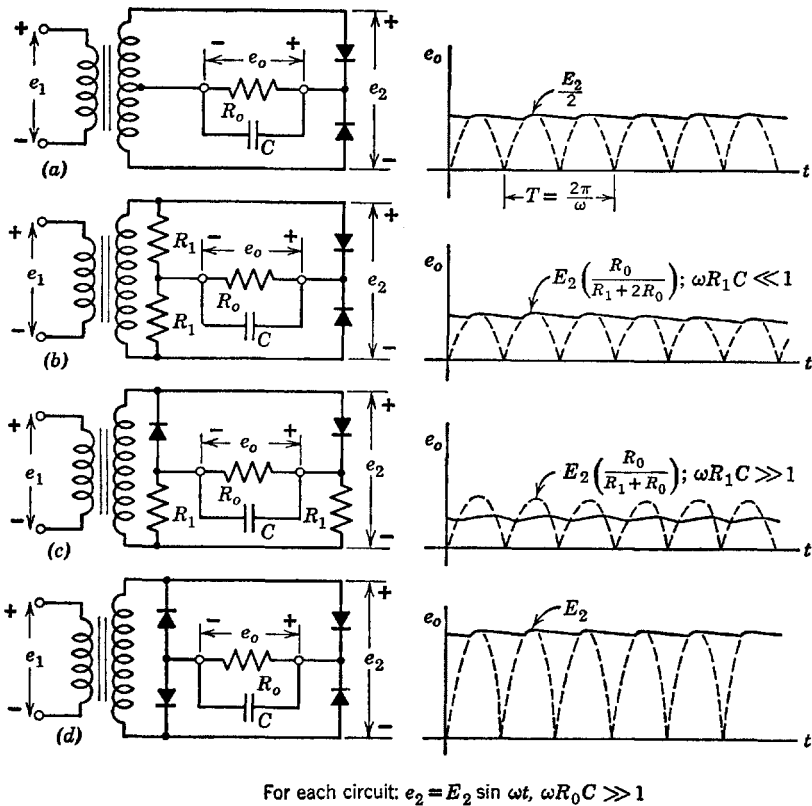
Fig. 4.12. Full-wave rectifier circuits with smoothing capacitor.

A very common form of the circuit is shown in Fig. 4.12(a). Here, half of the voltage available from the transformer secondary is applied to each half of the circuit. If the transformer center tap is not used, or if the source has no center tap available, the circuit may take the form shown in Fig. 4.12(b) or (c). These two forms are not well suited to power-rectifier applications because of the power lost in the resistors R_1. The circuit of Fig. 4.12(c) produces a slightly larger d-c voltage than the circuit of (b). When no center tap is available, the bridge circuit shown in Fig. 4.12(d) is preferable for power-rectifier applications. Of the circuits (b), (c), and (d), this one is unique in providing an output

that approaches the peak of the a-c supply voltage without wasting power in resistances (R_1).

4.11   Ripple Filter with Capacitance Input

A single smoothing capacitor used in conjunction with a rectifier is usually adequate for low-current loads such as those encountered in instrument applications, but when a d-c source is required to supply larger currents, additional filtering may be necessary. Typical examples are the power supplies that provide the d-c operating voltages for vacuum-tube electrodes in electronic systems. Voltages of 100 to 300 volts and currents of the order of 100 ma or more are common for low-power triode and pentode circuits. The effects of a poorly filtered power supply on an audio system can be heard directly in the form of an objectionable hum. The effects of ripple on the behavior of other types of circuits can be equally serious.

When ripple reduction is important, a full-wave rectifier — such as the one shown in Fig. 4.12(a) — is the logical choice. Consider the rectifier and filter circuit shown in Fig. 4.13, where an inductor and a capacitor have been added between the original smoothing capacitor and the resistive load. With these energy-storage elements in the circuit, a general solution is lengthy. However, in the design of a power-supply ripple filter we are obviously interested in making the ripple output very small compared with the d-c output. Using this as an assumption, we can reduce the problem to much simpler proportions.

It is convenient to resolve each current and voltage into a steady component (d-c) and a ripple component (a-c), as indicated in Fig. 4.13. This is possible for any periodic time function, regardless of the relative amplitudes of the d-c and a-c components.

We can establish the d-c condition by inspection. Because the capacitors cannot pass d-c

$$\overline{i_1} = \overline{i_2} = \overline{e_3}/R \qquad (4.8)$$

If we assume that the inductor L has zero resistance, then there can be no d-c drop across it, and

$$\overline{e_2} = \overline{e_3} \qquad (4.9)$$

The ripple waveforms can be readily derived if the assumption of small percentage ripple is carefully exploited. First, C_1 must be large enough to make e_{2r} much smaller than E_1, in which case the current i_1 will consist of brief pulses occurring every half-period. Since there are

Let $\quad i_1(t) = \bar{i}_1 + i_{1r}(t)$
$\quad\quad i_2(t) = \bar{i}_2 + i_{2r}(t)$
$\quad\quad i_3(t) = \bar{i}_3 + i_{3r}(t)$
$\quad\quad e_2(t) = \bar{e}_2 + e_{2r}(t)$
$\quad\quad e_3(t) = \bar{e}_3 + e_{3r}(t)$

Where, for example,

$$\bar{i}_1 = \frac{1}{T}\int_0^T i_1(t)\,dt\,,$$

and $\quad i_{1r}(t) = i_1(t) - \bar{i}_1\,,$

hence $\quad \dfrac{1}{T}\displaystyle\int_0^T i_{1r}(t)\,dt = 0$

Assume: $\quad \delta \approx 0$
$\quad\quad\quad E_1 \gg e_{2r} \gg e_{3r}$
$\quad\quad\quad i_{1r} \gg i_{2r} \gg i_{3r}$

Note:

$$\bar{i}_1 = \bar{i}_2 = \bar{i}_3 = \bar{e}_3/R$$

$$\bar{e}_3 = \bar{e}_2 = E_1 - \frac{T\bar{i}_1}{4C_1}$$

$e_1 = E_1 \sin \omega t$

$e_{2r} \approx \dfrac{1}{C_1}\int i_{1r}\,dt + K_1 \qquad \dfrac{Q_1}{C_1} = \dfrac{T\bar{i}_1}{2C_1} \approx \dfrac{E_1 T}{2RC_1}$

$i_{2r} \approx \dfrac{1}{L}\int e_{2r}\,dt + K_2$

$e_{3r} = Ri_{3r} \approx \dfrac{1}{C_2}\int i_{2r}\,dt + K_3$
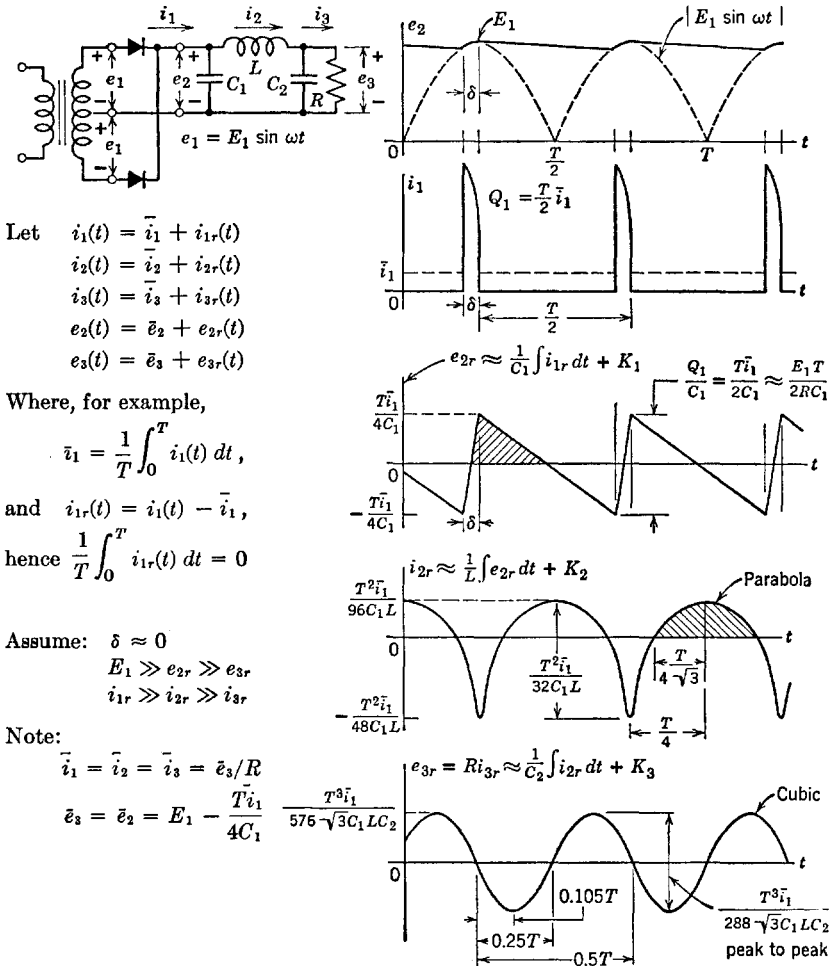
Fig. 4.13. Ripple filter with capacitance input.

two pulses in the period T the charge conducted during each pulse is related to the average current as follows:

$$Q_1 = \frac{T}{2}\bar{i}_1 \tag{4.10}$$

The amplitude of the equivalent rectangular current pulse of duration δ is

$$I_1 = \frac{T}{2\delta}\bar{i}_1 \tag{4.11}$$

Three additional closely related assumptions are necessary to complete the solution, all based on the assumption of adequate filtering. The ripple current i_{2r} must be much smaller than i_{1r}, so all of i_{1r} can be considered to flow through C_1. The ripple voltage e_{3r} must be much smaller than e_{2r}, so all of voltage e_{2r} can be considered to be across L. The ripple current through the load R must be much smaller than i_{2r}, so all of i_{2r} flows through C_2. Integration of the current waveform i_{1r} then gives the ripple voltage across C_1:

$$e_{2r} = \frac{1}{C_1} \int i_{1r}\, dt + K_1 \qquad (4.12)$$

The constant K_1 is adjusted to give e_{2r} zero mean value. Successive integrations yield i_{2r} and e_{3r} thus

$$i_{2r} = \frac{1}{L} \int e_{2r}\, dt + K_2 \qquad (4.13)$$

$$e_{3r} = \frac{1}{C_2} \int i_{2r}\, dt + K_3 \qquad (4.14)$$

Note that the peak value of e_{3r}, the output ripple, is directly proportional to the load current for this circuit.

The per-unit ripple (peak-to-peak) of the output voltage is

$$\frac{e_{3r}}{\bar{e}_3} = \frac{T^3}{288\sqrt{3}\,C_1 L C_2 R} \qquad (4.15)$$

or

$$\frac{e_{3r}}{\bar{e}_3} = \frac{0.5 \left(\dfrac{1}{\omega C_1}\right)\left(\dfrac{1}{\omega C_2}\right)}{\omega L R} \qquad (4.16)$$

where ω is the angular frequency of e_1.

Equation 4.15 shows that increasing any of the elements C_1, L, or C_2 will decrease per-unit ripple. These three elements must be chosen to yield the desired value of ripple. Since an increase in power-supply frequency eases the filtering problem, portable, mobile, or airborne electronic systems (which are not bound to a 60-cps a-c supply frequency) usually make use of much higher primary a-c supply frequencies, for example 400 cps or 1200 cps.

Note that because of the successive integrations, the final ripple waveform is very nearly sinusoidal. If the peaks occurred at $0.125T$ instead of $0.105T$ from the zero crossings, the wave would approach a sine wave. In any event, e_{3r} is close enough to sinusoidal to assume it to be a sine wave if further filter elements are to be added.

i_2 i_3

L C R

e_1 e_2 e_3

$e_1 = E_1 \sin \omega t$

$\omega = 2\pi / T$

$|E_1 \sin \omega t|$

E_m

E_2

e_{2r}

$\dfrac{2}{\pi} E_m$

$\dfrac{T}{2}$ T t

e_2

i_2

$\bar{i_2}$

$T/4$ $0.39T$ t

Assume: $i_2 > 0$

$\quad\quad i_{2r} \gg i_{3r}$

$\quad\quad e_{2r} \gg e_{3r}$

Note: $\bar{i_2} = \bar{i_3}$

$R\bar{i_3} = \bar{e_3} = \bar{e_2} = \dfrac{2}{\pi} E_1$

For $0 < t < T/2$,

$i_{2r} = \dfrac{E_1}{\omega L} \int \left(\sin (\omega t) - \dfrac{2}{\pi} \right) d\,(\omega t) + k_r$

$\quad = \dfrac{E_1}{\omega L} \left[-\cos \omega t - \dfrac{2\omega t}{\pi} + 1 \right]$

$i_{2r\,(\text{max})} = i_{2r}\,(t = 0.39T)$

$\quad\quad = 0.21\, E_1 / \omega L$

Assuming i_{2r} sinusoidal

$e_{3r}(\text{max}) = \dfrac{i_{2r}(\text{max})}{2\omega C} = \dfrac{0.105 E_1}{\omega^2 LC}$

$i_{2r} \approx \dfrac{1}{L} \int e_{2r}\, dt + k_1$

$i_{2r\,(\text{max})} = 0.21 E_1 / \omega L$

$i_{2r\,(a)} \approx -\dfrac{E_1}{\omega L} \cos \omega t$

$i_{2r\,(b)} \approx \dfrac{E_1}{\omega L} \left(1 - \dfrac{2\omega t}{\pi} \right)$

$e_{3r} = i_{3r} R \approx \dfrac{1}{C} \int i_{2r}\, dt + k_2$

$e_{3r\,(\text{max})} = \dfrac{0.105 E_1}{\omega^2 LC}$
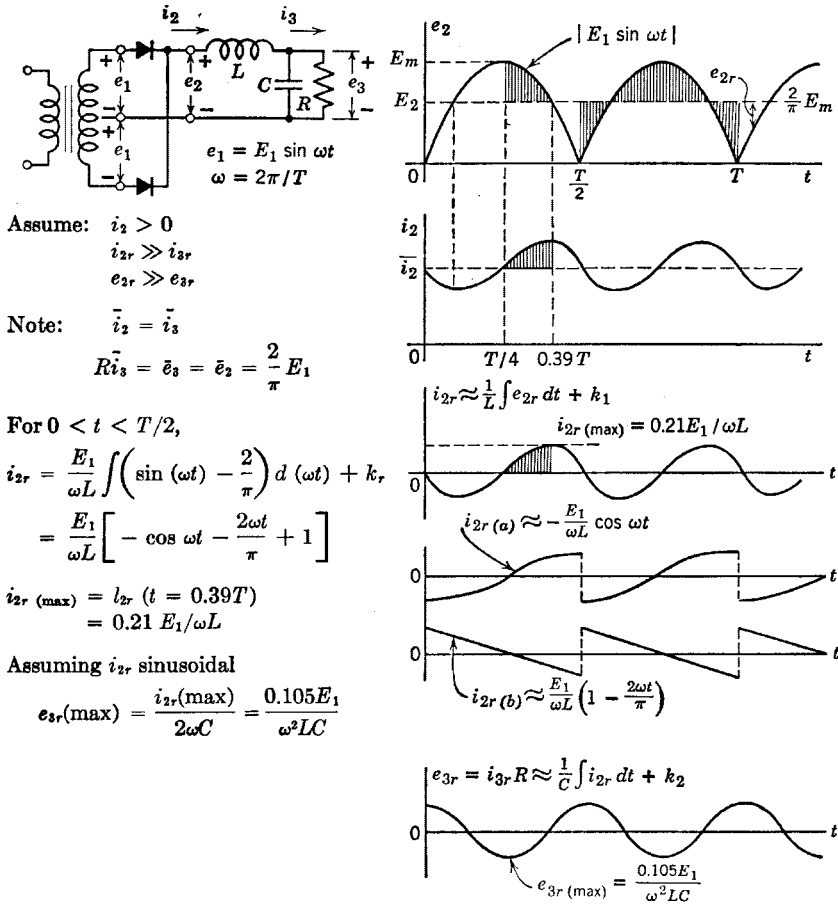
Fig. 4.14. Ripple filter with inductance input.

4.12   A Ripple Filter with Inductance Input

The capacitance-input filter described in the previous article requires large peak currents from the rectifying diodes in order to supply a nominal average or direct current. Since these peak currents are caused by the input capacitor C_1, let us remove C_1 from the filter of Fig. 4.13, leaving the configuration shown in Fig. 4.14. If the inductance L is large enough to make the ripple amplitude $i_{2r(\text{max})}$ smaller than the average value $\bar{i_2}$, then the current i_2 is always positive, and one diode or the other is always conducting. In addition to assuming i_2 always positive, let us assume effective filtering so that i_{2r} is much greater

than i_{3r} and e_{2r} is much greater than e_{3r}, as in the previous discussion. Also as before, we have

$$\overline{i_2} = \overline{i_3} = \frac{\overline{e_3}}{R} \qquad (4.17)$$

$$\overline{e_3} = e_2 \qquad (4.18)$$

But here

$$e_2 = \frac{2E_m}{\pi} \qquad (4.19)$$

As in the case of the capacitor-input filter, successive integrations of the ripple voltage e_{2r} yield the output ripple waveform e_{3r}:

$$i_{2r} = \frac{1}{L} \int e_{2r}\, dt + k_1 \qquad (4.20)$$

$$e_{3r} = \frac{1}{C} \int i_{2r}\, dt + k_2 \qquad (4.21)$$

A ripple filter with inductance input yields a d-c output voltage approximating the average value of the rectified waveform; whereas, with capacitance input, the d-c output approaches the peak value of the rectified sine wave.

If the total ripple current i_{2r} is assumed to be sinusoidal and much greater than i_{3r}, then it follows directly that the ripple output e_{3r} is approximately sinusoidal (at frequency 2ω) and has an amplitude

$$e_{3r\,\text{max}} = i_{2r\,\text{max}}/2\omega C$$

$$\approx 0.1\,\frac{E_1\,\dfrac{1}{\omega C}}{\omega L} \qquad (4.22)$$

where ω is the angular frequency of e_1.

The ripple voltage e_{3r} is nominally independent of load current I_3 for the inductance-input filter. This means that the inductance-input filter is well suited to applications requiring high output currents. In contrast, the ripple output from a capacitor-input filter is directly proportional to the useful load current or inversely proportional to load resistance (assuming that d-c voltage remains nearly constant).

The inductance-input filter requires a minimum load to sustain the mode of operation just described. As shown in Fig. 4.15, the d-c output voltage tends to rise sharply below a certain critical load current. Below this value of current, the diodes are both off during part of the cycle and

our analysis does not apply. For zero load current (infinite R), the d-c output voltage approaches E_1. The critical value of load resistance for which i_2 just reaches zero at one point in each half-cycle corresponds to the condition

$$\overline{i_2} = i_{2r\ max} \tag{4.23}$$

From Fig. 4.14

$$i_{2r\ max} = 0.21\,\frac{E_1}{\omega L} \tag{4.24}$$
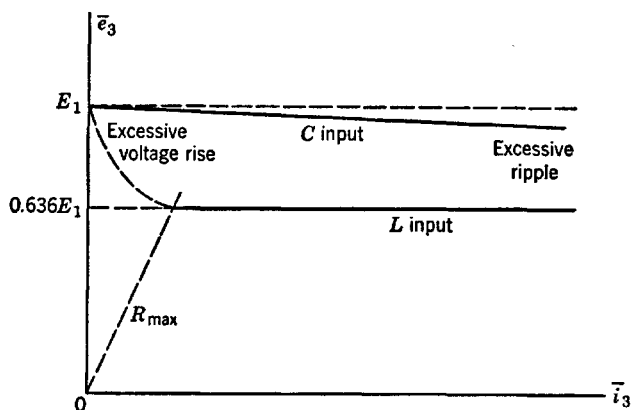
Fig. 4.15. Qualitative comparison of C-input and L-input filters.

The maximum value of load resistance which will maintain continuous diode conduction and thus good regulation and low ripple, as calculated from Eqs. 4.17–4.24, is

$$R_{max} \approx 3\omega L \tag{4.25}$$

Operation in the flat region of the curve of output voltage versus output current can be assured by permanently connecting a resistance R_{max} in parallel with the useful load. Then, even if the load resistance becomes infinite, the inductor current never goes to zero.

4.13   Amplitude-Modulation Detector

The basic rectifier circuit is also the basic amplitude-modulation (a-m) detector. Thus far we have considered several aspects of the behavior of this circuit, but always with a fixed-amplitude input signal. An amplitude-modulated sinusoid used as the input signal introduces

another variable: namely, the time-varying peak value. The waveform and the mathematical representation of a simple amplitude-modulated voltage are shown in Fig. 4.16(a). The wave has a sinusoidal "carrier" (frequency ω_c), the amplitude of which is varied or modulated by a signal

(a)
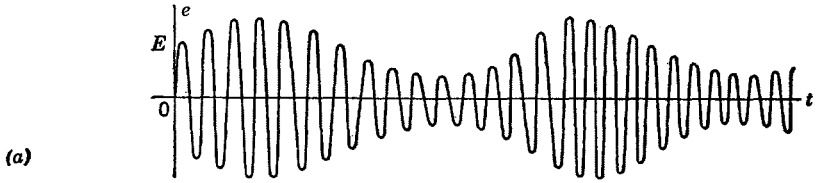
$$e = E(1 + m \sin \omega_m t) \sin \omega_c t$$

$$= E \left[\sin \omega_c t + \frac{m}{2} \cos (\omega_c - \omega_m)t - \frac{m}{2} \cos (\omega_c + \omega_m)t \right]$$

(b)

(c)

(d)

(e)
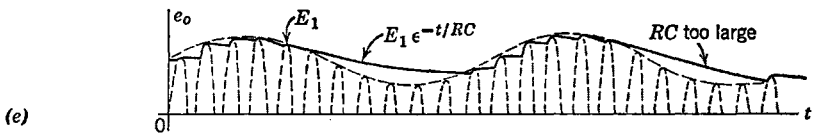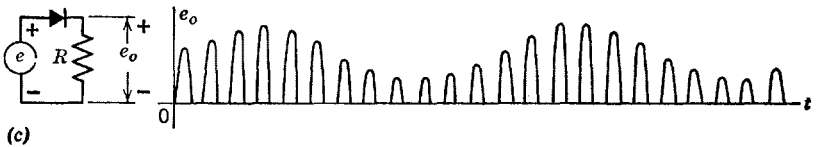
Fig. 4.16. Basic amplitude-modulation detector.

of lower frequency ω_m. The modulation index, m, is the fractional variation of E. For $m = 1$ (100 per cent modulation), the envelope rises to $2E_0$ and decreases to zero. The frequency spectrum shown in Fig. 4.16(b) is a graph of the amplitude versus frequency of the three components of a simple a-m wave.

The basic rectifier circuit is a widely used a-m detector; almost every a-m radio receiver uses one. The output of a resistive rectifier circuit with an a-m signal input is shown in Fig. 4.16(c). Either the peak amplitude (envelope) or the average value of each half-cycle contains the variation due to the modulating signal. The use of a capacitor to smooth out the high-frequency variations is shown in Fig. 4.16(d). To eliminate ripple and produce d-c output, we stated in Art. 4.3 that the capacitance C should be made as large as possible. We now desire to remove as much of the ripple as possible while preserving the modulation frequency. This imposes conflicting requirements on the size of capacitance C, and therefore requires a compromise. As the amplitude of the modulated wave builds up, the capacitance C is readily charged to nearly peak voltage by the diode. As the amplitude diminishes, the capacitance must discharge rapidly enough to follow the envelope. Too large a capacitance results in *envelope distortion*, as indicated in Fig. 4.16(e). Obviously, for $\omega_c \gg \omega_m$, the removal of ripple from the output can be made more complete without causing envelope distortion.

Let us consider in more detail the factors that influence the choice of an appropriate value for capacitance C. To avoid envelope distortion, the downward slope of the exponential discharge of capacitance C must exceed the downward slope of the modulation envelope. Assume the envelope to have reached a value E_a at time t_a during the downward swing as indicated in Fig. 4.17(a), and assume that the capacitor C has been charged to this peak value. Then the capacitor voltage during discharge is

$$e_c = E_a \epsilon^{-(t-t_a)/RC} \qquad (4.26)$$

and the initial slope is

$$\left. \frac{de_c}{dt} \right|_{t=t_a} = -E_a/RC$$

$$= \frac{-E(1 + m \sin \omega_m t_a)}{RC} \qquad (4.27)$$

Let us postulate that the design value of RC will prove to be much larger than the period of the carrier frequency ($RC \gg T_c = 2\pi/\omega_c$), so that the initial slope applies during the entire interval between carrier-frequency cycles. The slope of the modulation envelope is

$$\frac{de}{dt} = m\omega_m E \cos \omega_m t \qquad (4.28)$$

Modulation envelope:
$$e = E(1 + m \sin \omega_m t) \longrightarrow E_a = E(1 + m \sin \omega_m t_a)$$

Slope $\dfrac{de}{dt} = m\,\omega_m E \cos \omega_m t$

Exponential: $E_a \epsilon^{-\frac{t-t_a}{RC}}$

Slope: $-\dfrac{E_a}{RC}$

(a)

To avoid distortion: $\quad \dfrac{m\omega_m E \cos \omega_m t_a}{-E_a/RC} < 1; \qquad \dfrac{m\omega_m RC \cos \omega_m t_a}{-(1 + m \sin \omega_m t_a)} < 1$

(at t_a)

Peak-to-peak ripple: $E\,\Big\{(1 + m \sin \omega_m t_a)\,T_c/RC$
$\qquad\qquad\qquad - m\,[\sin \omega_m t_a - \sin \omega_m (t_a + T_c)]\Big\}$

(b)
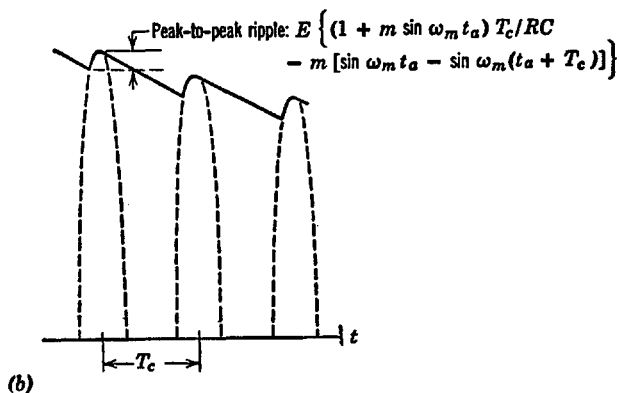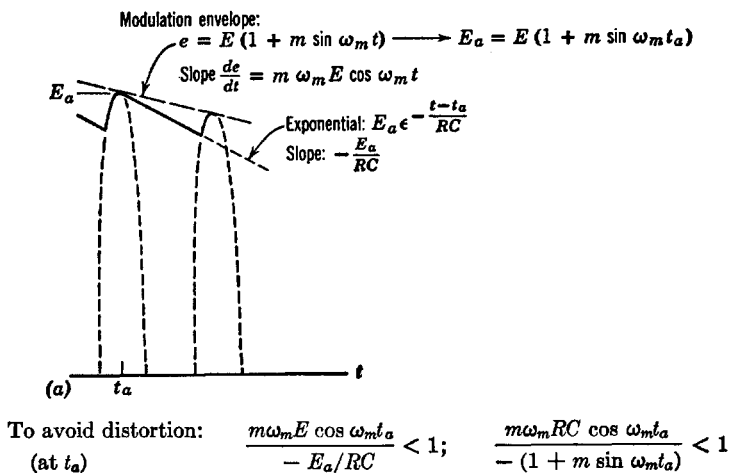
Fig. 4.17. Envelope distortion and ripple in diode detectors.

To avoid distortion, therefore,

$$-m\omega_m \cos \omega_m t_a < \frac{(1 + m \sin \omega_m t_a)}{RC} \tag{4.29}$$

For the least favorable value of $\omega_m t_a$,

$$RC \leqslant \frac{\sqrt{1 - m^2}}{m\omega_m} \tag{4.30}$$

Note that for $m = 1$ the inequality cannot be satisfied no matter how small RC is made ($RC = 0$ is of no practical value). For $m \approx 1$, a small amount of envelope distortion will inevitably occur near minimum envelope amplitude, since the exponential slope approaches zero linearly

with amplitude, whereas the slope of the envelope goes as a cosine function. The specific value of RC, then, is chosen on the basis of maximum modulation index m and the maximum ω_m for which freedom from distortion is desired.

For $\omega_c \gg \omega_m$, ripple is relatively small; but if this condition does not hold, the ripple may dictate a lower limit for RC. The peak-to-peak ripple calculation is shown in Fig. 4.17(b).

4.14 Square-Law Detection

Another method of detecting amplitude-modulated signals is called square-law detection. Any nonlinear curve can be represented by a power series expansion about a given point; and by proper selection of this operating point the square-law term can be enhanced. We shall therefore limit the present discussion to a square-law approximation. Introduction of higher-order terms refines the result, but adds to the algebraic complexity. Square-law detection is the fundamental process involved in frequency conversion or mixing. In addition, square-law analysis is useful for calculating harmonic distortion introduced by amplifiers or other circuits that are supposed to be linear.

Suppose the diode shown in Fig. 4.18 can be represented with reasonable accuracy by $i = Ke^2$ over some range of i and e. To obtain square-law detection the source and load resistances should be small so that the input voltage appears almost entirely across the diode and the current is proportional to the square of the voltage. (If R were made very large, current and voltage would be nearly linearly related.)

Let us first consider a sine wave applied to the circuit. As shown in Fig. 4.18, a d-c bias is used to insure operation in the square-law region of the diode curve. The voltage E_a produces a small d-c output voltage equal to RKE_a^2. Application of the sinusoidal input of frequency ω_c results in an output at ω_c, a double-frequency term at $2\omega_c$, and an additional d-c term due to carrier rectification.

The sinusoidal input corresponds to an amplitude-modulated wave with $m = 0$. Now, letting m have nonzero values, we obtain more complicated waveforms and trigonometric expressions, as shown in Fig. 4.19. Considering the expression for e_2, after squaring the expression within the bracket, we see that the first term $KRE_a^2 = I_aR$ merely represents the quiescent d-c output. The next term is the original amplitude-modulated sine-wave input multiplied by $2KRE_a$. For detection, the useful output comes from the last term designated by (A). The expansion for this term is shown in Fig. 4.19 as $e_2(A)$. The two terms in bold-face type represent the low-frequency output. The prod-

uct terms, if fully expanded, would yield frequencies $(2\omega_c + \omega_m)$, $(2\omega_c - \omega_m)$ and $(2\omega_c + 2\omega_m)$, $(2\omega_c - 2\omega_m)$. These, together with all
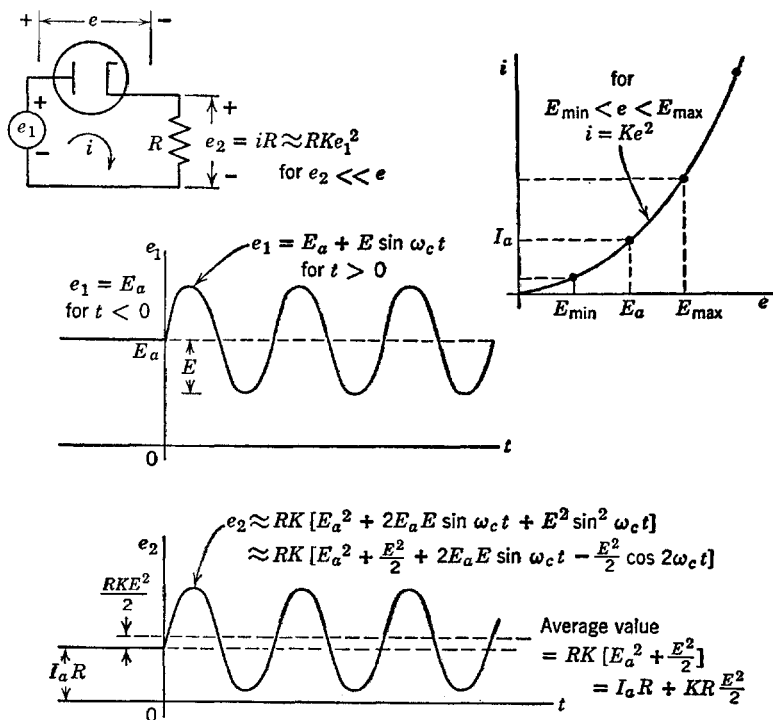

$e_2 = iR \approx RKe_1^2$

for $e_2 \ll e$

for
$E_{min} < e < E_{max}$
$i = Ke^2$

$e_1 = E_a + E \sin \omega_c t$
for $t > 0$

$e_1 = E_a$
for $t < 0$

$e_2 \approx RK [E_a^2 + 2E_a E \sin \omega_c t + E^2 \sin^2 \omega_c t]$
$\approx RK [E_a^2 + \frac{E^2}{2} + 2E_a E \sin \omega_c t - \frac{E^2}{2} \cos 2\omega_c t]$

Average value
$= RK [E_a^2 + \frac{E^2}{2}]$
$= I_a R + KR \frac{E^2}{2}$

Fig. 4.18. Square-law rectification of a sine wave.

terms other than the two low-frequency terms and the d-c term, can be removed from the output by means of a low-pass filter. A capacitor C with negligible reactance at carrier frequency $(1/\omega_c C \ll R)$, but large reactance at modulation frequency $(1/\omega_m C \gg R)$, would suffice. The frequency spectrum shown in Fig. 4.19 places in evidence each term of the trigonometric expansion of $e_2(t)$. It is much more readily drawn than the time function, and helps in visualizing the frequency distribution and relative amplitudes of the individual sinusoidal terms.

Actually, the desired output is $KRE^2 \, m \sin \omega_m t$. For a sinusoidal modulating signal, the second-harmonic term $KRE^2(m^2/4) \cos 2\omega_m t$ may often be largely eliminated from the output by a suitable filter. However, if the modulation extends over a range of frequencies (as, for example, the speech or music broadcast by a radio station), then harmonic distortion is inevitable. The ratio of the undesired term to the desired term is $m/4$; thus the square-law detector must be used in

$$e_2 = iR \approx RKe_1^2 \quad \text{for } e_2 \ll e$$

for $E_{min} < e < E_{max}$
$$i = Ke^2$$

$$e_1 = E_a + E(1 + m \sin \omega_m t) \sin \omega_c t$$

$$e_2 = RK[E_a + E(1 + m \sin \omega_m t) \sin \omega_c t]^2$$
$$= RK[E_a^2 + 2E_a E(1 + m \sin \omega_m t)(\sin \omega_c t)$$
$$\qquad + E^2(1 + m \sin \omega_m t)^2 (\sin \omega_c t)^2]$$

(A)
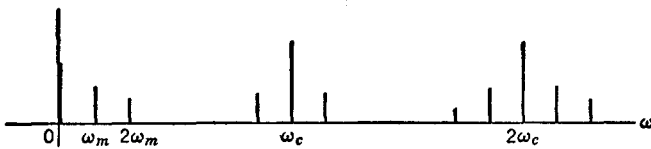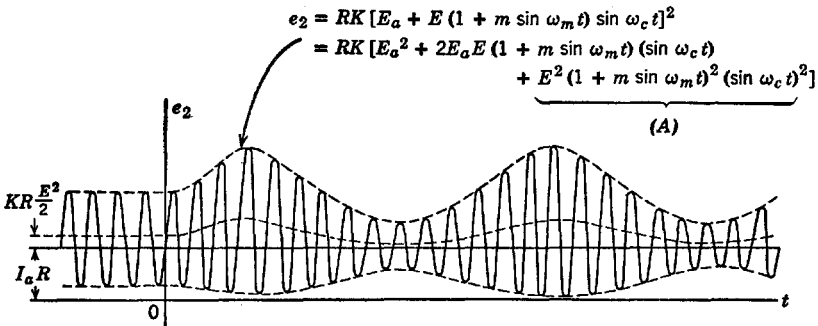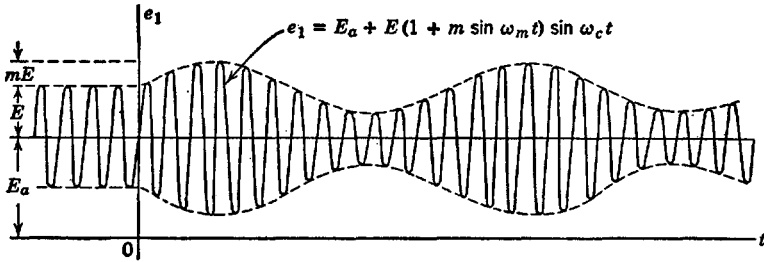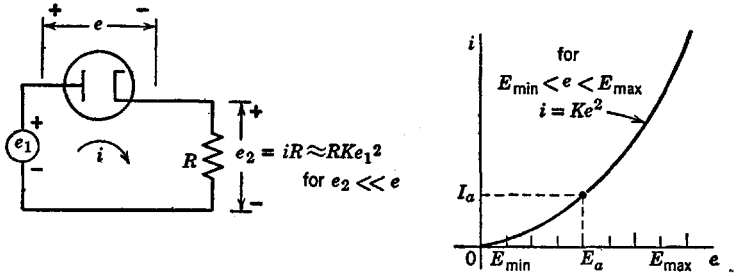
$$e_{2A} = RKE^2 \left[1 + 2m \sin \omega_m t + \frac{m^2}{2} - \frac{m^2}{2} \cos 2\omega_m t\right][\tfrac{1}{2} - \tfrac{1}{2} \cos 2\omega_c t]$$

$$= RKE^2 \left[\frac{1}{2} + m \sin \omega_m t + \frac{m^2}{4} - \frac{m^2}{4} \cos 2\omega_m t - \frac{1}{2} \cos 2\omega_c t\right.$$

$$\left. - m \sin \omega_m t \cos 2\omega_c t - \frac{m^2}{4} \cos 2\omega_c t + \frac{m^2}{4} \cos 2\omega_m t \cos 2\omega_c t\right]$$

Fig. 4.19. Square-law detection of an a-m sine wave.

applications where the modulation index of the applied wave is likely to be small.

4.15 Balanced Modulator Circuits

Modulation of a time function $e_1(t)$ by another time function $e_2(t)$ is a nonlinear process. The resulting modulated wave contains frequencies that were not present in the original time functions. If the two time
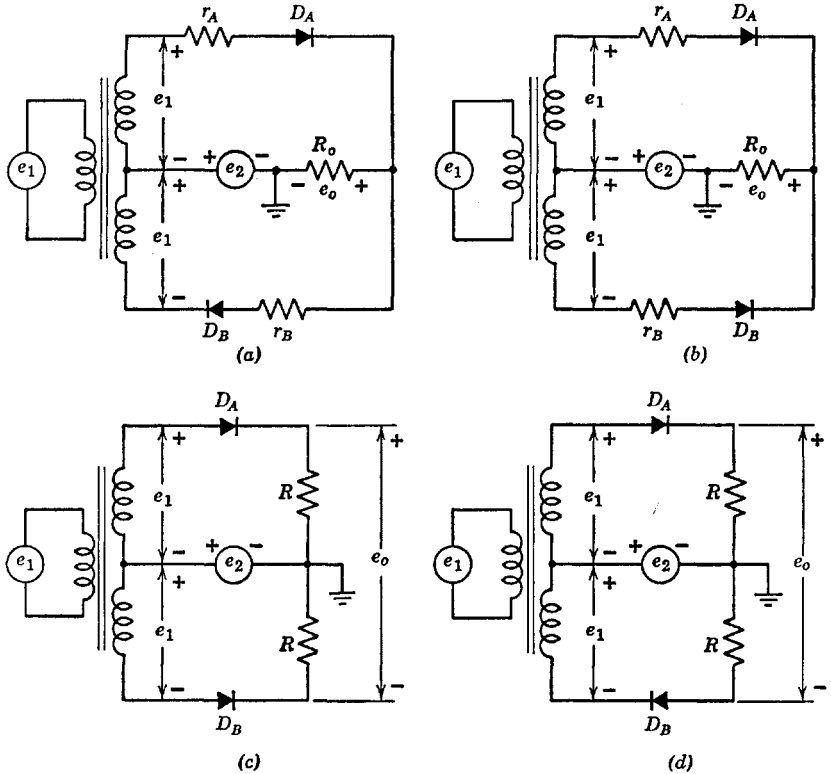
Fig. 4.20. Balanced modulator and demodulator circuits.

functions are sine waves of angular frequency ω_1 and ω_2, the modulated wave may contain frequencies ω_1 and ω_2, $(\omega_1 + \omega_2)$ and $(\omega_1 - \omega_2)$, $2\omega_1$, $2\omega_2$ and many others. The components present in the frequency spectrum of the output depend upon the circuit and the degree of nonlinearity.

Demodulation (detection) is a variation of the modulation process, usually aimed at recovering one of the signals in a modulated wave.

Both the diode detector and the square-law circuit carry out the demodulation process by rectifying and then filtering the modulated wave. When a low-frequency signal is to be recovered, the required

Fig. 4.21. Waveforms for the circuit of Fig. 4.20(a).

filter is low-pass and may be a simple smoothing capacitor. A band-pass filter is used if one desires to select a particular high-frequency com-

ponent of the modulated wave. Such a filter might well consist of a parallel RLC circuit resonant at the desired frequency.

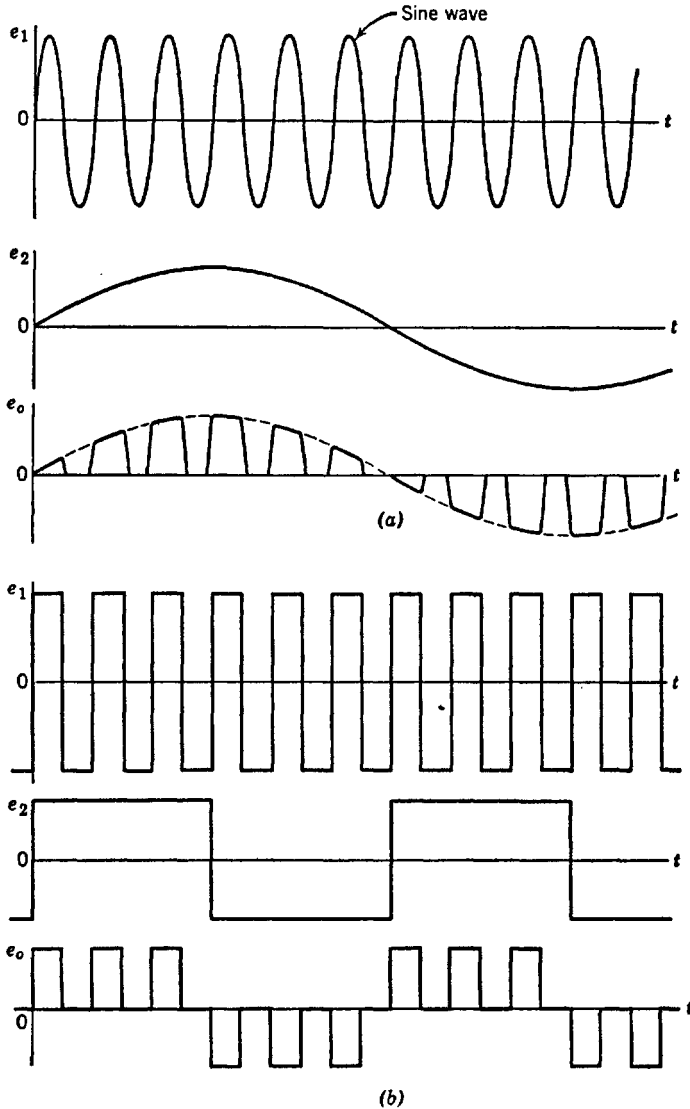Balanced modulator or demodulator circuits make use of symmetry to eliminate undesired frequency components. These circuits are sometimes called keyed rectifiers since they are full-wave rectifiers with two input signals. See Fig. 4.20. The important point to observe is that in all of the circuits one input is applied to the diodes symmetrically, whereas the other is applied antisymmetrically. The character of the input signals and the degree of diode nonlinearity determine which of several functions such a circuit performs. Among the possibilities are suppressed-carrier modulation, synchronous detection, pulse-amplitude modulation, waveform sampling, and phase sensing.

In the circuit of Fig. 4.20(a) the source e_1 alone tends to switch both diodes on and off together. The small resistance r_A and r_B may be actual diode resistances or resistances added to the circuit to improve balance so that e_1 alone yields no output. The source e_2 alone appears directly at the output (or nearly so for r_A and $r_B \ll R_0$). When e_2 is positive D_A conducts and when e_2 is negative D_B conducts.

With both e_1 and e_2 applied to the circuit at the same time we note that conduction through D_A depends on $(e_1 + e_2)$, whereas conduction through D_B depends on $(e_1 - e_2)$. Output waveforms for two different sets of input waveforms are shown in Fig. 4.21(a) and (b). In both cases $|e_1| > |e_2|$.

The circuit of Fig. 4.20(b) is like that of (a) but with ideal diode D_B reversed. Reversing a diode, in effect, interchanges the roles of e_1 and e_2, since now e_2 drives the two diodes on and off simultaneously, while e_1 alternately causes conduction through one diode and then the other. With e_2 alone applied, e_0 is a half-wave rectified version of e_2, whereas in the circuit of (a), e_2 appears directly at e_0. The circuits shown in Fig. 4.20(a) and (b) have a common ground terminal for the output and one of the sources. The circuits in (c) and (d) have a balanced output with respect to ground.

SUPPLEMENTARY READING

L. B. Arguimbau, R. B. Adler, *Vacuum Tube Circuits and Transistors,* John Wiley and Sons, New York, 1956.

B. Chance, *Waveforms,* McGraw-Hill, New York, 1948.

T. S. Gray, *Applied Electronics,* 2nd edition, John Wiley and Sons, New York, 1954.

E. A. Guillemin, *Introductory Circuit Theory,* John Wiley and Sons, New York, 1953.

J. Millman and H. Taub, *Pulse and Digital Circuits,* McGraw-Hill, New York, 1956.

Samuel Seely, *Electronic Engineering,* McGraw-Hill, New York, 1956.

Samuel Seely, *Radio Electronics,* McGraw Hill, New York, 1956.

PROBLEMS

4.1. Given the input $e_1(t)$ of Fig. P4.1, find the steady-state waveform of $e_2(t)$ for each of the circuits shown in Figs. P4.2–P4.7.
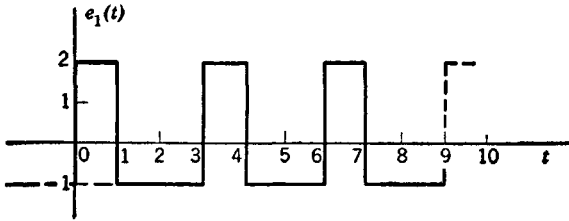
Fig. P4.1
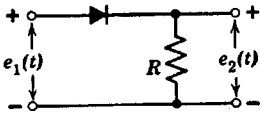
Fig. P4.2
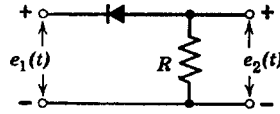
Fig. P4.3
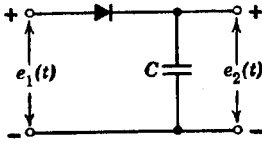
Fig. P4.4
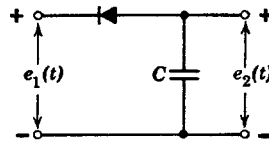
Fig. P4.5
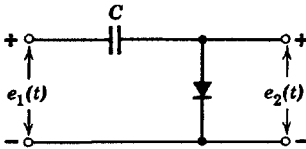
Fig. P4.6
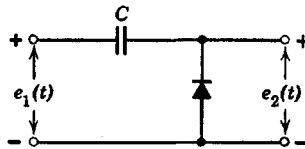
Fig. P4.7

4.2. Sketch $e_2(t)$ for the diode circuits shown in Figs. P4.8–P4.14, assuming that ωRC is much larger than unity so that the capacitor ripple voltage is negligible.
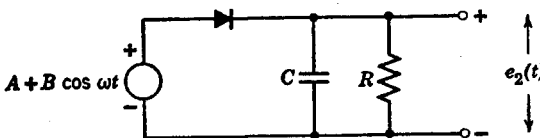
Fig. P4.8
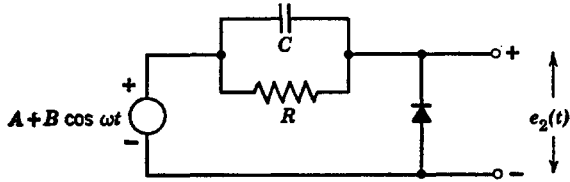
Fig. P4.9
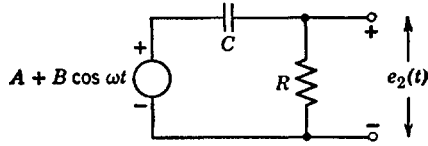
Fig. P4.10
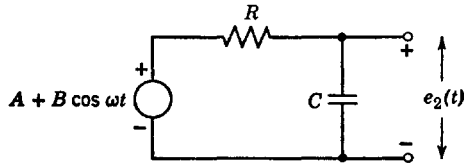
Fig. P4.11
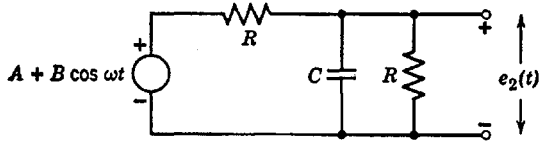
Fig. P4.12
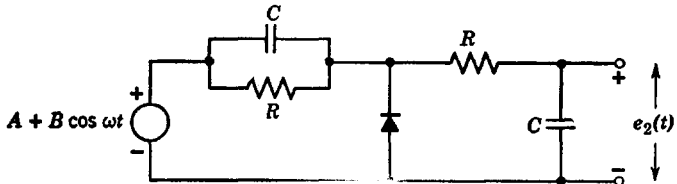
Fig. P4.13
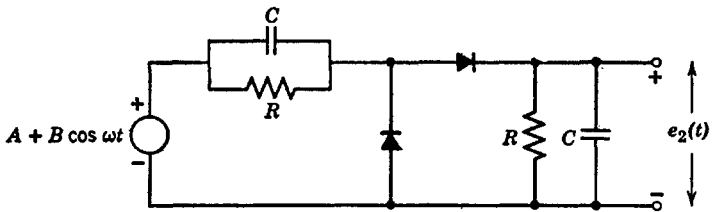
Fig. P4.14

4.3. The input voltage $e_1(t)$ is a periodic square wave of peak-to-peak amplitude $2E_1$ and period T. Assume that RC is many times larger than the period of the square wave. Find the approximate peak-to-peak ripple voltage in the output $e_2(t)$, expressed as a percentage of the d-c output voltage for (a) the circuit of Fig. P4.15, (b) the circuit of Fig. P4.15 with a resistance R in parallel with the capacitor, (c) the circuit of Fig. P4.15 with a resistance R in parallel with the diode, (d) the circuit of Fig. P4.16, and (e) the circuit of Fig. P4.16 with an additional resistance R across the e_2 terminals.
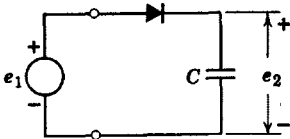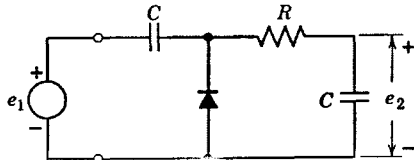
Fig. P4.15   **Fig. P4.16**

4.4. The voltage pulse $e_1(t)$ shown in Fig. P4.17 is applied to the circuit of Fig. P4.18. Find $e_2(t)$.
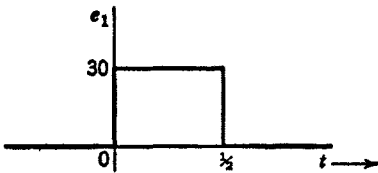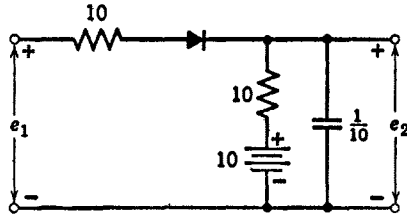
Fig. P4.17   **Fig. P4.18**

4.5. The waveform $e_1(t)$ in Fig. P4.19 is applied to the circuit of Fig. P4.20. Sketch $e_2(t)$.
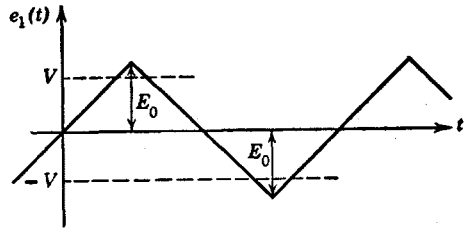
Fig. P4.19
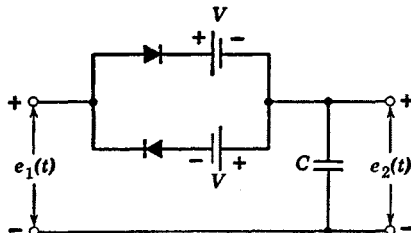
Fig. P.20

4.6. For the circuit of Fig. P4.21, sketch and dimension the waveforms $e_1(t)$ and $e_2(t)$ when a single brief pulse of current $i_1(t)$ is applied $(\delta_1 \ll \sqrt{LC_1})$.
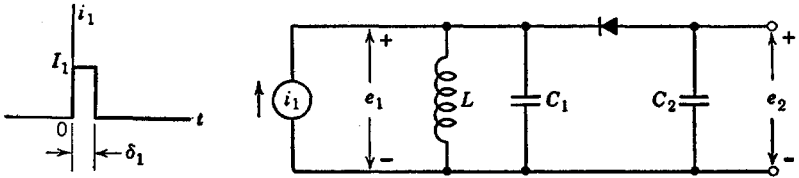
Fig. P4.21

4.7. A full-wave rectifier is driven by an ideal transformer, supplying 100 volts at a frequency of $200/\pi$ cps to each half of the circuit shown in Fig. P4.22. The diodes are assumed to be ideal rectifiers. The capacitor C is 1 microfarad. For a certain value of R each diode conducts for $\frac{1}{10}$ of each half period. (Note: $\sin x \approx x$ for small values of x.)
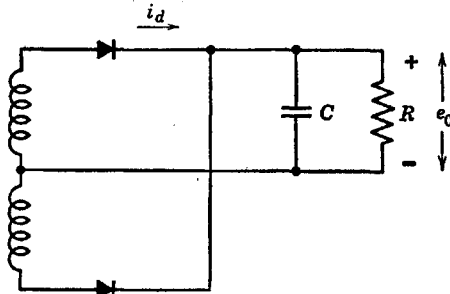
Fig. P4.22

(a) Sketch and dimension the diode current i_d as a function of time.
(b) Find the average value of i_d.
(c) Find the value of R.
(d) Find the ripple voltage.
(e) Find the rms value of i_d.

4.8. The circuit shown in Fig. P4.23 is used to demodulate an amplitude modulated signal e_1. The carrier frequency is 10^6 cps, the highest modulation frequency 5000 cps. The modulation factor m is never larger than 0.8.
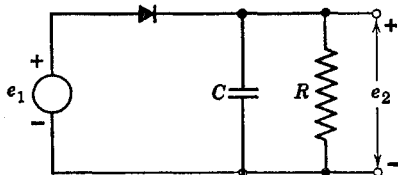
Fig. P4.23

(a) Determine the maximum allowable value for the product RC when the circuit is required to detect the modulation without incurring envelope distortion.

(b) For a carrier signal of 10 volts peak, modulated with 1000 cps at $m = 0.3$, list the approximate (average) ripple and audio output voltages under the following conditions:

$$RC = 0.5 \text{ times critical value}$$
$$RC = 0.1 \text{ times critical value}$$
$$RC = 0.01 \text{ times critical value.}$$

4.9. In Fig. P4.24, e_1 is a square wave that jumps between plus and minus 75 volts every $T/2$ seconds.

(a) Plot e_2 as a time function if $T \ll RC$.

(b) Repeat (a) if $T \gg RC$.
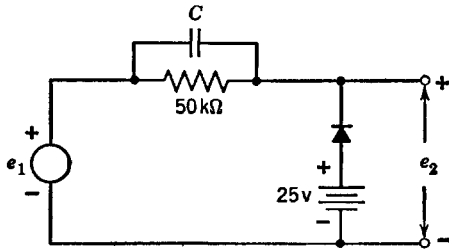
Fig. P4.24
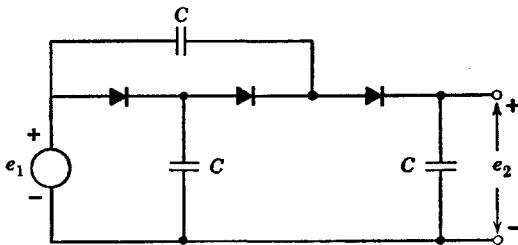
Fig. P4.25

4.10. In Fig. P4.25, e_1 is a rectangular wave that jumps between $\pm E_1$ every $T/2$ seconds starting at $t = 0$. Plot e_2 as a function of time until steady-state conditions exist. Assume ideal diodes, and the capacitors initially uncharged.

4.11. In the circuit shown in Fig. P4.26,
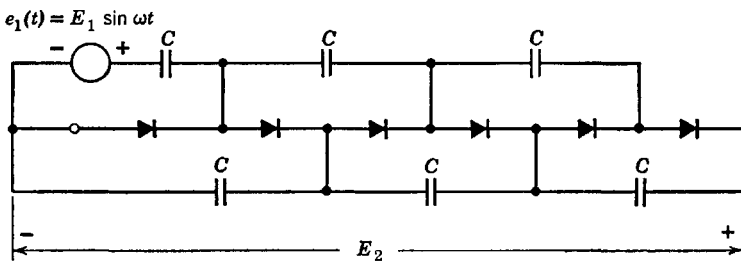
Fig. P4.26

(a) Find the steady-state value of E_2.

(b) What is the maximum peak inverse voltage the diodes have to stand?

(c) What is the maximum voltage on each capacitor?
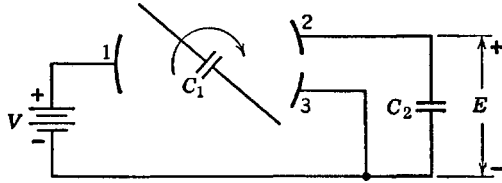
Fig. P4.27

4.12. The capacitor C_1 in Fig. P4.27 rotates about an axis normal to the plane of the diagram, its terminals alternately making contacts 1, 2 and 1, 3.

(a) Find the steady-state value of E.
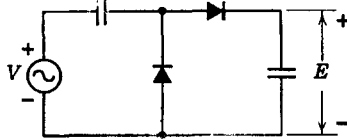
Fig. P4.28

(b) Show the analogy between the switching circuit of Fig. P4.27 and the rectifier circuit of Fig. P4.28, which produces a steady-state output E when attached to a sinusoidal voltage V.
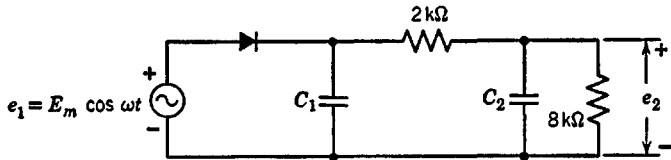
Fig. P4.29

4.13. The rectifier circuit shown in Fig. P4.29 produces a steady-state output voltage e_2 containing a time-average component $\overline{e_2}$ and a periodic ripple component. The design specifications require the peak-to-peak ripple to be no more than $0.001\overline{e_2}$. Consider two designs. First let C_2 be zero and find the required value of C_1. Second, find C_1 and C_2 such that the specifications are satisfied and at the same time $C_1 + C_2$ is minimized. Which design requires less total capacitance?

4.14. The circuit shown in Fig. P4.30 is at rest when the square wave shown in Fig. P4.31 is suddenly applied. Assume that the capacitor C is sufficiently large so that the periodic ripple in the output voltage is small and e_2 may be replaced by a ripple-free average value E_2. At an arbitrary time t the value of E_2 is undetermined. Formulate, in terms of E_1, r, R, T, and E_2,

(a) The net charge per cycle flowing through resistances r and R.

(b) The net charge per cycle delivered to C.

(c) Using the results of parts (a) and (b), formulate and solve a differential equation to find $e_2(t)$.

(d) What is the value of E_2?
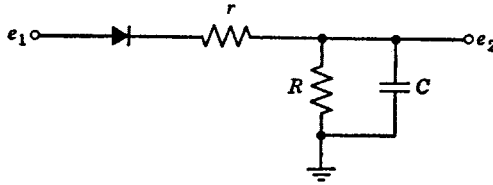
Fig. P4.30
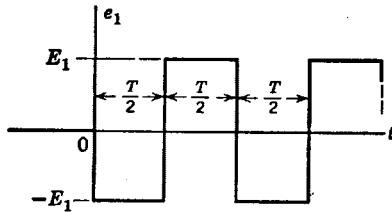
Fig. P4.31

(e) On the basis of the charge delivered to C during the first and second halves of the cycle, what is the approximate peak-to-peak steady-state ripple?

4.15. It is desired to supply a current i to an instrument m whose resistance is negligibly small. The direct component of i is to be as large as possible and the ripple (ratio of a-c amplitude to d-c) must not exceed some specified
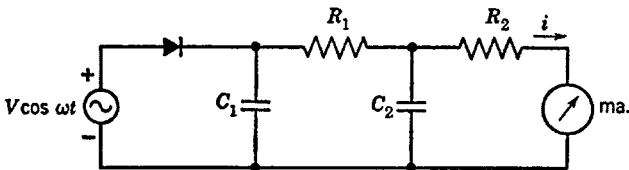
Fig. P4.32

limit, say one-tenth per cent. Quantities V, ω, and the total capacitance $C = C_1 + C_2$ are known constants, being fixed by cost and convenience considerations. Find optimum values of R_1, R_2, C_1, and C_2 in Fig. P4.32. Use any approximations which appear reasonable.

4.16. In the detector circuit of Fig. P4.33 the product ωRC is so large that capacitor ripple voltage may be ignored. Sketch and dimension the curve of

(a) E versus A, with B positive and fixed.

(b) E versus B, with A positive and fixed.

4.17. The circuit shown in Fig. P4.34 is a piecewise-linear model of a microwave detector. For simplicity of analysis, assume that the high-frequency input voltage $e_1(t)$ is a square wave of amplitude E_1. Capacitance C and inductance L are so large that the ripple voltage across C and the ripple

current in L are negligible. Hence $i_1(t)$ must be a square wave of some amplitude I_1. Derive relations giving I_1 and I_2 in terms of E_1 and E_2. Show that
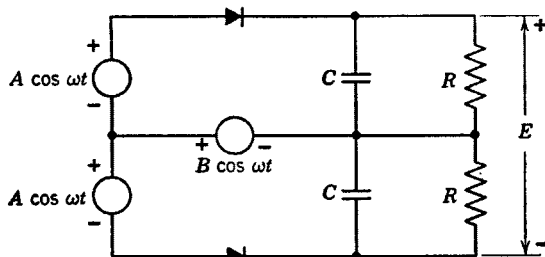
Fig. P4.33
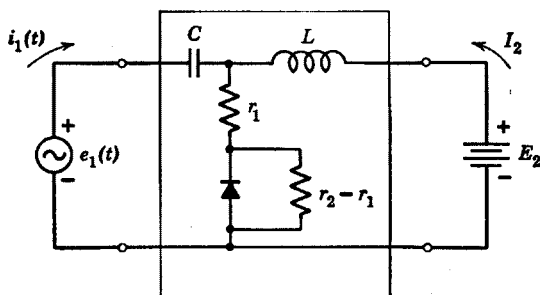
Fig. P4.34

these relations may be represented as the d-c circuit (Fig. P4.35). Find r_a, r_b, and r_c in terms of r_1 and r_2. Now let source $e_1(t)$ have a series resistance R_1 and let E_2 be replaced by a resistance R_2. Find R_1 and R_2 in terms of r_1 and r_2 so as to maximize the load power $I_2{}^2 R_2$ per unit available source power $E_1{}^2/4R_1$.
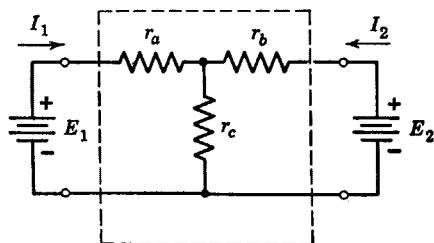
Fig. P4.35

4.18. The square waves of voltage $e_1(t)$ and $e_2(t)$ in Fig. P4.36 are applied to the circuit in Fig. P4.37. Sketch the output waveform $e_3(t)$.

4.19. For the circuit shown in Fig. P4.38 with $e_1(t) = 100 \sin \omega t$,

(a) Sketch and dimension as a function of time the voltage $e_2(t)$ with switch open.

(b) Find the approximate peak-to-peak ripple voltage component of e_3 after the switch has been closed a long time. (Assume $\omega RC \gg 1$.)
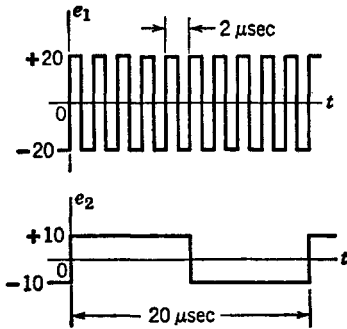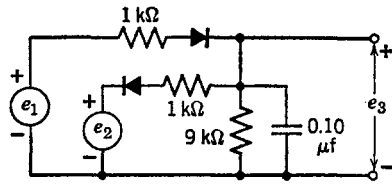
Fig. P4.36   Fig. P4.37
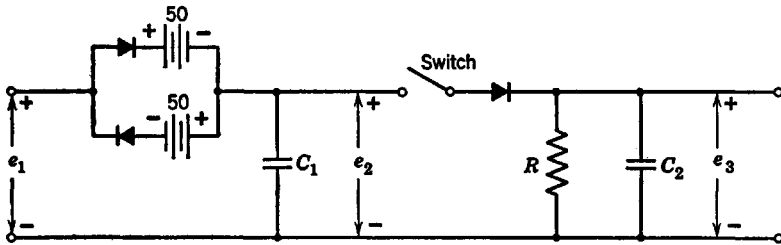
Fig. P4.38

4.20. From the circuit shown in Fig. P4.39:

(a) Determine approximately the secondary voltage required in the input transformer for $E_{d-c} = 400$ v and $I_{d-c} = 0.030$ amp.
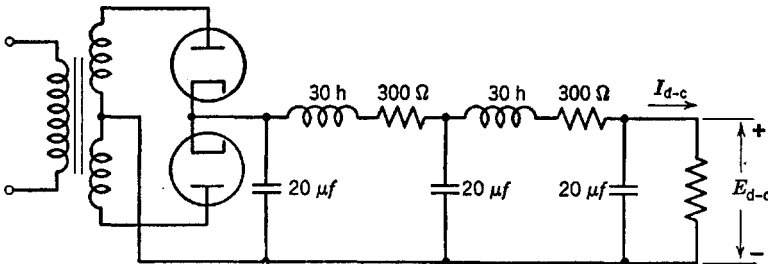
Fig. P4.39

(b) Determine approximately the output rms hum voltage for power frequencies of 60 cps and 800 cps.

Justify all approximations clearly.

4.21. From Figs. P4.40 and P4.41 (making reasonable approximations)

(a) Determine the direct output voltage for a load current of 125 ma.

(b) Determine the value of the 120 cps hum voltage.

(c) Determine the minimum load current for which E_{d-c} is relatively independent of I_{d-c} as the load resistance is varied.
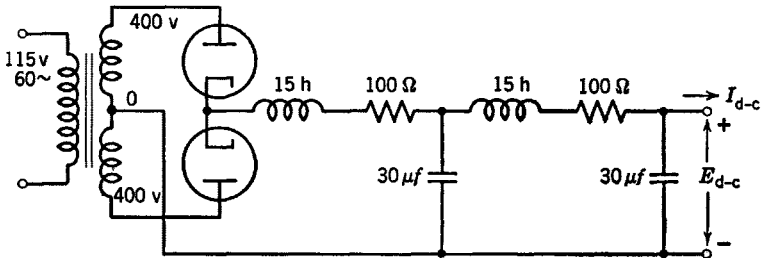
Fig. P4.40
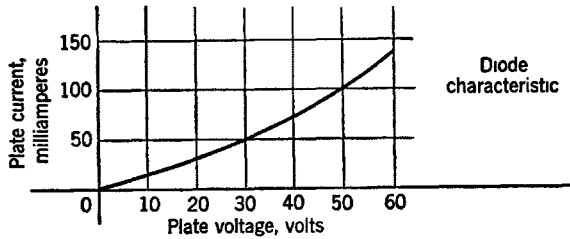
Fig. P4.41

4.22. For the circuit shown in Fig. P4.42, what can be said in general about the relationship of the output voltage e_0 to the three input voltages v_a, v_b, and v_c?
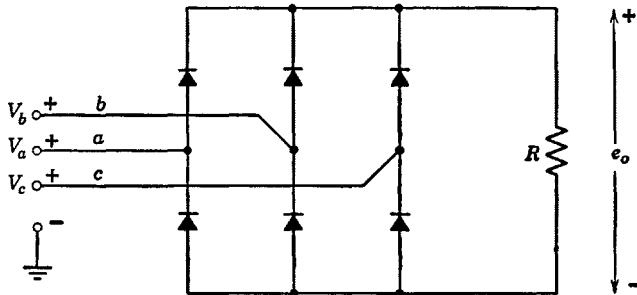
Fig. P4.42

4.23. (a) Sketch the output voltage e_0, for the circuit of Problem 4.22, if the input voltages are balanced three phase voltages, that is,

$$v_1 = v_b - v_a = V \sin \omega t$$

$$v_2 = v_c - v_b = V \sin (\omega t - 120°)$$

$$v_3 = v_a - v_c = V \sin (\omega t - 240°)$$

(b) What is the fundamental frequency of the ripple voltage in the output with the input of (a)?

(c) What is the per-unit ripple?

4.24. The elements in the circuit of Fig. P4.43(a) are ideal. In addition, $\sqrt{L/C}$ is small enough so that the tuned circuit voltage is, for all practical purposes, sinusoidal, and $\omega R_1 C_1 \gg 1$. As far as can be determined from bridge measurements over a range of frequencies in the vicinity of the tuned circuit resonant frequency, the terminal behavior of the circuit facing the current source is the same as that of the circuit shown in Fig. P4.43(b).
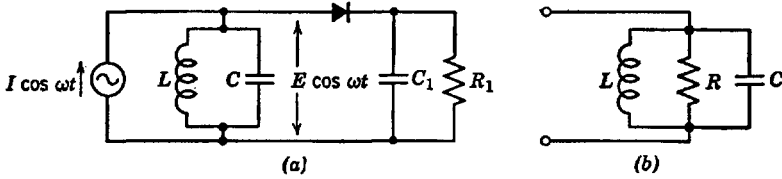
Fig. P4.43

(a) What power is dissipated in R_1?
(b) What power is dissipated in the diode?
(c) Relate I to E.
(d) What value of R was determined from the bridge measurements?

4.25. In Fig. P4.44, voltage e_1 is a square wave of amplitude E_1 and capacitor C is so large that ripple in the output voltage E_2 may be ignored. Derive the relation between E_1 and E_2. Show that $E_2 = f(E_1{}^2)$ is asymptotic to a

$$i = I_0(\epsilon^{v/V_0} - 1)$$
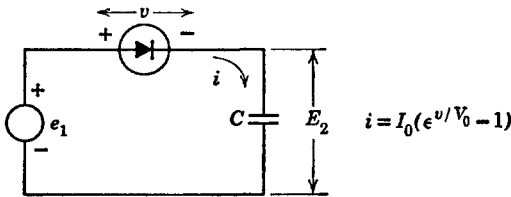
Fig. P4.44

parabola for small E_1 and to a straight line for large E_1. Plot the two asymptotes in the E_2/V_0 versus E_1/V_0 plane. Find one or two points in the transition region between asymptotes, and sketch the curve $E_2 = f(E_1)$.

4.26. The process called frequency conversion is a form of modulation used in superheterodyne radio receivers and in other parts of communication systems where the sum or difference of two frequencies is needed. The use of a square-law curve for the diode in Fig. P4.45 simplifies the determination of output voltage e_0. The square-law approximation can be justified by noting that this term (of a series representation of the diode curve) yields the desired difference frequency. In a practical circuit, resistance R is replaced by a filter to enhance the desired frequency component and reject all others, including those generated by the terms other than square law. Find all of the frequencies generated in the output, assuming $E_a > E_1 + E_2$.

4.27. For the balanced modulator circuit shown in Fig. P4.46, assume $e_1 = E_1 \sin \omega_1 t$ and $e_2 = E_2 \sin \omega_2 t$, with $E_2 \gg E_1$ and $\omega_2 \gg \omega_1$. Sketch the waveform of output voltage e_0.

$$i = ke^2 \text{ for } e > 0$$
$$i = 0 \text{ for } e < 0$$

$e_1 = E_1 \sin \omega_1 t$

$e_2 = E_2 \sin \omega_2 t$

E_a

$e_0 = iR \approx Rke^2 \text{ for } e_0 << e$
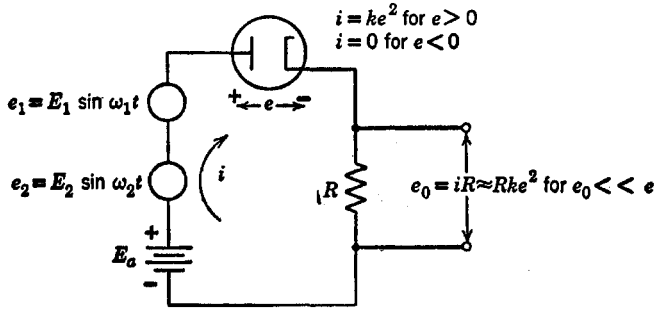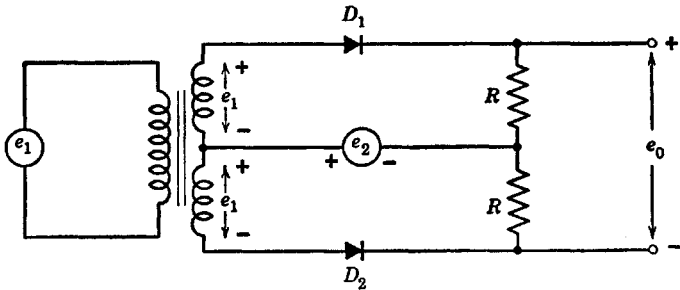
Fig. P4.45

Fig. P4.46

4.28. Repeat Problem 4.27 with e_2 consisting of brief rectangular pulses as shown in Fig. P4.47. Assume the repetition frequency $f_2 = 1/\delta_2$ exceeds
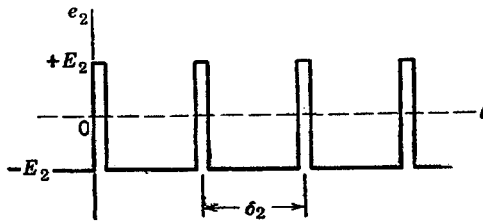
Fig. P4.47

$f_1 = \omega_1/2\pi$. Note that the output consists of "samples" of the input taken at the time of each rectangular pulse.

4.29. The phase sensitive detector shown in Fig. P4.48 is driven by a reference voltage v_r and a signal voltage v_s, where

$$v_r = V_r \cos 800\pi t$$

Assume ideal diodes and investigate the circuit's characteristics under the following conditions.

(a) Plot the output voltage as a time function if $v_r = 100 \cos 800\pi t$ volts and $v_s = 50 \cos 800\pi t$ volts.

(b) Plot the output voltage as a time function if $v_r = 100 \cos 800\pi t$ volts and $v_s = 50 \sin 800\pi t$ volts.
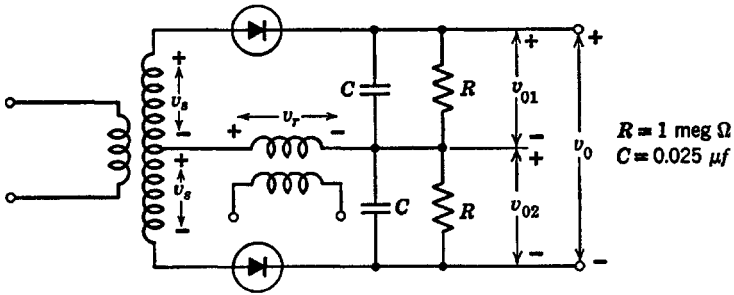
Fig. P4.48

(c) Plot a vector diagram that illustrates how the output of this circuit varies with the phase angle ϕ for

$$v_s = V_s \cos (800\pi t + \phi)$$
$$v_r = V_r \cos 800\pi t$$

Transistor Models and Circuits

5.1   Introduction

In its most familiar form the transistor is a semiconductor triode. Like the semiconductor diode, it was first developed as a point-contact device. The point-contact transistor consists of a base of semiconductor material, say n-type germanium, on which two fine, pointed wires rest in close proximity. One of these is called the emitter, and the other is the collector. As shown in Fig. 5.1 (a) the unit appears to be two point-contact diodes using the same base. The emitter and the collector are not identical, since the latter is treated by a process called "forming" to modify the base material near the collector.

The junction transistor, evolved several years after the point-contact transistor, can be described qualitatively as two junction diodes utilizing a common base. Junction transistors are made in a variety of ways. In one method a single crystal of germanium or other semiconductor is grown by applying a seed crystal to the top of a molten bath of the semiconductor material. Withdrawing the crystal at the proper rate results in cooling and crystallization. As the crystal is withdrawn, electron-donor or electron-acceptor impurities are added in carefully controlled amounts to form alternate layers of n-type and p-type semiconductors. The crystal is then cut into small slabs, and leads are attached by soldering. The result is a grown-junction tran-

sistor indicated in Fig. 5.1(b). The dimensions of the semiconductor slab are usually small fractions of an inch, and the thickness of the base layer is likely to be of the order of 0.001 inch. Locating the base layer and attaching the connecting lead to it requires precise fabrication techniques.
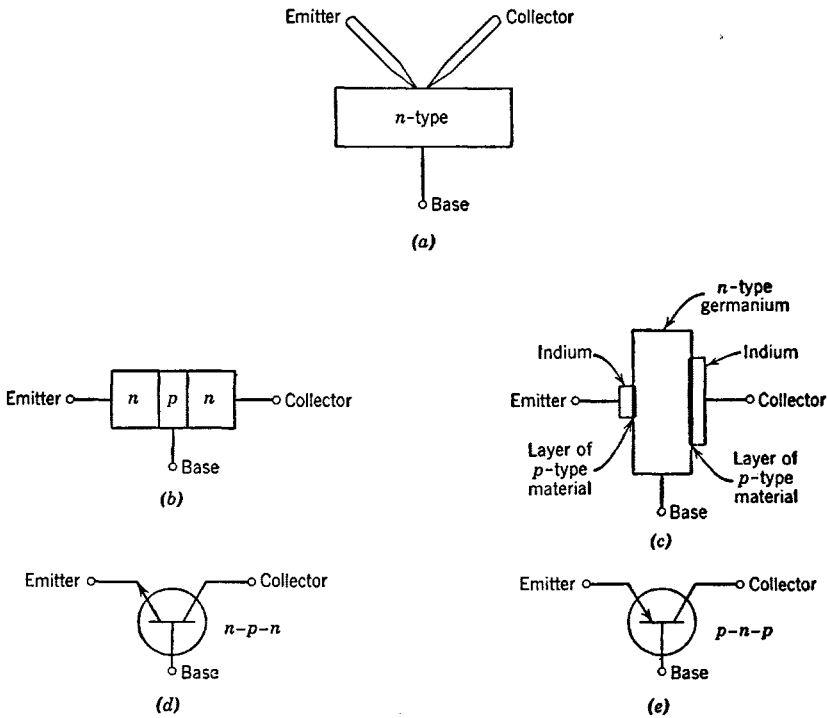
Fig. 5.1.   Transistor configuration.

The alloy-junction transistor, Fig. 5.1(c), is made by placing small amounts of an acceptor like indium on opposite sides of a piece of n-type semiconductor, say n-type germanium. When the unit is heated to high temperatures, an alloy of the indium and germanium is formed on each side of the germanium. The result is the formation of p-type material on each side of the n-type and, hence, a p-n-p transistor. A p-type semiconductor material can be alloyed with a donor impurity such as arsenic to make n-p-n transistors. Figure 5.1(d) and (e) shows circuit symbols.

There are numerous other fabrication processes used to make tran-

sistors, some of which result in substantially different kinds of units. The reader is referred to other books and to the literature for details, since the subject is already a broad one and continues to evolve as new techniques and new materials are exploited. A few references are listed at the end of the chapter.

The qualitative discussion of transistor behavior in the following article is based on the treatment of semiconductor diodes in Chapter 2 and will be limited to the junction transistor. The physics of the point-contact transistor is more involved, but it leads to a circuit model of the same type. This brief account is included as a means of relating the basic physical processes to the electrical behavior and to the circuit models.

5.2   Electrical Conduction in *p-n-p* Junction Transistors

In both *p*- and *n*-type semiconductor material, hole-electron pairs are created by thermal agitation. At equilibrium, just as many pairs disappear by recombination as are formed by thermal agitation. In *p*-type material, holes are the majority carriers and the free electrons that exist are called minority carriers, since they are outnumbered by the free holes provided by the acceptor impurity. In *n*-type material electrons are the majority carriers and the holes are minority carriers, since the donor impurity provides a large number of free electrons. Diffusion of the majority carriers across a *p-n* junction (holes from *p* to *n* and electrons from *n* to *p*) establishes a potential barrier which eventually brings the diffusion process into equilibrium in the absence of applied fields.

When a small forward bias voltage is applied to a *p-n* junction (plus to minus from *p* to *n*), the potential barrier is effectively lowered, and a large current flows as the result of both kinds of majority carriers moving in opposite directions. The mobility of electrons differs from that of holes, and the number of holes and electrons may be unequal; hence, the contributions of the two types of majority carriers to the total conduction current are not necessarily equal. In fact, the operation of a transistor is based on current flow from only one type of majority carrier. For a *p-n-p* transistor, let us make the first junction (*p-n*) from emitter to base with the *p*-type material having more acceptor impurity content than the *n*-type has donor impurity. Then the forward current from emitter to base will consist largely of holes (+) going from emitter to base, rather than electrons (−) going from base to emitter. This is illustrated in Fig. 5.2(*a*).

Now recall that a reverse bias voltage on a p-n junction results in a small current caused by the minority carriers. This current reaches a saturation value that is temperature dependent, but remains constant over a fairly wide range of reverse voltages, as shown in Fig. 5.2(b).
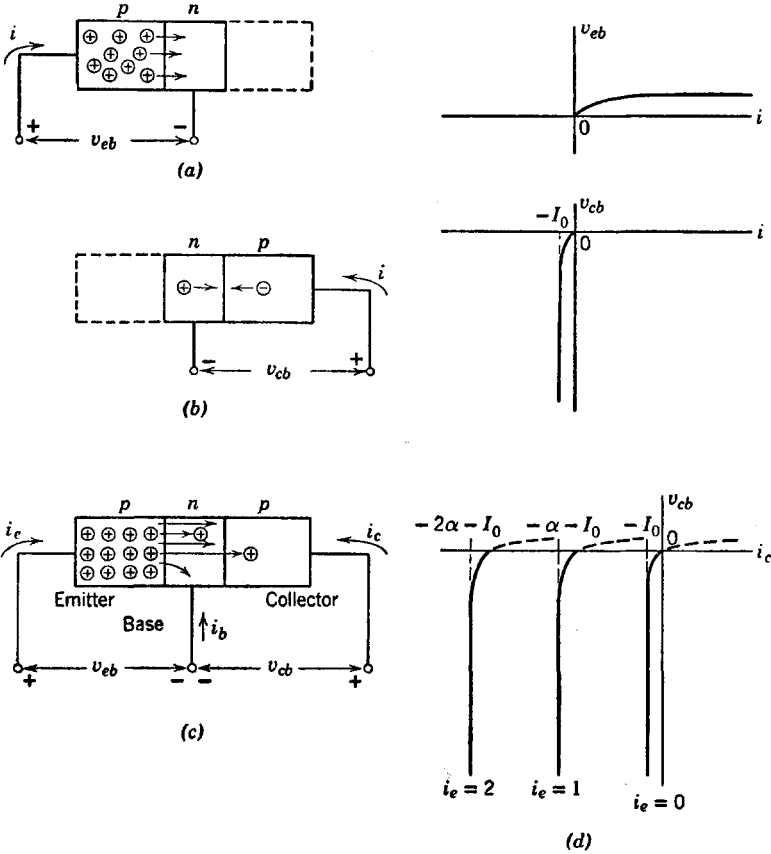
Fig. 5.2. Injection of holes by the emitter in p-n-p type transistors. (a) Forward bias ($v_{eb} > 0$); (b) Reverse bias ($v_{cb} < 0$); (c) p-n-p transistor; (d) Collector curves (v_{cb} vs. i_c).

Combining these two diodes to form a p-n-p transistor, as shown in Fig. 5.2(c), we find that the emitter injects holes into the base across the forward-biased p-n junction. Once across the junction, the holes diffuse through the base. Those that diffuse across to the n-p junction are then in a field of attraction caused by the reverse bias on that junction, and thus flow across to the collector, causing a large increase

in collector current. In most junction transistors, 95 per cent or more of the holes injected into the base from the emitter will reach the collector. The distance across the base is made very small, and in the alloy junction the collector cross-section is often made larger than that of the emitter in order to enhance the collection of the majority carriers.

The sketch of v_{cb} vs. i_c shown in Fig. 5.2(d) is typical of junction-transistor collector curves. If i_e equals 0, the current $-I_0$ is the same as for the diode shown in Fig. 5.2(b). When i_e is greater than 0, we can envisage the majority-carrier injection as providing a current source of magnitude αi_e in the collector circuit, producing a change in current $i_c = -\alpha i_e$. This shifts the v_{cb} vs. $-i_c$ curve parallel to itself, as shown for a number of values of i_e. For each value of i_e the curve of v_{cb} vs. i_c is just like a diode curve with a shift controlled by i_e.

5.3   Electrical Conduction in *n-p-n* Junction Transistor

In *n-p-n* transistors, electrons are the majority carriers; hence, a discussion of these units is like that for the *p-n-p* transistors provided all polarities are reversed and the roles of electrons and holes are interchanged. The possibility of exploiting either a positive or a negative charge carrier provides a great deal of flexibility in transistor circuit design.

Referring to Fig. 5.3 we see that the *n-p-n* unit can be envisioned as two *n-p* junctions, one from emitter to base and the other from collector to base. Reference directions for currents and voltages are the same as those in Fig. 5.2, so that v_{eb} in Fig. 5.3(a) must now be negative to cause the majority carriers (electrons) to migrate freely from emitter to base. Since conventional current (motion of positive charge) is opposite to electron current, the emitter current is negative in this case. To provide reverse bias for the collector junction, the voltage v_{cb} in Fig. 5.3(b) must be positive so that electrons are attracted to the collector after diffusing through the base. Electrons emerging from the collector terminal constitute a current into the terminal; hence, i_c is positive. The polarity changes are indicated by the curves of Fig. 5.3, which should be compared with the corresponding curves in Fig. 5.2.

Ability to deal with both *p-n-p* and *n-p-n* transistors is important in circuit design; but since the difference between them is really one of detail, no purpose is served by duplicate analyses of *p-n-p* and *n-p-n* transistor circuits. In the remainder of the chapter, we shall discuss circuit models and basic circuits for *p-n-p* transistors. The methods and results are readily transferred to *n-p-n* transistors.
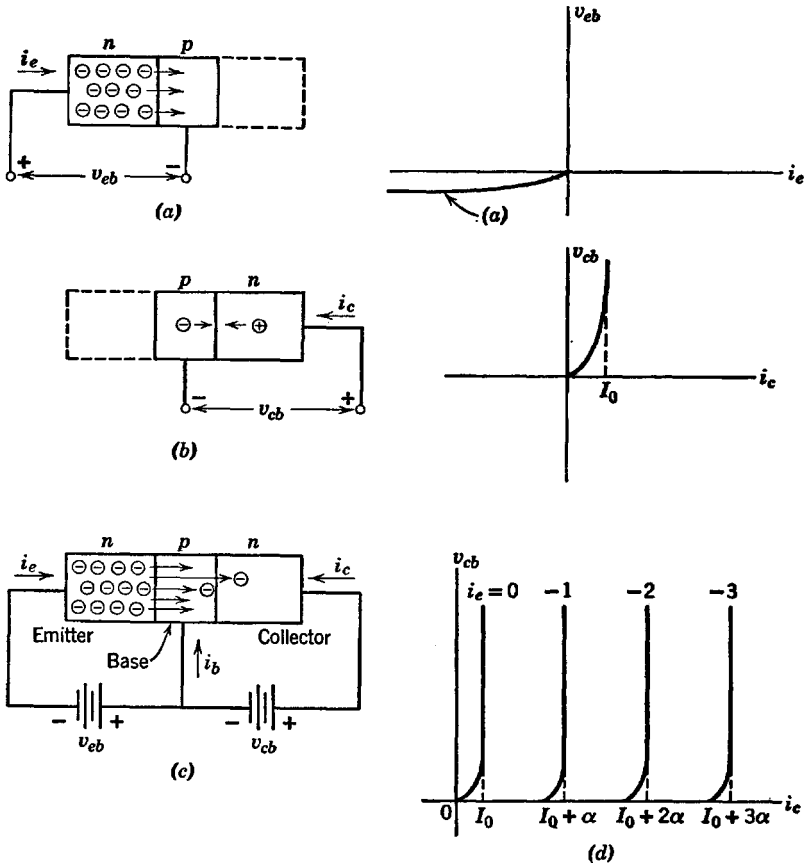
Fig. 5.3. Injection of electrons by the emitter in n-p-n type junction transistors. (a) Forward bias ($v_{eb} < 0$); (b) Reverse bias ($v_{cb} > 0$); (c) n-p-n transistor; (d) Collector curves (v_{cb} vs. i_c).

5.4 Variational Voltage Gain and Power Gain

The foregoing discussion indicated the basic structure of a transistor to be two diodes with a common element, namely the base. Furthermore, the curves relating collector current and collector-to-base voltage are merely diode curves biased or polarized by an amount depending on emitter current. Transistors are therefore transfer or control valves in contrast with diodes, which can be called self-operated valves. The property of a control valve that is of greatest interest is the capability for producing variational power amplification. The resistance r_e between

emitter and base with forward bias applied is likely to be as low as 50 to 100 ohms. On the other hand, the resistance between collector and base when the output circuit is reverse biased may be as large as 1 megohm or more. Thus a series resistor R_c of a few kilohms (say, $10\,\mathrm{k\Omega}$) could be inserted in the collector circuit, as in Fig. 5.4, without appreciably altering the current flow. In this example, the current gain might be 0.95. However, for a large ratio of R_c to R_{in}, the variational voltage gain and the variational power gain are both considerably greater than unity, as can be seen from the expressions in Fig. 5.4.
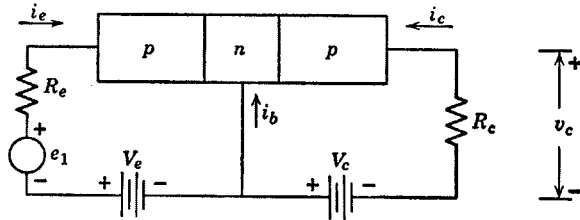
Variational current gain $= \alpha = -\Delta i_c / \Delta i_e$

Input resistance $= R_{in} = \Delta e_1 / \Delta i_e$

Variational voltage gain $= \dfrac{\Delta v_c}{\Delta e_1} \approx \dfrac{R_c \alpha \Delta i_e}{R_{in} \Delta i_e} \approx \dfrac{\alpha R_c}{R_{in}}$

Variational power gain $= \dfrac{-\Delta v_c \Delta i_c}{\Delta e_1 \Delta i_e} \approx \dfrac{R_c \alpha^2 (\Delta i_e)^2}{R_{in} (\Delta i_e)^2} \approx \dfrac{\alpha^2 R_c}{R_{in}}$

Fig. 5.4. Variational gain.

Variational power amplification leads to the more useful functions performed by transistors or other control valves. Voltage or current amplification is often an important function, but the existence of one or both is really an incidental by-product of power amplification. After all, an ordinary linear transformer with a turns ratio of n yields a voltage ratio of n (actually, slightly less) and a current ratio of $1/n$. If n is greater than 1, this corresponds to a voltage gain and current attenuation. If n is less than 1, the transformer has a current gain and voltage attenuation. In either case, the power gain $(e_2 i_2 / e_1 i_1)$ from primary to secondary is less than unity and approaches unity only for an ideal transformer. The power amplification attributed to transistors (or other control valves, such as vacuum triodes and pentodes) must satisfy the conservation-of-energy principle. Power must be supplied to the transistor and its associated circuit by means of bias-voltage sources. The transistor (or any other control valve) merely permits a small source of power to control a large source of power. Thus, the

input-to-output power gain exists in the variational sense but not in terms of total voltages and currents.

5.5   Ideal-Diode Model for *p-n-p* Junction Transistor

Let us postulate a simple circuit model to approximate the behavior of the *p-n-p* transistor discussed in Article 5.2.   The model shown in
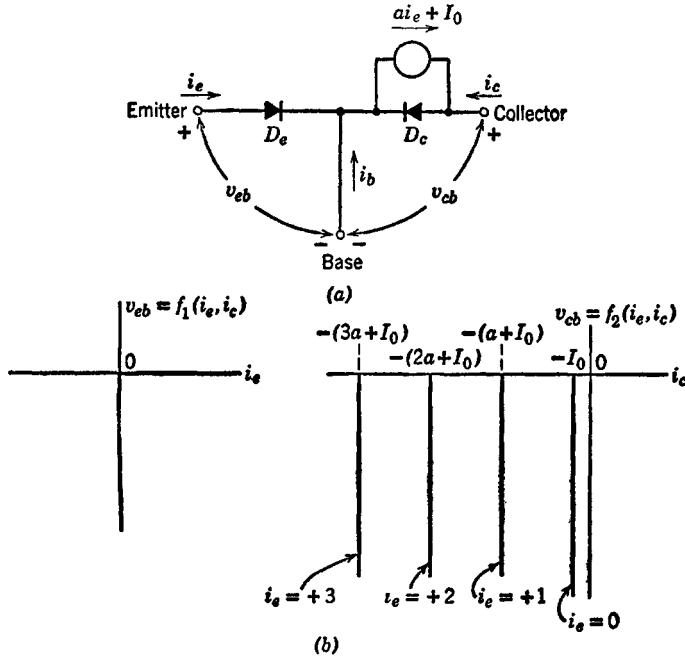
(a)

(b)

Note: Since $i_e + i_b + i_c = 0$, we have the following relations for D_e closed and D_c open:

$$i_c = -I_0 - ai_e, \qquad\qquad \Delta i_c/\Delta i_e = -a$$

$$i_b = I_0 - (1-a)i_e, \qquad \Delta i_e/\Delta i_b = -1/(1-a)$$

$$i_c = (-I_0 + ai_b)/(1-a), \quad \Delta i_c/\Delta i_b = a/(1-a)$$

(c)

Fig. 5.5.   Ideal-diode model for *p-n-p* junction transistor.

Fig. 5.5(*a*), based on the physical structure of the *p-n-p* junction transistor, contains two ideal diodes, one from emitter to base and one from collector to base.   The dependent current source ai_e accounts for the influence of the emitter current on the collector current.   The constant a

is chosen to match the variational current gain α ($\partial i_c/\partial i_e$ for v_{cb} constant) of the transistor being represented by the model. The model has a constant current of magnitude I_0 to represent the small reverse current in the collector diode. The constant current I_0 will be omitted from
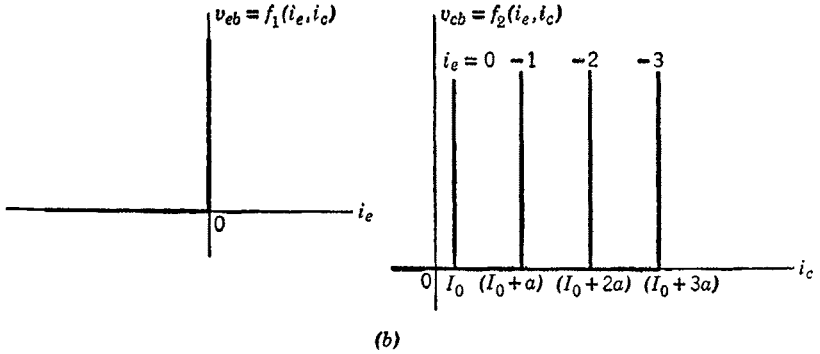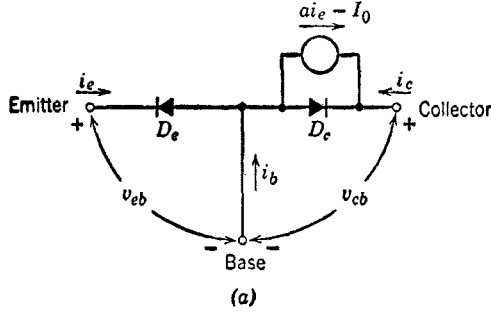
(a)

(b)

Note: Since $i_e + i_b + i_c = 0$, we have the following relations for D_e closed and D_c open:

$$i_c = +I_0 - ai_e, \qquad \Delta i_c/\Delta i_e = -a$$

$$i_b = -I_0 - (1-a)i_e, \qquad \Delta i_e/\Delta i_b = -1/(1-a)$$

$$i_c = (+I_0 + ai_b)/(1-a), \qquad \Delta i_c/\Delta i_b = a/(1-a)$$

(c)

Fig. 5.6. Ideal-diode model for n-p-n junction transistor.

the dependent generator for the sake of simplicity in some of the following articles. The piecewise-linear curves shown in Fig. 5.5(b), readily derived from the model by diode break-point analysis, are remarkably similar to typical transistor collector curves, such as those shown in Appendix A. The normal operating state of the transistor occurs when the emitter diode D_e is closed and the collector diode D_c is open. This state represents a region called the amplification region on the v_{cb} vs. i_c

plane, which for this model is the entire third quadrant to the left of $-I_0$ ($i_c < -I_0$, $v_{cb} < 0$). If D_e is opened, the transistor is said to be cut off, and in this simple model the cut-off region degenerates to the line $i_e = 0$. Similarly, the saturation state, when D_c is closed, is here represented by the $v_{cb} = 0$ axis. The corresponding ideal-diode model of an n-p-n transistor, and its curves, are shown in Fig. 5.6.
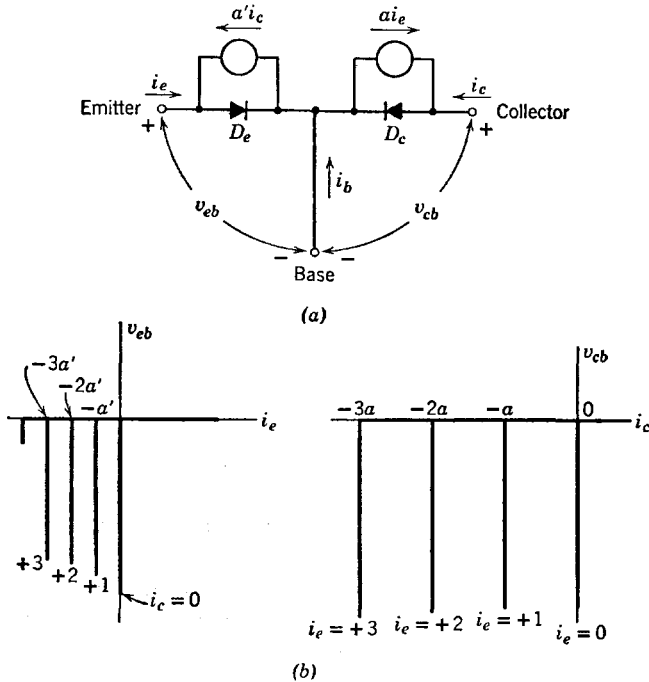
Fig. 5.7. Ideal-diode model with two dependent generators (neglecting reverse diode current I_0).

These ideal-diode models provide a convenient way of remembering the ideal current gains between various terminal pairs when the transistor is operating in the normal amplification state (emitter diode closed and collector diode open). As indicated in Fig. 5.5(c) and Fig. 5.6(c), the relations between the various currents lead directly to expressions for variational current gain. When resistive parameters are included in the circuit model, the current gains are smaller, but the ideal values given here provide limits for checking the more involved expressions.

Thus far we have considered positive emitter currents that cause holes to diffuse through the base to the collector. The physical structure of the transistor suggests the possibility of interchanging the roles of

emitter and collector by sending the hole current through the device in the opposite direction. If emitter and collector are made identical, as they are in a "symmetrical" transistor, the emitter and collector can be interchanged at will. Most transistors, however, are designed to have a much smaller current gain when emitter and collector are interchanged. Whether the transistor is symmetrical or not, it is apparent that the model shown in Fig. 5.7(a) is more general than the previous ones, since it includes the effect of collector current on the emitter current. The corresponding curves are shown in Fig. 5.7(b). A model of this type is useful when the voltages or currents applied to the device are likely to have large amplitudes of arbitrary polarity.

5.6   Ideal-Diode Model for Common Emitter Connection

In the models of Figs. 5.5, 5.6, and 5.7, the voltages are referred to the base. The plots can be expressed in functional notation as $v_{eb} = f_1(i_e, i_c)$ and $v_{cb} = f_2(i_e, i_c)$, a convenient representation for transistor circuits in which the base is the terminal common to input and output. However, in many transistor circuits, the emitter is the common terminal. Although the curves in the previous figures contain all the required information and thus can still be used for this "common emitter" connection, it may be more instructive to present the data in another form, namely base-to-emitter voltage versus base current and collector-to-emitter voltage versus collector current. These curves are shown in Fig. 5.8(b). It is desirable to redraw the model to the form shown in (a) to emphasize the fact that the base current i_b is now the control parameter. To calculate the current gain, the ratio of collector current to base current, assume that the emitter current is 1 ma. Then if D_c is open and D_e closed (forward gain or amplification region) the collector current is $-a$ ma, and, by summing currents at the common node, the base current must be $-(1 - a)$ ma. Thus the current gain for the common emitter connection is $a/(1 - a)$, which for a close to unity can be a large number. For example, if a is 0.98, the current gain is 49.

The transistor states, as defined in the previous article, are located on the v_{ce} versus i_c plane as follows:

Forward gain or amplification	Third quadrant
Saturation	$v_{ce} = 0$ line
Cutoff	$i_b = 0$ line, $v_{ce} < 0$

The figure shows a set of lines (dotted in the first quadrant), which
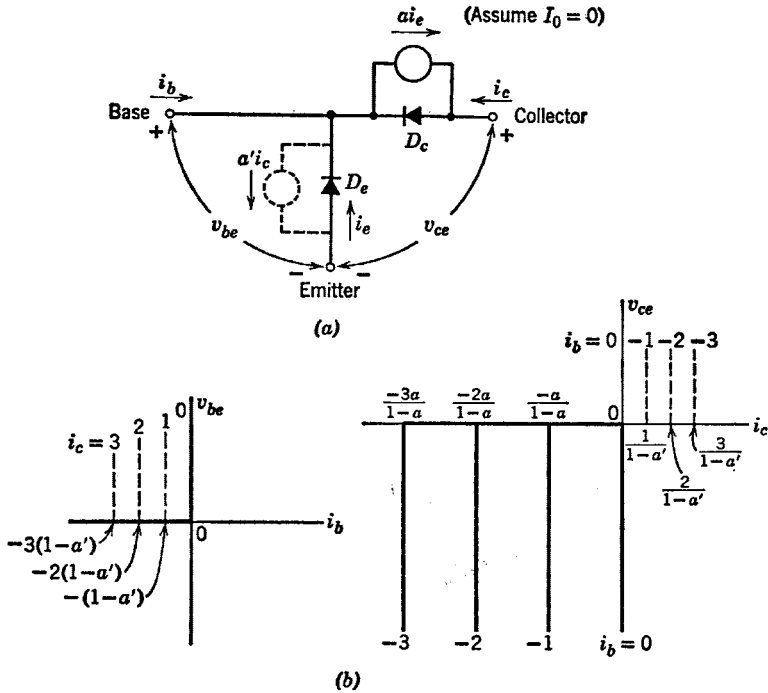
Fig. 5.8. Simple ideal-diode model for common emitter connection of p-n-p junction transistors (neglecting reverse current I_0).

have not appeared in previous curves. These lines represent the region where the reverse-current generator $a'i_c$ comes into play.

5.7 Effect of a Coupling Resistance in the Transistor Model

Making better resistive approximations to the junctions in a transistor tends to improve the accuracy of the model. Thus, we might replace each ideal diode by the exponential approximation to a junction diode as in Chapter 2. There are many cases in which an exponential diode model is useful. However, the linear-circuit methods are so powerful that we shall emphasize piecewise-linear models in order to use linear circuit theory in each diode state. For examining general circuit behavior, the accuracy will be adequate with a two-segment piecewise-linear approximation to the junctions. Resistive approximations for the diodes give the corresponding curves finite and nonzero slopes as might be expected by analogy to the diode models discussed in Chapter 3.

However, before including resistive diodes, let us consider first the effect of base resistance alone.

In the simple common-base transistor model, a current in the emitter circuit can influence the collector circuit only by means of the dependent current generator. In an actual transistor the ohmic resistance of the

(a)

(b)
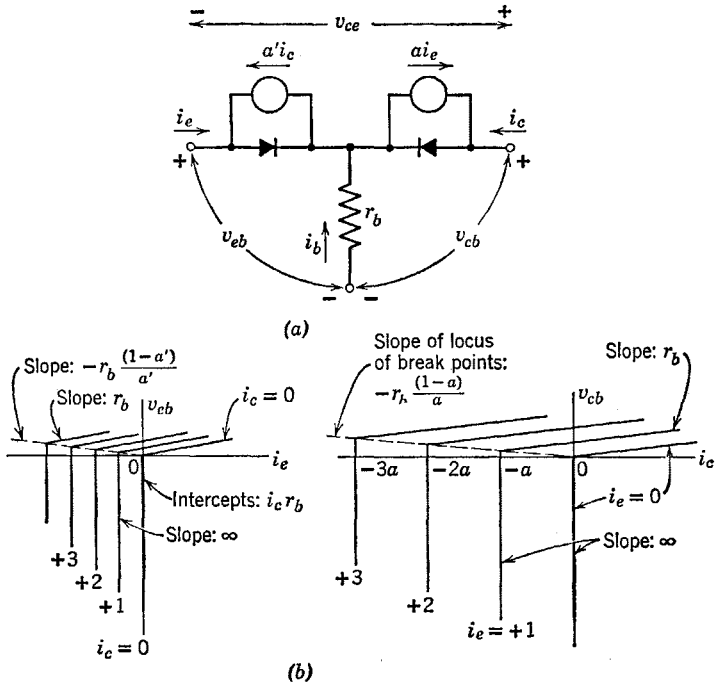
Fig. 5.9. Transistor model with ideal diodes and base resistance.

base material, 50 to 100 ohms for most transistors, causes a resistive coupling effect between the two circuits. Including a base resistance in the ideal-diode model as shown in Fig. 5.9(a) leads to the curves in (b) for the common-base connection. The collector curves are a reasonable approximation to the collector curves for an actual transistor.

If the polarizing voltages together with the applied signal do not drive the transistor into all of the diode states, it is well to use the less complete form for the model. Thus, if collector current is never likely to assume positive values in a given application, the $a'i_c$ generator can be omitted, since for negative values of i_c the $a'i_c$ generator is shorted by the ideal diode. Another way of saying this is that the two current generators ai_e and $a'i_c$ can usually be considered one at a time. If both i_e and i_c are positive, we simply have two diodes using the same

base, and the dependent generators are both shorted by the ideal diodes in the model.

(a)

(b)
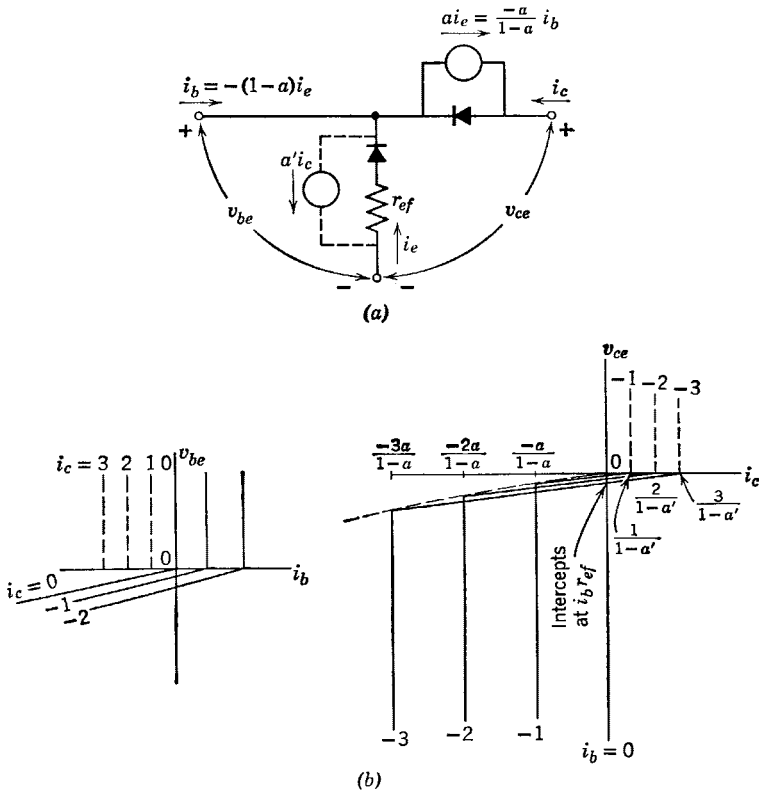
Fig. 5.10. Model for common-emitter connection with nonzero emitter-diode forward resistance.

Resistive coupling between the input and output circuits in the common emitter connection is caused by the resistance of the emitter diode. For simplicity let us add to the ideal diode model only a forward resistance r_{ef} in the emitter diode as shown in Fig. 5.10(a). The corresponding input and output curves are shown in (b).

5.8   Piecewise-Linear Resistive Model

A piecewise-linear resistive approximation to each of the semiconductor junctions in a p-n-p transistor yields the model shown in

(a)

(b)

(c)
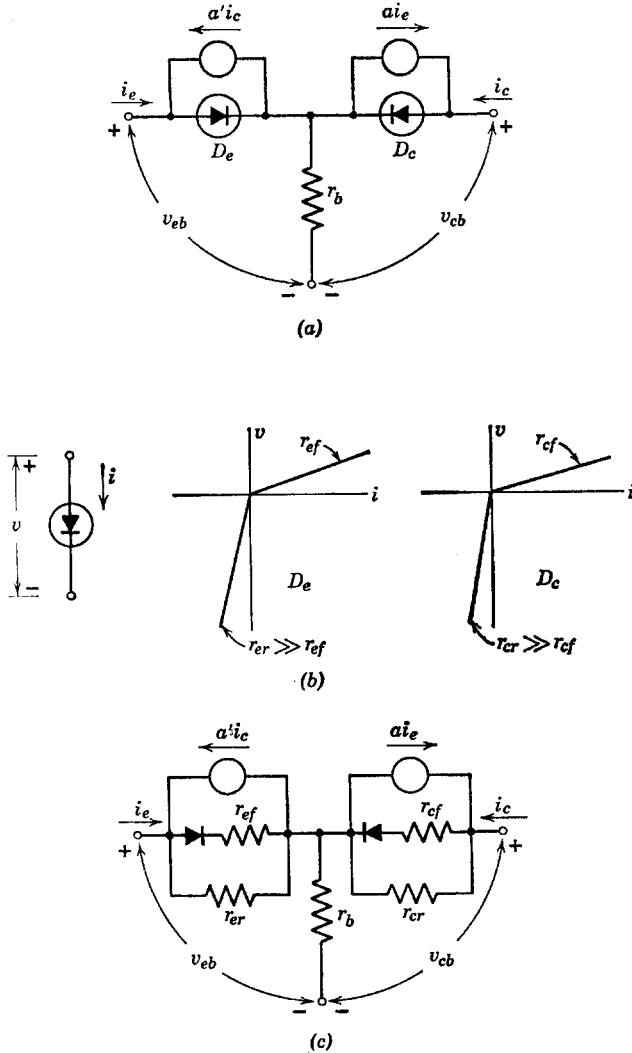
Fig. 5.11. Transistor model using piecewise-linear resistive diodes.

Fig. 5.11(a). The piecewise-linear diodes D_e and D_c are specified by forward and reverse resistances as indicated in Fig. 5.11(b). Typical values of forward resistances are 50 to 200 ohms, whereas reverse resistances are of the order of a megohm. The numerous parameters available in this model afford more than enough flexibility to approximate a set of transistor curves very well. In fact, the complete model

shown in Fig. $5.11(c)$ is more involved than necessary for most circuit analysis or design problems.
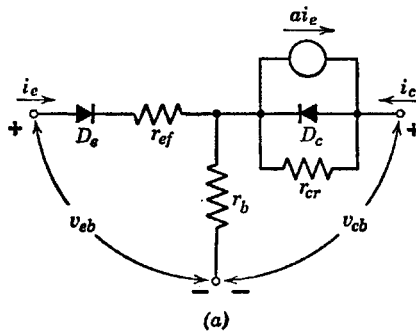
States	D_e	D_c
I Cutoff	Off	Off
II Forward Gain	On	Off
III Saturation	On	On
IV *	Off	On
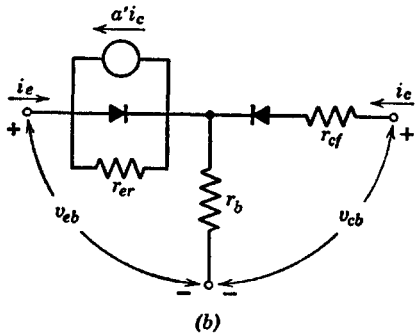
Fig. 5.12. Simplified piecewise-linear resistive model.

The analytical complexity encountered with the general model can be reduced by using a model in which the emitter and the collector diodes are made semi-ideal. As shown in Fig. $5.12(a)$ only forward emitter resistance and reverse collector resistance are included. To solve a given circuit problem, one should choose the simplest possible model that will give the required accuracy. Often the best approach is to solve the problem first with the ideal-diode model, and then add

---

* This is the inverted-gain state which corresponds to State II with emitter and collector interchanged. Use model of Fig. 5.9 or $5.12(b)$.

appropriate resistances to the model and refine the solution. Similarly, the model with the ai_e generator alone is convenient if collector current is always negative, but if i_c is positive at any time, the $a'i_e$ generator must also be included.
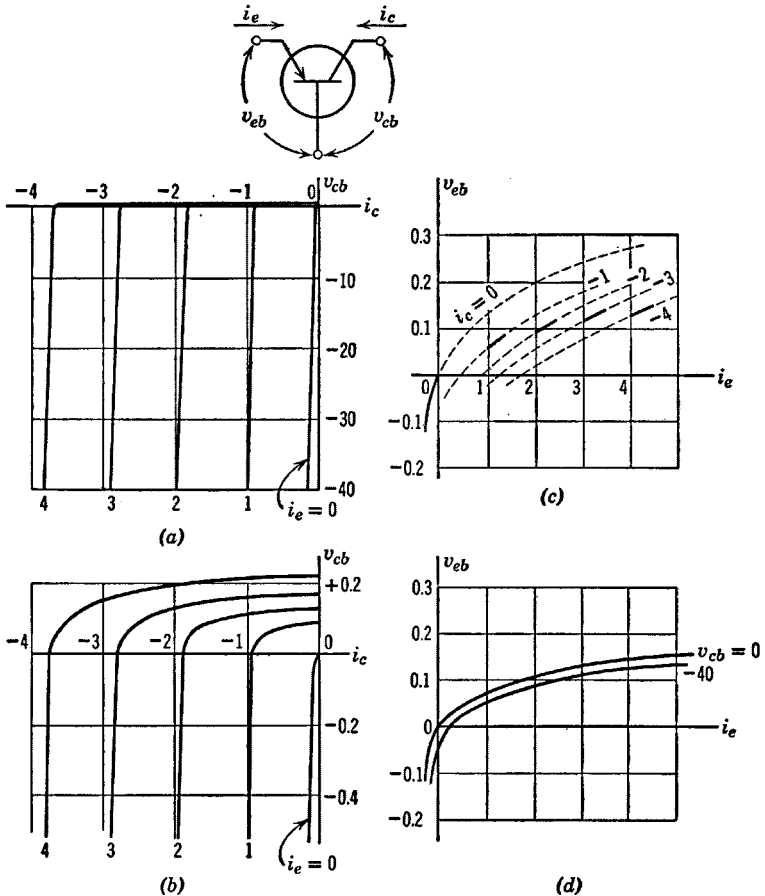
Fig. 5.13.   Typical curves for a p-n-p junction transistor for common-base connection.

5.9   Approximating Transistor Curves

Typical measured curves specifying the "static" or resistive terminal behavior of a p-n-p junction transistor are shown in Fig. 5.13. This is only one of a number of forms in which a manufacturer may present average characteristics. The family of collector curves shown in Fig. 5.13(a) is repeated in (b) with an expanded voltage scale to emphasize

the exponential character of the curves for small values of collector voltage. From the expanded scale it is apparent that each curve in the collector family resembles the curve for a semiconductor diode, with successive curves shifted along the current axis by an amount corresponding approximately to the majority-carrier injection provided by the emitter.
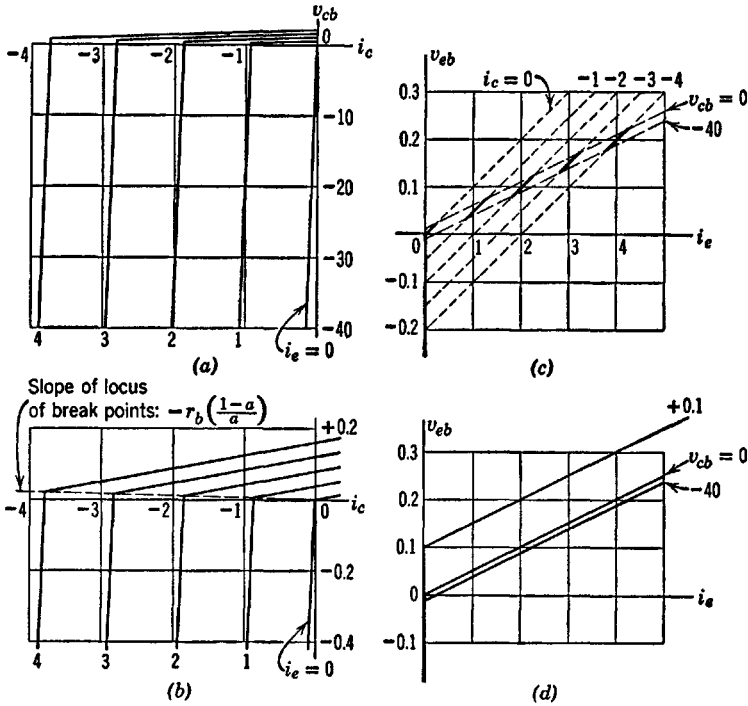
Fig. 5.14. Approximate curves for a *p-n-p* transistor using the model of Fig. 5.12, with $r_{ef} = 50$ ohms, $r_b = 50$ ohms, $r_{cr} = 200$ kΩ, $a = 0.95$.

Since two voltages and two currents are needed to specify the terminal behavior of the transistor, another plot is necessary. Either the plot in Fig. 5.13(c) or that in (d) links the fourth variable v_{eb} to the other three. The graphical data can also be presented in a variety of other forms, some of which will be introduced later.

Comparing the typical collector curves of Fig. 5.13(a) with those for the ideal diode model in Fig. 5.5(b), we see the possibility of reasonable quantitative agreement. The inclusion of base resistance in the ideal-diode model makes the collector family approximation somewhat better [compare curves in Fig. 5.13(b) and Fig 5.9(b)].

The piecewise-linear curves in Fig. 5.14 are plotted for the model in Fig. 5.10 using the following values: $r_{ef} = 50$ ohms, $r_b = 50$ ohms, $r_{er} = 200$ kilohms and $\alpha = 0.95$. The corresponding parts of Figs. 5.13 and 5.14 are seen to be in very close agreement except in the vicinity
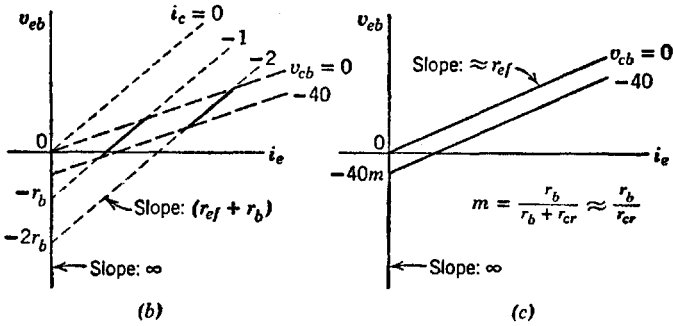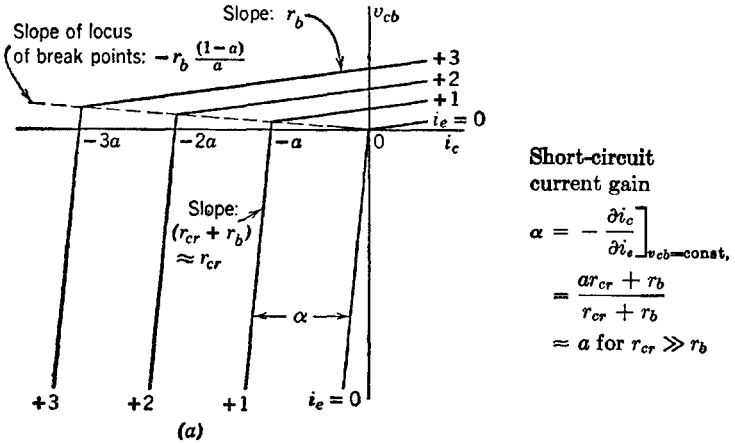
Fig. 5.15. Slopes and intercepts of common-base curves for p-n-p transistor, using the piecewise-linear model of Fig. 5.10. (a) v_{cb} vs. i_c, parameter i_e; (b) v_{eb} vs. i_e, parameter i_c; (c) v_{eb} vs. i_e, parameter v_{cb}.

of break points. This model then affords a fairly good approximation of the resistive behavior of a transistor. The heavy lines in Fig. 5.14(c) show the normal range of operation for the unit being described.

The curves of Fig. 5.14 were plotted from a break-point analysis of the model shown in Fig. 5.12(a). The curves are redrawn in Fig. 5.15 without regard to relative numerical values in order to show the general values of slopes and intercepts. These plots show the effects of r_b, r_{ef}, r_{cr},

and a, so that suitable values of these parameters can be chosen to approximate a set of actual curves.

An alternative method of analyzing circuits that contain models like those of Figs. 5.11 or 5.12 emphasizes the resistive character of the junction approximations. The model shown in Fig. 5.16(a) is purely resistive and, hence, applies in any one diode state when the appropriate
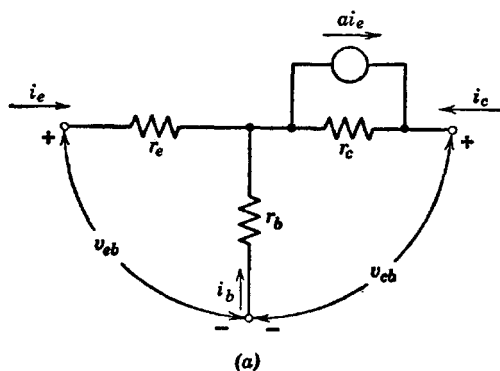
Note:

$r_e = r_{ef}$ or r_{er}

$r_c = r_{cr}$ or r_{cf}

(a)

$$v_{eb} = i_e(r_e + r_b) + i_c r_b$$

$$v_{cb} = i_e(r_b + ar_c) + i_c(r_c + r_b)$$

(b)

Fig. 5.16. Resistive representation of piecewise-linear model.

values are used for r_e and r_c, namely, r_{ef} or r_{er} and r_{cf} and r_{cr}. A single current generator is shown in the interests of simplicity. The need for including $a'i_c$ and removing ai_e in state IV is no more difficult to remember than is the appropriate value to be used for each parameter. Actually, when the forward diode resistances are nonzero, omission of the second generator modifies the results somewhat. The resistive model of Fig. 5.16(a) is convenient to use because the same equations apply to all of the circuit states.

5.10   Total and Incremental Models

The voltages and currents on graphical plots describing an electronic device like a transistor are total quantities. Thus, such symbols as i_e or v_{cb} refer to the total instantaneous values of the variables. However, it is often possible to base an analysis on only the variational portions of the currents and voltages. Circuit models for calculating only

variational currents and voltages are called incremental models or small-signal models.

Although the form of the model for any one state of a piecewise-linear approximation may be similar to that of an incremental model, the

(a)

$$v - V_a = + r(i - I_a)$$
$$\Delta v = r\Delta i$$

$$v = f(i)$$
$$= V_a + \frac{dv}{di}\Big|_a (i - I_a) + \cdots$$

(b)

$$v = f(i)$$
$$(V_a + \Delta v) = f(I_a + \Delta i)$$

(c)

$$(v - V_a) = r(i - I_a)$$
$$\Delta v = r\Delta i$$

(d)

$$v = E_0 + ri$$
$$V_a + \Delta v = E_0 + r(I_a + \Delta i)$$
where $r = \Delta v/\Delta i$
and $E_0 = (V_a - rI_a)$

(e)

For
$$i > 0, \qquad v = E_0 + ri$$
$$v < E_0, \qquad i = 0$$
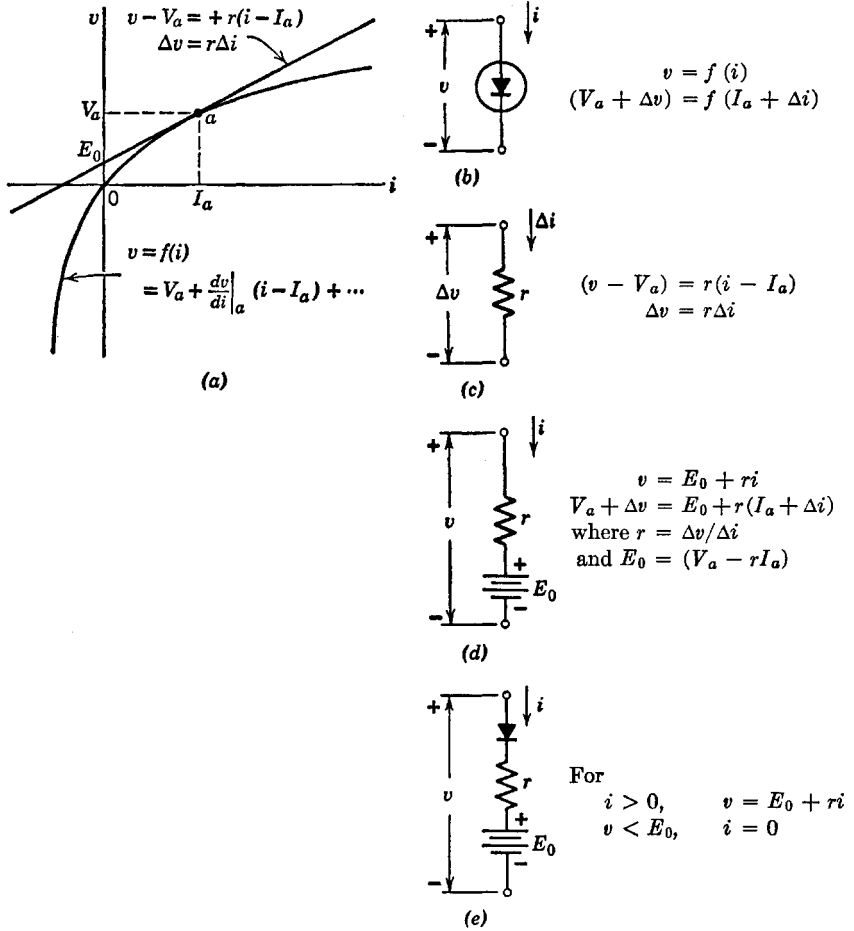
Fig. 5.17. Total and incremental circuit models for a nonlinear function of one variable, $v = f(i)$.

numerical values of the parameters may or may not be the same. The major difference in the two models stems from the use of total quantities in one and variational quantities in the other.

The mathematical basis for the incremental model is the Taylor

Series expansion for a nonlinear curve. The constant term fixes the "operating point" since this is the point about which the expansion is made. The first-order term gives the slope of the curve at the operating point and, hence, determines the change in current produced by a given change in voltage. The ratio $\Delta v/\Delta i$ is called the incremental resistance. For small signals the higher-order terms can be neglected.

The foregoing remarks are illustrated in Fig. 5.17. A nonlinear diode curve is shown in (a), together with the Taylor Series expansion about the point a and the expression for the tangent. The diode symbol is shown in (b) with voltages and currents expressed as variations or increments around the operating point. The incremental model shown in (c) effectively amounts to a shift of coordinates from the origin to the quiescent point. Once the circuit is reduced to linear incremental form, the symbol Δ has no significance, since the linear relation extends indefinitely.

The linear incremental relation between current and voltage effectively drops the fixed reference potential whereas a linear model based on total quantities preserves a unique origin of coordinates as indicated by Fig. 5.17(d). A piecewise-linear model like that in (e) places in evidence the range of values of v and i for which a linear approximation applies.

5.11   Linear Incremental Models for Transistors

The terminal behavior of any three-terminal device is specified by two voltages and two currents. As shown in Fig. 5.18(a), the general function is usually separated into two functions for convenience in graphical or analytical manipulations. If each of the variables is expressed as a constant plus an increment, the constant values of the independent variables determine the constant values of the dependent variables. This assumes the increments to be zero. The differentials of the equations relate the variational components as shown in Fig. 5.18(b). If the circuit is linear, the various partial derivatives are constants; that is, independent of the operating point set by the direct components of currents and voltages. These results are merely extensions of those given in Fig. 5.17 for the simpler function $v = f(i)$.

For the linear incremental three-terminal circuit, there are many possible ways of specifying the relations between currents and voltages, but all involve four constant parameters. For the choice of variables used here there are six basic forms the equations can take as indicated in Fig. 5.19. All of these are related by simple linear transformations. As a matter of convenience the symbol Δ is omitted in Fig. 5.19, leaving

the symbols for incremental voltages and currents the same as those used for total quantities. If both total and incremental values are involved in an analysis, primes or deltas or other distinguishing symbols must be used to avoid ambiguity.
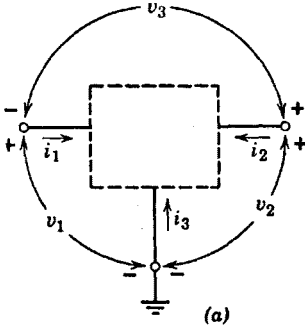
$$f(v_1, v_2, i_1, i_2) = 0$$

or

$$\begin{cases} v_1 = f_1(i_1, i_2) \\ v_2 = f_2(i_1, i_2) \end{cases}$$

(a)

$$\text{For} \begin{cases} v_1 = V_1 + \Delta v_1 \\ v_2 = V_2 + \Delta v_2 \\ i_1 = I_1 + \Delta i_1 \\ i_2 = I_2 + \Delta i_2 \end{cases}, \qquad \begin{aligned} V_1 + \Delta v_1 &= f_1\left[(I_1 + \Delta i_1), (I_2 + \Delta i_2)\right] \\ V_2 + \Delta v_2 &= f_2\left[(I_1 + \Delta i_1), (I_2 + \Delta i_2)\right] \end{aligned}$$

Letting $\Delta i_1 = \Delta i_2 = 0$
so that $\Delta v_1 = \Delta v_2 = 0$ yields $\begin{aligned} V_1 &= f_1(I_1, I_2) \\ V_2 &= f_2(I_1, I_2) \end{aligned}$

Incrementally,

$$\Delta v_1 = \frac{\partial f_1}{\partial i_1} \Delta i_1 + \frac{\partial f_1}{\partial i_2} \Delta i_2$$

$$\Delta v_2 = \frac{\partial f_2}{\partial i_1} \Delta i_1 + \frac{\partial f_2}{\partial i_2} \Delta i_2$$

(b)

Fig. 5.18. General equations for a three-terminal device.

The circuit models for the first two sets of equations in Fig. 5.19 involve driving-point resistances or conductances (r_{11}, r_{22}, g_{11}, g_{22}) and transfer resistances or conductances (r_{12}, r_{21}, g_{12}, g_{21}). The remaining sets involve parameters in various other categories. For instance, the h parameters involve a driving-point resistance h_{11}, reverse voltage gain h_{12}, forward current gain h_{21}, and a driving-point conductance h_{22}. These linear circuit formulations are not restricted to resistances and conductances, but also apply to impedances and admittances.

In Fig. 5.20 the resistive circuit of Fig. 5.16 is used as an incremental model of a transistor in the common-base connection. Since the model contains no d-c sources, the circuit and the equations are valid for either total or incremental calculations. Identification of corresponding

(a)

$$\begin{cases} v_1 = r_{11}i_1 + r_{12}i_2 \\ v_2 = r_{21}i_1 + r_{22}i_2 \end{cases}$$

(b)

$$\begin{cases} i_1 = g_{11}v_1 + g_{12}v_2 \\ i_2 = g_{21}v_1 + g_{22}v_2 \end{cases}$$

(c)

$$\begin{cases} v_1 = h_{11}i_1 + h_{12}v_2 \\ i_2 = h_{21}i_1 + h_{22}v_2 \end{cases}$$

(d)

$$\begin{cases} v_2 = k_{11}v_1 + k_{12}i_2 \\ i_1 = k_{21}v_1 + k_{22}i_2 \end{cases}$$

(f)

$$\begin{cases} v_1 = Av_2 - Bi_2 \\ i_1 = Cv_2 - Di_2 \end{cases}$$

(e)

$$\begin{cases} v_2 = Dv_1 - Bi_1 \\ i_2 = Cv_1 - Ai_1 \end{cases}$$

(g)
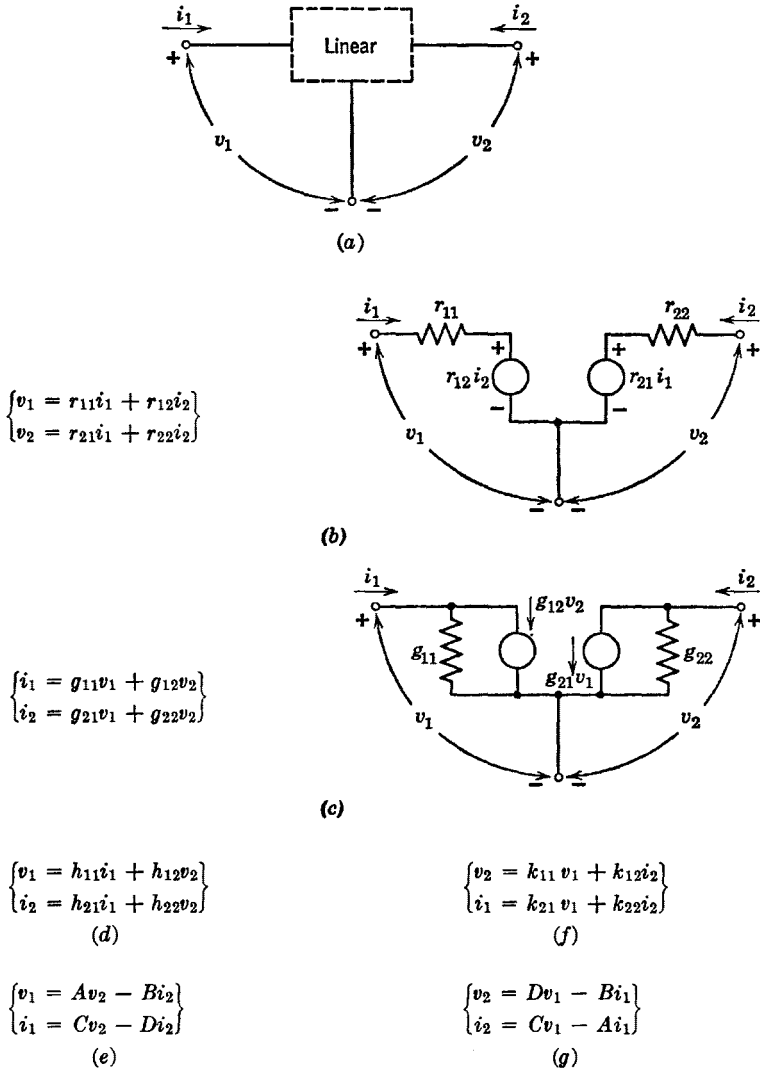
Fig. 5.19. Incremental models for a three-terminal device.

terms in the impedance formulation and in the T circuit relates the generalized parameters (r_{11}, r_{12}, r_{21}, r_{22}) to the T parameters (r_e, r_b, r_c, a).

Several formulations of the incremental model for the common-emitter connection are shown in Fig. 5.21 together with the corresponding equations. The circuit in (a) is the same as that of Fig. 5.20, but the

equations involve the common-emitter variables $(v_{be}, v_{ce}, i_b, i_c)$. The circuit and equations in (b) result when the dependent generator is specified in terms of base current. The model in (c) involves the so-called hybrid parameters or h parameters. One current and one voltage $(i_b$ and $v_{ce})$ are independent variables and the other current and voltage

$$\begin{cases} v_1 = r_{11}i_1 + r_{12}i_2 \\ v_2 = r_{21}i_1 + r_{22}i_2 \end{cases} \qquad \begin{cases} v_{eb} = i_e(r_e + r_b) + i_cr_b \\ v_{cb} = i_e(r_b + ar_c) + i_c(r_c + r_b) \end{cases}$$

$$r_{11} = r_e + r_b \qquad\qquad r_e = r_{11} - r_{12}$$

$$r_{12} = r_b \qquad\qquad r_b = r_{12}$$

$$r_{21} = r_b + ar_c \qquad\qquad r_c = r_{22} - r_{12}$$

$$r_{22} = r_c + r_b \qquad\qquad a = \frac{r_{21} - r_{12}}{r_{22} - r_{12}}$$
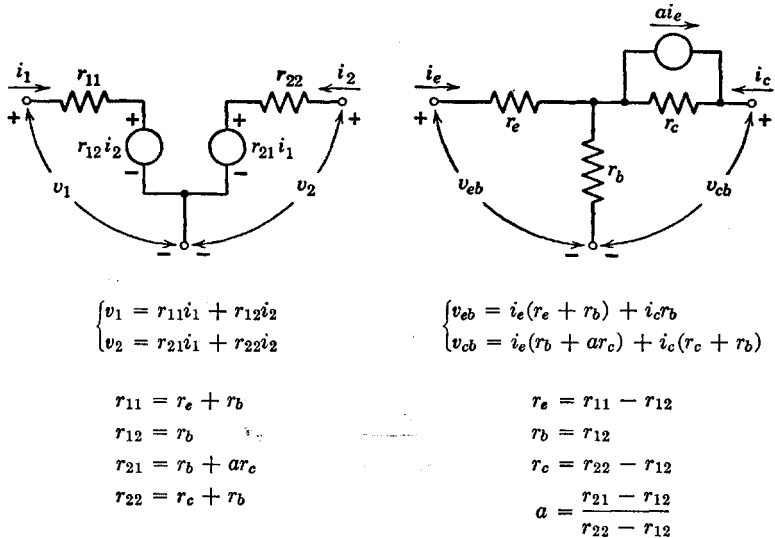
Fig. 5.20. Incremental models for a transistor in the common-base connection.

(v_{be}, i_c) are dependent variables. For typical numerical values $(r_c \gg r_e$ or $r_b)$ the hybrid parameters are as follows:

$$h_{11} \approx r_b + r_e/(1 - a) \tag{5.1}$$

$$h_{12} \approx r_e/r_d \tag{5.2}$$

$$h_{21} = \alpha_{cb} \approx a/(1 - a) \tag{5.3}$$

$$h_{22} = 1/(r_e + r_d) \tag{5.4}$$

The h parameter formulation for the common-emitter connection is not as obviously related to the physical conduction processes in the transistor as is the common-base T formulation involving $r_e, r_b, r_c,$ and a. However, the h parameters can be measured conveniently and accurately. The current gain is significant, but the other three constants are small and can often be neglected. For circuit design purposes, the h parameters

are very convenient. Each of the parameters in this formulation places in evidence a quantity significant in transistor amplifier circuits; namely,
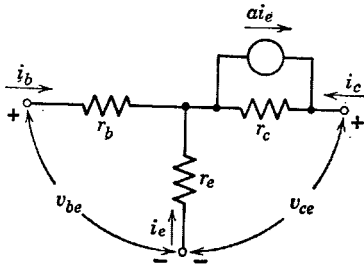
$$v_{be} = i_b(r_b + r_e) + i_c r_e$$
$$v_{ce} = i_b(r_e - ar_c) + i_c[r_e + r_c(1 - a)]$$

(a)
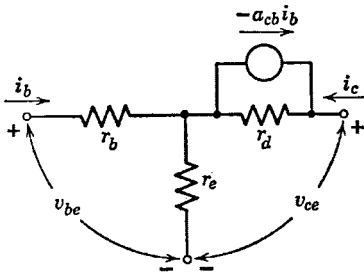
$$v_{be} = i_b(r_b + r_e) + i_c r_e$$
$$v_{ce} = i_b(r_e - a_{cb}r_d) + i_c(r_e + r_d)$$
Where $r_d = r_c(1 - a)$
$$a_{cb} = a/(1 - a)$$

(b)

$$v_{be} = i_b\left[r_b + \frac{r_e r_c}{r_e + r_c(1 - a)}\right] + v_{ce}\left[\frac{r_e}{r_e + r_c(1 - a)}\right]$$

$$i_c = i_b\left[\frac{-r_e + ar_c}{r_e + r_c(1 - a)}\right] + v_{ce}\left[\frac{1}{r_e + r_c(1 - a)}\right]$$
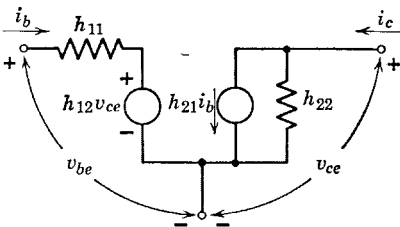
$$h_{11} \approx r_b + r_e/(1 - a)$$
$$h_{12} = r_e/r_d$$
$$h_{21} = \alpha_{cb} \approx a/(1 - a)$$
$$h_{22} = 1/(r_e + r_d)$$

(c)

Fig. 5.21. Incremental models for a transistor in the common-emitter connection.

an input impedance ($h_{11} = \partial v_{be}/\partial i_b$ with v_{ce} fixed), a reverse voltage gain ($h_{12} = \partial v_{be}/\partial v_{ce}$ with i_b fixed), a forward current gain ($h_{21} = \partial i_c/\partial i_b$ with v_{ce} fixed), and an output admittance ($h_{22} = \partial i_c/\partial v_{ce}$ with i_b fixed). Most manufacturers' data on transistors are presented in these terms.
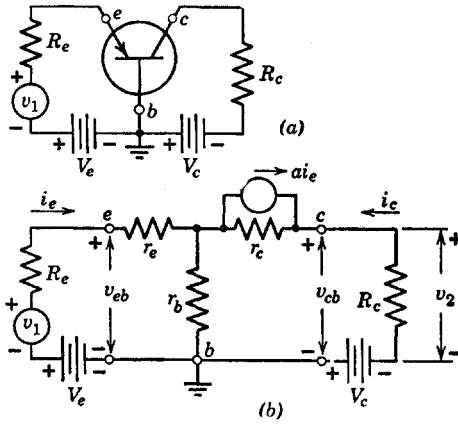
Assume: $a = 0.95$

$r_b = 50$ ohms

$r_{ef} = 50$ ohms

$r_{er} = \infty$

$r_{cr} = 200$ k Ω

$r_{cf} = 0$

$$\begin{cases} v_{eb} = i_e(r_e + r_b) + i_c r_b \\ v_{cb} = i_e(r_b + a r_c) + i_c(r_e + r_b) \end{cases}$$

$$\begin{cases} v_{eb} = V_e + v_1 - i_e R_e \\ v_{cb} = -V_e - i_c R_c \end{cases}$$

(c)
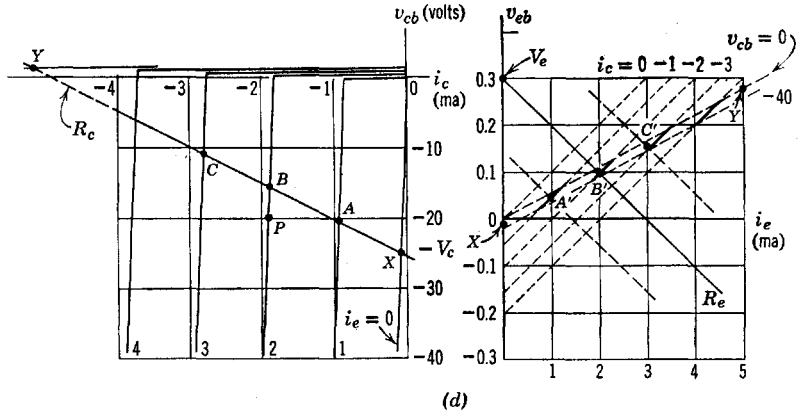
Fig. 5.22. Determination of operating point in common-base circuit.

5.12 Determination of Operating Point in the Common-Base Circuit

A common-base resistive transistor circuit is shown in Fig. 5.22(a). The piecewise-linear model is shown in (b); and since this is a total model, the variables v_1, v_2, v_{eb}, i_e, etc. represent total quantities. The

element values are the same as those assumed in Fig. 5.14. To determine the operating point of the transistor, we can write two equations pertaining to the input terminals. One expresses the constraint on terminal voltage and current imposed by the external circuit, and the other expresses the constraints imposed by the transistor. A similar pair of equations can be written for the output circuit. For the circuit in (b) the equations relating the terminal voltages to terminal currents in terms of the transistor parameters are:

$$v_{eb} = i_e(r_e + r_b) + i_c r_b \qquad (5.5)$$

$$v_{cb} = i_e(r_b + ar_c) + i_c(r_c + r_b) \qquad (5.6)$$

The equations relating the terminal voltages to currents in terms of the external circuit parameters are

$$v_{eb} = V_e + v_1 - i_e R_e \qquad (5.7)$$

$$v_{cb} = -V_c - i_c R_c \qquad (5.8)$$

Equations 5.5, 5.6, 5.7, and 5.8 can be solved simultaneously to find the four coordinates of the operating point.

Operating points of actual transistor circuits are usually determined graphically using the nonlinear characteristic curves for the device, but to illustrate the procedure, we will perform the graphical construction on the piecewise-linear curves of Fig. 5.22(d), which correspond to the model in (b). The output load equation (Eq. 5.8) is plotted on the v_{cb} vs. i_c plane, and the input load equation (Eq. 5.7) is plotted on the v_{eb} vs. i_e plane. The points A', B', and C' on the input plane correspond to A, B, and C on the output plane, so a line through A', B', and C' corresponds to the output load line transferred to the input plane. An exact solution of the operating point problem requires finding the intersection in the input plane of the input load line with the transferred output load line. However, the transistor curves are so restricted in this plane that the operating point can be determined with reasonable accuracy from the input load line alone. A more significant simplification results when V_e is of the order of V_c, and R_e is much larger than $(r_e + r_b)$. On the scale shown, the input load equation will appear as a nearly vertical line on the input plane. In other words v_{eb}, the drop across the transistor input terminals, can be neglected relative to V_e. Then the input load equation becomes

$$V_e + v_1 = i_e R_e \qquad (5.9)$$

and this equation, together with Eqs. 5.6 and 5.8, solved simultaneously or graphically on the output plane alone, determines the operating point.

With V_e, R_e, V_c, and R_c specified, the graphical or analytical solution for the quiescent operating point ($v_1 = 0$) is unique. As v_1 varies, the input load line actually translates parallel to itself, and the operating point moves along the output load line. This situation corresponds to the *analysis* of a specified circuit. The *design* or *synthesis* problem is not unique, since it is readily apparent that any number of lines can be passed through point B by varying V_c and R_c appropriately, and any number of lines can be passed through point B' by choosing different values of V_e and R_e. In some instances additional requirements on the circuit narrow the possible choice of values, but often there is still considerable freedom.

In the design of a transistor circuit, an operating point must be selected to suit the design conditions. To illustrate this kind of problem, suppose we have available a type of transistor specified by the curves given in Fig. 5.23 parts (a) and (c). The first set of curves are in terms of the variables most useful for the common-base circuit shown in (b), whereas the second set relates the variables pertinent to the common-emitter circuit shown in (d). Suppose we wish to develop a peak-to-peak a-c output of ten volts across resistance R_2 without distorting the input signal appreciably. For the common-base curves and circuit shown in Fig. 5.23(a) and (b) the entire range of design possibilities is readily envisioned. First of all we need a collector supply voltage $V_c > 10$ volts. We shall choose a value of 12 or more volts (say 15) to avoid the distortion that would be produced by the nonlinearity of the curves near $v_{cb} = 0$. The value of 2 kilohms for R_2 is the minimum that will produce the required output, assuming the maximum available emitter current is 6 ma. For this value of R_2 the operating point must be placed at P_1, where $i_e = 3$ ma. If the emitter and base resistances of the transistor can be neglected relative to R_1, we merely require the ratio V_e/R_1 to be 3 ma.

For $R_2 = 3.75$ kilohms the 10 volt peak-to-peak output is obtained with an emitter current variation of only about 3 ma. With the operating point set at P_2 (where $i_e = 2$ ma) we have i_e ranging from about 0.5 to about 3.5 ma to produce the required output. The location of the operating point in this case is somewhat less critical since the portion of the load curve shown as a solid line can be shifted along the load line somewhat. The same is true for $R_2 = 7.5$ kilohms. In either of these cases the required values of R_1 and V_e can be determined as before.

The choice between possible values of R_2 is determined by gain and frequency-response requirements. With a larger value of R_2 the circuit has more voltage gain or power gain but over a smaller range of frequencies. With a small value of R_2, the gain is less but the frequency response is better.
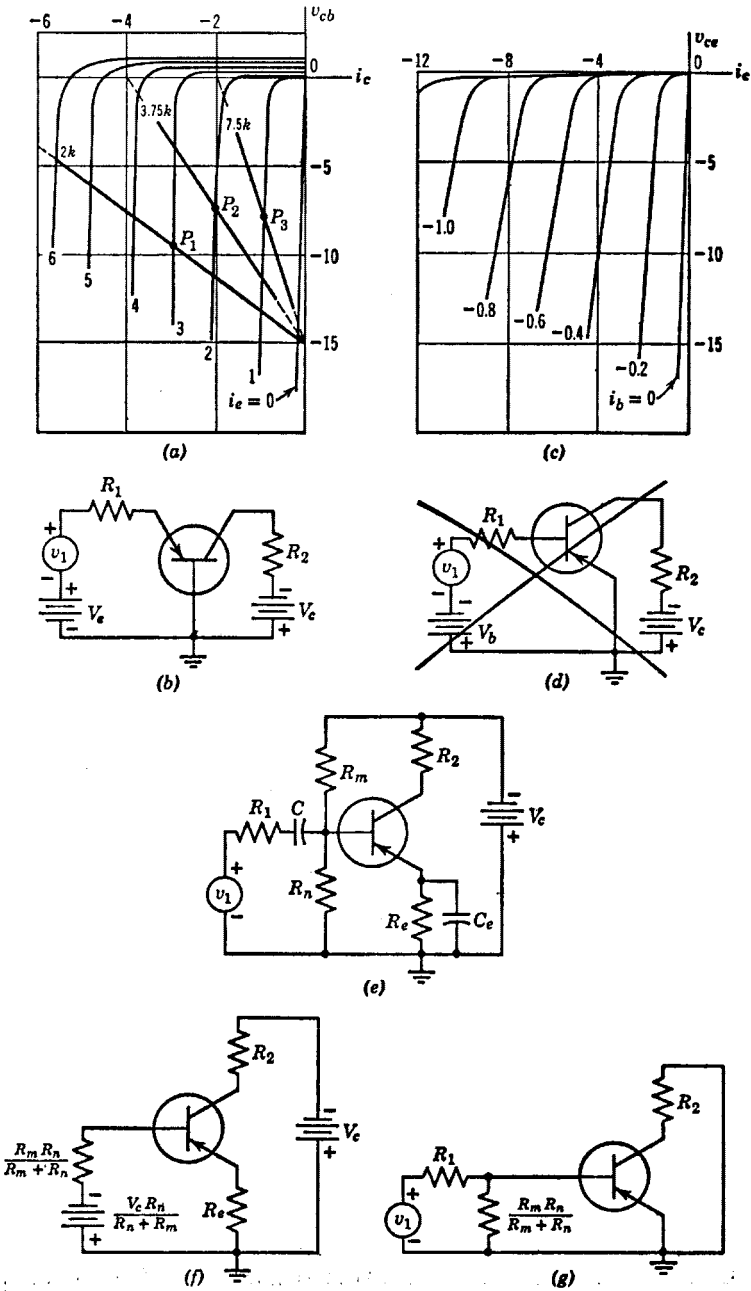
Fig. 5.23. Stabilization of bias for transistor amplifiers.

Referring to the curves of Fig. 5.23, part (c), we might proceed with a circuit like that shown in (d) in a manner analogous to that described for the common-base circuit. However, *the circuit shown in (d) is poor from a practical standpoint* because it exhibits a tendency toward instability of the operating point. In the common-base circuit (a), the current I_0 passes primarily through the collector-base loop. For the circuit of (d), most of I_0 will be driven through the emitter-collector loop by the polarizing battery V_c. The reverse current I_0 in a semiconductor junction is due to minority carriers. The number of such carriers available depends on the rate of hole-electron pair formation, which varies exponentially with temperature. Thus, any heating of the collector junction will increase I_0, which will further raise the temperature of the junction and increase I_0 still more. This may result in "thermal runaway." The result of collector junction heating may be a damaged transistor unless V_c and R_2 limit the current to a safe value. Even if the transistor is not damaged, an increase in I_0 shifts the entire set of collector characteristics so that the operating point is no longer properly located.

A circuit like that shown in Fig. 5.23(e) can be used to avoid this unstable operation. In this circuit C and C_e are assumed to be so large that they have negligibly small impedances for the lowest signal frequency involved. They are, nevertheless, open circuits for d-c. Thus for signal frequencies the incremental circuit reduces to a common-emitter circuit as shown in (g), but to represent the effects of the polarizing voltage and I_0 we have the circuit of (f). In (f) let the parallel combination of resistances in series with the base be designated R_b and let the equivalent battery be V_b. Assuming $R_b \gg r_b$, $R_e \gg r_{ef}$, and r_{cr} very large, we can write the equations

$$V_b + R_b i_b = R_e i_e \qquad (5.10)$$

and

$$i_b + i_e = a i_e + I_0 = -i_c \qquad (5.11)$$

These can be solved to obtain the dependence of i_c on V_b and I_0 as follows:

$$-i_c = \frac{(R_b + R_e)I_0 + aV_b}{(1 - a)R_b + R_e} \qquad (5.12)$$

For $R_e = 0$, we have $\partial i_c/\partial I_0 = -1/(1 - a)$, the situation that exists in the undesirable circuit of (d). With a very near unity (say 0.99) an increase of I_0 by an amount ΔI_0 will cause a change of $-100\Delta I_0$ in the value of i_c. The resulting increase in collector junction temperature may be sufficient to cause thermal runaway. With $R_e \neq 0$, we have the

situation that exists in the circuit represented by (e) or (f). If R_e equals R_b a change ΔI_0 yields only a change of $-2\Delta I_0$ in collector current. This represents a fifty-fold reduction in the sensitivity of i_c to changes in I_0. The point is that R_e must not be too small relative to R_b in the circuit of Fig. 5.23(f).
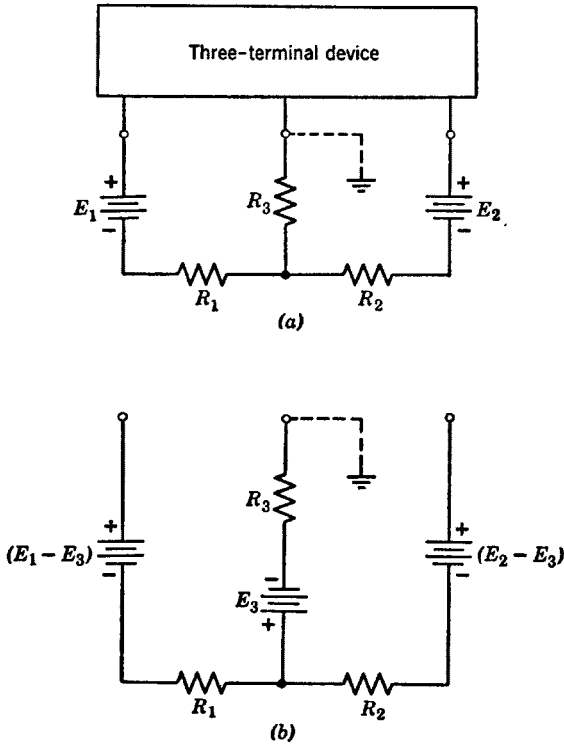
Fig. 5.24. General polarizing circuit.

Stability of the operating point is only one of many possible design considerations. The parameter values R_m, R_n, R_2, and R_e in the circuit of Fig. 5.23(e) determine (1) the d-c stability of the operating point, (2) the location of the operating point, (3) the a-c operating path upon the transistor curves, (4) the current and voltage gains, and (5) the input and output impedances of the amplifier. The circuit design starts with a trial specification of items such as (1) through (5) above and proceeds toward suitable values of the circuit resistances, usually through a number of successive approximations. A straightforward analytical solution of the problem is convenient only for very simple circuit con-

figurations. In any case, we must remember that individual transistor characteristics may deviate greatly from the published "average" curves, and that all other circuit elements are likewise known only within certain tolerances.

The circuits used to provide d-c polarizing voltages or currents for transistors or vacuum tubes appear to take many different forms. They may make use of one or more sources and resistive voltage-dividers. However, for any three-terminal device all polarizing circuits are varia-
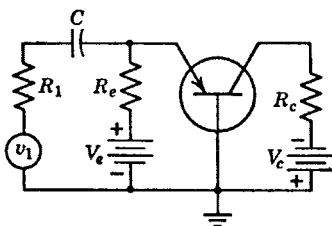
Fig. 5.25. Capacitive coupling of source.

tions of the general form shown in Fig. 5.24(a). The inclusion of a third battery as in (b) adds nothing significant, since an adjustment of the other two reduces the circuit to that shown in (a).

The use of capacitive coupling as shown in Fig. 5.23(e) permits separate determination of polarizing source and signal source requirements. In the circuit of Fig. 5.22(b), the d-c and signal currents must pass through the same resistance (R_1), so that the polarizing and signal source requirements are interdependent. The circuit of Fig. 5.25 shows a method of "isolating" the d-c and signal currents in a common-base circuit.

5.13 Driving-Point and Transfer Curves for Common-Base Circuit

The succession of operating points between X and Y in the input plane of v_{eb} versus i_e of Fig. 5.22(d), defines the driving-point curve seen at the emitter-to-ground terminals. This curve is shown in Fig. 5.26(a). Since $v_1 = v_{eb} - V_e + i_eR_e$, the corresponding curve of input voltage v_1 versus emitter current i_e can be obtained readily from the driving-point curve by shifting the intercept an amount V_e and increasing the slope by R_e, as shown in (a).

The transfer curve v_2 vs. v_{eb} and v_2 vs. v_1 can be obtained by solving

Eqs. 5.5, 5.6, 5.7, and 5.8 for i_c in terms of v_1 or v_{eb}. Then since

$$v_2 = -i_c R_c \tag{5.13}$$

the curve can be plotted by using the appropriate values of r_e and r_c for each state. This procedure yields the equations for each segment
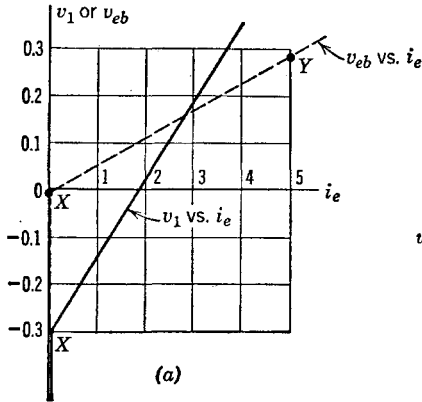
$$v_1 = i_e \left[R_e + r_e + \frac{r_b \{ R_c + r_c(1 - a) \}}{R_c + r_c + r_b} \right]$$
$$- \left[V_e + \frac{V_c r_b}{R_c + r_c + r_b} \right]$$
$$v_{eb} = v_1 + V_e - i_e R_e$$
$$= i_e \left[r_e + \frac{r_b \{ R_c + r_c(1 - a) \}}{R_c + r_c + r_b} \right]$$
$$- \left[\frac{V_c r_b}{R_c + r_c + r_b} \right]$$

(a)

(b)
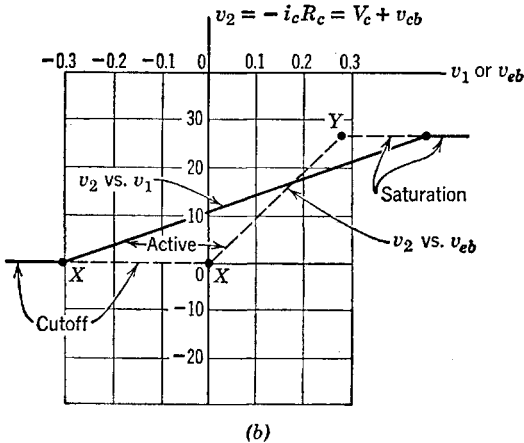
Fig. 5.26. Driving point and transfer curves for the circuit of Fig. 5.22.

of the curve. For the specific conditions illustrated in Fig. 5.22(d), graphical location of break points X and Y yields the transfer curves directly. At point X the emitter diode breaks, marking the transition between the cutoff and active or amplification region. At Y the collector diode breaks, marking the transition from active to saturation. The plot in Fig. 5.26(b) shows the resulting curves. The slopes are the voltage gains from the v_1 or v_{eb} terminals to the output terminals. The

effect of resistance R_e is to reduce the gain dv_2/dv_{eb} relative to dv_2/dv_1. For the numerical values used in the example, the reduction is approximately a factor of three.

Numerical values of the gains can be measured from the curves, and general gain expressions can be computed from the equations of Fig. 5.22(c). However, the incremental model is more convenient for gain calculations, since the constant terms do not enter the expressions.

5.14   Comparison of the Basic Transistor Circuits

Current gain, voltage gain, and power gain, as well as the input and output resistances of transistor circuits, are important in nearly all transistor applications. These properties of common-base, common-emitter, and common-collector circuits can be found by measuring the slopes of the appropriate transfer or driving-point curves as mentioned in Article 5.13. They can also be calculated by analysis of the appropriate incremental circuit, for example, the circuit in Fig. 5.20. However, the resulting expressions are usually cumbersome, and considerable simplification results when numerical values are substituted. Thus, rather than work out complicated exact expressions, it is more appropriate to calculate the approximate values of current gain, voltage gain, etc., using the ideal diode models to find approximate relative values of these quantities for the three basic transistor circuits in any one diode state. For example consider the most interesting state, namely that for which amplification is obtained.

The values of the gains and resistances tabulated in Fig. 5.27 are based on the use of an ideal incremental model for the transistor in the active amplification region ($r_e = r_b = 0$, $r_c = \infty$, and $a = \alpha$). The quantities in the table can be obtained directly from the simple circuits that result from this model.

The relative values of the emitter, base and collector currents provide the key to the determination of the gains and resistances. Since the value of emitter-to-collector current gain is slightly less than unity for a junction transistor, the common-base circuit has a current gain less than unity, whereas the current gains of the other two circuits are large (approximately 20 for $\alpha = 0.95$). The voltage gain is nearly R_2/R_1 for the common-base circuit, considerably higher for the common-emitter circuit and slightly less than unity for the common-collector circuit. The latter is sometimes called an "emitter-follower," since the incremental changes in emitter voltage follow incremental changes in base voltage quite closely.

In terms of power gain the common-emitter circuit is clearly superior to the common-base circuit. The power gain of the common-collector circuit also appears likely to be smaller than that of the common emitter, but the relative merits of the two depend on the ratio R_2/R_1. In fact, the common-collector circuit is most useful for coupling a high resistance

	Common Base	Common Emitter	Common Collector
Circuits			
Current gain $\dfrac{i_2}{i_1}$	$-a$	$\dfrac{a}{1-a}$	$-\dfrac{1}{1-a}$
Voltage gain $\dfrac{v_2}{v_1}$	$a\dfrac{R_2}{R_1}$	$-\left(\dfrac{a}{1-a}\right)\dfrac{R_2}{R_1}$	$\dfrac{R_2}{R_2+R_1(1-a)}\approx 1$
Power gain $-\dfrac{v_2 i_2}{v_1 i_1}$	$a^2\dfrac{R_2}{R_1}$	$\left(\dfrac{a}{1-a}\right)^2\dfrac{R_2}{R_1}$	$\left(\dfrac{1}{1-a}\right)\left(\dfrac{R_2}{R_2+R_1(1-a)}\right)$ $\approx\dfrac{1}{1-a}$
Input resistance $\dfrac{v_1}{i_1}$	$R_1+0=R_1$	$R_1+0=R_1$	$R_1+\dfrac{R_2}{1-a}$
Output resistance $\dfrac{v_2}{i_0}$	$(R_2\,\|\,\infty)=R_2$	$(R_2\,\|\,\infty)=R_2$	$R_2\,\|\,R_1(1-a)$

Fig. 5.27. Comparison of basic circuits using an ideal transistor model
$(r_e = r_b = 0,\ r_c = \infty)$.

source to a low resistance load $[R_2/R_1 \ll 1]$, since the circuit has a relatively high input resistance and a low output resistance. Under these conditions the power gain of the common-collector circuit becomes slightly larger than that of the common-emitter circuit.

The values given in Fig. 5.28 imply R_1 to be much greater than the input resistance contributed by the nonzero values of r_e and r_b. Also, R_2 must be much less than the effective output resistance. For a more accurate comparison of the three circuits, the complete expressions, involving all the transistor parameters, should be used.

For n-p-n transistors, diode polarities and the corresponding bias voltage polarities must be reversed. From the standpoint of the circuit models, the only difference between point-contact and junction transistors is in numerical values; for example, with point-contact transistors we may have a short-circuit-current gain from emitter to collector greater

than unity, so that values of a are in the vicinity of 2 or 3. Also, the characteristic curves are such that diodes in the piecewise-linear circuit model are likely to have slightly larger forward resistances (r_{ef} and r_{cf} of the order of a few hundred ohms) and considerably smaller reverse resistances (r_{er} and r_{cr} of the order of 20 kilohms to 100 kilohms). In addition, the individual curves of the family $v_{cb} = f(i_c, i_e)$ have greater variation in slope and spacing throughout the v_{cb} vs. i_c plane.

	Common Base	Common Emitter	Common Collector
Circuits			
Current gain $\dfrac{i_2}{i_1}$	$-a$	$\dfrac{a}{1-a}$	$-\dfrac{1}{1-a}$
Voltage gain $\dfrac{v_2}{v_1}$	$\dfrac{aR_2}{R_1 + R_b(1-a)}$	$\dfrac{-aR_2}{(1-a)R_1 + R_e}$	$\dfrac{R_2}{R_2 + R_1(1-a)}$
Power gain $\dfrac{v_2 i_2}{v_1 i_1}$	$\dfrac{a^2 R_2}{R_1 + R_b(1-a)}$	$\dfrac{a^2 R_2}{(1-a)[(1-a)R_1 + R_e]}$	$\dfrac{R_2}{(1-a)[R_2 + R_1(1-a)]}$
Input resistance $\dfrac{v_1}{i_1}$	$R_1 + R_b(1-a)$	$R_1 + \dfrac{R_e}{(1-a)}$	$R_1 + \dfrac{R_2}{1-a}$
Output resistance $\dfrac{v_2}{i_o}$	$R_2 \parallel \infty = R_2$	$R_2 \parallel \infty = R_2$	$R_2 \parallel R_1(1-a) \approx R_1(1-a)$

Fig. 5.28. Effects of coupling resistance in the basic circuits ($r_e = r_b = 0, r_c = \infty$).

The ideal-diode model, the piecewise-linear diode model, and the linear-incremental model have one very important feature in common. They all yield "linear" circuits and therefore linear differential equations when combined with energy-storage elements. To be sure, the diode models have multiple states and therefore will yield different linear equations for each state traversed during operation of the circuit.

Various resistive models can be devised for transistors in addition to those already discussed. As soon as we deviate from the use of linear segments to match the curves, we have many possible functions to consider. One that is often useful is based on the use of exponential diodes, and it is in keeping with the exponential relation between current and voltage in a semiconductor junction. Although exponential models (or other nonlinear models) may result in a better fit to the graphical curves or a better interpretation of the physical behavior of the transistor,

they are usually less useful in circuit analysis and will not be pursued here. It is important to keep in mind the fact that *no circuit model is the "right one" or the "best one" for all purposes.* A model that is too general may be very cumbersome for analysis of circuit behavior. It is usually preferable to use the simplest model that adequately describes a device in a specific operational environment.

Thus far no mention has been made of the effects of high-frequency variations of the currents and voltages applied to a transistor, nor has anything been said about the effects of temperature changes. Both of these materially influence transistor behavior.

5.15   High-Frequency Circuit Model

Resistive models are valid only for d-c or low-frequency operation. Modification of the models is necessary if they are to be used to determine circuit behavior at high frequencies. We shall here define high frequencies rather loosely as those for which the resistive models no longer represent the transistor adequately. Depending on the type of transistor involved and the circuit configuration, this may mean frequencies as low as tens of kilocycles per second, or as high as a few megacycles per second.

Semiconductor junctions exhibit an effective capacitance that influences transistor behavior when rapid variations of voltage or current occur. This capacitance depends upon the cross-sectional dimensions of the junction, the materials, the applied voltage, and other factors related to the method of fabrication. Typical values run from 10 to 100 $\mu\mu f$. The incremental model shown in Fig. 5.29 (a) includes such a capacitance C_c in the collector circuit to represent the base-to-collector junction. For frequencies where the capacitive reactance is appreciably less than the collector resistance, $1/2\pi f C_c \ll r_c$, the resistance can be omitted, as shown in (b). The corresponding capacitance across the emitter-to-base junction can usually be neglected if the emitter diode is closed, since this capacitor is shunted by the very small emitter resistance. The equations previously given for resistive models apply to the circuit in (a) if r_c is replaced by complex impedance z_c, i by I and v by V.

The presence of the capacitance at the collector junction can be visualized in terms of the charge and potential distributions. Diffusion of majority carriers across the base-to-collector junction builds up a potential difference, which is characteristic of a capacitance. The charge distribution, electric field and potential difference near the junction are shown in Fig. 5.30. The reverse bias applied for normal operation

$$V_{eb} = I_e(r_e + r_b) + I_c r_b$$
$$V_{cb} = I_e(r_b + az_c) + I_c(z_c + r_b)$$
$$y_c = \frac{1}{z_c} = \frac{1}{r_c} + j\omega C_c$$

$$V_{eb} = I_e(r_e + r_b) + I_c r_b$$
$$V_{cb} = I_e\left(r_b - j\frac{a}{\omega C_c}\right) + I_c\left(r_b - j\frac{1}{\omega C_c}\right)$$
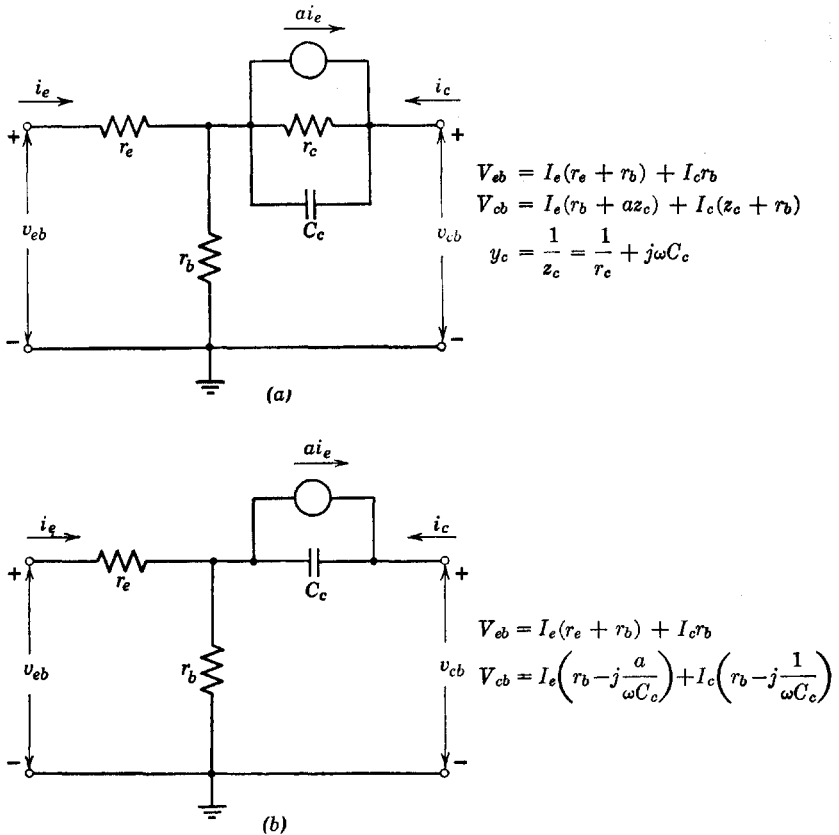
Fig. 5.29. Incremental model including collector capacitance.

merely serves to charge this capacitance further. Conversely the forward bias applied to the emitter-to-base junction results in a large conduction current that virtually swamps the capacitance effect.

5.16 Effect of Collector Capacitance on the Gain of the Grounded-Base Circuit

The collector capacitance tends to reduce the magnitude of the current gain of the transistor and also introduces a phase shift. Referring to the circuit of Fig. 5.29(a), it is evident that a load resistance R_c, connected between collector and ground, will receive almost the entire current ai_e at zero frequency and a small fraction $i_e r_b/(r_b + R_c)$ at infinite frequency where the capacitive reactance reaches zero.

A grounded-base circuit with collector capacitance in the transistor model is shown in Fig. 5.31(a), together with a vector diagram. The expression for complex current gain I_c/I_e is given in (b), and the absolute
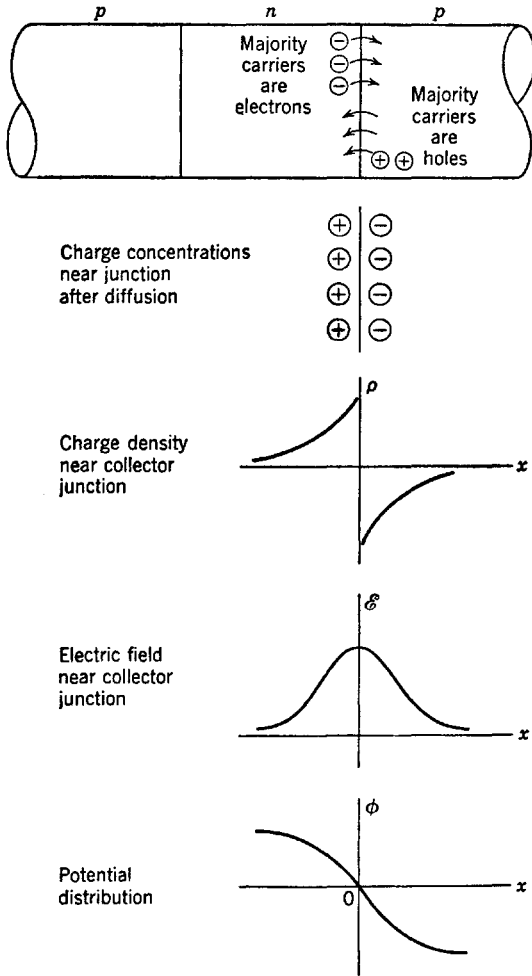
Fig. 5.30. Charge, electric field, and potential distribution near the collector junction of a p-n-p transistor.

magnitude $|I_c/I_e|$ is given in (c). The phase angle θ, representing the amount by which I_c leads I_e, is given in (d). Sketches of $|I_c/I_e|$ vs. ω, and θ vs. ω are given in (e), using rectangular coordinates for the plots. The polar coordinate sketch in (f) indicates $|I_c/I_e|$ as the length of the radius vector, and the phase angle appears as the polar angle. This is

(a)
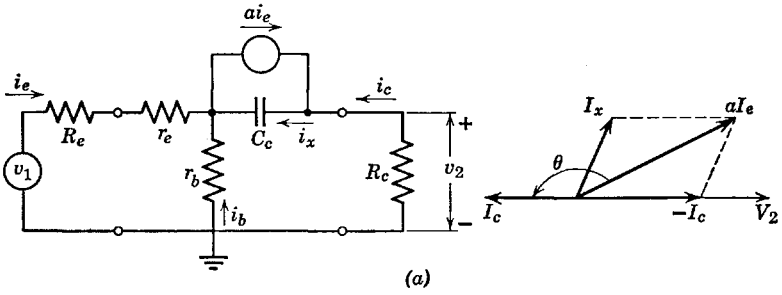
$$(b) \quad \frac{I_c}{I_e} = -\frac{r_b + a/j\omega C_c}{r_b + R_c + 1/j\omega C_c} = -\frac{\left[r_b(r_b + R_c) + \dfrac{a}{(\omega C_c)^2} \right] - j\left[\dfrac{aR_c - (1-a)r_b}{\omega C_c} \right]}{(r_b + R_c)^2 + (1/\omega C_c)^2}$$

$$(c) \quad \left| \frac{I_c}{I_e} \right| = \sqrt{\frac{r_b{}^2 + (a/\omega C_c)^2}{(r_b + R_c)^2 + (1/\omega C_c)^2}} = \sqrt{\frac{(r_b \omega C_c)^2 + a^2}{(r_b + R_c)^2(\omega C_c)^2 + 1}}$$

$$(d) \quad \tan \theta = \frac{-[aR_c - (1-a)r_b]\omega C_c}{[(\omega C_c)^2 r_b(r_b + R_c) + a]} \approx \frac{-1}{\dfrac{\omega C_c r_b}{a} + \dfrac{1}{\omega C_c R_c}} \text{ for } r_b \ll R_c$$

(e)

(f)

Fig. 5.31. Current gain versus frequency for grounded-base circuit.

drawn from a succession of vector diagrams for different frequencies; hence, ω is a parameter along the curve. This polar plot is called a Nyquist diagram and is frequently used to describe transfer ratios.

The dependence of current gain on frequency, due to collector capacitance, can be determined in a similar manner for the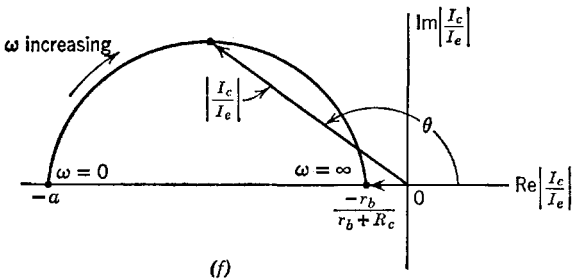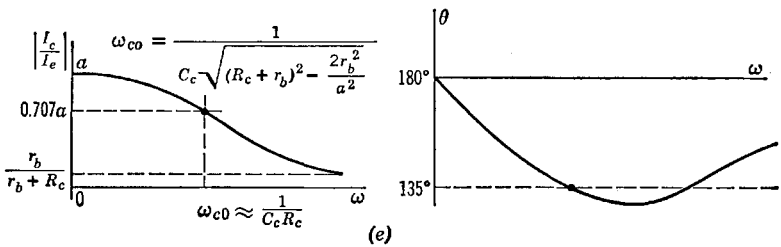 common-emitter and common-collector circuits. The effective reactance is modified in these circuits by the current gain. Thus, in the common-emitter circuit, since $r_d = r_c(1 - a)$, we have $C_d = C_c/(1 - a)$.

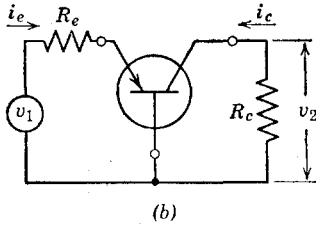## 5.17 Frequency Dependence of Current Gain

Another factor that influences the gain of transistor amplifiers at high frequencies is the frequency dependence of α, the short-circuit current gain of the transistor. This effect can be attributed in the p-n-p transistor to the time required for holes to diffuse through the base region. If the emitter current is varied incrementally at frequencies for which the period is comparable to the diffusion time, the collector current suffers a phase lag and a reduction in amplitude. The phase shift or time lag can be visualized as a delay in the effect of an increase (or decrease) of the hole current injected into the base. If the variation is sinusoidal, the peak (or any other reference point) in the waveform of i_e has passed before the corresponding variation is transmitted through the base by diffusion. The reduction in the effective amplitude can be attributed to dispersion of transit times of the individual carriers because of the randomness of the hole motion. Thus, for good high-frequency performance, the base layer should be kept very thin to keep diffusion time low.

A power series in frequency can be used to approximate an experimental curve of short-circuit current gain versus frequency. For most purposes the first two terms of the expansion suffice to give reasonable results. A circuit model with a frequency-dependent current generator, as shown in Fig. 5.32(a), yields an approximate representation of this effect. The quantity a_0 is the d-c or low-frequency value of a and corresponds to the a used in all previous circuit models. The frequency dependence can be included in all previous discussions by letting $a = a_0/[1 + j(\omega/\omega_0)]$. The frequency ω_0 is the frequency for which $|a| = 0.707\, a_0$.

If the frequency-dependent model is used in a grounded-base circuit, as shown in Fig. 5.32, the basic expressions in (c) lead to the expression for current gain shown in (d). The expressions for the grounded-emitter
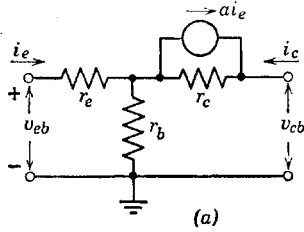
$$a = \frac{a_0}{1 + j\dfrac{\omega}{\omega_0}}$$

$$v_1 = i_e[R_e + r_e + r_b] + i_c r_b$$

$$0 = i_e[r_b + ar_c] + i_c[R_c + r_c + r_b]$$

(c)

$$\frac{I_c}{I_e} = -\frac{r_b + r_c\left[\dfrac{a_0}{1 + j\dfrac{\omega}{\omega_0}}\right]}{r_b + r_c + R_c}.$$

(d)

Fig. 5.32. Frequency dependence of current source.

and grounded-collector circuits can be modified in like manner by using the expression $a_0/[1 + j(\omega/\omega_0)]$ for the parameter a.

5.18 Variation of Transistor Parameters

A piecewise-linear circuit model implies constancy of the parameters that specify the transistor throughout a given state of the diodes in the model. Actual transistor curves deviate from this idealization to some extent, and if the deviation is important the variation of parameters must be considered. Figure 5.33 gives an example of the parameter changes that may occur as a function of the operating point location in the amplification state. Such data are useful in circuit design. For instance, we may wish to choose an operating point in a region where a parameter variation is smallest; or we may wish to obtain maximum

gain, or meet some other condition. The specific numerical values of the parameters at the appropriate point can then be used in the design calculations. In large-signal applications the parameter variations are usually less important, and a piecewise-linear model can be based on
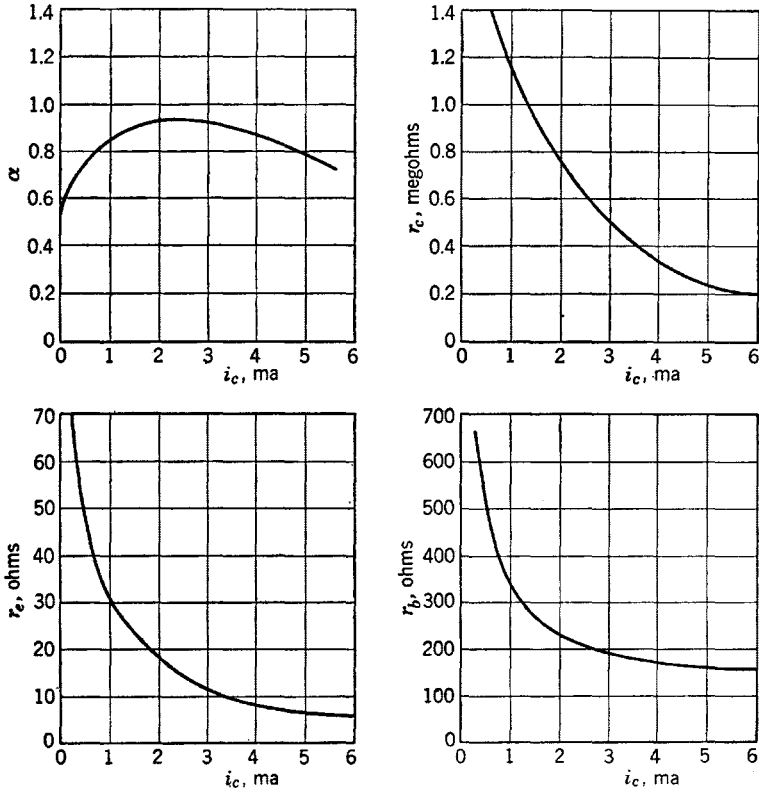
Fig. 5.33. Typical variations of transistor parameters with collector current.

average values of the parameters. In either case one of the design considerations is reduction of the effects of parameter variations.

The temperature dependence of transistor parameters poses another problem, since the environment in which a circuit is to be used may not be under the control of the circuit designer. The seriousness of temperature effects depends upon the semiconductor material and upon the fabrication process used by the transistor manufacturer. One critical effect of ambient temperature on transistor characteristics is a shift of the entire v_{cb} vs. i_c family, which results because the saturation current

I_{co} through the collector junction increases exponentially with tempera-
ture. This is the effect that was mentioned in connection with d-c
polarization of the common-emitter circuit. A semilog plot of a typical
variation with temperature of the reverse collector-junction current for
zero emitter current is shown in Fig. 5.34(a). We recall from the dis-
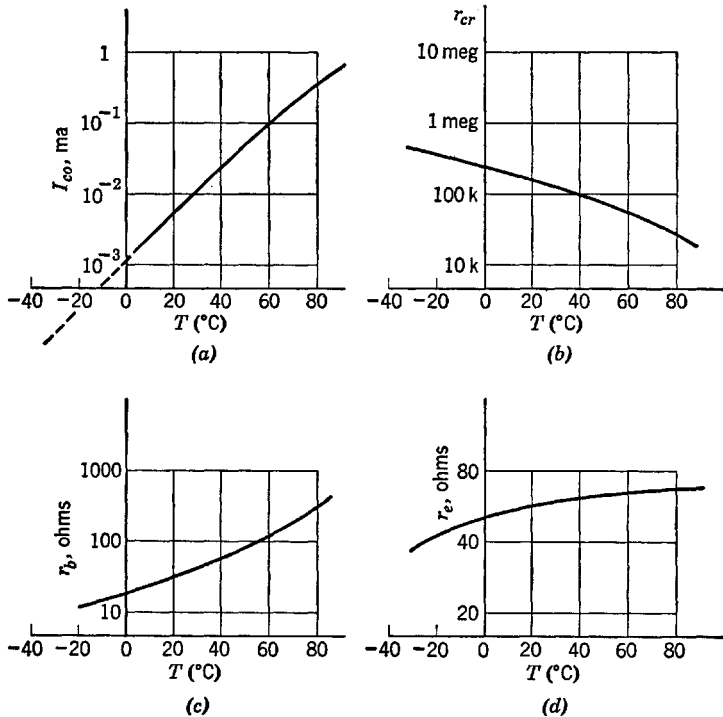
Fig. 5.34. Effect of temperature on parameters of typical germanium junction transistors.

cussion of semiconductor diodes in Chapter 2 that this reverse current
is due to minority-carrier conduction that depends on the rate of forma-
tion of hole-electron pairs by thermal agitation. Once formed, such
pairs exist (on the average) for a fraction of a millisecond before re-
combining. The number of charge carriers, and consequently the
conductivity, increases exponentially with temperature. This effect
places a definite temperature limitation on semiconductor devices
(60 to 80°C for germanium and 120° to 150°C for silicon).

A plot of conductivity versus temperature for minority-carrier con-
duction can be expected to resemble the curve of I_{co} vs. T. Corre-

spondingly, resistivity, or a resistance such as r_{cr}, should diminish with temperature, as indicated by the typical curve of r_{cr} vs. T in Fig. 5.34(b). Resistance due to majority-carrier conduction, such as r_{ef} or r_b, tends to rise with temperature, as indicated in Fig. 5.34(c) and (d), because the resistivity of p-type or n-type semiconductor material generally tends to rise with temperature. This is true only up to the temperature for which the conduction of the intrinsic material becomes comparable to that of the donor or acceptor charge carriers. For still higher temperatures, the resistivity of such materials decreases, as does the resistivity of intrinsic semiconductor material.

5.19   Transistor Curves

In plotting transistor collector curves we have used collector current as the independent variable, collector voltage as the dependent variable, and either emitter current or base current as the parameter. These choices are based on the fact that the transistor is a current-controlled device. The use of standard four-terminal-circuit reference directions for currents and voltages results in third-quadrant curves for the p-n-p types and first-quadrant curves for the n-p-n types. These forms are the natural result of combining the physical behavior of transistors with circuit theory standards. Transistor data supplied by most manufacturers usually places collector curves in the first quadrant by reversing reference directions (in effect) for p-n-p units. Furthermore, manufacturers' data usually appear as a plot of current versus voltage, for ease of comparison with vacuum triode and pentode data.

Electronic circuit design may involve transistors, vacuum tubes, and other devices specified by manufacturers' graphical data. The designer must be prepared to relate the data to his analysis or synthesis procedures. The transistor curves given in Appendix A for use in connection with some of the problems are patterned after typical manufacturers' curves.

SUPPLEMENTARY READING

L. B. Arguimbau, R. B. Adler, *Vacuum Tube Circuits and Transistors*, John Wiley and Sons, New York, 1956.

W. G. Dow, *Fundamentals of Engineering Electronics*, John Wiley and Sons, New York, 1952.

E. A. Guillemin, *Introductory Circuit Theory*, John Wiley and Sons, New York, 1953.

L. P. Hunter, editor, *Handbook of Semiconductor Electronics*, McGraw-Hill, New York, 1956.

Lo, Endres, Zawels, et al., *Transistor Electronics*, Prentice-Hall, New Jersey, 1955.

R. D. Middlebrook, *An Introduction to Junction Transistor Theory*, John Wiley and Sons, New York, 1957.

Samuel Seely, *Electronic Engineering*, McGraw-Hill, New York, 1956.

Shockley, *Holes and Electrons in Semiconductors*, Van Nostrand, New York, 1950.

Karl R. Spangenberg, *Fundamentals of Electron Devices*, McGraw-Hill, New York, 1957.

Robert L. Sproull, *Modern Physics: A Textbook for Engineers*, John Wiley and Sons, New York, 1956.

Aldert van der Ziel, *Solid State Physical Electronics*, Prentice-Hall, New Jersey, 1957.

Dewitt and Rossoff, *Transistor Electronics*, McGraw-Hill, New York, 1957.

PROBLEMS (See Appendix A for transistor curves)

5.1. The resistive-diode circuit of Fig. P5.1 contains two dependent sources.
(a) Find the breakpoints of the two diodes.
(b) Sketch and dimension i_1 vs. e_1.
(c) What is the incremental impedance, $\Delta e_1/\Delta i_1$, at $e_1 = -3$ volts.

Fig. P5.1

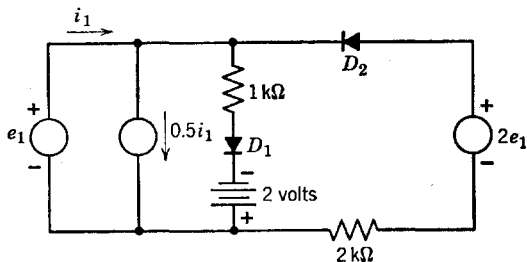**5.2.** Using an ideal-diode transistor model in Fig. P5.2:
(a) Sketch and dimension e_2 vs. e_1.
(b) Sketch and dimension i_2 vs. i_1.
(c) Determine the incremental voltage gain $\Delta e_2/\Delta e_1$ and the incremental current gain $\Delta i_2/\Delta i_1$ for the region of operation in which the emitter diode is closed and the collector diode is open.
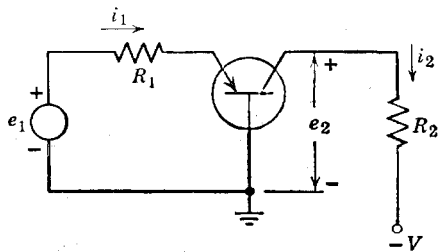

Fig. P5.2

5.3. Repeat Problem 5.2 for the common-emitter connection (Fig. P5.3).
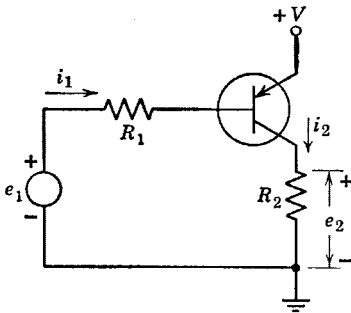
Fig. P5.3
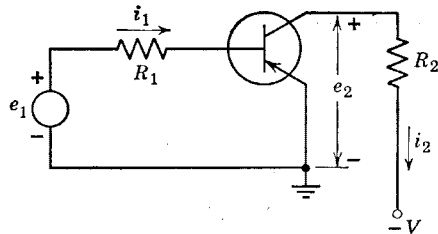
Fig. P5.4

5.4. Repeat Problem 5.2 for the common-emitter connection shown in Fig. P5.4.

5.5. Repeat Problem 5.2 for the common-collector connection shown in Fig. P5.5.
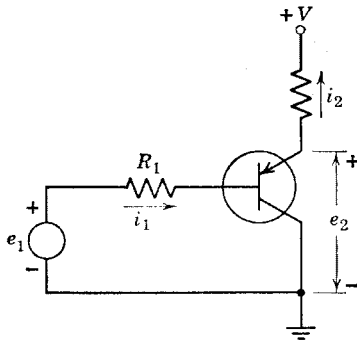
Fig. P5.5

5.6. For the Type I transistor operated with the base common determine the least load resistance that may be used with a collector supply voltage of -10 volts without exceeding the rated collector dissipation of 4×10^{-3} watts when the emitter current may be any value in the range 0 to 5 ma.

5.7. From the Type I curves estimate the parameters r_{ef}, r_{cr}, r_b, and α belonging to the piecewise-linear circuit model of Fig. 5.12(a).

5.8. For a transistor operated with the base common determine the variational power gain on the basis of these data:

(a) Model is the piecewise-linear one of Fig. 5.12(a) operated with the collector in the reverse region and the emitter in the forward region.

(b) $r_{ef} = 50$ ohms, $r_b = 50$ ohms, $r_{cr} = 2 \times 10^5$ ohms, $\alpha = 0.95$.

(c) Emitter source resistance $= 100$ ohms, load resistance $= 2000$ ohms.

5.9. For a transistor operated with the emitter common determine the variational power gain on the basis of the data of Problem 5.8.

5.10. (a) Show that all possible transistor d-c polarizing circuits are equivalent to the circuit shown in Fig. P5.6.

(b) For the idealized transistor model in Fig. P5.7, how does i_e depend on V_e, V_c, and the minority-carrier current I_0.

(c) Find $\Delta i_e/\Delta I_0$, and discuss the significance of the result in the design of a d-c polarizing network.
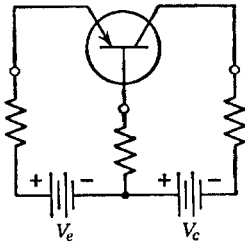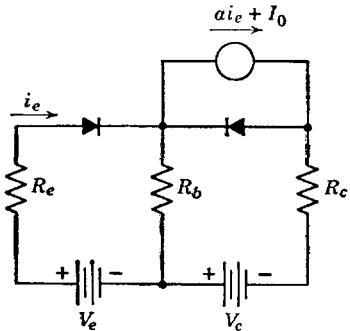
Fig. P5.6   **Fig. P5.7**

5.11. A common-base transistor circuit drives a load R_2 from a source with internal resistance R_1. For incremental voltages and currents, in the amplification region, the transistor model includes r_e, r_b, r_c, and a. Derive expressions for the following quantities in terms of R_1, R_2, r_e, r_b, r_c and a:

(a) Current gain.
(b) Voltage gain.
(c) Power gain.
(d) Input resistance.
(e) Output resistance.

5.12. Repeat Problem 5.11 for a common-emitter circuit.

5.13. Repeat Problem 5.11 for a common-collector circuit.
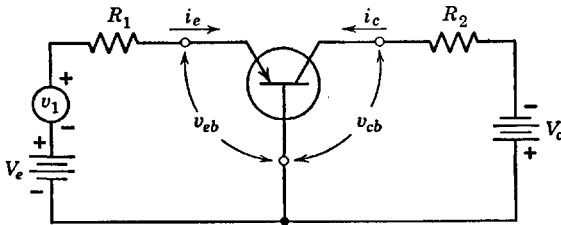
$V_e = 5$ volts
$V_c = 10$ volts
$R_1 = 2$ kilohms
$R_2 = 2$ kilohms

Fig. P5.8

5.14. The p-n-p junction transistor in the common-base amplifier circuit shown in Fig. P5.8 can be approximated by an ideal-diode model in which the emitter-to-collector current gain is given by $a = 0.95$. Neglect the reverse saturation current of the collector junction (I_0) and assume the current gain from collector to emitter is zero ($a' = 0$).

(a) Sketch and dimension the following transistor curves based on the ideal-diode model

v_{cb} vs, i_c with i_e as the parameter

v_{eb} vs. i_e with i_c as the parameter

(b) Find the values of i_e, i_c, v_{cb}, and v_{eb} at the quiescent operating point of the circuit ($v_1 = 0$).

(c) Plot the driving-point curve, v_1 vs. i_e.

(d) Plot the transfer curves v_{cb} vs. v_{eb} and v_{cb} vs. v_1.

$V_c = 10$ volts
$R_2 = 2$ kilohms
$R_1 = 2$ kilohms
$V_e = 5$ volts
$R_e = ?$
$C = ?$
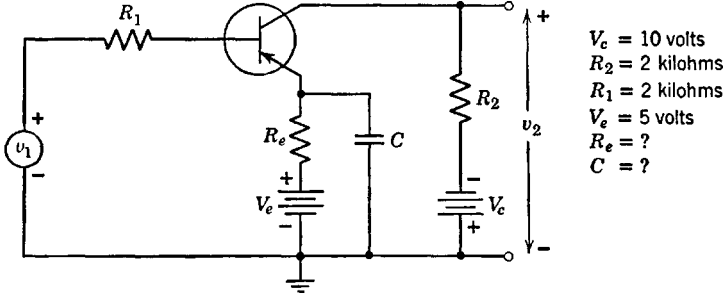
Fig. P5.9

5.15. The transistor of Problem 5.14 is connected in the common-emitter circuit shown in Fig. P5.9.

(a) Determine the value of R_e needed to make the emitter current (i_e) at the quiescent operating point ($v_1 = 0$) the same as in Problem 5.14.

(b) Plot the transfer curve v_2 vs. v_1, assuming v_e, R_e and C can be replaced by an appropriate battery.

(c) Determine the incremental voltage gain $\Delta v_2/\Delta v_1$ for the circuit assuming the increments in v_1 are small and the capacitance is very large.

(d) If $v_1 = V_1 \sin \omega_1 t$, find the maximum value of V_1 that permits linear amplification for the quiescent conditions established in (a).

(e) For $\omega_1 = 10^4$ radians/sec, determine the approximate value of C needed to hold the peak-to-peak variation of capacitor voltage to 2 per cent of the average capacitor voltage.
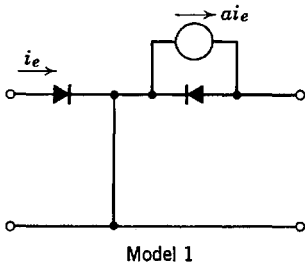
Model 1
Fig. P5.10
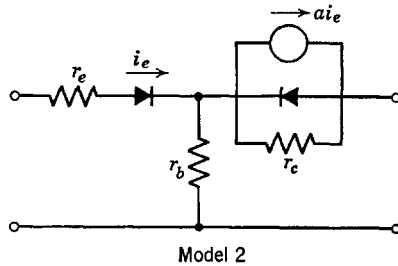
Model 2
Fig. P5.11

5.16. (a) From the information on the common-emitter characteristics of the Type II transistor determine the values of a, r_c, and r_e that you would use in each of the two models shown in Figs. P5.10 and P5.11 at an operating point $v_{ce} = -4$, $i_b = -0.6$. ($r_b \approx 100$ ohms from independent measurement)

(b) Using the circuit of Fig. P5.8, determine for which of the following calculations Model 1 (Fig. P5.10) can be used as a good approximation for the transistor rather than the more complete Model 2 (Fig. P5.11).

(1) Current gain $\Delta i_c / \Delta i_e$

(2) Voltage gain $\Delta v_{cb} / \Delta v_{eb}$, and $\Delta v_{cb} / \Delta v_1$

(3) Input impedance $\Delta v_{eb} / \Delta i_e$

(4) Output impedance $\Delta v_{cb} / \Delta i_c$

5.17. Repeat Problem 5.16(b) for the Type III *n-p-n* transistor.

Model for transistor

$V_c = -15$ volts
$R_1 = 1$ kilohm
$R_b = 750$ kilohms
$R_2 = 5$ kilohms
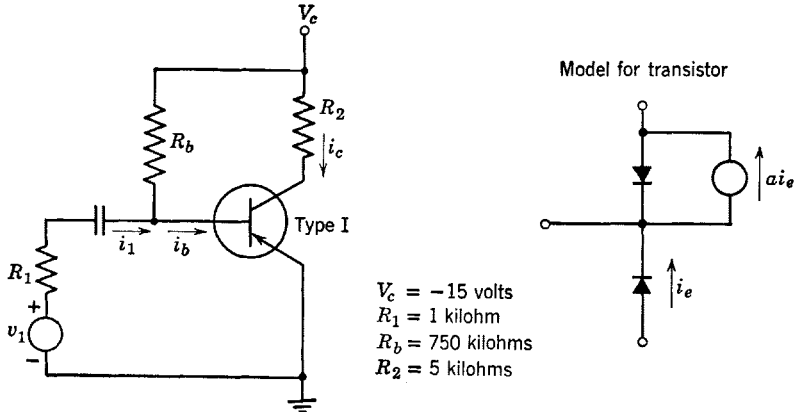
Fig. P5.12

5.18. The circuit of Fig. P5.12 uses a Type I transistor in the common emitter connection. The emitter bias is established by a base current I_b obtained through R_b. Thus V_c biases both the emitter and collector.

(a) Assuming that the base-to-emitter voltage is small compared to V_c, determine the quiescent collector voltage (V_{ce}) and collector current (I_c) from the characteristics of the common-emitter connection.

(b) Graphically determine the incremental current gain $\Delta i_c / \Delta i_1$. From this value of current gain, compute the value of a to be used in the idealized model shown in Fig. P5.12.
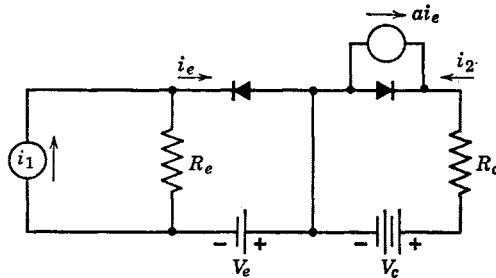
Fig. P5.13

5.19. Using the ideal-diode transistor model shown in Fig. P5.13, obtain the i_2 vs. i_1 characteristic. Repeat for the grounded emitter and grounded collector connections.

5.20. A Type IV point-contact transistor is used in the circuit of Fig. P5.14.

(a) State the condition of the diodes in the normal operating region (open or closed).

V_c = 50 volts V_e = 0.5 volt
R_c = 5 kilohms R_e = 100 ohms
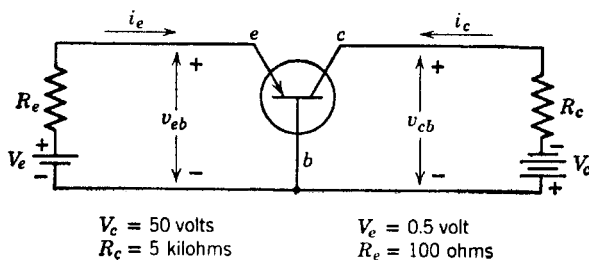
Fig. P5.14

(b) On the emitter characteristics (v_{eb} vs. i_e) of the Type IV transistor, plot the input load line equation. Similarly, on the collector characteristics plot the output load line.

(c) Locate the quiescent operating point by transferring the load line on the collector characteristics onto the emitter characteristics. Also locate the operating point on the collector characteristics.

(d) Determine the small signal parameters (r_e, r_b, r_c and a) at the operating point.

(e) Determine the small signal voltage gain $\Delta v_{cb}/\Delta v_{eb}$ at the operating point.
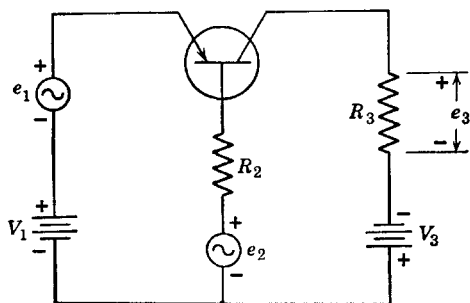
Fig. P5.15

5.21. In the circuit of Fig. P5.15, assume the transistor has negligibly small forward emitter resistance and negligibly small reverse collector conductance.

(a) With e_1 and e_2 set equal to zero, sketch and dimension, in the V_3 versus V_1 plane, the boundaries of the region within which the transistor is operating in the usual amplification region (emitter conducting and collector non-conducting).

(b) For operation in such region, find the a-c signal at e_3 when small a-c signals $e_1(t)$ and $e_2(t)$ are applied.

5.22. The variable resistance of Fig. P5.16, R_M (a carbon button microphone) has a resistance of 400 ohms ± 10 per cent ($R_M = R_0 + \Delta r$). What would

be a suitable value of R_1 for an operating point of 200 microamperes base current. Determine v_0 as a function of Δr. Use the data for the Type I transistor.
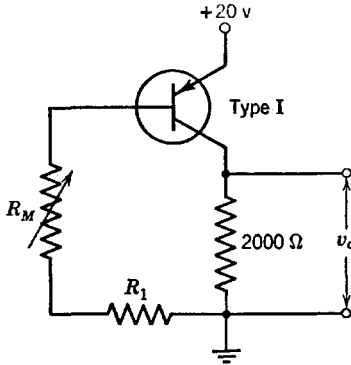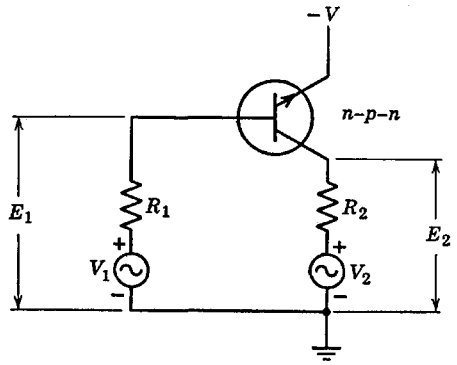
Fig. P5.16   **Fig. P5.17**

5.23. The circuit shown in Fig. P5.17 is used to measure the h-parameters of a transistor. Small a-c voltages of adjustable amplitudes V_1 and V_2 are applied as shown. The resulting a-c voltages E_1 and E_2 are measured with a high-impedance voltmeter. Assume that R_1 is much greater than $r_b + r_e/(1-a)$, and R_2 is much less than $r_c(1-a)$.

(a) Determine the h-parameters in terms of the external voltages and resistances.

(b) Determine r_e, r_b, r_c, and a.
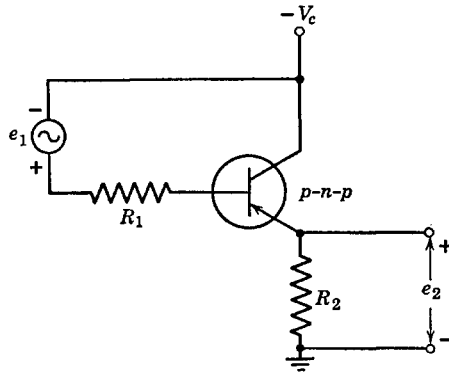
Fig. P5.18

5.24. An emitter-follower circuit using a p-n-p junction transistor is shown in Fig. P5.18. The transistor is biased in the forward gain region and may be represented by the idealized linear incremental circuit model with $r_e = r_b = 0$ and $r_c = \infty$.

(a) Find the incremental voltage gain, $\Delta e_2/\Delta e_1$.

(b) Find the incremental input resistance facing the signal source, e_1.

(c) Find the incremental output resistance facing the load resistance, R_2.

5.25. The circuit shown in Fig. P5.19 is used as the output stage of a transistor audio amplifier. Note that by combining p-n-p and n-p-n transis-
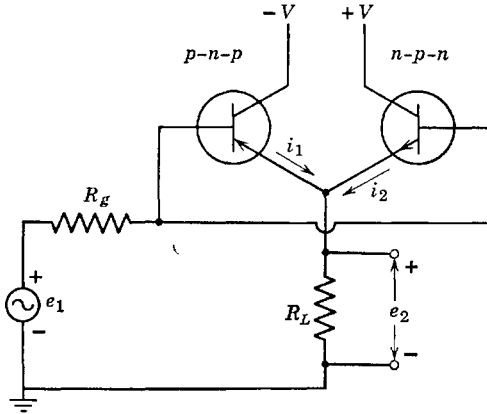
Fig. P5.19

tors, push-pull operation is obtained without the use of transformers. Assuming that in the transistor model $r_e = r_b = 0$ and $r_c = \infty$, $a = 0.95$.

(a) Plot and dimension the transfer curve e_2 vs. e_1.

(b) Assuming $e_1 = E_1 \sin \omega t$; plot e_2, i_1, and i_2 vs. time for $E_1 < V$.

5.26. Using piecewise-linear models for the n-p-n and p-n-p transistors in the circuit shown in Fig. P5.20, find e_2 as a function of e_1, and show the result in graphical form (transfer plot). Assume that $r_c = \infty$, $r_e = r_b = 0$, and that the a for both transistors is 0.95.
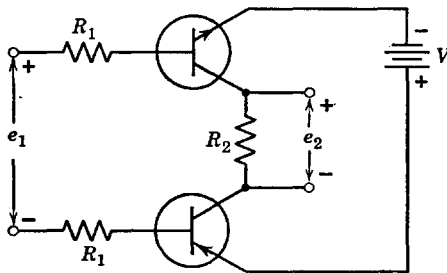
Fig. P5.20

5.27. The signal from a source having an internal resistance of 350 ohms is amplified by a single-stage amplifier which uses a p-n-p junction transistor in the common-base connection. The amplified signal is delivered to a purely resistive 5000-ohm load. Assume that the transistor is sufficiently well represented by the circuit of Fig. P5.21.

(a) At the d-c operating point, with $e_s(t) = 0$, compute the power flowing out of V_e and V_c, the power flowing into the load (R_L), the power flowing into the source terminals, and the power flowing into the transistor.

(b) When amplifying a signal from a source of square waves with an

open-circuit voltage varying between $+1$ and -1, compute the average power flowing into R_L, the average power flowing out of the source terminals, and the average power flowing into the transistor.

(c) Is the average power flowing into the transistor greater or less in part (b) than in part (a)? Explain briefly.

(d) What is the maximum possible power which could be supplied by the source if it were disconnected from the transistor amplifier and reconnected
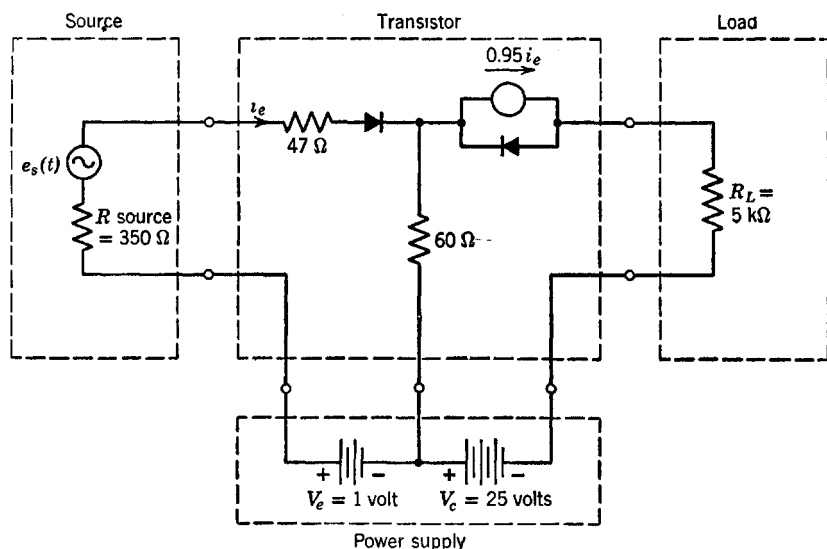
Fig. P5.21

to a resistance which matches its internal resistance? Is this greater or less than the power flow from the source in part (b)?

(e) Construct an incremental circuit for the source, transistor, and load, and compute the ratio of signal power in the load to the signal power flowing out of the source terminals in your equivalent circuit. This is one possible definition of power gain (called "network efficiency," and for most practical amplifiers is a number greater than unity and a function of the load impedance). Compute the ratio of signal power in the load to the maximum signal power available at the source terminals. This is the usual definition of "power gain."

CHAPTER SIX

Vacuum Triodes

6.1 Introduction

The vacuum triode can be described as a vacuum diode with a control element, called the grid, interposed between the heated cathode and the anode. Thermionic emission of electrons from the heated cathode provides the free electrical charges necessary for conduction through the vacuum. Although the physical laws of conduction in vacuum differ in many details from those applying to semiconductors, the gross behavior of vacuum triodes bears considerable resemblance to that of transistors. From the standpoint of electronic circuit analysis, both triodes and transistors can be classified as control valves that are essentially linear in restricted regions of operation. Both are primarily resistive in character, but with reactive properties at high frequencies. In the forward-gain region the operation of a transistor can be described qualitatively by considering the collector-to-base circuit as a semiconductor diode, normally polarized in the reverse direction. Control of conduction in this circuit is provided by charge carriers (holes or electrons) injected into the base by the emitter. The anode-to-cathode circuit of a vacuum triode is a vacuum diode, normally polarized in the forward direction. Control of anode current is accomplished by modifying the potential distribution between anode and cathode. The voltage applied to the grid relative to the cathode provides the mechanism for varying this potential distribution.

6.2 Triode Structure

Vacuum triodes are made in a wide range of sizes and have various electrode configurations. A subminiature tube is about one inch long and has a diameter comparable to that of an ordinary pencil. The dimensions of a large transmitting tube are more likely to be expressed
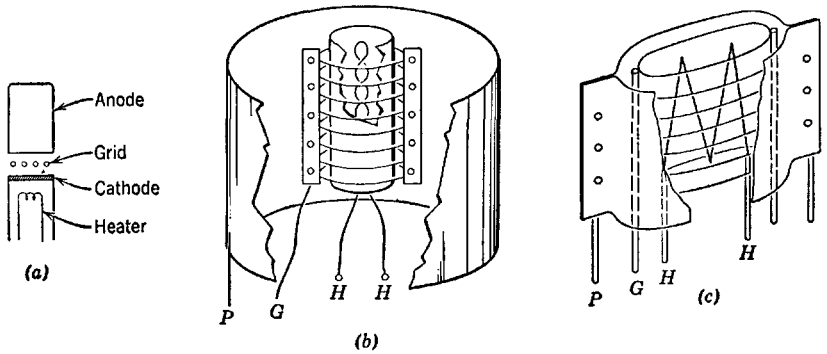
Fig. 6.1. Typical triode electrode structures.

in feet. The size is determined primarily by the power levels required in a given application. The great majority of electronic control valves operate at power levels of a few watts, but large high-power tubes used

P = plate or anode
G = grid
F = filament
(a)

P = plate or anode
G = grid
K = cathode
H = heater
(b)

(c)
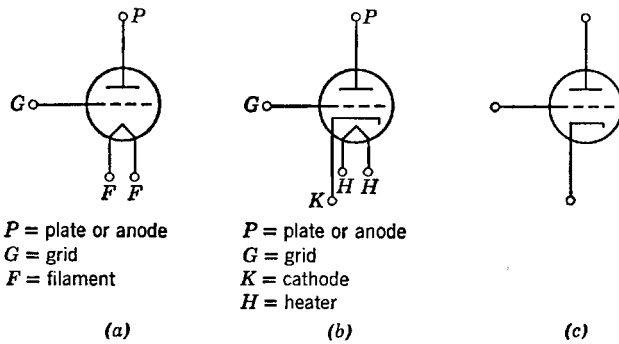
Fig. 6.2. Triode circuit symbols.

in broadcast transmitters may be required to handle many kilowatts of power.

A few of the more common electrode structures are shown in Fig. 6.1. The cathode, grid, and anode may be parallel planes, as shown in *(a)*, or

concentric cylinders, as in (b). Triodes may have a filamentary cathode, as shown in (c). Elliptical cross sections are also used, as are various hybrid combinations of simple geometrical shapes.

Circuit symbols commonly used to represent a vacuum triode are shown in Fig. 6.2. The symbol for a triode using a heated filament as the electron emitter is shown in (a). The symbol for a triode with an indirectly heated cathode is shown in (b). For our purposes, the abbreviated symbol shown in (c) is most convenient. The heater is implied but not shown, since the heater power merely maintains the cathode at an operating temperature suitable for thermionic emission of electrons.

6.3   Current Versus Voltage Relations

Conventional designations and reference directions for total electrode currents and voltages are shown in Fig. 6.3(a). The triode is a three-terminal circuit element, hence we can specify terminal behavior by means of four variables: e_b, e_c, i_b, and i_c. One commonly used functional representation of triode terminal relations is given in (b). The graphs of these functions, shown in (c), are static curves typical of the d-c or low-frequency behavior of a vacuum triode. The qualitative form is independent of triode size or electrode geometry. The plot of i_c vs. e_c is commonly called the "grid family," and the plot of i_b vs. e_b is called the "plate family."

For fixed values of e_c (zero or negative), the plate circuit conduction (i_b vs. e_b) is like that of a diode in series with a voltage source proportional to e_c. Thus, for a constant value of i_b, the voltage e_b is nearly linearly dependent on e_c, for e_c less than zero. For e_c greater than zero, the variation of e_b with e_c is smaller and not nearly as constant. Each line of the grid family also resembles a diode curve. The grid and cathode constitute an auxiliary diode whose conduction is somewhat influenced by the anode voltage e_b. The influence of e_b on i_c is considerably less than the influence of e_c on i_b.

On the basis of the above consideration, we can postulate an approximate model of the triode, as indicated in Fig. 6.3(d). The grid-to-cathode circuit can be represented by a simple vacuum diode, and the plate-to-cathode circuit by a vacuum diode in series with a source proportional to e_c. For e_c less than zero, the proportionality factor μ is very nearly constant.

The simple piecewise-linear models shown in Fig. 6.4, together with their graphical representations, are obtained by replacing the vacuum

$$i_c = f_c(e_c, e_b)$$
$$i_b = f_b(e_b, e_c)$$

$$\mu = - E_b/E_c \text{ for } e_c < 0$$
$$\mu \to 0 \text{ for } e_c \text{ increasingly positive}$$

$$e_b \approx - \mu e_c + (i_b/K)^{2/3}$$
$$\text{or} \quad i_b \approx K(e_b + \mu e_c)^{3/2}$$
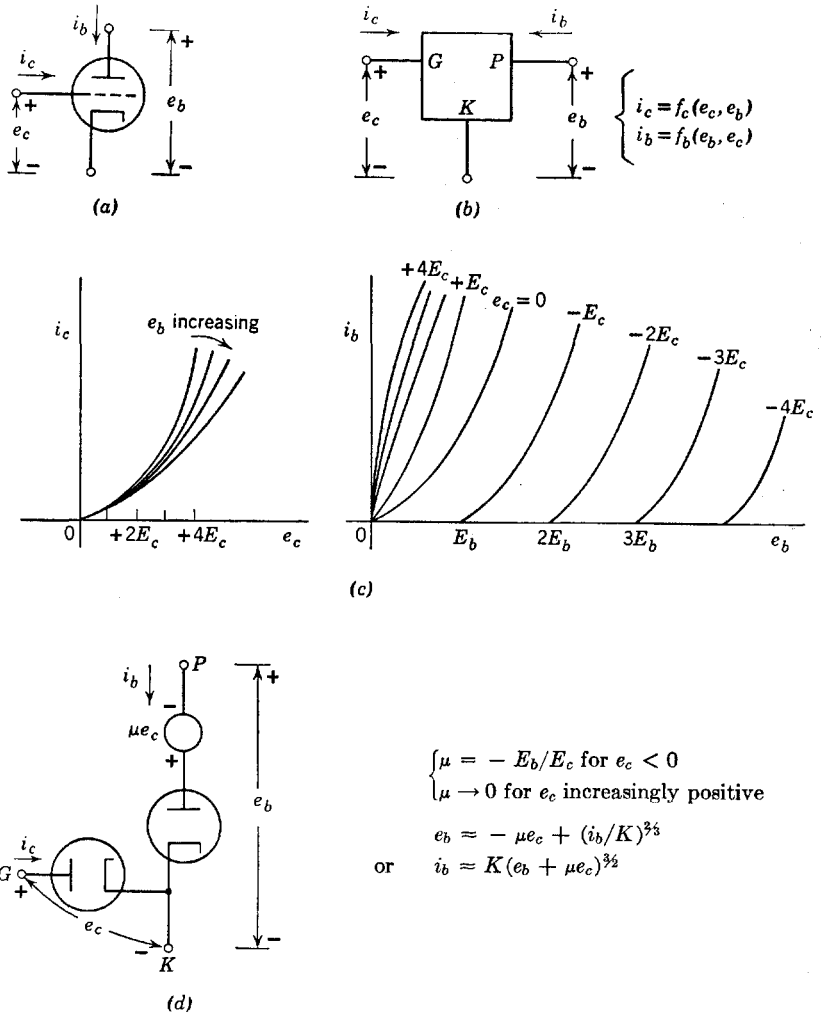
Fig. 6.3. Qualitative triode curves.

diodes by ideal rectifiers or piecewise-linear diodes. These curves and circuit models are basically the same as the piecewise-linear transistor curves and models, discussed in the preceding chapter.

Before proceeding with a more quantitative discussion of triode models and circuits, let us examine the physical basis for the current-versus-voltage relations. As we shall see in the succeeding articles, triode behavior can be explained by an extension of the results obtained in Chapter 2 for conduction in a vacuum diode.

$$\begin{cases} i_c = 0 \text{ for } e_c < 0 \\ e_c = 0 \text{ for } i_c > 0 \end{cases}$$

$$\begin{cases} e_b = -\mu e_c \text{ for } i_b > 0 \\ i_b = 0 \text{ for } e_b < -\mu e_c \end{cases}$$

(a)

$$\begin{cases} i_c = 0 \text{ for } e_c < 0 \\ e_c = i_c r_g \text{ for } e_c > 0 \end{cases}$$

$$\begin{cases} e_b = i_b r_p - \mu e_c \text{ for } i_b > 0, e_c \le 0 \\ i_b = 0 \text{ for } e_b < -\mu e_c \text{ or } e_b < 0 \end{cases}$$

(b)
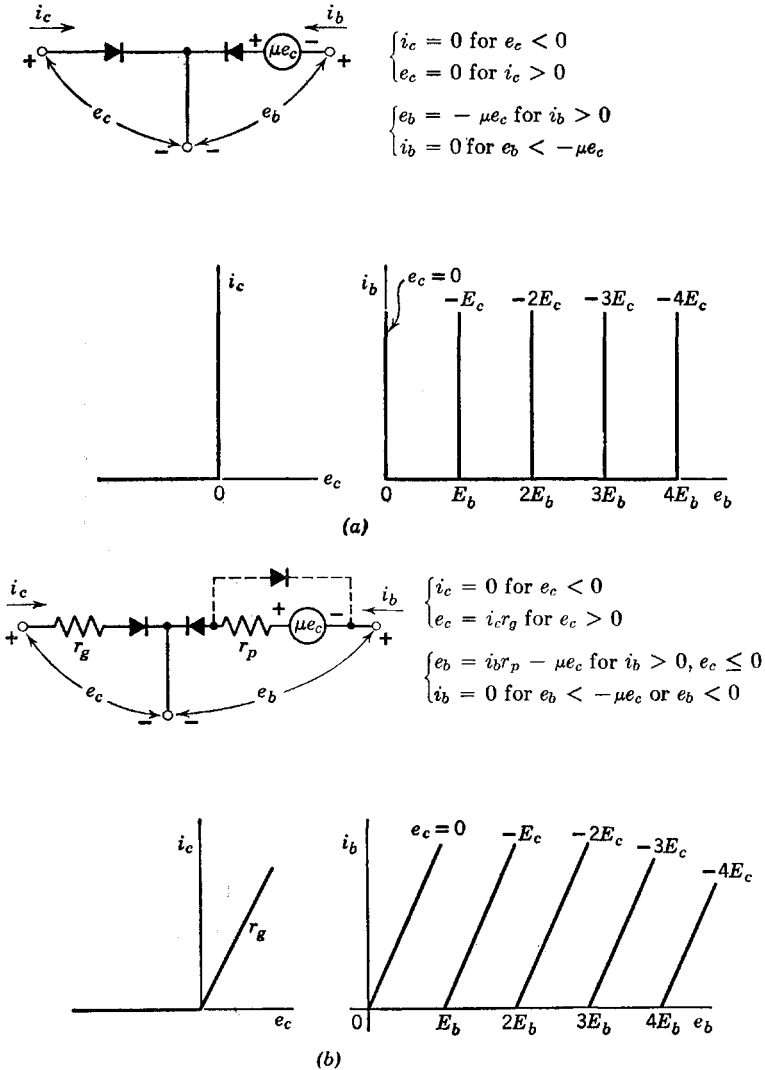
Fig. 6.4. Piecewise-linear triode models.

6.4 The Cold Parallel-Plane Triode

As mentioned above, the applied grid voltage affects the electric potential distribution in the interelectrode space and thereby influences the flow of electrons from cathode to plate. For a more detailed explana-

tion of this effect, let us first examine the electric potential distribution within a "cold" triode with the heater disconnected so that the cathode emits no electrons. The cold tube, having no interelectrode space charge, is essentially a three-electrode capacitor in which the inter-

Fig. 6.5. Potential distribution in a cold-cathode planar triode.

electrode electric fields are governed by linear electrostatic laws. The applied electrode voltages determine the interelectrode fields of the cold tube, from which we can deduce something about the amount of current which flows when the cathode is heated. For simplicity, we shall choose parallel-plane geometry with electrodes of infinite extent. This is an

idealization of the parallel-plane triode illustrated in Fig. 6.1(a). The anode and cathode of our structure are assumed to be metallic planes, and the grid structure consists of parallel rods or wires, as shown in Fig. 6.5(a). The grid wires are spaced a distance s apart and have a diameter d.

$$-C_{kg}e_c - C_{pk}e_b = Q_k = -C_k\phi_c$$
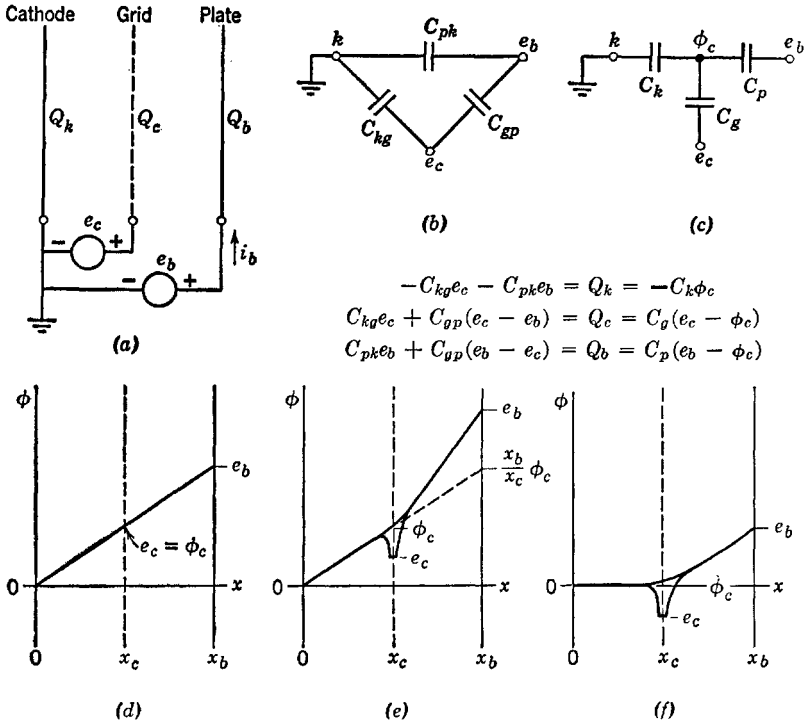
$$C_{kg}e_c + C_{gp}(e_c - e_b) = Q_c = C_g(e_c - \phi_c)$$

$$C_{pk}e_b + C_{gp}(e_b - e_c) = Q_b = C_p(e_b - \phi_c)$$

Fig. 6.6. The electrostatic model of the cold triode.

The grid wires are fixed at a potential e_c with respect to the cathode, and the anode is fixed at a potential e_b with respect to the cathode. In the absence of the grid wires, the potential varies linearly between cathode and anode, as shown by the dotted line ϕ' in Fig. 6.5(b). With the grid wires present, the potential $\phi(x,y)$ drops sharply in the vicinity of each wire ($y/s = 0, 1, 2, \ldots$), as shown in Fig. 6.5(b), and negative induced charge appears on the grid wires. Midway between grid wires ($y/s = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$), the potential curve is relatively smooth. At distance greater than s to the left or right of the grid plane, the potential approaches a curve corresponding to a uniform planar distribution of grid charge. A sketch indicating a uniform sheet of charge in the grid

plane is shown in Fig. 6.5(c), and the corresponding distribution of potential is shown in (d).

When voltages e_c and e_b are applied to the parallel-plane cold triode, Fig. 6.6(a), induced charges Q_k, Q_c, and Q_b appear on the electrodes. We assume here that each electrode has unit area and that the spacing is small, so that fringing at the edges of the electrodes is a negligible effect. Since the charge-voltage relationships are linear, we can always represent the cold tube as the system of three capacitors shown in Fig. 6.6(b). This is the *general circuit model* for any configuration of three perfect conductors in free space; one grounded and the other two driven by d-c voltages.

A three-terminal "mesh" or "delta" network of three capacitances is always replaceable by an equivalent "star" configuration, as indicated in Fig. 6.6(c). The pertinent relationships are

$$C_{gp} = C_g C_p / (C_g + C_p + C_k) \qquad (6.1)$$

and, by symmetry, two similar forms for C_{pk} and C_{kg}. These relationships can be deduced from the charge-voltage equations given in Fig. 6.6. Thus the cold triode can be represented by the circuit model in Fig. 6.6(c). We must now identify the potential ϕ_c appearing in the circuit model and relate this potential to the electric field at the cathode to deduce the amount of cathode current which will flow when the heater is reconnected.

Assume, first, that Q_c is zero. With no charge on the grid, the grid becomes "electrically transparent" and the potential curve has the same slope throughout the cathode-to-plate region as indicated in Fig. 6.6(d). Hence $e_c/x_c = e_b/x_b$. By inspection of the circuit model, we see that Q_c equal to zero implies that e_c is equal to ϕ_c, and this identifies ϕ_c in Fig. 6.6(d). Now let us *decrease* e_c exactly *one* volt and, at the same time, *raise* e_b just enough to keep ϕ_c unchanged, as shown in Fig. 6.6(e). This particular increase in e_b we shall define as μ, the *electrostatic amplification factor* or *screening factor* of the triode. By inspection of the circuit model, Fig. 6.6(c), we see that

$$\mu = C_g / C_p \qquad (6.2)$$

Moreover, it follows from the star-mesh relationships that

$$C_g / C_p = C_{gk} / C_{pk} \qquad (6.3)$$

Hence

$$\mu = C_{gk} / C_{pk} \qquad (6.4)$$

Finally, from the above definition of μ and from Fig. 6.6(e),

$$e_b - (x_b/x_c)\phi_c = \mu(\phi_c - e_c) \qquad (6.5)$$

so that

$$\phi_c = (e_b + \mu e_c)/[\mu + (x_b/x_c)] \tag{6.6}$$

Figure 6.6(f) offers a slightly different, but equivalent, interpretation of the triode parameter μ. We can say that μ is the plate voltage required to produce *zero* electric field at the cathode when the grid is *one* volt negative. It is apparent from this interpretation that μ remains unchanged if we move the cathode to a different location, keeping the grid and plate fixed. In other words, μ depends upon the grid-to-plate separation $x_b - x_c$ and upon the grid-wire diameter d and spacing s, but does *not* depend upon the grid-to-cathode separation. The calculation of μ from $x_b - x_c$, d, and s is a problem in electrostatics which, for certain geometries, is often very conveniently solved by conformal mapping methods.

6.5   The Thermionic Triode

When the cathode is heated to operating temperature, thermionic emission provides large numbers of free electrons in the space between cathode and anode. As in the case of the vacuum diode, the anode current of a triode in normal operation is limited by space charge rather than by temperature saturation. The space charge due to the cloud of electrons in the interelectrode space modifies the potential distribution that exists for the cold-cathode condition.

A qualitative sketch of potential distribution in a hot triode is shown in Fig. 6.7(a). The relative magnitude of ϕ_c and the grid-to-cathode spacing x_c have both been exaggerated in order to show the curvature of the potential distribution in this region. The small sharp potential dip just outside the cathode is attributable to initial emission velocities. The potential minimum normally occurs very close to the actual cathode and is only a few volts below cathode potential. This minimum is often called a *virtual cathode*. Since most electrons reach the potential minimum with relatively low velocity, the effect is like that of a slightly negative cathode located at the potential minimum and emitting zero velocity electrons. For this discussion we shall assume that the location and potential of the virtual cathode coincide with those of the actual cathode.

Under the usual operating conditions, e_b is large and positive, whereas e_c is small and negative, and the current is space-charge limited. Electrons are subject to a large acceleration in the grid-plate region because of the large positive value of e_b. They arrive at the anode with a velocity

v_b, determined by equating kinetic energy $(\frac{1}{2}mv_b{}^2)$ to potential energy (qe_b). The current density is the product of the electronic charge, the electron density and the electron velocity. A given anode current therefore produces little space charge in regions where velocity is high. Thus the potential curve near the anode is approximately a straight line, as it was for the cold (space-charge-free) triode.
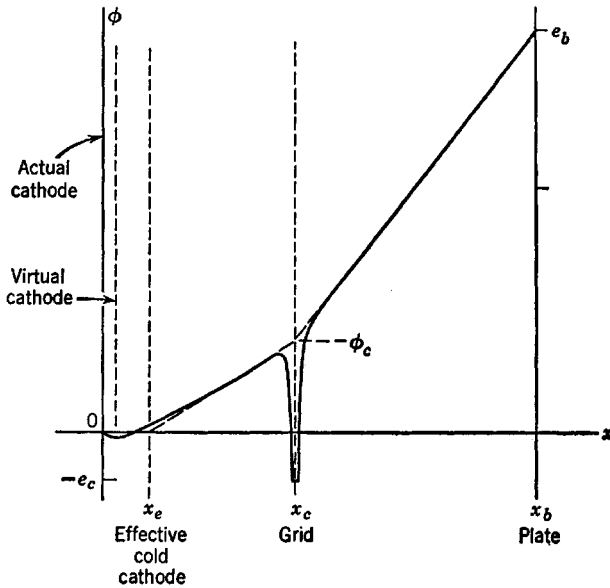
(b) $\quad \phi_c = \dfrac{e_b + \mu e_c}{\left[\dfrac{x_b - x_e}{x_c - x_e}\right] + \mu}$

(c) $\quad i_b = K(e_b + \mu e_c + k)^{3/2}$

Fig. 6.7. Potential distribution in a triode with heated cathode.

Near the heated cathode, the electron velocities are small and space charge is appreciable. The space charge produces curvature in the potential distribution as in the space-charge-limited diode. For the calculation of potentials farther to the right, the cloud of space charge near the cathode can be represented by a concentrated sheet of charge located at some point x_e, as indicated in Fig. 6.7(a). In other words, a heated cathode can be replaced by a cold cathode located closer to the grid without altering the potential distribution in the grid-plate region

where space charge is negligible. For plane-parallel geometry, the potential ϕ in a hot space-charge-limited triode varies approximately as the four-thirds power of the distance from the cathode. (This is a consequence of the plane-parallel-diode law. Since the current density is independent of x, we have $\phi^{3/2}$ proportional to x^2, so that ϕ is proportional to $x^{4/3}$.) It follows that $x_e = x_c/4$. Hence the equation for ϕ_c given in the preceding article applies to the triode with heated cathode, provided x_c and x_b are each decreased by an amount x_e, as indicated in Fig. 6.7(b).

The region between the cathode and grid of the triode forms a diode whose anode voltage is the effective grid-plane potential ϕ_c. Thus the cathode current i_k of the triode is proportional to the three-halves power of $(\phi_c + \phi_0)$, where ϕ_0 is a correction voltage accounting for the small initial-velocity potential dip, and also for contact-potential differences between the electrodes. For zero grid current (e_c negative) we have i_b equal to i_k so that anode current i_b is proportional to $(\phi_c + \phi_0)^{3/2}$. Since ϕ_c is proportional to $e_b + \mu e_c$, as indicated in Fig. 6.7(b) the anode current i_b can be expressed by the relation given in Fig. 6.7(c), where K, k, and μ are constants of the tube. This relation holds for any triode geometry (parallel-plane, cylindrical, etc.), provided (1) the grid current is negligible, (2) the anode current is space-charge limited, (3) the grid-to-plate space charge is negligible, and (4) the grid-wire spacing is uniform and small compared with interelectrode spacing. The quantity k may amount to several volts, since it includes a grid-correction voltage multiplied by μ. However, a satisfactory approximation is usually obtainable without including k, so that

$$i_b = K(e_b + \mu e_c)^{3/2} \qquad (6.7)$$

6.6 Plate-Current Curves for a Typical Triode

The relation for plate current given in Fig. 6.7(c) confirms the plate-circuit model postulated in Fig. 6.3(d). For a fixed value of grid voltage e_c, the plate current varies as in a biased vacuum diode. The plot given in Fig. 6.8(a) shows this relation for several values of e_c. Note that the curves are spaced a distance μE along the plate-voltage axis, where E is the grid-voltage increment. On this plot the correction voltage k (for initial velocities and contact potential) is assumed to be zero, so the curve for $e_c = 0$ passes through the origin.

The plot of Fig. 6.8(b) shows experimental plate-current curves for a typical triode, with the theoretical three-halves power curves of (a) shown as dotted lines for comparison. The constants K and μ must be

chosen appropriately to fit the curves of a given triode. The difference between actual measured curves and the theoretical curves can be visualized more readily from the plot of $i_b^{2/3}$ vs. e_b shown in (c). Here
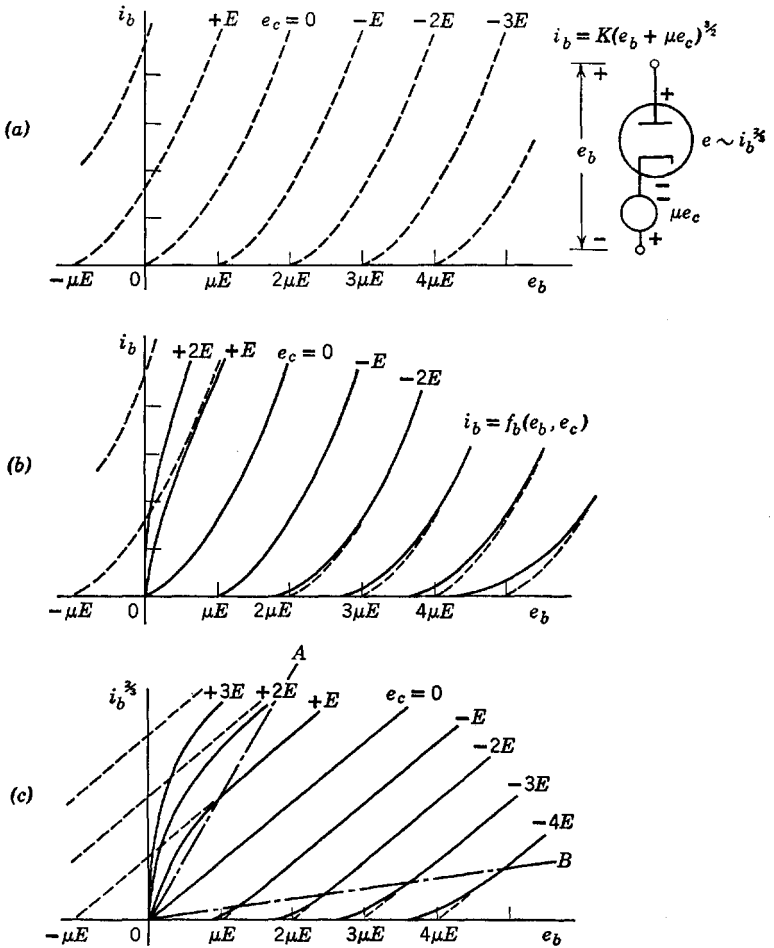

Fig. 6.8. Theoretical and actual plate-current curves.

the theoretical curves are straight lines. Within the central region contained between lines OA and OB, the experimental curves follow the theoretical curves very well. Above line OA, the actual plate current is less than that given by the three-halves power idealization because the potential distribution is modified by the presence of substantial space charge in the grid-to-plate region. When e_b is negative, the plate current vanishes, since the emitted electrons are retarded to zero velocity

($\phi = 0$) before reaching the plate. The entire cathode current is then collected by the grid. Below line OB in Fig. 6.8(c), the actual current lies above that predicted by the idealized theory. This departure can be attributed primarily to unavoidable nonuniformities in grid-wire spacing and grid-to-plate spacing. In effect, the "local" μ varies slightly from one part of the grid to another. The plate current can be visualized as the sum of the currents in a number of smaller triodes connected in parallel, each with a slightly different value of μ. The total current of such a parallel combination does not cut off abruptly as the plate voltage is lowered, because of the dispersion of cutoff voltages for the various small triodes. In some tubes the grid-wire spacing is purposely made nonuniform to make the plate-current cutoff gradual. These tubes, called remote-cutoff tubes (in contrast to sharp cutoff), are used extensively in applications where control of gain by means of grid bias is desired. A common example of such use is an automatic gain control circuit in a radio receiver.
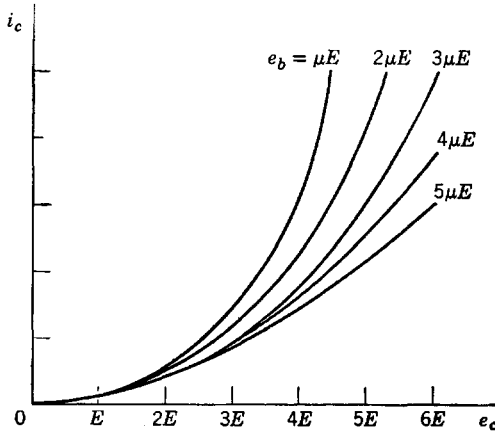
6.7   Grid-Current Curves for a Typical Triode

The grid and cathode of a triode form a subsidiary diode which conducts current for positive values of grid voltage. Grid-current curves for a typical triode are shown in Fig. 6.9(a), where plate voltage e_b is the parameter on a plot of grid current i_c versus grid voltage e_c. For e_c small and e_b large, the grid current for constant e_c is nearly independent of plate voltage, and a simple diode model represents the grid-to-cathode behavior. Figure 6.9(b) shows another way of presenting the same data. Below line OA the grid current is essentially independent of plate voltage. Either of these presentations of the input data can be combined with the plate-current curves to specify the resistive behavior of a vacuum triode. Another useful form for the plate-circuit data is a plot of e_b vs. e_c, with i_b as a parameter. Nonlinear vacuum-tube curves are useful for locating operating points and for determining numerical values for the parameters that appear in circuit models. However, piecewise-linear or linear-incremental models will be our main concern in this chapter.

6.8   Piecewise-Linear Approximations to Triode Curves

The qualitative character of typical triode curves suggested the resistive models postulated in Article 6.3. These included the simple piecewise-linear models shown in Fig. 6.4. Our brief discussion of the

theoretical basis for the resistive properties of triodes indicates that the model of Fig. 6.4 (b) should provide a fairly good quantitative approxima-

(a)
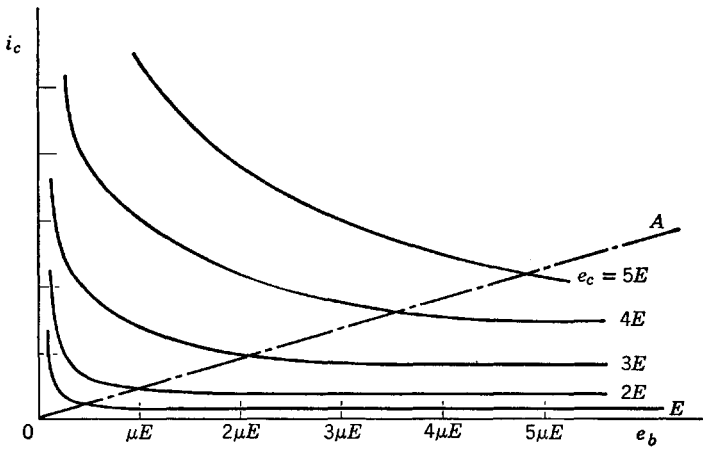
(b)

Fig. 6.9. Typical grid-current curves.

tion to triode curves. If a better fit is desired, we can make use of the fact that the triode has the same plate current versus plate voltage relation as the vacuum diode. Thus any of the vacuum diode models developed in Chapter 3 can be easily modified for use as triode models.

For example, the model of Fig. 6.10(a) is an adaptation of the model in Fig. 3.13(a). The curves in Fig. 6.10(b) have sufficient flexibility to match triode plate curves over a large portion of the i_b vs. e_b plane. Letting the intercept E_0 equal zero reduces the model to the simpler one postulated in Fig. 6.4(b). It is repeated for convenient reference in
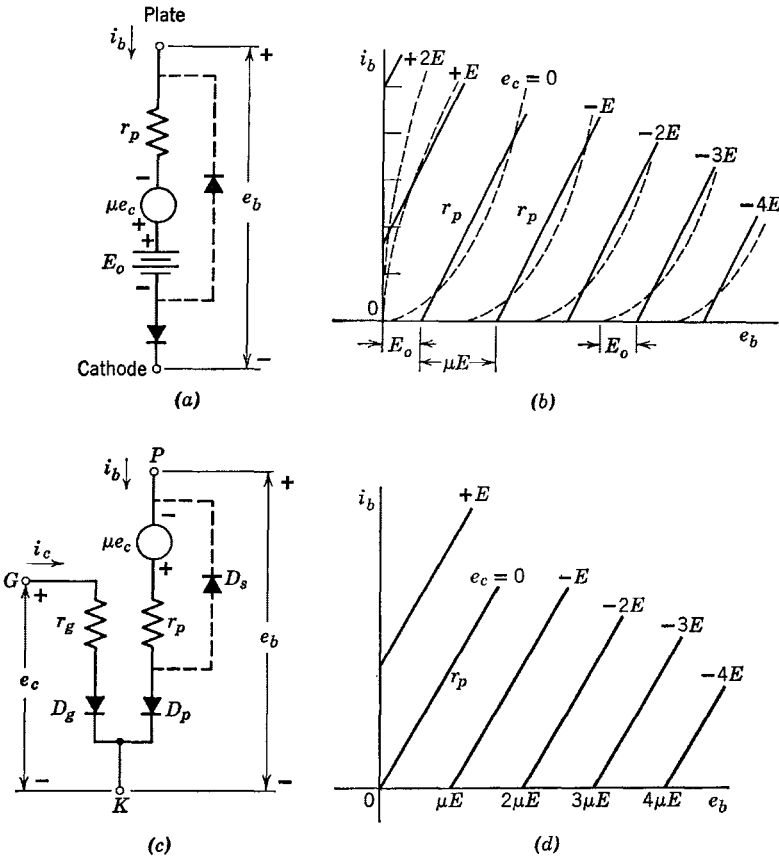

Fig. 6.10. Approximating triode curves with piecewise-linear models.

Fig. 6.10(c). The parameters μ and r_p can usually be chosen to provide a good fit to the actual curves over most of the negative-grid region. The model is not particularly accurate in the positive-grid region, nor does it do more than provide a crude approximation to the grid-current curves. Models more suitable for the positive-grid region can be devised, but the simple one suffices for most analysis or design problems involving triode circuits.

Referring to the model of Fig. 6.10(c), we note that r_g and ideal rectifier D_g form a piecewise-linear approximation to the grid-cathode diode, whereas r_p and D_p approximate the plate-cathode diode. The dependent voltage source μe_c shifts the plate-circuit curve along the plate-voltage axis in the manner predicted by the effect of e_c upon ϕ_c and therefore upon i_b. The ideal rectifier D_s is included to prevent the dependent source from driving e_b negative. This diode roughly approximates the reduction in the effective value of μ which takes place in the actual triode for low values of plate voltage. The ideal rectifier D_s is connected by dotted lines in Fig. 6.10(c), because it is somewhat less directly related to the physical behavior of triodes than are the other elements in the model. We could realize the same set of curves [Fig. 6.10(d)] by omitting D_s from the model and specifying $\mu = 0$ for $e_b < 0$. This restriction on the source μe_c is, in effect, a piecewise-linear (actually a stepwise-linear) approximation to the gradual reduction in μ as e_b approaches zero.

With three diodes in the model, there are eight diode states; four with D_s open, and four with D_s closed. We can eliminate the four with D_s closed as being of no practical interest. Three of the remaining four diode states define the regions of major interest in triode operation. With D_g and D_p off, both grid current and plate current are zero. This current cutoff (state I) exists for e_c more negative than $-e_b/\mu$. For values of e_c between 0 and $-e_b/\mu$, D_p is *on* and D_g is *off*. This results in plate current but no grid current, and is the normal forward-gain or amplification region (state II). For e_c greater than 0, both D_p and D_g are *on*; hence both plate current and grid current exist. This is variously called the grid-current, positive-grid, or saturation region, and can be designated as state III. The remaining possibility (state IV) calls for D_g *on* and D_p *off*. In this state, which cannot be achieved without applying an appropriate source between plate and cathode to make e_b negative, a reverse gain occurs in the actual triode, and our simple model does not describe the behavior adequately. Since this state is of interest only in rather special instances, it will not be discussed in detail. It can be represented by a generator $\mu' e_b$ in the grid circuit.

6.9 Linear Incremental Triode Models

In some circuit applications the range of values for terminal currents and voltages is contained within a single state of the piecewise-linear model, that is, D_p and D_g are at all times either *on* or *off*, and can be

replaced by short circuits or open circuits. The piecewise-linear model then reduces to a linear model. The linear model can be further simplified by ignoring the fixed voltages and currents at the operating point (established by d-c polarizing sources) and considering only the incre-
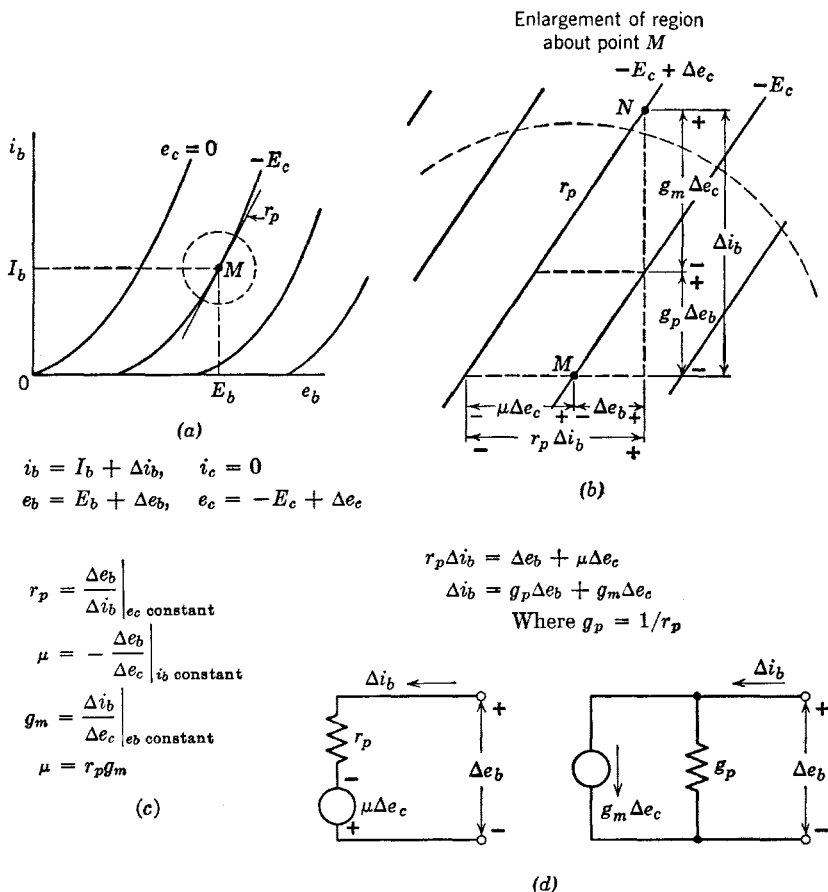

Enlargement of region about point M

(a)

$i_b = I_b + \Delta i_b, \quad i_c = 0$

$e_b = E_b + \Delta e_b, \quad e_c = -E_c + \Delta e_c$

(b)

$$r_p = \frac{\Delta e_b}{\Delta i_b}\bigg|_{e_c \text{ constant}}$$

$$\mu = -\frac{\Delta e_b}{\Delta e_c}\bigg|_{i_b \text{ constant}}$$

$$g_m = \frac{\Delta i_b}{\Delta e_c}\bigg|_{e_b \text{ constant}}$$

$$\mu = r_p g_m$$

(c)

$r_p \Delta i_b = \Delta e_b + \mu \Delta e_c$

$\Delta i_b = g_p \Delta e_b + g_m \Delta e_c$

Where $g_p = 1/r_p$

(d)

Fig. 6.11. Linear incremental model of plate circuit.

mental variations about the operating point. The shift from total quantities to incremental quantities amounts to shifting the origin of coordinates to the operating point. In terms of the plate curves, the operation takes place in a small region containing the operating point M, as shown in Fig. 6.11(a). The plate curves in this restricted region appear as nearly straight, parallel, equidistant lines, as shown in (b).

The values of r_p and μ can be chosen for maximum accuracy in the small region involved. In general, these values may differ slightly from those used in the piecewise linear model, which must provide a reasonable approximation over the entire first quadrant of the i_b vs. e_b plane.

The slope of the plate curves determines the parameter r_p in the incremental model. This slope is usually taken as the tangent to the curve passing through the operating point. The horizontal spacing between lines can be specified by the parameter μ, or the vertical spacing can be specified by a parameter g_m. The definitions of r_p, μ, and g_m in terms of incremental quantities are given in Fig. 6.11(c). From the expressions, or from the geometry indicated in (b), it is evident that $\mu = r_p g_m$.

Let us now consider the relations among the incremental quantities as the instantaneous values of currents and voltages change. Referring to Fig. 6.11(b), assume that the change corresponds to a shift from point M to the nearby point N. From the geometry of the figure, we see:

$$\Delta e_b = r_p \, \Delta i_b - \mu \, \Delta e_c \qquad (6.8)$$

and

$$\Delta i_b = g_p \, \Delta e_b + g_m \, \Delta e_c \qquad (6.9)$$

The two circuits shown in Fig. 6.11(d) are pictorial representations of the two equations. Either one can be taken as an incremental model for the triode plate circuit, since the two circuits are equivalent.

The incremental circuit can also be obtained by application of differential calculus. Referring to the general piecewise-linear model of Fig. 6.10(a) and (b), we see that

$$e_b = r_p i_b - \mu e_c + E_0 \qquad (6.10)$$

expresses the relation between e_b, i_b, and e_c. Assume now that both μ and r_p have been chosen to match slope and spacing in some particular region. Total differentiation yields

$$de_b = r_p \, di_b - \mu \, de_c \qquad (6.11)$$

which is the same as the corresponding incremental relation, Eq. 6.8.

The incremental plate-circuit equation can also be obtained by expanding the plate current in a Taylor's series about the desired operating point $(E_b, -E_c)$. This derivation is given in Fig. 6.12. The series approach is not needed here, but it is important because of its generality. It is useful for incremental analysis when linear analysis is inadequate, since higher-order terms can be included. It is also helpful in analyzing the linear incremental behavior of tubes having more electrodes (tetrodes, pentodes, etc.). Since the behavior of these

tubes involves more variables, a simple geometrical interpretation such as that of Fig. 6.11(b) is not possible.

At this point we shall introduce symbols that have become associated with incremental grid and plate variables through common usage.

$$i_b(e_b, e_c) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_{jk}(e_b - E_b)^j(e_c + E_c)^k$$

(a)
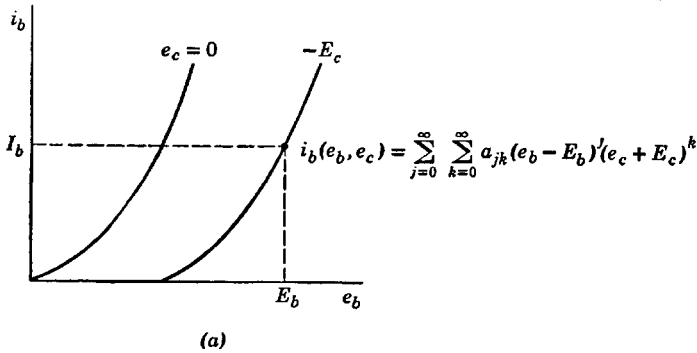
(b)

$$a_{00} = i_b(E_b, -E_c) = I_b$$

$$a_{10} = \frac{\partial i_b(E_b, -E_c)}{\partial e_b} = g_p$$

$$a_{01} = \frac{\partial i_b(E_b, -E_c)}{\partial e_c} = g_m$$

For small excursions from the operating point,

(c) $$i_b(e_b, e_c) \approx a_{00} + a_{10}(e_b - E_b) + a_{01}(e_c + E_c)$$

(d) $$i_b(e_b, e_c) - I_b \approx g_p(e_b - E_b) + g_m(e_c + E_c)$$

(e) $$\Delta i_b \approx g_p \Delta e_b + g_m \Delta e_c$$

Fig. 6.12. Series representation of triode plate curves.

These symbols and their equivalents are:

$$i_g = \Delta i_c \qquad (6.12)$$

$$e_g = \Delta e_c \qquad (6.13)$$

$$i_p = \Delta i_b \qquad (6.14)$$

$$e_p = \Delta e_b \qquad (6.15)$$

Let us now consider more complete models than those just shown for the plate circuit alone. The circuit of Fig. 6.13(a) is a linear resistive incremental model for a triode whose grid-to-cathode voltage is negative; the circuit of (b) is for positive values of e_c. Here we have added an incremental representation for the grid circuit. Of course, the shunt

model for the plate circuit (g_p in parallel with source $g_m e_g$) could be used instead of the series model.

More complete linear models for negative and positive values of e_c are shown in Fig. 6.13(c) and (d). These include interelectrode capacitances (grid-to-cathode C_{gk}, grid-to-plate C_{gp}, and plate-to-cathode
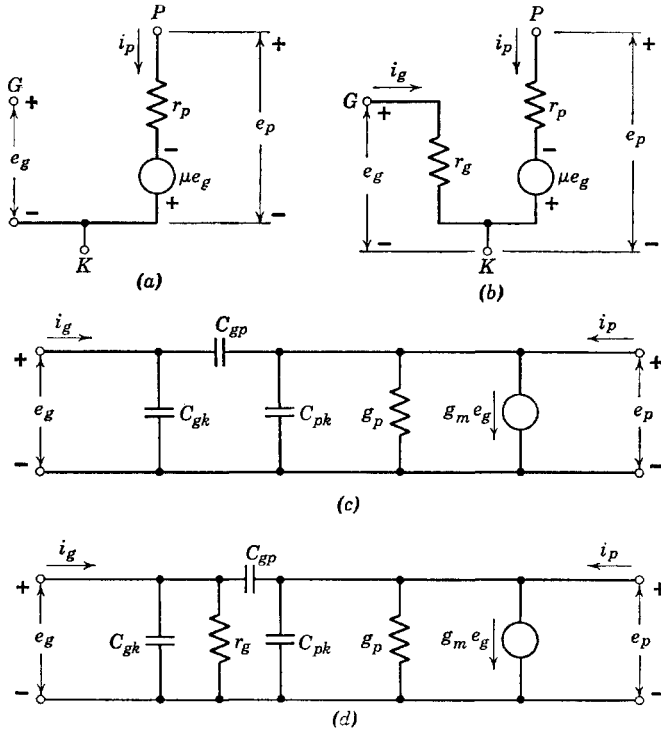
Fig. 6.13. Incremental triode models. Note: $e_g = \Delta e_c$, $i_g = \Delta i_c$, $e_p = \Delta e_b$, and $i_p = \Delta i_b$.

C_{pk}), which become important when voltage variations occur at high frequencies. As frequency is raised to still higher values, it may become necessary to introduce circuit elements representing the inductance of the leads that support the electrodes. For sufficiently high frequencies, it is also necessary to use complex values of g_m or μ in order to introduce a phase lag that accounts for the transit time of electrons through the interelectrode space. We have been vague in specifying frequencies for which interelectrode capacitances, lead inductances, or transit-time effects become important. Quantitative discussion of these effects is better deferred until required in connection with specific

circuit problems. The relative influence of the capacitances and inductances depends heavily on the values of external circuit parameters as
well as on frequency. Transit-time effects depend on tube geometry
and operating voltages.

6.10   Grounded-Cathode Amplifier Circuit

The simplest triode circuit is the grounded-cathode amplifier, sometimes called the plate-loaded amplifier. The circuit shown in Fig. 6.14 (a)
includes a triode, the plate polarizing battery E_{bb}, and the plate load
resistance R_b. Resistance R_c may be considered as part of the amplifier,
or it may be associated with the source e_1. Either the source e_1 or
auxiliary circuitry provides a d-c polarizing grid voltage $-E_{cc}$ which
establishes a quiescent operating point Q.

Let the triode plate-circuit behavior be specified by the curves of
Fig. 6.14(b). Then the load resistance R_b and source E_{bb} impose the
linear load constraint, which can be expressed by the relation

$$i_b = (E_{bb} - e_b)/R_b \qquad (6.16)$$

This equation plots as a straight line on the triode curves and is designated by R_b. The quiescent operating point Q is determined by the
condition $e_c = -E_{cc}$, together with the load constraint. By assuming
various values for e_c, we can plot the transfer curves i_b vs. e_c and e_b vs. e_c
shown in (c) and (d). Point P, the intersection of the load line and the
$e_c = 0$ line, marks the boundary between the amplification region and
the grid-current region. If we wish to relate i_b and e_b to the input voltage
e_1, rather than e_c, we must consider the effects of the source resistance
R_c. The curve of e_b vs. e_1 shown in (f) is identical with the e_b vs. e_c
curve of part (d) in the region where e_c is negative, since i_c is then
negligible and R_c has no effect. However, where e_c is positive the curves
will differ, and the triode input curves shown in part (e) must be used
to effect a solution in this region. Since $e_c = e_1 - i_c R_c$, the resistance
R_c imposes a linear constraint that can be plotted on the input curves
by locating two points. A given value of e_c determines e_b on the plate
curves. Then e_b and e_c locate a point such as A on the input curves.
A line of slope $1/R_c$ through this point locates the value of e_1 corresponding to the assumed value of e_c. In Fig. 6.14 (f) the dotted line designated
$R_c = 0$ is the same as the curve in (d), since $e_1 = e_c$ when $R_c = 0$.
The dotted line $R_c = \infty$ indicates perfect limiting at $e_c = 0$. For any
value of R_c, the positive-grid portion of the transfer curve must lie
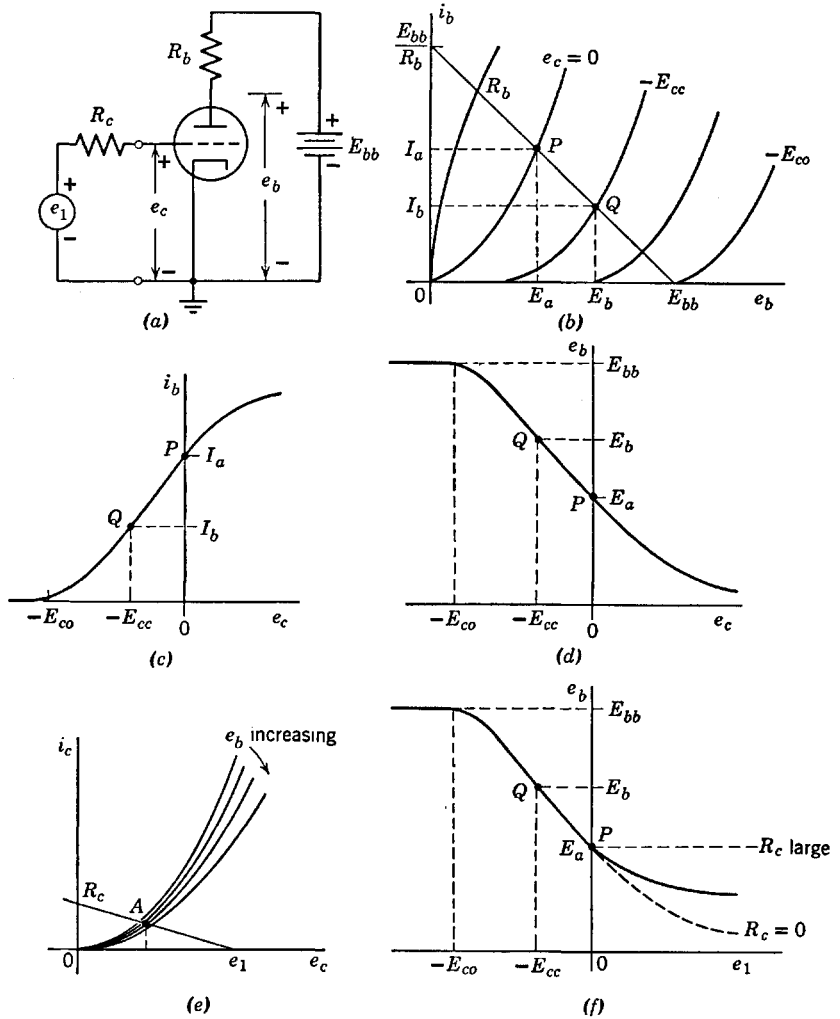between these two extremes.

Fig. 6.14. Grounded-cathode amplifier—graphical analysis using nonlinear triode curves.

The nonlinear relations shown in Fig. 6.14(c), (d), and (f) must be plotted point by point in order to obtain reasonable accuracy. The graphical triode curves, furnished by the manufacturer or measured experimentally, are the basis for determining any such relations between current and voltage variables.

For comparison with the nonlinear graphical analysis, let us carry out a graphical analysis of the same circuit, using a piecewise-linear
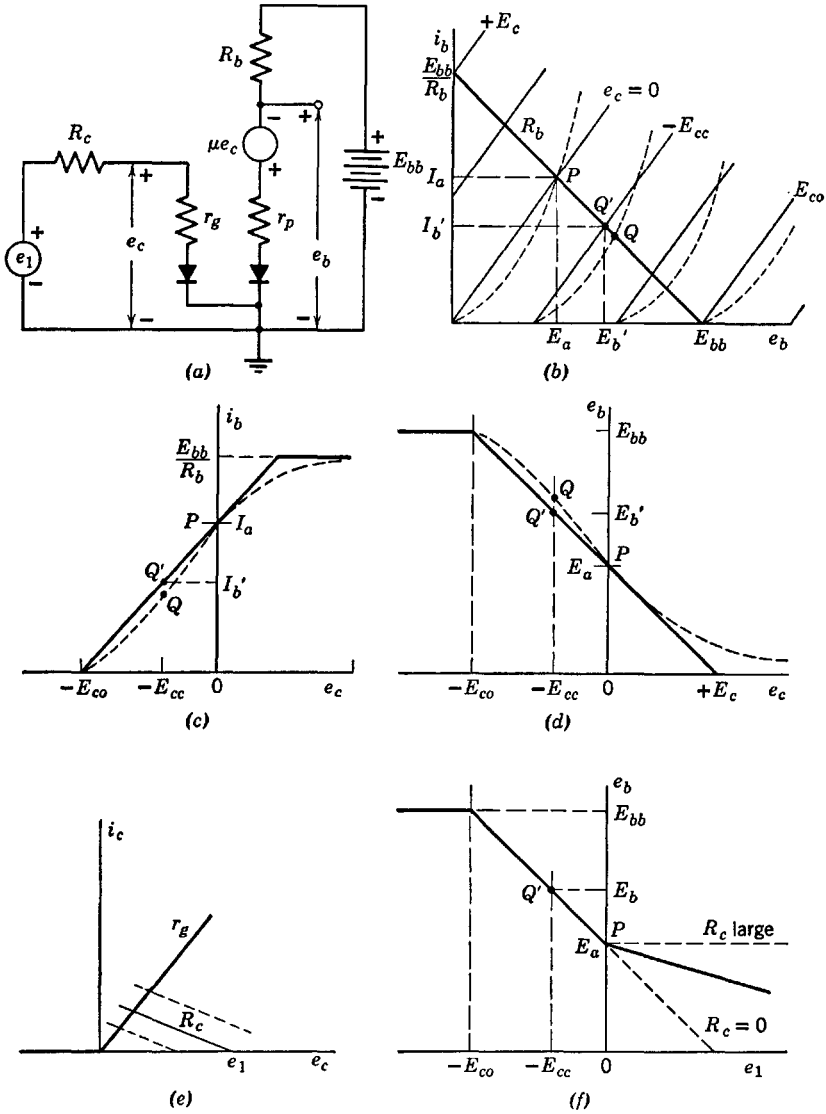
Fig. 6.15. Graphical analysis of grounded-cathode circuit using a piecewise-linear triode model.

model for the triode, as indicated by Fig. 6.15. The circuit is shown schematically in (a), and the piecewise-linear triode curves are given in (b). The actual triode curves are shown by dotted lines. In making this particular piecewise-linear approximation to the actual tube curves,

the slope r_p was made such that the $e_c = 0$ line intersects the actual $e_c = 0$ curve at the point P. Thus I_a and E_a will have the same values in both solutions. The value of μ was adjusted to yield the same grid voltage $(-E_{co})$ for plate-current cutoff when $e_b = E_{bb}$. We have thereby matched the approximation to the actual curves at the transition between states I and II and between II and III. This yields a fairly good average match over the region of major interest: namely, the amplification region (state II).
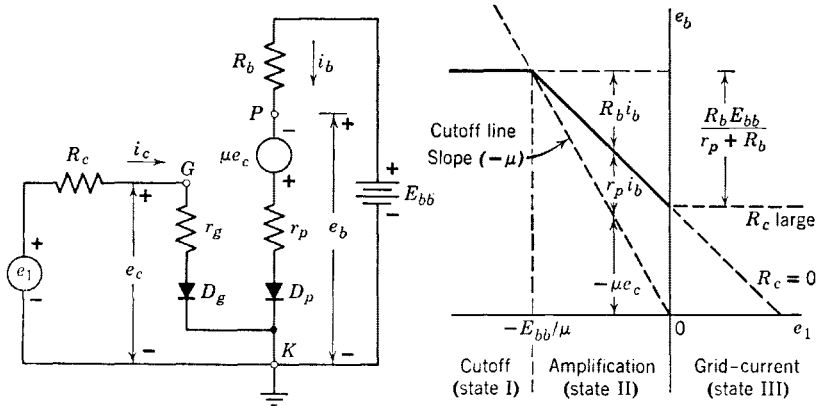
State	I	II	III
D_g	Off	Off	On
D_p	Off	On	On
$i_c =$	0	0	$e_c/r_g = (e_1 - e_c)/R_c = e_1/(R_c + r_g)$
$i_b =$	0	$\dfrac{E_{bb} + \mu e_c}{r_p + R_b}$	$\dfrac{E_{bb} + \mu e_c}{r_p + R_b} = \dfrac{E_{bb} + \left(\dfrac{\mu r_g}{r_g + R_c}\right) e_1}{r_p + R_b}$
$e_b = E_{bb} - i_b R_b =$	E_{bb}	$\dfrac{E_{bb}r_p - \mu e_c R_b}{r_p + R_b}$	$\dfrac{E_{bb}r_p - \mu e_c R_b}{r_p + R_b} = \dfrac{E_{bb}r_p - \left(\dfrac{R_b \mu r_g}{r_g + R_c}\right) e_1}{r_p + R_b}$

Fig. 6.16. Relations between total voltages and currents in grounded-cathode circuit (using a piecewise-linear triode model).

The curves of Fig. 6.15(c), (d), (e), and (f) can be compared directly with the corresponding curves of Fig. 6.14. Note that the graphical constructions for the piecewise-linear circuits are greatly simplified by the fact that all variables must be related by piecewise-linear curves, so that two points (break points) suffice to plot each segment. In

Fig. 6.15(c) and (d), the dotted lines show the nonlinear solution obtained from Fig. 6.14(c) and (d). The quiescent point Q' for the approximation does not coincide with Q for the actual curves. For some applications it may be preferable to match the piecewise-linear model to the actual curves at the desired quiescent point rather than at P, and to let other portions of the approximate transfer curves deviate from the actual ones.

The use of piecewise-linear models permits simple algebraic solution of resistive-circuit problems. Nevertheless, free-hand graphical sketches provide a worth-while visual supplement. For the grounded-cathode circuit, algebraic results can be written by inspection. They are presented in tabular form in Fig. 6.16. The sketch of e_b vs. e_1 in Fig. 6.16 shows the influence of the various circuit parameters on the dimensions of the transfer curve e_b vs. e_1. Observe from the tabulation that state I corresponds to removal of the triode from the circuit, and is analytically trivial. Nevertheless, in many circuit applications the cutoff state occurs during the cycle of operation.

6.11 Polarizing the Grounded-Cathode Circuit

The grounded-cathode circuit is useful for virtually all of the functions performed by electronic circuits; namely, voltage amplification, power amplification, wave shaping, waveform generation, modulation, and demodulation. Various modes of operation are determined by the relative values of polarizing voltages and signal voltages applied to the circuit. Energy-storage elements also play a vital role in the behavior of the circuit but detailed consideration of these effects is deferred to Chapters 8 and 9. For the present we shall use capacitances only as "ideal coupling elements" to illustrate practical methods of providing polarizing voltages and of coupling one amplifier circuit to another.

Ideal coupling implies a capacitance (or an inductance) so large that the stored energy does not change appreciably due to a-c signal variations. The capacitance is then effectively a short circuit for a-c but an open circuit for d-c. A reasonable model for the capacitor in this case is a battery whose voltage equals the capacitor voltage. Such a battery provides a zero resistance path for a-c and holds the average current at the proper value (zero).

If a polarizing battery or voltage source carries direct current in the reverse direction, the source can be replaced by a parallel combination of resistance and capacitance. The resistance is chosen to produce a voltage drop equal to the battery voltage when the specified d-c current
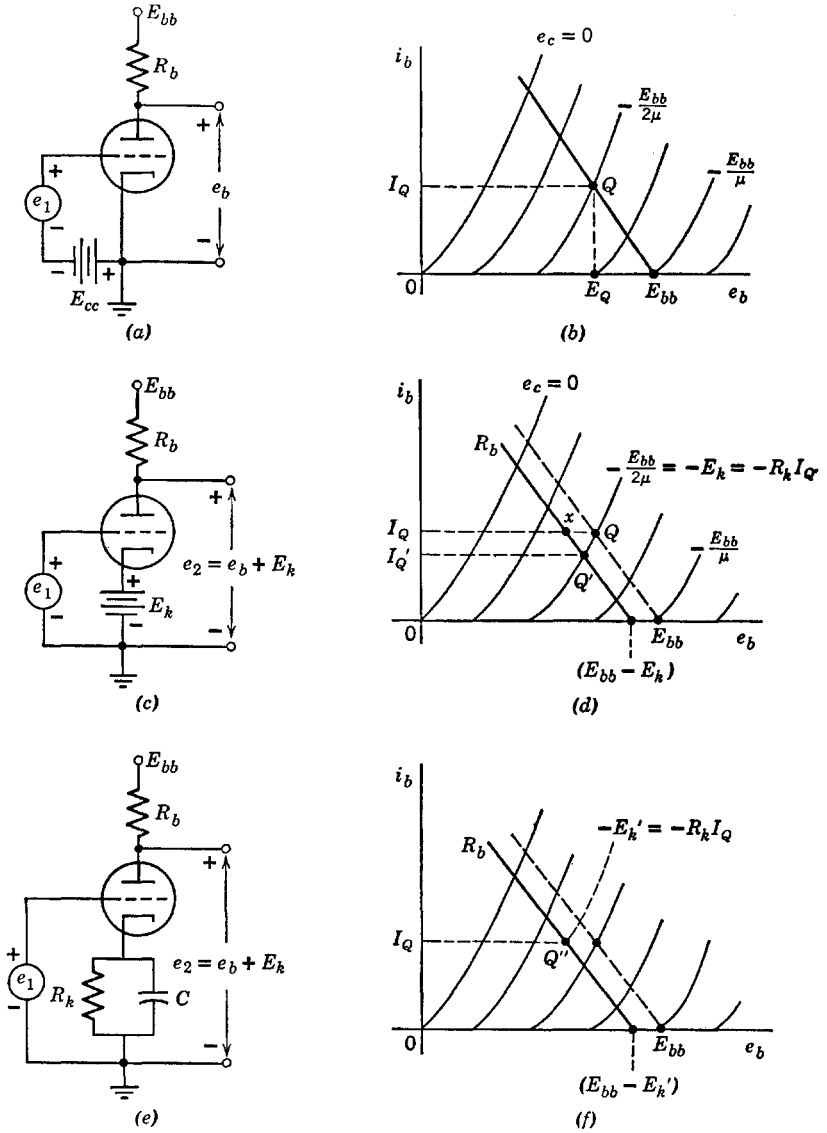
Fig. 6.17. Grounded-cathode circuit polarized for linear amplification.

is flowing. The capacitance provides a path for alternating currents and must be large enough to maintain a nearly constant voltage. Grid polarizing voltages are normally provided in this way, since it is more economical than a separate source.

The circuit shown in Fig. 6.17(a) uses a separate source E_{cc} to polarize the grid somewhere between plate-current cutoff and grid-current conduction. To provide the maximum range of linear operation the value of E_{cc} should be approximately $E_{bb}/2\mu$, which results in the quiescent point Q shown in (b) of Fig. 6.17. The circuit shown in (c) yields essentially the same result if we let E_k have the same value as E_{cc}. Since $E_{bb}/2\mu$ is appreciably smaller than E_{bb}, the operating point
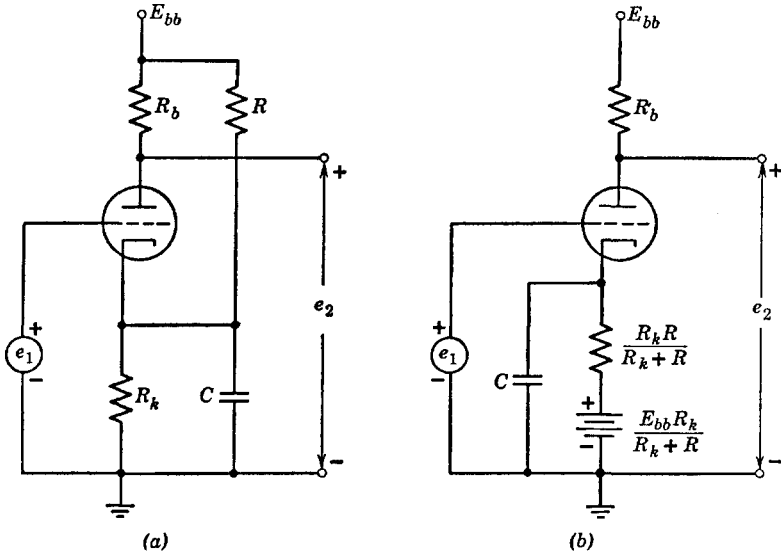
Fig. 6.18. Polarizing circuit that permits plate-current cutoff.

Q' shown in (d) differs only slightly from that of (b). Since plate current passes through E_k we can replace the battery by a resistance $E_k/I_{Q'}$ in parallel with a large capacitance, as shown in (e). Alternatively, the operating point can be maintained at the current I_Q by reducing E_k to the value E_k' as shown in (f). The value of R_k required is then E_k'/I_Q.

The values of R_k, R_b, and E_{bb} in Fig. 6.17(e) can be chosen to place the operating point at nearly any point in the first quadrant. However, since the cathode voltage depends on plate current, this circuit can never produce plate-current cutoff under quiescent conditions. Current i_b is positive for all finite values of R_k. A circuit that yields a greater range of possibilities in this regard is shown in Fig. 6.18(a). The equivalent form (b) places in evidence the effective battery produced by the current through R and R_k with plate current equal to zero.

6.12   Incremental Analysis of Grounded-Cathode Amplifier

Within a given state of the piecewise-linear model, we can further simplify calculations by using a linear incremental model. The equations relating incremental voltages and currents in Fig. 6.19(a) and (b) can be obtained by total differentiation of the appropriate equations in

$$i_g = 0; \quad e_1 = e_g$$

$$i_p = \frac{\mu e_g}{r_p + R_b}$$

$$e_p = -\frac{\mu e_g R_b}{r_p + R_b}$$

$$G = \frac{e_p}{e_g} = -\frac{\mu R_b}{r_p + R_b}$$

$$R_0 = \frac{e_p}{i_0} = \frac{R_b r_p}{R_b + r_p}$$

$$R_i = \frac{e_1}{i_g} = \infty$$

(a)

$$i_g = \frac{e_g}{r_g} = \frac{e_1 - e_g}{R_c} = \frac{e_1}{R_c + r_g}$$

$$i_p = \frac{\mu e_g}{r_p + R_b}$$

$$e_b = -\frac{\mu e_g R_b}{r_p + R_b}$$

$$G = \frac{e_p}{e_g} = \frac{-\mu R_b}{r_p + R_b}$$

$$G_1 = \frac{e_p}{e_1} = -\left(\frac{\mu R_b}{r_p + R_b}\right)\left(\frac{r_g}{r_g + R_c}\right)$$

$$R_0 = \frac{e_p}{i_0} = \frac{R_b r_p}{r_p + R_b}$$

$$R_i = \frac{e_1}{i_g} = r_g + R_c$$

(b)
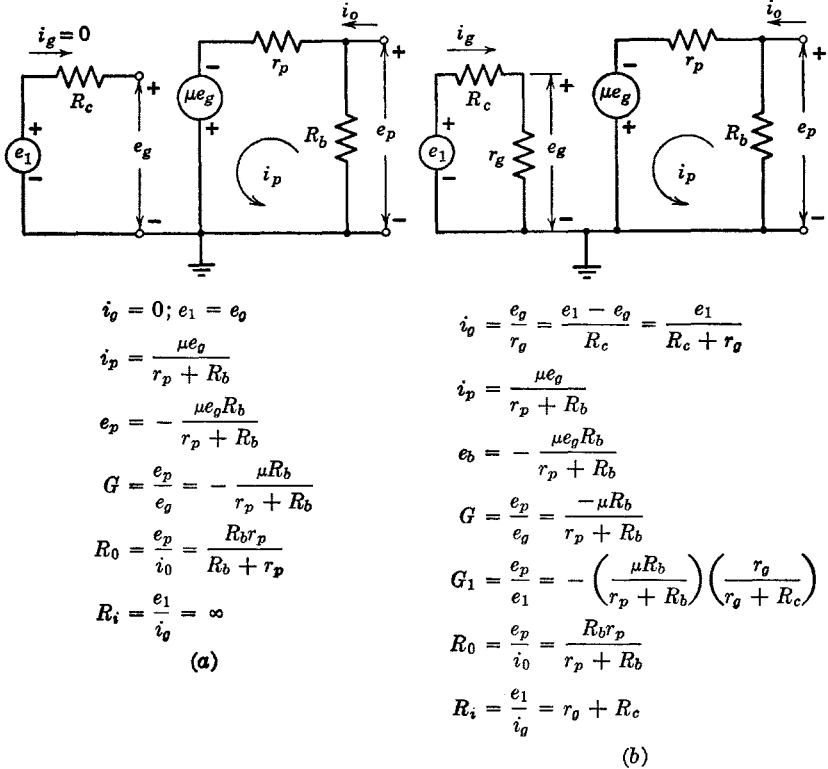
Fig. 6.19. Incremental analysis of grounded-cathode circuit. (a) Negative grid, $e_c < 0$ (state II). (b) Positive grid, $e_c > 0$ (state III).

Fig. 6.16 and replacement of di_b by i_p, de_b by e_p, etc. The circuit models in Fig. 6.19 follow directly from these incremental equations. When the equations relating total quantities are totally differentiated, the direct components of voltage and current always disappear. Thus the *linear incremental* model can be obtained directly from the linear model by short-circuiting all d-c voltage sources and open-circuiting all d-c current sources.

The voltage gain

$$G = de_b/de_c = e_p/e_g \tag{6.17}$$

is the slope of the transfer curve e_b vs. e_c, and

$$G_1 = de_b/de_1 = e_p/e_1 \tag{6.18}$$

is the slope of e_b vs. e_1. Referring to the graphical plots of these transfer curves [Fig. 6.15(d) and (f)], we see that in state I

$$G = G_1 = 0 \tag{6.19}$$

and in state II

$$G = G_1 = \frac{-\mu R_b}{R_b + r_p} \tag{6.20}$$

whereas in state III

$$G = \frac{-\mu R_b}{R_b + r_p} \tag{6.21}$$

and

$$G_1 = G \frac{r_g}{r_g + R_c} \tag{6.22}$$

For the nonlinear transfer curves e_b vs. e_c, or e_b vs. e_1, we see that the slope varies smoothly instead of abruptly as the triode operation passes from one region to another. Such discrepancy is the price we must pay in return for the convenience of linear approximation. When necessary, a better approximation can be made at the transition between two regions of operation.

Expressions for incremental input resistance R_i and output resistance R_0 are also given in Fig. 6.19. Assuming R_c to be part of the amplifier, we see that the incremental input resistance faced by the source e_i is

$$R_i = r_g + R_c \tag{6.23}$$

If R_c is considered as part of the source, the input resistance to the amplifier will be simply

$$R_i = e_g/i_g = r_g \tag{6.24}$$

The output resistance R_0 may be interpreted as the increment of output voltage produced by a unit change of external current i_0. This is clearly equal to the output conductance $G_0 = 1/R_0$. Alternatively, we may calculate R_0 by applying Thevenin's theorem to the output terminals, whence R_0 is the ratio of open-circuit voltage $\mu e_g R_b/(R_b + r_p)$ to short-circuit current $\mu e_g/r_p$ at the output terminals. In either case the result is

$$R_0 = R_b r_p/(R_b + r_p) \tag{6.25}$$

Voltage amplification for the grounded-cathode circuit is seen to be a fraction of the triode amplification factor μ. The range of values of μ in different types of readily available triodes is quite large, say from 3 to 100. The incremental plate resistance is apt to be nearly proportional to μ for triodes of comparable size and power rating. In other words, $g_m = \mu/r_p$ is largely determined by the power rating of the tube. For triodes operating at power levels of a few watts, g_m ranges from 1000 to 10,000 micromhos. The value of plate-load resistance R_b, relative to r_p, determines the fraction of μ available as voltage amplification. For triode amplifier circuits R_b is commonly made comparable to r_p.

6.13   The Cathode-Follower Circuit

Slight rearrangement of the circuit elements used in the plate-loaded amplifier yields the cathode-follower circuit shown in Fig. 6.20(a). The load resistance R_k and the polarizing battery E_{bb} are connected in series with the plate and cathode, just as before. The terminal common to input and output voltage is at a different position in the plate-circuit loop, but this does not affect the determination of plate current i_b for a given grid-to-cathode voltage e_c.

Assume that the triode plate current is specified by the curves of Fig. 6.20(b). The linear load constraint is given (as before) by

$$i_b = (E_{bb} - e_b)/R_k \qquad (6.26)$$

When plotted on the tube curves, this constraint is exactly the same (for $R_b = R_k$) as for the plate-loaded case. Accordingly, for negative values of e_c the transfer curve i_b vs. e_c shown in Fig. 6.20(c) is the same as that of Fig. 6.14(c). For positive e_c the grid current i_c must be taken into account, since $i_b + i_c$ is the current in the load resistor R_k. The effect on the transfer curve is shown by the dotted line in Fig. 6.20(c). The curve of e_2 vs. e_c in (d) corresponds to the curve of e_b vs. e_c in Fig. 6.14(d), except that it rises from zero ($e_2 = i_b R_k$) instead of falling from E_{bb} as e_c increases from $-E_{co}$. In the positive-grid region,

$$e_2 = (i_b + i_c)R_k \qquad (6.27)$$

hence, the curve can rise above the value E_{bb}. If e_2 is greater than E_{bb}, then e_b is less than 0 and $i_b = 0$, so that

$$e_2 = i_c R_k = e_1 R_k/(R_k + r_g + R_c) \qquad (6.28)$$

This circuit state, with the grid-cathode diode conducting and the plate-
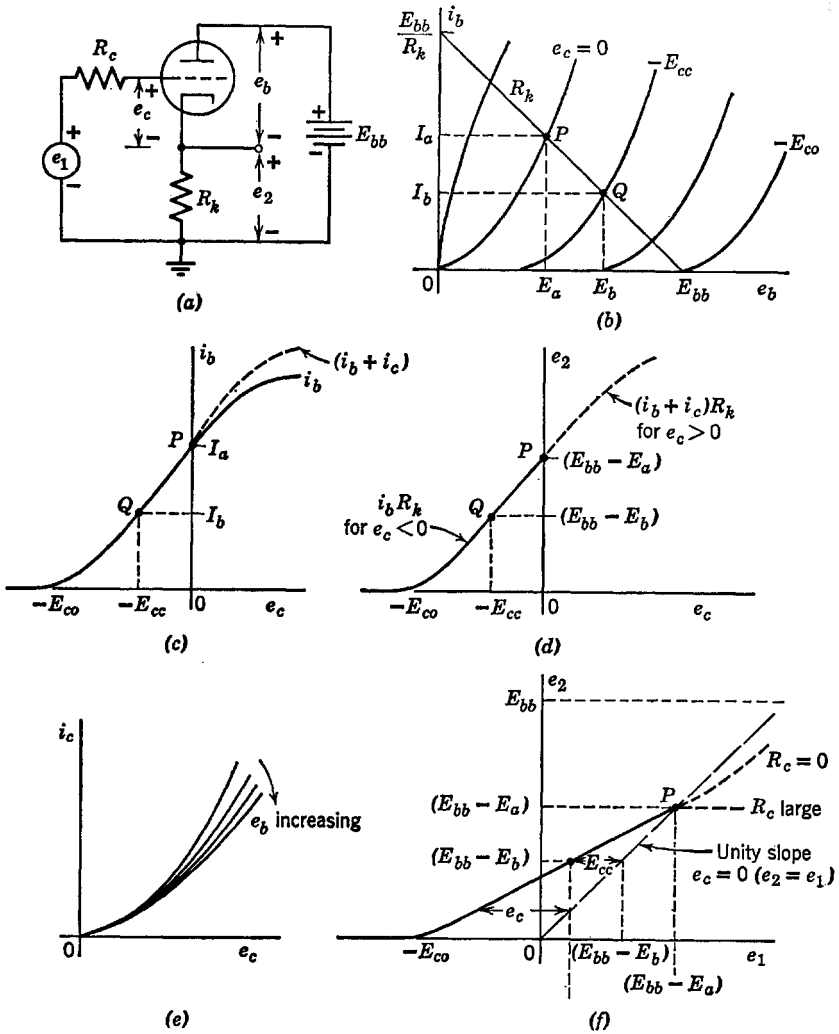
Fig. 6.20. Cathode-follower circuit-transfer curves.

cathode diode cut off, was not realizable in the plate-loaded circuit merely by varying e_1. Although possible in the cathode follower, it is neither a useful nor a likely mode of operation. In the positive-grid region, grid-current curves such as those shown in Fig. 6.20(e), or some equivalent grid-circuit data, are required for quantitative determination of the $(i_b + i_c)$ vs. e_c or e_2 vs. e_c curves.

The over-all voltage transfer curve (e_2 vs. e_1) for the cathode follower

is plotted in Fig. 6.20(f). The voltage $e_1 = e_c + e_2$ is obtained from the curve of e_2 vs. e_c plotted in (d) by adding output voltage e_2 to e_c to find the new abscissa e_1. This effectively shears the original curve to the right. The unity-slope dashed line through the origin indicates the amount by which the curve in (d) is sheared to the right for various values of e_2; for example, $e_2 = 0$ means $i_b = 0$; hence $e_1 = e_c = -E_{co}$, a common point on the two curves. The curve of e_2 vs. e_1 can also be derived directly from the i_b vs. e_b curve (b) as follows:

1. Assume a value of e_c.
2. From Fig. 6.20(b) find the corresponding value of i_b.
3. Find e_2 using $e_2 = i_b R_k$.
4. Find e_1 using $e_1 = e_c + e_2$.

At point P, $e_c = 0$ and $e_1 = (E_{bb} - E_a)$. For e_1 greater than this value, the curve is horizontal if R_c is large, since e_c remains zero regardless of e_1. If $R_c = 0$, the slope of the curve remains about the same, or it may actually increase slightly because of grid current through R_k. If r_g is negligibly small compared with R_k, then for $R_c = 0$

$$e_2 = e_1 R_k/(R_k + r_g) \approx e_1 \qquad (6.29)$$

In this case the curve would approach the unity-slope line very closely.

The transfer curve e_2 vs. e_1 shown in Fig. 6.20(f) specifies the circuit behavior over a range of values of the input voltage. If only a single operating point is sought (say for $e_1 = E_1$), we can determine e_2 by the procedure outlined in Fig. 6.21. Assuming e_c is less than zero, we must satisfy the three conditions

$$E_1 = e_c + e_2 \qquad (6.30)$$

$$i_b = \frac{E_{bb} - e_b}{R_k} \qquad (6.31)$$

$$i_b = f(e_b, e_c) \qquad (6.32)$$

In Fig. 6.21, the load line (Eq. 6.31) is designated by R_k. Equation 6.30 can be plotted on the tube curves by substituting $i_b R_k$ for e_2, so that

$$E_1 = e_c + i_b R_k \qquad (6.33)$$

or

$$i_b = (E_1 - e_c)/R_k \qquad (6.34)$$

Equation 6.34 thus assumes the same form as Eq. 6.31, but with e_c appearing in place of e_b. We might therefore designate this as a transfer line, since one input and one output variable are involved. The relation, shown by dotted lines on the tube curves of Fig. 6.21, is commonly called a "bias line." Intersection of this line with the plate-circuit load

line determines the required operating point. If E_1 is greater than $(E_{bb} - E_a)$, so that e_c is greater than zero at the required operating point, any of several similar procedures can be followed to effect a solution. Note that in this case

$$E_1 = e_c + e_2 = e_c + (i_b + i_c)R_k \qquad (6.35)$$

and

$$(i_b + i_c) = (E_{bb} - e_b)/R_k \qquad (6.36)$$
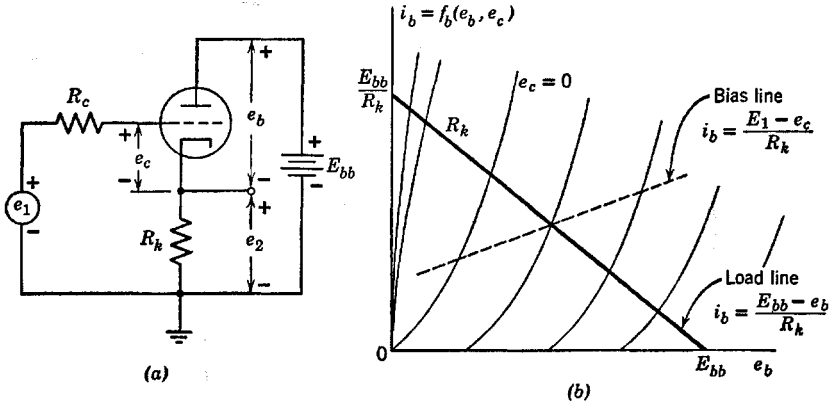
Fig. 6.21. Cathode-follower operating-point determination.

Thus, if we have available the graphical plots of $i_b = f_b(e_b, e_c)$ and $i_c = f_c(e_b, e_c)$, we can replot the curves as $i_k = (i_b + i_c) = f_k(e_b, e_c)$. The load line and bias line can then be plotted on the curves of i_k vs. e_b as before to locate the operating point. If we do not replot the triode data, a cut-and-try solution is required. Since this proceeds best if values of e_c are assumed, it is almost as simple to form the complete transfer curve as to find a single operating point.

A piecewise-linear triode model in a cathode-follower circuit is shown in Fig. 6.22(a). Plate-current curves corresponding to the model are shown in (b), and the simple grid-current curve appears in (c). The voltage transfer curve e_2 vs. e_1 is shown in (d), and pertinent algebraic relations for each diode state are tabulated in (e). A fifth state, with D_g ON, D_p OFF, and D_s ON, is not listed in the table. In state V the circuit reduces to a simple voltage divider.

6.14 Incremental Analysis of the Cathode Follower

Having determined various modes of circuit operation from the relations between total voltages and total currents, analytical simplifi-
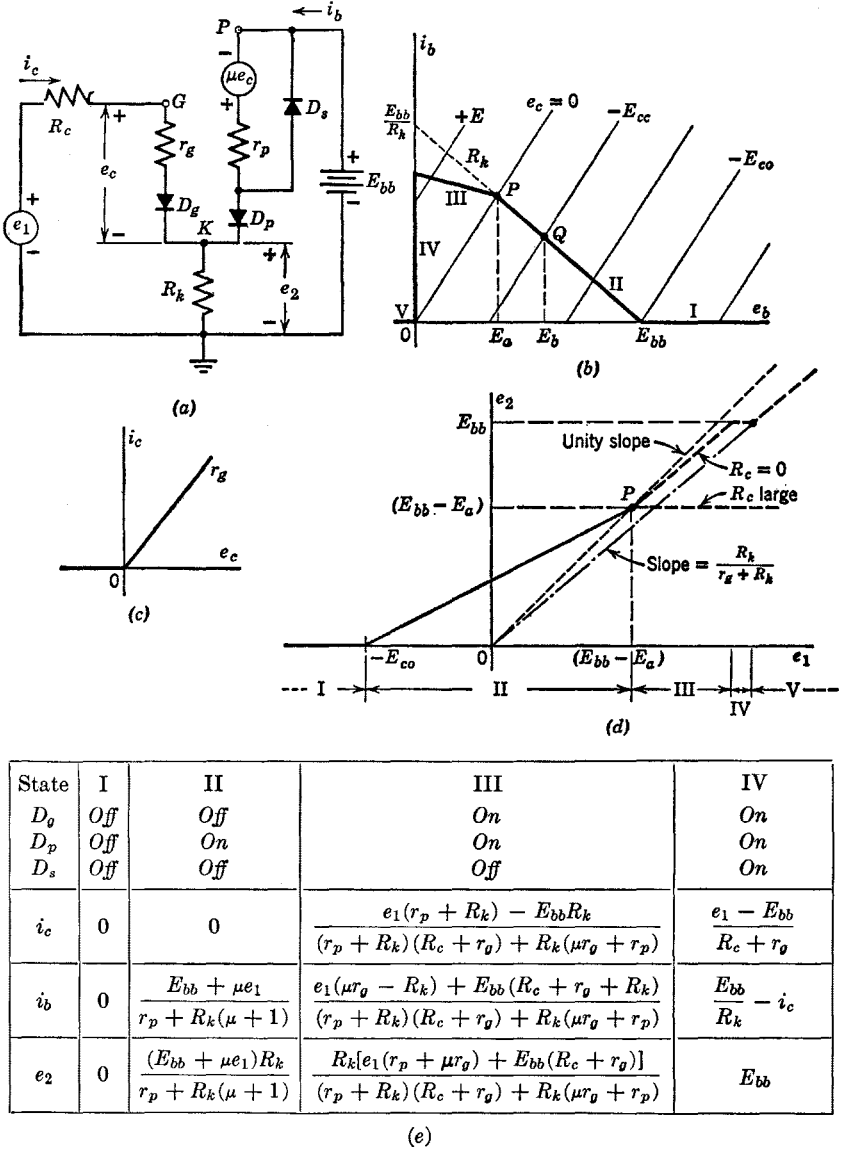
Fig. 6.22. Piecewise-linear analysis of the cathode follower.

State	I	II	III	IV
D_g	Off	Off	On	On
D_p	Off	On	On	On
D_s	Off	Off	Off	On
i_c	0	0	$\dfrac{e_1(r_p + R_k) - E_{bb}R_k}{(r_p + R_k)(R_c + r_g) + R_k(\mu r_g + r_p)}$	$\dfrac{e_1 - E_{bb}}{R_c + r_g}$
i_b	0	$\dfrac{E_{bb} + \mu e_1}{r_p + R_k(\mu + 1)}$	$\dfrac{e_1(\mu r_g - R_k) + E_{bb}(R_c + r_g + R_k)}{(r_p + R_k)(R_c + r_g) + R_k(\mu r_g + r_p)}$	$\dfrac{E_{bb}}{R_k} - i_c$
e_2	0	$\dfrac{(E_{bb} + \mu e_1)R_k}{r_p + R_k(\mu + 1)}$	$\dfrac{R_k[e_1(r_p + \mu r_g) + E_{bb}(R_c + r_g)]}{(r_p + R_k)(R_c + r_g) + R_k(\mu r_g + r_p)}$	E_{bb}

(e)

cation again results from the use of incremental quantities. We cannot emphasize too often the fact that nonlinear graphical analysis or its substitute, the piecewise-linear approximation, implicitly precedes incremental analysis. The total voltages and currents determine the

operating region (the diode state in the piecewise-linear model), and on this basis an appropriate incremental model can be chosen.

Referring to the tabulation in Fig. 6.22(e), it is apparent that the incremental model for state I is completely trivial, since the triode is effectively removed, whereas the circuit for state IV merely amounts to a resistive voltage divider consisting of R_c, r_g, and R_k in series.

The differentials of the expressions for state II lead to the equations relating the incremental components of the variables for negative-grid operation. These equations are given in Fig. 6.23, together with incremental circuit models that represent the equations. Symbols e_1' and e_2' are used here to designate incremental quantities. The incremental circuit based on these equations does not represent the output behavior of the circuit, if an external load is connected, because R_k appears as part of the source resistance.

Dividing numerator and denominator of the equations in Fig. 6.23(a) by the quantity $(\mu + 1)$ leads to the equations and incremental circuit shown in (b). Since R_k now appears only at the output, this model is reasonable, provided we can justify the modification in apparent plate resistance and source value. Reference to the actual circuit [Fig. 6.21(a)] helps us to visualize the physical basis for a low output resistance. Suppose that the circuit is operating in the negative-grid region and that the output voltage is raised an increment Δe_2 by connecting a battery $E_k + \Delta e_k$ to the output terminals. If e_1 is held constant, then e_c must become more negative by an amount Δe_k. The net change in voltage across r_p must then be

$$-r_p \Delta i_b = \Delta e_k + \mu \, \Delta e_k \tag{6.37}$$

causing a corresponding decrease in plate current

$$-\Delta i_b = \frac{(\mu + 1)\Delta e_k}{r_p} \tag{6.38}$$

The net change in external current (in the $E_k + \Delta e_k$ battery) is

$$-\Delta i_b + \frac{\Delta e_k}{R_k} = \frac{(\mu + 1)\Delta e_k}{r_p} + \frac{\Delta e_k}{R_k} \tag{6.39}$$

The incremental output resistance is therefore

$$\frac{\Delta e_k}{-\Delta i_b + \dfrac{\Delta e_k}{R_k}} = \frac{1}{\dfrac{\mu + 1}{r_p} + \dfrac{1}{R_k}} = \frac{R_k \left(\dfrac{r_p}{\mu + 1} \right)}{R_k + \dfrac{r_p}{\mu + 1}} \tag{6.40}$$

An alternate method of calculating output resistance is indicated in Fig. 6.23(c).

$$di_b = \frac{\mu de_1}{r_p + R_k(\mu+1)} \text{ or } i_p = \frac{\mu e_1'}{r_p + R_k(\mu+1)}$$

$$de_2 = \frac{\mu R_k de_1}{r_p + R_k(\mu+1)} \text{ or } G = \frac{e_2'}{e_1'} = \frac{\mu R_k}{r_p + R_k(\mu+1)}$$
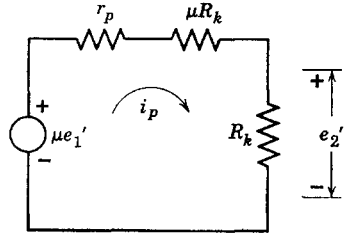
(a)

$$i_p = \frac{[\mu/(\mu+1)]e_1'}{R_k + r_p/(\mu+1)}$$

$$G = \frac{e_2'}{e_1'} = \frac{[\mu/(\mu+1)]R_k}{R_k + r_p/(\mu+1)}$$
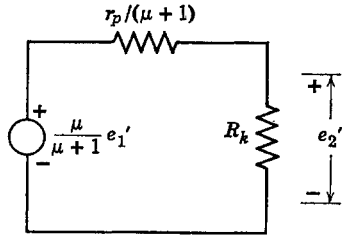
$$\frac{r_p}{\mu+1} = \frac{1}{g_m + g_p} \approx \frac{1}{g_m} \text{ for } \mu \gg 1$$

(b)

(c)

$$\frac{\mu e_g}{r_p} = \frac{-\mu e_2'}{r_p} = -g_m e_2'$$
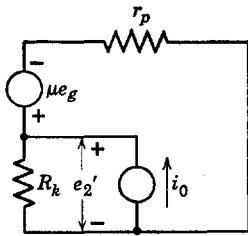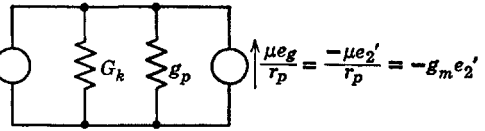
$$(i_0 - g_m e_2') = e_2'(G_k + g_p)$$

$$G_0 = \frac{i_0}{e_2'} = G_k + g_p + g_m = G_k + \frac{\mu+1}{r_p}$$

$$R_0 = 1/G_0 = (R_k \| r_p \| 1/g_m) = \frac{R_k r_p/(\mu+1)}{R_k + r_p/(\mu+1)}$$
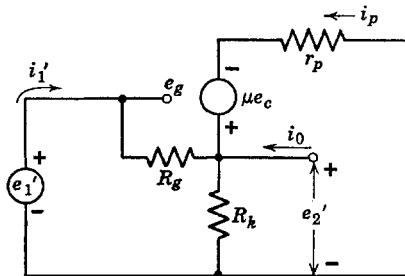
for $R_g \gg R_k$ (or $i_1' \ll i_p$)

$$e_g = i_1' R_g = e_1' - e_2' = e_1'(1 - G)$$

$$\therefore R_i = \frac{e_1'}{i_1'} = \frac{R_g}{1 - G}$$

(d)

Fig. 6.23. Incremental analysis of cathode follower for $-E_{bb}/\mu < e_c < 0$ (state II).

For e_c less than 0, the incremental input resistance of the cathode-follower circuit is theoretically infinite, since the grid-cathode diode does not conduct. Suppose, however, that a resistance R_g is connected between grid and cathode, as indicated in the incremental circuit shown in Fig. 6.23(d). Let us assume that R_g is much greater than R_k and r_p, or that i_1' is much less than i_p. Then the incremental current i_1' will not appreciably affect the gain or other quantities previously calculated for the cathode follower. Since an increment of input voltage e_1' results in a nearly equal increment in output voltage $e_2' = Ge_1'$, we have

$$e_g = e_1' - e_2' = e_1'(1 - G) \qquad (6.41)$$

But $e_g = i_1'R_g$; hence the input resistance is

$$R_i = e_1'/i_1' = R_g/(1 - G) \qquad (6.42)$$

Since G is nearly unity, the input resistance appearing between grid and ground is R_g multiplied by the large factor $1/(1 - G)$. If R_g is connected between the grid and any arbitrary tap point on R_k, the effective input resistance is $R_g/(1 - G_t)$, where G_t is the gain from input e_1' to output e_{2t}', with the latter measured between the tap point and ground.

An incremental circuit for positive-grid operation (state III) is shown in Fig. 6.24(a), together with the voltage equations for the two loops of the circuit. The incremental expressions for the currents and the gain are shown in (b).

The incremental input resistance R_i given in Fig. 6.24(c) can be obtained directly from the first equation in (b). The incremental output resistance is determined by the circuit and equations of Fig. 6.24(d). In making this calculation, the incremental source e_1' is set equal to zero, and the circuit is effectively being driven by a current source i_0.

6.15   Properties of the Cathode Follower

The important properties of the cathode follower circuit are its linearity, its low output resistance, and its high input resistance. Although the incremental voltage gain is always less than unity (see Fig. 6.24) the circuit yields power amplification since there is a current gain. The cathode follower is often used as an isolating circuit since it permits a source (even of relatively high internal resistance) to drive a low resistance load without the load reacting on the source.

To make maximum use of the linear range $(0 > e_c > -E_{bb}/\mu)$, the cathode follower can be polarized near $e_c = 0$ for signals that are

essentially negative-going, near $-E_{bb}/\mu$ for signals that are essentially positive-going, and near the center of the range for symmetrical signals like sine waves. These possibilities are illustrated in Fig. 6.25. The

(a)
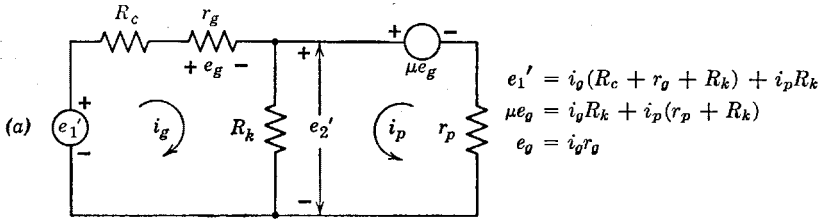
$$e_1' = i_g(R_c + r_g + R_k) + i_p R_k$$
$$\mu e_g = i_g R_k + i_p(r_p + R_k)$$
$$e_g = i_g r_g$$

(b)

$$i_g = \frac{e_1'(r_p + R_k)}{(r_p + R_k)(R_c + r_g) + R_k(\mu r_g + r_p)}$$

$$i_p = \frac{e_1'(\mu r_g - R_k)}{(r_p + R_k)(R_c + r_g) + R_k(\mu r_g + r_p)}$$

(c)

$$G = \frac{e_2'}{e_1'} = \frac{R_k(i_p + i_g)}{e_1'} = \frac{R_k[\mu r_g + r_p]}{(r_p + R_k)(R_c + r_g) + R_k(\mu r_g + r_p)}$$

$$R_i = \frac{e_1'}{i_g} = R_c + r_g + \frac{R_k(\mu r_g + r_p)}{r_p + R_k}$$

(d)
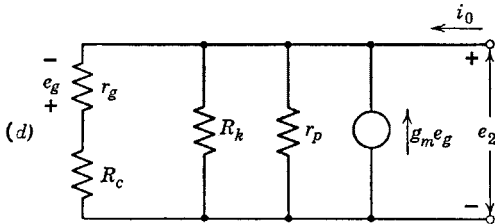
$$-e_g = e_2' r_g/(r_g + R_c); \text{ Let } G_1 = 1/(R_c + r_g)$$

$$i_0 - g_m e_2' r_g/(r_g + R_c) = e_2'(G_k + g_p + G_1)$$

$$G_0 = i_0/e_2' = \frac{1}{R_0} = \frac{1}{R_k} + \frac{1}{r_p} + \frac{1 + g_m r_g}{R_c + r_g}$$

Fig. 6.24. Incremental analysis of cathode follower for $e_c > 0$ (state III).

capacitors are assumed to be so large that they are ideal coupling elements. The determination of R_2 for the symmetrical case illustrated in (c) is no different than the corresponding problem for the grounded-cathode circuit. [See graphical constructions of Fig. 6.17(d) and (f).]
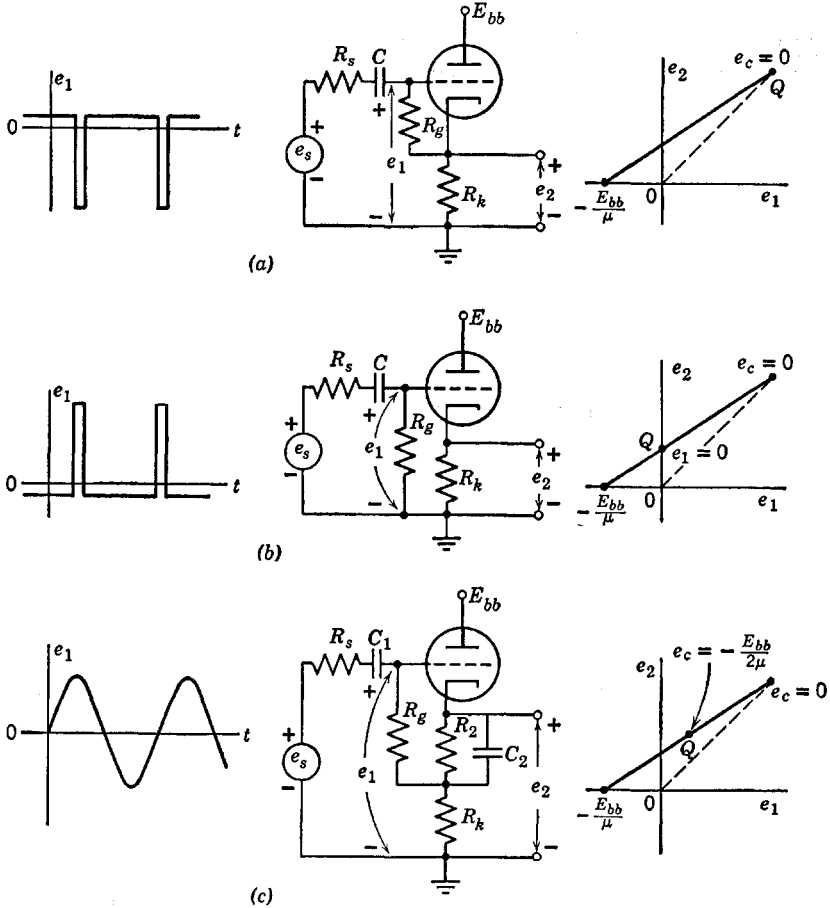
Fig. 6.25. Polarizing the cathode follower.

6.16 General Grid-Driven Triode Circuit

The cathode-follower circuit (grounded plate) and the plate-loaded circuit (grounded cathode) are special cases of a more general circuit with both plate and cathode loads. The circuit is shown schematically in Fig. 6.26(a) with outputs indicated between cathode and ground as well as between plate and ground. A simplified schematic diagram for the circuit is shown in Fig. 6.26(b). The voltages E_{bb}, e_1, e_2, and e_3 designated at the various terminals are assumed to be measured from ground. Voltages like e_b and e_c that are not measured from ground must

be designated as before. This form of circuit diagram can, of course, be applied to the cathode follower, the plate-loaded amplifier, transistor circuits, etc. It will be used extensively in subsequent articles and chapters.
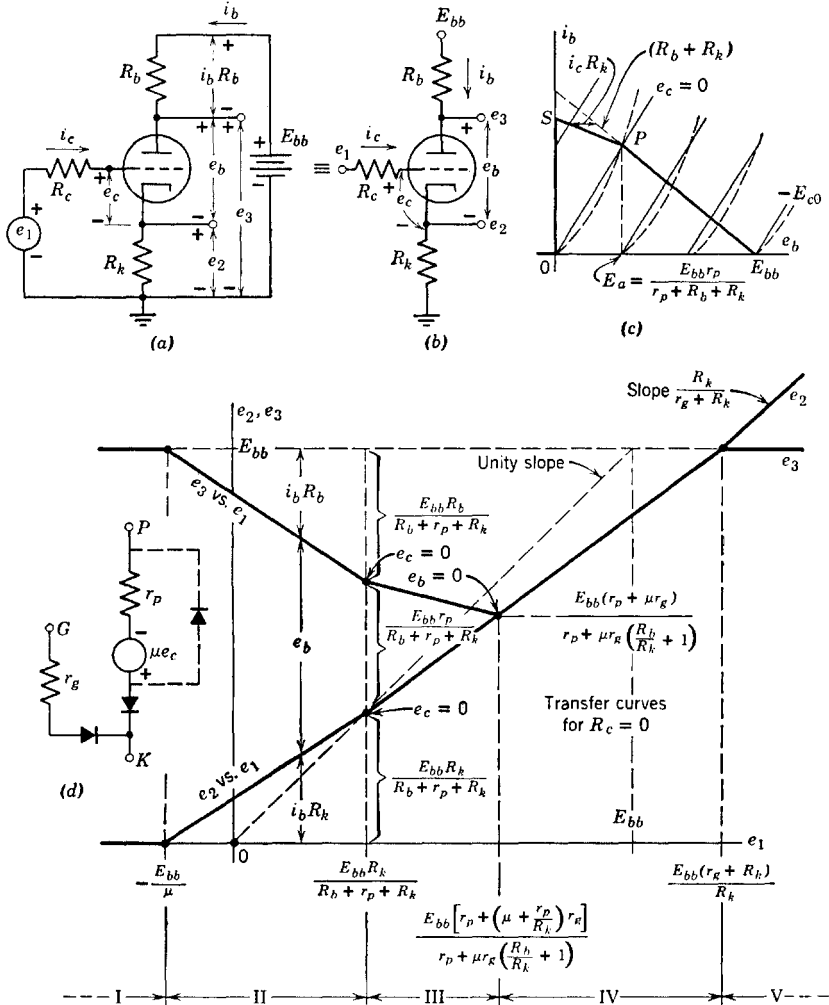
Fig. 6.26. Piecewise-linear analysis of grid-driven circuit with plate and cathode load.

The piecewise-linear analysis, Fig. 6.26(d), and the incremental analysis, Fig. 6.27, follow the same pattern as the corresponding analyses in the preceding articles. If we let $R_b = R_k$, then from Fig. 6.27 we see

(a)
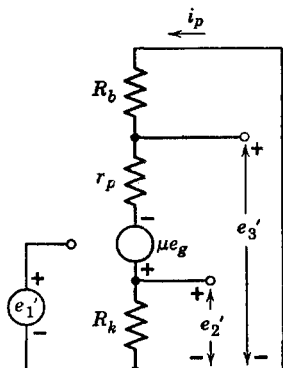
$$\begin{cases} e_g = e_1' - i_p R_k \\ \mu e_g = i_p (R_b + r_p + R_k) \end{cases}$$

$$i_p = \frac{\mu e_1'}{R_b + r_p + R_k(\mu + 1)}$$

(b)

$$e_2' = i_p R_k$$

$$G_2 = \frac{e_2'}{e_1'} = \frac{\mu R_k}{R_b + r_p + R_k(\mu + 1)}$$

$$= \frac{[\mu/(\mu + 1)]R_k}{(R_b + r_p)/(\mu + 1) + R_k}$$

(c)

(d)

$$e_3' = -i_p R_b$$
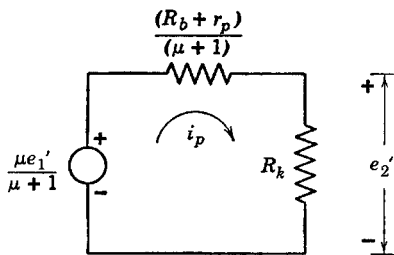
$$G_3 = \frac{e_3'}{e_1'} = \frac{-\mu R_b}{R_b + r_p + R_k(\mu + 1)}$$

(e)
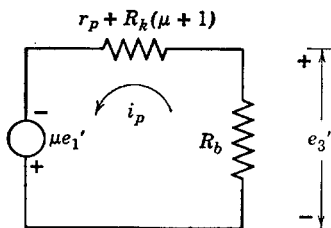
(f)

Fig. 6.27. Incremental analysis of circuit with plate and cathode load for $e_c < 0$ (state II).

that G_2 equals G_3 for e_c less than zero. With this choice of element values the circuit is used extensively in audio amplifiers to drive balanced circuits.
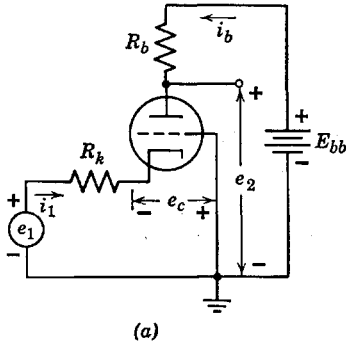
Note: $e_k = -e_c$

$i_1 = -i_k = -i_b (\text{for } e_k > 0)$

(a)

(b)

(c)

(d)

(e)
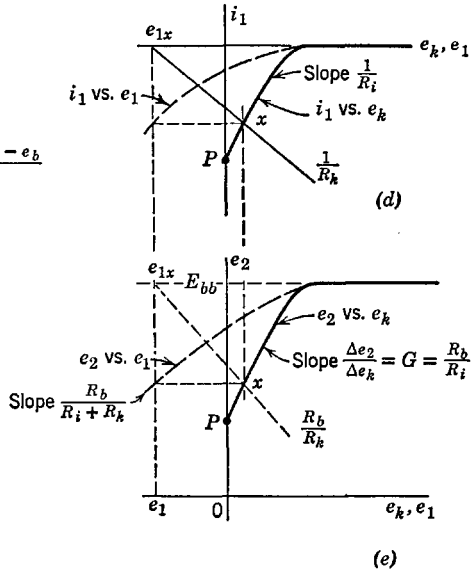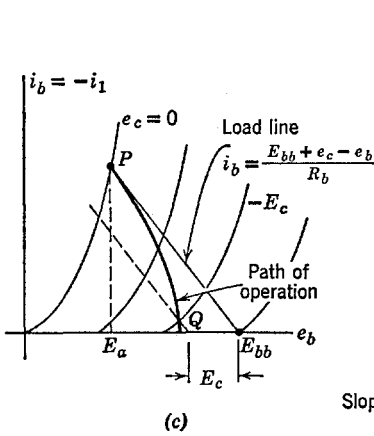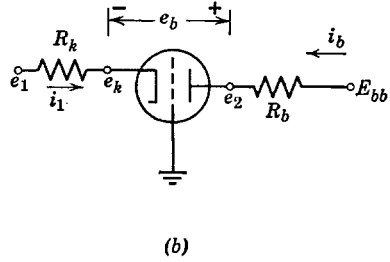
Fig. 6.28. Grounded-grid amplifier.

6.17 Grounded-Grid Amplifier

The grounded-grid amplifier shown in Fig. 6.28(a) and (b) can also be described as a cathode-driven triode with plate load. For negative-grid operation, we see from Fig. 6.28(b) that

$$e_b = E_{bb} - e_k - i_b R_b \tag{6.43}$$

Solving this equation for i_b we obtain

$$i_b = (E_{bb} - e_k - e_b)/R_b \tag{6.44}$$

or

$$i_b = (E_{bb} + e_c - e_b)/R_b \qquad (6.45)$$

Equation 6.45 can be plotted on the i_b vs. e_b plane, as indicated in Fig. 6.28(c). Note that the load-line moves as e_c is varied. For $e_c = 0$, the operating point P is determined by the solid load-line. For $e_c = -E_c$ ($e_k = +E_c$) the operating point is at Q, as indicated by the dotted load-line. The path of operation, shown by the heavy line, is the locus of intersections of the tube curves with load-lines that shift to the left as e_c is made more negative.

Since $i_b = -i_1$, this locus of the operating point can be replotted as i_1 vs. e_k. The plot of i_1 vs. e_k, shown in Fig. 6.28(d), represents the driving-point relation for the circuit. The slope of the curve determines the incremental resistance at the cathode terminal:

$$R_i = \Delta e_k/\Delta i_1 \qquad (6.46)$$

The resistance between the input terminal and ground is merely the sum of R_i and R_k.

The transfer curve (e_2 vs. e_k) shown in Fig. 6.28(e) is related to the input curve by the linear equation

$$e_2 = E_{bb} + i_1 R_b \qquad (6.47)$$

The gain $G_{2k} = de_2/de_k$ is slightly higher than the corresponding gain $G_{2c} = de_2/de_c$ for a grounded-cathode amplifier, assuming the same tube, load resistance, and operating point. This fact may be seen from Fig. 6.28(c), where a given change of e_c produces a slightly larger change in i_b along the PQ locus than along the load-line.

The incremental form of Eq. 6.47 is

$$\Delta e_2 = \Delta i_1 R_b \qquad (6.48)$$

Solving Eqs. 6.46 and 6.48 for $\Delta e_2/\Delta e_k$ yields

$$\frac{\Delta e_2}{\Delta e_k} = G_{2k} = \frac{R_b}{R_i} \qquad (6.49)$$

Equation 6.49 indicates that if $G_{2k} = 10$ and $R_b = 10k$, then $R_i = 1k$. We see, therefore, that the grounded-grid connection exhibits a relatively low incremental input resistance.

If the cathode voltage e_k is produced by a source e_1 driving the circuit through a resistance R_k, the transfer curve of interest is e_2 vs. e_1. From the circuit diagram of Fig. 6.28(a) or (b), we see that

$$e_1 = e_k + i_1 R_k \qquad (6.50)$$

Equation 6.50 is plotted on the i_1 vs. e_k plane in Fig. 6.28(d). The

intersection x with the input curve relates e_1 to e_k and therefore also relates e_2 to e_1, since the transfer curve in Fig. 6.25(e) relates e_2 to e_k. Since the input and transfer curves have the same general shape, this

(a)

$$\begin{cases} e_1 = e_k + i_1 R_k \\ e_c = -e_k \\ i_1 = -i_b, \text{ since } i_c = 0 \text{ for } e_c < 0 \\ E_{bb} = i_b(R_b + r_p) - \mu e_c + e_k \\ \quad = i_b(R_b + r_p) + (\mu + 1)e_k \end{cases}$$

(b)

$$R_i = \frac{\Delta e_k}{\Delta i_1} = \frac{r_p + R_b}{\mu + 1}$$

$$G_{2k} = \frac{\Delta e_2}{\Delta e_k} = \frac{(\mu + 1)R_b}{r_p + R_b} = \frac{R_b}{R_i}$$

$$G_{21} = \frac{\Delta e_2}{\Delta e_1} = \frac{(\mu + 1)R_b}{r_p + R_b + R_k(\mu + 1)}$$

$$= \frac{R_i}{R_i + R_k} \cdot \frac{(\mu + 1)R_b}{r_p + R_b} = \frac{R_i}{R_i + R_k} G_{2k}$$

$$= \frac{R_b}{R_i + R_k}$$

(e)

(c)

(d)
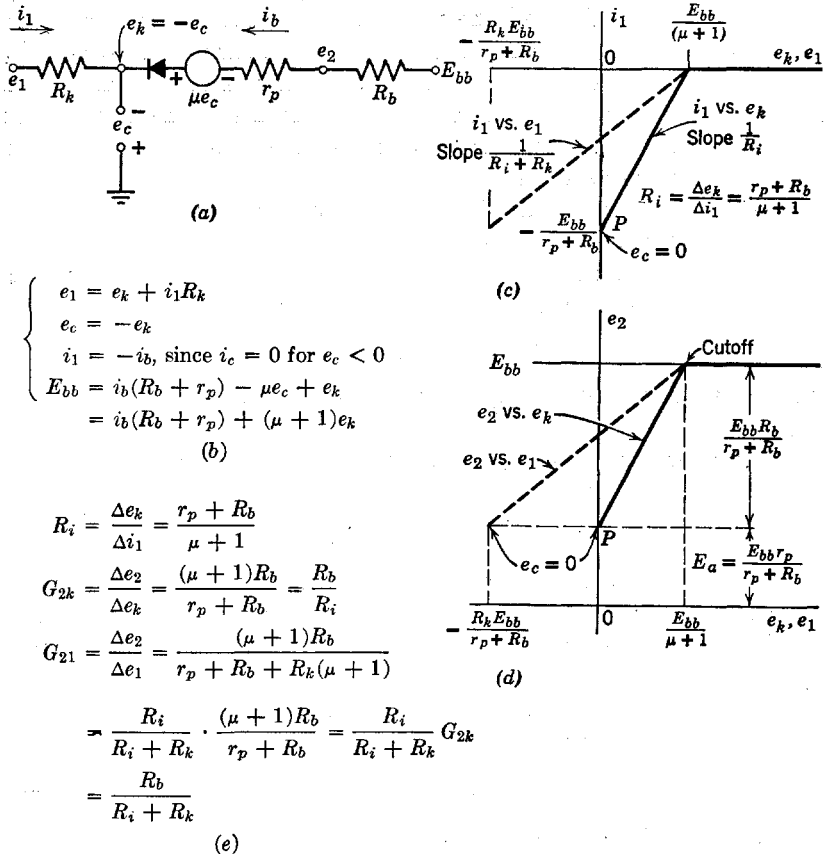
Fig. 6.29. Piecewise-linear analysis of grounded-grid circuit for $e_c < 0$.

source line can be transferred directly to the e_2 vs. e_k plane, where its slope is determined by R_b/R_k, and e_1 is its intercept on the line $e_2 = E_{bb}$.

Reference to the curves shown in Fig. 6.28(d) and (e) shows, for a given operating point, that the incremental gains are related as follows:

$$G_{k1} = de_k/de_1 = R_i/(R_i + R_k) \tag{6.51}$$

and

$$G_{2k} = de_2/de_k = R_b/R_i \tag{6.52}$$

Therefore

$$G_{21} = de_2/de_1 = G_{2k}R_i/(R_i + R_k) = R_b/(R_i + R_k) \tag{6.53}$$

A piecewise-linear model of a grounded-grid amplifier, applicable in the negative-grid region ($e_c < 0$), is shown in Fig. 6.29(a). Inspection of this circuit yields the equations given in (b). Application of the break-point method of analysis leads directly to the driving-point curve plotted in (c) and the voltage transfer curve given in (d).

The incremental quantities given in Fig. 6.29(e) are for state II which is determined by the conditions

$$E_{bb}/(\mu + 1) > e_k > 0 \qquad (6.54)$$

These limits on state II are readily determined from the dimensions of the curves in (c) and (d). Note particularly the fact that the circuit has a very low input resistance R_i, which enters directly into the expression for incremental voltage gain.

Although the grounded-grid amplifier requires driving power at the input terminals, the input signal controls a much larger incremental power variation in the load. Because the input and load currents are identical (for $e_c < 0$), the ratio of incremental output power to incremental input power is numerically equal to the voltage gain $G_p = G_{21}$.

The more general piecewise-linear model for the grounded-grid circuit shown in Fig. 6.30(a) applies to both the negative-grid and the positive-grid regions. The equations given in (b) are written for e_c greater than 0. The transfer curve e_2 vs. e_k, plotted in Fig. 6.30(c), shows the negative-grid behavior in dashed lines.

The grounded-grid circuit is less widely used than the grounded-cathode or grounded-anode (cathode follower) circuits. It is used in the input stage of the intermediate-frequency amplifier of many receivers. It is also used in high-frequency amplifier circuits where the grounded grid provides a natural isolation between input and output circuits. The electrostatic shielding action of a grounded grid is discussed further in Chapter 7.

In conjunction with a cathode-follower, the grounded-grid circuit forms a commonly used two-triode circuit called a cathode-coupled amplifier. This circuit is widely used as an inverting amplifier and as the basis for a relaxation oscillator called the cathode-coupled multivibrator.

6.18   Direct-coupled Grounded-Cathode Circuits

The basic circuits discussed above can be coupled together in a variety of ways when more than one triode is to be used. A few of the possible two-triode configurations will be described here and in succeeding articles.

(a)
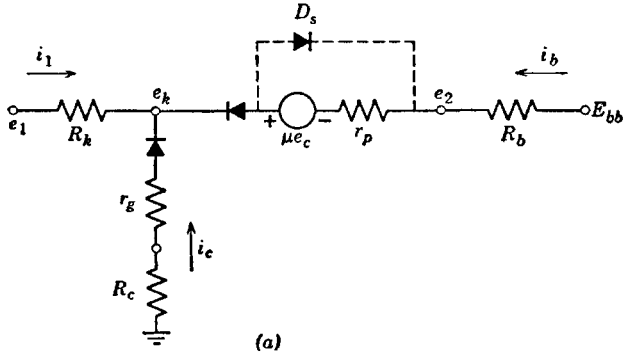
$$i_1 = -(i_b + i_c);$$
$$e_c = i_c r_g$$
$$e_1 = -i_b R_k - i_c(R_c + r_g + R_k)$$
$$E_{bb} = i_b(R_b + r_p) - i_c(R_c + r_g) - \mu e_c$$
$$= i_b(R_b + r_p) - i_c[R_c + r_g(\mu + 1)]$$

(b)

(c)
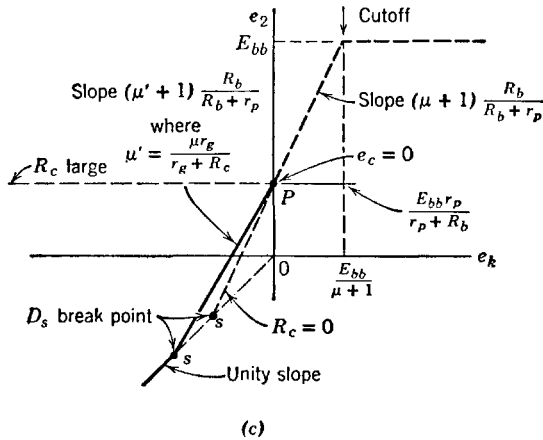
Fig. 6.30. Piecewise-linear analysis of grounded-grid circuit for $e_c > 0$.

In this chapter, we shall restrict ourselves to interstage coupling networks that involve only resistive circuit elements: namely, resistances, batteries, and diodes. Such circuits are said to be direct-coupled, since they respond to frequencies down to and including zero; in other words, they can be used to amplify d-c as well as a-c signals. The assumption of "ideal coupling" by a large capacitor, as mentioned in Article 6.11, does

not apply for d-c voltages. The use of energy-storage elements for inter-stage coupling is discussed in Chapter 8.

One of the main reasons for using two or more triodes (or other electronic devices) in a circuit is to obtain greater amplification than a single triode can supply. The particular combination of basic single-triode circuits chosen for this purpose usually depends upon factors other than amplification. Among these may be linearity, dynamic range, input resistance, output resistance, interchangeability of triodes, etc.

A simple two-triode circuit is shown in Fig. 6.31(a). It consists of two plate-loaded, grounded-cathode circuits coupled together by means of a battery E. The transfer curve e_{b1} vs. e_1 for the first triode circuit, shown in (b), assumes the battery E to be disconnected. The transfer curve e_2 vs. e_{c2} for the second circuit alone is similar in form and is shown in (c).

Since the maximum value of e_{b1} is E_{bb}, the maximum value of e_{c2} is $(E_{bb} - E)$ when the two circuits are coupled together. If we choose E equal to E_{bb}, then the maximum value of e_{c2} is equal to zero, a con-venient choice for preliminary consideration, since the second circuit never loads the first one. Thus, the curve of e_{c2} vs. e_1 shown in (d) is the same as (b) with a change in reference level. The over-all transfer curve e_2 vs. e_1, plotted in (e), is obtained from the curves in (c) and (d). The break point P corresponds to plate-current cutoff in the first triode, and Q corresponds to cutoff in the second. The choice of a grid-polariz-ing voltage E that avoids the grid-current states of both of the triodes is highly desirable because μ and r_p are more nearly constant in the nega-tive-grid region. Thus, selecting the voltage E in this manner favors linearity in the amplification state of the circuit (both triodes con-ducting plate current) and yields good clipping (due to cutoff).

A variation of this circuit is shown in Fig. 6.32(a). If we again wish to have zero as the maximum excursion of e_{c2}, as in the previous circuit, we require that E_{cc}/R_g equal E_{bb}/R, assuming R_{b1} is much less than R. The individual transfer curves e_{b1} vs. e_1 and e_2 vs. e_{c2} are the same as those shown in Fig. 6.31(b) and (c). However, in this case there is a loss of amplification in going from the plate of the first triode to the grid of the second:

$$\frac{\Delta e_{c2}}{\Delta e_{b1}} = A = R_g/(R_g + R) \qquad (6.55)$$

The over-all transfer curve is shown in Fig. 6.32(b) with break points and slopes dimensioned. The over-all gain of this amplifier in the range between break points P and Q is G_1G_2A, where G_1 and G_2 are the gains of the individual stages and A is the resistive attenuation. These values

apply for signal frequencies sufficiently low to permit omission of the effects of interelectrode capacitances.   The inadequacy of resistive-
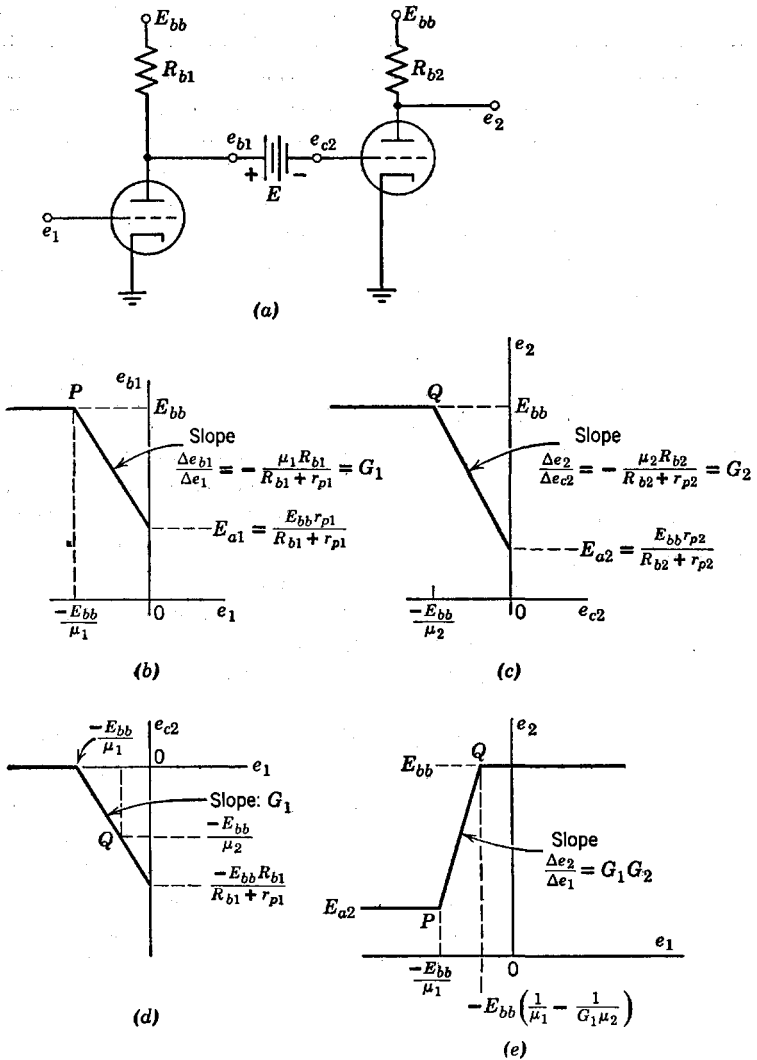
Fig. 6.31.   Direct-coupled grounded-cathode circuits.

circuit models for representing high-frequency circuit behavior is considered in Chapter 8.

$$\text{Slope } \frac{\Delta e_2}{\Delta e_1} = G_1 G_2 A = \left(\frac{\mu_1 R_{b1}}{R_{b1} + r_{p1}}\right)\left(\frac{\mu_2 R_{b2}}{R_{b2} + r_{p2}}\right)\left(\frac{R_g}{R_g + R}\right)$$

$$E_{a2} = \frac{E_{bb} r_{p2}}{R_{b2} + r_{p2}}$$

$$-E_{bb}\left(\frac{1}{\mu_1} - \frac{1}{G_1 A \mu_2}\right)$$

(b)
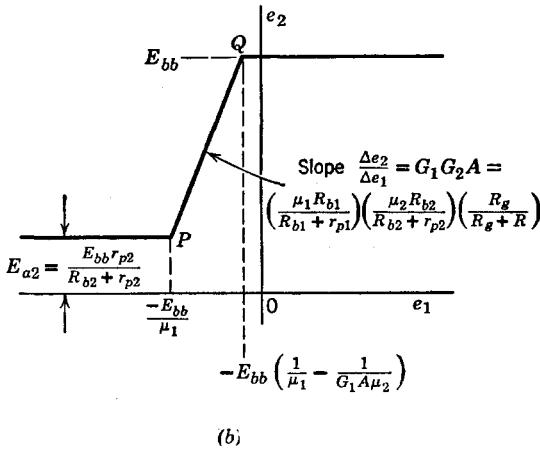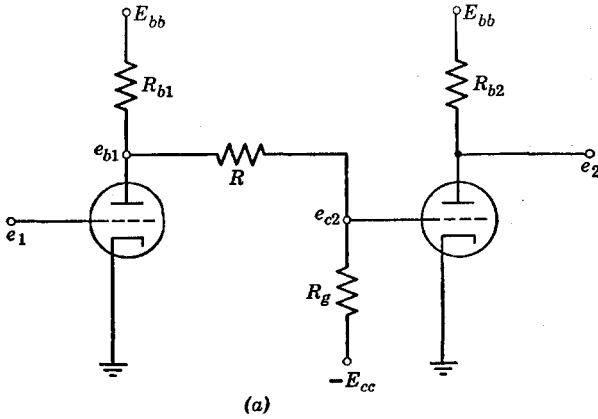
Fig. 6.32. Grounded-cathode circuits coupled by resistance attenuator.

6.19 Cathode-Coupled Circuit

The cathode-coupled circuit shown schematically in Fig. 6.33(a) can be considered to consist of a cathode follower driving a grounded-grid amplifier. The plate load R_{b1} has only a minor effect on the over-all transfer curve e_3 vs. e_1, since it merely augments the plate resistance r_{p1}. However, inclusion of R_{b1} makes the circuit somewhat more general, since an additional output voltage e_2 may be taken from the plate of the first triode. The grid of the second triode is shown connected to a polarizing source E_{cc}. For incremental voltages, this amounts to
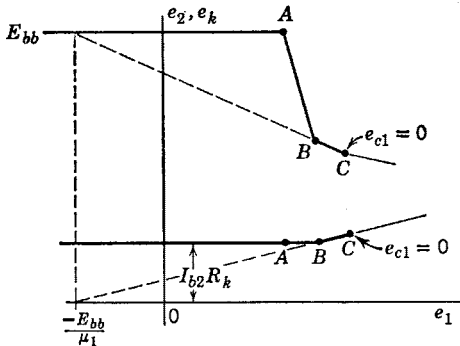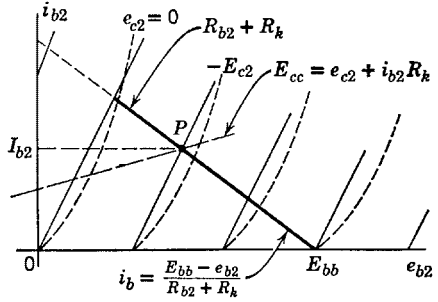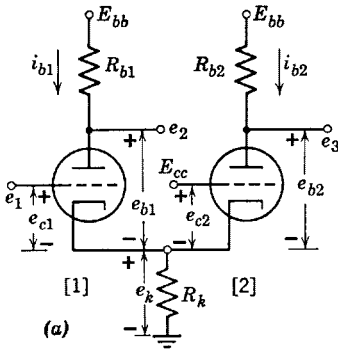
for $e_{c2} < 0$, $i_{b1} = 0$

$$\begin{cases} E_{cc} = e_{c2} + i_{b2}R_k \\ E_{bb} = i_{b2}(R_{b2} + r_{p2} + R_k) - \mu_2 e_{c2} \end{cases}$$

(b)

$$\begin{cases} I_{b2} = \dfrac{E_{bb} + \mu_2 E_{cc}}{R_{b2} + r_{p2} + R_k(\mu_2 + 1)} \\ -E_{c2} = E_{cc} - I_{b2}R_k \end{cases}$$

(c)

At break point A:

$$\begin{cases} e_k = I_{b2}R_k \\ e_2 = E_{bb} \\ e_3 = E_{bb} - I_{b2}R_{b2} \\ e_1 = [I_{b2}R_k(\mu_1 + 1) - E_{bb}]/\mu_1 \end{cases}$$

(d)
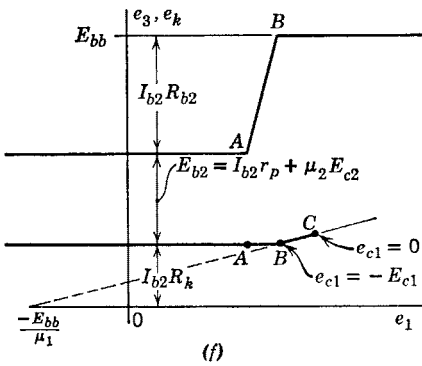
At break point B:

$$\begin{cases} e_k = E_{kB} = I_{b1}R_k \\ e_2 = E_{bb} - I_{b1}R_{b1} \\ e_3 = E_{bb} \\ e_1 = -E_{c1} + I_{b1}R_k \end{cases}$$

where $I_{b1} = \dfrac{E_{kB}}{R_k} = \dfrac{(E_{bb} + \mu_2 E_{cc})}{R_k(\mu_2 + 1)}$

and $E_{c1} = \dfrac{E_{bb} - I_{b1}(R_{b1} + r_{p1} + R_k)}{\mu_1}$

(e)

Fig. 6.33. Cathode-coupled circuit.

grounding the grid. The effects of E_{cc} on the behavior of the circuit will become apparent as the analysis proceeds.

Use of the break-point method facilitates the determination of the transfer curves e_2 vs. e_1 and e_3 vs. e_1. Although we shall use piecewise-

linear models for the triodes, the graphical sketches could be applied equally well to the actual triode curves if a nonlinear resistive analysis is required. As always, some judicious physical reasoning about the circuit behavior results in a considerable reduction of algebraic effort. Rather than analyze all possible states, let us first determine which ones are likely to occur. We can then calculate break points from the simplest of these and thus avoid making any calculations at all for the more complicated circuit-model configurations. For example, let e_1 take on a negative value sufficiently large to cut off plate current in the first triode. Then the second triode conducts at a fixed operating point P, determined only by E_{bb}, E_{cc}, R_{b2}, and R_k.

As e_1 is made less negative, carried through zero, and finally made positive, no changes can occur until the first triode conducts. This must occur when e_1 is slightly less than E_{cc}, since the conduction of the second triode establishes a quiescent value of e_k near E_{cc}. Now if e_1 is increased further, the conduction of the first triode tends to increase e_k. Since E_{cc} is fixed, this tends to make e_{c2} more negative and will eventually cause plate-current cutoff in the second triode. At that point the first triode operates as a basic grid-driven circuit with plate and cathode load. We see, therefore, that for large positive or negative values of e_1 the circuit behavior is determined by single-triode circuits, which we have already analyzed. Break points between each of these circuit states and the one in which both triodes conduct plate current can be determined from the analysis of the two single-triode circuits. Since the transfer relations e_3 vs. e_1 or e_2 vs. e_1 must be piecewise-linear, we can join the break points to complete the curve.

The transfer curves e_2 vs. e_1 and e_3 vs. e_1 are shown in Fig. 6.33(f). In each case, e_k vs. e_1 is plotted on the same coordinates. The differences between the two curves on each plot are e_{b1} and e_{b2}, respectively. Note that e_3 is independent of e_1 to the left of break point A and to the right of break point B. The break point B lies on the transfer curves of e_2 vs. e_1 and e_k vs. e_1 for the first triode alone (with R_{b1} and R_k as loads). For values of e_1 beyond those at break point C ($e_{c1} = 0$), the detailed behavior of e_2 and e_k may be obtained from Fig. 6.26. One is usually interested in having a linear range between A and B; hence E_{cc} and the circuit elements are chosen so that break point B occurs before the grid-current point for tube 1. The quiescent conduction level of tube 2 is usually set at $e_{c2} \leq 0$; hence the grid-current states do not affect the output e_3. The voltage amplification in the region AB can be expressed in terms of the break-point coordinates; for example,

$$G_{31} = \Delta e_3/\Delta e_1 = (e_{3b} - e_{3a})/(e_{1b} - e_{1a}) \qquad (6.56)$$

In a numerical problem, the voltages at the break points yield the gain very readily. The literal expressions for G_{31} and G_{21} are somewhat cumbersome.

$$\mu_1 e_{g1} = i_{p1}(R_{b1} + r_{p1} + R_k) + i_{p2}R_k$$
$$\mu_2 e_{g2} = i_{p1}R_k + i_{p2}(R_{b2} + r_{p2} + R_k)$$
$$e_{g1} = e_1 - (i_{p1} + i_{p2})R_k$$
$$e_{g2} = -e_k = -(i_{p1} + i_{p2})R_k$$

$$(b)$$
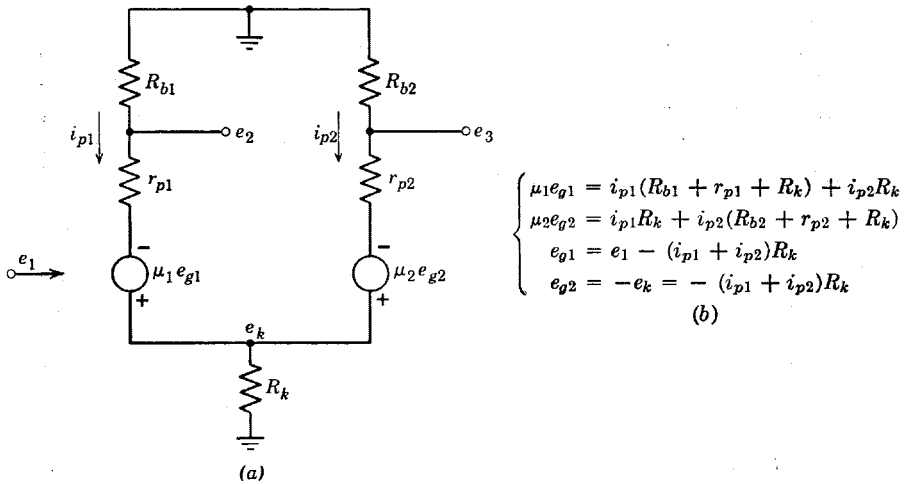
$$\mu_1 e_1 = i_{p1}[R_{b1} + r_{p1} + R_k(\mu_1 + 1)] + i_{p2}R_k(\mu_1 + 1)$$
$$0 = i_{p1}R_k(\mu_2 + 1) + i_{p2}[R_{b2} + r_{p2} + R_k(\mu_2 + 1)]$$

$$(c)$$

$$G_{21} = \frac{e_2}{e_1} = \frac{-i_{p1}R_{b1}}{e_1}$$

$$= \frac{-\mu_1 R_{b1}[R_{b2} + r_{p2} + R_k(\mu_2 + 1)]}{[R_{b1} + r_{p1} + R_k(\mu_1 + 1)][R_{b2} + r_{p2} + R_k(\mu_2 + 1)] - R_k^2(\mu_1 + 1)(\mu_2 + 1)}$$

$$G_{31} = \frac{e_3}{e_1} = \frac{-i_{p2}R_{b2}}{e_1}$$

$$= \frac{\mu_1 R_{b2}R_k(\mu_2 + 1)}{[R_{b1} + r_{p1} + R_k(\mu_1 + 1)][R_{b2} + r_{p2} + R_k(\mu_2 + 1)] - R_k^2(\mu_1 + 1)(\mu_2 + 1)}$$

$$G_{k1} = \frac{e_k}{e_1} = \frac{(i_{p1} + i_{p2})R_k}{e_1}$$

$$= \frac{\mu_1(R_{b2} + r_{p2})R_k}{[R_{b1} + r_{p1} + R_k(\mu_1 + 1)][R_{b2} + r_{p2} + R_k(\mu_2 + 1)] - R_k^2(\mu_1 + 1)(\mu_2 + 1)}$$

$$(d)$$

Fig. 6.34. Incremental analysis of cathode-coupled amplifier.

The circuit of Fig. 6.34(a) is a linear incremental model for the cathode-coupled amplifier in the region AB, where both triodes conduct plate current but with negative grid-to-cathode voltages. The four equations given in (b) are readily reduced to the two given in (c), and

these in turn lead to the desired expressions for gain given in (d). Although the two triodes are normally of the same type, subscripts are used to designate μ and r_p, so that the contributions of each triode can be traced in the final results.

SUPPLEMENTARY READING

L. B. Arguimbau, R. B. Adler, *Vacuum Tube Circuits and Transistors*, John Wiley and Sons, New York, 1956.

B. Chance, *Waveforms*, McGraw-Hill, New York, 1948.

W. G. Dow, *Fundamentals of Engineering Electronics*, John Wiley and Sons, New York, 1952.

D. V. Geppert, *Basic Electron Tubes*, McGraw-Hill, New York, 1951.

T. S. Gray, *Applied Electronics*, 2nd edition, John Wiley and Sons, New York, 1954.

E. A. Guillemin, *Introductory Circuit Theory*, John Wiley and Sons, New York, 1953.

Willis W. Harman, *Fundamentals of Electronic Motion*, McGraw-Hill, New York, 1953.

Samuel Seely, *Electron-Tube Circuits*, McGraw-Hill, New York, 1950.

Samuel Seely, *Electronic Engineering*, McGraw-Hill, New York, 1956.

Karl R. Spangenberg, *Fundamentals of Electron Devices*, McGraw-Hill, New York, 1957.

Robert L. Sproull, *Modern Physics: A Textbook for Engineers*, John Wiley and Sons, New York, 1956.

G. E. Valley and H. Wallman, *Vacuum Tube Amplifiers*, McGraw-Hill, New York, 1948.

PROBLEMS (See Appendix B for triode curves)

6.1. (a) Two high-vacuum triodes have the same grid-to-plate spacing, but triode B has a larger cathode-to-grid spacing than triode A. What are the qualitative differences between their plate curves?

(b) Two high-vacuum triodes are identical except that the grid wires of triode B are smaller in diameter than the grid wires of triode A, the center-to-center grid-wire spacing being the same in both tubes. What are the qualitative differences between their plate curves?

(c) Triode B is exactly twice as large in all dimensions as triode A. What are the differences between their plate curves?

6.2. In what regions of the e_b vs. e_c plane is the approximation $i_b = K(e_b + \mu e_c)^{3/2}$, where K and μ are constants, not valid for a high-vacuum triode, and why? Consider the entire plane.

6.3. Describe the measurements necessary to determine the capacitances C_{pk}, C_{kg}, and C_{gp} in Fig. 6.6.

6.4. The volt-ampere characteristics of a triode vacuum tube are characterized by the expression

$$i_b = 10^{-5}(e_b + 20e_c)^{3/2} \quad \text{(amperes, volts)}$$

(a) Plot on graph paper a family of plate characteristics for this tube over the region

$$0 < e_b < 400 \text{ volts}$$
$$0 < i_b < 14 \text{ ma}$$
$$e_c = 0, -2, -4, \cdots -16 \text{ volts}$$

(Note that when one curve for a given grid potential has been constructed, the other curves can be obtained by properly displacing this curve along the e_b axis.)

(b) From the graphical plot of part (a), determine the values of μ, g_m, and r_p at the following operating points

(i) $i_b = 2.5$ ma   (ii) $i_b = ?$ (iii) $i_b = 13$ ma
    $e_b = 360$ volts $e_b = 240$ volts $e_b = ?$
    $e_c = ?$ $e_c = -8$ volts $e_c = 0$ volts

Does $\mu = g_m r_p$ at each point?

(c) Determine from the analytical function above, expressions for r_p and g_m of the form

$$r_p = f(i_b)$$
$$g_m = g(i_b)$$

Plot both quantities as a function of i_b on one set of coordinates over the range $0 < i_b < 14$ ma. Locate on this plot the values of r_p and g_m determined graphically in part (b).

6.5. From the Type I triode plate characteristics:

(a) Plot constant-current characteristics (e_b vs. e_c) for $i_b = 0.2, 2, 4, 6, 8,$ and 10 ma.

(b) Plot transfer characteristics (i_b vs. e_c) for $e_b = 20, 100, 200,$ and 300 volts.

At the operating point $e_b = 200$, $e_c = -6$, determine:

(c) r_p and μ from the plate characteristics.

(d) g_m from the transfer characteristics.

(e) Check the relation $r_p g_m = \mu$.
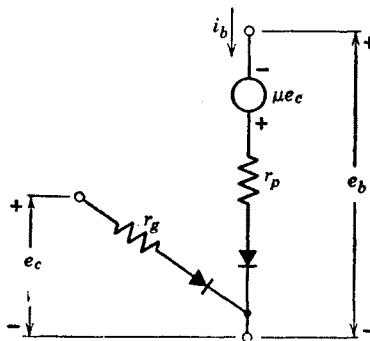
Fig. P6.1

6.6. A Type I triode is to be approximated by the piecewise-linear model shown in Fig. P6.1.

(a) Choose the constants r_p and μ, so that the model draws the same plate current as the actual tube at each of the two points $(e_b = 80,\ e_c = 0)$ and $(e_b = 360,\ e_c = -16$ volts$)$.

(b) Plot the plate curves of the model upon the plate curves of the tube. Outline the area within which the plate current in the model is in error by about 1 ma.
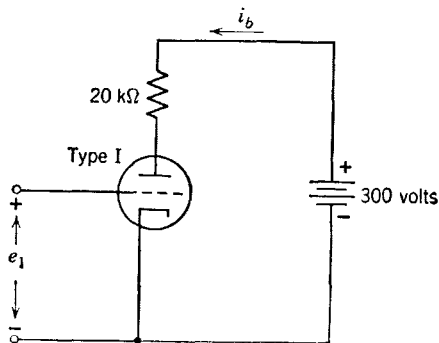
Fig. P6.2

6.7. Using the circuit of Fig. P6.2 and the results obtained in Problem 6.6 (a) plot the the curve of i_b versus e_1 for the negative-grid-voltage region of operation.
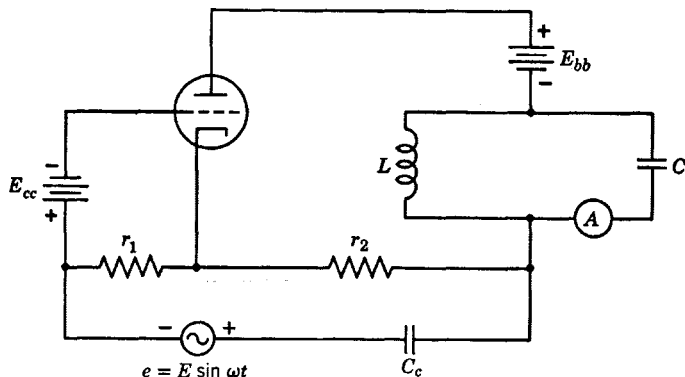
Fig. P6.3

6.8. With the circuit of Fig. P6.3 the μ of a triode can be measured by adjusting r_1 and r_2 until the ammeter A reads zero. Assume that amplitude E is sufficiently small so that operation is linear. Assume $\omega L \gg 1/\omega C$ so that ammeter A reads all of the alternating component of plate current. The reactances of C and C_c are very small at the frequency of operation. The ammeter offers negligible resistance.

(a) Draw the a-c linear incremental equivalent circuit of the system.

(b) Suppose that resistances r_1 and r_2 are adjusted to yield zero ammeter current. Under this condition, derive an expression for μ in terms of the parameters of the equivalent circuit.

(c) If $r_1 = 200$ ohms, $r_2 = 20$ kilohms and $r_p = 80$ kilohms when the ammeter A reads zero, find the μ of the tube.
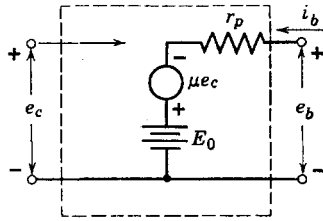
Fig. P6.4

6.9. Determine the values of the constants in the linear circuit model of Fig. P6.4 such that it behaves like the Type II triode in the neighborhood of $e_b = 200$ volts, $e_c = -2$ volts.

(b) Sketch and give values for the corresponding current-source model.

6.10. A certain triode which can be described by the equations

$$\frac{i_b}{I_0} = \left(\frac{e_b + 18e_c}{E_0}\right)^2, \text{ for } e_b > 0, \; e_b + 18e_c \gtrless 0$$

$$\frac{i_b}{I_0} = 0, \text{ for } e_b < 0$$

$$\frac{i_b}{I_0} = 0, \text{ for } e_b + 18e_c \lessgtr 0$$

is to be used as an amplifier in the circuit shown in Fig. P6.5 (E_0 and I_0 are positive constants).
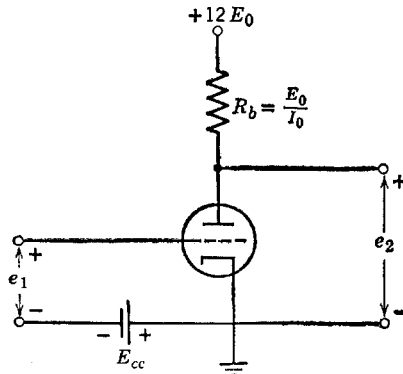
Fig. P6.5

(a) Sketch and dimension the e_b/E_0 vs. i_b/I_0 characteristics for this tube and find a *linear* model that describes the tube in the vicinity of the operating point for $E_{cc} = \frac{1}{3}E_0$.

(b) Determine the incremental gain $\Delta e_2/\Delta e_1$ for $E_{cc}/E_0 = \frac{1}{3}$, and sketch this incremental gain as a function of E_{cc}/E_0.

6.11. Using the Type I triode characteristics, plot the load line and locate the operating point for the circuit in Fig. P6.6. Plot the transfer curve e_2 versus e_1 for negative grid-to-cathode voltage only.
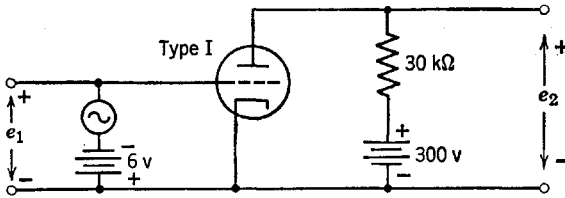
Fig. P6.6

6.12. Construct a set of piecewise-linear curves which approximates the Type I curves. Make the piecewise-linear curve for $e_c = -6v$ tangent to the actual $e_c = -6v$ curve at the operating point.

Locate the load line and operating point of the circuit shown in Fig. P6.6 on these piecewise-linear curves. Plot the transfer curve e_2 versus e_1 obtained from these curves on the same graph used in Problem 6.11.

Draw the piecewise-linear model for total currents and voltages, using the values for the triode parameters determined above. Determine the slope of the e_2 versus e_1 curve.
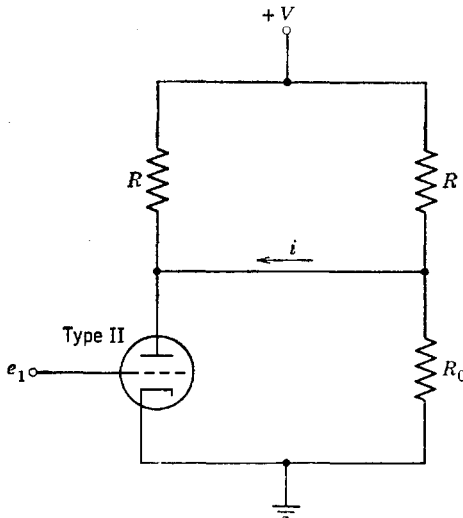
Fig. P6.7

6.13. (a) Referring to the Type II triode curves, show how you would choose R_0 in Fig. P6.7 to make i have the value zero when e_1 has the value zero.

(b) Using a piecewise-linear triode model, adjusted to match the conditions of part (a), and assuming that R_0 is fixed at the value found in part (a), sketch and dimension a curve of i versus e_1. Show only the region in which grid

voltage is negative and plate current is positive. How does this approximate curve differ (qualitatively) from the actual i vs. e_1 relationship?

(c) For small values of e_1 a linear incremental circuit model may be used to calculate i from e_1. Explain how you would determine suitable values for the tube parameters r_p and μ to be used in this incremental model.

(d) Calculate i/e_1 in terms of the constants of the incremental circuit model.
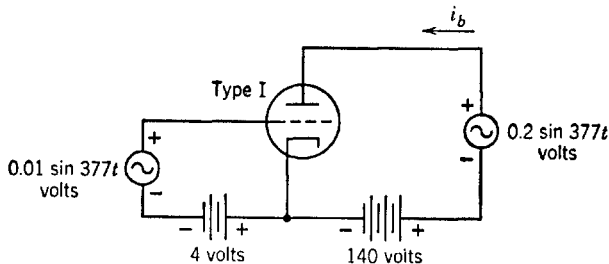
Fig. P6.8

6.14. Write an expression for i_b in milliamperes as a function of time for the circuit shown in Fig. P6.8. Use a piecewise-linear model for the Type I triode.
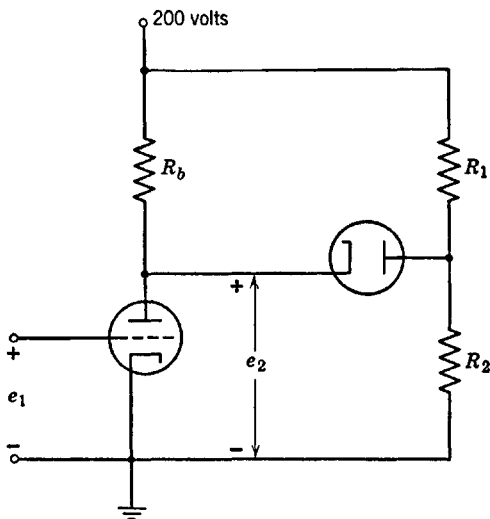
Fig. P6.9

6.15. Assume the triode in the circuit shown in Fig. P6.9 can be represented by linear characteristics with $\mu = 20$ and $r_p = 10$ kilohms. The following conditions are to be met:

For $-10 < e_1 < -5$, $G = |\Delta e_2/\Delta e_1| = 15$

For $-5 < e_1 < 0$, $G = |\Delta e_2/\Delta e_1| = 10$

The diode may be assumed to be ideal. Determine the required values of R_b, R_1, and R_2.

6.16. In the circuit of Fig. P6.10 the transformer is assumed to be ideal, that is $e_1{}^{dc} = 0$, $e_2{}^{ac} = e_1{}^{ac}/n$, and $e_1{}^{ac}/i_1{}^{ac} = n^2 R_2$. Also let C_k be so large that the cathode potential is nearly constant. Assume $\mu = 40$, $r_p = 12$ kilohms.
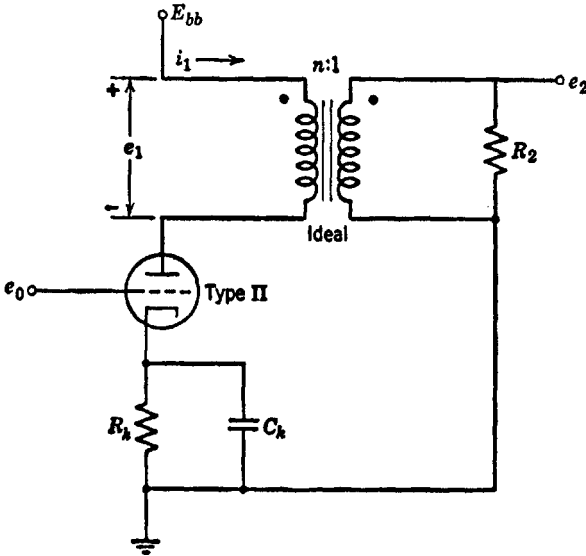
Fig. P6.10

(a) With e_0 a small sine wave voltage of fixed amplitude E_0, find the value of the turns-ratio n which maximizes the a-c output voltage e_2.

(b) With n and E_0 fixed, find the value of R_2 for maximum power delivered to R_2. Compare with the results of part (a).

6.17. As the grid voltage of a triode passes from positive to negative, the grid current becomes very small. In some applications even an exceedingly small grid current is significant, however, and the following scheme is proposed for finding out something about the low grid currents flowing in the negative-grid-voltage region of operation. The triode is connected as a grounded-cathode amplifier and a point-by-point plot of e_b versus e_c is made. The adjustable input source is then removed and an uncharged 0.1 μf capacitor is connected from ground to grid at time zero. The capacitor was previously checked for low leakage by observing that it would hold a charge for many minutes without appreciable diminution. Voltage e_b is recorded and plotted as a function of time. How do you find the grid-current versus grid-voltage curve of the amplifier from the available data and what portions of the curve are reliably determined by this method?

6.18. Sketch and dimension the static e_2 vs. e_1 curve of the amplifier shown in Fig. P6.11. Assume piecewise-linear behavior of the tube. The external diode is to be treated as an ideal rectifier.

6.19. Using the Type I triode characteristics, find e_0 for the circuit shown in Fig. P6.12 when (a) $e_1 = 0$ and (b) $e_1 = 50$ volts.
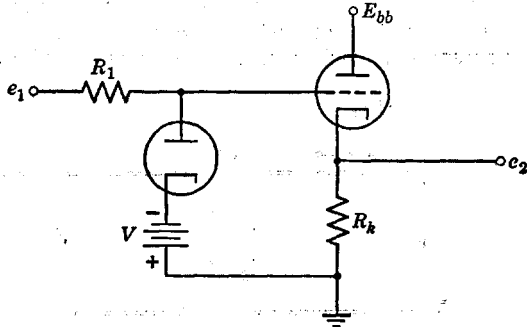
Fig. P6.11
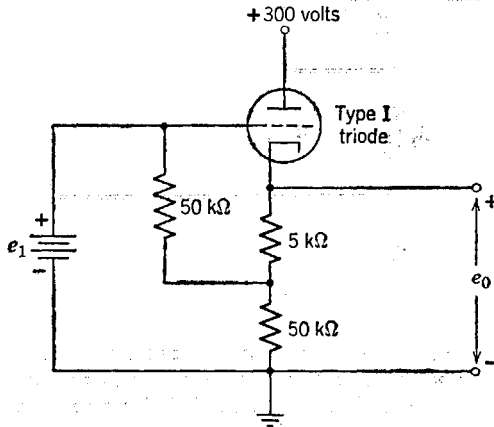
Fig. P6.12

6.20. (a) Plot the load line and find the operating point of the cathode follower circuit in Fig. P6.13 on the Type II triode curves.
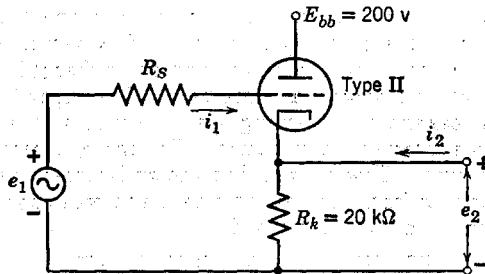
Fig. P6.13

(b) Determine graphically and plot the transfer characteristic e_2 vs. e_1, (1) assuming that $R_s = 10$ kilohms and that the grid-to-cathode circuit

of the tube can be represented by a 1-kilohm resistance in series with an ideal diode, and

(2) assuming $R_s = 0$, $r_g = 1$ kΩ.

6.21. From Fig. P6.14:

(a) If $e_1(t) = 250 \sin \omega t$, sketch and dimension $e_2(t)$.

(b) If $e_1(t) = 0.1 \sin \omega t$, sketch and dimension $e_2(t)$.
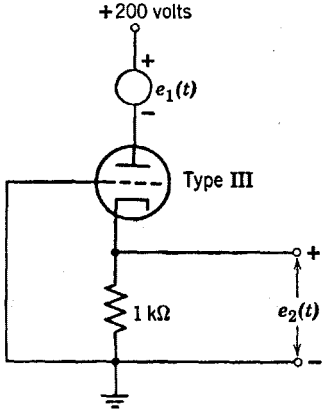
Fig. P6.14
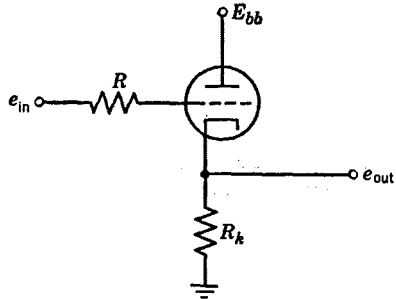
Fig. P6.15
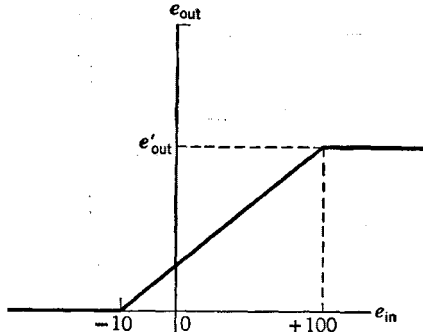
Fig. P6.16

6.22. Approximating the tube in Fig. P6.15 by linearized plate characteristics having $r_p = 10$ kilohms, $\mu = 20$, and cutoff at $e_b = -20e_c$, find the values of R_k and E_{bb} which yield the output-input characteristic shown in Fig. P6.16. Assume that the limiting resistance R is large. What is the value of e'_{out} in Fig. P6.16?

6.23. The cathode-follower circuit shown in Fig. P6.17 is required to handle input signals with peak-to-peak amplitudes up to 140 volts. Grid limiting and cutoff clipping are to be avoided. The total cathode load $(R_1 + R_2)$ is 20 kilohms. Determine appropriate values of R_1 and R_2. The transfer characteristic may be assumed linear between cutoff and $e_c = 0$.
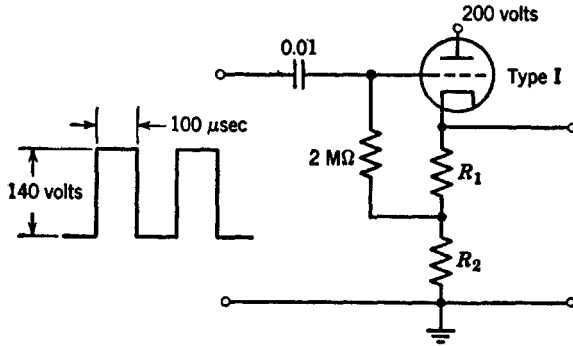
Fig. P6.17

6.24. Assume $i_b = 2e_c + 0.1e_b$ for the triode in Fig. P6.18, where units are volts and milliamperes.

(a) Find that value of resistor R for which the curve of meter current i versus input voltage e passes through the origin in the i versus e plane. Use
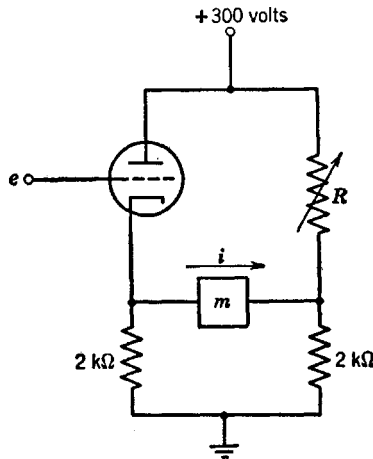
Fig. P6.18

this value of R in parts (b) and (c) of the problem. Remember that e is measured from ground.

(b) Calculate the sensitivity $\Delta i/\Delta e$. Assume that the meter resistance is negligible.

(c) Sketch and dimension the curve of i vs. e, showing the entire region within which grid voltage is negative and plate current is positive.

6.25. For the circuit shown in Fig. P6.19, sketch the transfer characteristic e_2 as a function of e_1 for V_b positive and also for V_b negative. Give the slope of the characteristic in terms of the circuit parameters.

6.26. For the circuit shown in Fig. P6.20, plot the path of the operating point on the e_b vs. i_b plane as e_s is varied from $-\infty$ to $+\infty$. Assume a piecewise-linear triode model where $r_p = 1000$ ohms, $\mu = 5$.
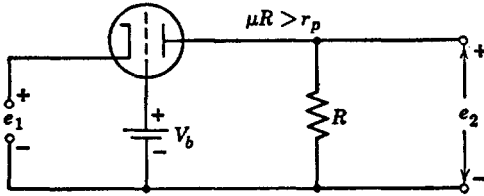
Fig. P6.19
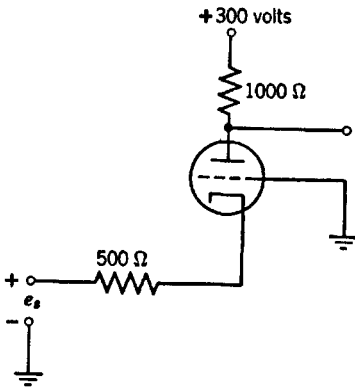
Fig. P6.20
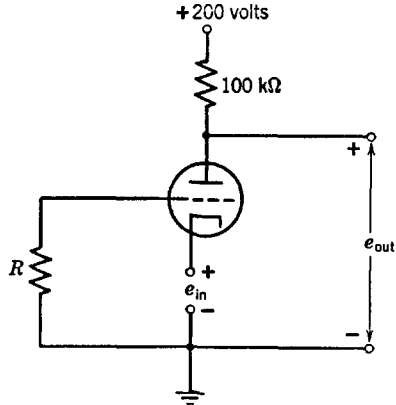
Fig. P6.21

6.27. For the circuit of Fig. P6.21 plot the complete output-input character-istic for an input voltage range from -250 to $+50$ volts. Assume the plate characteristics are piecewise-linear with $\mu = 100$, $r_p = 50$ kilohms. When grid current flows, the voltage drop through resistance R is sufficiently large to hold e_c at a value very nearly equal to zero. Show the locus of operation on the plate characteristics.
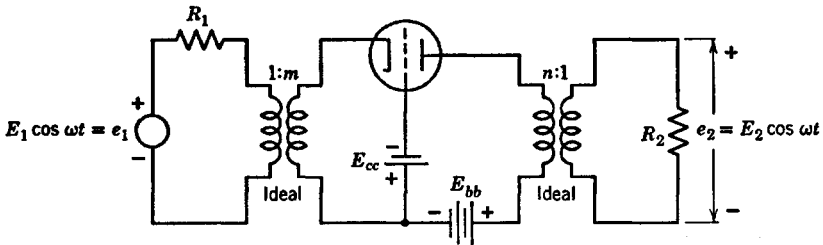
Fig. P6.22

6.28. The transformers shown in Fig. P6.22 are ideal at frequency ω but each winding is a short circuit for d-c. Assume that E_1 is small, so that the operation of the circuit is linear.

(a) Derive an expression for the power gain $G = P_L/P_A$, where the available source power P_A and the load power P_L are defined as: $P_A = E_1{}^2/4R_1$, $P_L = E_2{}^2/R_2$. Gain G is a function of R_1, R_2, m, n, μ, and r_p.

(b) The maximum gain obtainable by adjustment of turns-ratio n is defined as the available gain G_A. Find G_A in terms of R_1, m, μ, and r_p. Plot $G_A/(\mu + 1)$ vs. $(M^2 R_1)/r_p/(\mu + 1)$.

(c) The maximum gain obtainable by simultaneous adjustment of both m and n is defined as the maximum available gain G_{Am}. Find G_{Am} in terms of μ and r_p.

(d) Show that for fixed E_1, R_1, and R_2, but adjustable m and n, the power-gain maximization also maximizes the voltage amplification E_2/E_1.

6.29. Sketch the e vs. i characteristics of the circuit shown in Fig. P6.23. A piecewise-linear model should be assumed.
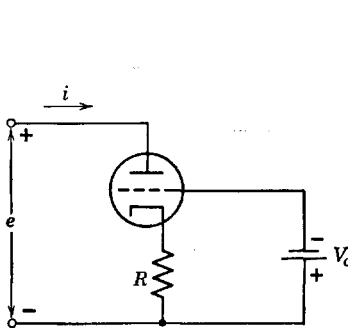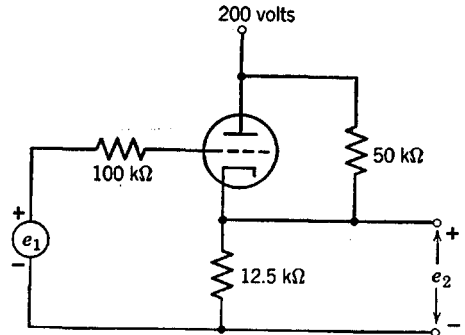
Fig. P6.23   **Fig. P6.24**

6.30. (a) Sketch and dimension the static transfer curve e_2 vs. e_1 for the cathode follower circuit shown in Fig. P6.24. Assume a piecewise-linear triode with $\mu = 20$, $r_p = 10$ kilohms, and $r_g = 0$.

(b) Determine $e_2(t)$ for the circuit shown in Fig. P6.24 with $e_1(t) = +35 + 50 \sin \omega_0 t$.

(c) If $e_1(t) = E_1 + 40 \sin \omega_0 t$ is applied to the circuit of Fig. P6.24, choose a value of E_1 that will result in a sinusoidal output $e_2 = E_2 + E_m \sin \omega_0 t$. Determine E_2 and E_m.

6.31. Sketch approximate piecewise-linear curves of e_2 and e_3 versus e_1, on the same coordinates, for the circuit shown in Fig. P6.25. Show only the negative grid-voltage region. Choose the constants μ and r_p for the piecewise-linear triode model in order to give a reasonable approximation to the Type I triode curves. Let $E_{bb} = 300$ volts and let both R_b and R_k equal 10 kilohms.

6.32. Tabulate the states of each of the three diodes in Fig. 6.26 for the five regions of operation.

6.33. A Type I triode is connected as shown in Fig. P6.26. Confining attention to the negative-grid-voltage region of operation, plot on the plate curves the locus of (a) varying e_1 for $R_b = 15$ kilohms, $R_k = 5$ kilohms, (b) varying R_b for $e_1 = 25$ volts, $R_k = 5$ kilohms, (c) varying R_k for $e_1 = 25$ volts, $R_b = 15$ kilohms, (d) What is the value of i_b for $e_1 = 25$ volts, $R_b = 15$ kilohms, $R_k = 5$ kilohms? (e) Plot i_b vs. e_1 for $R_k = 5$ kilohms, $R_b = 15$ kilohms.
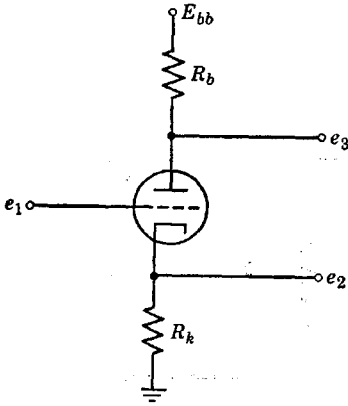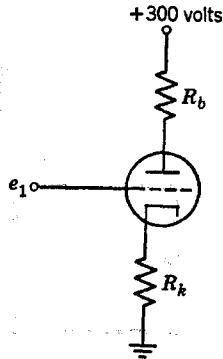
Fig. P6.25   Fig. P6.26

6.34. (a) For the circuit shown in Fig. P6.27, graphically find the quiescent operating point.
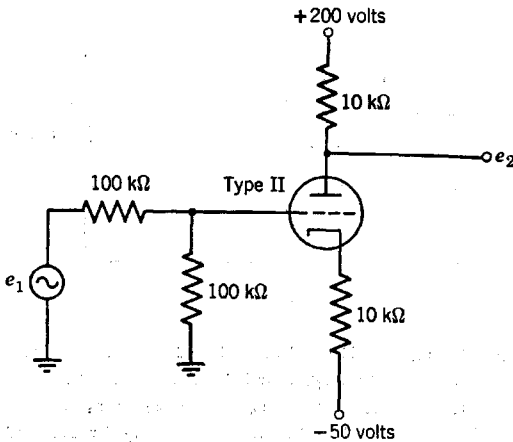
Fig. P6.27

(b) At the quiescent operating point, calculate the incremental parameters μ, r_p, and g_m. Show the piecewise-linear model of the circuit of Fig. P6.27.

(c) For a small signal Δe_1, calculate the incremental gain of the circuit.

6.35. For the Type II triode connected as in Fig. P6.28, plot the complete output-input characteristic for an input voltage range from -20 to $+30$ volts. When grid current flows, the voltage drop through resistance R is sufficiently large to hold e_c at a value very nearly equal to zero.

6.36. Find the operating point $(i_c, i_b$ and $e_b)$ for the Type I triode in the circuit of Fig. P6.29. Use the plate and grid characteristics for positive values of grid voltage. A cut-and-try process is necessary.

6.37. (a) Using the Type I triode characteristics, plot the output-input
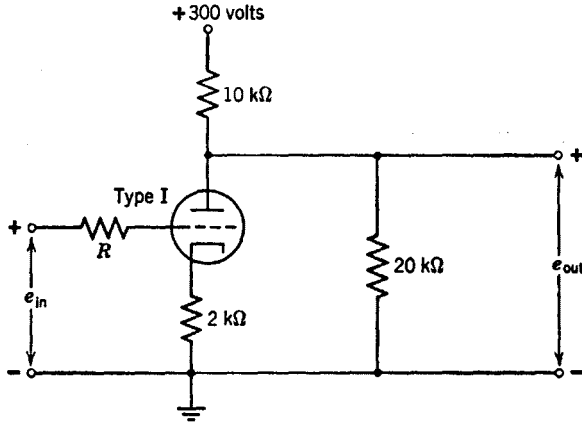
Fig. P6.28
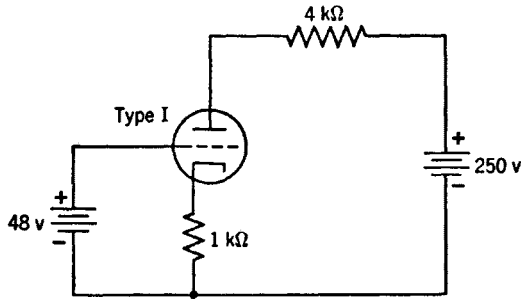
Fig. P6.29

characteristic (e_{out} vs. e_{in}) of the circuit shown in Fig. P6.30 over the range $-20 < e_{in} < +10$ volts. The suggested method is to make a table of values of e_c, e_b, e_k, $(300 - e_b)/50$, $e_{in} = e_c + e_k$, $e_{out} = e_b + e_k$, for $e_c = 0$, -4, -8, -12, -14, -16, from which the characteristic may be plotted.

(b) At $e_{in} = 0$, compute the plate dissipation $e_b i_b$ and the power supplied by the 300-volt source.

6.38. (a) For an input voltage $e_{in} = 8 \cos 1000t$ volts in Fig. P6.30, make a sketch (not a plot) of the manner in which the output voltage waveform is distorted from a true sinusoidal shape.

(b) If harmonics higher than the second are ignored, the output voltage may be written in the form

$$e_{out} = A_0 + A_1 \cos(1000t + \theta_1) + A_2 \cos(2000t + \theta_2)$$

Evaluate the constants by sampling the wave at $1000t = -\pi, -\pi/2, 0, \pi/2, \pi$. (Hint: show that $\theta_1 = \theta_2 = 0$.) Express the ratio of A_2 to A_1 as a percentage second-harmonic distortion.

(c) Ignoring the second harmonic and approximating the plate voltage by

the expression

$$e_b = \left(\frac{e_{b\ max} + e_{b\ min}}{2}\right) - \left(\frac{e_{b\ max} - e_{b\ min}}{2}\right) \cos 1000t$$

compute the time-average a-c power delivered to the 49-kilohm load, and also the time-average plate dissipation.
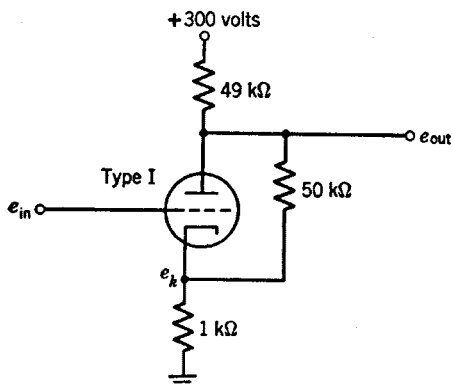
Fig. P6.30

(d) How does the time-average plate dissipation compare with the no-signal value found in Problem 6.37? Explain the difference.

6.39. For an input voltage $e_{in} = 0.8 \cos 1000t$ volts in Fig. P6.30, the output voltage may be written as $e_{out} = A_0 + A_1 \cos 1000t$.

(a) Using the appropriate linear incremental equivalent circuit, find the value of A_1.

(b) How does the voltage amplification $A_1/0.8$ compare with the corresponding ratio $(A_1/8)$ in Problem 6.38?

(c) How does the voltage amplification compare with the slope of the output-input characteristic found in Problem 6.37? At what point should the slope be compared?

(d) How does the value of A_0 compare with the value found in Problem 6.38? Explain the difference.

6.40. In Fig. P6.31, a triangular waveform $e_1(t)$ is to be converted into a trapezoidal wave $e_2(t)$ by the circuit shown in Fig. P6.32. Assume a piecewise-linear model for the triode ($\mu = 20$, $r_p = 10$ kilohms, $r_g = 0$), and assume grid current is negligible in comparison with plate current. Determine the values required for R_b, R_k, and R.

6.41. (a) For the circuit shown in Fig. P6.33, locate the operating point of the tube on Type I triode characteristics. Use a systematic procedure and not pure trial and error. Assume $i_1 = 0$.

(b) For the circuit of Fig. P6.33, determine the tube parameters at the operating point and calculate incremental voltage gain $\Delta e_2/\Delta e_1$, input impedance $\Delta e_1/\Delta i_1$, and output impedance.

6.42. The circuit in Fig. P6.34 is called a phase inverter. It provides two outputs e_2 and e_3. The input voltage e_1 is sinusoidal and of a high enough
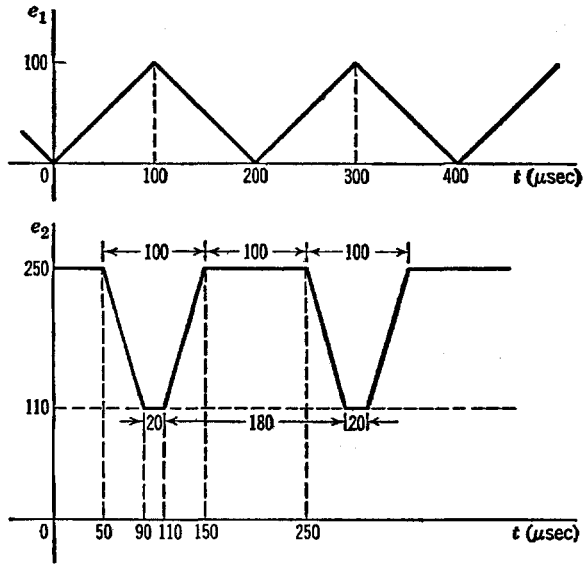
Fig. P6.31
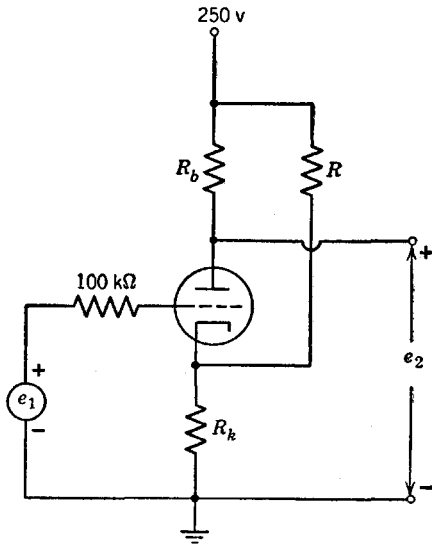
Fig. P6.32
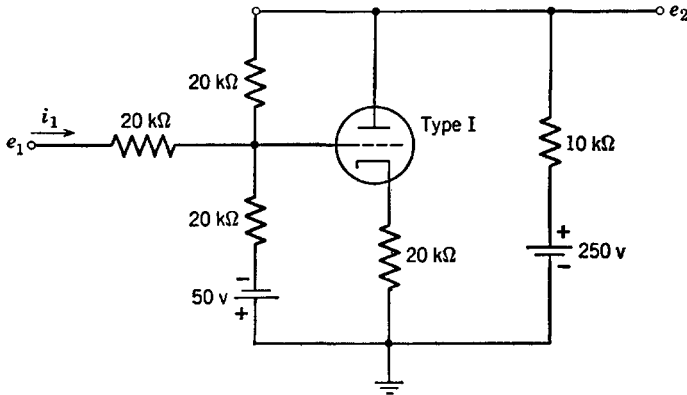
Fig. P6.33

frequency that the reactance of the capacitors C_1 can be neglected. The circuit operates only in the negative-grid region. Using a linear incremental circuit ($\mu = 100$, $r_p = 50$ kilohms) find the gains e_2/e_1 and e_3/e_1, and the phase angle between e_2 and e_3.
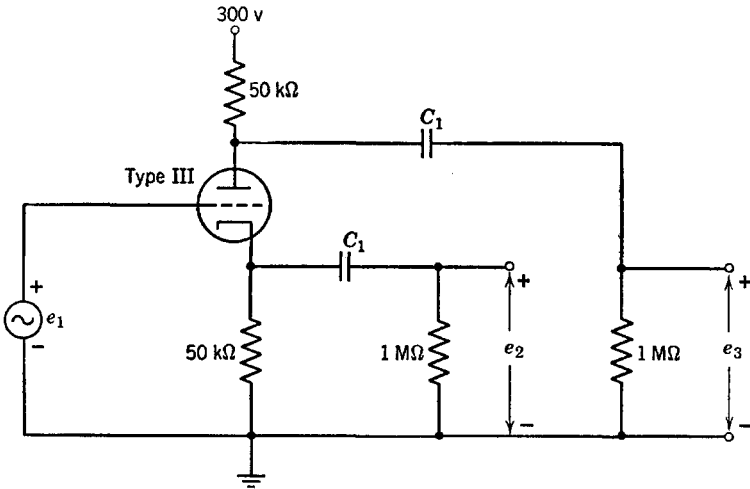
Fig. P6.34

6.43. For the cascode circuit of Fig. P6.35, in which the tubes are assumed to be identical and piecewise-linear:

(a) Sketch and dimension the boundary of the region in the V/E_{bb} versus e_1/E_{bb} plane within which both tubes conduct plate current but no grid current.

(b) Sketch and dimension the curve of e_2 vs. e_1. Use the value of V giving the widest range of e_1 between cutoff and grid current. What is the value of $A = \Delta e_2/\Delta e_1$?
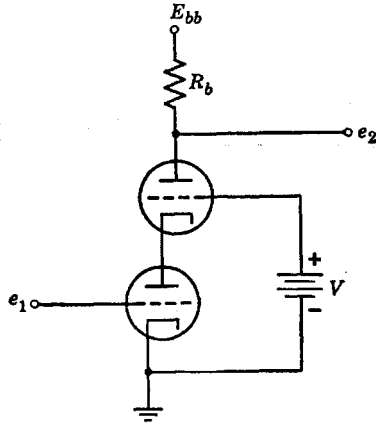
Fig. P6.35

6.44. When operating in the negative grid-voltage region the vacuum triodes in the circuit shown in Fig. P6.36 can both be represented by a linear circuit model with plate resistance r_p, and amplification factor μ. Find the
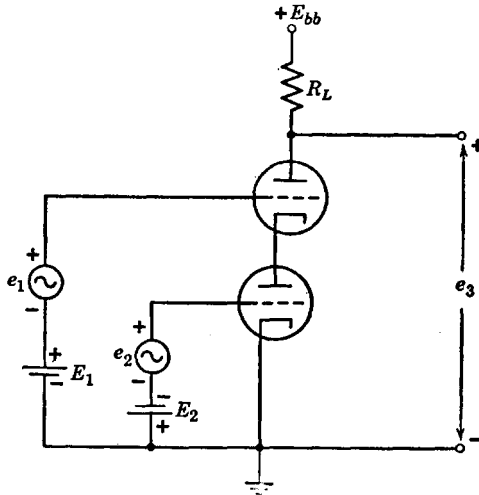
Fig. P6.36

incremental output voltage as a function of the input voltages e_1 and e_2. Assume that the biasing voltages are such that the triodes operate in the negative grid region.

6.45. Assuming that the tubes are operating linearly and have equal μ and r_p, determine the gain e_2/e_1, and the output impedance of the circuit shown in Fig. P6.37.
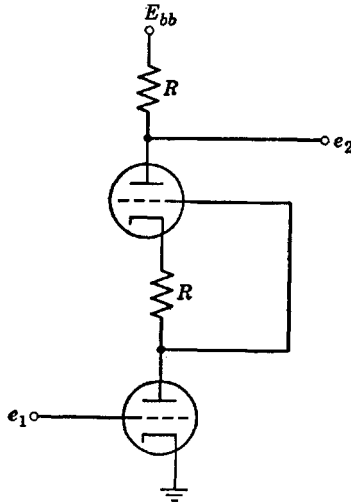
Fig. P6.37

6.46. Locate the static operating point of the circuit in Fig. P6.38.
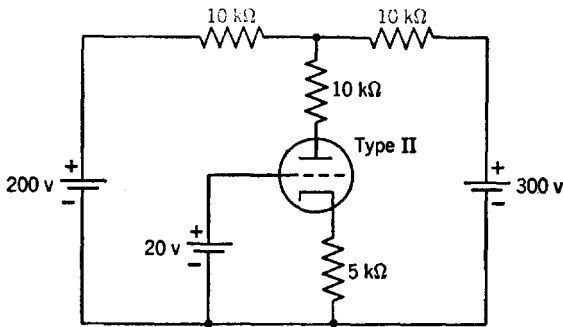
Fig. P6.38

6.47. In Fig. P6.39, determine the incremental amplification e_2/e_1 as a function of the circuit constants. The two tubes are identical. Find the output impedance.

6.48. Using the tube characteristics find the operating points by trial and error for the two Type I triodes in the cathode coupled amplifier shown in Fig. P6.40. Find the incremental parameters for the two tubes at their respective operating points and using these values, determine the incremental output impedance of the amplifier. (R_2 is to be regarded as the output load.) Derive your results using suitable literal symbols for the circuit coefficients. Substitute numerical quantities only after obtaining your results in literal form.
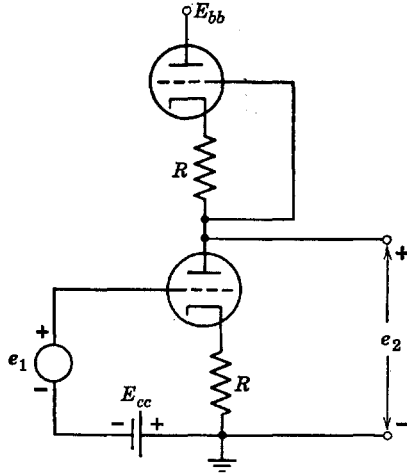
Fig. P6.39
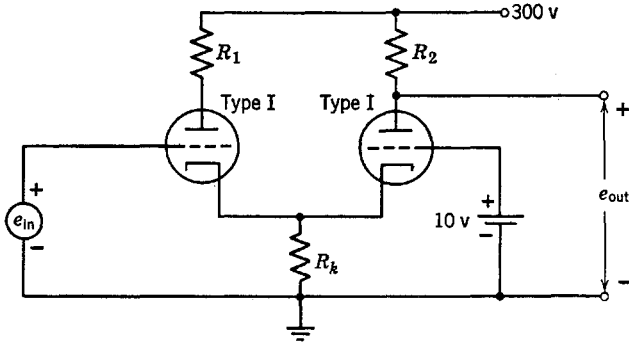
$R_1 = 5,000$ ohms
$R_2 = 10,000$ ohms
$R_k = 1,000$ ohms
$e_{in} = +8v$ d-c

Fig. P6.40
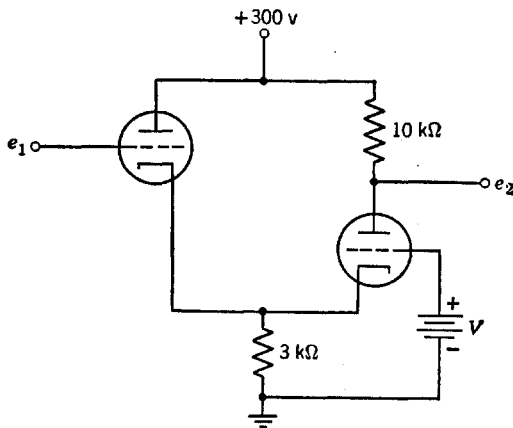
Fig. P6.41

6.49. Using a piecewise-linear triode model with $\mu = 40$, $r_p = 40$ kilohms, plot e_2 versus e_1 for the cathode-coupled amplifier shown in Fig. P6.41. Use that value of V which yields the widest amplification region (both tubes conducting) subject to the restriction that neither tube shall draw grid current in this region.
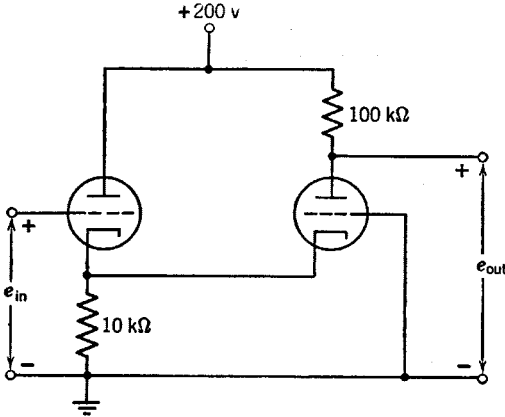
Fig. P6.42

6.50. For the circuit of Fig. P6.42, plot the complete output-input characteristic for an input voltage range from -15 to $+10$ volts. Both tubes have $\mu = 50$, $r_p = 40$ kilohms.

6.51. Using a piecewise-linear model of a triode with $\mu = 20$ and $r_p = 10$ kilohms, design single-stage amplifiers to fulfill the following requirements:

Amplifier No. 1: The incremental voltage gain is to be 19. The input signal, derived from a generator with a 20-kilohm source impedance, has a maximum amplitude of 15 volts peak to peak.

Amplifier No. 2: For a fixed value of e_1, this amplifier containing no transformers must deliver the maximum possible power to a 500-ohm load.

Amplifier No. 3: This amplifier is to have an incremental gain of 5, and must be designed to reduce the dependence of the gain on the μ of the tube.
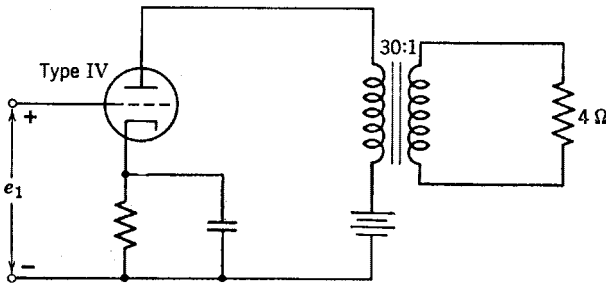
Fig. P6.43

6.52. The quiescent operating point of the Type IV triode in the amplifier circuit of Fig. P6.43 is at $i_b = 60$ ma, $e_c = -60$ volts. For the purpose of

this problem, the transformer may be considered ideal and the capacitor to have zero impedance.

(a) Determine the quiescent plate dissipation.

(b) Assuming essentially linear operation, determine the amplitude of the sinusoidal grid signal required to develop $\frac{1}{2}$ watt in the 4-ohm resistor.

(c) Determine the average plate dissipation when the signal specified in part (b) is applied. (The slight change in average plate current which occurs when the signal is impressed may be neglected.)

$E_{bb} = 250$ v
$R_k = 200\ \Omega$
$R_L = 4\ \Omega$
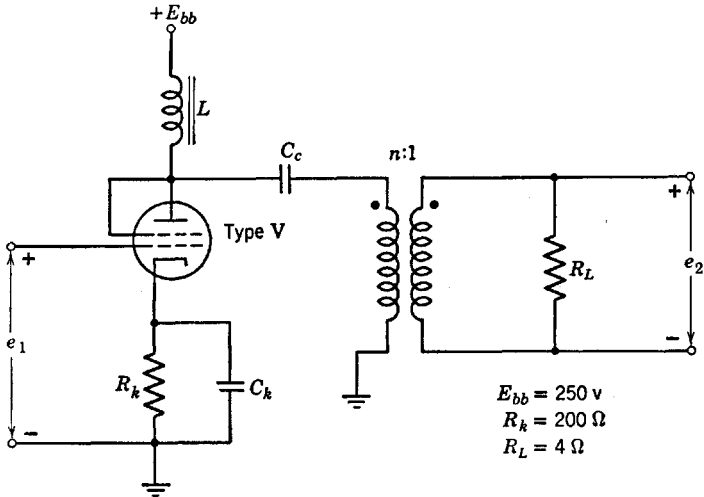
Fig. P6.44

6.53. In Fig. P6.44 is shown the circuit of a shunt-fed amplifier. At the frequency of operation, the reactance of the inductor L is very large, and the reactances of the capacitors C_K and C_c are very small. The inductor resistance is negligible and the transformer is ideal.

(a) Determine on the Type V triode characteristics the operating point (e_b and i_b) when $e_1 = 0$.

(b) If $e_1 = 0.1 \sin \omega t$, what is the turns ratio n for maximum power transfer to the load R_L?

(c) Draw on the characteristics the operating path of the tube when $e_1 = 5 \sin \omega t$.

(d) Determine from the incremental equivalent circuit the voltage gain of the circuit, $A = \Delta e_2/\Delta e_1$, at the operating point.

6.54. The circuit shown in Fig. P6.45 provides an adjustable supply voltage V for a resistive vacuum-tube load that draws a current varying between 0 and 10 ma. Determine the value of E_1 required to make $V = 150$ volts when $I = 5$ ma. With this value of E_1, determine V for $I = 0$ and 10 ma, and calculate the percent regulation.

6.55. The circuit shown in Fig. P6.46 provides an adjustable supply voltage for the variable load R. Assume the triode to be linear with $\mu = 40$ and $r_p = 20$ kilohms. For a given value of E_1 and with $e_c \leqslant 0$ the circuit provides some voltage regulation. Determine (a) and (b) as functions of E_1.
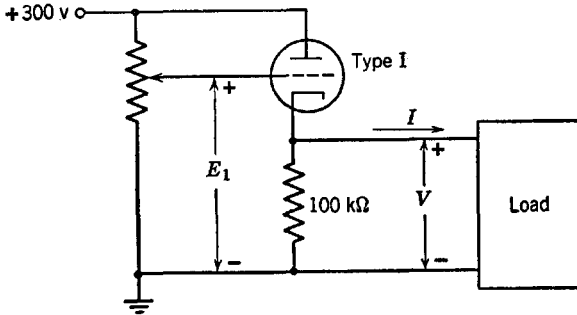
Fig. P6.45

(a) Open-circuit voltage available at the load.

(b) Maximum load current for which reasonable voltage regulation is obtained.
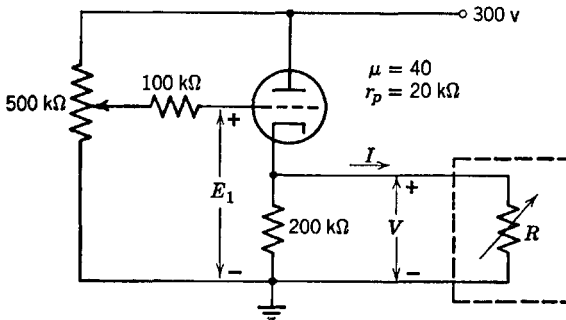
Fig. P6.46
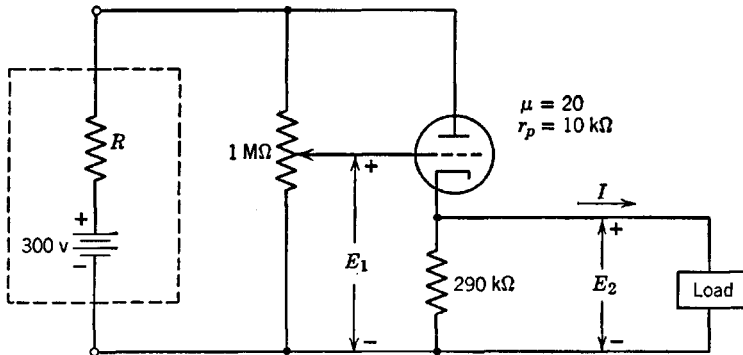
Fig. P6.47

6.56. In the circuit shown in Fig. P6.47, assume the internal resistance of the power supply to be zero $(R = 0)$.

(a) With $I = 0$, determine the value of E_1 needed to make $E_2 = 250$.

(b) With a resistive load connected, plot E_2 vs. I.

(c) Over what range of I is the voltage "regulated"?

6.57. Repeat Problem 6.56 with $R = 1$ kilohm.

6.58. The series regulator circuit shown in Fig. P6.48 is used to maintain a fixed output voltage essentially independent of supply and load fluctuations. The questions below are concerned with the operation and limitations of this regulator.

Type IV Maximum Ratings
$i_{b \, avg} = 120$ ma
$e_{b \, avg} = 300$ volts
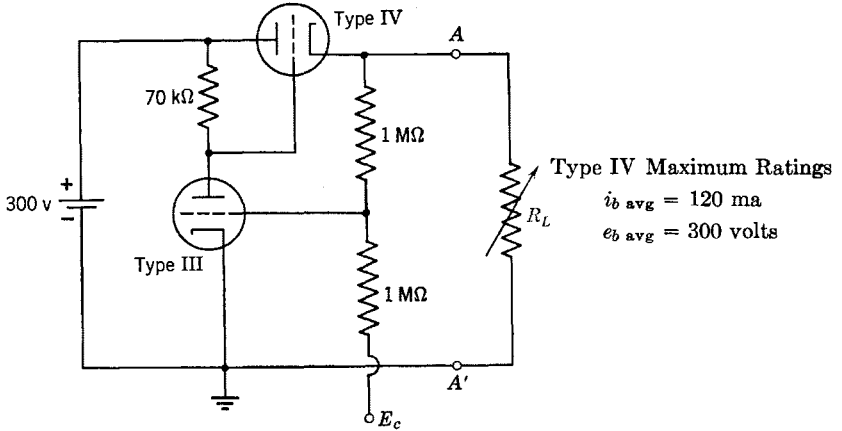
Fig. P6.48

(a) Determine the quiescent operating points and the bias, E_c, necessary to deliver 200 volts at 70 ma to the load.

(b) For the conditions of (a) determine the Thevenin equivalent of the circuit to the left of terminals $A - A'$ (i.e., the circuit seen by R_L) valid for small changes in load current.

(c) For a load voltage of 200 volts, what is the maximum power output that can be obtained without exceeding the tube ratings?

Other Control Valves and their Circuit Models

7.1 Introduction

In the preceding chapters the properties of diodes, vacuum triodes, and transistors have been treated in some detail. There are, of course, many other electronic devices, each with its own peculiarities. In this chapter some of the fundamental properties of control valves are discussed and three specific devices, the vacuum pentode, the cryotron,* and the thyratron, are considered as examples. The purpose of this chapter is to suggest the generality and usefulness of a few simple ideas in the process of characterizing an electronic device.

7.2 The Control Valve

By the word "valve" we mean a physical device which imposes some constraint or relationship between two physical variables whose product is power. If the relationship does not involve time, then the valve is said to be "amnesic" or "memoryless." Figure 7.1 shows (a) an amnesic valve for which the flow-versus-pressure relationship may be either linear, as in (b), or nonlinear, as in (c). In each case, (b) or (c), the valve is *passive*, since the curve never enters the second or fourth quad-

---

* D. A. Buck, The Cryotron — A Superconductive Computer Component, *Proc. I.R.E.*, Vol. 44, p. 482 (1956).

rants of the flow-pressure plane, and the instantaneous power absorbed by the valve is therefore never negative.

By a *control valve* we mean a valve whose flow-versus-pressure relationship is influenced by some other physical variable called the control
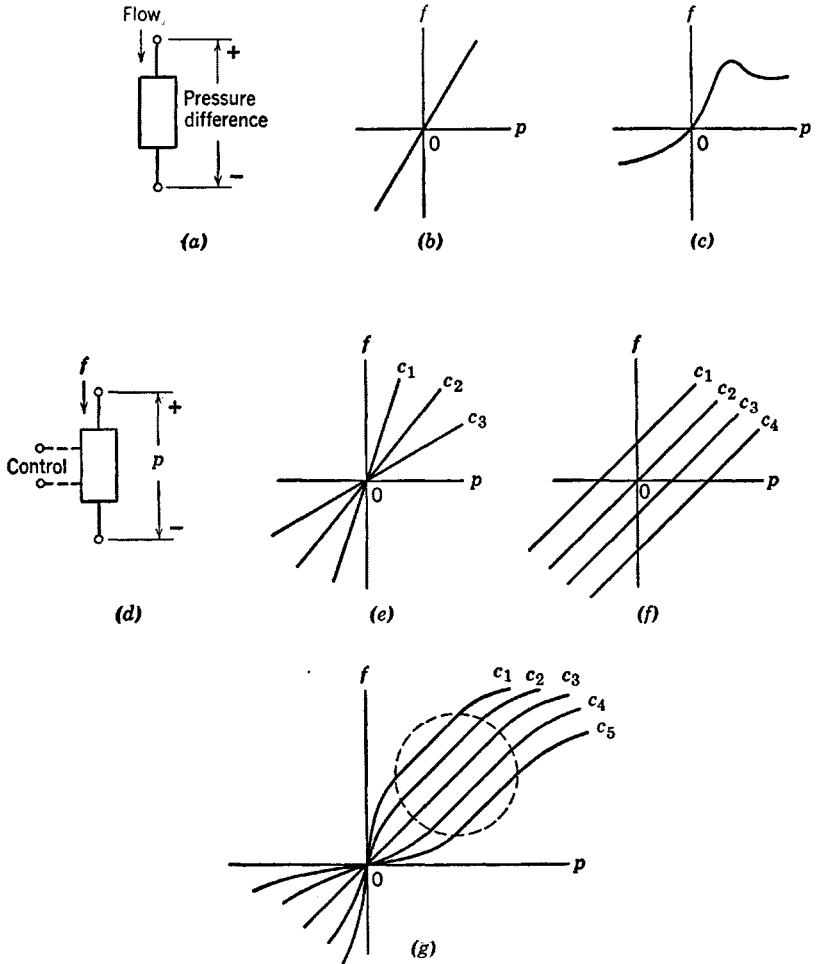
Fig. 7.1. Valves and valve curves.

parameter or control variable. The control variable is itself either a pressure or a flow and is applied to the valve at some auxiliary terminal-pair as indicated in Fig. 7.1 (*d*). Hence the control variable is one of a pair of variables whose product is power, and the relationship between

these two variables is fixed by the characteristics of the control valve. We shall say that the control valve is *unilateral* if the control-variable-pair relationship is unaffected by the value of the pressure or flow in the main part of the valve. In other words, the control influences the valve but there is no reaction or influence in the reverse direction. If application and maintenance of the control signal require negligible power, then the valve has *ideal control*.

An *ideal unilateral amnesic control valve* is completely described by a family of valve curves such as that shown in Fig. 7.1(e), one curve for each of a number of different values of the control parameter c, with the tacit assumption that we can interpolate between these curves for other values of c. (If the valve is not unilateral, we require another such family of curves to relate the two variables at the control point, with either f or p playing the role of a control parameter in the reverse direction.)

The ideal unilateral control valve of Fig. 7.1(e) is passive (at least for the range of c shown), since none of the curves enter the second or fourth quadrants. For any fixed c, the relationship between f and p is linear, but the complete function $f(p, c)$ is nonlinear. An ideal multiplier $(f = pc)$ or divider $(f = p/c)$ has characteristic curves like those in Fig. 7.1(e). The carbon microphone and the kitchen faucet fall roughly into this category.

If the function $f(p, c)$ is linear, then the valve curves are parallel straight lines with equal spacing for equal increments of the control parameter, as shown in Fig. 7.1(f). It is apparent, therefore, that an ideal unilateral control valve cannot be passive if it is linear.

Many devices exhibit a control effect arising from physical laws that are inherently linear over a significant region of operation. In the vacuum triode, for example, grid control is governed, essentially, by electrostatic field equations. The uniform horizontal spacing μ of the triode curves in the central region of operation is a consequence of the linearity of these field equations. Similarly, the emitter-to-collector current-control ratio of a transistor is essentially a constant dependent upon the transistor geometry, with the result that the transistor collector curves are very nearly parallel and equally spaced in the usual region of interest. Thus we have some reason to be concerned with control devices which are regionally-linear as indicated by the curves within the dashed contour in Fig. 7.1(g). The point is that if an ideal unilateral amnesic control valve is linear in some region, then it must necessarily become nonlinear outside this region in order to remain passive, as illustrated in Fig. 7.1(g).

The above discussion gives us the background for a better under-

standing of the circuit models for such devices. In the next two articles we shall see how the circuit model of an *ideal unilateral amnesic passive regionally-linear control valve* can be represented as a combination of two parts, one providing ideal unilateral linear control and the other providing the nonlinearities necessary to insure passivity.

7.3   The Energy Valve

Figure 7.2(*a*) shows the ideal diode or ideal rectifier, a circuit element that has appeared repeatedly in the preceding chapters. The ideal diode is not only passive but also lossless, since the product vi is always identically equal to zero. The ideal diode can be completely characterized by statements that its current is never negative, its voltage is never positive, and the product vi is always zero. The combination of two diodes indicated in Fig. 7.2(*b*) is a two-terminal-pair "filter" having the marvelous property that energy can flow through the filter from left to right but *never from right to left!* To see this, first observe that i_1 is never negative and that v_1 cannot be negative when i_1 is positive. Thus, the instantaneous power $v_1 i_1$ can never be negative. Now, since each of the diodes is lossless, the complete circuit must be lossless and the output power $v_2 i_2$ must be equal to the input power $v_1 i_1$. Hence the circuit shown in Fig. 7.2(*b*) is a trapdoor or valve through which energy can pass in *only one direction.*

Parts (*c*) and (*d*) of Fig. 7.2 offer two additional forms of the lossless energy valve. The lattice structure (*d*) is recognizable as a diode bridge circuit and is obtainable by connecting two circuits of the type shown in part (*c*) with their outputs in parallel and their inputs in series opposition.

The energy valve is a useful building block in the construction of a circuit model for a passive device. No matter what we may connect to the output terminals, the complete circuit, as viewed from the input terminals, v_1 and i_1, is automatically passive.

7.4   Circuit Models for Control Valves

Figure 7.3(*a*) shows the circuit model for an ideal unilateral linear (but not passive) control valve. The voltage source V is assumed to be unilaterally controlled by some physical variable at an ideal control point not shown in the figure. The associated valve curves are indicated in (*b*). The model is obviously not passive since power can be *extracted*

(a)

$i \geq 0, v \leq 0$

$vi \equiv 0$

(b)

$i_1 \geq 0$

$v_1 \geq 0$ for $i_1 > 0$

$v_1 i_1 \equiv v_2 i_2 \geq 0$

(c)

$v_1 \geq 0$

$i_1 \geq 0$ for $v_1 > 0$

$v_1 i_1 \equiv v_2 i_2 \geq 0$

(d)

$i_1 \geq 0$ for $v_1 > 0$

$i_1 \leq 0$ for $v_1 < 0$

$v_1 i_1 \equiv v_2 i_2 \geq 0$
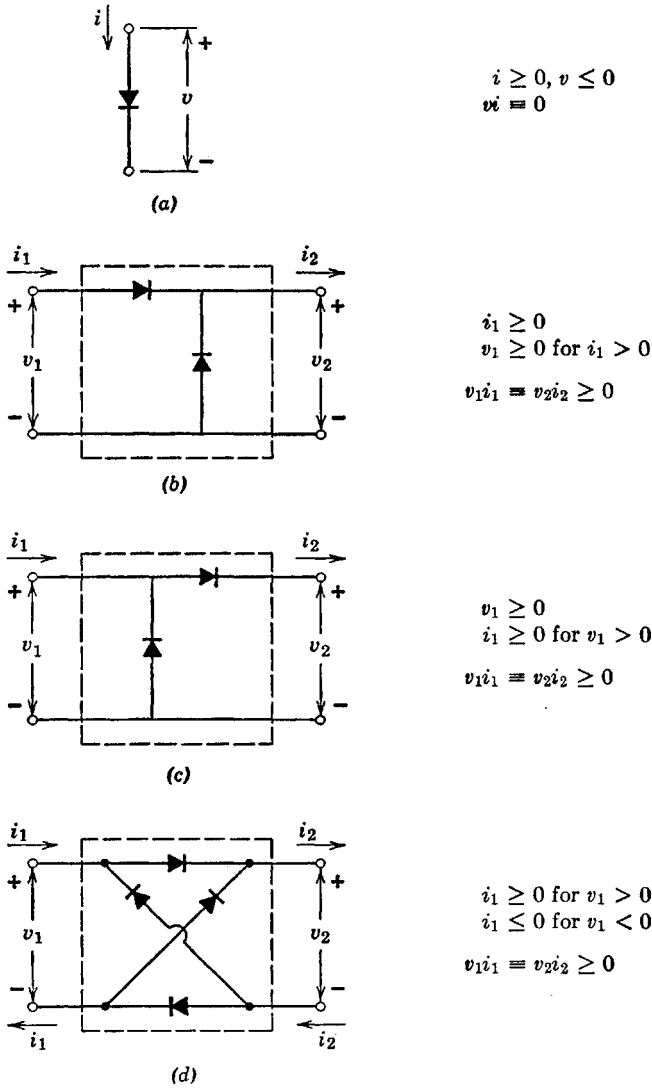
Fig. 7.2. Ideal valves. (a) The ideal diode valve, into which no energy flows. (b), (c), (d) Energy valves, through which energy can flow in only *one* direction.

from it for any nonzero V by proper adjustment of an applied voltage v or current i.

Insertion of an energy valve, as shown in Fig. 7.3(c), assures passivity without disturbing the character of the valve curves within the first

quadrant of the i vs. v plane. The energy valve introduces the non-linearity indicated in (d) and makes the second and fourth (nonpassive) quadrants forbidden ground. With V replaced by the negative of μe_c,
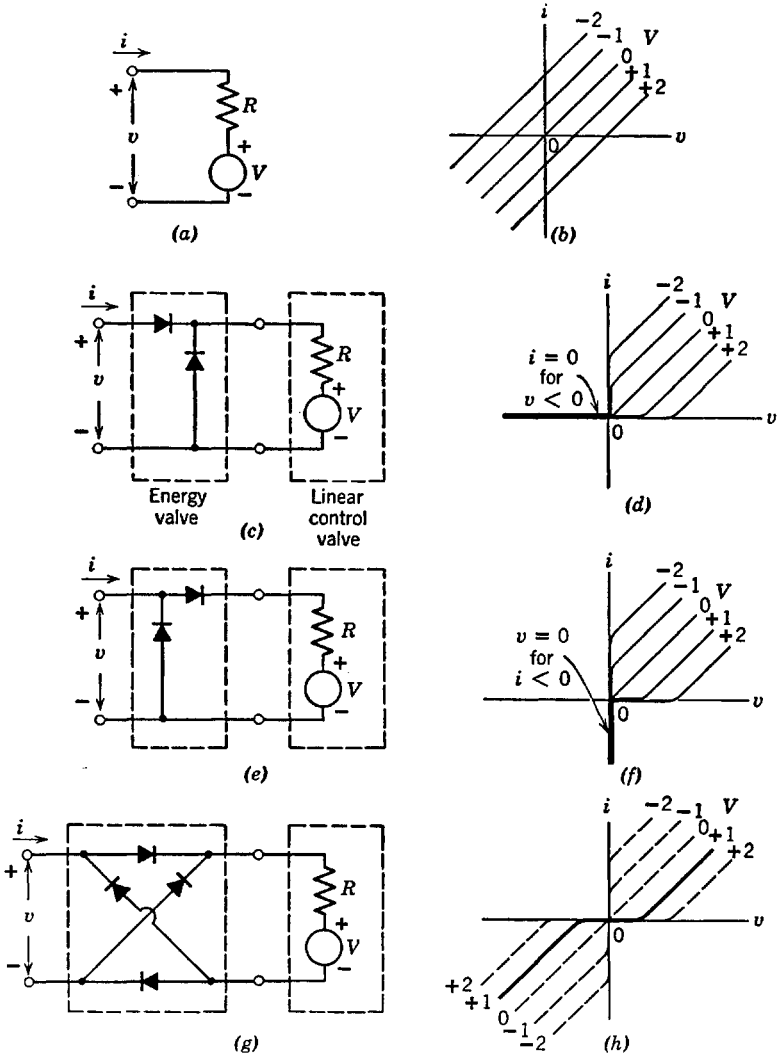
Fig. 7.3. Ideal control valves and associated curves.

where e_c is the grid control voltage of a vacuum triode, we recognize Fig. 7.3(d) as the piecewise-linear approximation to the plate curves of that triode [see Fig. 6.10(d)].

Two added possibilities, corresponding to the other basic forms of the energy valve, are shown in Fig. 7.3(e) and (g), and the associated curves are given in (f) and (h). The main point here is that resolution of the circuit model into two parts, each having an easily interpretable function, makes the complete circuit appear much less formidable and also suggests the generality of such models for the representation of a wide variety of devices.

7.5 The Pentode

From generalities we now shift to the discussion of some specific electronic devices. The first of these is the pentode, a triode with two additional grids. Development of the pentode arose from the desire to overcome certain disadvantages inherent in the vacuum triode. Figure 7.4(a) shows a simple triode amplifier with the grid-to-plate inter-electrode capacitance C placed in evidence as a circuit element. At high frequencies the susceptance of this capacitance becomes appreciable.

(a)   (b)   (c)
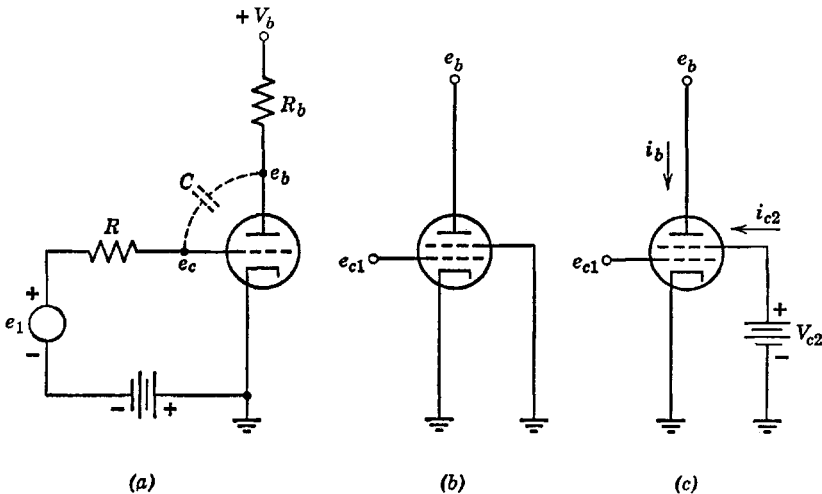
Fig. 7.4. Transition from triode to tetrode. (a) Grid-to-plate interelectrode capacitance which reduces triode gain at high frequencies; (b) Introduction of a screen grid to shield the plate from the control grid; (c) Addition of a screen-grid supply voltage to maintain the flow of plate current.

Signal current flowing in the capacitor must also flow through the source resistance R and the resulting voltage drop from e_1 to e_c reduces the voltage amplification between e_1 and the output e_b. The loss of amplification is aggravated by the so-called "Miller effect." The essence of

the Miller effect is as follows. For one volt of a-c signal at the grid a considerably larger a-c signal appears at the plate, say A volts, and the polarity of the plate signal is opposite to that of the grid signal. Hence the voltage across the capacitance C is larger than the grid signal by a factor $1 + A$ and the capacitive current is therefore $(1 + A)$ times greater than it would be with the same capacitance connected between grid and cathode. The capacitance C therefore has an effect comparable to that of a larger capacitance C $(1 + A)$ between grid and cathode. Thus we see that grid-to-plate capacitance severely limits the operation of a vacuum triode as a high-frequency voltage amplifier.

The capacitance between grid and plate can be greatly reduced by the insertion of an electrostatic shield or screen between the grid and the plate as indicated in Fig. 7.4(b). This shield cannot, of course, be a solid sheet, for then no current could flow between cathode and plate. Hence a perforated screen is used and this is called the screen grid. If we experiment upon the circuit shown in Fig. 7.4(b), we find that the capacitance between the control grid and the plate is indeed very small, but that hardly any plate current flows unless the plate voltage is raised to a tremendously high value. The reason is that the grounded screen grid not only shields the plate from the control grid but also further shields the plate from the cathode, thereby eliminating the motive force for normal current flow in the tube. The remedy is shown in Fig. 7.4(c). Here the screen grid is held at some fixed positive potential which produces much the same electric field in the region between the screen grid and the cathode as that in a triode. The cathode, control grid, and screen grid *are* in fact a triode; a triode with holes in its "plate." The large majority of electrons pass through the openings between the screen-grid wires into the drift space beyond and are eventually collected by the true plate of the tube.

The tube in Fig. 7.4(c) has four electrodes and is therefore called a tetrode. In examining its properties we shall find reasons for the introduction of still another grid and the tube will then become a pentode. Figure 7.5(a) shows the variation of electric potential between the cathode and plate of a plane-parallel tetrode. Conditions within the "electron gun" portion of the tetrode are much the same as those in a triode. The "gun" shoots electrons through the drift space and into the plate. The "muzzle velocity" of the electron gun is mainly determined by the screen-grid voltage V_{c2} whereas the number of bullets per second (the cathode current) is mainly fixed by the control grid voltage. For the purposes of this discussion we shall hold the control grid at zero voltage.

Because of the electrostatic shielding effect of the two grids, the

(a)

(b)
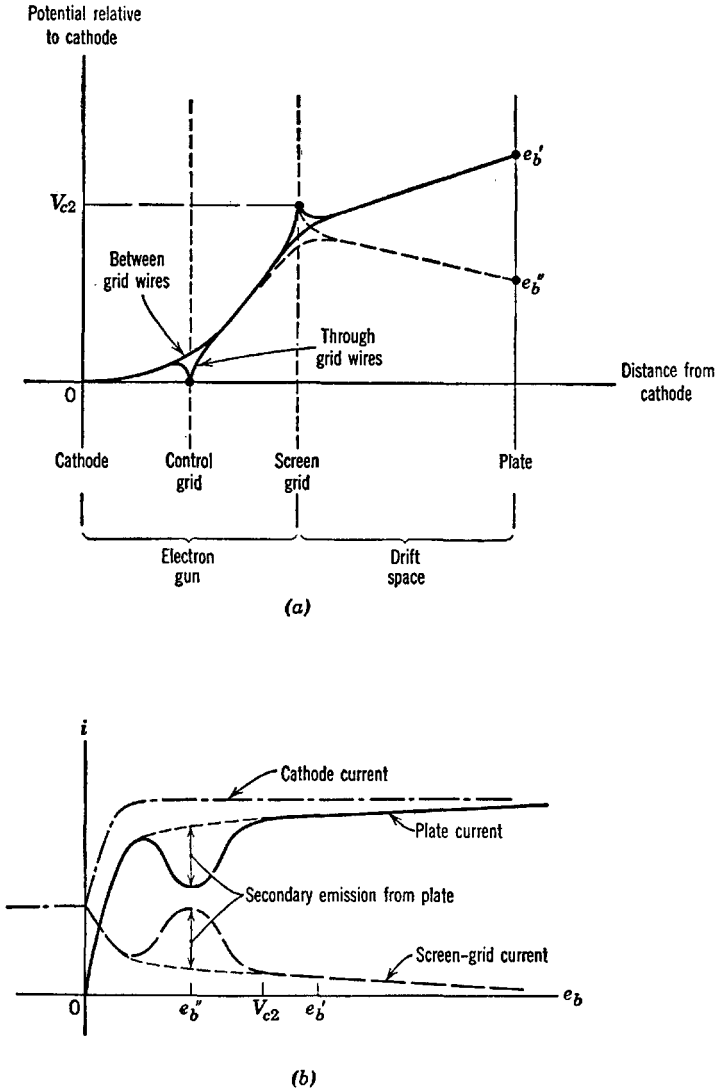
Fig. 7.5. Tetrode curves. (*a*) Electric potential in a plane-parallel tetrode; (*b*) Current versus plate voltage for $e_{c1} = 0$ and $e_{c2} = V_{c2} =$ constant.

cathode current is influenced very little by variations in the plate voltage, as indicated by the cathode-current curve in Fig. 7.5(*b*). The screen-grid current decreases slightly with increasing plate voltage because of minor modifications in the electric field pattern near the screen-grid

wires. When the plate beckons more strongly, the electrons are not quite as likely to be captured by a screen-grid wire. Hence, plate current increases slightly with increasing plate voltage, screen-grid current decreases, and cathode current remains essentially unchanged, as shown in Fig. 7.5(b).

The kinetic energy of an electron is proportional to the height of the potential curve in Fig. 7.5(a), hence the electrons strike the plate with considerable speed. When an electron enters the plate, it can "splash"
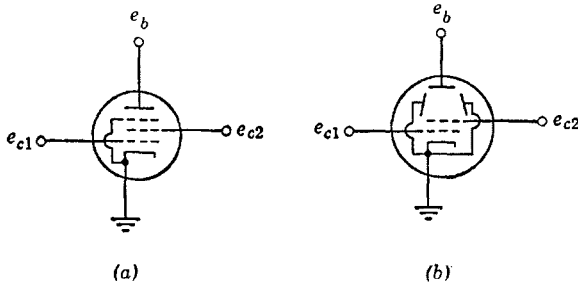
Fig. 7.6. Pentode symbols. (a) Pentode, with suppressor grid internally connected to the cathode; (b) Beam tetrode, with beam-forming electrodes internally connected to the cathode.

several other electrons out of the plate. Such liberation of electrons by bombardment is called secondary emission. The secondary electrons are emitted with less kinetic energy than that of the impinging electron but this energy frees the secondaries from the attraction of the plate. When the plate voltage is higher than the screen-grid voltage, as in the solid curve of Fig. 7.5(a), the secondary electrons find themselves in an electric field which quickly returns them to the plate. However, if the plate voltage is reduced from e_b' to a lower value e_b'', the secondary electrons ride upward along the potential curve and are eventually collected by the screen grid. The result is a dip in the plate-current curve as shown in Fig. 7.5(b). At still lower plate voltages the electrons are greatly decelerated in the drift space and arrive at the plate with insufficient energy to produce appreciable secondary emission. Finally, as the plate voltage is reduced to zero, the plate current is cut off and all "bullets" fall back into the "muzzle" of the "gun." As the electrons are decelerated and turned back from the plate, they return to the screen grid at high velocity and may oscillate through the screen many times before being collected. The electron cloud produces a relatively heavy negative space charge that depresses the potential curve and reduces the cathode current as indicated in Fig. 7.5(b).

The effects of secondary emission can be suppressed by the introduction of a third grid, called the suppressor grid, between the screen grid
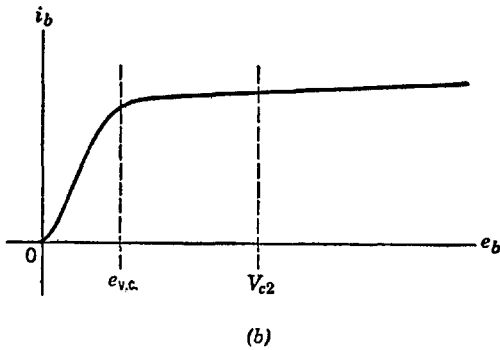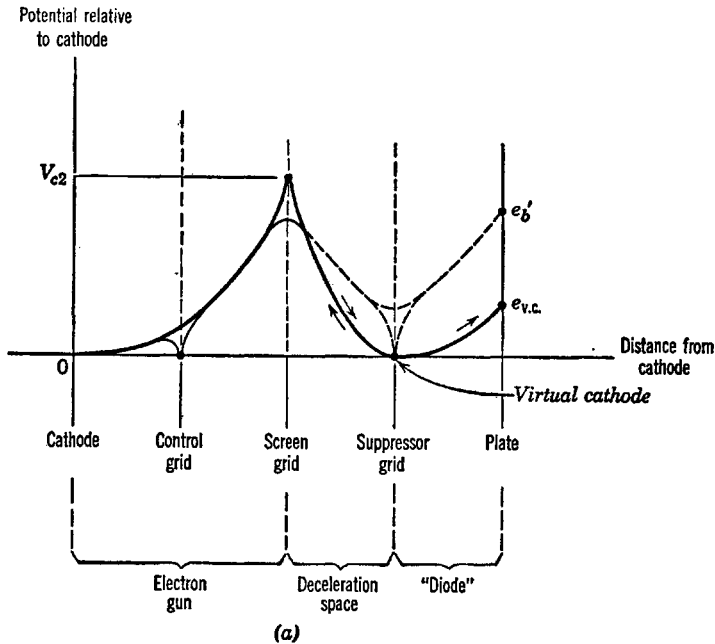
Fig. 7.7. Pentode curves. (*a*) Formation of a virtual cathode near the suppressor grid at low plate voltage; (*b*) Pentode plate curve for $e_{c1} = 0$ and $e_{c2} = V_{c2}$.

and the plate as shown in Fig. 7.6 (*a*). In most pentodes the suppressor is internally connected to the cathode. The induced negative charge on

the suppressor grid wires depresses the potential curve as indicated by the broken line in Fig. 7.7(a) and thereby prevents secondary electrons from travelling to the screen grid. At the same time the suppressor grid does not interfere with the passage of high-energy electrons from the gun. Unless one of these electrons happens to be aimed directly at the center of a suppressor-grid wire, it will reach the suppressor-grid plane with sufficient velocity to carry it between the wires and onward to the plate.

The beam tetrode shown in Fig. 7.6(b) accomplishes the same result in a different way. Beam-forming electrodes are located in such a position that their induced negative charge focusses the electron stream into a relatively compact beam. This increases the space-charge density within the electron stream and therefore depresses the potential curve. Also, by aligning screen and control-grid wires, electrons are directed between screen wires, thereby reducing screen current.

Figure 7.7(b) shows the pentode plate curve for zero control-grid voltage. Just as in the tetrode the plate current decreases slightly with decreasing plate voltage, but the secondary-emission dip is no longer present. When the plate voltage drops to some critical value $e_{v.c.}$ the plate current begins to drop much more sharply. The onset of this break point can be explained as follows. As e_b decreases, electrons are decelerated more drastically in the drift space between the screen grid and the plate. Deceleration means that the electrons bunch more closely together, with an attendant increase in their space-charge density. This increase in space-charge density further depresses the potential curve which, in turn, enhances the deceleration. The effects are cumulative, and, as e_b decreases, a point is reached at which the potential curve drops to zero in the neighborhood of the suppressor grid. Under these conditions the electrons arrive at the suppressor grid with zero velocity, forming a local reservoir of stationary electrons. Some of them fall backward to the screen and some fall toward the plate. The situation is really the same as that which exists just outside a hot cathode under conditions of space-charge-limited current flow. Some of the electrons are returned to the cathode and some proceed onward past the potential minimum. A reservoir of stationary electrons at some surface in space is equivalent to an ideal cathode located on that surface. When the potential within the pentode drops to zero, we say that a *virtual cathode* has been formed at the suppressor grid. The region between the suppressor and the plate then behaves much like a high-vacuum diode. For plate voltages below the critical value $e_{v.c.}$ the diode is unsaturated; that is, the diode demands less current than that supplied by the electron gun. For values of plate voltage above the critical value, the diode is

saturated. All of the "emission current" available at the virtual cathode is exhausted and the "diode curve" levels off, as shown in Fig. 7.7(b).

7.6 Pentode Curves and Pentode Circuit Models

Since the electron-gun portion of the pentode behaves like a triode, we might expect the plate current to be proportional to the three-halves

(a)

(b)
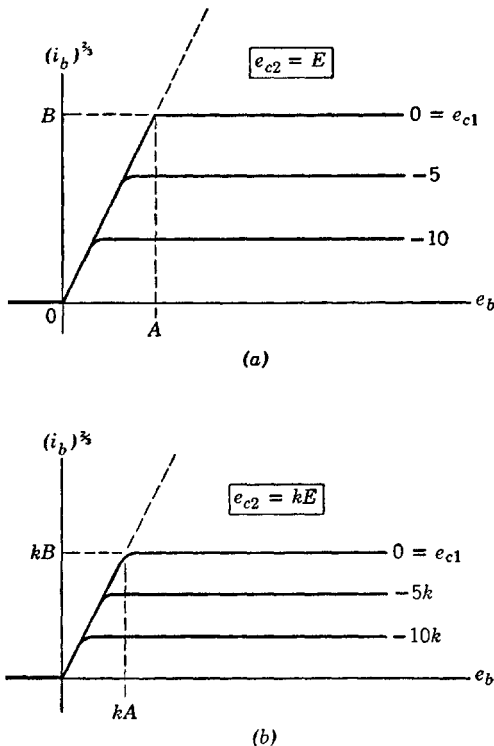
Fig. 7.8. Idealized pentode curves, showing the effect of a change in screen-grid potential. Note that $i_b{}^{2/3}$ is the scale, not i_b.

power of the quantity $(e_{c2} + \mu' e_{c1})$, where μ' is the amplification factor of the electron gun. This is a fair approximation, and it leads to the idealized pentode curves plotted in Fig. 7.8 with the two-thirds power of i_b, rather than i_b itself, as the vertical co-ordinate. The idealized curves are horizontal and equally spaced. In the virtual-cathode region the curves all join on a line of slope B/A, representing the three-halves

power law for unsaturated operation of the diode portion of the pentode. Comparison of Fig. 7.8(a) and (b) shows the effect of a change in the screen-grid supply voltage. To convert a set of pentode curves from
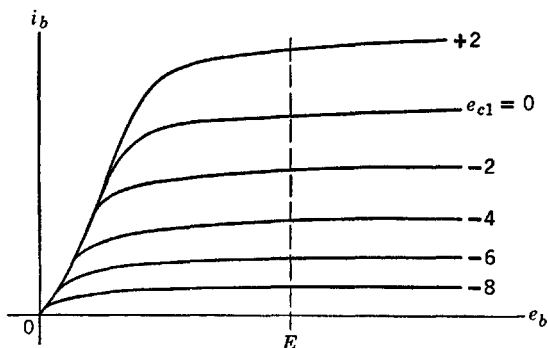
Fig. 7.9. A typical family of pentode plate curves ($e_{c2} = E$ = constant).

one fixed value of screen-grid voltage to another, simply multiply each of the voltage scales by the scale factor k and multiply the current scale by the three-halves power of the scale factor k.

Figure 7.9 shows a family of pentode plate curves for some fixed screen grid voltage. Here the first power of the current is plotted and the curves therefore exhibit a crowding at low values of current. In the central region of the curves the control-grid-to-plate transconductance g_{1p} is approximately proportional to the cube root of the plate current, for if $i_b^{2/3} \sim (e_{c2} + \mu_{21}e_{c1})$, then $g_{1p} = \partial i_b/\partial e_{c1} \sim (i_b)^{1/3}$.

Figure 7.10 shows a typical pentode amplifier circuit. The by-pass capacitors C_k and C_{c2} are assumed to have negligible reactance at all signal frequencies of interest. Hence the cathode and screen-grid potentials are essentially constant. In such a circuit the pentode operates as a good high-frequency voltage amplifier. The pentode has one disadvantage not present in the triode. The cathode current divides into two portions; a small fraction to the screen grid and the remainder to the plate. The slight inherent random fluctuation in this division ratio produces an additional fluctuation in plate current which is not present in the cathode current. This extra component of plate current is called "partition noise" or "division noise" and, other things being equal, it makes the pentode slightly noisier than the triode. Nevertheless the pentode has found wide use in high-frequency applications where extreme sensitivity is not required.

In some special purpose pentodes, called "gate" tubes, the suppressor-

grid terminal is externally available as shown in Fig. 7.11(a). If the suppressor-grid voltage is made several volts negative with respect to the cathode, the interelectrode potential between screen-grid and plate
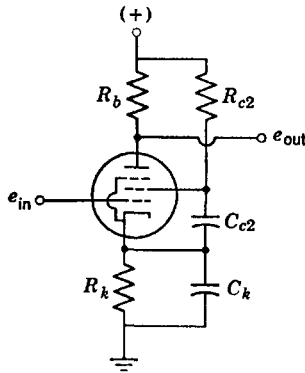
Fig. 7.10. A pentode amplifier operating from a single supply voltage.

(a)   *(b)*
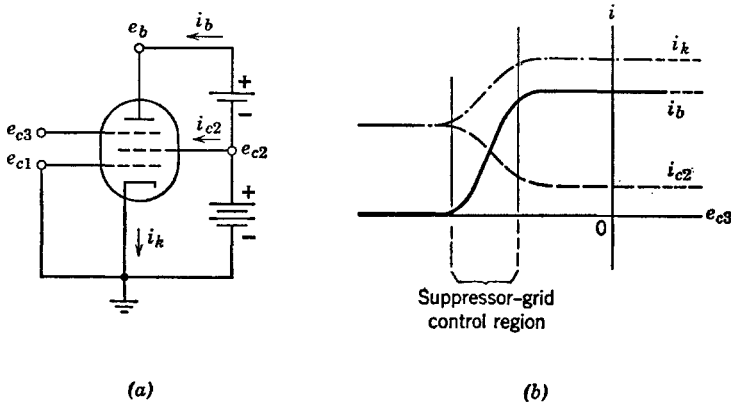
Fig. 7.11. Gate circuit. (a) A "gate" tube, having an externally available suppressor-grid terminal; (b) Control action at negative values of suppressor-grid voltage.

comes down like a curtain and begins to cut off the flow of plate current. Hence a region exists within which the suppressor grid acts like a control grid, as illustrated in Fig. 7.11(b). A sufficiently large negative signal at the suppressor grid can therefore be utilized to "close the gate" and interrupt the flow of signals from the control grid to the plate in a pentode amplifier.

Figure 7.12(a) shows a piecewise-linear model which approximates

the behavior of the pentode plate circuit. We assume here that the suppressor grid is connected to the cathode. The associated plate curves are shown in Fig. 7.12(c). If parameter g is allowed to approach infinity, the curves occupy the entire first quadrant of the i_b vs. e_b plane, and the circuit model becomes equivalent to the one shown in Fig. 7.3(c). For

(a)   (b)

(c)
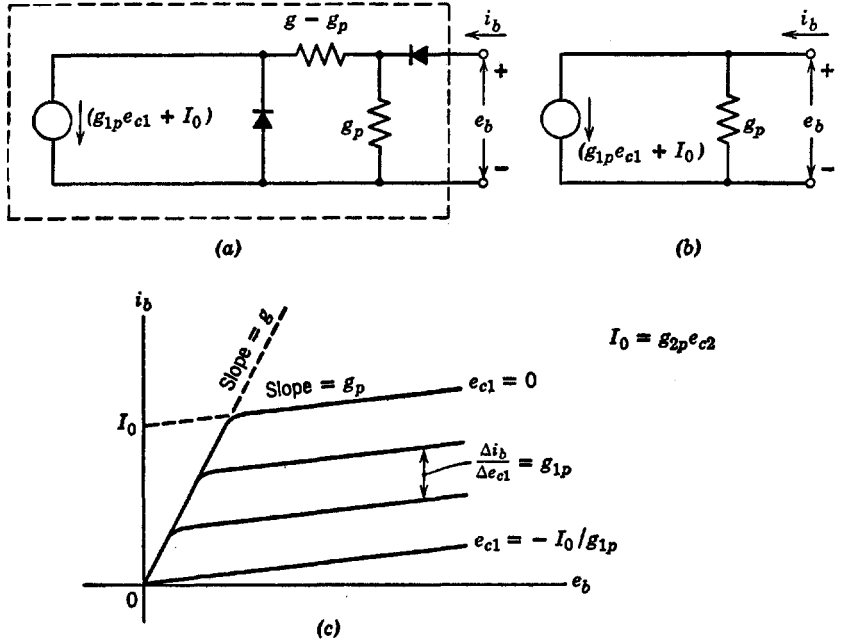
Fig. 7.12. Pentode model and curves. (a) A piecewise-linear model for the pentode plate circuit ($e_{c3} = 0$, e_{c2} = constant); (b) The linear model, applicable in the central region of the plate curves, to which (a) reduces with the shunt diode open and the series diode closed; (c) The associated piecewise-linear plate-current curves.

operation in the central region of the plate curves, a suitable linear model can be obtained from Fig. 7.12(a) by short-circuiting the series diode and open-circuiting the shunt diode. Also, the conductance $(g - g_p)$ may be ignored in this case, since it is in series with a current source. In short, the linear model for the plate circuit of a pentode amplifier consists of a transconductance current $g_{1p}e_{c1} + I_0$ in parallel with a conductance g_p. This is the same as the current-source form of the triode model, except for the fact that g_p is considerably smaller in the pentode than in the triode, and except for the additional term I_0. Current I_0 is determined largely by the screen-grid voltage e_{c2}. For a gross approxi-

mation to this effect we can write $I_0 = g_{2p}e_{c2}$, where g_{2p} is the screen-grid-to-plate transconductance of the model, a constant chosen to approximate the actual tube data over a range of e_{c2}.

7.7 The Cryotron

The cryotron, shown in Fig. 7.13, is an ideal control valve which depends for its operation upon the phenomenon of superconductivity. As the temperature of a material is decreased, its resistance in general
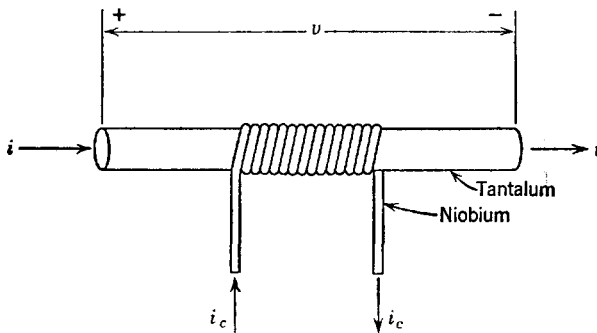
Fig. 7.13. The cryotron.

also decreases. For a certain class of materials called superconductors, the curve of resistivity versus absolute temperature levels off at a few degrees above absolute zero and then drops sharply and completely to zero, as indicated in Fig. 7.14. In the superconducting region lying below the critical temperature K_c the resistivity is not merely small but actually zero insofar as the most advanced and refined measurements can show. Moreover, the application of a magnetic field H to the material alters the critical temperature, and it is upon this effect that the operation of the cryotron is based.

A cryotron can be constructed by simply winding a "control" coil of insulated niobium wire upon a short length of tantalum wire. Let us now enclose the cryotron in a container surrounded by liquid helium which reduces its temperature below the critical value for both tantalum and niobium. If we pass a current through the tantalum "valve" wire (by means of leads brought out through the mouth of the container), the voltage drop across the valve will remain at the value zero until the current is raised to some critical value. At this point the self-induced magnetic field within the wire reaches the superconductivity limit for

that temperature and the superconductivity is suddenly destroyed. What happens is that one small portion of the wire near the surface first becomes resistive. This distorts the lines of current flow and increases the magnetic field nearby. The result is a self-propagating wave of increased resistivity which quickly destroys superconductivity throughout the short piece of valve wire. Let us assume that the short length of tantalum valve wire is electrically connected to the remainder of the system through leads of niobium hook-up wire. The hook-up wire
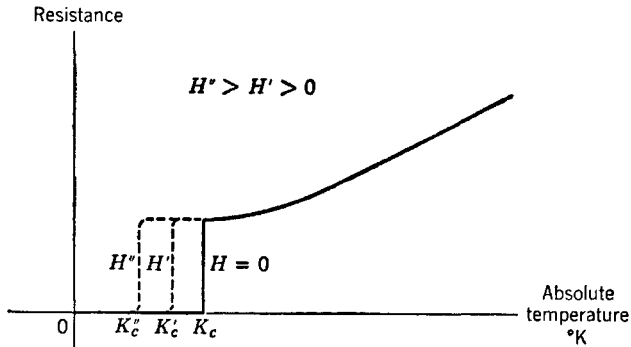
Fig. 7.14. Resistance of a piece of wire versus absolute temperature °K, showing superconductivity below some critical temperature, K_c. A magnetic field H within the wire reduces the critical temperature.

and the control winding will remain superconducting since niobium has a critical temperature roughly twice that of tantalum. This gives us the zero-control-current curve shown in Fig. 7.15. Let us now return the valve current to zero and repeat the experiment with a nonzero value of the control current i_c. For a nonzero value of control current, the control winding produces an initial magnetic field within the valve wire. As the valve current i is again increased from zero, the critical magnetic field will be reached at a smaller value of valve current. The two components of magnetic field caused by the two currents i and i_c are at right angles to each other and the critical magnetic field of which we speak is the vector sum of these two components. Hence we may expect to find a quadratic relationship between the critical valve current I and the control current i_c. This relationship is given in Fig. 7.15. A straightforward calculation of the magnetic field due to the two currents yields the values of the constants I_0 and i_0. The values are $\pi d H_0$ and H_0/n, respectively, where H_0 is the critical value of H for tantalum at the experimental temperature, d is the diameter of the valve wire, and n is

the number of turns per unit length in the control winding. The constant i_0 is that value of control current required to destroy superconductivity with no valve current flowing, and the constant I_0 is the critical value of valve current when no control current is flowing.

Because superconducting niobium can tolerate a much higher magnetic field than superconducting tantalum, the control winding remains
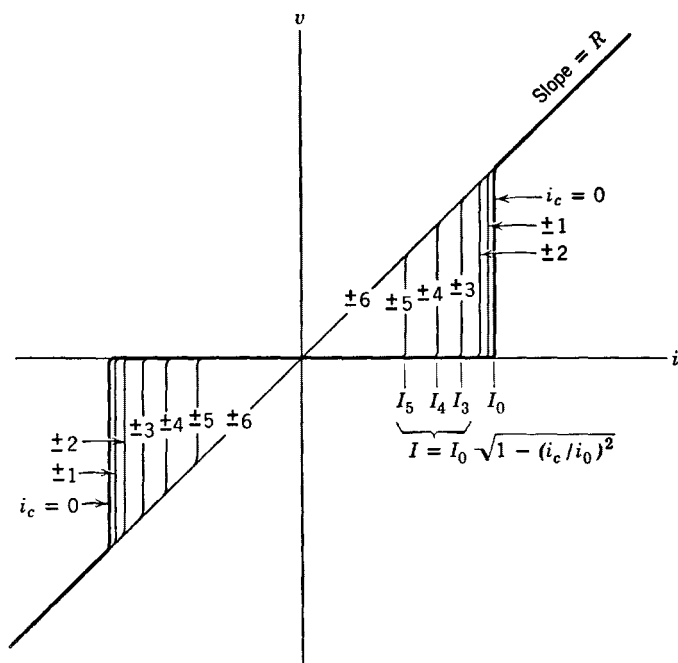
Fig. 7.15. Cryotron valve curves.

superconducting throughout the ranges of operation indicated in Fig. 7.15. The cryotron therefore is truly an ideal unilateral control valve.

Economical modern methods of producing liquid helium in large quantities make practical cryotron systems a reality. Cryotrons are very small and hundreds of them can be connected together to form a complicated valve system such as a computer network. The entire network can then be enclosed in a relatively small container, with the cost of maintaining the helium no worse than the cost of air-conditioning a similar computer made up of semiconductor or vacuum tube components. At present, the main drawbacks of cryotron systems are that the cryotron cannot be switched or pulsed as rapidly as some other

electronic devices and that a cryotron system requiring many different access leads to the outside world presents some problems of heat insulation.
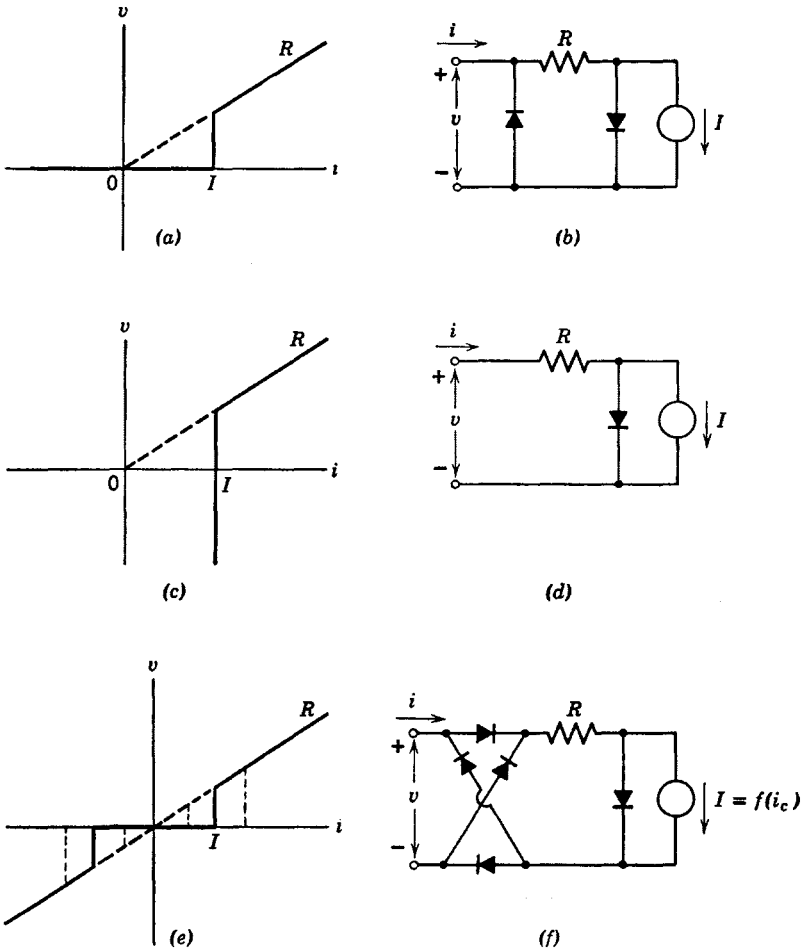
Fig. 7.16. Construction of a piecewise-linear model for the cryotron.

For a circuit model of the cryotron let us begin with the curve shown in Fig. 7.16(a), which can be synthesized as the circuit indicated in (b). Connection of two such circuits in series opposition yields a single symmetric cryotron curve. It is interesting to observe that the circuit in Fig. 7.16(b) is a very close cousin of that in Fig. 7.3(e). If the current

source I were connected across the resistance R instead of across the diode, the two circuits would be equivalent. We really should not expect them to be twins because the cryotron curves do not sweep out the entire passive first and third quadrants of the valve plane.

An alternative synthesis of the circuit model, which utilizes more diodes but only one controlled current source I, begins with the curve of Fig. 7.16(c) and the associated elementary circuit indicated in (d). Connection of this circuit to the bridge-type energy valve yields a satisfactory cryotron model (f) which has the representative curve shown in (e).

7.8   The Thyratron

The thyratron is a hot-cathode gas triode which is used as a switch or relay. In preparation for the discussion of the thyratron, Fig. 7.17 re-iterates the characteristics of a simple gas diode circuit. The black dot in the circuit symbol for the diode represents a gas molecule and indicates that the gas pressure, though much smaller than atmospheric pressure, is nevertheless much greater than that in a good vacuum diode. The applied voltage e_b' must be connected to the diode through some load resistance R_b in order to limit the value of plate current when the tube conducts. The presence of resistance R_b in Fig. 7.17(a) also permits us to describe the operation of the circuit in terms of the applied voltage variable e_b' alone, as illustrated in Fig. 7.17(b). If the tube is not already conducting, then it fires upon crossing *from* region B *into* region C; and if it is not already nonconducting, then it extinguishes upon crossing *from* region B *into* region A. In the memory region B the tube retains either state A or state C, whichever of the two was experienced most recently.

The explanation is as follows. With no gas in the tube the diode curve has a shape indicated by the dashed curve in Fig. 7.17(c). When gas is introduced, any plate current larger than some critical value i_{b0} produces significant ionization of the gas molecules. These positive ions are heavy and remain in the interelectrode space for a relatively long time, thereby producing a heavy positive space charge which helps the plate voltage draw electrons out of the cathode. The hysteresis or "memory" exhibited by the circuit becomes apparent when we plot the applied voltage e_b', rather than the plate voltage e_b, as shown in Fig. 7.17(d). Increasing the applied voltage from zero, we reach point a, at which the tube fires and the operating point jumps to b. A decrease in the applied voltage then brings us to point c, whence a slight further

decrease causes the tube to extinguish and the operating point returns to the low-current portion of the curve at point d.

The basic thyratron circuit is depicted in Fig. 7.18(a). With no gas in the tube we should have the ordinary vacuum-triode curves shown in (d). For values of plate current less than the critical ionization
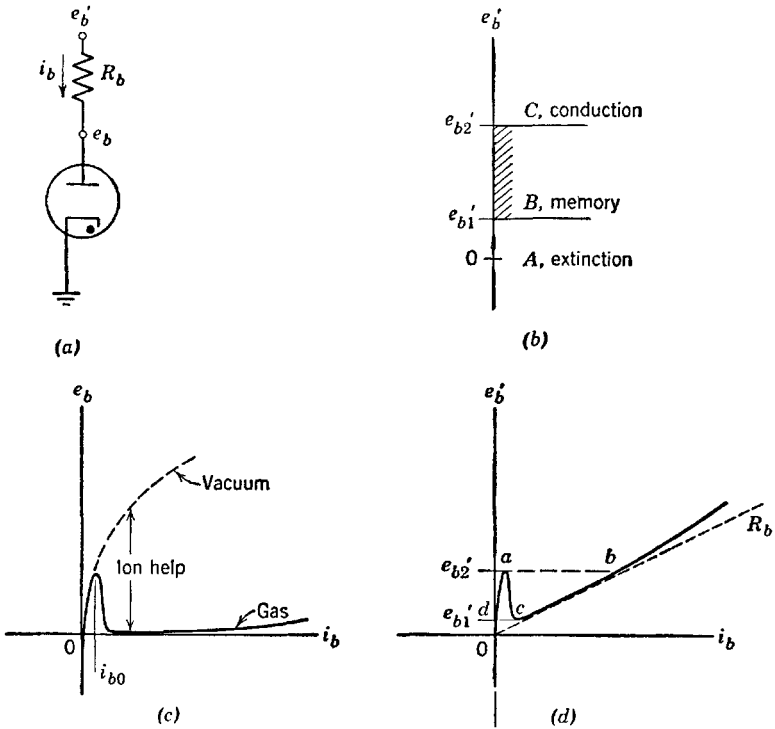
Fig. 7.17. Gas diode properties. (a) Gas diode; (b) State-transition voltages; (c) The diode curve; (d) The curve of applied voltage, showing transition thresholds.

current i_{b0} the vacuum-triode curves (d) and the thyratron curves (c) are very much the same. However, once there is appreciable ionization, the thyratron curves drop rapidly to a very low value of plate voltage. Once the thyratron fires or conducts, a sheet of positive ions forms around the grid wires and the grid therefore losses control—a control which cannot be regained until the thyratron is once more extinguished by lowering the applied plate voltage e_b' to a sufficiently small value. Figure 7.18(b) shows the three regions of operation in the plane of

applied plate voltage e_b' versus applied control voltage e_c'. For negative values of e_c', a larger e_b' is required in order to fire the tube. Just as before, the tube remains nonconducting until the operating point crosses the boundary from region B into region C and then remains conducting
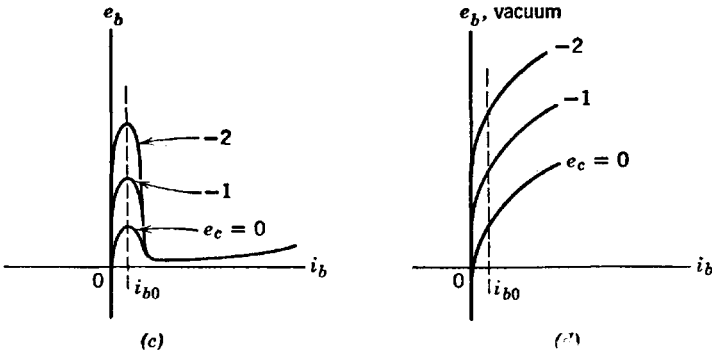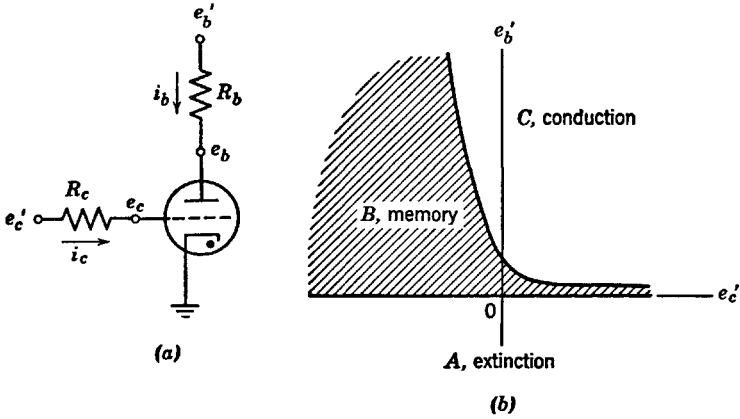
Fig. 7.18. Thyratron properties. (a) The thyratron; (b) State-transition curves in the plane of applied voltages; (c) Thyratron curves; (d) Vacuum-triode curves, for comparison.

until the operating point next crosses from region B into region A. Thus the thyratron is a self-holding relay, its advantage over similar electromechanical relays being its speed of operation.

Parts (a) and (b) of Fig. 7.19 show two possible combinations of

applied voltage with which thyratrons are sometimes operated. In each case the locus of operation is indicated by a dashed line. In Fig. 7.19(a) the tube fires at point p and extinguishes at point q, hence the conduction angle θ can be changed by varying the d-c component E of the

$$e_b' = V_b \cos \omega t$$
$$e_c' = + V_c \sin \omega t - E$$

(a)

$$e_b' = V_b \cos \omega t$$
$$e_c' = + V_c \cos \omega t - E$$

(b)

(c)
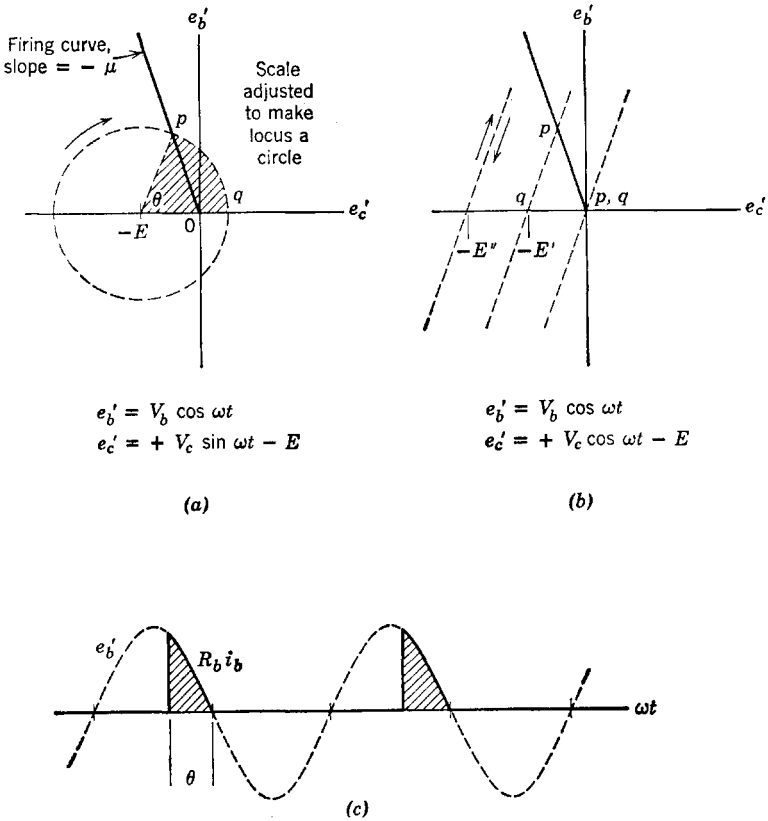
Fig. 7.19. (a), (b) Loci of operation in the applied-voltage plane; (c) Waveform of plate-load voltage $R_b i_b$, showing the conduction angle θ.

control signal e_c'. In this manner the average power delivered to the load R_b is adjustable by varying the control signal E.

When the two applied a-c signals are in phase, the locus of operation is a straight line as indicated in Fig. 7.19(b). However, the tube still fires at point p and then extinguishes upon next reaching point q, so that the conduction angle is still under the influence of the control signal E. For zero E in Fig. 7.19(b), conduction takes place over the complete

positive half cycle of e_b' and maximum power is delivered to the load. As E becomes larger in Fig. 7.19(b), the conduction angle decreases to one quarter cycle and then drops suddenly to zero when the locus of operation no longer touches the firing curve. The scheme in Fig. 7.19(a) has the advantage of continuous control of the conduction angle through all values from zero to π.

Figure 7.19(c) shows the waveform of plate current for a resistive load R_b. Thyratrons are often used for speed control of small d-c motors, in which case the load (the motor winding) is not representable as a pure resistance. Nevertheless, the signal at the thyratron grid still effects wide control over the time-average current supplied to the motor.

SUPPLEMENTARY READING

L. B. Arguimbau, R. B. Adler, *Vacuum Tube Circuits and Transistors*, John Wiley and Sons, New York, 1956.

H. Bruining, *Physics and Applications of Secondary Electron Emission*, McGraw-Hill, New York, 1954.

W. G. Dow, *Fundamentals of Engineering Electronics*, John Wiley and Sons, New York, 1952.

D. V. Geppert, *Basic Electron Tubes*, McGraw-Hill, New York, 1951.

T. S. Gray, *Applied Electronics*, 2nd edition, John Wiley and Sons, New York, 1954.

Willis W. Harman, *Fundamentals of Electronic Motion*, McGraw-Hill, New York, 1953.

Samuel Seely, *Electron-Tube Circuits*, McGraw-Hill, New York, 1950.

Samuel Seely, *Electronic Engineering*, McGraw-Hill, New York, 1956.

Karl R. Spangenberg, *Fundamentals of Electron Devices*, McGraw-Hill, New York, 1957.

Robert L. Sproull, *Modern Physics: A Textbook for Engineers*, John Wiley and Sons, New York, 1956.

G. E. Valley and H. Wallman, *Vacuum Tube Amplifiers*, McGraw-Hill, New York, 1948.

PROBLEMS (See Appendix C for Pentode Curves)

7.1. If we imagine that each electron in the interelectrode space of a pentode is replaced by two electrons, then (by superposition) the electric potential is everywhere doubled and, since the electrons are moving freely under the action of the electric field, their kinetic energy at a chosen point in the tube must also double. Hence electron velocity (at a given point on the electron path) is multiplied by the square root of two. Now, in the light of the above discussion, what can you say about the change in the electrode currents of a pentode if all electrode potentials, relative to the cathode, are doubled?

7.2. Pentodes A and B are identical but B is operated at twice the screen-grid voltage of A. Sketch and compare their plate curves. (The suppressor grid is internally connected to the cathode.)

7.3. (a) Sketch a pentode curve of i_b vs. e_{c1} for fixed e_b and e_{c2}. (e_{c3} is zero.) Use a value of e_b only slightly to the right of the knee of the zero-control-grid-voltage plate curve. Carry your sketch far enough positive and negative in e_{c1} to show three distinctly different portions of the curve.

(b) The central portion may be closely approximated by a three-halves power law $i_b = K_1(e_{c1} + k_2)^{3/2}$. Why?

(c) How do you account for the two outer portions of the curve?

(d) How does the incremental transconductance g_m (the slope of the curve) vary with i_b in the central portion?

(e) If i_b reaches cutoff when e_{c1} is about 10 volts negative and if i_b equals 30 ma when e_{c1} is zero, what is the incremental transconductance at 20 ma? At 2.5 ma?

7.4. Suppose that the control grid, suppressor grid, and cathode of a pentode are grounded, the screen grid is held at a fixed positive potential, and the plate voltage is slowly increased, starting with a value several volts negative. Sketch the curve of i_b vs. e_b and explain its shape. For each of the three principal regions of this curve, describe the motion of an electron after it leaves the cathode.

7.5. The plate current of a pentode with the suppressor connected to the cathode is a function of three voltages $i_b = f(e_{c1}, e_{c2}, e_b)$. The general three-halves power law states that if we change these voltages in proportion, the current will vary as the three-halves power of any of the voltages. In short, if $i_b' = f(e_{c1}', e_{c2}', e_b')$, and if any three of the four ratios

$$\frac{e_{c1}'}{e_{c1}} \; ; \; \frac{e_{c2}'}{e_{c2}} \; ; \; \frac{e_b'}{e_b} \; ; \; \left(\frac{i_b'}{i_b}\right)^{2/3}$$

are equal, then the fourth equals the other three.

(a) Using Type II pentode plate curves measured at $e_{c2} = 250$ volts, determine i_b, r_p, and g_m at the operating point $e_{c1} = -6$, $e_{c2} = 150$, $e_b = 180$ volts.

Fig. P7.1

(b) Give directions for relabeling the i_b, e_b, and e_{c1} scales of a set of pentode plate curves in order to make the same set of curves applicable at a new fixed e_{c2}.

7.6. Two identical triodes are to be used in the circuit shown in Fig. P7.1 to replace a pentode. (This is often called a "cascode" circuit.) For purpose of design it is desirable to know the approximate composite characteristics of this circuit.
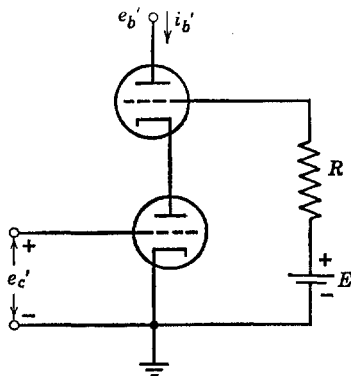
(a) Sketch and dimension the composite curves e_b' vs. i_b', with e_c' as a parameter, using piecewise-linear models for the triodes. Use idealized curves for the individual triodes with $\mu = 50$, $r_p = 40$ kilohms.

(b) For the amplifier circuit shown in Fig. P7.2, what is the incremental voltage gain $\Delta e_2/\Delta e_1$? Be sure to justify your assumptions as to the conditions of the diodes in the piecewise-linear models. The by-pass capacitor C is a short circuit for frequencies of interest.
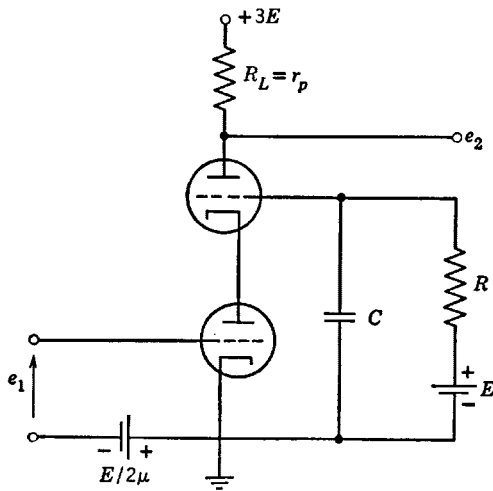
Fig. P7.2

7.7. (a) Using the plate characteristics for the Type I pentode, determine the values and proper algebraic signs of the constants I_0, g_m, and r_p such that the linear equivalent circuit of Fig. P7.3 closely represents the actual tube in the neighborhood of the following operating point:

$$e_b = 125 \text{ volts}, \ e_{c1} = -1 \text{ volt}, \ e_{c2} = 100 \text{ volts}, \ e_{c3} = 0$$
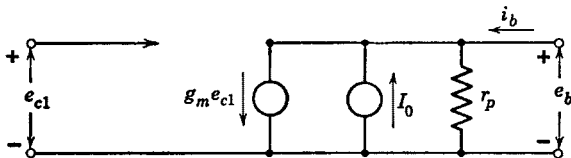
Fig. P7.3

the last two being fixed voltages. Be sure to specify proper units, as well as the numerical values and algebraic signs.

(b) Repeat part (a) at the operating point:

$$e_b = 187.5 \text{ volts}, \ e_{c1} = -1.5 \text{ volts}, \ e_{c2} = 150 \text{ volts}, \ e_{c3} = 0$$

the last two being fixed voltages. Use the generalized three-halves power law.

7.8. In the circuit of Fig. P7.4, quantity A is a positive constant determining the dependent current source Ai_0 in terms of current i_0, and quantity I_2 is a fixed positive supply current.

(a) Sketch and dimension the curve of i_2 vs. i_1 for the region in which i_0 is positive.
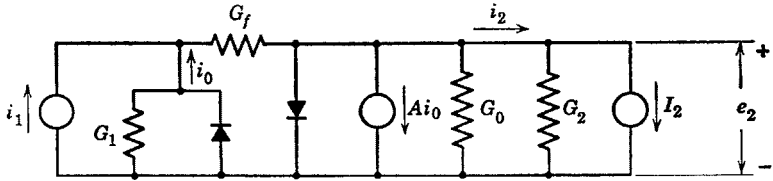
Fig. P7.4

(b) Simplify your result for the special case in which the conductance G_f is zero.

(c) Compare your results with the voltage transfer curves of certain vacuum-tube amplifiers.

7.9. In some pentodes (called gate tubes) the suppressor grid is brought to an external terminal instead of being internally connected to the cathode.

(a) Using physical reasoning, sketch a qualitative curve of plate current i_b vs. suppressor-grid voltage e_{c3} for e_b and e_{c2} fixed positive and e_{c1} zero. Show only the negative e_{c3} region.

(b) How would you connect the pentode and how would you bias the control and suppressor grids so that a positive signal pulse applied to one of them will produce a negative output pulse at the plate if and only if a positive "gating" pulse is simultaneously applied to the other grid?

(c) Discuss the "pedestal" effect, where part of the gating waveform appears at the output along with the waveform to be gated.
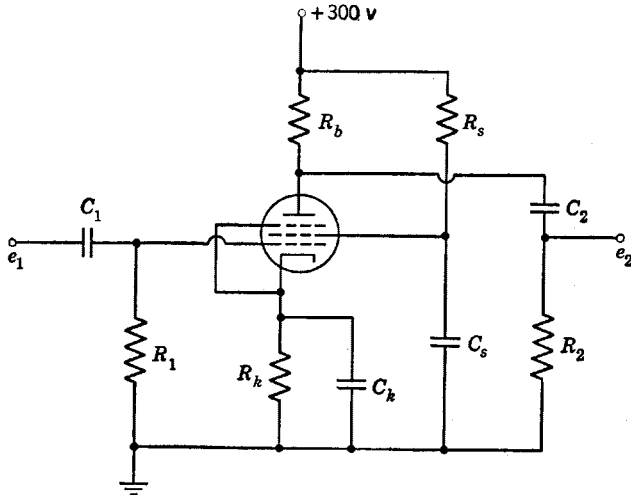
Fig. P7.5

7.10. In the pentode circuit of Fig. P7.5, e_1 is a small a-c signal and the capacitors shown are large enough so that their a-c reactances are negligible at the frequency of e_1.

(a) How would you choose R_1, R_k, R_s, R_2, and R_b to obtain a desired d-c operating point?

(b) How would you choose suitable values of C_1, C_k, C_s, and C_2?

(c) How would you determine the constants r_p and g_m of the linear incremental pentode model appropriate to this operating region?

(d) Calculate the incremental a-c voltage gain of the amplifier in terms of appropriate constants. Assume that interelectrode and stray wiring capacitance offers negligibly small susceptance at the frequency of interest.
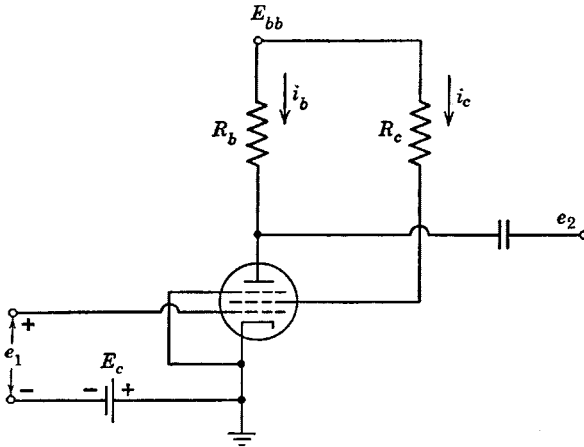
Fig. P7.6

7.11. A pentode is connected as shown in Fig. P7.6. Find the gain $\Delta e_2/\Delta e_1$ in terms of R_b, R_c, and tube parameters.
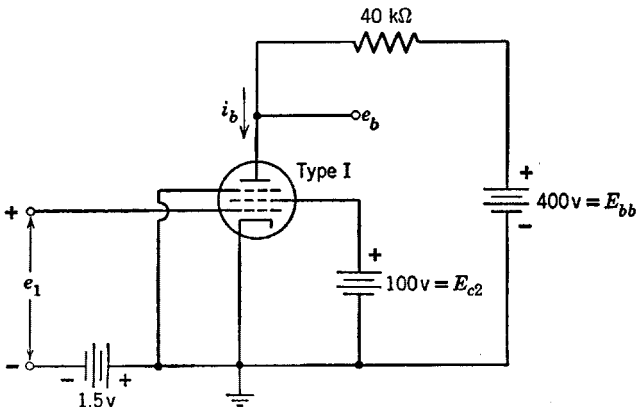
Fig. P7.7

7.12. A Type I pentode is connected in the voltage amplifier circuit shown in Fig. P7.7.

(a) When e_1 is a sinusoidal voltage of 1 volt rms, determine the maximum and minimum values of the plate voltage e_b.

(b) When e_1 is a sinusoidal voltage of 0.01 volt rms, determine the maximum and minimum values of the plate voltage e_b.

(c) If the screen supply voltage E_{c2} should fall to 90 volts, find the values of plate voltage e_b and plate current i_b for $e_1 = 0$.
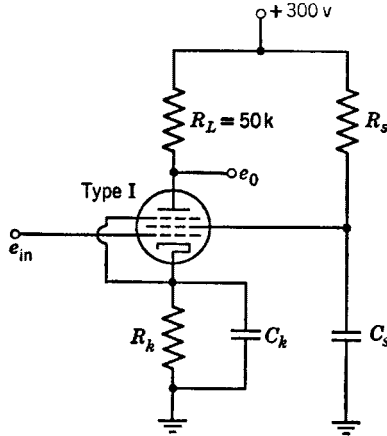
Fig. P7.8

7.13. The cathode bias resistor R_k in the pentode circuit shown in Fig. P7.8 produces a grid bias voltage of -1.5 volts and the quiescent screen-grid current through R_s reduces the screen grid voltage to 100 volt. Find:

(a) The quiescent plate current and plate voltage.

(b) The quiescent screen-grid current on the assumption that the ratio of the screen grid current to the plate current is independent of the negative grid-bias voltage for any plate voltage well above that at the knee of the curves.

(c) The required value of R_s.

(d) The required value of R_k.

(e) The gain e_0/e_{in}.

Assume C_k and C_s are short circuits for alternating currents. Use the Type I pentode characteristics.

7.14. The thyratron shown in Fig. P7.9 may be assumed ideal so that it is replaceable by the model shown in Fig. P7.10. The relay closes only when
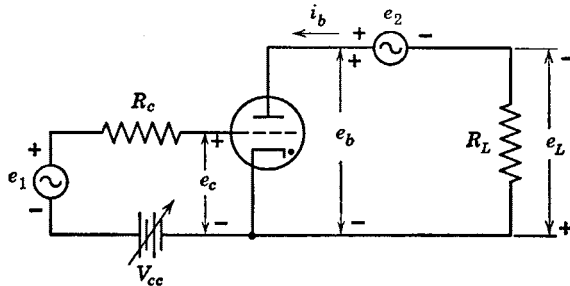
Fig. P7.9

current flows in either or both of the coils. Coil inductance and resistance are negligible in this idealization. Let

$$e_1 = V_1 \sin \omega t$$

$$e_2 = V_2 \cos \omega t$$
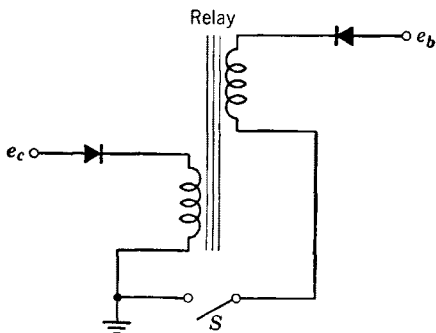
Fig. P7.10

(a) Plot the time-average value of e_L as a function of V_{cc} over the control range.

(b) Repeat part (a) for $e_1 = -V_1 \sin \omega t$.

7.15. A negative-grid thyratron has a firing curve specified by the following data:

E_c (volts)	E_b (volts)
+2	40
0	62
−2	110
−4	200
−6	310
−8	440

The thyratron is used in the circuit shown in Fig. P7.11. Find the firing angle of the thyratron. Find the average d-c current through the 100-ohm load resistor assuming that the thyratron when conducting has a constant voltage
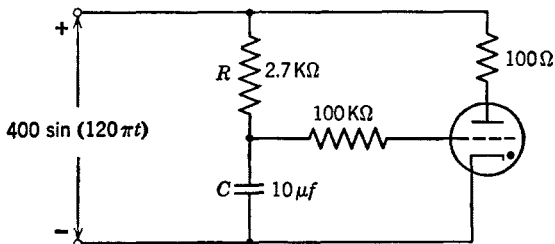
Fig. P7.11

drop of 10 volts. Draw the waveform of the current through the load. If the positions of R and C are interchanged, what is the new firing angle? The average load current?

7.16. In the thyratron rectifier circuit shown in Fig. P7.12, it is desired to achieve plate current conduction for exactly one-twelfth of each cycle of e_1. The applied potentials are:

$$e_1 = 300 \sin \omega t$$
$$e_2 = -10 + 10 \sin(\omega t + \varphi)$$
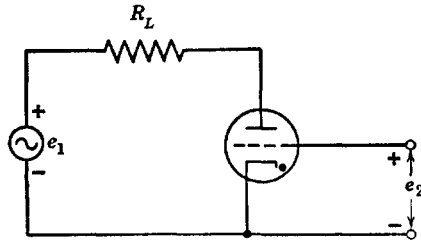
Fig. P7.12

The critical grid potential is given as $e_{1\,\text{crit}} = -e_b/50$. Neglect the tube drop during conduction.

(a) Sketch the locus of firing points in the e_2 vs. t plane as φ varies.

(b) Find the value of φ to produce plate conduction for exactly one-twelfth of each cycle of e_1.

(c) If the plate circuit conducts for one-twelfth of each cycle of e_1, find the rms value of the current in R_L.
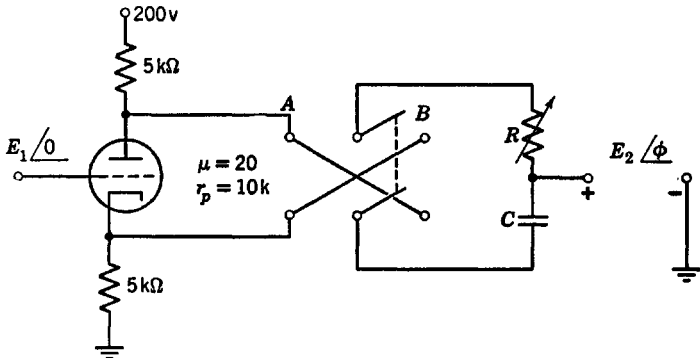
Fig. P7.13

7.17. The circuit shown in Fig. P7.13 can be used for phase-shifting the grid voltage of a thyratron. Derive an expression for phase shift as a function of the variable resistance for switch position A. Sketch the general appearance of the curve of ϕ vs. R for switch positions A and B. Let the input frequency be 60 cycles and let the capacitive reactance be 100 kilohms. Assume the triode operates in the linear amplification region and neglect the loading effects of the RC circuit.

7.18. A pair of phototubes are used in a bridge circuit (Fig. P7.14) to obtain an output response dependent on the position of a light beam.

When the beam is centrally positioned with respect to the phototubes, 0.08 lumen falls on each tube and the output voltage is zero. Suppose the

45v

Output

45v

Light beam
0.16 lumen

Phototube characteristic

Anode microamperes

Light flux = 0.10 lumen
0.08
0.06

0   20   40   60   80
Anode volts
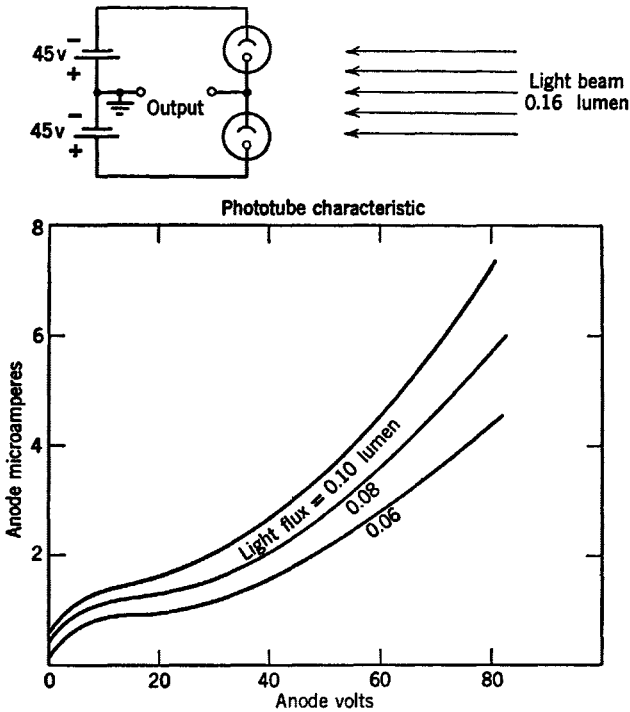
Fig. P7.14

beam is shifted so that 0.10 lumen falls on the upper phototube, the remainder falling on the lower tube.   Determine the magnitude of the output voltage and its polarity with respect to ground potential.

Wave Shaping and Amplification

8.1 Introduction

Wave shaping describes the function performed by a wide variety of electronic circuits. For example, a limiter may produce a rectangular output waveform from a sinusoidal input waveform. An integrating circuit converts a rectangular wave to a triangular wave. Linear amplification of voltage or current time functions calls for an increase in amplitude without distortion of waveform and can be regarded as a special case of wave shaping. The basic aspects of linear amplification are considered here, along with wave shaping, because many of the circuits that modify or shape a waveform can be used to amplify without modifying the waveform. The difference in function depends on the values of circuit elements and on the input waveform. Analysis of the wave-shaping properties of a circuit determines the conditions necessary for linear amplification.

In the piecewise-linear approximation of a resistive circuit, the condition for linear amplification amounts to restricting the operation to a single state of the ideal diodes in the circuit model. This can be done by restricting both the direct component and the alternating component of the input voltage to suitable values. With one or more energy-storage elements in the circuit, the wave-shaping properties are frequency dependent, in which case both the frequency and the amplitude of the input waveform affect the result.

8.2   The Role of Energy-Storage Elements

In this chapter we shall concern ourselves with the basic effects of a single energy-storage element on nonlinear or piecewise-linear circuit operation. The results of our calculations will apply to more complicated circuits when the effects of a single energy-storage element can be isolated in frequency or in time; for example, if a circuit includes two capacitors, one very large and one very small, the circuit behavior can usually be specified with adequate accuracy by considering the effects of each capacitor separately.

Since electronic devices normally require polarizing voltages, the interconnection of circuits becomes a problem. Direct coupling with resistances and batteries is cumbersome and costly. When the direct component of a signal need not be preserved, the d-c isolation provided by a capacitor or a transformer results in considerable circuit simplification. We shall examine simple coupling circuits and their influence on wave shaping and amplification. We shall also examine the effects of small energy-storage elements (such as interelectrode capacitances) which are inherent in all physical devices. Adding these appropriately to a resistive model extends the range of frequencies for which the model is valid.

8.3   Wave Shaping with Nonlinear Resistive Circuits

The wave-shaping properties of resistive circuits depend upon the non-linearities in driving-point or transfer curves. Many of the circuits discussed in the previous chapters illustrate amplitude limiting, clipping, or slicing. These terms denote only a difference in degree, since they all imply a removal of some portion of the input time function.

The basic limiter circuit is shown in Fig. 8.1 (a). The corresponding transfer curve, (b), has slopes of zero and $R_2/(R_1 + R_2)$. When the slope between the break points is required to be greater than unity, linear amplification is necessary. In comparing Fig. 8.1 (c) and (d) with (e) and (f), note that the end result depends upon the order in which limiting and amplifications are performed. For linear operations the end result is not affected by reversing the order.

To avoid considering resistive loading effects, assume that the amplifier has a very large input resistance and negligible output resistance. The amplifier required in (c) must, of course, have a greater range of linear operation than the one in (e). We note from the curve in (d) that the
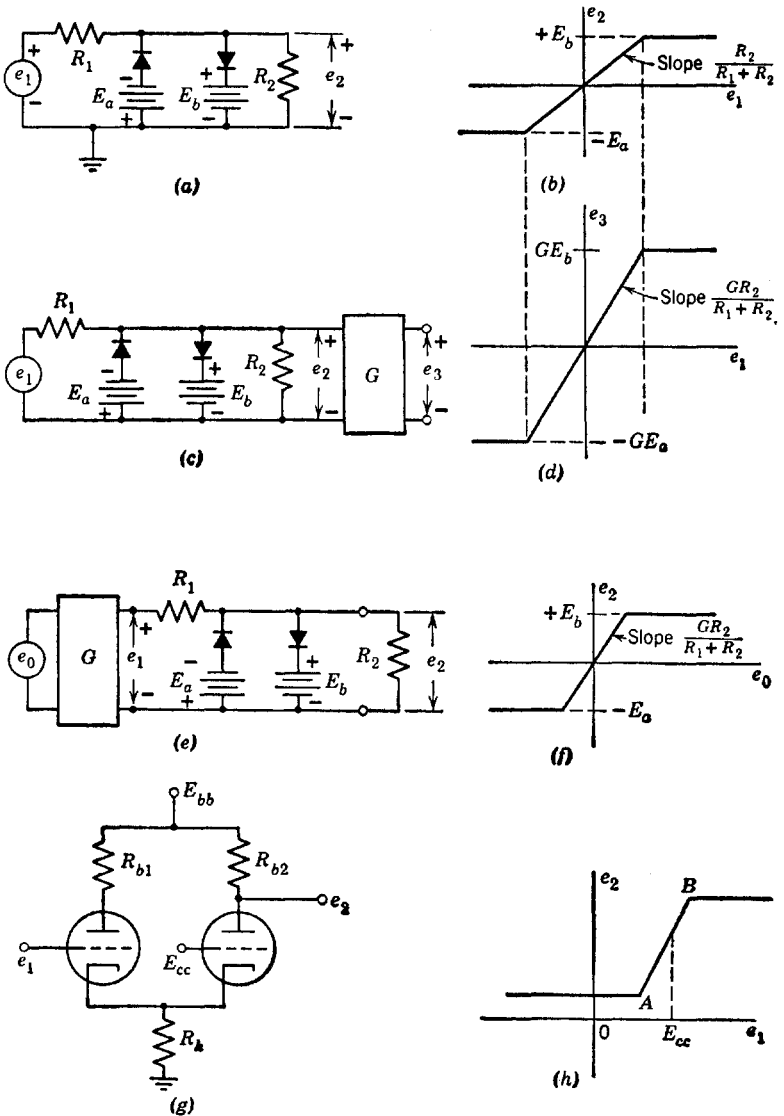
Fig. 8.1. Amplitude limiters.

input voltage range between break points has not changed but the output range has expanded. On the other hand, in the curve of (f) the output range is unchanged, and the range of input voltage between break points is compressed. The cathode-coupled circuit shown in Fig. 8.1 (g) is often used as a limiter. Its transfer curve is shown in (h).

The effects of a limiter on a variety of electrical time functions $e(t)$ are shown in Fig. 8.2. Note that the amplitude of the waveform, together with the d-c level, determines the character of the output wave. For
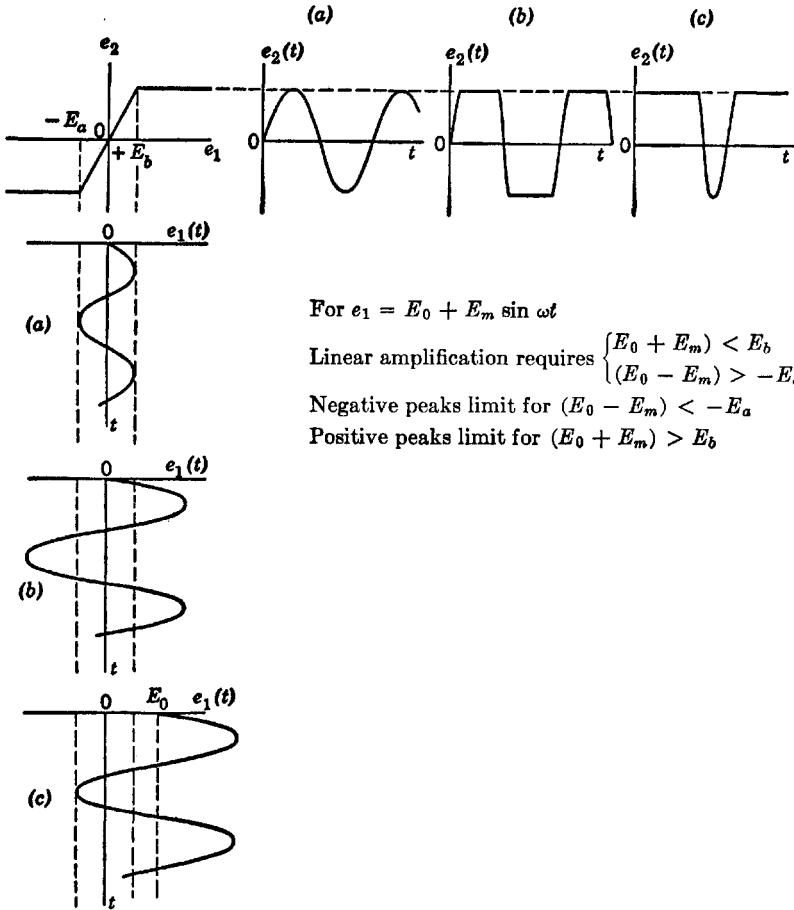
For $e_1 = E_0 + E_m \sin \omega t$

Linear amplification requires $\begin{cases} (E_0 + E_m) < E_b \\ (E_0 - E_m) > -E_a \end{cases}$

Negative peaks limit for $(E_0 - E_m) < -E_a$

Positive peaks limit for $(E_0 + E_m) > E_b$

Fig. 8.2. Linear amplification and limiting.

this limiter-amplifier curve, linear amplification is obtained if the time-varying signal can be fitted between the break points, as shown in part (a). Symmetrical and unsymmetrical limiting or clipping are illustrated by (b) and (c). If the distance between break points along the input axis is very small compared to the total amplitude of the input signal, the operation is called "slicing." This amounts to passing only a thin slice of the input waveform, and results in a very nearly rectangular output waveform. Slicing amounts to "hard limiting," and effectively

determines the instant at which the input time function crosses a given voltage level.

8.4 Wave Shaping with Linear Energy-Storage Elements

The relation between current and voltage for a resistance is independent of time. The terminal relations between current and voltage for inductance or capacitance are time dependent ($v = L \, di/dt$ and $i = C \, dv/dt$). The simple voltage and current waveforms shown in Fig. 8.3(b) and (c) illustrate the wave-shaping properties of inductance and capacitance. Whether we consider current to be the cause and voltage the effect, or vice versa, the shapes of the time functions $v(t)$ and $i(t)$ are
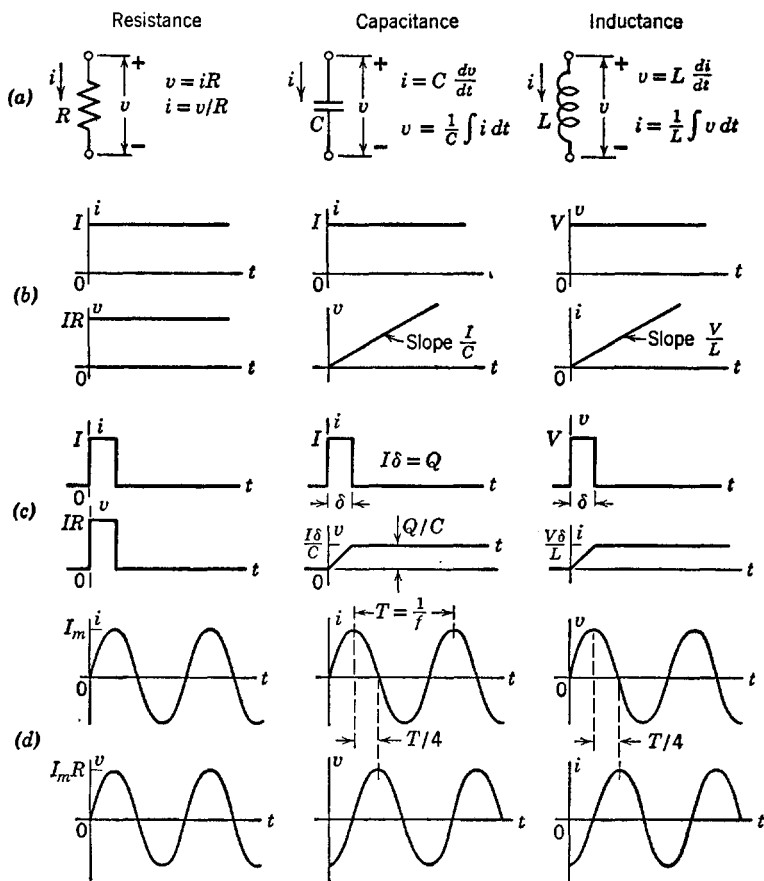
Fig. 8.3. Current-voltage relations for R, L, and C.

different for L and C, but identical for R. For the idealized waveforms shown, the necessary differentiation or integration can be carried out by inspection. In (d), note that $v(t)$ or $i(t)$ is subject to a time delay (or advance) rather than a change in shape.

8.5   Series *RC* Circuit with Sine-Wave Input

The response of a series RC circuit to a sinusoidal input is summarized in Fig. 8.4. This circuit is of considerable importance, since it represents the basic capacitive coupling network used for d-c isolation. The direct component of the input voltage charges the capacitance to the value $Q_0 = CE_0$, but does not affect the steady-state a-c voltage across the resistor; nor does the static capacitor voltage E_0 affect the variational capacitor voltage. The separability of these d-c and a-c effects is a consequence of the fact that the principle of superposition applies to
• linear systems. The voltage waveforms appear in (d). We have tacitly assumed that the d-c and a-c components of e_1 were applied to the circuit at some time in the past, and that steady-state conditions already existed at our arbitrary time reference $t = 0$.

The vector diagram [Fig. 8.4(e)] shows the relative amplitudes and phase relations of the three voltages involved. The direction of the current vector I_m coincides with that of $E_2 = I_m R$. This vector diagram is drawn for a particular value of angular frequency ω. Since the current is common, a 90° phase relation exists between resistive voltage and capacitive voltage for any value of frequency. Thus the locus of point P is a semicircle if ω is varied from zero to infinity while E_m is held constant. The vector diagram also shows that the magnitude of e_x or e_2 cannot exceed that of e_1, since the diameter of the circle is larger than any chord.

8.6   Step Response of a Series *RC* Circuit

Let us now review the short-time or transient effect of suddenly applying a d-c voltage to a series RC circuit, Fig. 8.5(a). The basic equations are given in (b). The requirement that e_1 shall equal zero for t less than zero, and equal E for t greater than zero, is met by the circuit given in (c), where switch S closes at $t = 0$ to apply the battery E to the RC combination. This is equivalent to a time-varying generator $e_1(t)$ whose output waveform is a step of voltage from 0 to E at $t = 0$. If the time t is plotted in units $RC = \tau$ (the time constant), the voltage wave-

(a)

(b)
$$e_1 = e_x + e_2$$
$$= \frac{1}{C} \int i \, dt + iR$$
or
$$\frac{de_1}{dt} = \frac{i}{C} + R\frac{di}{dt}$$
For $e_1 = E_0 + E_m \sin \omega t$
$$\frac{de_1}{dt} = \omega E_m \cos \omega t$$

(c)
$$e_1 = E_0 + E_m \sin \omega t$$
$$i = I_m \sin (\omega t + \phi)$$
Where $I_m = \dfrac{E_m}{\sqrt{R^2 + \left(\dfrac{1}{\omega C}\right)^2}}$ and $\tan \phi = \dfrac{1}{\omega CR}$
$$e_2 = I_m R \sin (\omega t + \phi)$$
$$e_x = E_0 + \frac{I_m}{\omega C} \sin \left(\omega t + \phi - \frac{\pi}{2}\right)$$

(e)

(d)
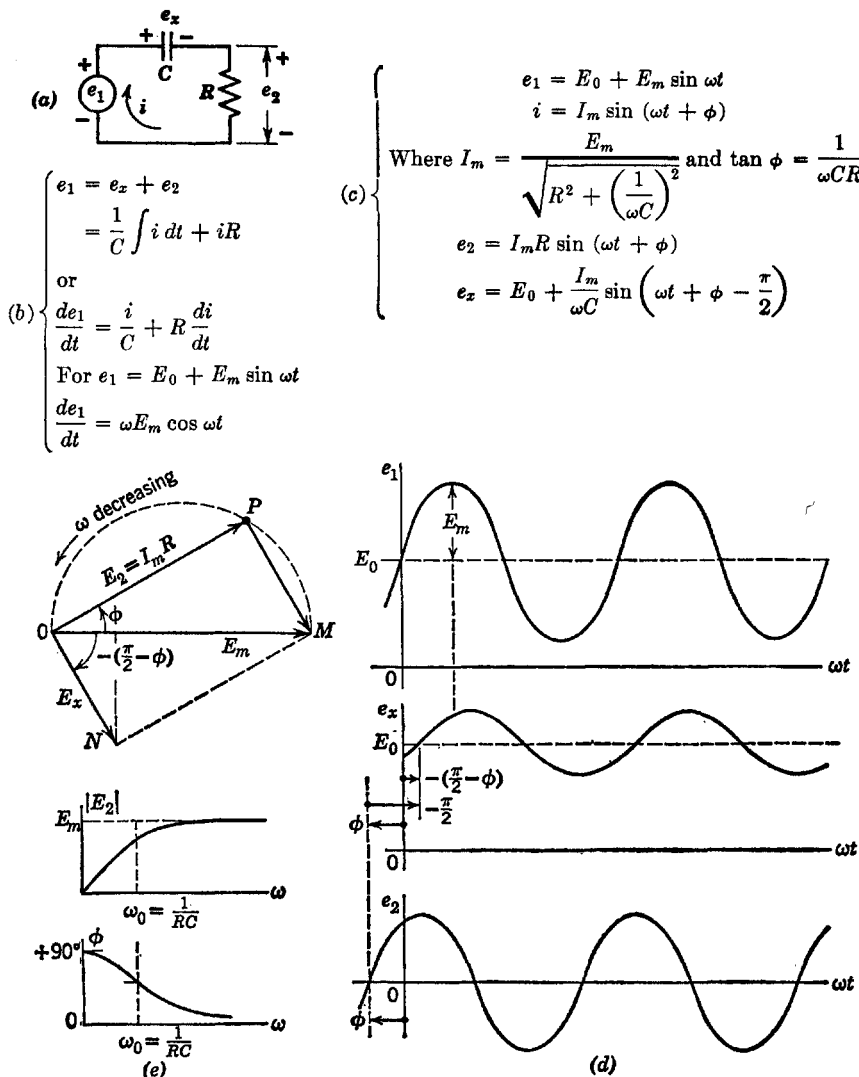
Fig. 8.4.   Linear RC circuit with d-c and a-c voltages applied.

forms are independent of R and C, and the current waveform involves R only in the amplitude scale. Alternatively, we may plot normalized time $(x = t/RC = t/\tau)$, in which case the units are 1, 2, 3 etc., instead of τ, 2τ, 3τ, etc. (See exponential curves in Appendix D.)

The slope of an exponential waveform at any point is the remaining amplitude divided by the time constant. Thus, at $t = 0$ the slope has

a magnitude E/τ; and if the initial slope is extended to $t = \tau$, the final value is reached.

Integration of the current waveform from $t = 0$ to $t = \infty$ yields $E\tau/R = EC = Q$, the charge required to change e_x from zero to E. It

(a)

(b)
$$e_1 = e_x + e_2$$
$$= \frac{1}{C} \int i \, dt + iR$$
Let $e_1 = 0$ for $t < 0$
$e_1 = E$ for $t > 0$

Closes at $t = 0$

(c)

(d)
For $t > 0$
$$\frac{de_1}{dt} = 0 = i/C + R \, di/dt$$
$$di/i = -dt/RC$$
$$\ln i = -t/RC + \ln A$$
$$i = A\epsilon^{-t/RC}$$
$$= \frac{E}{R} \epsilon^{-t/RC}$$
$$e_2 = E\epsilon^{-t/RC}$$
$$e_x = E(1 - \epsilon^{-t/RC})$$
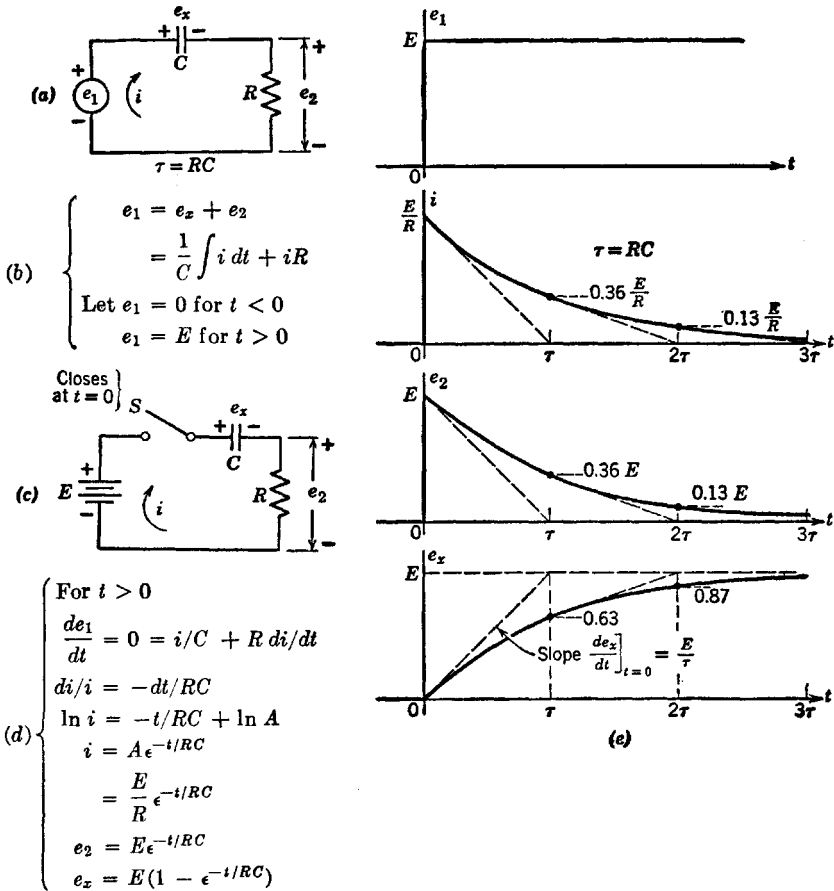
Fig. 8.5. Step response of series RC circuit.

follows that the total area under the waveform of e_2 is this value times R: namely, $E\tau$. With reference to the waveform of e_2 in part (e), this means that half of the area under the e_2 curve is under the dotted line between $t = 0$ and $t = \tau$, while the remaining half is bounded by the waveform, the dotted line, and the horizontal axis.

Since superposition applies to linear circuits, it is convenient to

normalize the amplitude as well as the time scale.   Thus, if we divide
all voltage and current values by E, the results are those that would be

(a)

(b)

(c)

(d)

(e)
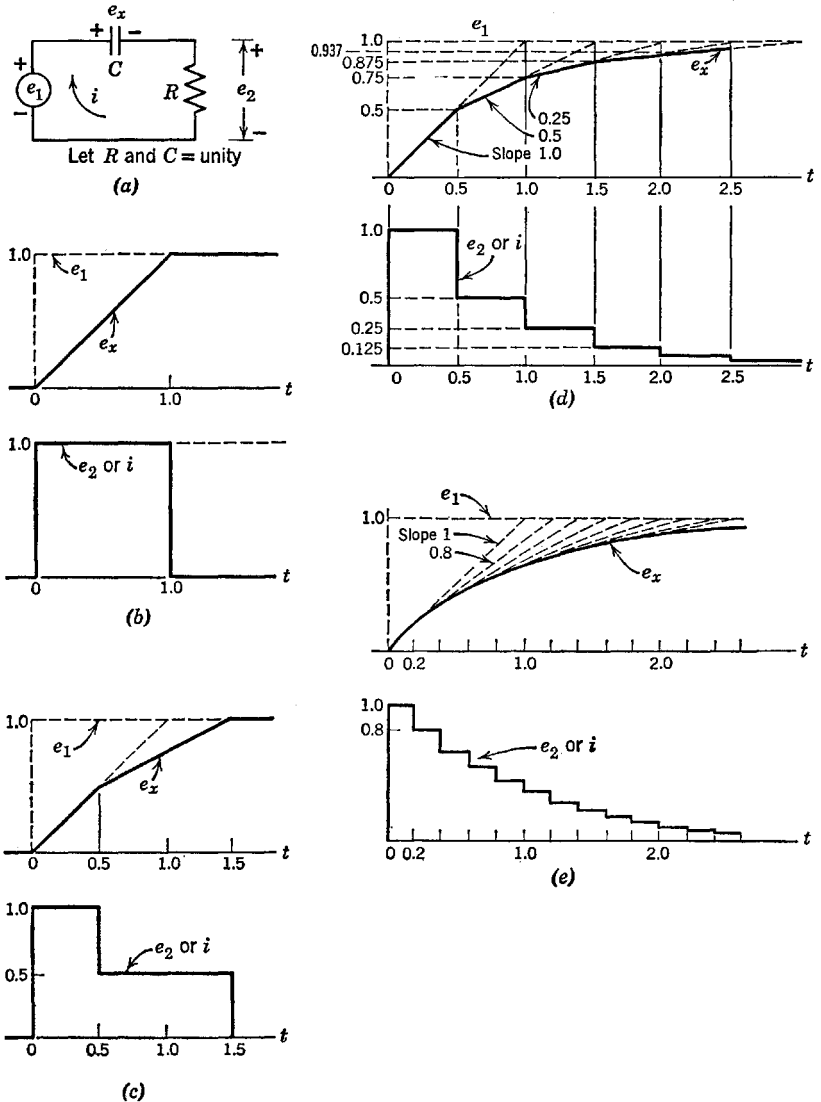
Fig. 8.6.   Approximating the step response of the series RC circuit.

obtained with a step of unit amplitude applied to the circuit.   Then,
using a normalized time variable $(x = t/RC)$, we have initial slopes

equal to unity (or minus one), the area under e_2 equal to unity, and the area under the current waveform equal to $1/R$.

The step and ramp waveforms for individual elements (R and C) given in Fig. 8.3(b) can be used to approximate the behavior of a circuit such as the series RC circuit of Fig. 8.6(a). For convenience, we shall assume R is 1 ohm and C is 1 farad. Also, let the input step have unit amplitude.

The sketches in Fig. 8.6(b) represent a very crude approximation to the exponential curves. The upper sketch shows waveforms of e_1 (dotted line) and e_x (solid line), whereas the second sketch shows e_2 or i. Application of a unit step e_1 causes unit voltage e_2, since e_x cannot change instantaneously unless there is an impulse of current (a finite charge deposited on the capacitor by means of an infinite current lasting for zero time). The unit voltage e_2 causes unit current i. This current causes e_x to increase with unit slope ($de_x/dt = i/C$). If this situation persists for unit time, the capacitor voltage reaches unity, the steady-state value, which corresponds to a steady-state current $i = 0$. Thus we are assuming a rectangular pulse of current i, which leads to the ramp e_x. However, adding e_x to e_2 yields an input voltage which exceeds the unit step by the ramp voltage e_x. This is obviously a very poor representation of the step input. The pulse and ramp are equally poor approximations to the desired exponentials.

The waveforms shown in Fig. 8.6(c) begin in the same fashion as those in (b), but persist only until $t = 0.5$. Now, since $e_x = 0.5$ at $t = 0.5$, the current i should correspondingly be reduced to 0.5, since $i = (e_1 - e_x)/R$. If this current persists for one time constant, e_x rises with slope 0.5 and reaches unity at $t = 1.5$. The current waveform is still very crude, but the integration has given e_x some semblance of realism. The waveforms in Fig. 8.6(d) and (e) are more accurate approximations, obtained by simple extension of the above techniques.

8.7 Step, Ramp, and Pulse Waveforms

The ramp waveform, like the step, is convenient to use in electronic circuit analysis. As indicated in Fig. 8.7(a) and (b), the integral of a step with an amplitude of E volts is a ramp with a slope of E volts per unit time. The impulse waveform indicated in Fig. 8.7(c) can be regarded as the derivative of the step waveform. Since the step rises to amplitude E in zero time, the slope is infinite. Interpreting the impulse as the derivative of the step requires that it have infinite amplitude for zero time with an area equal to the step amplitude E. This is more

readily visualized if we regard the step as the limiting case of a truncated ramp with an amplitude E. The sequence of waveforms in Fig. 8.7 (d) indicates the trend of the pulse toward the impulse as the slope of the

(a) $x(t) = 0$ for $t < 0$
       $x(t) = Kt$ for $t > 0$
       Slope K
       $x(t) = \int y(t)\, dt$

(b) $y = \frac{dx}{dt}$ $y(t) = 0$ for $t < 0$
                             $y(t) = K$ for $t > 0$
     K
                             $y(t) = \int z(t)\, dt$

(c) $z = \frac{dy}{dt} = \frac{d^2x}{dt^2}$ $z(t) = 0$ for $t \gtrless 0$
                                               $\dot{z}(t) = \infty$ for $t = 0$
     Impulse of
     area K $\int_{-\infty}^{\infty} z(t)\, dt = K$

(d)

$y = \frac{dx}{dt}$
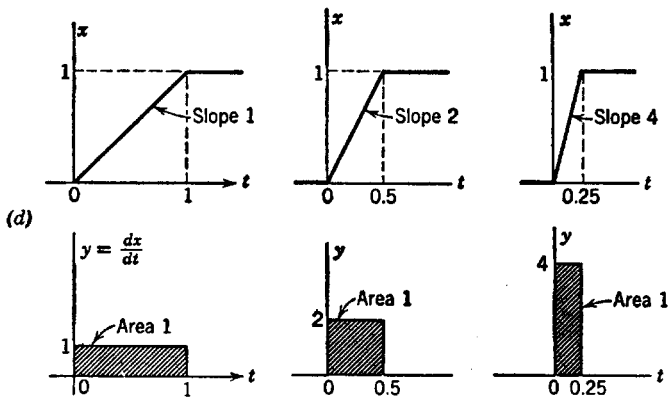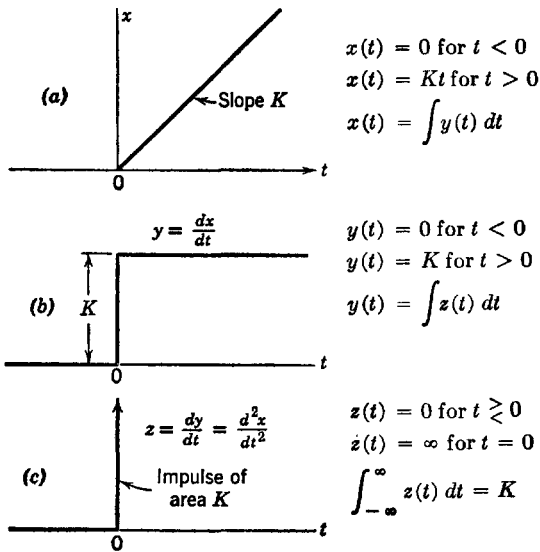
Fig. 8.7. Relations among step, ramp, and impulse waveforms.

limited ramp approaches that of the step. The area of the pulse remains constant and equal to the amplitude E of the ramp waveform as the limiting process takes place.

The impulse waveform is very useful in linear circuit analysis. In analyzing an electronic circuit, we can replace a pulse of amplitude E and duration δ by an impulse of area $E\delta$, provided the amplitude E is

(a)

$$e(t) = 0 \text{ for } t < 0$$
$$e(t) = E \text{ for } 0 < t < \delta$$
$$e(t) = 0 \text{ for } t > \delta$$

(b)

$$\left.\begin{array}{l} e_a(t) = 0 \text{ for } t < 0 \\ e_a(t) = E \text{ for } t > 0 \\ \\ e_b(t) = 0 \text{ for } t < \delta \\ e_b(t) = -E \text{ for } t > \delta \end{array}\right\} e_a + e_b = e$$

(c)

$$e(t) = 0 \text{ for } t < 0$$
$$e(t) = Et/\delta \text{ for } 0 < t < \delta$$
$$e(t) = E \text{ for } t > \delta$$

(d)

$$\left.\begin{array}{l} e_a(t) = 0 \text{ for } t < 0 \\ e_a(t) = Et/\delta \text{ for } t > 0 \\ \\ e_b(t) = 0 \text{ for } t < \delta \\ e_b(t) = -E(t-\delta)/\delta \text{ for } t > \delta \end{array}\right\} e_a + e_b = e$$

(e)
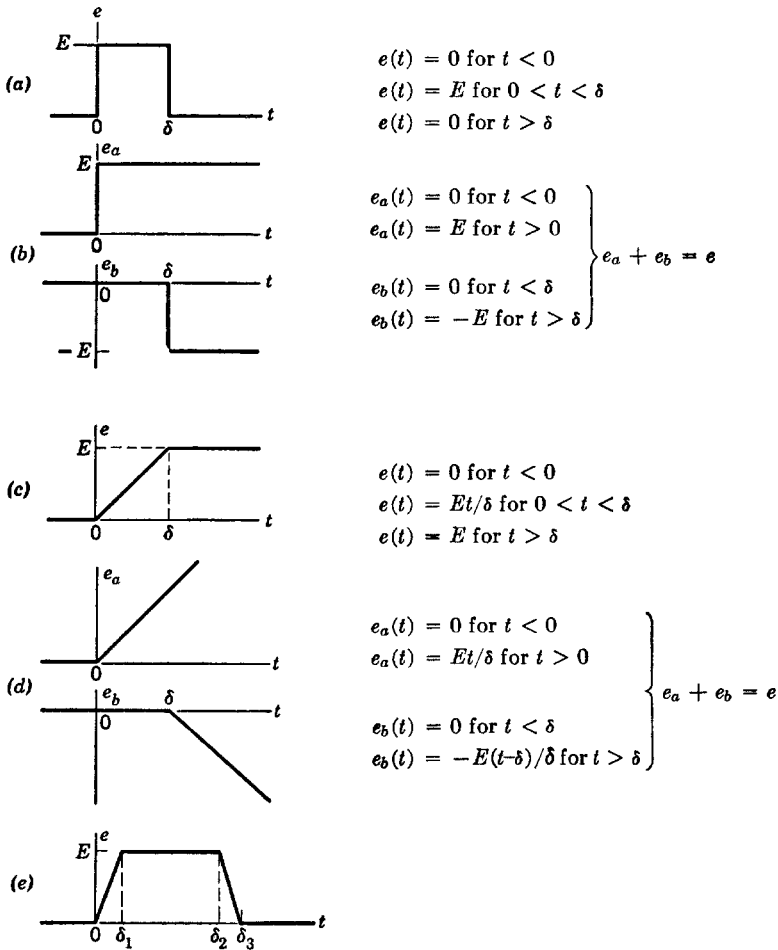
Fig. 8.8. Composite waveforms.

such that the electronic device operates essentially linearly, and the duration δ is small compared with the time constants associated with the energy-storage elements.

The pulse $e(t)$ shown in Fig. 8.8(a) is an example of a composite waveform, since it can be made up of a positive step $e_a(t)$ and a negative

step $e_b(t)$ delayed by the pulse duration δ. The limited ramp is another composite waveform. As shown by (c) and (d), it is equivalent to a positive continuous ramp and a delayed negative ramp. Similarly, the waveform given in Fig. 8.8(e) can be composed of continuous ramp functions.
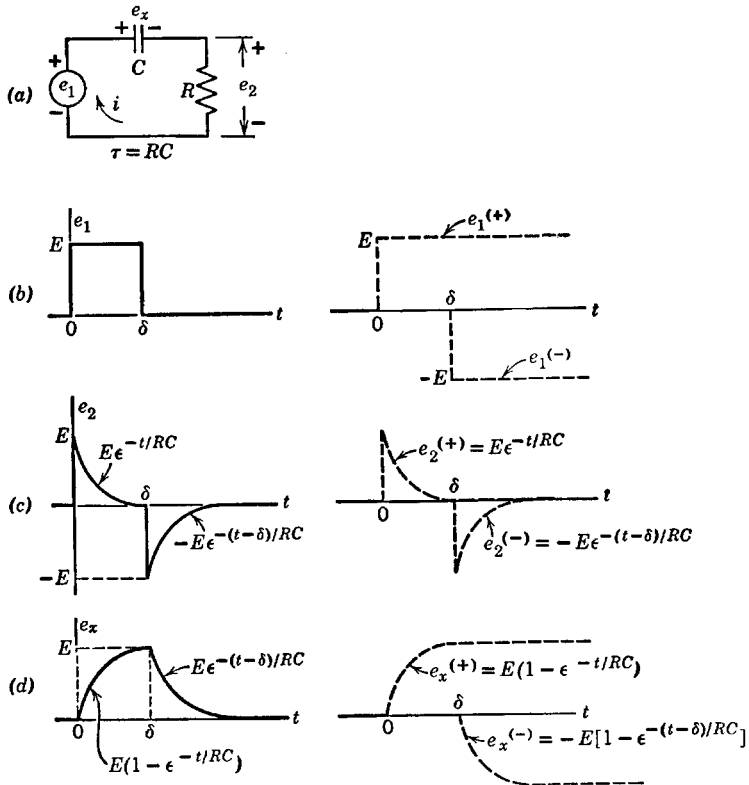
Fig. 8.9. Response of a series RC circuit to a rectangular pulse having a duration δ much greater than the time constant ($\delta > 5RC$).

8.8 Pulse Response of a Series RC Circuit

Since superposition applies to linear circuits, the response of a given circuit to a composite waveform can be expressed as the sum of the responses to the basic elements of the driving waveform. Thus, if the input voltage e_1 for the circuit of Fig. 8.9(a) consists of a rectangular pulse, we can use the results given in Fig. 8.5(d) and (e) to find the output voltage e_2.

The waveforms shown in Fig. 8.9 correspond to the case in which the circuit time constant is much less than the duration of the input pulse, hence the transient due to the positive step is "complete" before the delayed negative step occurs. The capacitor voltage E at $t = \delta$ merely
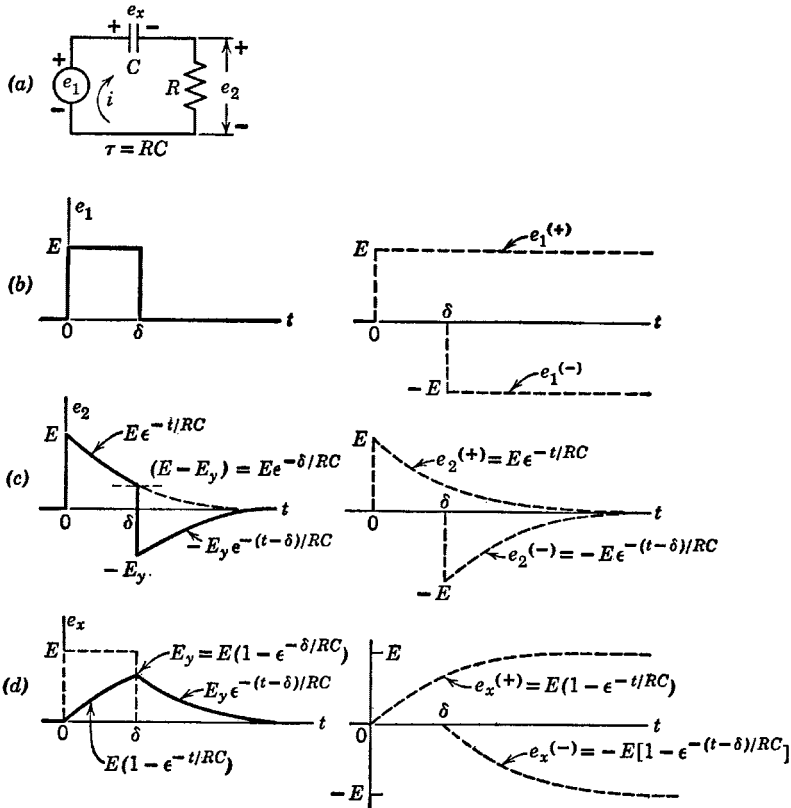
Fig. 8.10. Response of a series RC circuit to a rectangular pulse ($\delta < 5RC$).

shifts the zero level for the exponential resulting from the negative step. The waveforms for e_1, e_2, and e_x are shown in composite form in (b), (c), and (d). The separate components resulting from individual step responses are shown at the right of each waveform. The analytic expression for the composite of e_2 is at every point equal to one component or the other, since only one transient is in progress at any one time.

When the pulse duration is insufficient to permit the first transient to die out before the next one begins, the results are slightly more complicated. The waveforms for this case are shown in Fig. 8.10. During

the pulse, the expressions for e_2 and e_x are the same as before. Numerical values and the shapes of the waveforms differ from those of Fig. 8.9, since the transients do not reach their final values. In general, the value reached by the capacitor voltage at $t = \delta$ is

$$E_y = E(1 - \epsilon^{-\delta/RC}) \tag{8.1}$$

At this point e_2 is $(E - E_y)$. For t greater than δ, the remaining portions of the waveforms due to $e_1{}^+$ subtract from those initiated at $t = \delta$ by $e_1{}^-$. Simple algebraic manipulations reduce the results of the summing operation to the forms shown on the composite waveforms; that is, for t greater than δ

$$e_x = E_y\epsilon^{-(t-\delta)/RC} \tag{8.2}$$

and

$$e_2 = -E_y\epsilon^{-(t-\delta)/RC} \tag{8.3}$$

The fact that $e_2 = -e_x$ for t greater than δ is immediately obvious, since $e_1 = 0$ here. More important is the fact that the final expressions involve only E_y in the amplitude factor. Thus, for t greater than δ, the waveforms depend only on the capacitor voltage at $t = \delta$. If the source voltage e_1 is nonzero for t greater than δ, the waveforms will include another term corresponding to the source voltage. *Regardless of the source applied for t greater than δ, the residual effects of anything that transpired for t less than δ are all accounted for by the capacitor voltage E_y at t = δ.* We can thus conveniently use a new time reference for each discrete section of a waveform. For the waveforms of e_2 and e_x in Fig. 8.10, we could express the voltage variations somewhat more simply for t greater than δ by counting time from $t = \delta$: in other words, replace $(t - \delta)$ by t' or simply by t. The latter amounts to considering $\delta = 0$, in which case δ disappears from the expressions. The numeric E_y computed in the interval $0 < t < \delta$ is, of course, not subject to this modification. Although the dropping of one symbol may seem trivial here, the extension of this idea to repetitive waveforms saves time and effort during the course of a calculation. A delay factor can be reinserted in the final result to refer to a specific time reference if desired.

8.9   Ramp Response of a Series RC Circuit

The series RC circuit with a ramp input is shown in Fig. 8.11, and the differential equation for the current is established in (b). The solution for the current is

$$i = KC(1 - \epsilon^{-t/RC}) \tag{8.4}$$

For $t \gg RC$, the current approaches KC, the value that would exist if

the ramp were applied to the capacitance alone. Expressions for the waveforms of voltages e_2 and e_x follow immediately from the expression for current, since e_2 equals iR and e_x equals $(e_1 - e_2)$. The RC circuit response to a limited ramp is shown in Fig. 8.12 for δ greater than 5τ and δ less than 5τ, where $\tau = RC$.
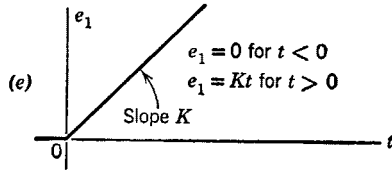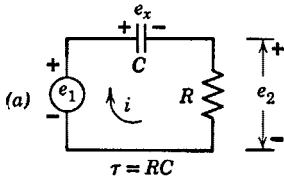
For $t > 0$

$$(b) \begin{cases} e_1 = Kt = \dfrac{1}{C}\displaystyle\int i\,dt + Ri \\[2mm] \dfrac{de_1}{dt} = K = \dfrac{i}{C} + R\dfrac{di}{dt} \\[2mm] \dfrac{di}{dt} + \dfrac{i}{RC} = \dfrac{K}{R} \end{cases}$$
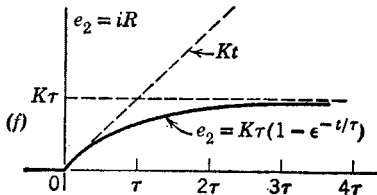
$$(c) \begin{cases} \dfrac{dy}{dx} + Py = Q \\[2mm] y = \epsilon^{-\int P\,dx}\displaystyle\int Q\epsilon^{\int P\,dx}\,dx + k \end{cases}$$

$$(d) \begin{cases} i = \epsilon^{-t/RC}\left(\dfrac{K}{R}\right)[RC\epsilon^{t/RC}]_0^t \\[2mm] = KC\epsilon^{-t/RC}[\epsilon^{t/RC} - 1] \\[2mm] = KC[1 - \epsilon^{-t/RC}] \end{cases}$$

(a) $\tau = RC$

(e) $e_1 = 0$ for $t < 0$
       $e_1 = Kt$ for $t > 0$
       Slope K

(f) $e_2 = iR$
       Kt
       $e_2 = K\tau(1 - \epsilon^{-t/\tau})$

(g) $e_x = e_1 - e_2$
       Kt
       $e_x = K[t - \tau(1 - \epsilon^{-t/\tau})]$
       $K(t - \tau)$
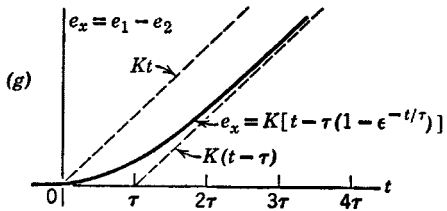
Fig. 8.11.   Response of a series RC circuit to a ramp waveform.

8.10   Integration and Differentiation of Excitation and Response Waveforms

Differentiation of the expressions for e_1, i, e_2, and e_x in Fig. 8.11 leads to the expressions obtained for a step waveform of e_1 [Fig. 8.5(d)] if we identify the constant K with the amplitude of the step. This calculation is summarized in Fig. 8.13. Conversely, the ramp response waveforms can be obtained directly from those for the step response by integrating each of the expressions in Fig. 8.5.

The generalization of this statement may be expressed as follows: If, for a linear system, the response function for a given excitation function
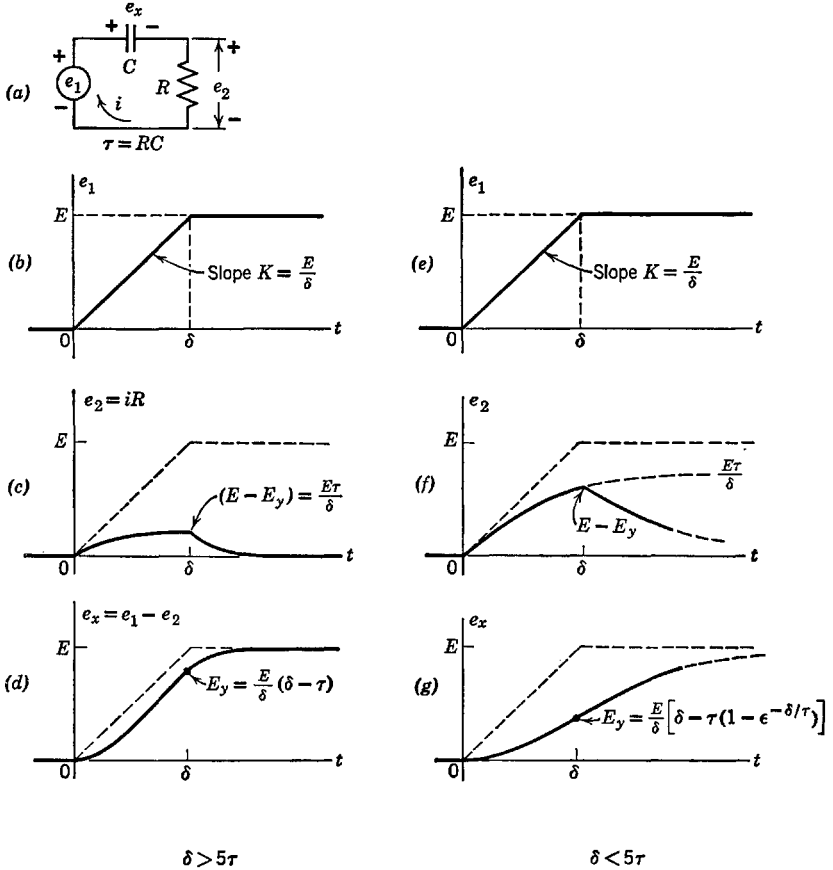
Fig. 8.12. Response of a series RC circuit to a limited ramp.

is known, the integrals of each—or the derivatives of each—form additional excitation-response pairs for that system. By mathematical induction we see that the process can be repeated so that the 2nd, 3rd, $\cdots$, nth integrals or derivatives also form excitation-response pairs of the system. The validity of these statements is one of the many useful consequences of linearity.

8.11 Square-Wave Response of a Linear Series *RC* Circuit

The response of the series RC circuit to a step waveform led directly to the response for a single pulse. This solution, in turn, can be extended

$$e_1 = Kt \qquad e_1' = \frac{de_1}{dt} = K$$

$$i = KC[1 - \epsilon^{-t/\tau}] \qquad i' = \frac{di}{dt} = \frac{K}{R}\epsilon^{-t/\tau}$$

$$e_2 = K\tau[1 - \epsilon^{-t/\tau}] \qquad e_2' = \frac{de_2}{dt} = K\epsilon^{-t/\tau}$$

$$e_x = K[t - \tau(1 - \epsilon^{-t/\tau})] \qquad e_x' = \frac{de_x}{dt} = K(1 - \epsilon^{-t/\tau})$$
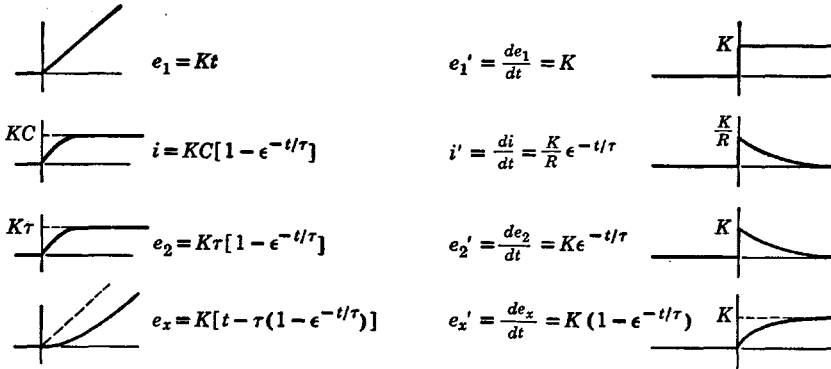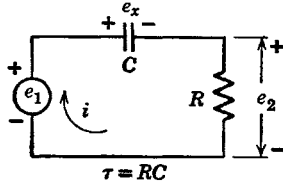
Fig. 8.13. Relation between ramp and step response of a linear RC circuit.

to the square wave. The circuit and the input waveform are shown in Fig. 8.14(a). Let us assume that the square wave has been applied for a long time prior to $t = 0$, so that steady-state conditions exist. Since the average value of current i is zero, the average value of e_2 must be zero. Since the average value or direct component of e_1 is $E/2$, the average value of e_x must also be $E/2$. Let us now consider the character of the waveforms of e_2 and e_x for various values of the time constant.

For the waveforms of e_x and e_2 shown in Fig. 8.14(b), the RC time constant is chosen to be much less than δ, so that the transient goes essentially to completion in each interval. The capacitor voltage can not change at the instant of switching; therefore $\Delta e_2 = \Delta e_1 = \pm E$ at the transitions of the square wave. As the time constant is made smaller, relative to the interval δ, the waveform of e_x approaches more closely the waveform of e_1, while e_2 approaches a succession of alternately positive and negative spikes. Since these spikes have a finite amplitude E, their area $E\tau$ approaches zero as τ approaches zero. In the limit, therefore, these spikes have a vanishingly small area.

When the time constant τ is comparable to the interval δ, neither the charge nor the discharge of the capacitor goes to completion. The waveforms shown in Fig. 8.14(c) have been drawn for δ less than τ, the actual value being $\delta \approx 0.7\tau$. In this case the capacitor charges toward
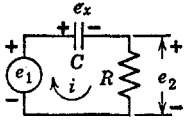
(a)   Let $\tau = RC$, $A = \epsilon^{-\delta/\tau}$
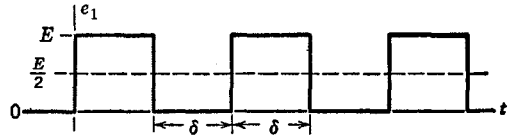
(b)   For $\delta > 5\tau$
e_x charges to E
and discharges to 0
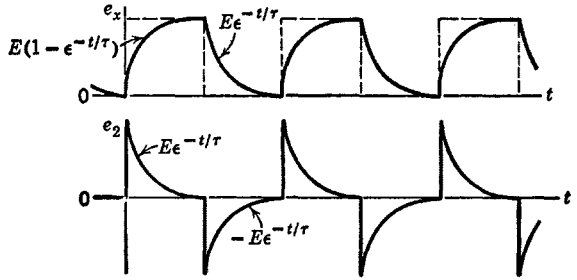
(c)   For $\delta \approx \tau$
e_x charges to
$$E_y = E/(1 + A)$$
$$= \frac{E}{2} + \frac{E}{2}(1 - A)/(1 + A)$$
and discharges to
$$E_x = AE/(1 + A)$$
$$= \frac{E}{2} - \frac{E}{2}(1 - A)/(1 + A)$$
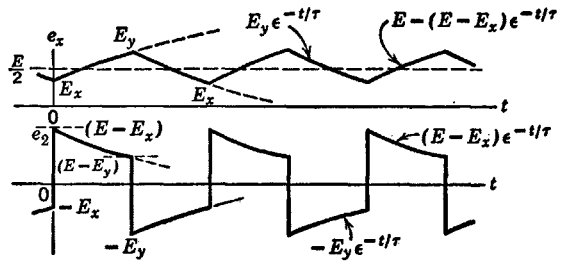
(d)   For $\delta \ll \tau$; $e_x \approx$ constant
$$A = \epsilon^{-\delta/\tau} \approx 1 - \delta/\tau,$$
$$E_y \approx \frac{E}{2}\left(1 + \frac{\delta}{2\tau}\right)$$
$$E_x \approx \frac{E}{2}\left(1 - \frac{\delta}{2\tau}\right)$$
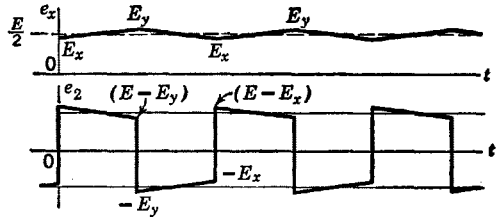
Fig. 8.14. Square-wave response of linear RC circuit.

E, but reaches only the value E_y. It then discharges toward zero, but falls only as low as E_x. With $t = 0$ taken at the beginning of a discharge interval, the expression for capacitor voltage during that interval is

$$e_x = E_y \epsilon^{-t/\tau} \tag{8.5}$$

The final value at $t = \delta$ is

$$E_x = E_y \epsilon^{-\delta/\tau} \tag{8.6}$$

$$E_x = A E_y \tag{8.7}$$

where

$$A = \epsilon^{-\delta/\tau} \tag{8.8}$$

If we now take $t = 0$ at the beginning of the charge interval, the expression for the capacitor voltage during charge is

$$e_x = E - (E - E_x)\epsilon^{-t/\tau} \tag{8.9}$$

or

$$E - e_x = (E - E_x)\epsilon^{-t/\tau} \tag{8.10}$$

At the end of the charge interval, $e_x = E_y$, so that Eq. 8.10 becomes

$$E - E_y = A(E - E_x) \tag{8.11}$$

Solving Eqs. 8.7 and 8.11 simultaneously we obtain

$$E_y = \frac{E}{1 + A} \tag{8.12}$$

$$E_x = \frac{AE}{1 + A} \tag{8.13}$$

If we express E_y as $E/2 + \Delta_1$, and E_x as $E/2 - \Delta_2$, we find

$$\Delta_1 = \Delta_2 = \frac{E(1 - A)}{2(1 + A)} \tag{8.14}$$

Thus the peak-to-peak amplitude of the variation in e_x is

$$\Delta e_x = E_y - E_x = \frac{E(1 - A)}{(1 + A)} \tag{8.15}$$

The sum of E_y and E_x is the input amplitude E, and the average of E_y and E_x is $E/2$. Since the exponentials have the same time constant and consequently the same form in the charge and discharge intervals, $E/2$ is also the time average of e_x.

An exponential is completely specified by the time constant τ and the values at two known instants of time. Hence, the waveform of e_2 in Fig. 8.14(c) is completely determined once the values of E_y and E_x have been found, since these—together with the input waveform—fix the initial and final values of e_2 in each interval.

If RC is increased so that δ/τ further decreases, the waveforms of e_x and e_2 tend toward the shapes indicated in Fig. 8.14(d). The constant A approaches unity, both E_x and E_y approach $E/2$; and the exponential variations become nearly linear. The waveform of e_x therefore approximates a triangular wave. For small values of δ/τ, the constant A can be approximated by $(1 - \delta/\tau)$. Substitution of this value into the expressions in Fig. 8.14(c) yields the approximate expressions given in

(d). The peak-to-peak amplitude of capacitor voltage is approximately $E\delta/2\tau$ in this case. The waveform of e_2 is nearly a reproduction of e_1, since the input waveform e_1 is closely coupled to the output. However, there is a change in average value or direct component. The direct component of e_1 appears across the capacitor.

8.12 Rectangular-Wave Response of Piecewise-Linear Circuits

The considerations in Article 8.11 imply no reasons for restricting the intervals for capacitor charge and discharge to equal values. Also, since the interval δ and the time constant τ enter the waveform expressions only as the ratio δ/τ, the time constant need not have the same value during the charge and discharge intervals. Finally, there is no reason to impose any restrictions on the direct component of the input waveform. It follows that the previous analysis can be extended to the circuit and input waveform shown in Fig. 8.15(a). The source e_1 can be envisioned as a switch that alternately applies the constant voltages E_1 and E_2 to the circuit for the intervals δ_1 and δ_2, respectively. The polarity of e_2 (or the direction of current flow) determines which diode conducts, and therefore determines the time constant for charge or discharge. The resistances R_1 and R_2, together with the two ideal rectifiers, might represent a piecewise-linear approximation to a semi-conductor diode curve or similar nonlinear resistance.

For the input waveform shown, the circuit of Fig. 8.15(a) can be represented as in (b) during the interval δ_2, and as in (c) during the interval δ_1. Since these circuits differ in no respect from the linear circuit already analyzed, we can postulate the general shape of the waveforms of e_2 and e_x to be similar to those given in Fig. 8.14(c). We are considering here the most general case: the one for which the exponential transients do not reach completion in their respective intervals, thus E_y and E_x are unknowns. The exponential expressions for e_x are given in (b) and (c). Using the methods of Article 11, we find the expressions for capacitor voltage at the transitions to be

$$E_x = E_1 + \frac{EA_1(1 - A_2)}{1 - A_1A_2} \qquad (8.16)$$

$$E_y = E_1 + \frac{E(1 - A_2)}{1 - A_1A_2} \qquad (8.17)$$

As before, the end points of the waveform are fixed when E_y and E_x are

known. The upward and downward jumps in e_2 are the same as the jumps in e_1. Between initial and final values, the exponential variation

$\tau_1 = R_1C \quad A_1 = \epsilon^{-\delta_1/\tau_1}$
$\tau_2 = R_2C \quad A_2 = \epsilon^{-\delta_2/\tau_2}$

(a)

$\bar{e}_1 = (E_1\delta_1 + E_2\delta_2)/(\delta_1 + \delta_2)$
$\quad = E_1 + E\delta_2/(\delta_1 + \delta_2)$

$e_x = E_x$ at $t = 0$

During interval δ_2

$(E_2 - e_x) = (E_2 - E_x)\epsilon^{-t/\tau_2}$
$(E_2 - E_y) = (E_2 - E_x)A_2$

(b)

$e_x = E_y$ at $t = 0$

During interval δ_1

$(e_x - E_1) = (E_y - E_1)\epsilon^{-t/\tau_1}$
$(E_x - E_1) = (E_y - E_1)A_1$

(c)

(d) $E_x = \dfrac{E_2A_1(1 - A_2) + E_1(1 - A_1)}{1 - A_1A_2} = E_1 + \dfrac{EA_1(1 - A_2)}{1 - A_1A_2}$

$E_y = \dfrac{E_2(1 - A_2) + E_1A_2(1 - A_1)}{1 - A_1A_2} = E_1 + \dfrac{E(1 - A_2)}{1 - A_1A_2}$

(e) $\bar{e}_2 = \dfrac{C(E_y - E_x)R_2 - C(E_y - E_x)R_1}{\delta_1 + \delta_2} = C(E_y - E_x)(R_2 - R_1)f_r$

$\bar{e}_x = \bar{e}_1 - \bar{e}_2$
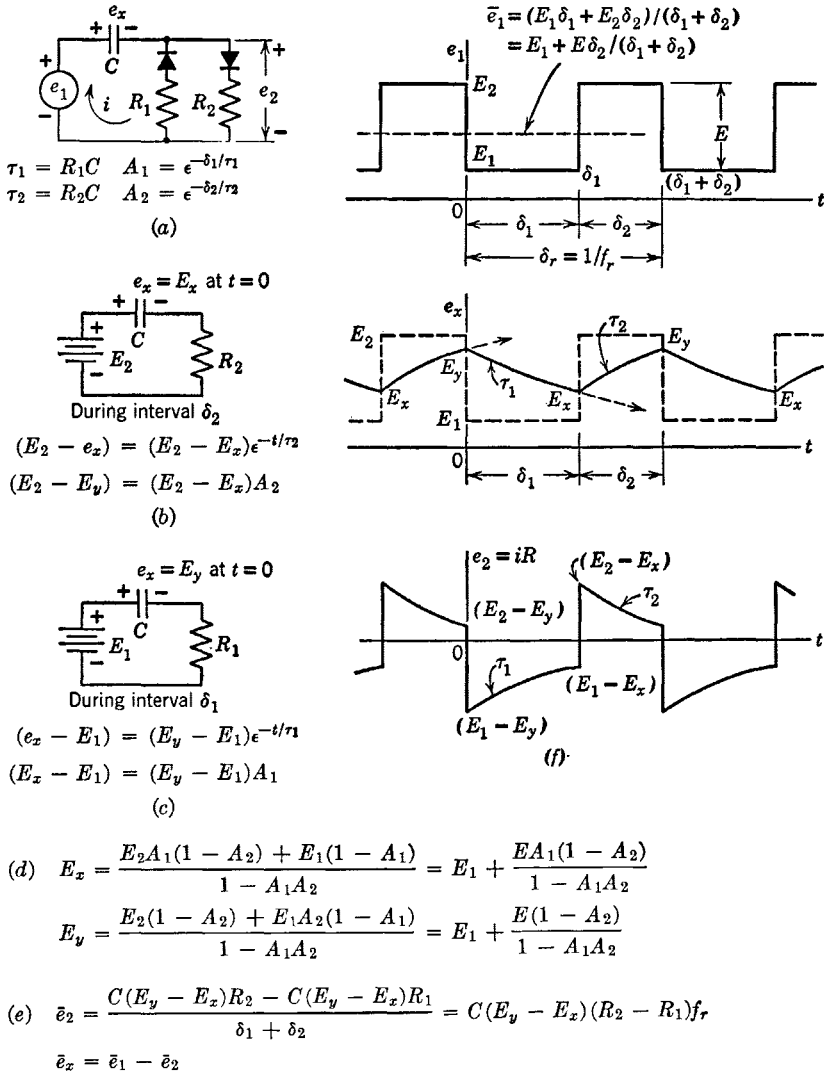
Fig. 8.15. Rectangular-wave response of a piecewise-linear RC circuit.

is in accordance with the appropriate time constant (τ_1 during δ_1 and τ_2 during δ_2).

The average current through the capacitor must still be zero. How-

ever, there is no longer any restriction on the average value of the output voltage e_2, which now is

$$\overline{e_2} = \frac{1}{\delta_r} \left[\int_0^{\delta_1} iR_1 \, dt + \int_{\delta_1}^{\delta_1 + \delta_2} iR_2 \, dt \right] \qquad (8.18)$$

But

$$\int_0^{\delta_1} i \, dt = \Delta Q_{(\text{discharge})} = C(E_x - E_y) \qquad (8.18a)$$

and

$$\int_{\delta_1}^{\delta_1 + \delta_2} i \, dt = \Delta Q_{(\text{charge})} = C(E_y - E_x) \qquad (8.18b)$$

Substituting Eqs. 8.18a and 8.18b in Eq. 8.18:

$$\overline{e_2} = f_r(R_2 - R_1)(E_y - E_x)C \qquad (8.19)$$

Since E_y is greater than E_x, this average value will be positive for R_2 greater than R_1 and negative for R_2 less than R_1.

The calculations outlined in Fig. 8.15 can be simplified for the special cases in which either or both exponential transients essentially reach completion. The function $\epsilon^{-t/\tau}$ is unity at $t = 0$, and 0.006 at $t = 5\tau$. Thus we can consider the transient to be essentially complete after an interval of five time constants. Waveforms for three examples are shown in Fig. 8.16. When $\delta_1/\tau_1 > 5$, as in Fig. 8.16(b), we have $E_x = E_1$ and, hence, E_x is known immediately. The value of E_y can then be determined directly from the charge interval. Similarly, when $\delta_2/\tau_2 > 5$, as in Fig. 8.16(c), we have $E_y = E_2$ and, hence, E_y is known, and E_x can be determined easily. For the waveforms of Fig. 8.16(d), $\delta_1/\tau_1 > 5$ and $\delta_2/\tau_2 > 5$; hence the initial and final values for both intervals are known, and the exponentials can be sketched directly by using the known end points and the known time constants.

8.13   Other Simple RC and RL Circuits

The results of the foregoing analyses can be extended to any RC circuit that can be reduced to the same form by means of Thevenin's theorem or other network transformations. Since piecewise-linear circuit models are linear within a given state of the ideal rectifiers, these methods of linear circuit analysis can also be applied to the individual states of piecewise-linear circuits. For example, the previous results can be applied to any circuit which includes one linear capacitor and any number of linear or two-segment piecewise-linear resistances. Alternatively, the circuit may include one such resistance and any number of
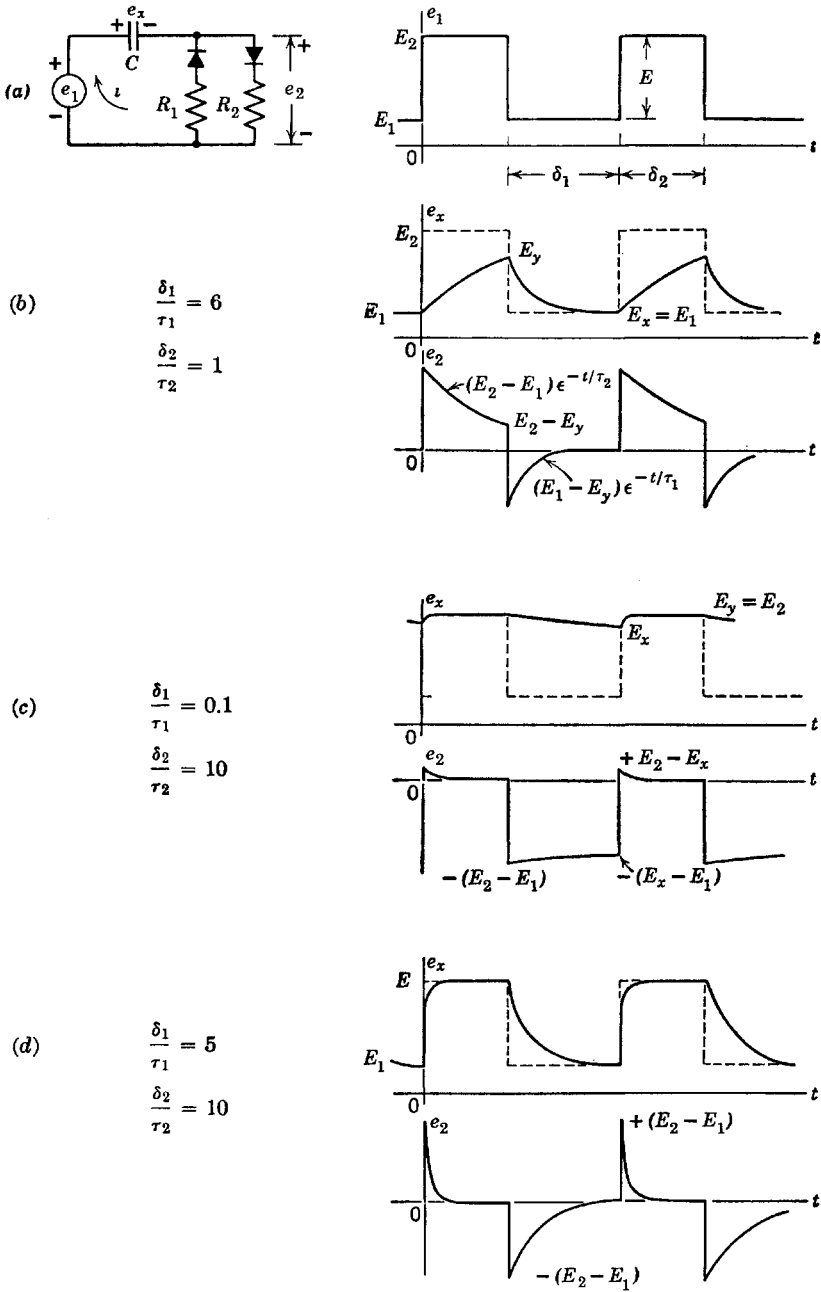
Fig. 8.16. Typical waveforms for piecewise-linear circuit with rectangular-wave input.
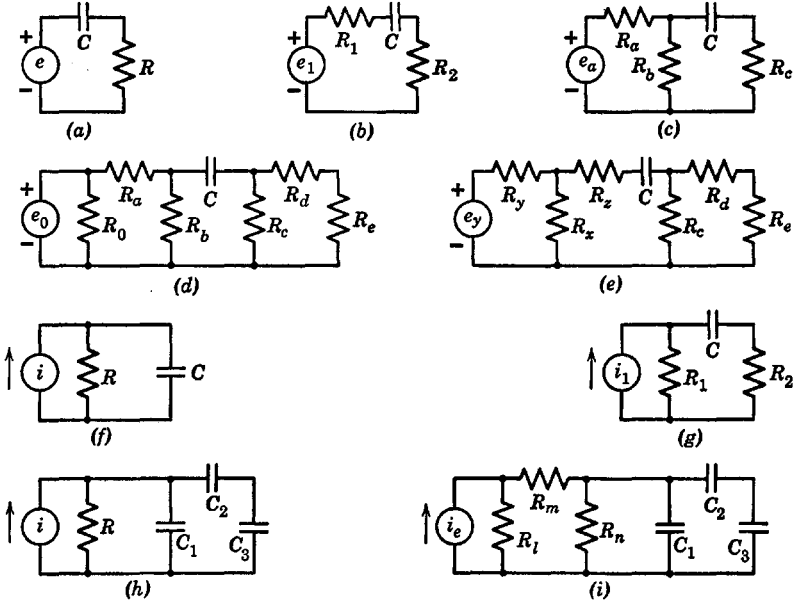
Fig. 8.17. Variations of the simple RC circuit.

For equivalence;

$$e = iR = e_0 + i_0R$$

or $\quad i = \dfrac{e}{R} = \dfrac{e_0}{R} + i_0$

(a)

For equivalence;

$$e = \frac{1}{C}\int i\,dt = e_0 + \frac{1}{C}\int i_0\,dt$$

or $\quad i = C\,de/dt = C\dfrac{de_0}{dt} + i_0$

(b)

For equivalence;

$$e = L\frac{di}{dt} = e_0 + L\frac{di_0}{dt}$$

or $\quad i = \dfrac{1}{L}\int e\,dt = \dfrac{1}{L}\int e_0\,dt + i_0$

(c)
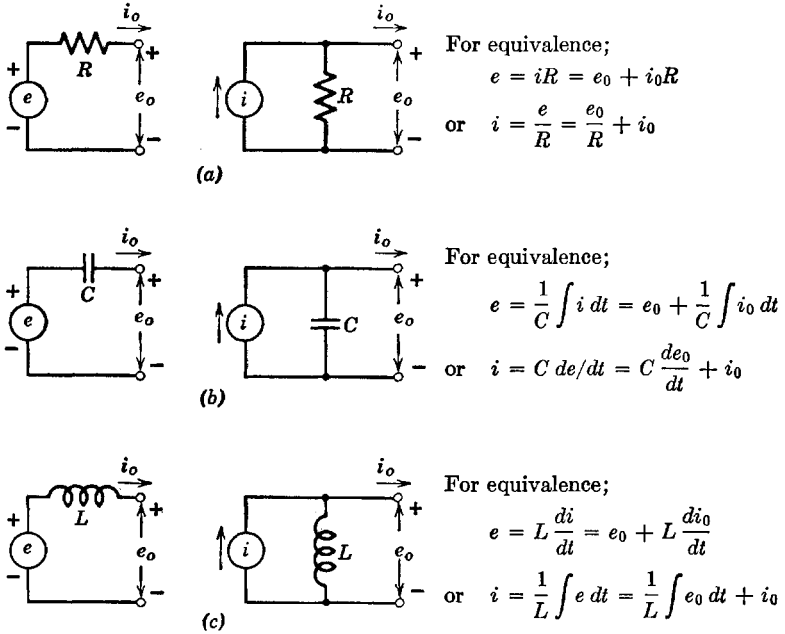
Fig. 8.18. Summary of equivalent two-element linear source models.

linear capacitances. Several examples of circuits which can be reduced to the series RC circuit form are shown in Fig. 8.17. The Thevenin equivalents shown in Fig. 8.18 can be used to reduce the circuits in Fig. 8.17(f), (g), (h), (i) to the original voltage-source forms.

Circuit Duals

Voltage	Current
Resistance	Conductance
Inductance	Capacitance
Series	Parallel
Loop	Node

(a)

$$e = Ri + L\frac{di}{dt}$$

$$i = Ge + C\frac{de}{dt}$$

(b)

$$e = Ri + \frac{1}{C}\int i\,dt$$

$$i = Ge + \frac{1}{L}\int e\,dt$$

(c)

$$e = L\frac{di}{dt} + Ri + \frac{1}{C}\int i\,dt$$

$$i = C\frac{de}{dt} + Ge + \frac{1}{L}\int e\,dt$$
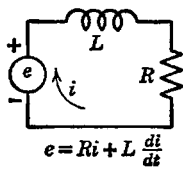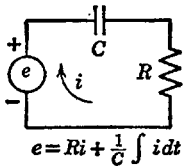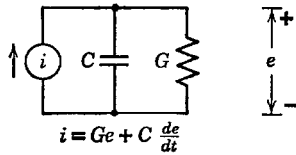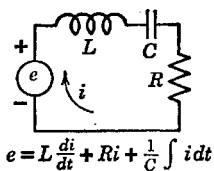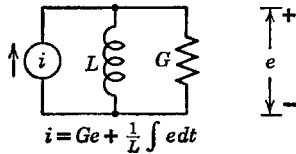
(d)
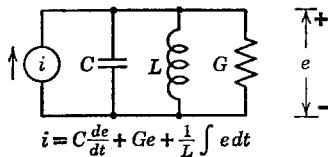
Fig. 8.19. Duality in simple circuits.

The methods thus far applied to RC circuit analysis can be used equally well to analyze circuits containing a linear inductance and linear or piecewise-linear resistances. The student may wish to repeat some of the analyses for simple RL circuits in order to gain familiarity with the

methods. However, it is possible to apply the concept of duality* to extend the results of RC circuit analysis to RL circuits. Duality makes it possible to transfer the results of one analysis to a different situation merely by appropriate definition of dual quantities. A few useful pairs of dual quantities and concepts are listed in Fig. 8.19, together with dual pairs for simple three-element and four-element circuits. Any of the dual pairs in the columns of Fig. 8.19(a) can be interchanged.

8.14   Graphical Analysis of Nonlinear Circuits

The nonlinear analysis of a simple diode circuit with rectangular-wave input is outlined in Fig. 8.20. Consider diode D in Fig. 8.20(a) to be a semiconductor junction diode with the curve $i = f(e_2)$, as shown in (b). The voltage equation for the circuit is

$$e_1 = e_r + e_x + e_2 \qquad (8.20)$$

For a specific value of input voltage, say E_b, and an arbitrary value e_x for the capacitor voltage, the graphical solution of the voltage equation specifies the operating point. The load-line form of the voltage equation is

$$\frac{(e_1 - e_x) - e_2}{R} = i \qquad (8.21)$$

In this form, $(e_1 - e_x)$ is the effective driving voltage on the series combination of R and D. With a rectangular wave input, e_1 remains constant during each interval δ_a or δ_b, but e_x is a function of time; hence the load line moves with time. If the capacitor charges and discharges completely during the intervals δ_b and δ_a, the capacitor voltage varies between $-E_a$ and $+E_b$. The voltage across the resistive elements R and D varies, therefore, from $(E_b + E_a)$ to zero during the interval δ_b, and from $-(E_b + E_a)$ to zero during δ_a. Correspondingly, the current and therefore the diode operating point moves from point Q to point 0, and from point P to point 0, as indicated in Fig. 8.20(b).

In carrying out a graphical integration process, we are essentially replacing differentials by increments: namely, replacing $i = C \, de_x/dt$ by $i = C \, \Delta e_x/\Delta t$. We assume that current i remains constant for an increment of time Δt, during which e_x changes by the amount Δe_x. This results in a new value of current which is held constant for a second increment of time during which a second incremental change occurs in

---

* E. A. Guillemin, *Introductory Circuit Theory*, John Wiley and Sons, New York, 1953, pp. 42–51, 241–243, 251–253, 373–374, 535–539.

e_x. Repetition of this process determines e_x as a smooth waveform which is a piecewise-linear approximation to the true waveform of e_x. The waveform of current i is a stepwise approximation. Carrying out the
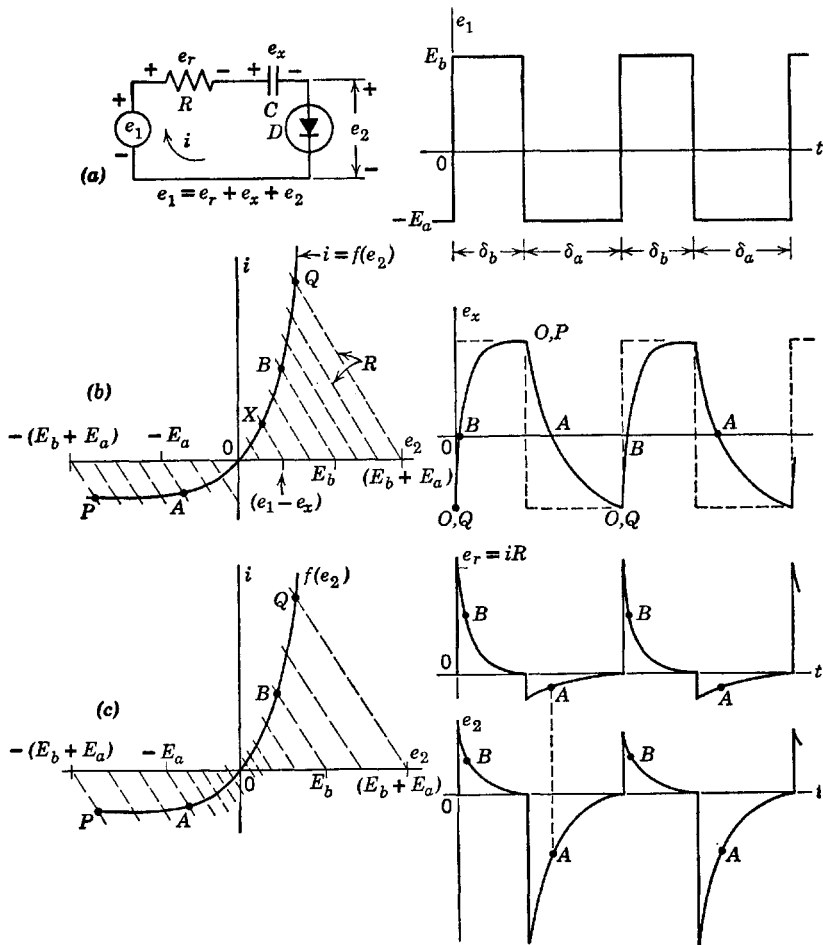
Fig. 8.20. Graphical analysis of a nonlinear diode circuit.

analysis in systematic fashion, we can take equal increments for Δe_x, as indicated by the sequence of dashed lines in Fig. 8.20(b); or we can take equal increments in Δt, as implied by Fig. 8.20(c). The former provides a better approximation to the true waveforms near the beginning of each interval, while the latter provides more points near the end. A combination of the two can also be used. In any case, it is readily seen that

the process is time-consuming even for this extremely simple circuit problem. To obtain reasonable accuracy, the nonlinear curve and waveform plots must be constructed carefully and drawn to a fairly large scale.

8.15   Triode with Parallel *RC* Plate Load
and Large-Amplitude Rectangular-Wave Input

The instantaneous values of terminal current and voltage specify the locus of the operating point for a control valve like a transistor, triode or pentode. When a circuit contains energy-storage elements, the locus depends on time scale and amplitude of the input waveform as well as the circuit parameters. Although waveforms of current and voltage [$i_b(t)$ and $e_b(t)$ for a triode] contain all the necessary information, the locus of the instantaneous operating point (i_b vs. e_b) plotted directly on the i_b vs. e_b plane aids in visualizing the conditions imposed on the valve by the circuit and input signal. The relations between the path of operation of a triode and the current and voltage waveforms will be illustrated in terms of a simple plate-loaded triode with a capacitor in parallel with the load resistor. The same methods will then be applied to other basic circuits. For continuity and ease of comparison, the vacuum triode is used in a series of related examples of waveshaping circuits. The methods and circuits apply equally well to other control valves such as transistors, pentodes, etc.

The two forms of the circuit shown in Fig. 8.21 (a) and (b) are identical insofar as the triode is concerned. There is only a difference in the direct component of the capacitor voltage for any particular set of circuit conditions. Since the physical behavior of the two circuits is identical, we shall confine our remarks to the circuit in (b).

Suppose the triode in the circuit of Fig. 8.21 (b) is described by the plate curves shown in Fig. 8.21 (c). The quiescent voltage E_a for $e_1 = 0$ is determined by drawing the resistive load line R_b on the i_b vs. e_b plane. Let the rectangular waveform of e_1 shown in (d) have sufficient amplitude to result in plate-current cutoff during the interval δ_2, and assume both δ_1 and δ_2 to be of sufficient duration to permit complete charge and discharge of capacitor C. Under these conditions, the voltage and current waveforms will be as shown in Fig. 8.21 (e) and (f). The dotted line in (e) is the waveform of e_2 in the absence of the capacitor C. The locus of the operating point (i_b vs. e_b) for the triode is shown on the plate curves in Fig. 8.21 (c). The points designated as A, B, C, and D on the locus correspond to similarly marked points on the waveforms. The transitions from A to B and from C to D are instantaneous. The path
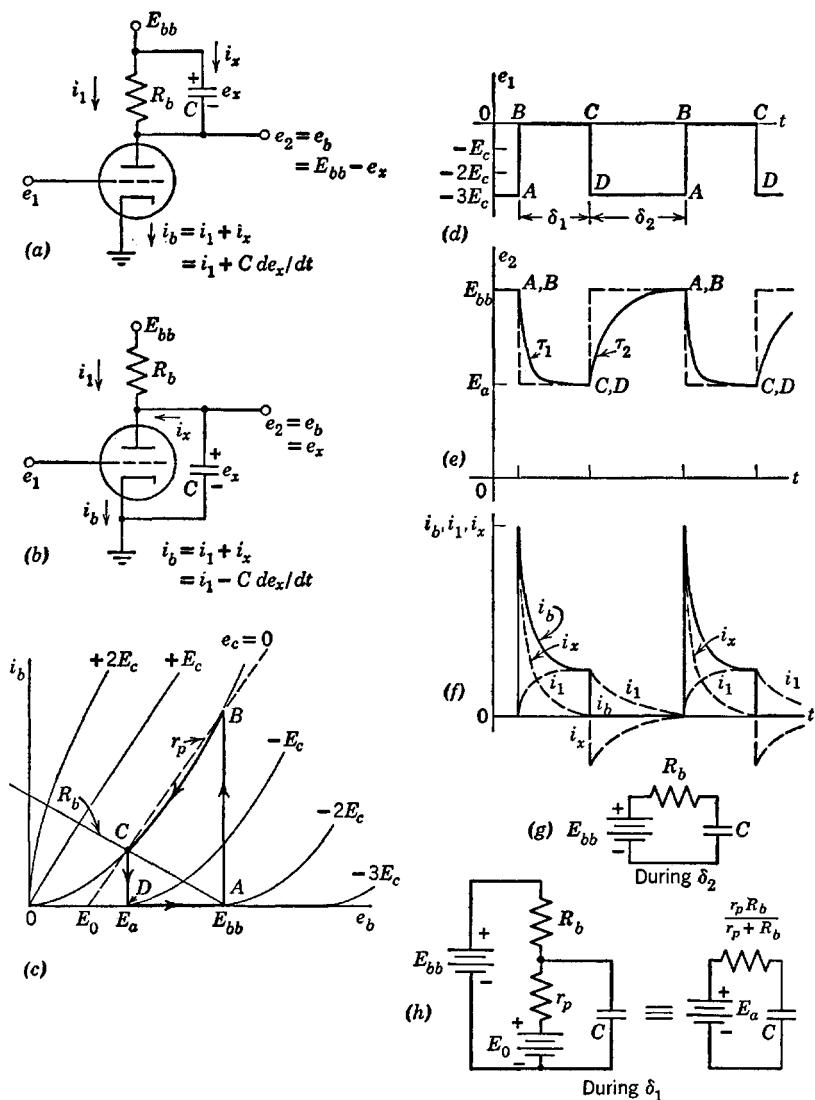
Fig. 8.21. Triode with parallel RC load. Rectangular-wave input
$$(\delta_1 > 5\tau_1, \ \delta_2 > 5\tau_2).$$

from B to C is traversed at a rate corresponding to the discharge of the capacitor through the triode plate resistance in parallel with load resistance R_b. The time along the path from D to A depends upon the exponential recharge of capacitor C to E_{bb} through R_b, as shown in

Fig. 8.21(g). The triode does not affect this portion of the path, since plate current is zero.

A step-by-step graphical calculation of the waveform of e_2 could be made during the interval δ_1. However, the use of a piecewise-linear
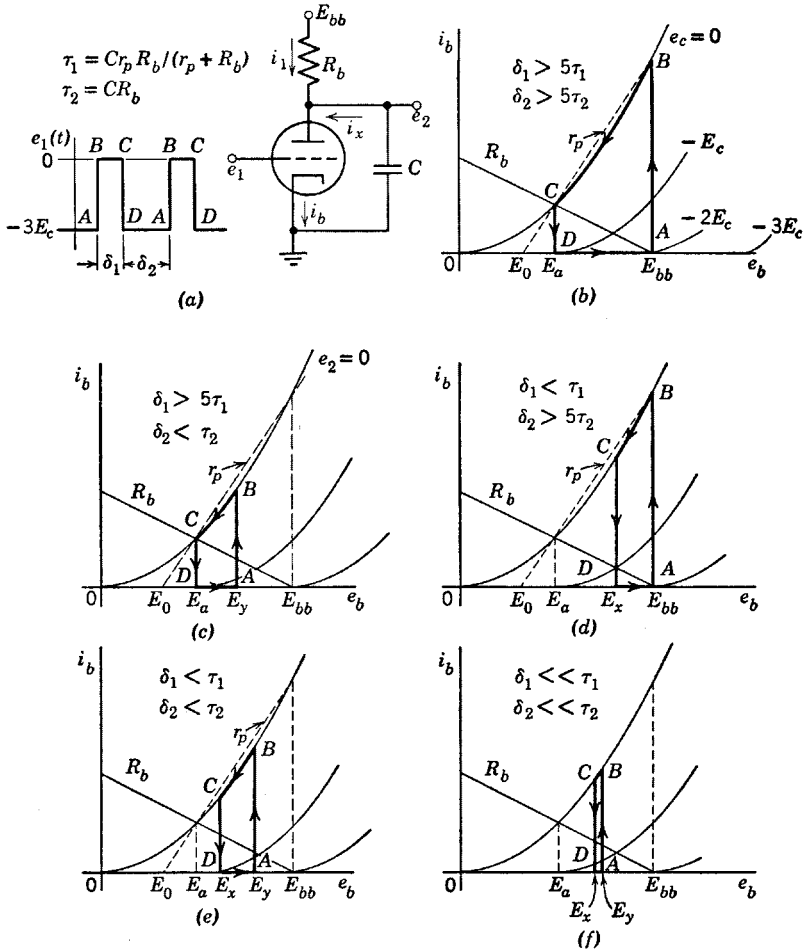
Fig. 8.22. Locus of operating point for various values of δ_1/τ_1 and δ_2/τ_2.

approximation to the triode yields a result that is well within the tolerances of the average plate characteristics. Let us approximate the $e_c = 0$ curve by passing a line [shown dotted on Fig. 8.21(c)] through points B and C, thereby matching initial points and end points on the waveforms. During δ_1 we now have the circuit model shown in

Fig. 8.21(h). We see, from Fig. 8.21(g) and (h), that the circuit is identical to the piecewise-linear RC circuit already analyzed.

Referring to Fig. 8.21(g) and (h), we see that the time constant of the exponential charge during interval δ_2 is

$$\tau_2 = R_b C \qquad (8.22)$$

For the discharge of capacitor C during interval δ_1, the time constant is

$$\tau_1 = C \frac{R_b r_p}{(r_p + R_b)} \qquad (8.23)$$

The sketches in Fig. 8.22 indicate the effects of changes in δ_1 and δ_2 relative to τ_1 and τ_2. The locus given in (b) is the same as that in Fig. 8.21(c)—namely, for complete charge and discharge. The locus shown in Fig. 8.22(c) is the intermediate case in which only the discharge goes to completion, and the locus in (d) is for complete charge only. The locus in (e) is the general case where neither charge nor discharge is completed. In the limit, as δ_1 becomes very small relative to τ_1, and δ_2 becomes very small relative to τ_2, the locus closes up, since E_y and E_x approach $\overline{e_2}$.

8.16   Effect of Rectangular-Wave Amplitude on a Triode with Parallel RC Plate Load

Let us now examine the locus of the operating point of the triode with parallel RC load for various values of the upper and lower values of the input waveform. For simplicity, the operating paths shown in Fig. 8.23 are sketched assuming $\delta_1 > 5\tau_1$ and $\delta_2 > 5\tau_2$ so that charge and discharge of the capacitor go to completion. The several input waveforms shown in the figure lead to the operating paths indicated on the plate curves. Note that, for each input waveform, the resistive output waveform (with capacitor removed) specifies the upper and lower bounds of capacitor voltage attainable. These bounds are indicated by the dotted line on the waveforms of e_2. For any one locus specified by the input waveform amplitude bounds, the effect of smaller intervals δ_1 and δ_2 will be to shrink the width of the locus, as in Fig. 8.22(c), (d), (e), and (f).

8.17   Triode with Parallel RC Plate Load and Sine-Wave Input

For the triode circuit with parallel RC load, the locus of the operating point when sinusoidal excitation is applied is closely related to the
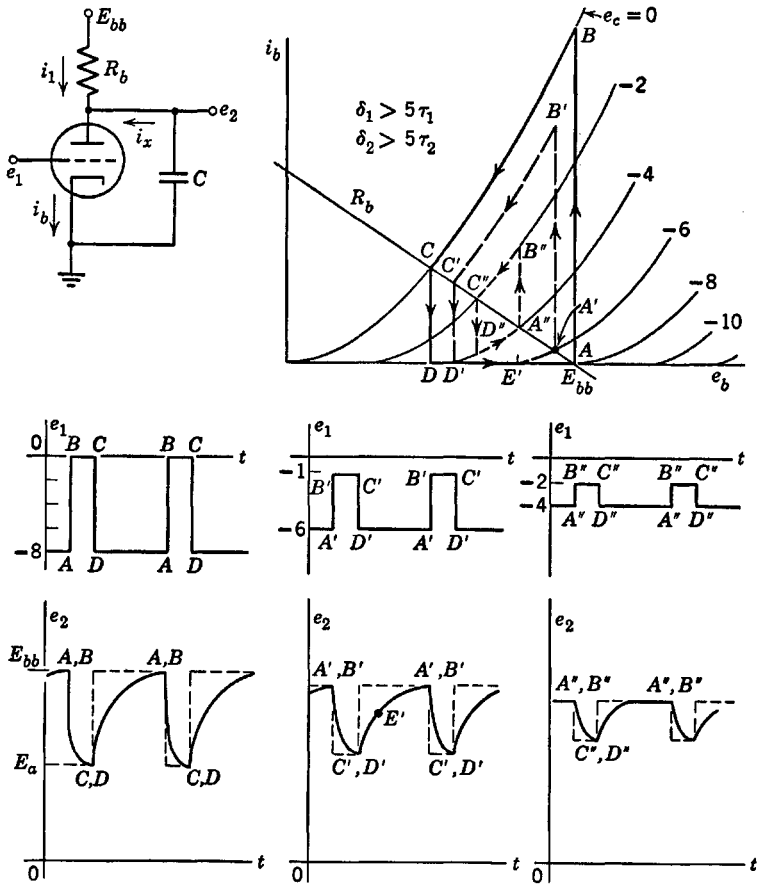
Fig. 8.23. Locus of operating point and waveforms for various values of input amplitude $(\delta_1 > 5\tau_1$ and $\delta_2 > 5\tau_2)$.

square-wave or rectangular-wave cases discussed in the preceding article. Let us consider an input voltage of the form

$$e_1 = -E_0 + E \sin \omega t \qquad (8.24)$$

where E_0 polarizes the grid negative with respect to cathode to a value about midway between zero and cutoff. Let us also assume E small enough to insure operation in the normal amplification region of the triode. The circuit and plate-current curves are shown in Fig. 8.24(a).

The locus of the operating point for a low-frequency square wave is indicated by the letters $ABCDA$, as before. If the square-wave frequency is raised to a sufficiently high value, the parallel dotted lines

AB and CD move together, and in the limit they merge into the single vertical line through quiescent point Q.* If each interval of the square wave is greater than five time constants, the locus reaches maximum size. Since the low-frequency square wave includes both very slow and

$$e_1 = -E_0 + E \sin \omega t$$
$$i_b = i_1 + i_x$$
$$= \frac{E_{bb} - e_b}{R_b} - \frac{C de_b}{dt}$$

(a)

(b)
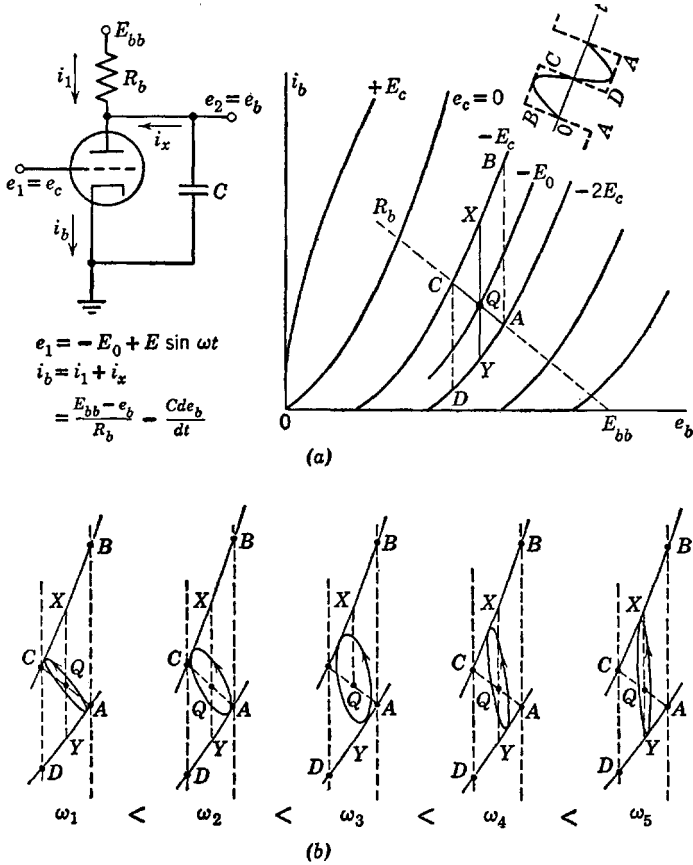
Fig. 8.24. Locus of operation for triode with parallel RC load and sine wave input. (a) Zero and infinite frequency locus of operating point. (b) Effect of successively increasing frequency.

very fast variations of input voltage, it actually represents many frequencies. During any one interval the square wave has a constant value and therefore changes as little as any other waveform could change

* This statement assumes linearity. Nonlinearities cause rectification and therefore a shift in the quiescent values (direct components).

within that interval. At each transition, the square wave changes as rapidly as any other waveform could. It can therefore be looked upon as the limiting condition for any waveform with the same upper and lower amplitude bounds.

If we now consider a sinusoidal input voltage with a very low frequency, ω, the capacitor current $i_x = -C \, de_2/dt$ is negligible, since the rate of change of voltage is negligible. The locus of the operating point for the limiting case as ω approaches zero is therefore along the R_b load line $QCQAQ$. For a sinusoidal input voltage of very high frequency, the locus will be the vertical line through point Q. Thus the high-frequency limit is the same for both sine waves and square waves.

So long as the input sine-wave amplitude is held constant, the locus for sine-wave operation must have a point of tangency on each of two lines of constant grid voltage. We can therefore postulate that the points of tangency must slide along from C to X and from A to Y as ω is increased from zero to infinity. Referring to Fig. 8.24(b), we see that for a very low frequency ω_1 the ellipse represents only a slight departure from the zero-frequency locus. For a slightly higher value ω_2, the ellipse has fattened, and the points of tangency have moved away from A and C. This tendency continues to ω_3. At ω_4, the ellipse has narrowed again, but the points of tangency have continued to move toward X and Y. At ω_5, the locus is beginning to approach the line XY, which is the limiting condition as ω approaches infinity. These qualitative considerations are extended by the following incremental analysis.

8.18 Incremental Analysis of the Triode with RC Plate Load and Sine-Wave Input

For small-signal operation in the vicinity of the quiescent point Q, shown in Fig. 8.24(a), the triode circuit can be represented by the incremental model given in Fig. 8.25(a). We are dealing with incremental quantities here instead of total quantities, and for convenience we shall use complex notation. If only output voltage and capacitor current are of interest, the equivalent model given in Fig. 8.25(b) simplifies calculations. The current-source model given in Fig. 8.25(c) is an alternative to the description given in (a) and the corresponding simplified equivalent is given in Fig. 8.25(d). These current-source models are almost always used as incremental models for pentodes since r_p is so high that it does not appreciably shunt R_b, and G simplifies to G_b.

The expressions for various voltages and currents in terms of e_g are
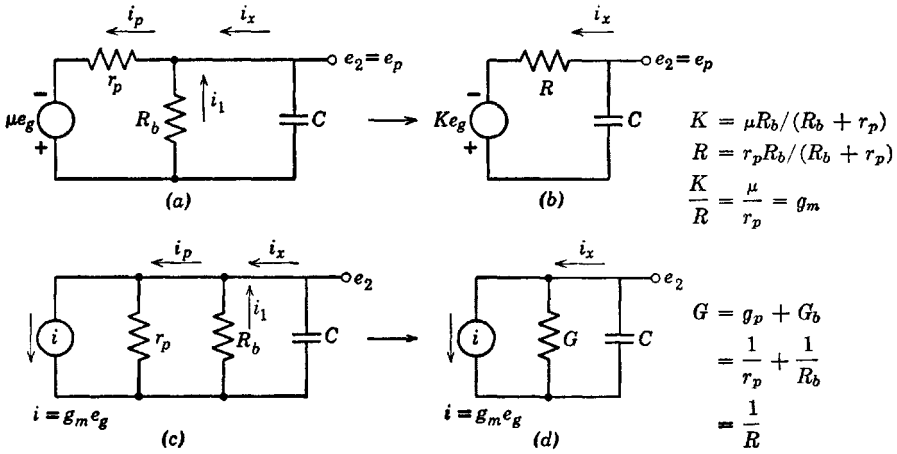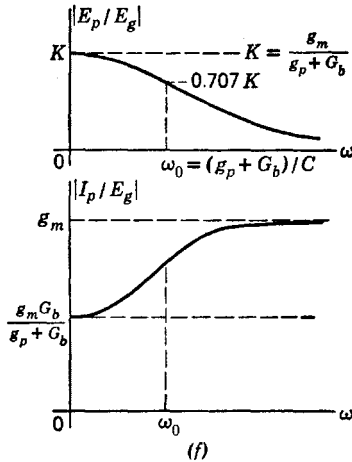
Fig. 8.25. Incremental analysis of triode with parallel RC load and sine wave input.

given in Fig. 8.25(e). Plots of $|E_p/E_g|$ and $|I_p/E_g|$ vs. ω are drawn in Fig. 8.25(f) for $G_b = g_p$. The voltage gain at $\omega = 0$ has a magnitude

$$K = \frac{\mu R_b}{R_b + r_p} = \frac{g_m}{g_p + G_b} \tag{8.25}$$

and diminishes to zero as ω approaches infinity. The bandwidth, ω_0, is defined as the point at which the gain E_p/E_g has diminished to $0.707K$,

This point occurs when

$$\omega_0 C = (g_p + G_b) \tag{8.26}$$

Since power gain is proportional to the square of voltage gain, this point is usually called the half-power point.
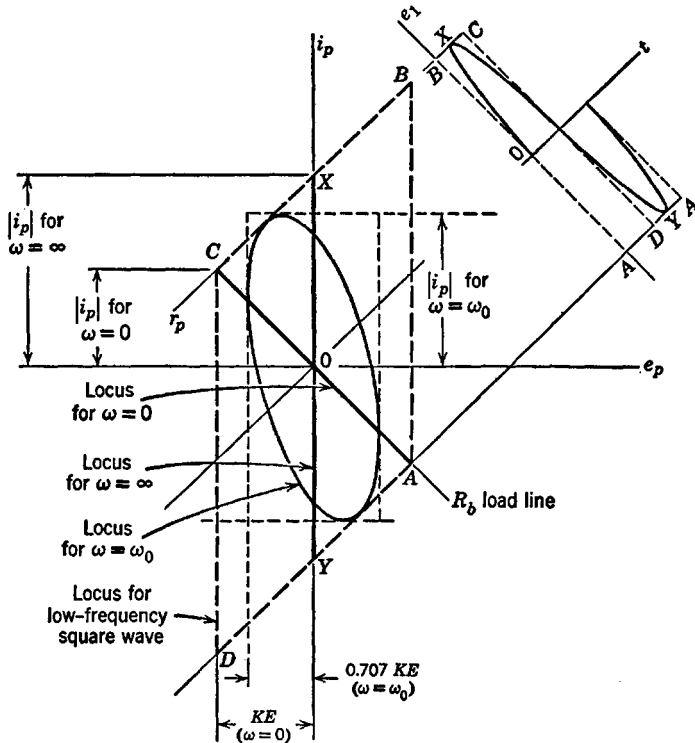
Fig. 8.26. Locus of operating point for incremental analysis.

In the locus plot (i_p vs. e_p) given in Fig. 8.26, the origin 0 for incremental quantities corresponds to the quiescent point Q for total quantities. This plot is drawn for $G_b = g_p$, and shows again the bounding locus $ABCDA$ corresponding to a square wave with a low repetition frequency and the same peak-to-peak amplitude as the sine wave.

The elliptical locus shown is drawn for $\omega = \omega_0$. The peak value of e_p is $0.707KE$, and the peak value of i_p is determined from the expression for i_p with $\omega = \omega_0$. The ellipse must be tangent to the lines of constant e_c corresponding to the peak-to-peak input voltage variation. The maximum horizontal excursion of the ellipse occurs on the R_b load line.

Since the plate voltage is not changing at the voltage peaks ($de_b/dt = 0$), the capacitor current is zero and the plate current is determined solely by R_b.

8.19 RC-Coupled Triodes with Rectangular Input Waveform (Intervals Large Compared with Time Constants)

The analysis of the triode with parallel RC load is readily extended to the RC-coupled amplifier. The coupling capacitor serves primarily to provide d-c isolation between coupled stages. This eases the problem of providing appropriate polarizing voltages for the electrodes of successive stages. If d-c isolation is the only function performed by the capacitor, the time constants are made large compared with the intervals, and hence the calculations are needed only to determine the distorting effects of the capacitor on the output waveform. In other cases the value of the capacitor is a design parameter calculated to produce a given wave-shaping effect.

In the circuit of Fig. 8.27 (a), we are merely connecting a resistance in series with the capacitor instead of connecting the capacitor directly to ground as in Fig. 8.23 or 8.24. The load on the plate circuit of the first triode is therefore the d-c load resistance R_b in parallel with a series combination of C and the piecewise-linear resistance R, where

$$R = R_g \qquad (e_c < 0) \qquad (8.27)$$

$$R = \frac{R_g r_g}{R_g + r_g} \qquad (e_c > 0) \qquad (8.28)$$

This piecewise-linear resistive load is indicated graphically by the sketch of i_x vs. e_2.

Pertinent equations for a graphical analysis of the circuit are given in Fig. 8.27 (b). The total tube current i_b is here expressed in terms of i_1 and i_x. The current i_1 is represented graphically by the d-c load line plotted in Fig. 8.27 (c). The capacitor current

$$i_x = \frac{e_x - e_b}{R} \qquad (8.29)$$

is represented by a piecewise-linear load line drawn from the voltage e_x, as indicated by dotted lines in Fig. 8.27 (c). Since the capacitor voltage is time-varying, this load line slides back and forth along the horizontal axis. With a rectangular input voltage, indicated by the waveform e_1, the maximum excursion of e_x lies between E_{bb} and E_a.

Let us now examine the operation of the circuit, assuming that the intervals δ_1 and δ_2 are long enough to permit completion of the RC transients. At point A the tube is cut off; hence $i_b = 0$, $e_b = E_{bb}$, $e_x = E_{bb}$, and $e_2 = 0$. The instantaneous transition to point B, where

(a)
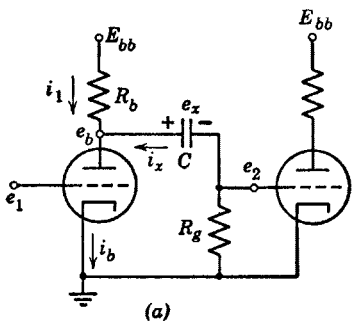
$$e_2 = e_b - e_x$$
$$i_b = i_1 + i_x$$
$$r_g \| R_g = \frac{E_{bb} - e_b}{R_b} + \frac{e_x - e_b}{R}$$

Where
$$R = R_g \text{ for } e_2 < 0$$
and
$$R = r_g \| R_g \text{ for } e_2 > 0$$

(b)

(c)
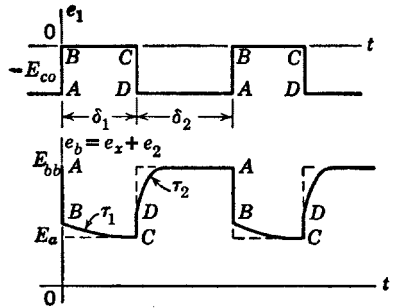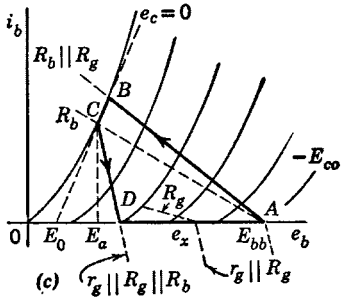
During δ_1
e_x discharges
from E_{bb} to E_a

During δ_2
e_x charges
from E_a to E_{bb}
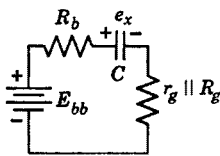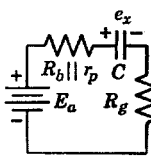
$\tau_1 = C(R_g + R_b\|r_p)$ $\tau_2 = C(R_b + r_g\|R_g)$

(d)   (e)

(f)
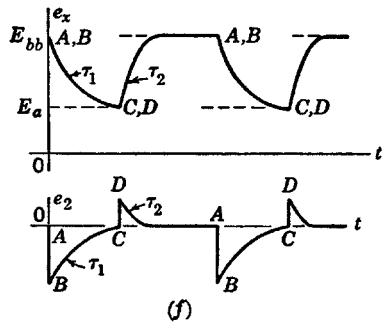
Fig. 8.27. Rectangular-wave behavior of RC coupling circuit ($\delta_1 > 5\tau_1$, $\delta_2 > 5\tau_2$).

$e_c = 0$, occurs along the load line AB. The change in plate voltage (Δe_b) must equal the change in the grid voltage Δe_2, since $\Delta e_x = 0$. Thus e_2 is driven negative.

The input voltage remains constant from B to C. Since the plate

voltage is less than the capacitor voltage, i_x is positive, and capacitor C discharges in accordance with the circuit shown in Fig. 8.27(d). As e_x decreases from E_{bb} to E_a, the R_g load line for i_x slides to the left, and the composite load line $R_b||R_g$ shifts downward parallel to itself. Since $e_c = 0$ during the interval δ_1, the operating point moves along the line $e_c = 0$ from B to C.

At the end of interval δ_1, e_1 changes instantaneously (C to D) to a value more negative than is required to effect plate-current cutoff. This change causes the plate voltage to rise toward E_{bb}. Again, since $\Delta e_x = 0$, we must have $\Delta e_b = \Delta e_2$ at the transition. The circuit given in Fig. 8.27(e) applies during the interval δ_2. For the instantaneous transition from C to D, the variational resistive load on the tube is $r_g||R_g||R_b$, because e_2 is positive during the interval δ_2. As the capacitor recharges, e_x rises from E_a to E_{bb} (D to A) to complete the operating cycle. The triode does not enter into the calculations during the interval δ_2, since $i_b = 0$ and therefore $i_x = -i_1$. The operating point moves along the $i_b = 0$ axis as the i_x load line slides to the right, and maintains $i_1 = -i_x$ as e_x increases.

The circuit models for the two intervals δ_1 and δ_2 shown in (d) and (e) are similar to those discussed in Art. 8.12, so the dimensions on the e_2 waveform in (f) may be found using the method discussed in that article. For the time constants chosen here, the e_2 waveform is a series of alternately positive and negative spikes at the transitions of the input waveform. If τ_1 and τ_2 are made very much smaller than δ_1 and δ_2, these spikes approximate the derivative of the input waveform. Under these conditions the circuit is also called a peaking circuit. In the following article we shall consider the long-time-constant case for which the transients do not go to completion.

8.20   RC Coupled Triodes with Rectangular Input Waveform (Intervals Small Compared with Time Constants)

Using the same circuit as that of Fig. 8.27(a), let us consider the case where δ_1 is less than $5\tau_1$ and δ_2 is less than $5\tau_2$. A piecewise-linear model of the circuit is shown in Fig. 8.28(a), and the two linear circuits that apply during the intervals δ_1 and δ_2 are shown in Fig. 8.28(b) and (c), respectively. The rectangular input waveform is again assumed to drive the first tube into the grid-current region during interval δ_1 and into the plate-current cutoff region during δ_2. Correspondingly, the capacitor voltage will charge to a value E_y, and discharge to a value E_x. These values can be calculated from the two linear circuits by the

method outlined in Fig. 8.15. Note that output voltage e_2 does not reach zero at the end of each interval, since the exponential charge and dis-
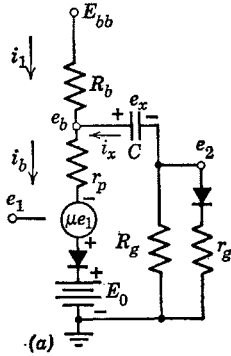
Open circuit value of e_b:

$$E_a = E_0 + \frac{(E_{bb} - E_0)r_p}{r_p + R_b}$$

Capacitor voltages E_x and E_y calculated as in Fig. 8.15

During δ_1
e_x discharges
from E_y to E_x

During δ_2
e_x charges
from E_x to E_y
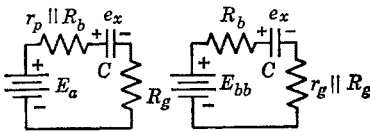
$$\tau_1 = C[R_g + r_p \| R_b] \qquad \tau_2 = C[R_b + r_g \| R_g]$$

(b) (c)
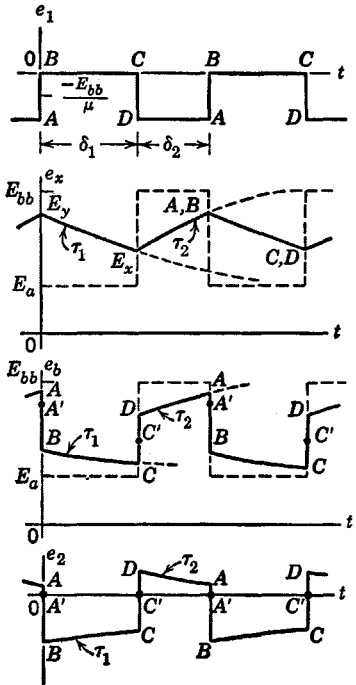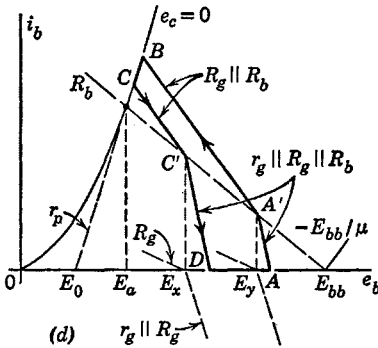
Fig. 8.28. Locus and waveforms for coupling circuit with rectangular-wave input ($\delta_1 < 5\tau_1$ and $\delta_2 < 5\tau_2$).

charge currents do not reach zero. For the same reason, e_b reaches neither E_a nor E_{bb}. These effects may be seen both from the waveforms and from the locus plot given in Fig. 8.28(d). Since e_x does not reach

its extreme values E_a and E_{bb}, the locus is restricted to a smaller portion of the i_b vs. e_b plane.

As e_1 switches from A to B, the plate voltage decreases. Since e_x remains constant, e_2 falls by the same amount as e_b. When e_2 crosses zero at point A', we have $e_b = e_x = E_y$. At this point the effective load resistance for the triode changes from $r_g||R_g||R_b$ to $R_g||R_b$. Note that whenever i_b exceeds i_1, the locus is above the R_b load line and i_x is positive. The piecewise-linear load lines shown by dotted lines at E_x and E_y represent the contribution of i_x to i_b for these specific values of e_x. For intermediate values of e_x, the load line for i_x slides along the horizontal axis. Correspondingly, the piecewise-linear load line determining total current i_b slides along R_b with the break moving from A' toward C' as time elapses from B to C. During this time, the operating point is sliding along the $e_c = 0$ line from B to C. When e_1 switches from C to D, the load transition at C' corresponds to $e_2 = i_x = 0$ and $e_b = e_x = E_x$. Closing the locus from D to A, we have e_x increasing from E_x to E_y, with $i_b = 0$ and $i_x = i_1$.

Determination of pertinent voltage values on the waveforms of e_b and e_2 follow readily, once E_x and E_y are calculated. The circuits do not differ from those of Fig. 8.15, except that the resistive voltage drop is divided into two parts here. Thus the total resistances of the charge and discharge circuits in Fig. 8.28 correspond to R_1 and R_2 in Fig. 8.15.

8.21   Linear Amplification with RC-Coupled Triodes

For the conditions described by Figs. 8.27 and 8.28, the waveform of e_2 has a negative average value because the resistance between grid and cathode of the second triode is R_g for negative values of e_2, and $r_g||R_g$ for positive values of e_2. This average voltage (sometimes called grid-leak bias) can be computed from E_x and E_y, as indicated in Fig. 8.15(e). The change in loading of the plate circuit of the first triode under these conditions, changes the voltage gain during the cycle and therefore results in waveform distortion. To amplify the input waveform without distortion, each triode must be operated in a single state: namely, with e_c between 0 and $-E_{bb}/\mu$. In order that this requirement may be satisfied, the amplitude of the input waveform must be suitably restricted, and a negative polarizing voltage must be applied to the grid of the second triode. In addition, the coupling time constant must be large enough to prevent distortion of the input waveform.

The circuit shown in Fig. 8.29(a) is like that of Fig. 8.27(a), with a polarizing voltage added for the grid of the second triode. This voltage
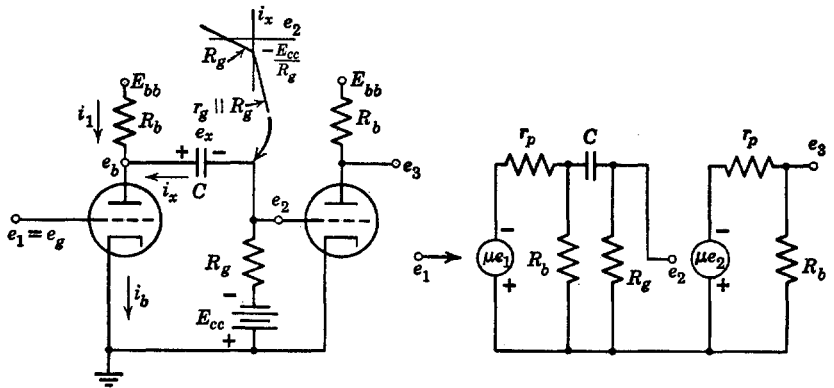
$$i_b = \frac{E_{bb} - e_b}{R_b} + \frac{(e_x - E_{cc}) - e_b}{R_g} \text{ for } e_2 < 0$$

(a) Linear amplifier circuit

For C very large $(\Delta e_x = 0)$ incremental gain is as follows:

$$\frac{e_3}{e_1} = \frac{e_2}{e_1} \cdot \frac{e_3}{e_2} = \frac{\mu(R_b \| R_g)}{r_p + (R_b \| R_g)} \cdot \frac{\mu R_b}{r_p + R_b}$$

(c) Incremental model
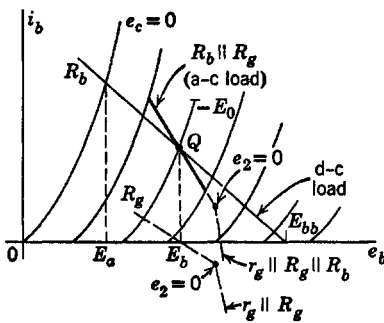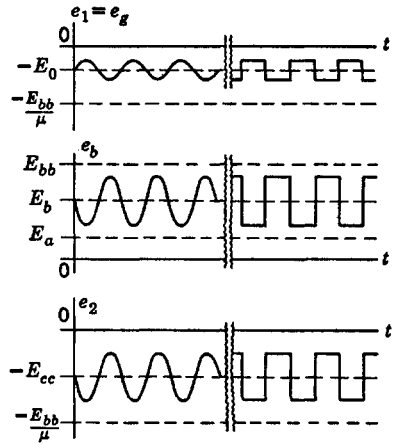
Fig. 8.29. Small-signal linear amplification with RC-coupled triodes. (a) Linear amplifier circuit; (b) Plate circuit of first triode for $e_1 = e_g = -E_0 + E \sin \omega t$; (c) Incremental model; (d) Waveforms (assuming e_x approximately constant at $E_b + E_{cc}$).

has the effect of shifting the current at the break point of the piecewise-linear load R to the point

$$i_z = -\frac{E_{cc}}{R_g} \qquad (8.29)$$

The locus plot for the plate circuit of the first triode, Fig. 8.29(b), assumes that the variational amplitude E is small enough to preserve linear operation throughout the circuit. When the piecewise-linear load line is plotted on the i_b vs. e_b plane, the break point occurs at a voltage $E_b + E_{cc}$. The polarizing voltage E_{cc} must be sufficiently negative to shift the break point of the broken load line outside the range of variation of e_b shown by the heavy line.

The incremental circuit model for the two triode stages is shown in Fig. 8.29(c). If the time constant

$$\tau = C[R_g + r_p \| R_b] \qquad (8.30)$$

is many times greater than the period of the input waveform, the changes in e_x will be very small during the cycle. In the limit, the capacitor voltage is constant (zero in the incremental circuit). In the circuit of Fig. 8.29(a) this amounts to replacement of the coupling capacitor C by a battery whose voltage is $E_b + E_{cc}$.

Waveforms for linear operation are given in Fig. 8.29(d) for both sinusoidal and square-wave input voltage variations. The waveform of e_3 (not shown) would be an inverted and amplified replica of e_2.

8.22 Effects of Shunt Capacitance on Coupling-Circuit Behavior

The model of the RC coupled circuit used in the preceding articles has represented only the explicit circuit components (vacuum tubes, resistors, and capacitors). However, any physical structure consisting of such components will also include implicit stray elements that have a significant effect on circuit operation when high frequencies or rapid changes are present in the input voltage. For example, the lines connecting elements on a circuit diagram usually signify ideal conductors that offer no impedance to the passage of current. Actual interconnecting wires have both resistance and inductance.

The major circuit elements themselves embody stray parameters (for example, the interelectrode capacitance of the triode). The capacitance between any component and the metal chassis (called ground) may also be important. The total capacitance between almost any node of the

circuit and the chassis will usually be several micromicrofarads for typical circuit construction. The principal effect of stray capacitances can be determined by simple approximate calculations.

An incremental model for the RC coupling circuit is shown in Fig. 8.30(a). This circuit assumes that the major effects of the stray elements are adequately represented by a capacitance C_1 from the plate of the first triode to ground, and by a capacitance C_2 from the grid of the second triode to ground. This model represents the actual circuit fairly well, particularly if the tubes used are pentodes.

The complex expression for E_2 in terms of E_1 and the circuit parameters is given in Fig. 8.30(a). Although the result can be written by inspection of the circuit, it is cumbersome for purposes of calculation or plotting. This expression can be reduced to three very simple expressions by making a few assumptions. In a practical circuit we shall nearly always have C much larger than C_1 or C_2. If C is large, it has a very low a-c impedance for all but low frequencies. If C_1 and C_2 are small, they have a very high impedance for all but high frequencies. We therefore expect to find a range of frequencies (usually called midband) for which the resistive elements alone yield a good approximation to e_2. Thus

$$E_2 = \frac{-g_m E_1}{g_p + G_g + G_b} \tag{8.31}$$

for ω in the vicinity of $\sqrt{\omega_c \omega_s}$. The circuit and the corresponding plot of $|E_2/E_1|$ vs. ω are shown in Fig. 8.30(b).

For values of ω approaching zero, we know that the reactance of coupling capacitor C cannot be neglected; hence the circuit and curve given in Fig. 8.30(c) are more representative of the true situation for low frequencies. The equation for E_2 becomes

$$E_2 = \frac{-g_m E_1 R_g}{(g_p + G_b)\left(R_g + \dfrac{1}{j\omega C}\right) + 1} \text{ for } \omega < \sqrt{\omega_c \omega_s} \tag{8.32}$$

Similarly, for high frequencies the shunting effect of the decreasing impedances of C_1 and C_2 cannot be neglected. The circuit and curve shown in Fig. 8.30(d) are therefore more accurate for high frequencies. Here

$$E_2 = \frac{-g_m E_1}{G + j\omega C_s} \text{ for } \omega > \sqrt{\omega_c \omega_s} \tag{8.33}$$

where

$$G = g_p + G_b + G_g \tag{8.34}$$

For a sinusoidal input,

$$E_2 = \frac{-g_m E_1 \left[\dfrac{1}{G_g + j\omega C_2}\right]}{\left[(g_p + G_b + j\omega C_1) + \left(\dfrac{1}{\dfrac{1}{j\omega C} + \dfrac{1}{G_g + j\omega C_2}}\right)\right]\left[\dfrac{1}{j\omega C} + \dfrac{1}{G_g + j\omega C_2}\right]}$$

$$(a)$$

$$E_2 = E_1 = -\frac{g_m E_1}{g_p + G_b + G_g}$$

$$(b)$$

$$\omega_c = \frac{1}{\tau_c} = \frac{1}{C[R_g + r_p \parallel R_b]}$$

$$E_2 = \frac{-g_m E_1}{g_p + G_b} \cdot \frac{R_g}{R_g + \dfrac{1}{j\omega C} + \dfrac{1}{g_p + G_b}}$$

$$(c)$$

$$\omega_s = \frac{1}{\tau_s} = \frac{G}{C_s}$$

$$E_2 = \frac{-g_m E_1}{G + j\omega C_s}$$

$$(d)$$
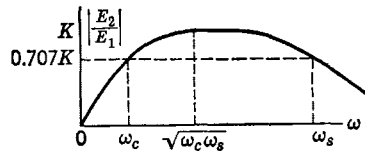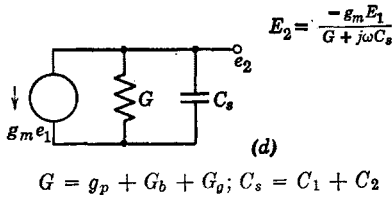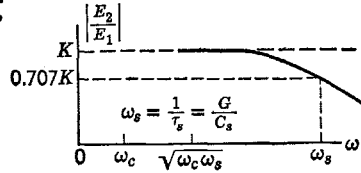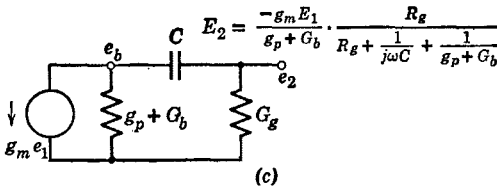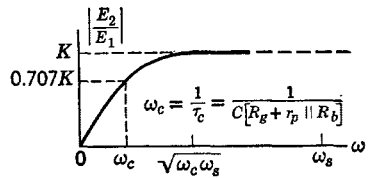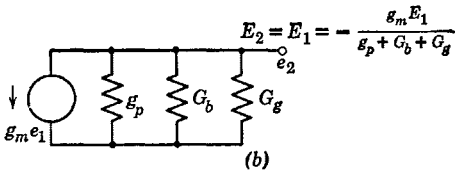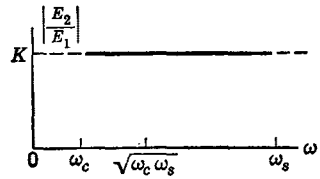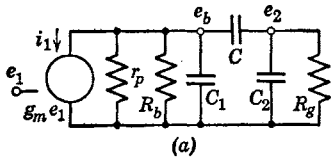
$$G = g_p + G_b + G_g; \quad C_s = C_1 + C_2$$

for $\omega > \sqrt{\omega_c \omega_s}$

(e) Combined effects of
   C and C_s for $\omega_s \gg \omega_c$

Fig. 8.30. Frequency response of the RC coupling circuit.

$$E_2 = \frac{-\mu E_1 (R_b \| R_g)}{r_p + (R_b \| R_g)} = \frac{-g_m E_1}{g_p + G_b + G_g}$$

$$(b)$$

$$e_2 = E_2 \epsilon^{-t/\tau_c}; \quad \tau_c = C[R_g + (r_p \| R_b)]$$

$$(c)$$

$$e_2 = E_2(1 - \epsilon^{-t/\tau_s}); \quad \tau_s = C_s/G$$

$$C_s = C_1 + C_2; \quad G = g_p + G_b + G_g$$

$$(d)$$

Combined effects of

C and C_s for $\tau_c \gg \tau_s$

$$(e)$$
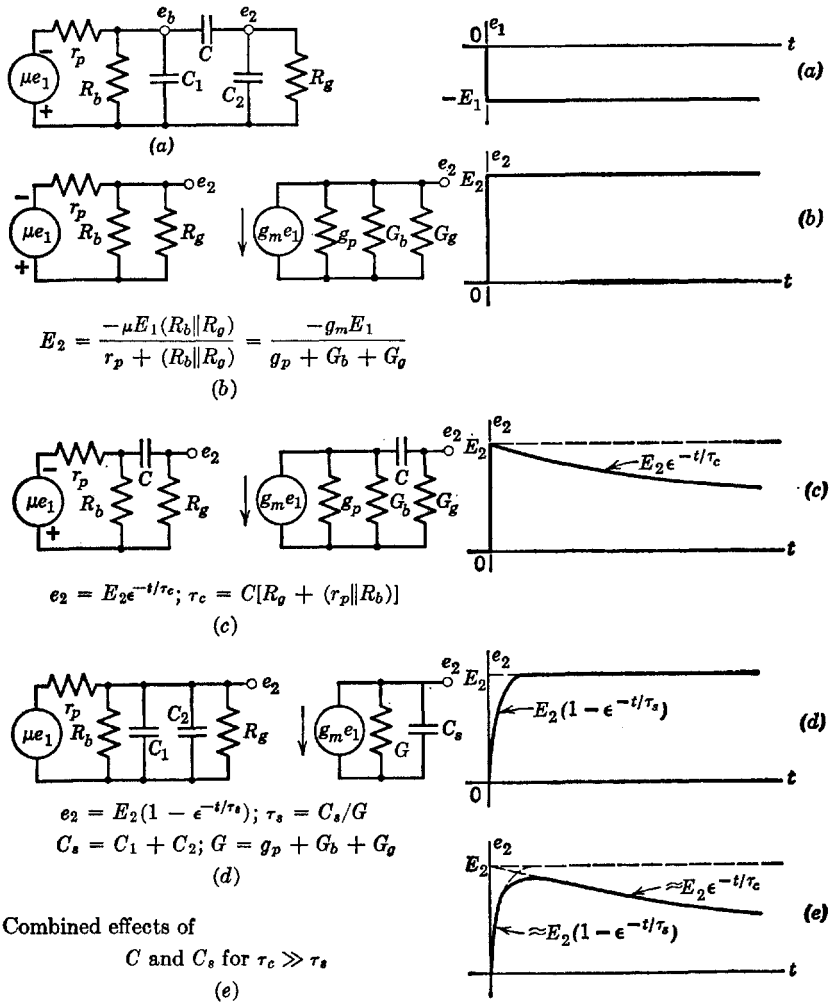
Fig. 8.31. Approximating the step response of a linear RC coupling circuit including stray capacitances.

The combined effects of coupling capacitor and shunt capacitors appear in the curve shown in Fig. 8.30(e). This curve, obtained very easily from the three separate circuits, is a good approximation to the curve obtained by plotting the expression given in Fig. 8.30(a). If C, C_1, and C_2 are comparable in magnitude, ω_c and ω_s are not widely separated, and the approximations are not as good.

The step response of the coupling circuit can be calculated in a similar manner, as outlined in Fig. 8.31. Since the circuit has two independent

capacitor voltages, we expect to find a transient solution with two exponential terms, each of which has a different time constant.

For the resistive approximation given in Fig. 8.31(b), the step is reproduced with no change other than a polarity reversal and a change in amplitude. The circuit with the coupling capacitor alone, Fig. 8.31(c), produces a slowly decaying exponential. The stray capacitances in the circuit of Fig. 8.31(d) inhibit the instantaneous rise of the step input. The combined effects of C and C_s are shown in Fig. 8.31(e); and for $C \gg C_s$, this is a very good approximation to the actual waveform. Calculating e_2 from the complete circuit yields an expression consisting of the sum of two exponentials with time constants slightly different from τ_c and τ_s. The exact plot would fall below the approximate one during the rapid rise and slightly above it during the long decay. Our approximation assumes no appreciable change in voltage on the coupling capacitor during the rise and neglects the current through the shunt capacitances during the decay.

8.23 Locus of Operation with Coupling and Shunt Capacitances

When C is much greater than C_1 and C_2, it is convenient to combine the two shunt capacitances into a single capacitance C_s where

$$C_s = C_1 + C_2 \qquad (8.35)$$

which appears only on one side of the coupling capacitor, as indicated in Fig. 8.32(a). This simplification makes only a minor difference in the voltage e_x across the coupling capacitor. The extreme error in e_x could not exceed $(C_2/C)(\Delta e_2)_{\max}$. This is the change in e_x required to change the voltage across C_2 over its maximum range of variation.

For the circuit of Fig. 8.32(a), with the input waveform shown, the locus of operation is indicated in (b), (c), and (d) for three values of coupling capacitor time constant. The d-c load line with slope R_b is shown as a dotted line. The a-c load line, or locus without C_s, is shown by the solid lines $AB'CD'$. With shunt capacitance C_s added to the circuit, instantaneous changes in e_b are prohibited. In Fig. 8.32(b), the operating point moves instantaneously from A to B, and then rapidly to point C as the shunt capacitance discharges. Conditions are then very nearly quiescent at point C until the grid voltage switches negative to D. This causes the instantaneous change CD and the rapid charge from D to A, where the operating point remains (approximately) until the next switch from A to B.

Actually, the coupling capacitor charges and discharges slightly and for a somewhat larger value of δ/τ_c the effect is more pronounced, as shown in Fig. 8.32(c). Here the operating point moves instantaneously from A to B, very rapidly from B to B', and then slowly from B to C.
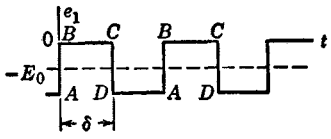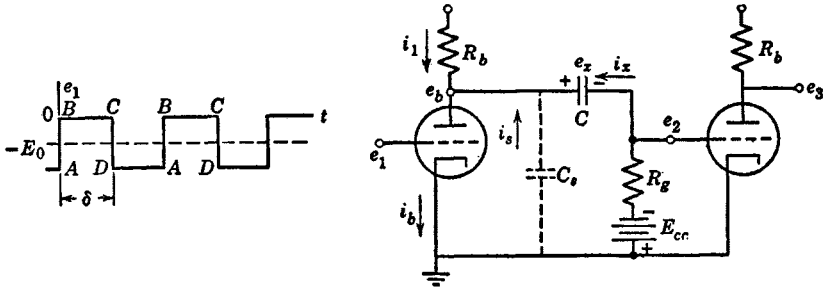
$$i_b = i_1 + i_x + i_s$$

$$i_b = \frac{E_{bb} - e_b}{R_b} + \frac{(e_x - E_{cc}) - e_b}{R_g} - C\frac{de_b}{dt}$$

for $e_2 < 0$ and $C \gg C_s$

$$\tau_c \approx C[R_g + r_p\|R_b]$$

$$\tau_s \approx C_s[r_p\|R_b\|R_g]$$

(a)

(b) $\quad \tau_c \gg \delta$ $\qquad$ (c) $\quad \tau_c \approx \delta$ $\qquad$ (d) $\quad \tau_c < \frac{1}{5}\delta$

$\tau_s \ll \delta$ $\qquad\qquad$ $\tau_s \ll \delta$ $\qquad\qquad$ $\tau_s \ll \delta$
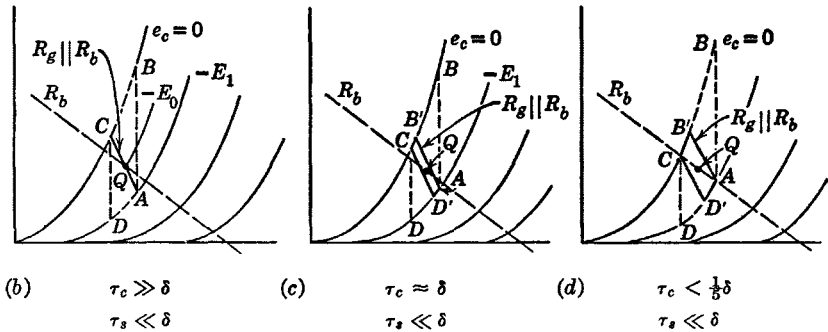
Fig. 8.32. Locus of operation with shunt capacitance and various degrees of coupling capacitor charge and discharge ($C \gg C_s$ in all cases).

When C is reached, there is an instantaneous transition to D, a rapid shift to D', and a slow shift back to A. Then the cycle repeats. The same general statements apply to Fig. 8.32(d), except that motion from B' to C and from D' to A is relatively more rapid, since δ/τ_c is even larger. Also, since the coupling capacitor may charge and discharge to completion, the operating point may repose at C and A for some time prior to each switching transition. In this case, C and A are true quiescent points (on R_b load line), whereas in Fig. 8.32(b) these points were only "quasi-quiescent" as a result of the small value of δ/τ_c.

The methods we have used for determining waveforms and loci of

operation apply to any of the basic circuits with inductive or tuned loads as well as capacitive loads. A few of the many possible combinations will be discussed in the following articles.

$$\tau_1 = \frac{L}{r_p \| R_b} \qquad \tau_2 = \frac{L}{R_b}$$

(a)

(b)

(c)
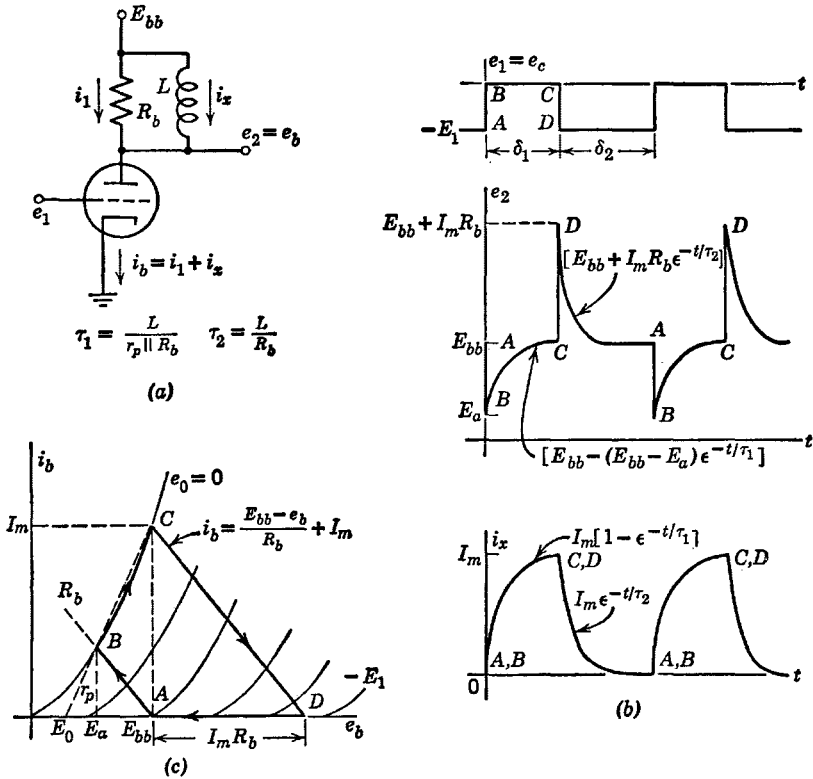
Fig. 8.33. Triode with parallel RL plate load and rectangular-wave input
$(\delta_1 > 5\tau_1, \delta_2 > 5\tau_2)$.

8.24   Triode with Parallel RL Plate Load

A triode with parallel RL load is shown in Fig. 8.33(a). Let the input waveform be as shown in (b), with intervals long compared to time constants. When triode conduction is initiated by the transition from A to B, the current in the inductance is zero; hence the apparent triode load is R_b, as indicated by the locus plot in Fig. 8.33(c). As current builds up in the inductance with a time constant

$$\tau_1 = \frac{L}{R_b \| r_p} \tag{8.36}$$

(if we use a piecewise-linear model for the triode), the voltage across the load decreases toward zero and plate voltage increases toward E_{bb}. The triode current during this time is

$$i_b = \frac{E_{bb} - e_b}{R_b} + i_x \qquad (8.37)$$

The build-up of i_x corresponds to a translation of the R_b load line parallel to itself, so that the operating point moves along the $e_c = 0$ line from B to C. In effect, i_x is a time-varying current source acting on the resistive portion of the circuit. When i_x has reached the quiescent value I_m, the voltage across the load is zero and $e_b = E_{bb}$. During the transition from C to D, the inductor current remains constant at I_m; hence the variational load again appears to be R_b. The current I_m is transferred from the triode to R_b as the triode cuts off, and results in an overshoot of e_b above the value E_{bb} by the amount $I_m R_b$. The current now decays exponentially to zero with the time constant

$$\tau_2 = \frac{L}{R_b} \qquad (8.38)$$

If the intervals δ_1 and δ_2 are not sufficiently long to permit completion of transients, we must solve a pair of simultaneous algebraic equations as in the RC case. The current i_x will vary between a maximum of I_y and a minimum of I_x instead of I_m and zero. Once these values have been found from the two linear circuits that apply during δ_1 and δ_2, respectively, the waveforms or locus of operation can be plotted very readily. The analysis is similar to the one outlined in Fig. 8.15 for the RC circuit.

The waveforms and locus of operating point shown in Fig. 8.33 apply for an input waveform that has a sufficiently negative peak $(-E_1)$ to drive the triode to plate-current cutoff at the transition from C to D. This requirement specifies a value of e_1 more negative than $-(E_{bb} + I_m R_b)/\mu$. For a negative excursion less than this value the triode plate current is not driven to zero instantaneously.

The locus and waveforms in Fig. 8.34(a) and (b) apply to the circuit of Fig. 8.33 for $-E_1 = -E_{bb}/\mu$. Assuming the intervals δ_1 and δ_2 long enough to permit completion of transients, the waveforms of e_2 and the locus are the same as before for the transition A to B and the interval δ_1 from B to C. However, since triode plate current is not cut off by the transition from C to D, the positive overshoot on the waveform of e_2 has a time constant $\tau_1 = L/(R_b \| r_p)$. In the previous example the time constant was $\tau_2 = L/R_b$. The exponential decay of inductor current

i_x (waveform not shown) also has the time constant τ_1. For $-E_1 = -E_{bb}/\mu$, as in this example, the triode plate current just reaches cutoff at the end of the overshoot exponential (D to A).
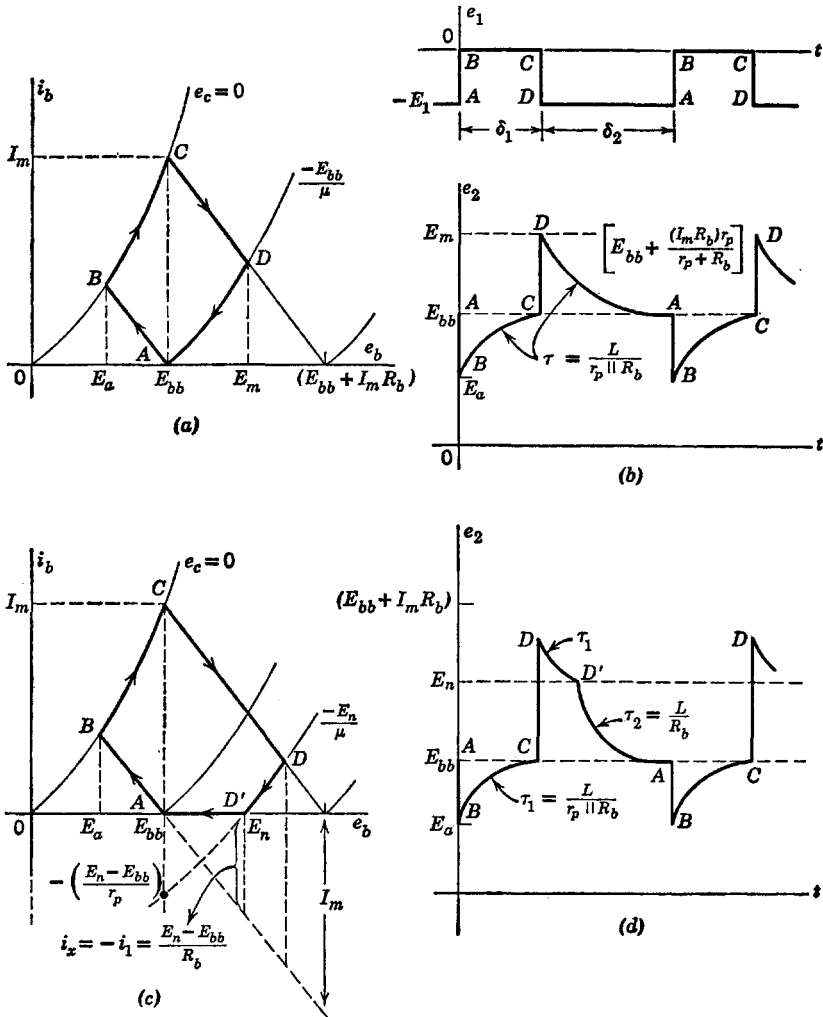
Fig. 8.34. Effect of input amplitude on triode with parallel RL load
$$(\delta_1 > 5\tau_1, \quad \delta_2 > 5\tau_1).$$

The locus and waveforms shown in Fig. 8.34(c) and (d) have been drawn for a value of $-E_1$ between $-E_{bb}/\mu$ and $-(E_{bb} + I_m R_b)/\mu$. For this intermediate case the overshoot waveform consists of an exponential with a time constant τ_1 from D to D' at which point the plate

current cuts off and the time constant is τ_2 from D' to A. The current i_x in the inductance (waveform not shown) decreases from I_m with a time constant τ_1 and is heading toward a final value $-(E_n - E_{bb})/r_p$. [See locus plot in Fig. 8.34(c).] However, when plate current reaches zero, we have $i_x = -i_1 = (E_n - E_{bb})/R_b$ and from this value i_x heads toward zero with a time constant τ_2. Although the slope of waveform e_2 changes at the break point, di_x/dt must be continuous at the break, since e_2 is continuous.

8.25   Triode with Series *RL* Plate Load

With a series resistance and inductance as the plate load on a triode, the response to a rectangular wave is as shown in Fig. 8.35. The effects
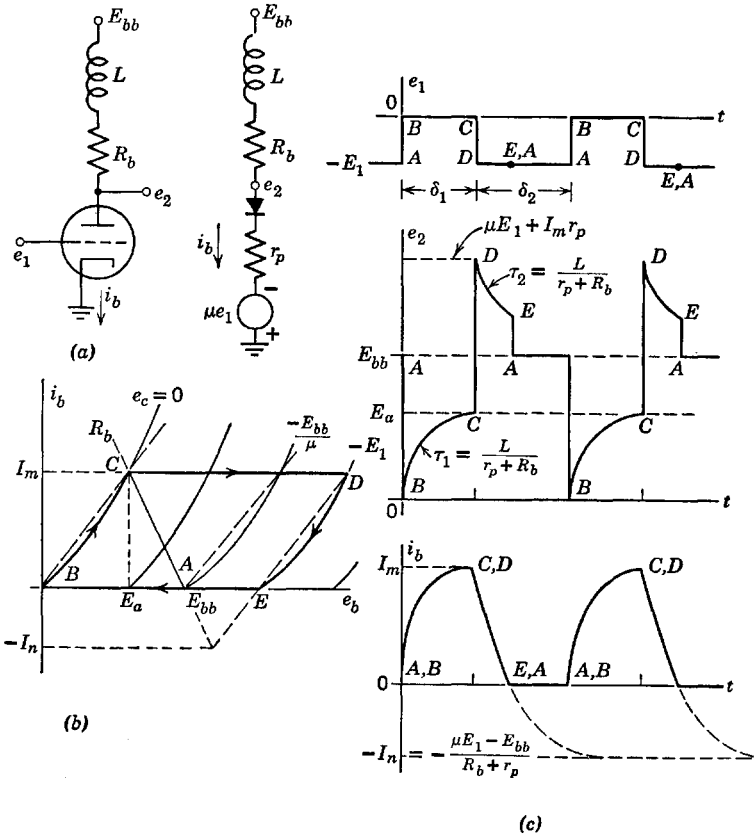
Fig. 8.35.   Triode with series RL plate load $(\delta_1 > 5\tau_1$ and $\delta_2 > 5\tau_2)$.

of stray capacitances are neglected, and the waveforms and locus are drawn for intervals δ_1 and δ_2 sufficiently long to permit completion of the transients.

There is a great deal of similarity between this case and the previous one. If, for example, in the circuit of Fig. 8.35(a) we consider resistance R_b as part of the triode plate resistance, we effectively have a triode with plate resistance $r_p' = R_b + r_p$ and an inductive load L. This, then, corresponds to the circuit of Fig. 8.33(a) with the resistance $R_b = \infty$. In the series circuit of Fig. 8.35(a), the plate current cannot be cut off no matter how negative E_1 is made.

8.26   Cathode-Follower Circuit with Series *RL* Load

A cathode follower with a series RL load is shown in Fig. 8.36(a). A piecewise-linear model is shown in (b), and an incremental model is shown in (c). The rectangular waveform of input voltage shown in (d) results in the output waveform also sketched in (d) and the locus of operation shown in (e). The amplitude of the input waveform has been chosen deliberately in order to avoid positive-grid operation; in fact, E_1 was chosen equal to $(E_{bb} - E_a)$. The intervals δ_1 and δ_2 are here being considered large compared with time constants, so that transients go to completion.

With $e_1 = 0$, the quiescent current is

$$I_0 = \frac{E_{bb}}{r_p + R_k(\mu + 1)} \tag{8.39}$$

which may be found either analytically or graphically from the intersection of the load line and the bias line

$$e_1 = 0 = e_c + i_b R_k \tag{8.40}$$

Note that the inductance is effectively absent under quiescent conditions. However, when the input voltage is switched from A to B, the inductance L plays a predominant role. The current cannot change instantaneously; therefore, the incremental load impedance appears to be infinite. As a result, the point B on the locus can be located from the open-circuit voltage gain of the cathode follower:

$$\Delta e_2 = \mu \, \Delta e_1/(\mu + 1) = \mu E_1/(\mu + 1) \tag{8.41}$$

For the value of E_1 chosen here, $(E_{bb} - E_a)$, the build-up of current leads to the quiescent point C at the intersection of the R_k load line and the $e_c = 0$ line. The variation in i_b proceeds from I_0 to I_m exponentially
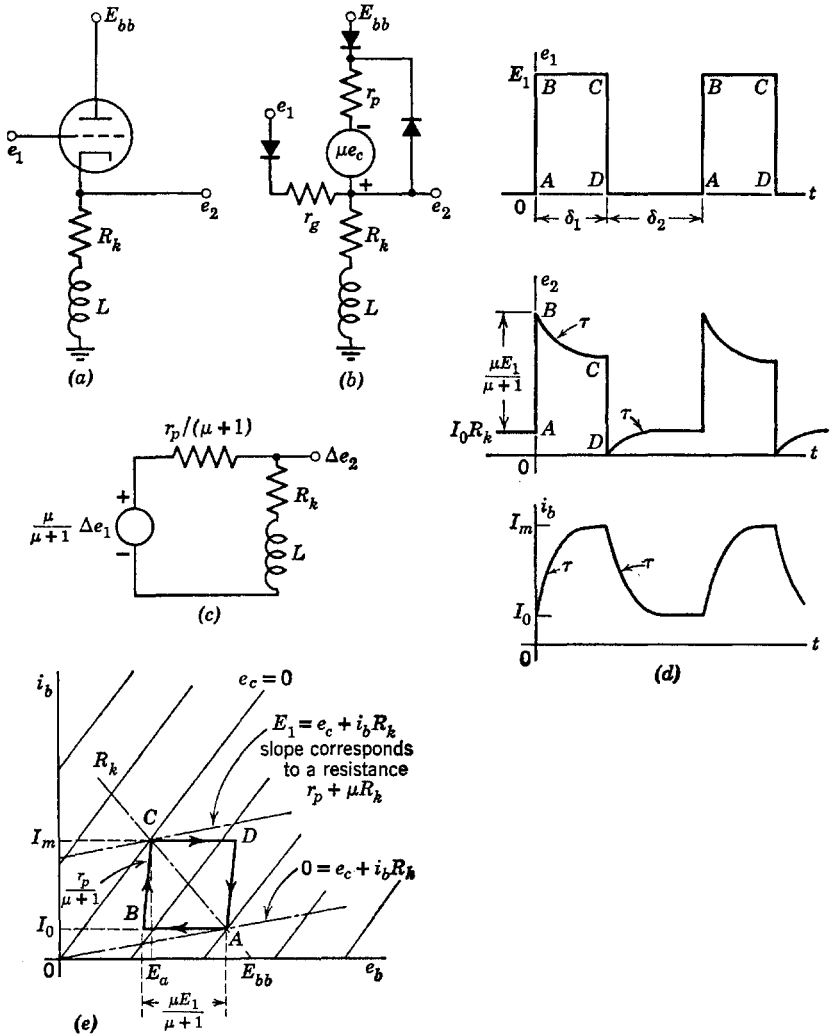
Fig. 8.36. Cathode follower with series RL load.

with the time constant

$$\tau = \frac{L}{R_k + r_p/(\mu + 1)} \tag{8.42}$$

On the locus plot, the variation from B to C corresponding to this build-up is a straight line of slope $r_p/(\mu + 1)$.

An alternate solution to this problem can be found by plotting a new set of i_b vs. e_b curves with e_1 as parameter, rather than e_c. These curves, readily derivable from the standard triode curves, are useful in the analysis of the cathode-follower circuit with any cathode load.
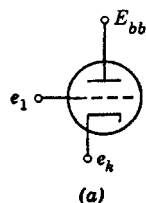
$$e_1 - e_k = e_c$$
$$e_b = E_{bb} - e_k$$
$$= (E_{bb} - e_1) + e_c$$
$$(b)$$
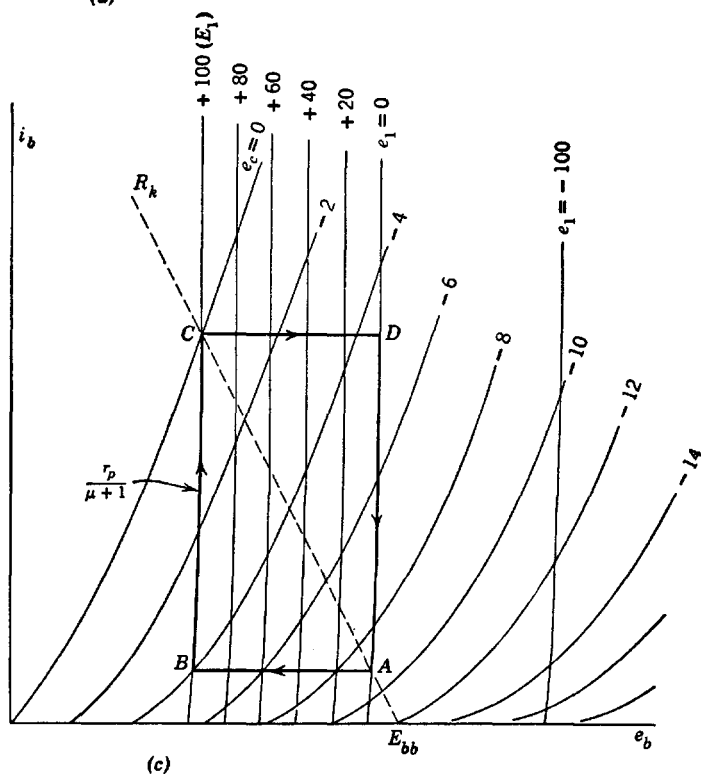
Fig. 8.37. Construction of cathode-output curves.

Consider the triode shown in Fig. 8.37(a). If the plate is held at a fixed potential E_{bb}, then from (b),

$$e_b = (E_{bb} - e_1) + e_c \qquad (8.43)$$

To plot the $e_1 = 0$ line, locate the intersection of the grid-voltage lines e_c

with the corresponding plate voltage $e_b = E_{bb} + e_c$. The line joining these intersections is the $e_1 = 0$ line and, as shown in Fig. 8.37(c), it has a slope of $r_p/(\mu + 1)$.

The locus of operation can now be easily determined. At point A on the input waveform, $e_1 = 0$; and if previous transients have died out, the operating point will be located at the intersection of the load line and the $e_1 = 0$ line. When the input voltage is switched from zero to $+E_1$ (A to B), the current cannot change instantaneously; so the locus is a horizontal line from $e_1 = 0$ to $e_1 = E_1$. During the interval δ_1, e_1 remains constant at E_1, the inductor current increases exponentially to I_m, and the operating point moves up the $e_1 = E_1$ curve. After the transient has died out, the operating point is at C, the intersection of the R_k load line with the $e_1 = E_1$ curve. Similar arguments hold for the transitions from C to D and D to A.

8.27 Cathode Follower with Parallel *RL* Load

The circuit shown in Fig. 8.38(a) can be approximated by the piece-wise-linear model given in (b). For grid-to-cathode voltages between zero and plate current cutoff, the incremental models given in (c) can be used to determine Δe_2 and Δi_x.

Assume that the grid voltage e_1 at point A is much more negative than the cutoff voltage $-E_{bb}/\mu$. Then the point A is located on the i_b vs. e_b plane at $i_b = 0$, $e_b = E_{bb}$, as shown in Fig. 8.38(d). The input transition from A to B drives the operating point along the R_k load line, since i_x remains zero instantaneously. The point B is located at the intersection of the R_k load line and the $e_1 = 0$ line. As the current i_x builds up exponentially to $I_m = E_{bb}/r_p$ with a time constant

$$\tau_1 = \frac{L}{R_k \left|\left| \left(\dfrac{r_p}{\mu + 1} \right) \right.}$$ (8.44)

the output voltage e_2 falls exponentially to zero. As $e_1 = 0$ during this interval (B to C), the operating point moves up the $e_1 = 0$ line to the $e_c = 0$ line (point C). Note that e_b equals E_{bb} at this point.

The input transition from C to D again drives the operating point along a line of slope corresponding to R_k. The apparent large increase in E_{bb} results because the current through the inductance, $i_x = I_m$, instantaneously flows through R_k, driving the cathode to a large negative voltage $-I_m R_k$. If e_1 is sufficiently negative at the point D to cut the tube off in spite of this greatly increased plate voltage, the operating
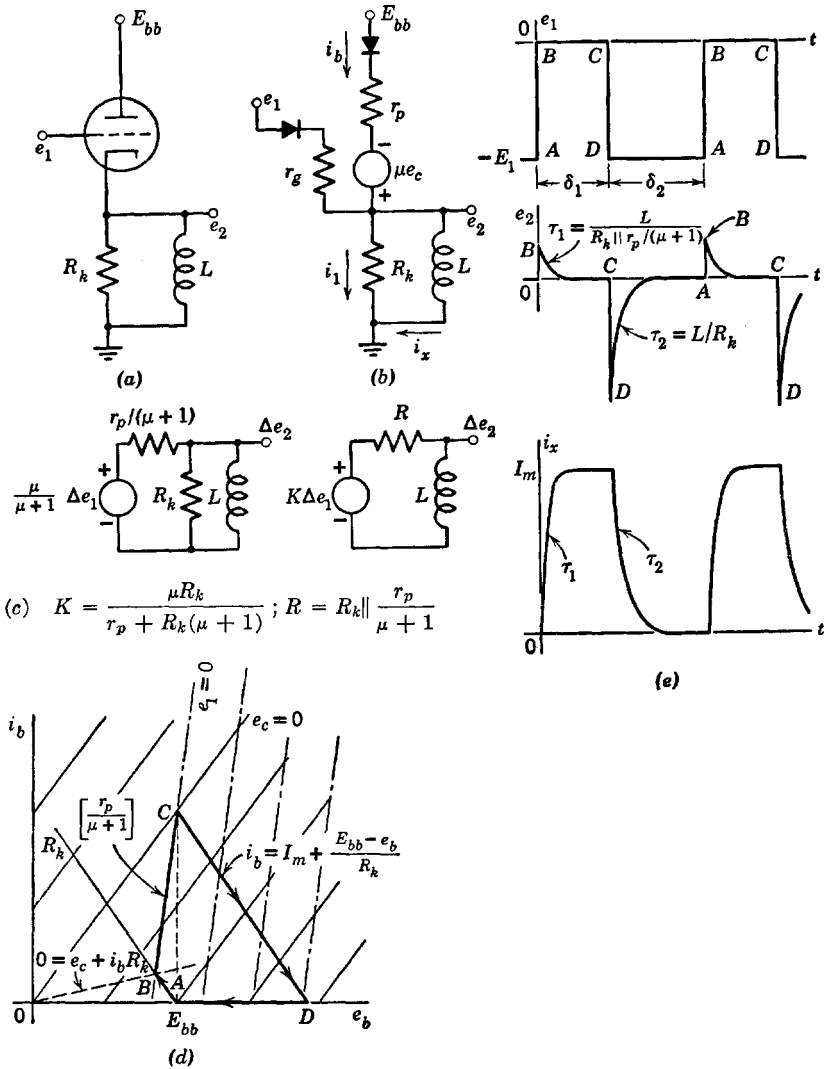
Fig. 8.38. Cathode follower with parallel RL load.

point will follow the R_k line down to the $i_b = 0$ axis. From the graph it can be seen that the required negative value of e_1 for cutoff is

$$e_1 = -E_1 = -I_m R_k - \frac{(E_{bb} + I_m R_k)}{\mu} \qquad (8.45)$$

During the interval δ_2 (D to A), the current i_x and voltage e_2 decay to zero, and the locus returns to A along the $i_b = 0$ axis.

Figure 8.38(d) shows both the e_c lines and the e_1 lines. The locus can be plotted using either of the methods discussed in Article 26.
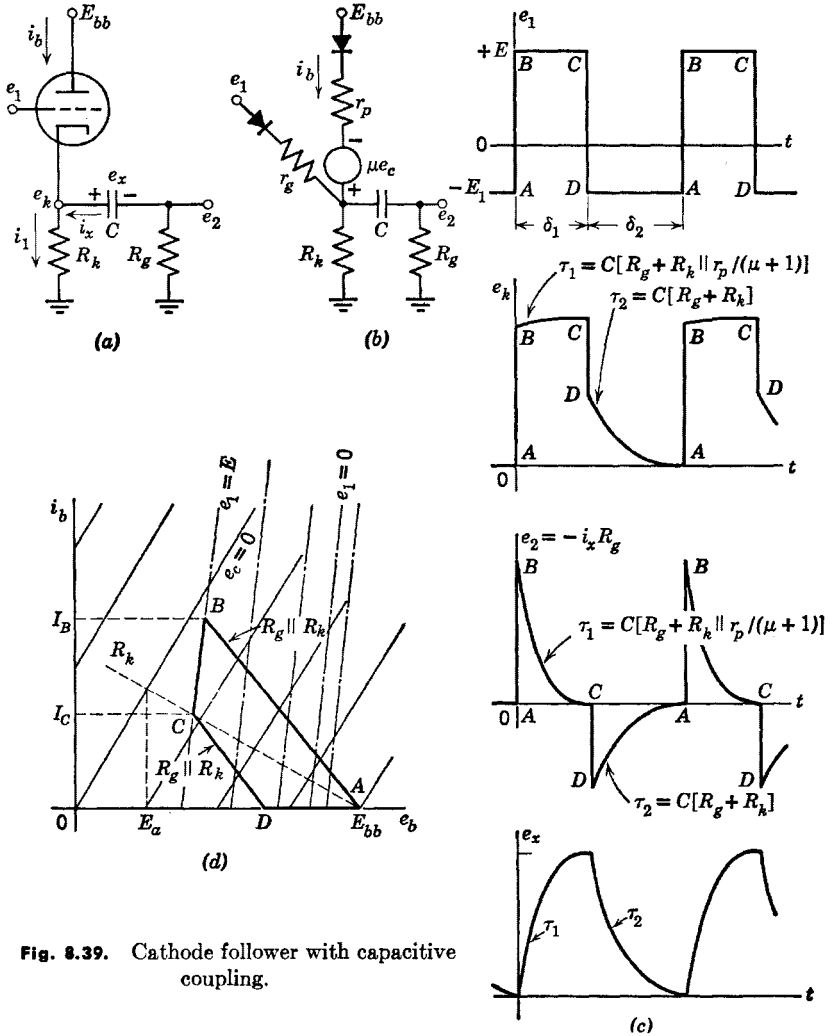
Fig. 8.39. Cathode follower with capacitive coupling.

8.28 Cathode Follower with Capacitive Coupling

The circuit of Fig. 8.39(a) shows a cathode follower with d-c load resistance R_k and capacitive coupling to another resistance R_g. In the

interests of brevity, we shall again confine our attention to the locus of operation with a rectangular input waveform having intervals large compared with the corresponding time constants. Recall that, for a given range of input waveform amplitudes, this case yields the bounding locus within which all others fall.

Using the piecewise-linear model shown in Fig. 8.39(b) and the input waveform shown in (c), we can determine pertinent waveforms and the locus of the operating point. The waveforms of cathode voltage e_k, output voltage e_2, and the capacitor voltage e_x are shown.

Referring to the locus of operation in Fig. 8.39(d), we see that the transition from A to B follows a line whose slope is determined by $R_g || R_k$, since e_x is zero at this time and cannot change instantaneously. In this example, the value of $+E$ has been chosen to avoid the positive-grid region. Point B is located at the intersection of the $R_g || R_k$ load line and the line $e_1 = E$. As the capacitor begins to charge, the current i_x through R_g decreases, and the operating point moves down the $e_1 = E$ line to the quiescent point C on the R_k load line. Again during the transition from C to D, the locus follows a line whose slope depends on $R_g || R_k$. The closure from D to A occurs with $i_b = 0$ and $i_1 = i_x$.

Note that e_k changes only very slightly between B and C, whereas e_2 changes very markedly. This is due to the fact that the incremental resistance between cathode and ground when the triode is conducting is $R_k || [r_p/(\mu + 1)]$, which is much smaller than R_g.

As R_g is decreased, the locus transitions from A to B and from C to D tend to approach the vertical. In this case the upward transition (A to B) can drive the triode far into the positive-grid region.

8.29 Cathode Follower with Parallel RC Load

An important consideration in the operation of cathode followers at high frequencies is the effect on circuit performance of cathode-to-ground capacitance. In Fig. 8.40(a), C could be an actual capacitor, or it may represent tube and wiring capacitances.

The piecewise-linear model is shown in Fig. 8.40(b). The waveforms in (c) are drawn for two different input amplitudes, one small enough to keep the tube in the linear region, and the other sufficiently large to cut the tube off for a portion of the cycle. Let us treat the linear case first.

Point A on the locus, Fig. 8.40(d), is located at the juncture of the R_k load line and the $e_1 = E_1$ line. When the input voltage jumps to E_2, the cathode voltage cannot change instantaneously; so the operating point moves vertically to the $e_1 = E_2$ line. As the capacitor charges,

the operating point is constrained to the $e_1 = E_2$ line, moving down to point C at the intersection of the R_k load line. The return cycle CDA is constructed in a similar manner. As the tube is always conducting in
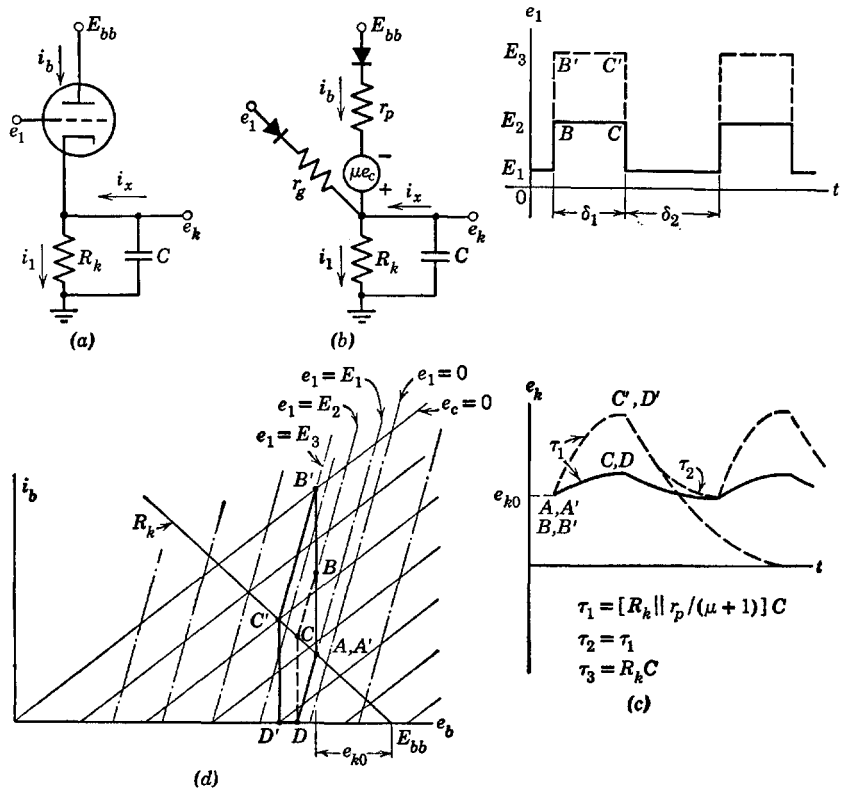
Fig. 8.40. Cathode follower with parallel RC load.

this case, the time constants involved in the e_k waveform are

$$\tau_1 = \tau_2 = \frac{CR_k\left(\dfrac{r_p}{\mu + 1}\right)}{R_k + \left(\dfrac{r_p}{\mu + 1}\right)} \tag{8.46}$$

The effect of shunt capacitance on linear cathode-follower operation can be seen by inspection of the locus. The vertical rise AB and vertical drop CD have the effect of greatly decreasing the magnitude of the input voltage that can be handled by the circuit if linear operation is desired.

If the input waveform amplitude is increased, the locus expands $(A'B'C'D')$; and if $E_3 - E_1$ is sufficiently large, the tube will be cut off immediately following the fast drop in input voltage $(C'D')$.   In this case the time constant of discharge will be CR_k until the tube again begins to conduct.   Note, however, that although the time constant is larger when the tube is cut off, the final value of this exponential is zero rather than e_{k0}.   The capacitor will be discharged to a given voltage more rapidly if the tube is cut off than if it remains conducting.   This fact can be demonstrated by writing the current equation for the cathode circuit.   From Fig. 8.40(a),

$$i_x = i_1 - i_b = \frac{e_k}{R_k} - i_b \qquad (8.47)$$

and

$$i_x = -C\frac{de_k}{dt} \qquad (8.48)$$

Therefore

$$\frac{de_k}{dt} = -\frac{e_k}{CR_k} + \frac{i_b}{C} \qquad (8.49)$$

Since i_b can never be negative, the capacitor will be discharged to a given voltage most rapidly if $i_b = 0$; that is, if the tube is cut off.

The loci shown in Fig. 8.40(d) extend into the positive-grid region if e_1 is greater than E_3.   In this case the grid diode closes and the capacitor is charged by grid current in addition to plate current.

8.30   Grounded-Base Transistor with Inductive Collector Load

The methods for determining waveforms and locus of operation for a triode circuit can be applied to pentode or transistor circuits.   For example, let the p-n-p junction transistor in the grounded-base circuit of Fig. 8.41(a) be represented by the ideal-diode model shown in (b). The shunt diode D_z and source V_z, shown by dotted lines between collector and ground, provide the voltage limiting due to the avalanche effect in the semiconductor.   The collector curves (v_{cb} vs. i_c) for this model are shown in (c).   The locus of the operating point ($ABCDEA$) corresponds to the rectangular waveform of current i_1 shown in (d), where waveforms of i_c and v_{cb} are also shown.   Intervals δ_1 and δ_2 are assumed large compared with the time constant L/R.

The idealization of the transistor model greatly simplifies the locus and waveform determinations.   When the input current switches from

A to B, $i_c = i_x$ is held at zero instantaneously because of inductance L. Since i_c remains at zero, the source ai_e closes the collector diode D_c; hence voltage v_{cb} switches instantaneously to zero. This voltage remains zero until i_c builds up exponentially ($\tau_1 = L/R$) to the value $-V_c/R$.
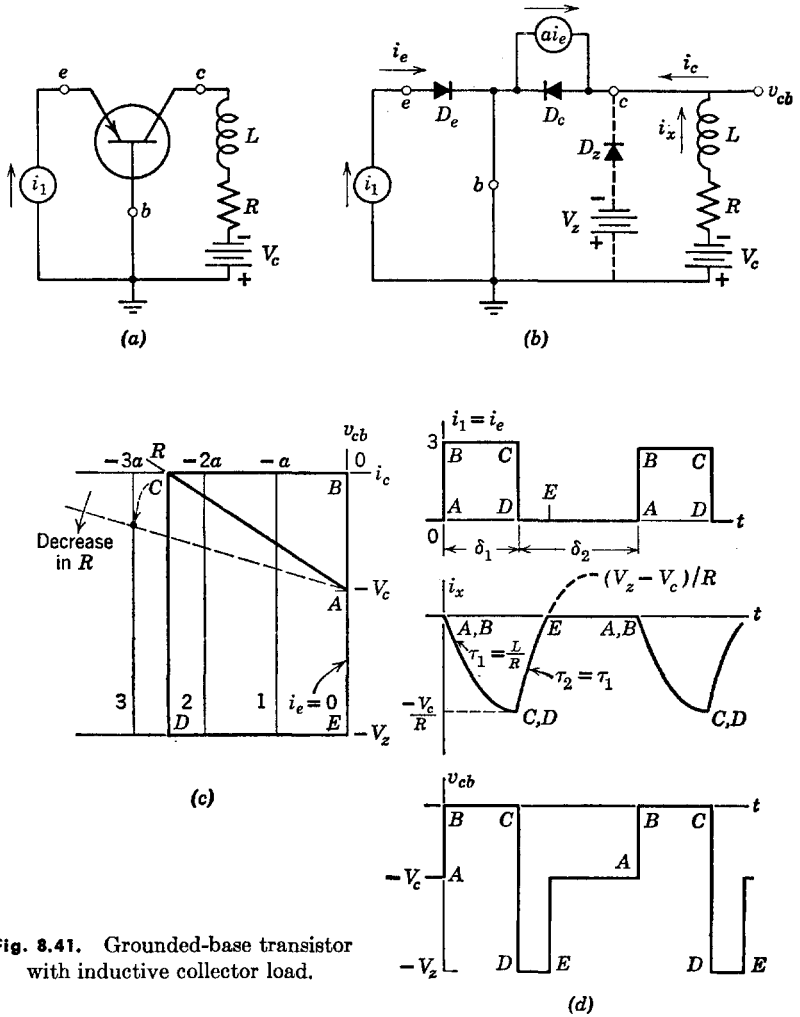
Fig. 8.41. Grounded-base transistor with inductive collector load.

The input transition from C to D reduces ai_e to zero and therefore opens D_c. Now the collector current is zero, but the inductor current i_x tries to remain constant at $-V_c/R$; hence the voltage v_{cb} drops to $-V_z$, thus closing D_z. The current i_x now varies exponentially ($\tau_2 = \tau_1 = L/R$) from $-V_c/R$ toward $+(V_z - V_c)/R$. When i_x reaches zero,

however, the diode D_z opens and permits v_{cb} to rise instantaneously from $-V_z$ to $-V_c$ in order to complete the cycle.
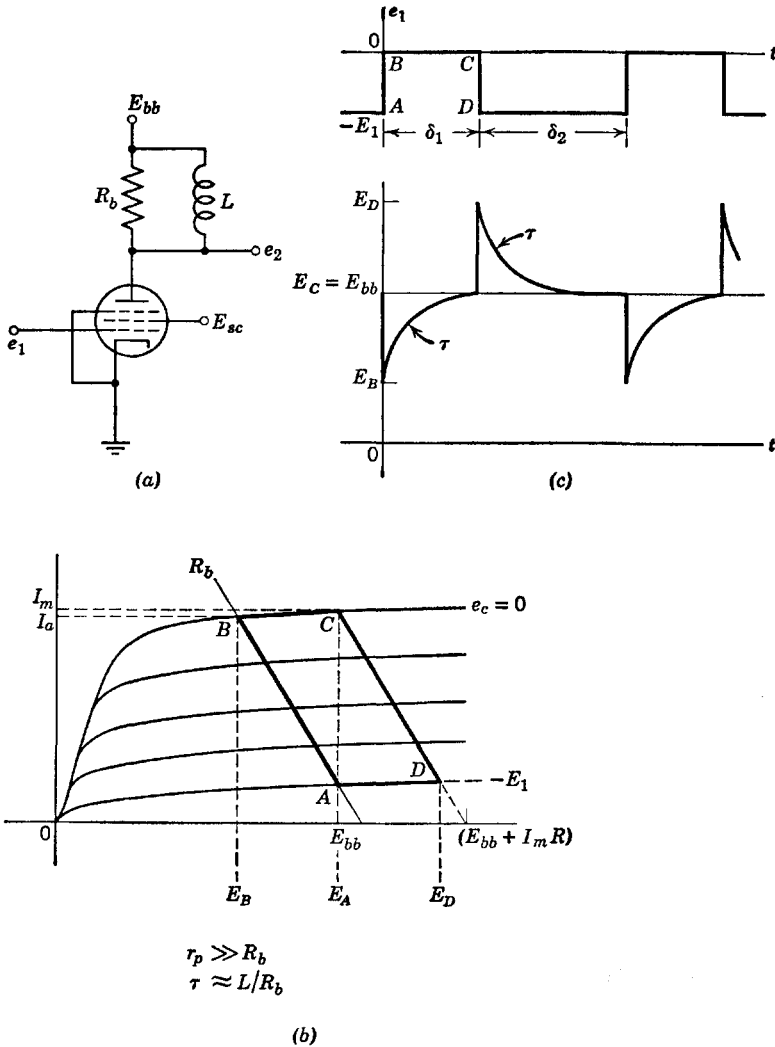
Fig. 8.42. Pentode with parallel RL plate load (δ_1, $\delta_2 > 5\tau$).

8.31   Pentode with Parallel *RL* Load

With the exception of the preceding article, we have used triodes in the single-time-constant circuits analyzed in this chapter. It was con-

venient to use one type of device in order to show the effects of load and input waveform variations with a minimum of other changes. However, once we specify a device by means of characteristic curves or a circuit model, the analysis proceeds in the same manner for any of the devices we have discussed. By way of illustration, Fig. 8.42 shows a pentode with parallel RL load. The locus and output waveform should be compared with those in Figs. 8.33 and 8.34. The shape of the locus appears to be different principally because the pentode is very nearly a voltage-controlled current source, whereas a triode more nearly resembles a voltage-controlled voltage source.

Since R_b is much less than the plate resistance of a pentode, the apparent resistance faced by the inductance is very nearly R_b whether or not the pentode conducts.

Another difference that may be important is the magnitude of inter-electrode capacitances. The grid-to-plate capacitance of the triode couples some of the input signal to the output (and vice versa). In the pentode, this capacitance is small enough to be neglected at all but very high frequencies.

As in the case of the RC coupling circuit we can approximate the effects of tube and wiring capacitances fairly readily. If the small capacitance (say 10 $\mu\mu f$) is assumed to exist between plate and ground in Fig. 8.41 the load on the pentode is effectively a parallel RLC circuit. For such a small value of C the waveform of e_2 is only slightly modified. The instantaneous jumps become finite slopes and the sharp peaks are slightly rounded.

8.32   Pentode with Parallel *RLC* Load

A pentode with a parallel RLC load is shown in Fig. 8.43(a). For a typical pentode in the normal region of operation the plate current is nearly independent of the plate voltage. Thus the rectangular waveform of plate current shown in Fig. 8.43(d), as determined by the resistive load alone, remains virtually unchanged when the inductance and capacitance are connected. The behavior, therefore, can be determined from the general linear circuit shown in (c). Whether or not the input waveform causes plate-current cutoff makes relatively little difference in incremental behavior, since the plate resistance for a pentode is very large compared with load resistance R. Thus we need only examine the behavior of the parallel RLC circuit with a rectangular waveform of current applied as indicated in (d). The differential equa-

tion for the circuit is

$$\frac{d^2v}{dt^2} + \frac{1}{RC}\frac{dv}{dt} + \frac{v}{LC} = 0 \tag{8.50}$$

The roots of the characteristic equation are

$$s_1 = -\frac{1}{2RC} + \sqrt{\left(\frac{1}{2RC}\right)^2 - \frac{1}{LC}} \tag{8.51}$$

$$s_2 = -\frac{1}{2RC} - \sqrt{\left(\frac{1}{2RC}\right)^2 - \frac{1}{LC}} \tag{8.51a}$$

The first three of the following modes correspond to positive, zero, and negative values of the radical, for R positive,* while the fourth or un-damped mode occurs for $R = \infty$

1.   Overdamped: $R < \dfrac{1}{2}\sqrt{\dfrac{L}{C}}$

2.   Critically damped: $R = \dfrac{1}{2}\sqrt{\dfrac{L}{C}}$ (8.52)

3.   Underdamped: $R > \dfrac{1}{2}\sqrt{\dfrac{L}{C}}$

4.   Undamped: $R = \infty$

The waveforms in Fig. 8.43(e) and (f) correspond to underdamped cases with (e) being near the critically damped case. The locus in (b) is drawn for the critically damped mode of operation. The locus corresponding to the waveform in (f) will be like that shown in (b) except that the horizontal portions will extend nearly equally on both sides of the vertical line at E_{bb}.

The expressions for the incremental voltage across the tuned circuit are given below. Adding v to E_{bb} or subtracting it from E_{bb} yields e_2 in the two intervals.

1.   Overdamped

$$v = \frac{I_a/C}{s_1 - s_2}[\epsilon^{s_1 t} - \epsilon^{s_2 t}] \tag{8.53}$$

2.   Critically damped

$$v = \frac{I_a}{C}t\epsilon^{-t/2RC} \tag{8.54}$$

* If R is negative there are three negatively-damped modes. These are pertinent to oscillator theory and are treated in Chapter 10.
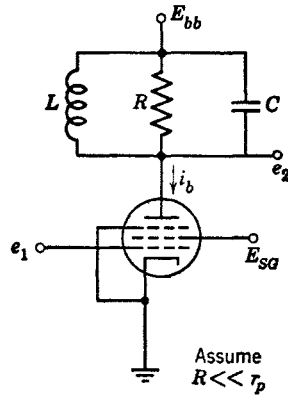
(a)
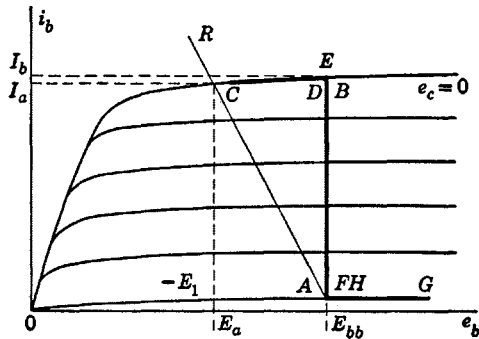
(b)
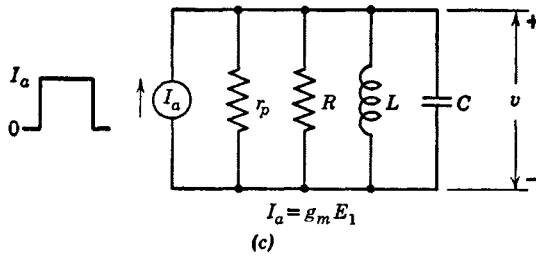
(c)

Fig. 8.43.   Rectangular-wave response of
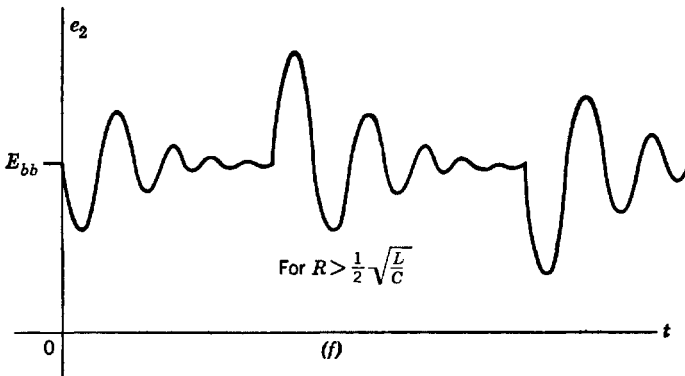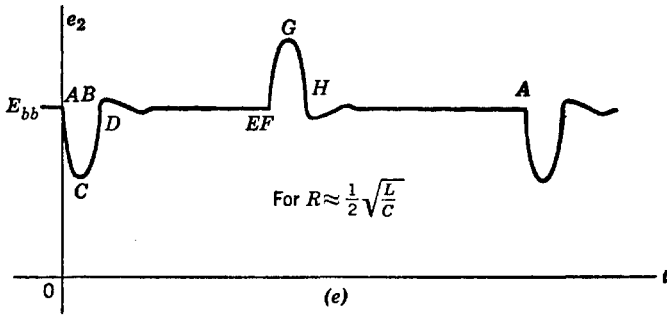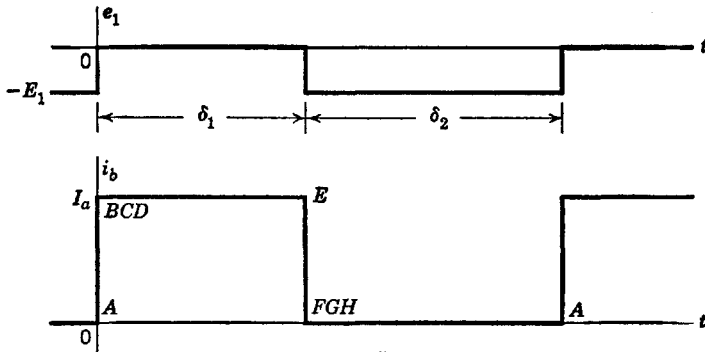
pentode with parallel RLC load.
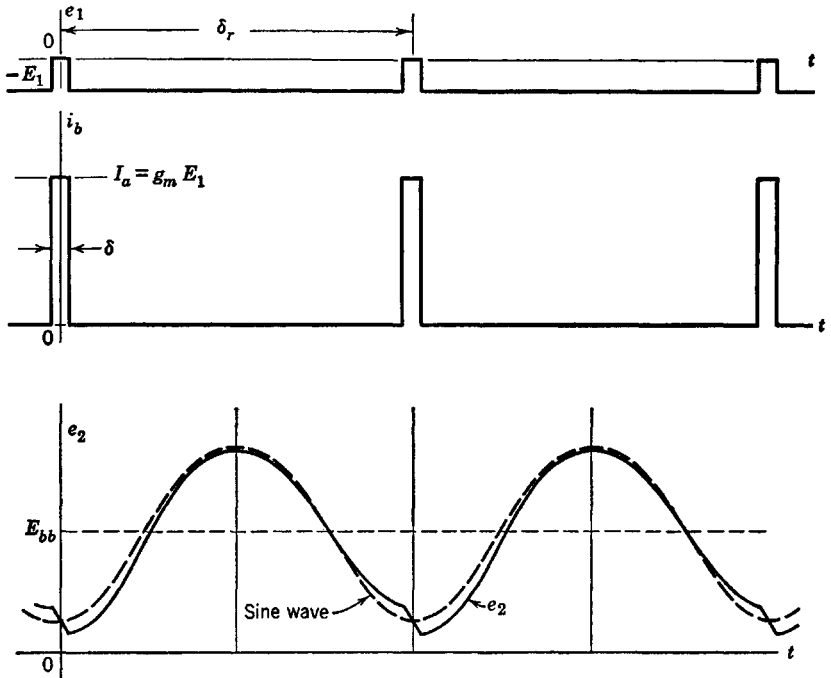
Fig. 8.44. Pulse response of pentode with RLC load for $\delta_r = 2\pi/\omega_0 = 1/f_0$.

3. Underdamped

$$v = \frac{I_a/C}{\beta} \epsilon^{-\alpha t} \sin \beta t \qquad (8.55)$$

where $\alpha = \dfrac{1}{2RC}$

and $\beta = \sqrt{\dfrac{1}{LC} - \left(\dfrac{1}{2RC}\right)^2}$

4. Undamped

$$v = \frac{I_a}{C} \sqrt{LC} \sin \sqrt{\frac{1}{LC}}\, t \qquad (8.56)$$

In the nearly-undamped case, the transient lasts a long time. An important application of the RLC load with relatively small damping is the tuned, overdriven amplifier used both for power amplification and frequency multiplication. A simplified representation of the behavior of a nonlinear power amplifier is shown in Fig. 8.44. The rectangular

wave has been replaced by brief pulses and the LC circuit is tuned to a frequency ω_0 whose period is the same as the repetition interval of the pulses. If the RLC circuit is nearly undamped, the waveform of e_2 is
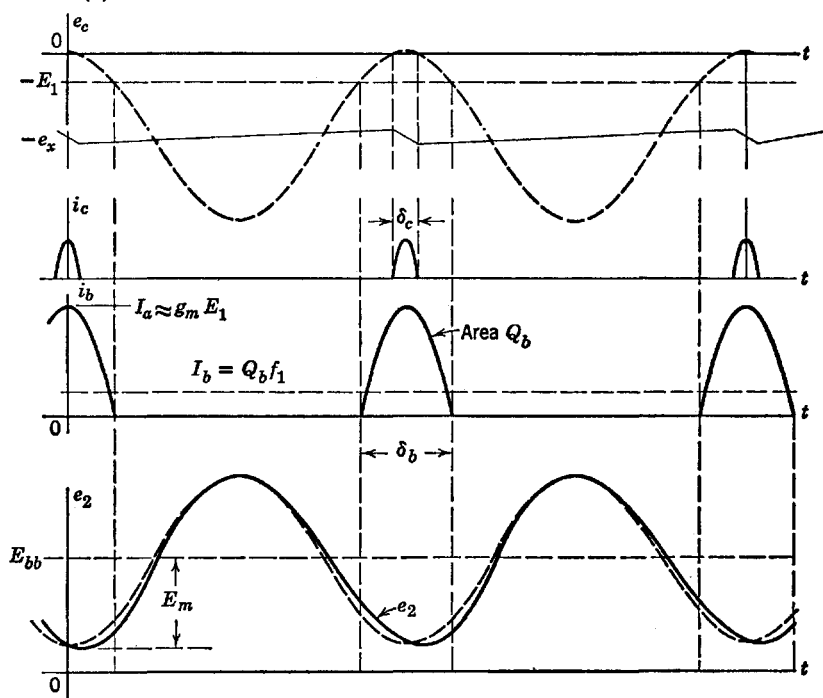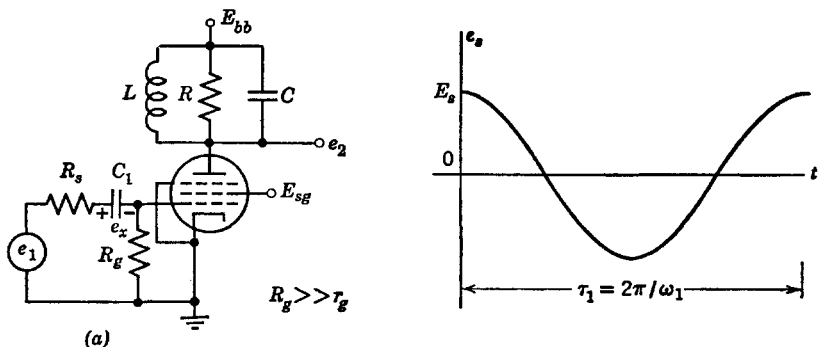
Fig. 8.45. Overdriven sine wave amplifier waveforms ($\omega_1 = \omega_0 = 1/\sqrt{LC}$).

nearly sinusoidal. The brief pulses add charge to the capacitor and bring the amplitude above the sine wave value. Equilibrium amplitude occurs when each pulse adds an amount of energy just equal to that

dissipated in R per cycle. In particular, the voltage across the tuned circuit $(E_{bb} - e_b)$ is near its maximum E_m and therefore nearly constant during the plate-current pulse interval. Thus the energy delivered to the tuned circuit per cycle is $E_m Q_b$, where Q_b is the area of one plate-current pulse. Energy dissipated per cycle is $(E_m{}^2/2R)\tau_0$, where $\tau_0 = 2\pi \sqrt{LC} = 2\pi/\omega_0$. Thus $E_m Q_b = E_m{}^2 \tau_0/2R$, or

$$E_m = 2RI_b, \qquad (8.57)$$

where $I_b = Q_b/\tau_0$ is the d-c plate current.

The practical embodiment of these ideas is the tuned, overdriven amplifier shown in Fig. 8.45. The large amplitude input waveform tends to drive the grid positive with respect to cathode. If R_g is very large, the input circuit is basically a clamper and develops a net charge on C_1 corresponding very nearly to the amplitude of the input waveform. Thus the positive peaks of the wave $e_c(t)$ are clamped at zero. The plate-current pulses then produce the output waveform shown qualitatively in Fig. 8.45. The peak amplitude E_m cannot exceed E_{bb}, because the amplitude of the plate-current pulse decreases markedly if minimum e_b falls below the knee of the pentode plate curves.

With the parallel-resonant circuit tuned to a frequency $\omega_0 = n\omega_1$, where n is an integer, the overdriven amplifier becomes a so-called ratio-n frequency multiplier. In this case, a current pulse is applied to the resonating circuit only once every n cycles, so that the average plate current is now $I_b = Q_b/n\tau_0$.

8.33   Frequency Response of Parallel *RLC* Circuit

If the input waveform for the circuit in Fig. 8.43 (a) is a sine wave that varies the grid voltage between zero and $-E_1$, the pentode is essentially linear and the plate current varies sinusoidally with an rms value I. The admittance of the tuned circuit can be expressed as

$$Y = \frac{1}{R} + \frac{1}{j\omega L} + j\omega C = \frac{1}{R}\left\{1 + jQ\left[\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right]\right\} \qquad (8.58)$$

where the resonant frequency $\omega_0 = 1/\sqrt{LC}$ and the figure of merit[*] $Q = R/\sqrt{L/C}$.

---

[*] Figure of merit or quality factor Q for a resonant circuit is defined as

$$Q = 2\pi \frac{\text{Peak stored energy}}{\text{Energy dissipated per cycle}}$$

(a)

(b)
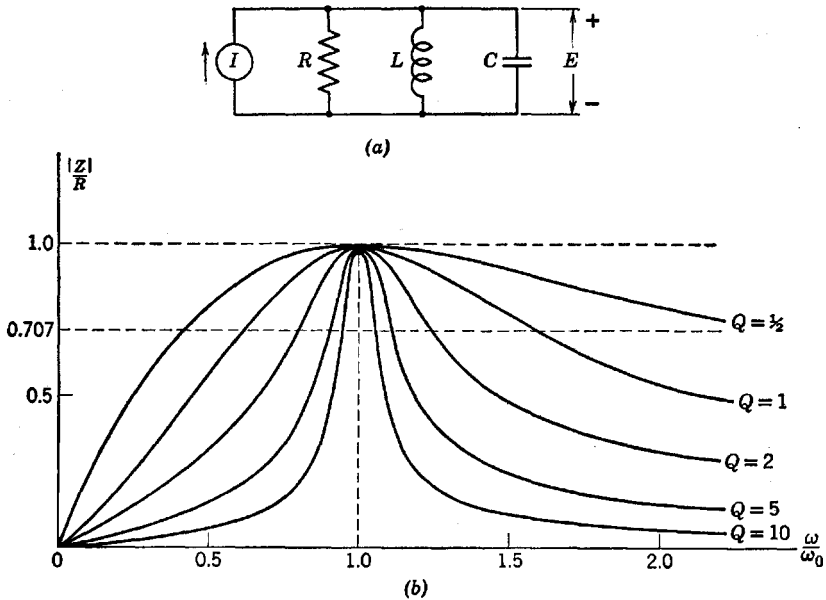
Fig. 8.46. Frequency response of pentode with parallel RLC load.

The impedance $(Z = E/I)$ is plotted as a function of ω in Fig. 8.46(b) for several values of Q. The RLC circuit is useful in frequency-selective amplifiers, since the output voltage $E(\omega)$ has the same form as $Z(\omega)$ when the circuit is driven by a current source.

SUPPLEMENTARY READING

L. B. Arguimbau and R. B. Adler, *Vacuum-Tube Circuits and Transistors*, John Wiley and Sons, Inc., New York, 1956.

B. Chance, *Waveforms*, McGraw-Hill, New York, 1948.

E. A. Guillemin, *Introductory Circuit Theory*, John Wiley and Sons, New York, 1953.

L. P. Hunter, editor, *Handbook of Semiconductor Electronics*, McGraw-Hill, New York, 1956.

J. Millman and H. Taub, *Pulse and Digital Circuits*, McGraw-Hill, New York, 1956.

J. F. Reintjes and G. T. Coate, *Principles of Radar*, 3rd edition, McGraw-Hill, New York, 1952.

Samuel Seely, *Electron-Tube Circuits*, McGraw-Hill, New York, 1950.

G. E. Valley and H. Wallman, *Vacuum Tube Amplifiers*, McGraw-Hill, New York, 1948.

V. K. Zworykin and G. A. Morton, *Television: The Electronics of Image Transmission in Color and Monochrome*, 2nd edition, John Wiley and Sons, New York 1954.

PROBLEMS (See Appendices for device curves)

8.1. Specify the waveform of Fig. 8.8(c) in terms of a combination of continuous ramp functions.
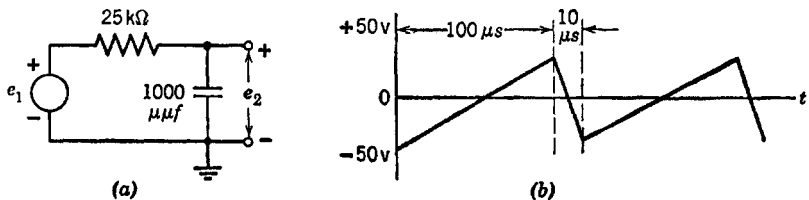
Fig. P8.1

8.2. For the circuit shown in Fig. P8.1(a), sketch and dimension the waveform of voltage e_1 required to produce the linear sawtooth voltage e_2 shown in Fig. P8.1(b).
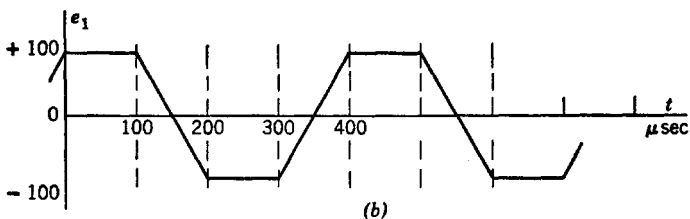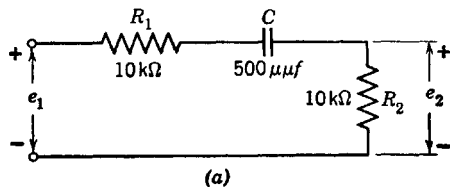
Fig. P8.2

8.3. Sketch and dimension the waveform of output voltage e_2 from the circuit of Fig. P8.2(a) for the input waveform shown in Fig. P8.2(b). Repeat with $R_1 = 0$.

8.4. The pulse transformer in the circuit of Fig. P8.3(a) has a magnetizing inductance $L_m = 80$ mh referred to the primary winding. Leakage inductances and stray capacitances may be neglected. Determine the waveform of input voltage e_1 which produces a single rectangular output pulse e_2 [Fig. P8.3(b)] having an amplitude of 20 volts and a duration of 2 microseconds.

8.5. Sketch and dimension the output $e_2(t)$ for the circuit of Fig. P8.4(a) with input waveform $e_1(t)$ as shown in Fig. P8.4(b). Let the triode be piecewise-linear with $\mu = 20$, $r_p = 10$ kilohms, and $r_g = 1$ kilohm.
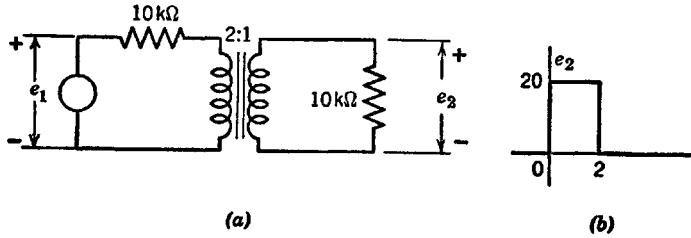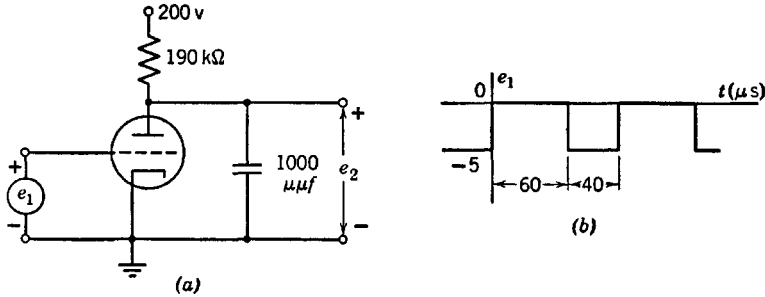
Fig. P8.3

Fig. P8.4

8.6. In the circuit of Fig. P8.5 the capacitance C is so large that the ripple in e_2 is negligible. The input voltage e_1 is an audio-frequency square wave of amplitude E_1. Use a piecewise-linear triode model with $r_p = 30$ kilohms, $r_g = 0$, and $\mu = 50$.
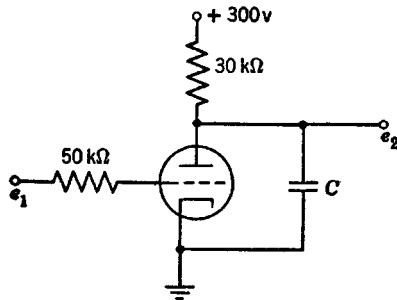
Fig. P8.5

(a) Sketch and dimension the operating path of the tube on the piecewise-linear plate curves.

(b) Plot the d-c output voltage E_2 vs. E_1.

8.7. (a) What is the minimum value of E in Fig. P8.6(b) for which the tube in Fig. P8.6(a) remains cut off during the interval $0 < t < 0.01$ sec.?

(b) With E equal to the value found in part (a), sketch and dimension the output voltage waveform, indicating the values of time constants on exponentials.
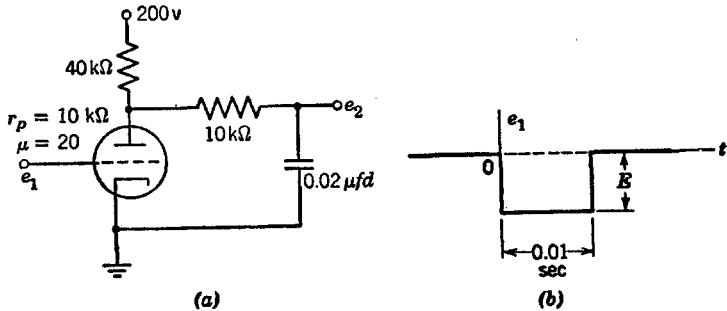
Fig. P8.6

8.8. In the circuit of Fig. P8.7, assume a piecewise-linear triode model with plate resistance r_p and cutoff for $e_b + \mu e_c$ negative. The input voltage e_1 is zero until a negative step of magnitude E_1 is applied at time zero.
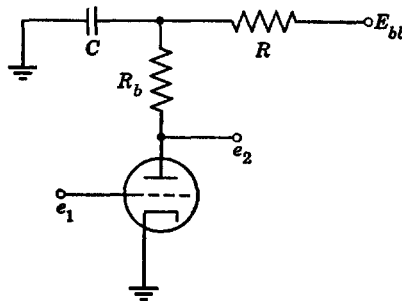
Fig. P8.7

(a) What is the maximum value of E_1 for which the tube never cuts off?

(b) What is the minimum value of E_1 for which the tube is permanently cut off after the step is applied?

(c) Sketch $e_2(t)$ for a value of E_1 somewhere between the two values found in (a) and (b).

8.9. Determine the plate power dissipated by the triode with and without C present in the circuit of Fig. P8.8(a) when the repetitive waveform $e_1(t)$ shown in Fig. P8.8(b) is applied. The triode may be represented by a piecewise-linear model with $\mu = 20$ and $r_p = 10$ kilohms.

8.10. Input and output voltage waveforms for the circuit of Fig. P8.9(a) are as shown in Fig. P8.9(b). Derive an expression for the average value of e_2 in terms of the time intervals, circuit parameters and voltages shown.

8.11. In the circuit of Fig. P8.10(a) the triode may be represented by a piecewise-linear model with $\mu = 20$, $r_p = 10$ kilohms, and $r_g = 0$. Determine the waveform of the output voltage e_0 with e_1 as shown in Fig. P8.10(b).

8.12. The circuit shown in Fig. P8.11(a) is used as a pulse-width discriminator. For an input pulse duration $\delta_1 \leqslant 4 \, \mu\text{sec}$, the variational output is to be zero. For $\delta_1 \geqslant 6 \, \mu\text{sec}$, the output pulses are to have an amplitude of 60 volts. Find the values required for C, R_{b2}, and E_k. Let the triodes be represented

Fig. P8.8

Fig. P8.9

(a)

(b)
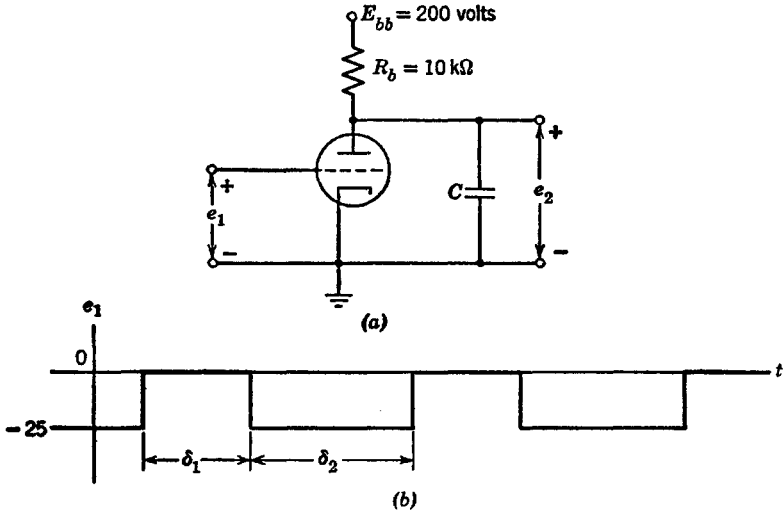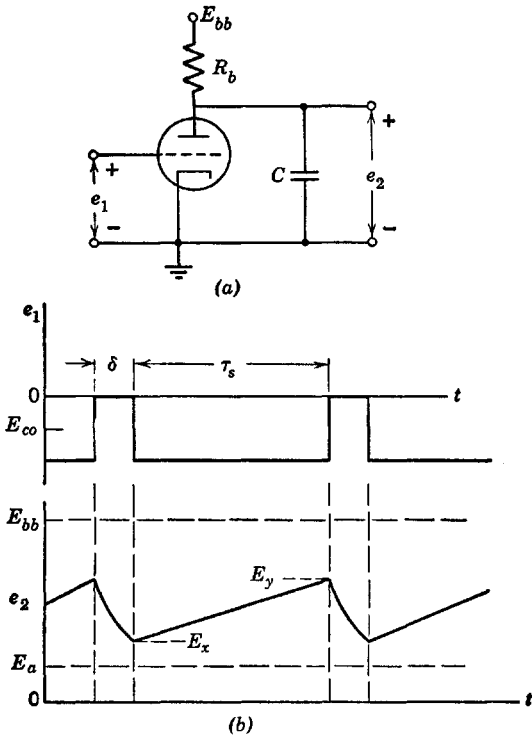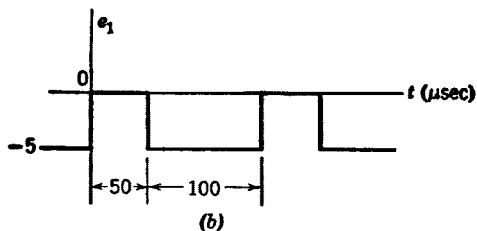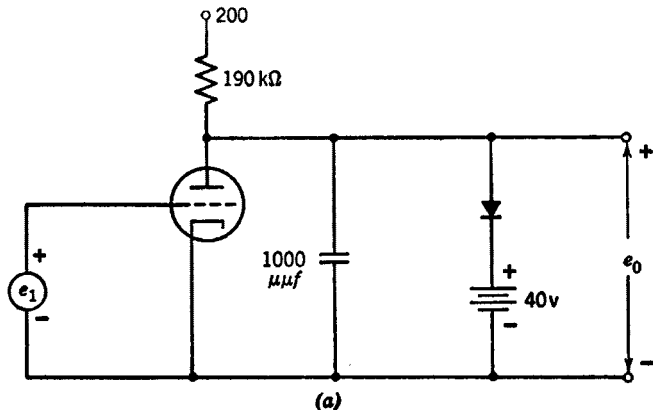
Fig. P8.10

by a piecewise-linear model with $\mu = 20$, $r_p = 10$ kilohms, $r_g = 1$ kilohm. Make simplifying approximations but check their validity or indicate the limitations they impose on the accuracy of the solution.

8.13. The 10 microsecond input pulse for the circuit of Fig. P8.12 has a repetition frequency of 5 kc. Sketch and dimension the waveform of output voltage e_2.

8.14. The triodes in the circuit of Fig. P8.13(a) can be approximated by a piecewise-linear model with $\mu = 25$, $r_p = 20$ kilohms, and $r_g = 1$ kilohm. The input waveform is as shown in Fig. P8.13(b). In terms of the circuit and triode parameters, what is the maximum ratio δ_1/δ_2 that permits complete charging of capacitor C (5 time constants) during δ_2 while maintaining $e_{c2} < -E_{bb}/\mu$ during δ_1?

8.15. Referring to the circuit and input waveform in Fig. P8.13, express the positive overshoot on e_{c2} in terms of circuit parameters.

8.16. Referring to Problem 8.14, let $\delta_1 = 40$ μsec, $\delta_2 = 60$ μsec, $R_{b1} = 20$ kilohms, $R_g = 200$ kilohms and $C = 1000$ $\mu\mu$f. Sketch and dimension the waveforms $e_{b1}(t)$, $e_x(t)$ and $e_{c2}(t)$.

8.17. For the conditions of Problem 8.16, sketch the locus of the triode operating point (i_b vs. e_b).

Fig. P8.11

Fig. P8.12

(a)

(b)
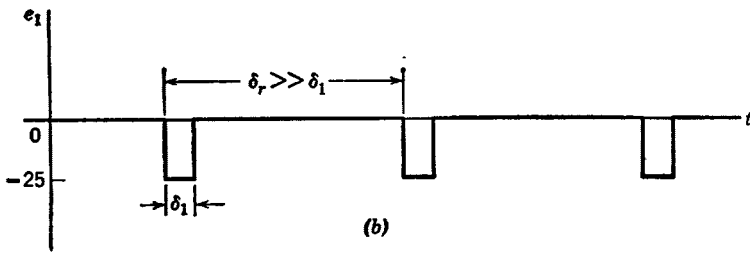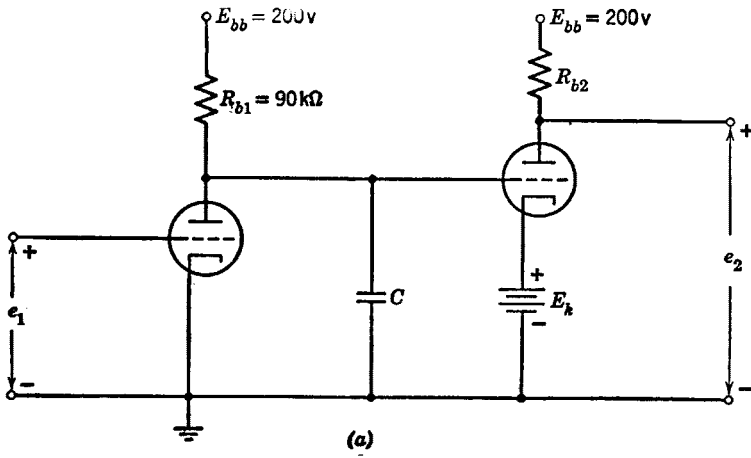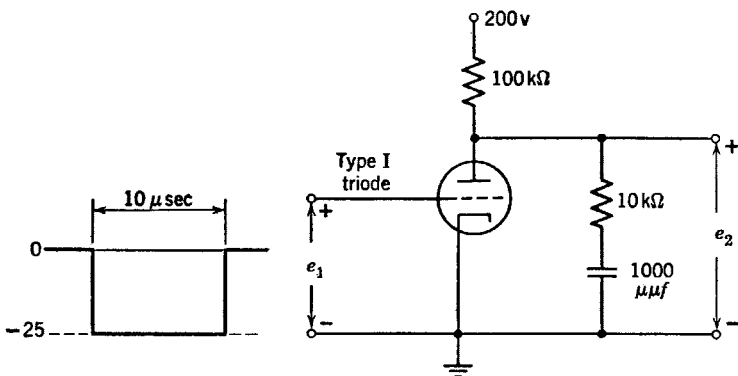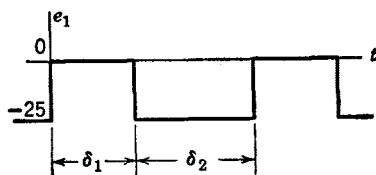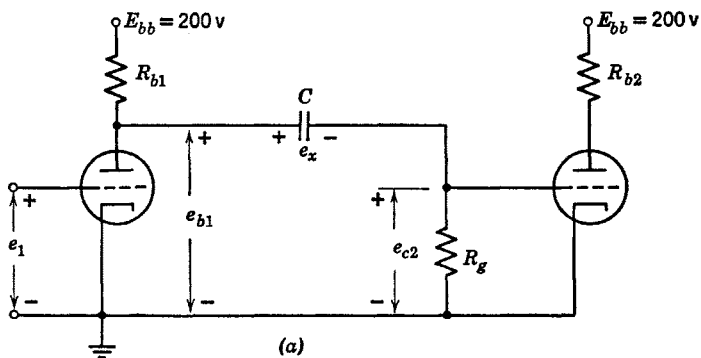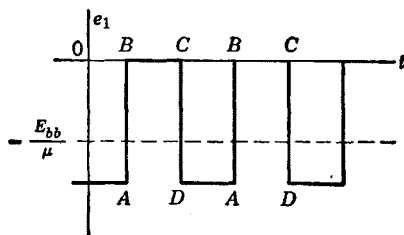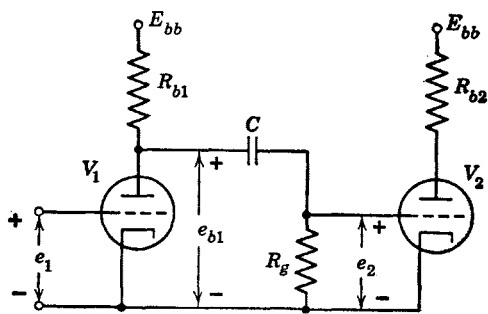
Fig. P8.13

Fig. P8.14

Fig. P8.15

8.18. The four-segment path in Fig. P8.16 is the locus of the operating point of triode V_1 in the i_b vs. e_b plane when the periodic waveform $e_1(t)$ shown in Fig. P8.14 is applied between the grid and cathode. The intervals BC and DA are both long compared with the time constants of the circuit of Fig. P8.15.
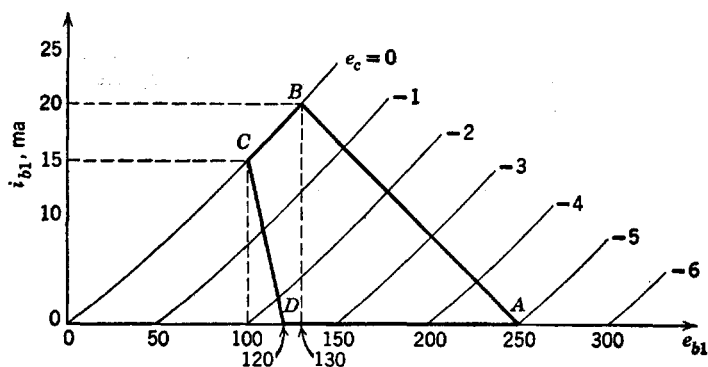
Fig. P8.16

(a) What is the value of R_{b1} in ohms?

(b) What is the value of R_g in ohms?

(c) During which segment of the locus is capacitor C charging?

(d) By what voltage does the voltage on capacitor C change during locus segment CD?

(e) What is the resistance in ohms which determines the charging time constant?
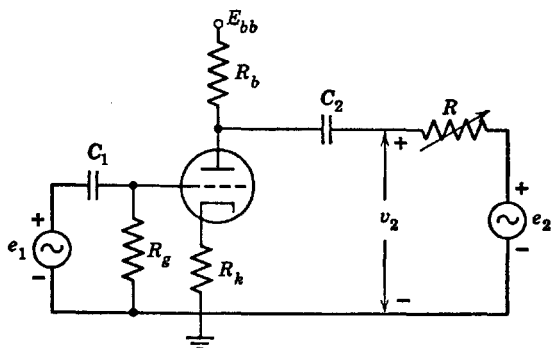
Fig. P8.17

8.19. With the a-c sources e_1 and e_2 both set equal to zero, the triode circuit shown in Fig. P8.17 operates at certain d-c values of e_b and i_b. Show how you would find these d-c values

(a) From the approximate piecewise-linear triode model, and

(b) By graphical construction on the Type III triode curves. Let r_p and μ be the constants of the linear incremental triode model valid in the neighborhood of the d-c operating point.

(c) With e_1 a small fixed-amplitude a-c signal and with e_2 set equal to zero, suppose that R is adjusted to make v_2 (as measured by an a-c voltmeter) take exactly half of its open-circuit (infinite-R) value. Find this value of R in terms of μ, r_p, R_b, and R_k. Assume that the frequency of $e_1(t)$ is high enough to give negligible ripple voltage across C_1 and C_2, but not so high as to make tube and stray wiring capacitances important.

(d) With e_1 set equal to zero and with e_2 a small fixed-amplitude a-c signal at the same frequency as in part (c), suppose that R is adjusted to make v_2 take exactly half of its short-circuit (zero-R) value. Find this value of R in terms of μ, r_p, R_b, and R_k.

(e) How are the values of R found in (c) and (d) related to the so-called incremental output resistance of the amplifier?

(f) Repeat with the left-hand terminal of capacitor C_2 connected to the cathode instead of the plate.
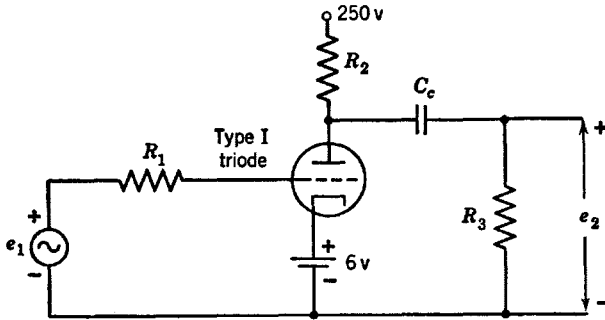
Fig. P8.18

8.20. In Fig. P8.18 assume that the grid-cathode capacitance including wiring is 12 $\mu\mu$f. The grid-plate capacitance including wiring is 8 $\mu\mu$f. Neglect the plate-cathode capacitance.

(a) Using the Type I triode plate characteristics, find μ, g_m, and r_p at the quiescent operating point for $R_2 = 20$ $k\Omega$.

(b) Calculate the effective grid-to-ground capacitance assuming R_1 and R_3 large.
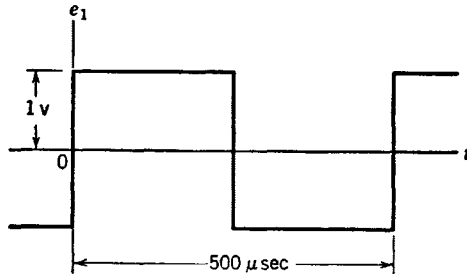
Fig. P8.19

(c) With e_1 a periodic square wave, as shown in Fig. P8.19, plot the output waveform e_2, assuming that C_c is infinite and C_g has the value calculated in (b).

(d) With e_1 as in Fig. P8.19, plot the output waveform e_2, assuming that C_g is zero and $C_c = 0.015~\mu f$.

(e) Sketch the waveform of e_2 if C_g has the value in (b) and $C_c = 0.015~\mu f$.

(f) Show on the plate characteristics the operating locus of the tube.

8.21. Using the same circuit as in Problem 8.20, apply an input voltage $e_1 = 2 \cos \omega t$.

(a) On semi-log paper plot the voltage gain in decibels as a function of frequency relative to the midband gain. What is the upper half-power frequency? What is the lower half-power frequency?

Plot on the same coordinates the phase characteristics as a function of frequency.

(b) Plot on polar coordinates the complex locus of the voltage gain. Show the location of points corresponding to the midband frequency, the lower half-power frequency, the upper half-power frequency.
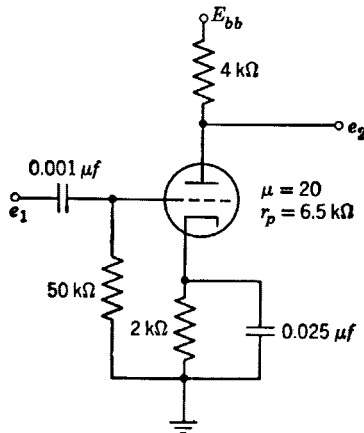
Fig. P8.20

8.22. From the circuit of Fig. P8.20, sketch and dimension:

(a) the response curve $|A(j\omega)|$;

(b) the transient response to a 1-volt step.

8.23. Using an appropriate linear incremental circuit model, calculate the small-signal a-c voltage amplification $A = e_0/e_1$ of the circuit shown in Fig. P8.21. Assume that capacitances C_1 and C_k are large.

8.24. With the circuit of Fig. P8.21 initially at rest, a negative pulse of magnitude E_1 and short duration T is applied at the input terminals. Plot the magnitude of the output voltage pulse E_0 as a function of E_1.

8.25. In Fig. P8.21 all parameters except E_{bb} and R_k are to remain as given. E_{bb} and R_k are to be determined for the following operating conditions.

A periodic square wave having a peak-to-peak amplitude of 16 volts is to be applied to the input of the amplifier. Under steady-state conditions, the minimum allowable plate current is specified as $i_b = 2$ ma and the maximum allowable grid voltage is specified as $e_c = 0$ volts to avoid cutoff and grid current.

Find the peak-to-peak amplitude of the output voltage waveform and the required values of R_k and E_{bb}.
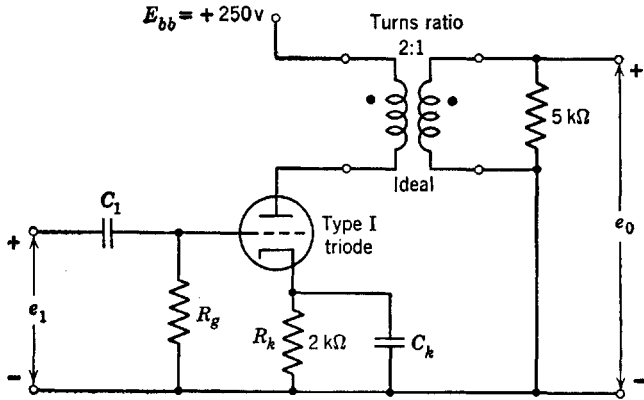
Fig. P8.21

8.26. For the circuit and input waveform shown in Figs. P8.22(a) and (b), sketch and dimension the output waveform, using a Type IV triode. Compare the results with those obtained when a Type V triode is used.
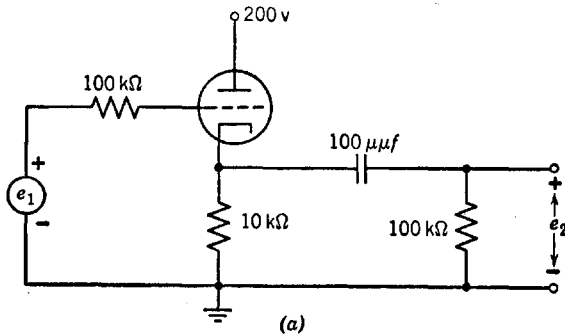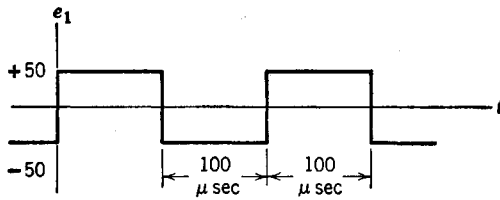
Fig. P8.22

8.27. Using the Type II triode characteristics, plot the locus of the operating point for the circuit and input waveform shown in Fig. P8.23.
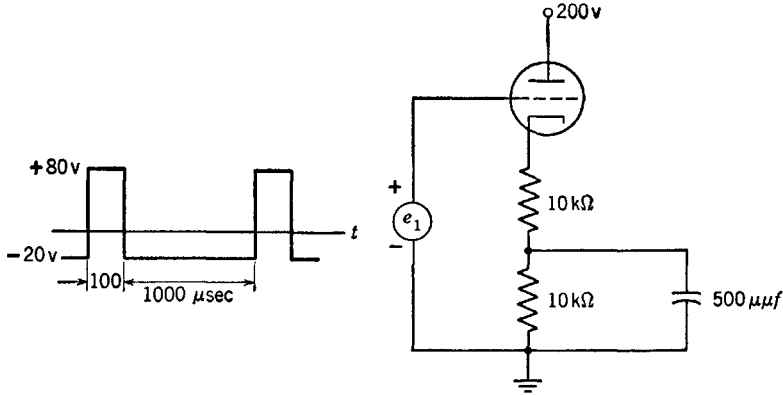
Fig. P8.23

8.28. The triode in the cathode-follower circuit shown in Fig. P8.24 is Type I.

(a) With $e_s = 0$, graphically determine the quiescent operating point.

(b) Determine the values of μ, r_p, and g_m at the operating point of part (a).
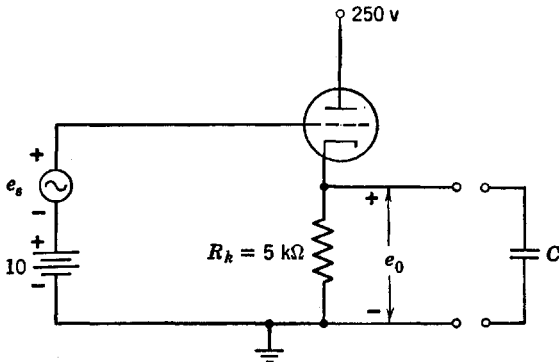
Fig. P8.24

(c) Derive the expression for the incremental voltage gain $\Delta e_0/\Delta e_s$ in terms of μ, r_p, and R_k.

(d) Determine the expression for complex voltage gain if a capacitor C is now connected across R_k.

8.29. Assume a piecewise-linear triode model in the circuit of Fig. P8.25 with R very much larger than r_p and ωRC very much larger than unity.

(a) Sketch a qualitative curve of steady-state d-c output voltage E_2 vs. the amplitude E_1 of the a-c input signal.

(b) Sketch and dimension a quantitative curve of E_2 vs. E_1 based upon the piecewise-linear triode model.

(c) How would you expect an experimental curve to differ from that of part (b)?
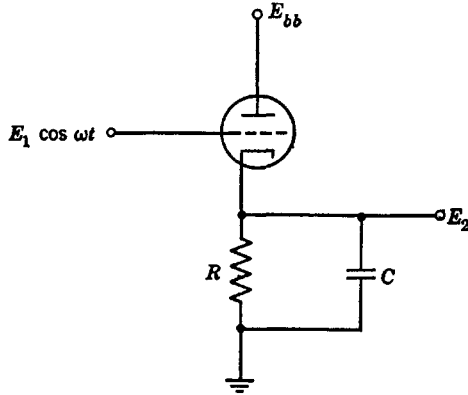
Fig. P8.25

8.30. The tube represented by the linearized characteristics of Fig. P8.26 is used as an "infinite-impedance" detector in the circuit shown in Fig. P8.27. With the circuit at rest, the two pulses shown in Fig. P8.28 are applied to the input.

(a) Plot carefully the output voltage over the time interval $0 < t < 100$ μsec.
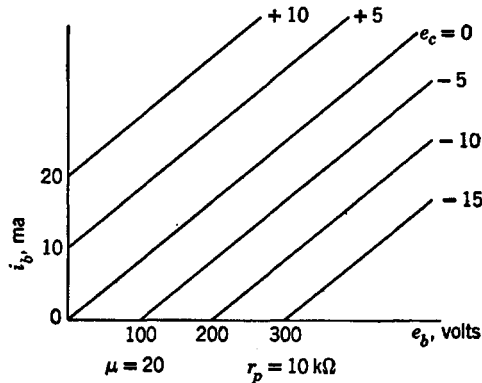
Fig. P8.26

(b) Sketch the tube characteristics and indicate the operating locus for the tube during the interval $20 < t < 55$ μsec. Locate, on the path, points corresponding to the times $t = 20^-, 20^+, 40^-, 40^+, 50^-$ μsec. (Show an enlarged section of the characteristics, if necessary.)

8.31. The circuit shown in Fig. P8.29(a) has the negative-step input indicated in Fig. P8.29(b).

(a) Justify the piecewise-linear model given in Fig. P8.29(c).

(b) Using this model, sketch and dimension $e_2(t)$ for E_1 less than V/μ.

(c) Repeat part (b) for E_1 greater than V/μ. Try to deduce the answers directly from the circuit model.
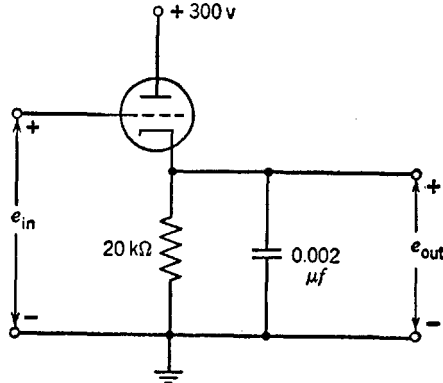
Fig. P8.27
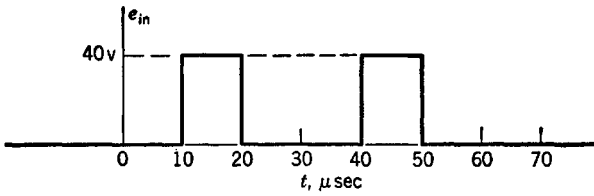
Fig. P8.28
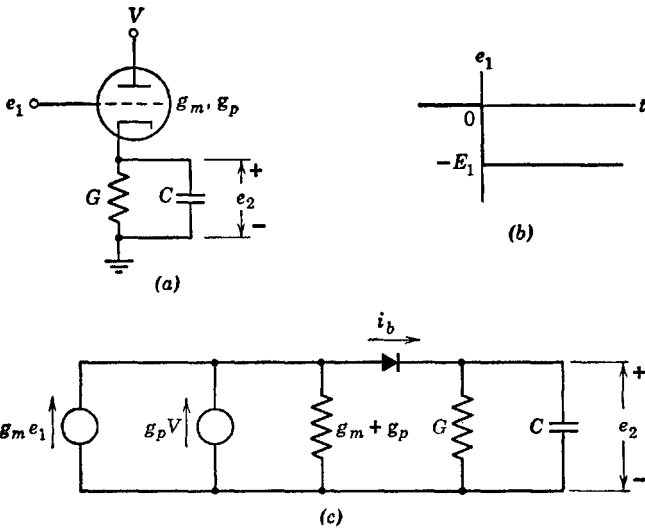
Fig. P8.29

8.32. A cathode follower is used to couple pulses from a source to a capacitive load. The present design, shown in Fig. P8.30(a), yields a satisfactory rise time but the output fall time is too large. The circuit shown in Fig. P8.30(b) has been proposed as one with better response for this application. Sketch and dimension the waveform $e_2(t)$ for each of the two circuits.
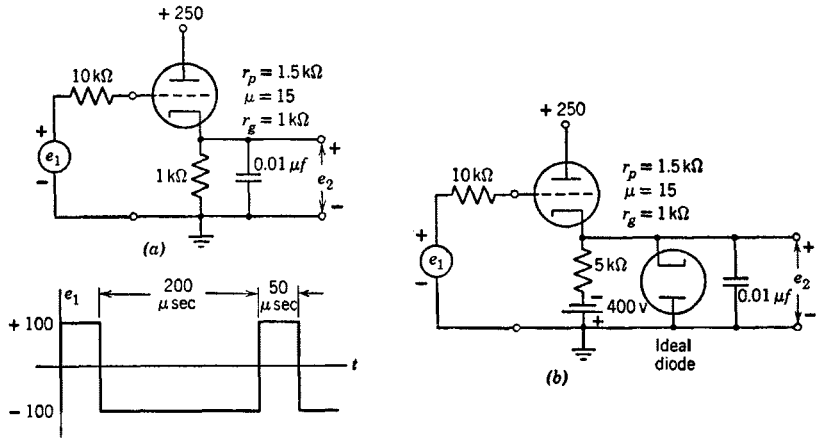
Fig. P8.30

8.33. A triode is operated as a cathode follower with a supply voltage $E_{bb} = 250$ volts. The circuit is used to drive a coaxial transmission line which can be represented here as a resistance $R_k = 2$ kilohms in parallel with a capacitance $C_k = 100$ $\mu\mu$f, the parallel combination being connected between cathode and ground. Other capacitances in the circuit are negligible. The input voltage e_1 (from grid to ground) is originally zero and the system is at rest when a negative step of height E_1 is applied at $t = 0$. Assume that the tube is piecewise-linear with $r_p = 10$ kilohms and $\mu = 20$.

(a) Sketch and dimension the cathode voltage $e_k(t)$ for $E_1 = 1$ volt.

(b) Repeat for $E_1 = 20$ volts.

(c) What is the greatest initial rate of decrease (volts/sec) of $e_k(t)$ obtainable with this circuit for any value of E_1?

8.34. In the circuit of Fig. P8.31, the input voltage $e_1(t)$ is zero for negative time and minus one-tenth volt for positive time. Choose a suitable circuit model to represent the operation of the triode in this problem and explain how to determine suitable values for the parameters in the model. Calculate the output voltage $e_2(t)$ for R_k not equal to zero, and R_k equal to zero.

8.35. A small negative step voltage of height E_1 is applied to the input terminal of the amplifier shown in Fig. P8.32. The input voltage is zero (ground) before the step is applied.

(a) Sketch and dimension the total output voltage $e_2(t)$ as a function of time. Assume that triode plate curves are available and that capacitor C is the only important energy-storage element in the circuit.

(b) Sketch and dimension the log-log plot of E_2 versus frequency when e_1 is a sinusoid of small fixed amplitude and E_2 is the amplitude of the sinusoidal component of the output voltage.
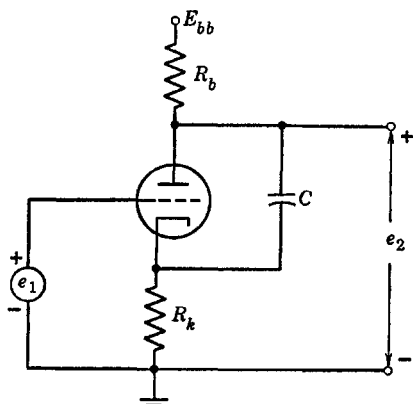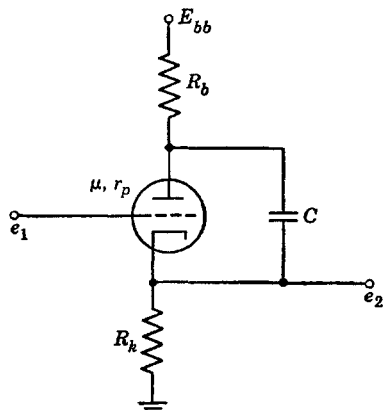
Fig. P8.31   Fig. P8.32

(c) Sketch and dimension $e_2(t)$ for a small negative step applied at e_1, using a piecewise-linear triode model instead of the nonlinear plate curves and with capacitor C replaced by an ideal inductor L.
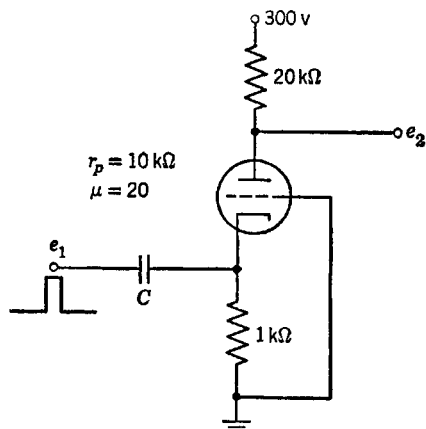
Fig. P8.33

8.36. A single 0.1-volt pulse of 100 μsec duration is applied at e_1 in the circuit of Fig. P8.33. Sketch and dimension:

(a) The cathode voltage $e_k(t)$.

(b) The output voltage $e_2(t)$, for a value of C which makes the time constant τ five times the pulse duration. What is the value of C? What is the qualitative effect of plate-to-cathode capacitance?

8.37. (a) In discussing the response $e_2(t)$ of the circuit in Fig. P8.34(a) to the input wave $e_1(t)$ shown in Fig. P8.34(b), what time constant τ is relevant, assuming that T is large enough so that interelectrode capacitances may be neglected, and that E is such that the tube operates in the linear region of its characteristics? Express τ in terms of R_1, R_2, C, μ, r_p.
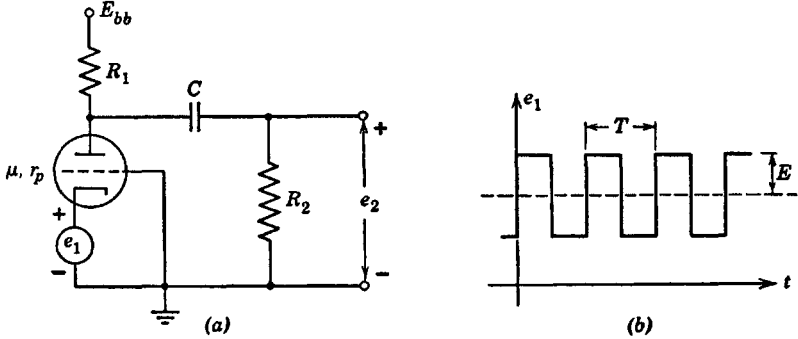
Fig. P8.34

(b) If $\tau = 100T$, i.e., if we can assume ideal coupling, sketch and dimension the output wave $e_2(t)$. Consider $e_1(t)$ to have been on for a long time so that the circuit is in the steady state.

8.38. A triode is operated as a grounded-cathode amplifier with a plate load $R_b = 20$ kilohms. Assume linear operation with $r_p = 20$ kilohms and $\mu = 40$. Use the linear incremental equivalent circuit. A capacitance $C_{gp} = 100$ $\mu\mu$f is purposely connected between the grid and plate terminals and a small input step voltage of height E_1 is suddenly applied at the grid. Other capacitances may be ignored. Sketch and dimension the input current $i_1(t)$ drawn from the input voltage source. On the basis of this waveform deduce an equivalent passive RC circuit which would draw the same input current as the actual tube circuit. Evaluate the constants of this circuit both literally (in terms of r_p, μ, R_b, C_{gp}) and numerically. How much larger than C_{gp} is the capacitance appearing in the equivalent input circuit? (This aggrandization of capacitance is called the Miller effect.)
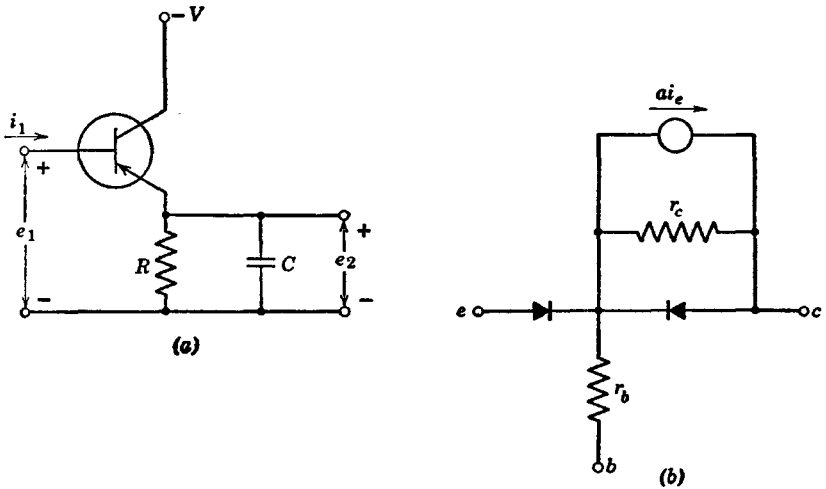
Fig. P8.35

8.39. A transistor is to be used as a common-collector amplifier to drive a capacitive load, as shown in Fig. P8.35(a). If energy storage in the transistor is neglected:

(a) What is the emitter voltage waveform for a small negative step of base current i_1?

(b) What is the small-signal frequency response?
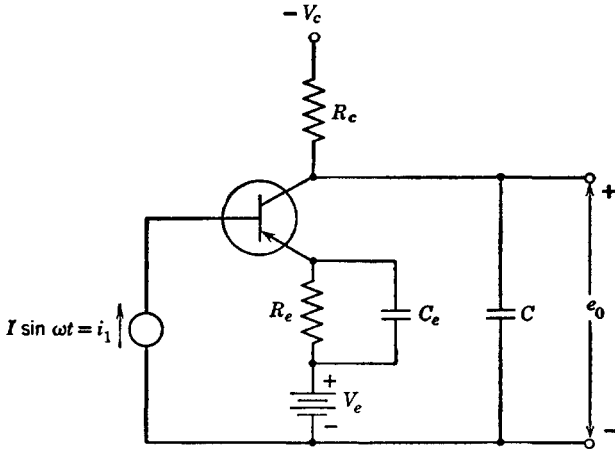
Fig. P8.36

8.40. From the circuit shown in Fig. P8.36, find the frequency at which the output voltage drops to 0.707 of its midband value (that is, its value when capacitor C can be neglected), in terms of r_e, r_b, r_c, a and the circuit parameters. Assume C_e is large enough to be considered an a-c short circuit for all frequencies being considered.
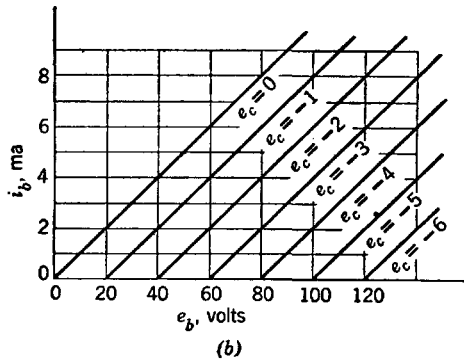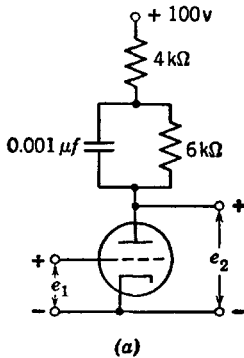
Fig. P8.37

8.41. The plate characteristics of the tube used in the circuit of Fig. P8.37(a) are shown in Fig. P8.37(b). A -4 volt step is applied at e_1 at $t = 0$.

(a) Sketch the operating path of the tube on the plate characteristics.

(b) Sketch and dimension in voltage and time the response $e_2(t)$.

8.42. A Type I triode is used in the amplifier circuit shown in Fig. P8.38. The tube is shunt fed and load is capacity coupled to the plate. A cathode resistor is used to obtain self bias. Determine the operating point for the tube under zero signal conditions.
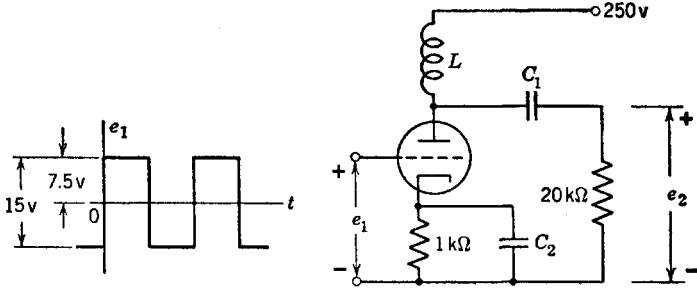
Fig. P8.38

Application of a large grid signal causes the operating point to shift from its quiescent value. Find the new operating point when a square wave of 15 volts peak-to-peak amplitude is applied to the grid. It may be assumed that C_1 and C_2 are sufficiently large so that there is no variation in voltage across their terminals during the signal cycle. Similarly, L is so large that there is no current variation in the inductor during the signal cycle.
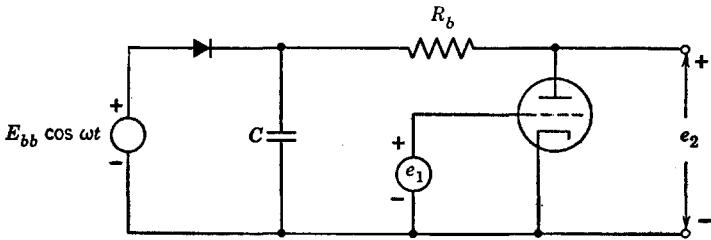
Fig. P8.39

8.43. A triode d-c amplifier has a plate supply voltage generated by a simple rectifier as shown in Fig. P8.39. Use a piecewise-linear triode model.

(a) Calculate the approximate peak-to-peak value of the ripple appearing at the output e_2. Notice that the amplitude of this hum or ripple will depend upon the value of the d-c input voltage e_1. Assume that C is large enough so that the diode conducts during a very small portion of the cycle.

(b) For $E_{bb} = 300$ v, $R_b = 20$ kilohms, $r_p = 10$ kilohms, and $\omega/2\pi = 60$ cycles, what capacitance would you use to keep the ripple below one-tenth volt over the entire range of amplifier operation lying between cutoff and grid current?

(c) Where in the circuit would you insert an additional R and C or an additional L and C in order to obtain appreciable ripple reduction (or the same ripple with smaller total capacitance)?

8.44. In the circuit of Fig. P8.40, $e_1(t)$ is a square wave of period T, maximum value 0, and minimum value -2 volts. Sketch the operating path in the

i_b vs. e_b plane and indicate the rates at which the instantaneous operating point moves about on this locus. Assume that the period T is equal to twice the time constant associated with C_c and twenty times as large as the time
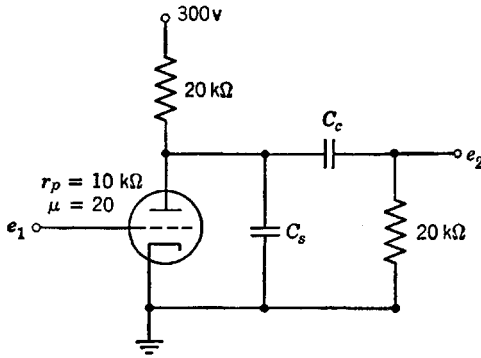
Fig. P8.40

constant due to C_s. Also sketch the actual a-c waveforms of $e_b(t)$ and $e_2(t)$ superimposed upon the idealized waveform which would result for $C_c = \infty$ and $C_s = 0$.

8.45. In Fig. P8.41:

(a) How large must E be in order to cut off the tube during the pulse shown in Fig. P8.42?
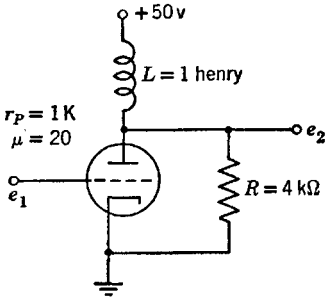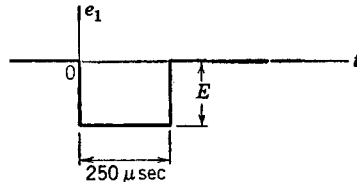
Fig. P8.41   **Fig. P8.42**

(b) With E equal to or greater than the value found in part (a), sketch and dimension the output voltage waveform.

8.46. From Fig. P8.43, determine the amplification e_2/e_1 in terms of the circuit parameters. Neglect C_c and the biasing filter (R_kC_k) but not R_2.

8.47. Figure P8.44 shows an amplifier whose input is a negative 2-volt step.

(a) Derive an incremental expression for the transient of e_0.

(b) By graphical construction determine the initial and final values for the variation of e_0 and superimpose these on the solution to part (a).

8.48. The amplifier in Fig. P8.45 is to deliver $\frac{1}{4}$ watt to the load. The constraints are as given on page 414.
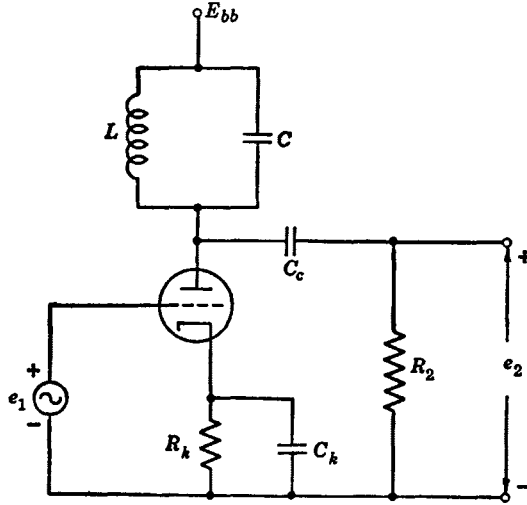
Fig. P8.43
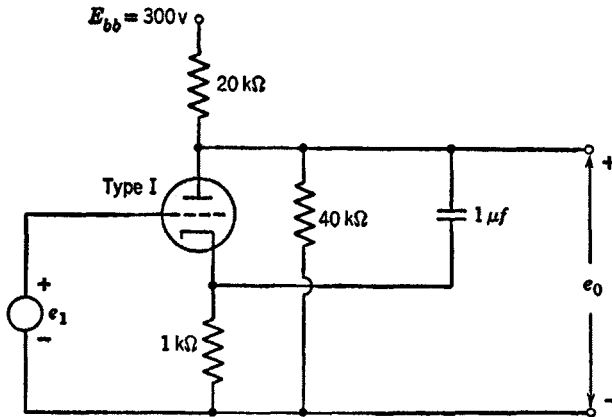
Fig. P8.44

(a) No grid current.

(b) A minimum plate current of 3 ma.

(c) A maximum d-c plate power dissipation of 2 watts.

Determine E_{bb} and n.

8.49. Using a piecewise-linear model for the triode in the circuit shown in Fig. P8.46 with $\mu = 25$, $r_p = 10$ kilohms, determine the operating point. The signal source yields a sinusoidal voltage $v_s = 6 \cos \omega t$. Find the following:

(a) The amplitude of the sinusoidal component of the anode voltage (the so-called anode voltage swing).

(b) The amplitude of the anode current swing.
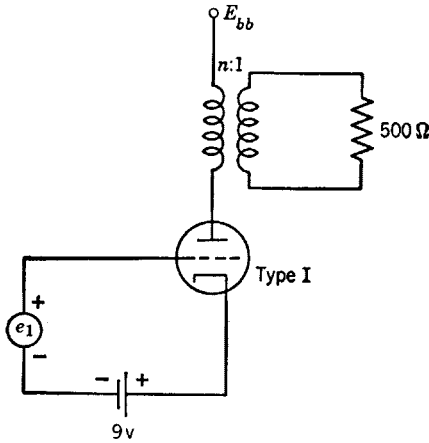
Fig. P8.45

(c) The signal power P_s dissipated in R_a.
(d) The total power dissipated in R_a.
(e) The total power P_t delivered by the plate supply.
(f) The power dissipated in the tube.
(g) The efficiency P_s/P_t.
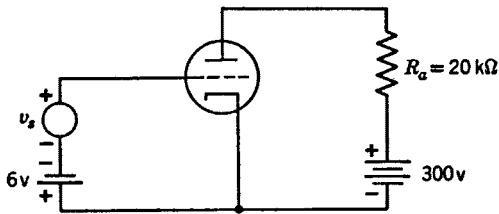
Fig. P8.46

8.50. The triode of Prob. 8.49 is coupled to the load by means of an ideal $1:1$ transformer, as in Fig. P8.47. The value of E_{bb} is chosen so that the tube has the same d-c operating point as in Prob. 8.49. Find, for the same v_s:

(a) The amplitude of the anode voltage swing.
(b) The amplitude of the anode current swing.
(c) The (a-c) power P_s dissipated in R.
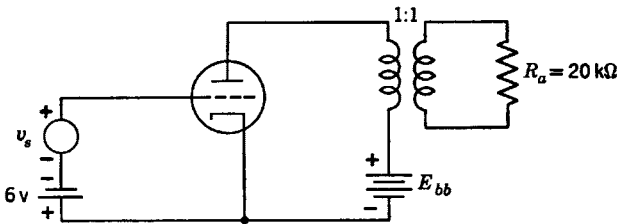
Fig. P8.47

(d) The power P_t delivered by the plate supply.

(e) The power dissipated in the tube.

(f) The efficiency, P_s/Pt.
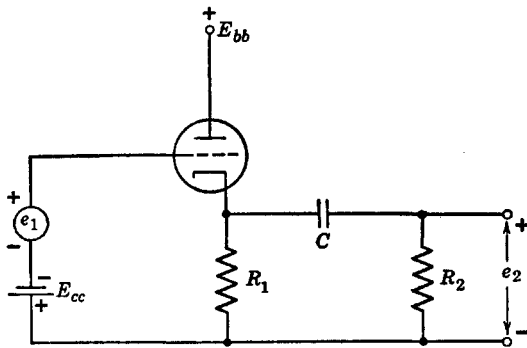
Fig. P8.48

8.51. The triode shown in Fig. P8.48 is to be represented by an incremental model. Operating conditions are such as to prevent flow of grid current.

(a) Assume $R_2 = \infty$. Calculate the incremental gain e_2/e_1 in terms of μ, g_m, and R_1.

(b) Find the value of R_2 for which the gain e_2/e_1 is half of the value found in (a).

(c) From these results, find the internal resistance of the cathode-follower circuit facing C and R_2.

(d) Check this value with the equivalent resistance of R_1 and $r_p/(\mu + 1)$ in parallel. Neglect the effect of C.
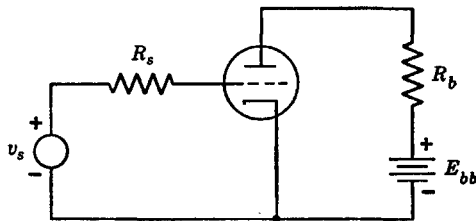
Fig. P8.49

8.52. The triode shown in Fig. P8.49 has zero d-c bias at the grid. It is to be represented by a piecewise-linear model. The value of R_s is equal to r_g for the model. Sketch the anode voltage when the grid circuit is driven by a sinusoidal voltage v_s.

8.53. From Fig. P8.50, sketch e_2 as a function of time, when e_1 is a sinusoidal voltage. No grid current flows and the time constant RC is considerably larger than the period of the input signal.

8.54. In the circuit of Fig. P8.51 the triodes are Type I and the input waveform is repetitive. Assume $r_g = 1$ kilohm for $e_c > 0$.
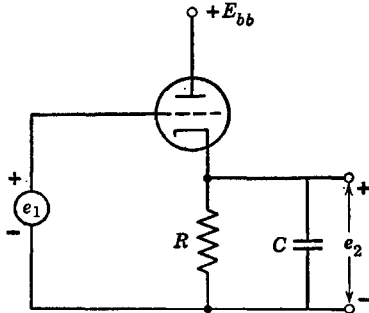
Fig. P8.50

(a) Determine the minimum value of C_0 that will maintain plate-current cutoff in the first triode during interval δ_s.

(b) Determine values for E_k and C to cause the second triode to be driven from plate-current cutoff to the grid-current point during the interval δ_s,
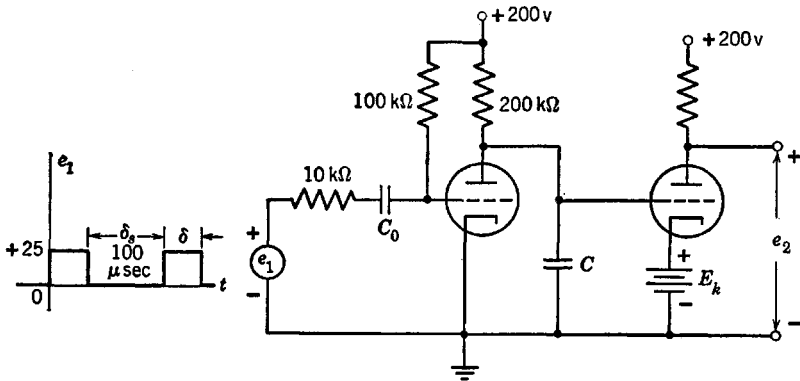
Fig. P8.51

assuming the capacitor C discharges essentially completely during the interval δ.

(c) What is the minimum value of δ that will yield this condition?

(d) Estimate the deviation from linearity of $e_2(t)$ during the interval δ_s.

8.55. A small-amplitude input signal $e_1 = E_1 \cos \omega t$ is applied to the circuit of Fig. P8.52. Assume that $\omega R_k C$ and $\omega L/R_b$ are very large. Sketch the operating path in the i_b vs. e_b plane.

8.56. In the circuit of Fig. P8.53 assume that the susceptances of stray and interelectrode capacitances and the reactances of the large coupling capacitances C_1, C_2, and C_k are negligibly small. Use a piecewise-linear triode model with $r_p = 10$ kilohms, $r_g = 2$ kilohms, and $g_m = 2000$ microhms. The input voltage e_1 is a square wave of amplitude E_1. The resulting output voltage, under the above assumptions, is also a square wave, of amplitude E_2.

(a) For $E_1 = 0$, show that the static operating point lies at the intersection
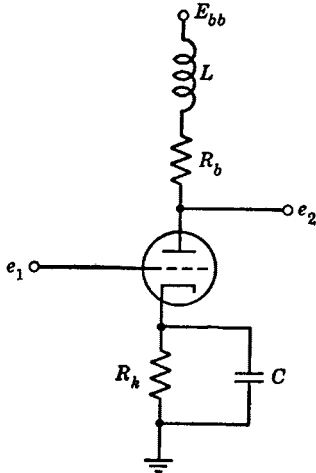
Fig. P8.52
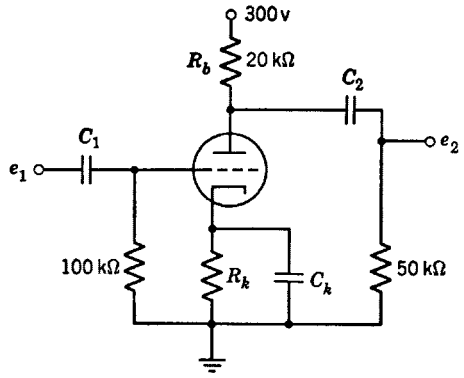
Fig. P8.53

of the d-c load line $E_{bb} = e_b + (20 + R_k)i_b$ and the d-c bias curve $e_c + R_k i_b = 0$. On the piecewise-linear plate curves, what is the slope of the bias curve in terms of r_p, μ, and R_k?

(b) Choose R_k so that the tube will reach cutoff and grid current for the same value of E_1. What is this value of E_1? Sketch and dimension the operating path of the tube on the piecewise-linear plate curves.

(c) Using the a-c incremental model, find the gain $A = E_2/E_1$.

(d) With $R_k = 0$, plot E_2 vs. E_1. The grid circuit now acts as a rectifier. Carry the plot beyond the value of E_1 for which cutoff occurs during the cycle.
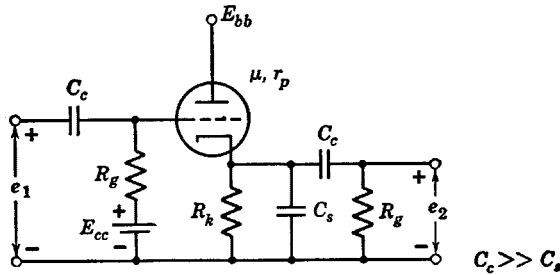
Fig. P8.54

8.57. The circuit of Fig. P8.54 is a cathode-follower amplifier with a total effective shunt capacitance C_s.

(a) What is the midband gain $\Delta e_2/\Delta e_1$?

(b) What is the lower half-power frequency ω_1?

(c) What is the upper half-power frequency ω_2?

(d) Sketch and dimension $\log |E_2/E_1|$ vs. $\log \omega$ over the region of interest.

8.58. Find the operating point for the Type II triode in the circuit of Fig. P8.55. Find the incremental parameters of the tube at this operating point.

Use the incremental parameters to find the amplitude and phase of the a-c components of i_b, e_b, and e_k that are developed when a small sinusoidal signal
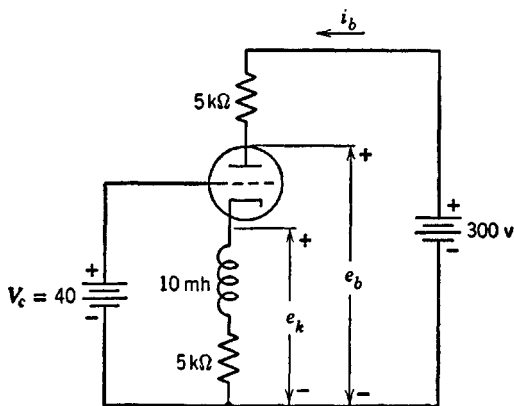
Fig. P8.55

$e_0 \sin \omega t$ is connected in series with the voltage source V_c. Summarize your results by drawing an appropriate vector diagram.

The sinusoidal signal connected in series with V_c is replaced by a step function generator which produces a one-volt step at $t = 0$. Employing incremental analysis, find the expression for e_b as a function of time.
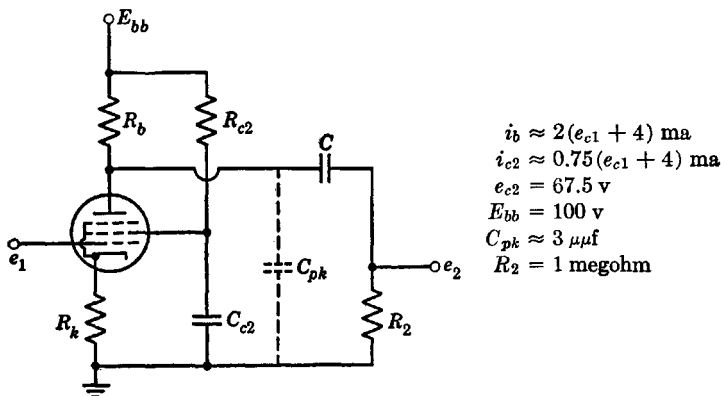
$$i_b \approx 2(e_{c1} + 4) \text{ ma}$$
$$i_{c2} \approx 0.75(e_{c1} + 4) \text{ ma}$$
$$e_{c2} = 67.5 \text{ v}$$
$$E_{bb} = 100 \text{ v}$$
$$C_{pk} \approx 3 \text{ μμf}$$
$$R_2 = 1 \text{ megohm}$$

Fig. P8.56

8.59. The pentode circuit shown in Fig. P8.56 is to be used to amplify rectangular pulses of duration T and height E_1. Choose R_k and R_{c2} so that the quiescent operating point ($e_1 = 0$) lies halfway between $e_{c1} = 0$ and cutoff and so that $e_{c2} = 67.5$ volts.

(a) With C_{c2} and C assumed arbitrarily large, the output voltage e_2 rises with a time constant τ_r toward a value E_2, in response to the leading edge of a negative pulse of height E_1. Discuss the choice of R_b for maximum gain

$A = E_2/E_1$ and minimum rise time τ_r. For a gain of 12.5, what are R_b and τ_r? What are the quiescent values of plate voltage and plate current? What is the minimum pulse length T which can be amplified with reasonable fidelity?

(b) Still assuming C_{c2} arbitrarily large, but ignoring C_{pk}, explain the effect of C upon the shape of the output voltage response when a very long pulse is applied. Suggest approximate values for C_{c2} and C such that a pulse 100 μsec in length shall not droop by more than 20 per cent.
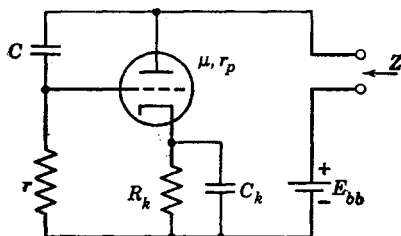
Fig. P8.57

8.60. Determine the incremental impedance Z for Fig. P8.57. Neglect the bias filter $R_k C_k$.
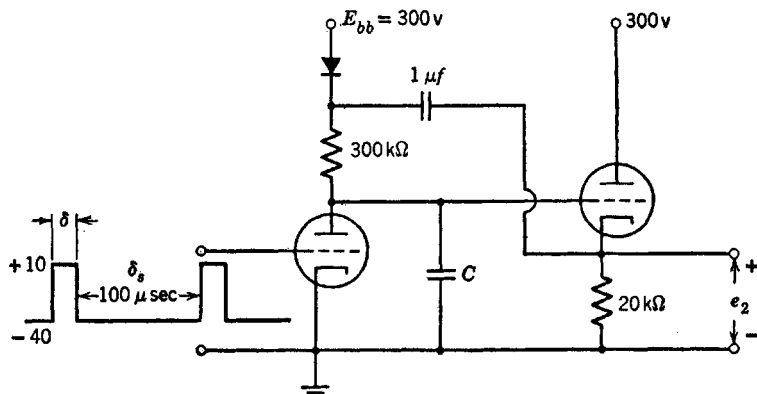
Fig. P8.58

8.61. The circuit of Fig. P8.58 is expected to yield a nearly linear sweep voltage during the interval $\delta_s = 100$ μsec. The triodes can be represented by a piecewise-linear model with $\mu = 20$, $r_p = 10$ kilohms and $r_g = 1$ kilohm.

(a) What is the maximum linear sweep amplitude E_s available at the output terminals?

(b) What value of C will yield this amplitude E_s in the interval $\delta_s = 100$ μsec?

(c) What is the sweep linearity? (Ratio of slope at $t = \delta_s$ to slope at $t = 0$.)

8.62. The circuit shown in Fig. P8.59 is to be used as a frequency meter. Input voltage consists of a square wave that alternately drives the grid voltage to zero and beyond plate-current cutoff. The diodes have a resistance of 1 kilohm when conducting. The milliammeter may be assumed to indicate

average current and the meter resistance is included in the 1-kilohm resistor shown. The effects of tube and wiring capacitances may be neglected.
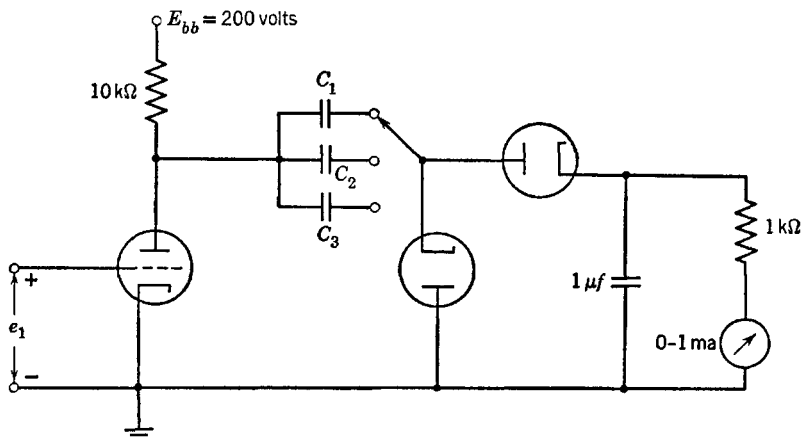
Fig. P8.59

Determine the values of C_1, C_2, and C_3 required to yield full-scale meter deflection for input frequencies of 10^5, 10^4, and 10^3 cps, respectively.

8.63. Referring to Problem 8.62, assume the milliammeter to be linear and to have a scale that can be read to within 0.5 per cent of the full scale reading.

Determine the maximum percentage error in frequency determination resulting from an assumed linear relationship between frequency and meter current.

Determine and plot a correction curve (deviation from linearity versus per cent of full scale deflection) for the high-frequency scale. Is the same curve applicable to the other two scales.

8.64. The triodes in the circuit of Fig. P8.60 are Type I. Sketch and dimension the waveform of voltage e_3 for input waveforms e_1 and e_2 as shown. The diodes may be assumed to have 1-kilohm forward resistance and infinite back resistance.

8.65. The 2:1 step-down transformer in the circuit of Fig. P8.61 has a magnetizing inductance $L_m = 100$ mh referred to the primary winding. Neglecting the effects of leakage inductance and stray capacitances, sketch and dimension the waveform of output voltage e_2. The 2-μsec input pulses occur at a repetition frequency of 500 per second.

8.66. The triode in the circuit of Fig. P8.62(a) can be represented by a simple piecewise-linear model with $\mu = 20$, $r_p = 10$ kilohms, and $r_g = 1$ kilohm. The 2:1 turns-ratio pulse transformer has a magnetizing inductance $L_m = 100$ mh referred to the primary winding. Leakage inductance and stray capacitance can be neglected.

The desired output is the repetitive flat-topped waveform shown in Fig. P8.62(b).

(a) Sketch and dimension a suitable input waveform to yield the flat-topped output pulse.

(b) Determine the amplitude and time constant of the negative overshoot on the output waveform.
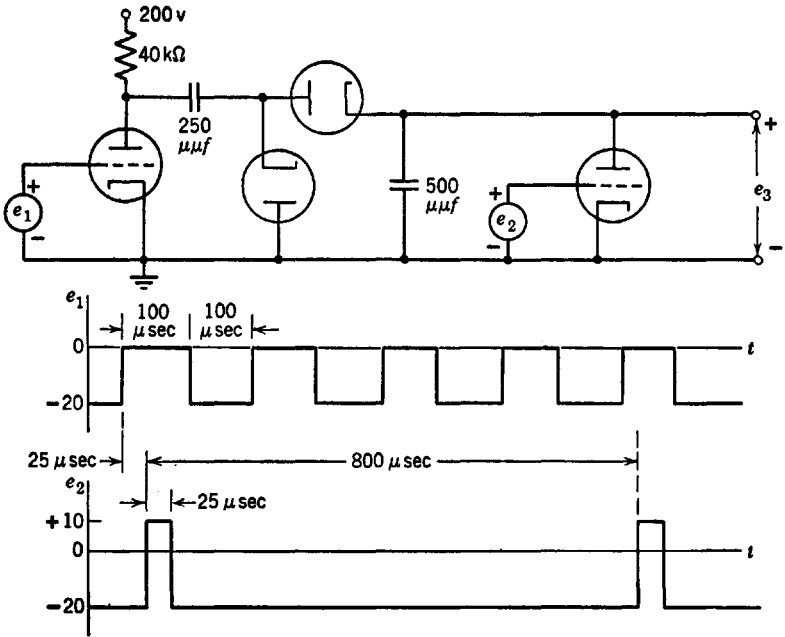
Fig. P8.60
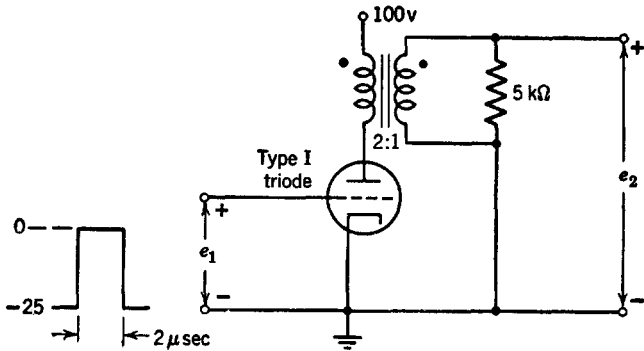
Fig. P8.61
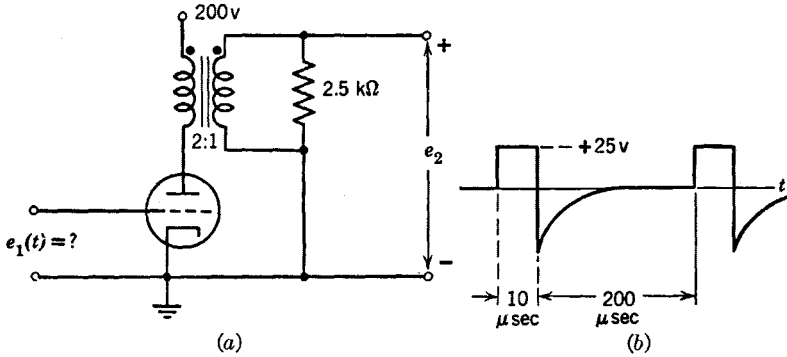
Fig. P8.62

8.67. Using the Type I triode curves, plot the locus of the operating point in the i_b vs. e_b plane of the first stage of the circuit shown in Fig. P8.63. Assume $L_m = 100$ mh as seen from the primary of the 2:1 pulse transformer.
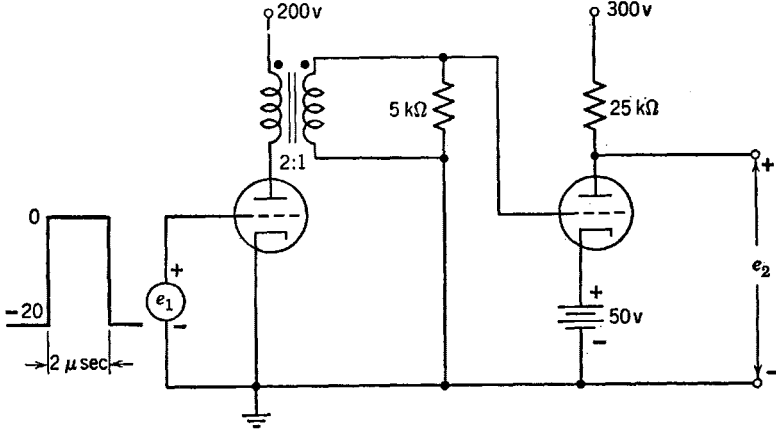
Fig. P8.63

Neglect leakage inductances and stray capacitances. The input waveform shown has a repetition rate of 1000 per second. Assume $r_g = 1$ kilohm for $e_c \geqslant 0$.

8.68. The input voltage e_1 for the circuit of Fig. P8.64 cuts off T_1 for an interval $\tau_s = 100$ μsec and during the remainder of the 1000 μsec repetition interval T_1 conducts with grid-cathode voltage equal to zero. The push-pull output, direct-coupled to the cathode-ray tube deflection plates, is required to produce a properly centered 4-inch sweep deflection. (Deflection factor is 50 volts per inch.) Using the Type I triode characteristics determine graphically the required value of E_3 and calculate the value of C required.

8.69. The ringing circuit shown in Fig. P8.65 is coupled to a cathode follower to minimize loading of the tuned circuit. The rectangular input wave e_1 has a repetition frequency of 5000 cycles per second. Type I triodes are used.

(a) Determine the values of L and C required to produce a ringing oscillation which has a frequency of 100 kc and a peak-to-peak amplitude of 100 volts (neglecting damping) at the output e_2.

(b) Determine the proper position for the bias tap on the cathode follower load resistor (total $R_k = 10$ kilohms) to avoid clipping.

(c) Determine the minimum amplitude E required of the input pulse.

(d) Determine the additional circuit elements required to provide the ringing tube plate voltage (100 volts) from a 200-volt supply.
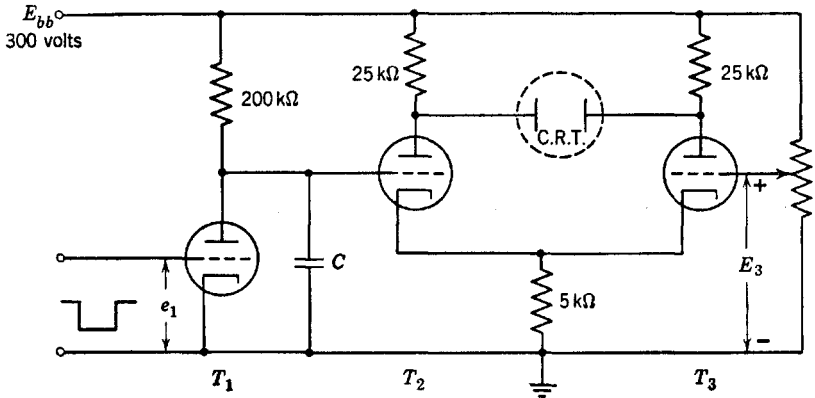
Fig. P8.64
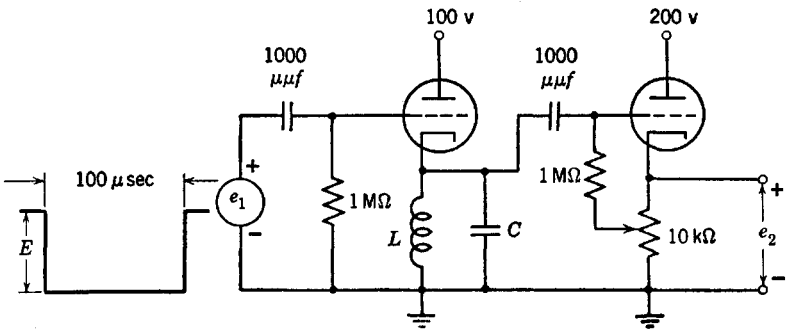
Fig. P8.65

8.70. Assume the diode in the circuit of Fig. P8.66 to be ideal (forward resistance zero and back resistance infinite). At $t = 0$, current in the inductance L is I_0 and capacitor voltages are zero. Sketch and dimension the waveforms of e_1 and e_2 for $t > 0$.

8.71. The 100-μsec rectangular input pulse for the ringing circuit shown in Fig. P8.67 has a repetition frequency of 2000 cps. The output voltage is to be squared and peaked to form timing markers from the positive-going edge of each ringing cycle.
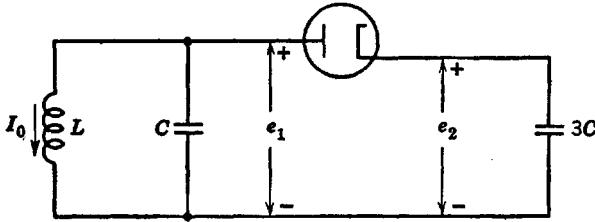
Fig. P8.66

(a) What value of C is required to produce pulses at 10-μsec intervals?

(b) What is the minimum value of input pulse amplitude E necessary for satisfactory operation?
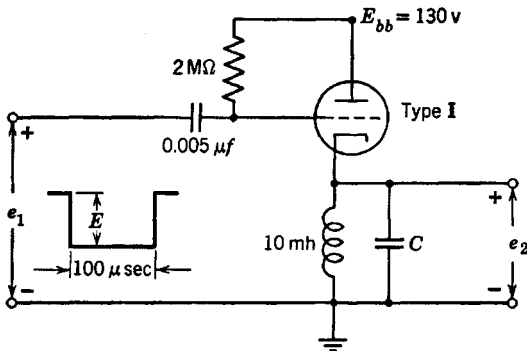
Fig. P8.67

8.72. In the circuit shown in Fig. P8.68, current through the deflection coil is required to vary linearly with time from 10 ma to 110 ma during the interval $\tau_s = 100$ μsec. The repetition interval is $\tau_r = 1000$ μsec.

Using the Type I triode and Type II pentode characteristics, determine the values of R and C required in the sweep generator circuit.

8.73. Sketch and dimension the output voltage waveform $e_2(t)$ for the circuit shown in Fig. P8.69. Magnetizing inductance seen from the primary of transformer T_1 is $L_m = 100$ mh. Neglect leakage inductances and stray capacitances. Assume $r_g = 1$ kilohm for $e_c > 0$. The diode has 1 kilohm forward resistance and infinite back resistance.

8.74. From the circuit given in Fig. P8.70, find, for the curve $|A(j\omega)|$,

(a) resonant frequency (cps);

(b) Q;

(c) bandwidth (cps);

(d) A at resonance.

8.75. An inductor is representable by a pure inductance L in series with a resistance R_L. Consider the impedance $Z(j\omega)$ of the parallel combination of the inductor, a pure capacitance C, and a resistance R_C. Let $\omega_0 = 1/\sqrt{LC}$,
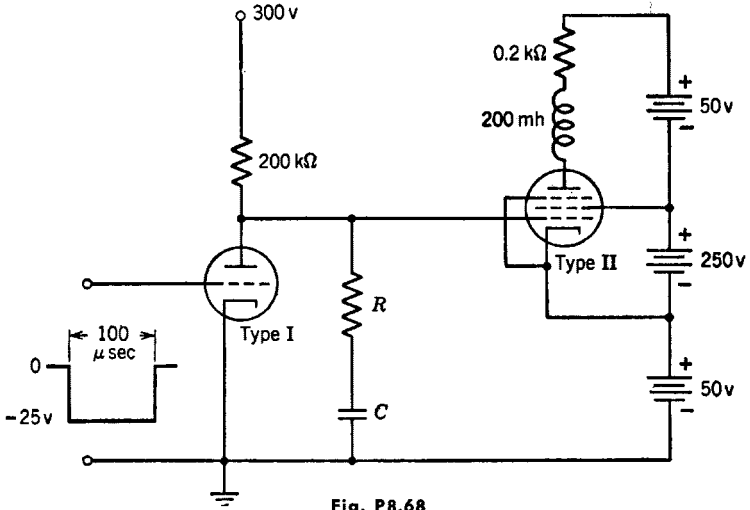
Fig. P8.68
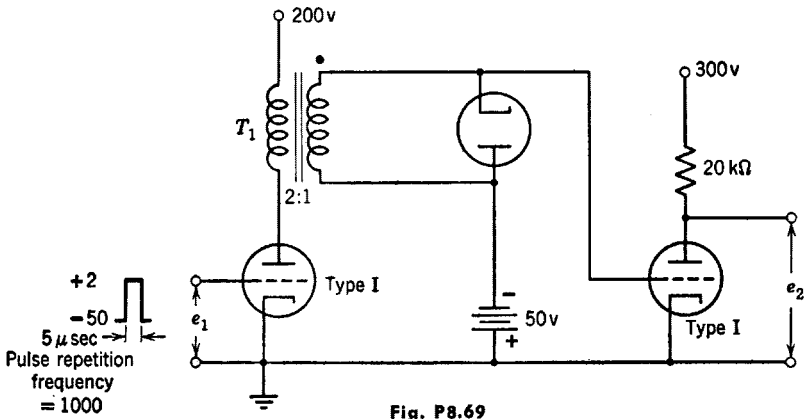
Fig. P8.69
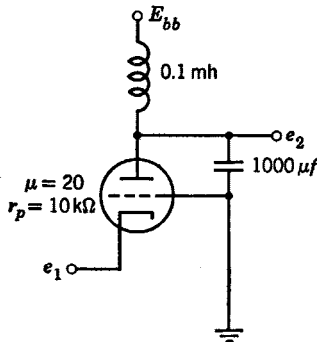
Fig. P8.70

$R_0 = \sqrt{L/C}$, and assume $R_L \ll R_0 \ll R_C$. For frequencies near ω_0, show that

$$Z(j\omega) \approx \frac{R_0 Q}{1 + jQ\left(\dfrac{\omega}{\omega_0} - \dfrac{\omega_0}{\omega}\right)} \approx \frac{R_0 Q}{1 + j2Q\left(\dfrac{\omega}{\omega_0} - 1\right)}$$

where

$$\frac{1}{Q} = \frac{1}{Q_L} + \frac{1}{Q_C}, \qquad Q_L = \frac{R_0}{R_L}, \qquad Q_C = \frac{R_C}{R_0}$$
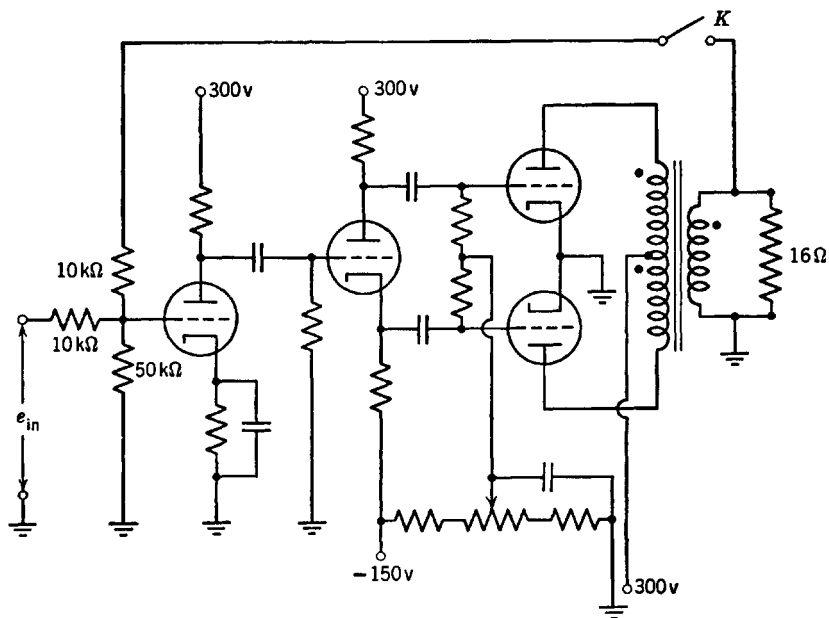
Fig. P8.71

8.76. Complete the design of the amplifier circuit shown in Fig. P8.71 by specifying the tube types, all resistors (ohms and watts), capacitors and turns ratio of the output transformer. The amplifier is to fulfill the following specifications:

(a) The final stage is designed to deliver maximum power with small distortion and no grid current, this power to be 5–10 watts.

(b) The transformer when in the circuit is expected to cause half-power frequencies of 50 cps and 10,000 cps. Make the rest of the amplifier flat in frequency response from 5 cps to at least 100,000 cps.

After designing the amplifier calculate the value of maximum power into the load and the magnitude of sinusoidal voltage e_{in} necessary to produce it.

8.77. (a) Assume the amplifier of Problem 8.76 to be linear. Calculate the equivalent circuit faced by the 16-ohm load resistor.

(b) Do the same with switch K closed.

(c) If the gain of the driver stage should change (due to tube aging) 10 per cent, what per cent change in output power would this 10 per cent change cause for switch K open? Switch K closed?

Waveform Generation

9.1   Introduction

The circuits that generate the various waveforms used in electronic systems are called oscillators. Very good approximations to the idealized waveforms such as rectangular waves and sine waves are readily generated by rather simple oscillator circuits. Other waveforms can be derived by wave-shaping operations performed on one of the basic waveforms. It is apparent that there is some overlap between the functions called wave shaping and actual waveform generation. We shall distinguish between them by considering the generator or oscillator to be the simplest circuit that will generate a sustained repetitive waveform of voltage or current. Additional circuitry then falls into the category of wave-shaping equipment.

9.2   Properties of Oscillators

Oscillators can be designed to produce power outputs ranging from a few microwatts to many kilowatts, frequencies ranging from a fraction of a cycle per second to thousands of megacycles per second, and waveforms ranging from sine waves to markedly discontinuous waves. The

three components required for generating self-sustained oscillations are a power source, an energy-storage system, and a control valve capable of releasing power from the source to the energy-storage system. Control of the power source may be exercised by vacuum tubes, gas-filled tubes, transistors, or any other device that can provide incremental power gain. Inductors, capacitors, electromechanical transducers (such as piezoelectric crystals or magnetostriction rods), or combinations of these, can be used as energy-storage systems. In electromechanical transducers the oscillations are actually in the form of mechanical vibrations, and the energy storage is in the form of potential and kinetic energy. Nevertheless, such devices display an equivalent inductance and capacitance at their electrical terminals. For our purposes we need not go beyond the electrical description of the energy-storage system in terms of inductance and capacitance.

The energy-storage system plays a major role in determining the waveform of the oscillator output. For example, a single inductance or capacitance tends to produce distinctly nonsinusoidal waveforms called relaxation oscillations. When the energy-storage system includes both inductance and capacitance, the oscillator tends to produce a nearly sinusoidal waveform. An ideal lossless LC circuit will sustain sinusoidal oscillations if the capacitor is initially charged or if the inductance has an initial current. Since any physical circuit includes resistance, the natural oscillations are damped sinusoids for a physical inductor and capacitor. The incremental power gain provided by the control device must supply the loss of power due to dissipation in the resistance associated with the LC circuit, plus any power to be delivered to the oscillator load. The behavior of oscillator circuits with three- or four-terminal energy-storage circuits is conveniently represented by an amplifier coupled to the energy-storage elements as shown in Fig. 9.1(a). The energy-storage elements are very likely to be integrally associated with the amplifier circuit. Detailed analysis of such oscillators evolves quite naturally from a consideration of the stability of linear amplifiers.

An oscillator circuit in which the major energy-storage elements are connected between a single pair of terminals can be analyzed most conveniently in terms of incremental negative resistance. In contrast with positive resistance which dissipates power, negative resistance can be considered as a source of incremental power, thus providing a mechanism for sustaining oscillations. A physical circuit that produces a negative resistance must be nonlinear since a linear negative resistance implies an infinite source of power.

In Chapters 5 through 8, we discussed a number of simple circuits that exhibit incremental power gain. When this transfer property exists in

a circuit, a negative driving-point resistance may also be obtained. If we include transformers among our circuit elements, any circuit that can provide incremental power gain from one pair of terminals to another pair can be modified to yield a negative incremental resistance
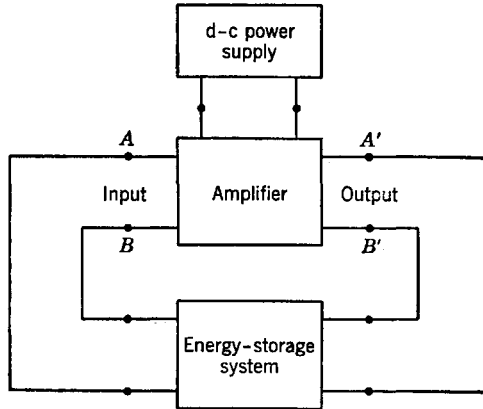
(a)
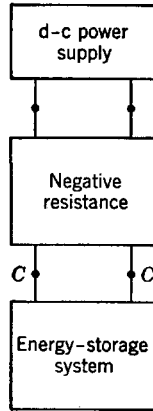
(b)

Fig. 9.1. Waveform generation based on incremental power gain or incremental negative resistance.

over a part of the driving-point curve that exists at a single pair of terminals.

Relaxation oscillators and many types of sine wave oscillators fall into the category shown in Fig. 9.1(b) and hence are readily analyzed by

using the negative-resistance concept. Before proceeding to the details of oscillator behavior, let us examine a number of circuits that produce a negative resistance.

9.3 Negative Resistance in the Common-Base, Point-Contact Transistor Circuit

A simple illustration of negative incremental resistance is provided by a common-base, point-contact transistor circuit. As shown in Fig. 9.2(a) it is a common-base circuit with an external resistance R_b added between base and ground. Because the short-circuit current gain of a point-contact transistor is greater than unity ($\alpha = 2$ to 3 for $i_e > 0$) any of the single-transistor circuits can exhibit a negative incremental resistance. For a circuit that includes only one junction transistor ($\alpha < 1$) a transformer is necessary to obtain negative resistance.

The expressions given in Fig. 9.2 for slopes, intercepts, break points, and the relation between v_1 and i_e can be derived by break-point analysis.

The curve shown in Fig. 9.2(d) determines voltage uniquely if the current i_e is specified, but yields multiple values of current i_e for some values of voltage v_1. Such a curve is called a *current-controlled negative resistance*.

9.4 Negative Resistance in a Common-Emitter, Point-Contact Transistor Circuit

The input curve of a common-emitter circuit using a point-contact transistor provides another example of negative resistance. The schematic diagram is shown in Fig. 9.3(a), and the circuit model is shown in (b). The equation in (c) relating v_1 and i_b can be readily derived from the circuit model. The circuit differs only in numerical values (particularly that of current-gain α) from the corresponding junction transistor circuit.

In this circuit, no values of the input variables v_1 and i_b ever yield the conditions for state I; hence, only II, III, and IV are pertinent. The curve of Fig. 9.3(d) shows that the current i_b is a single-valued function of v_1, whereas the voltage can have multiple values over a part of the range of i_b. The unique specification of i_b in terms of v_1 leads to the designation of this type of curve as a *voltage-controlled negative resistance*.

When piecewise-linear models are used to approximate nonlinear devices, break-point analysis must be supplemented by physical reasoning. Suppose two different circuits yield driving-point curves such as

(a)

(b)

$$(c) \qquad v_1 = i_e \left[r_e + \underbrace{\frac{(r_b + R_b)[R_c + r_c(1-a)]}{r_b + R_b + r_c + R_c}}_{\text{Slope } R} \right] - \underbrace{\frac{(r_b + R_b)V_{cc}}{r_b + R_b + r_c + R_c}}_{\text{Intercept } V}$$

State	r_e	r_c	a
I	r_{er}	r_{cf}	0
II	r_{er}	r_{cr}	0
III	r_{ef}	r_{cr}	a
IV	r_{ef}	r_{cf}	a
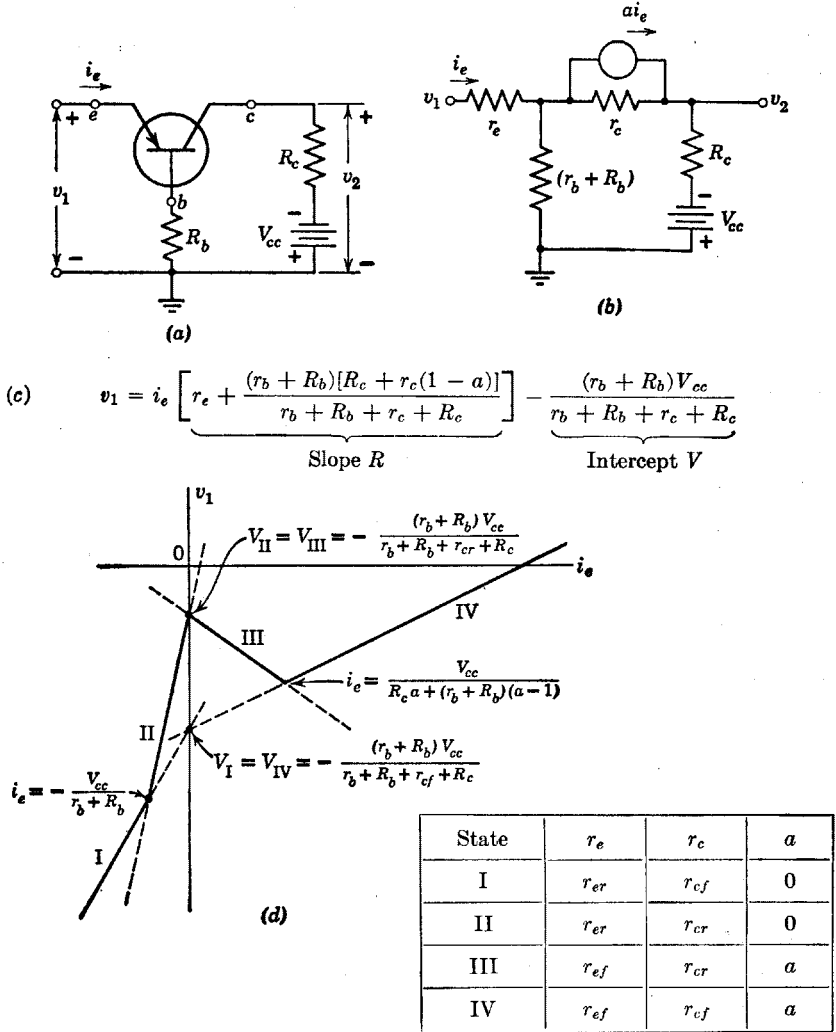
Fig. 9.2. Common-base point-contact transistor circuit.

those shown in Fig. 9.4 (a) and (b). Analysis of these circuits might lead to the curves shown in (c) and (d) where break-point coordinates and slopes (resistances) are much the same. From the three state lines alone, it is impossible to tell whether the curve is voltage controlled as in (e) or current controlled as in (f). We must consider the physical behavior of the circuit to determine the sequence of linear states as current or voltage is varied.

(a)   (b)
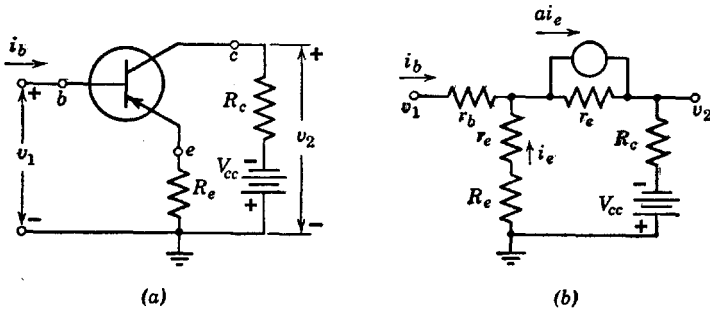
$$(c) \quad v_1 = i_b \left[r_b + \frac{(r_e + R_e)(r_c + R_c)}{r_e + R_e + r_c(1 - a) + R_c} \right] - \frac{V_{cc}(r_e + R_e)}{r_e + R_e + r_c(1 - a) + R_c}$$
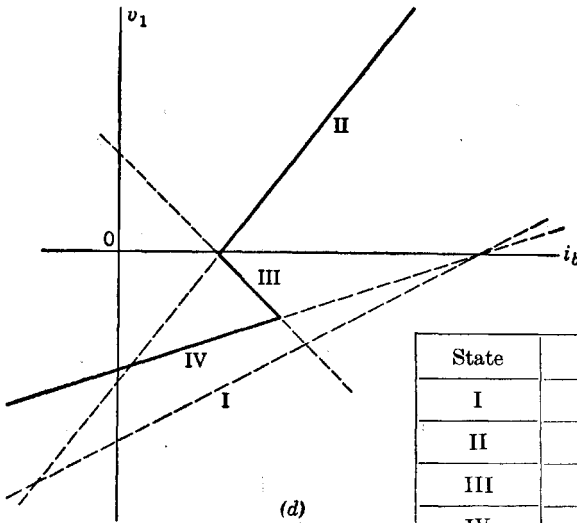
State	r_e	r_c	a
I	r_{er}	r_{cf}	0
II	r_{er}	r_{cr}	0
III	r_{ef}	r_{cr}	a
IV	r_{ef}	r_{cf}	a

Fig. 9.3. Common emitter point-contact transistor circuit.

9.5 Negative Resistance in a Series Triode Circuit

The series-connected triodes shown in Fig. 9.5(a) provide another example of a simple circuit which exhibits a negative-resistance region in a driving-point curve. If the triodes are represented by piecewise-linear models, the break-point method can be used to determine the state lines in the curve of i vs. v. Since total plate current (i_{b1} or i_{b2}) is unidirectional, the only possible direction of current in either of the resistors R is downward. Thus e_{c1} and e_{c2} cannot exceed zero, and the

only triode states that are of interest are those for $e_c \leqslant 0$ as tabulated in Fig. 9.5(b).

The circuit models shown in Fig. 9.5(c) and (d) specify the relation between v and i in states I and II. It is apparent from these models that triode T_2 conducts with zero grid voltage when the circuit is in state I.
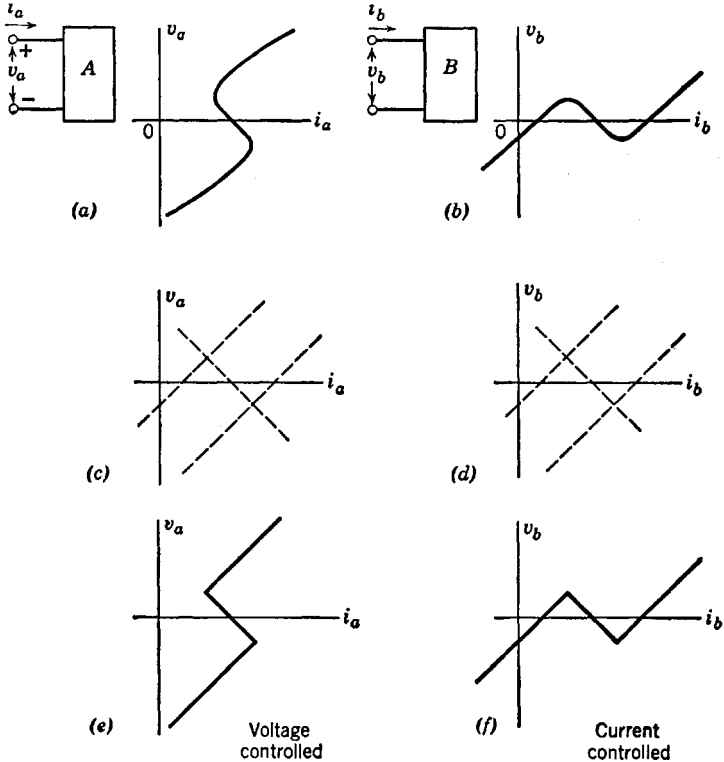
Fig. 9.4. Voltage-controlled and current-controlled negative resistance.

and T_1 conducts with zero grid voltage when the circuit is in state II, The plate-current cutoff condition

$$e_{c1} = -\frac{e_{b1}}{\mu} \tag{9.1}$$

marks the break point between states I and III, while

$$e_{c2} = -\frac{e_{b2}}{\mu} \tag{9.2}$$

(a)

States $(e_c \leq 0)$	T_1	T_2
I	Off	On
II	On	Off
III	On	On

(b)

(c)

(d)

(e)

$$v = \frac{E_{bb}(R + r_p)}{r_p + R(\mu + 1)}$$

State II $\quad i = \frac{v}{R + r_p}$

State III $\quad i = \frac{v - E_{bb}/2}{[r_p - R(\mu - 1)]/2}$

$\frac{E_{bb}}{r_p + R(\mu + 1)}$

$-\frac{E_{bb}}{r_p + R(\mu + 1)}$

$\frac{E_{bb}}{2}$

E_{bb}

$$v = E_{bb}\left[1 - \frac{R + r_p}{r_p + R(\mu + 1)}\right]$$

State I $\quad i = \frac{v - E_{bb}}{R + r_p}$
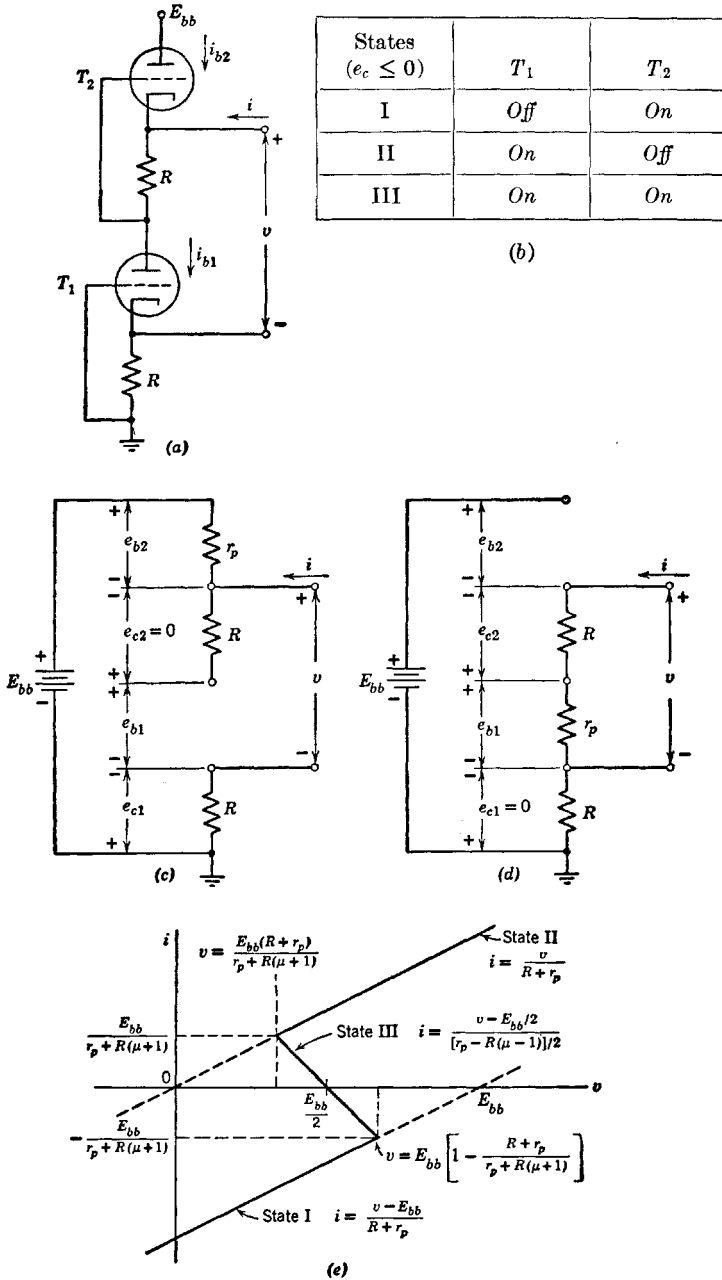
Fig. 9.5. Series-connected triodes.

marks the break point between states II and III. The i vs. v curve plotted in Fig. 9.5(e) is obtained by drawing the lines for states I and II, locating the break points, and joining them to specify the state III line.
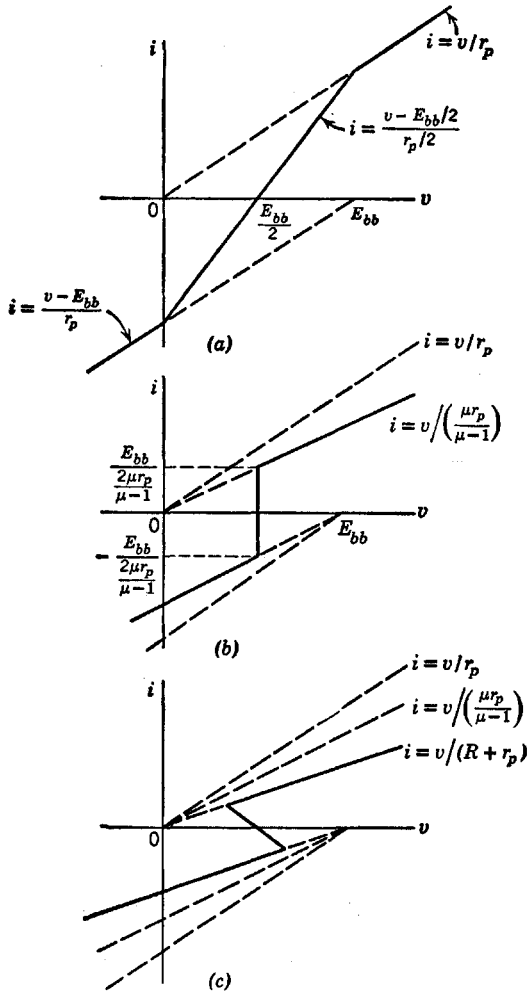
Fig. 9.6. Effects of variations in R on the i vs. v curve of the series-connected triode circuit. (a) $R = 0$; (b) $R = r_p/(\mu - 1)$; (c) $R > r_p/(\mu - 1)$.

The equation for i vs. v in state III can be written directly from the break-point coordinates. The intercept at $E_{bb}/2$ is logical, in view of the circuit symmetry with $i = 0$ in state III. The portions of the

state I and II lines that are pertinent are found by considering the values of v and i needed to hold the circuit in the specified state.

The slope of the state III line will be negative for R greater than $r_p/(\mu - 1)$. With R equal to $r_p/(\mu - 1)$, the apparent resistance in state III is zero; and for R less than this value, the resistance becomes positive, as shown in Fig. 9.6. The variation of R shows the effect of a circuit parameter on the negative-resistance curve. The effect of changing μ or r_p can be considered in a similar manner.

$$R_T = \frac{dv}{di} = R_N + R_2$$
$$= -R_1 + R_2$$

(a)
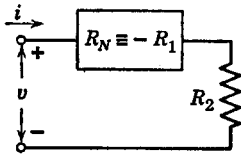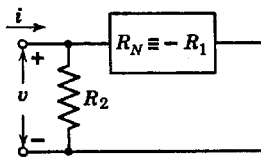
$$R_T = \frac{dv}{di} = \frac{R_N R_2}{R_N + R_2}$$
$$= \frac{R_1 R_2}{R_1 - R_2}$$
$$G_T = \frac{di}{dv} = G_N + G_2$$
$$= -G_1 + G_2$$

(b)

Fig. 9.7. Series and parallel combination of positive and negative linear resistances.

9.6 Some General Properties of Negative Resistance

The properties of negative incremental resistance are clarified when contrasted with those of positive resistance. A current produces a voltage drop in a positive resistance, and a voltage rise in a negative resistance. Positive resistance dissipates power, hence negative resistance acts as a power source in an incremental circuit model. Conservation of energy demands that any circuit that exhibits a negative resistance at a pair of terminals must draw more power from a primary source than the negative resistance can supply to a load. From this statement we infer that for any physical circuit a driving-point curve that has a negative resistance region must be nonlinear. Thus the driving-point curve at a pair of terminals may exhibit negative resistance over a limited range of current or voltage, but the apparent resistance must eventually become positive as current or voltage is increased either in the positive or negative direction. However, for the purposes of circuit

analysis, it is no more artificial to postulate a linear negative resistance than to postulate a linear positive resistance.

Resistive circuit theory is readily extended to include negative resistance. For example, the sketches shown in Fig. 9.7(a) and (b) indicate series and parallel connections of positive and negative resistances. In

For $t > 0$

$$v = E_0 \epsilon^{-t/RC}$$

$$i = \frac{E_0}{R} \epsilon^{-t/RC}$$

For $t > 0$

$$v = E_0 \epsilon^{-t/R_N C} = E_0 \epsilon^{t/RC}$$

$$i = \frac{E_0}{R_N} \epsilon^{-t/R_N C} = -\frac{E_0}{R} \epsilon^{t/RC}$$

(a)   *(b)*
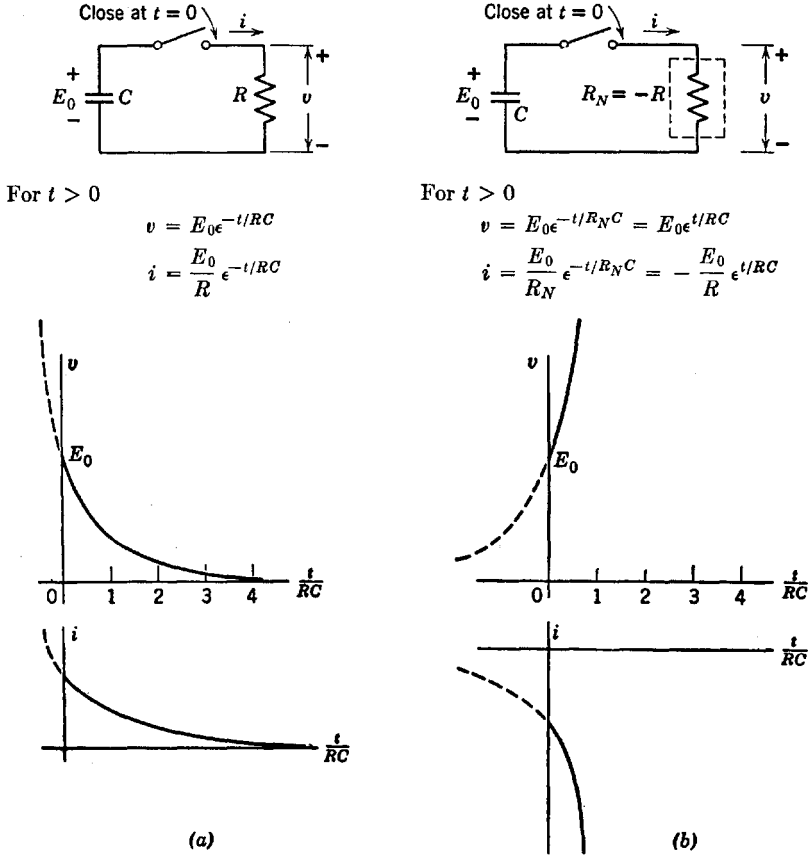
Fig. 9.8. Simple RC transients with positive and negative resistance.

order to draw attention to a point that sometimes causes confusion, the linear negative resistance is designated both as R_N and as $-R_1$. Here the symbol R_N contains the negative sign implicitly, while the symbol R_1 is assumed to be a positive numeric, such as 10 kilohms. The designation with sign *implicit* is convenient because it leaves formulae unchanged. The designation with sign *explicit* is sometimes preferred

because it emphasizes the presence of the negative resistance. From the equations we see that the total series resistance is negative if the magnitude of the negative resistance exceeds that of the positive resistance. The total parallel resistance is negative if the magnitude of the negative resistance is smaller than that of the positive resistance (negative conductance exceeds positive conductance).

The transient behavior of the basic RC circuit with negative R is indicated in Fig. 9.8. It is apparent from the voltage waveforms in (a) and (b) that changing the sign of the resistance is mathematically equivalent to changing the sign associated with time t. The waveforms of current show an additional difference. The current for positive resistance is in the positive reference direction and is a decaying exponential for t greater than zero. Thus it is a replica of the voltage waveform except for the scale factor. For negative resistance, the current is negative, since the resistance $R_N = -R$ appears in the scale factor as well as in the exponent. The direction of the current flow bears out the fact that a negative resistance behaves like a power source, since current emerges from the positive voltage terminal just as it does from a battery.

(a)

(b)

$j = $ I, II, III
$v = E_j + iR_j$

Fig. 9.9.   Representation of a piecewise-linear driving-point resistance.

The corresponding transient behavior for a circuit consisting of inductance and negative resistance is readily specified by duality.

The general behavior of a circuit consisting of a negative resistance and one or more energy-storage elements can be explained without reference to the details of the electronic devices and associated circuit elements from which the driving-point curve is derived. The "black-box" representation of a negative resistance circuit is shown in Fig. 9.9(a). For graphical calculations, the piecewise-linear curve shown in

Fig. 9.9(b) is convenient. The analytical counterpart of this curve is the circuit model shown in (c), where the subscript j represents the various circuit states. This is the linear model for any one region of either a current-controlled or a voltage-controlled curve.
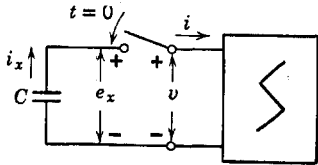
For $t < 0$, $\ i = f(v)$; $\ i_x = -C\,de_x/dt$

For $t \geq 0$, $i_x = i$ and $e_x = v$

(a)

(b)

$i = i_C\epsilon^{-t/R_{II}C}$

$i = -i_A\epsilon^{-t/R_IC}$

(d)

(c)

(e)

$$\delta_1{}' = R_I C \ln(i_A{}'/i_B) = R_I C \ln\left[(E_I)/(E_I - v_B)\right]$$
$$\delta_2 = R_{II} C \ln(i_C/i_D) = R_{II} C \ln\left[v_C/v_D\right]$$
$$\delta_1 = R_I C \ln(i_A/i_B) = R_I C \ln\left[(E_I - v_A)/(E_I - v_B)\right]$$

(f)
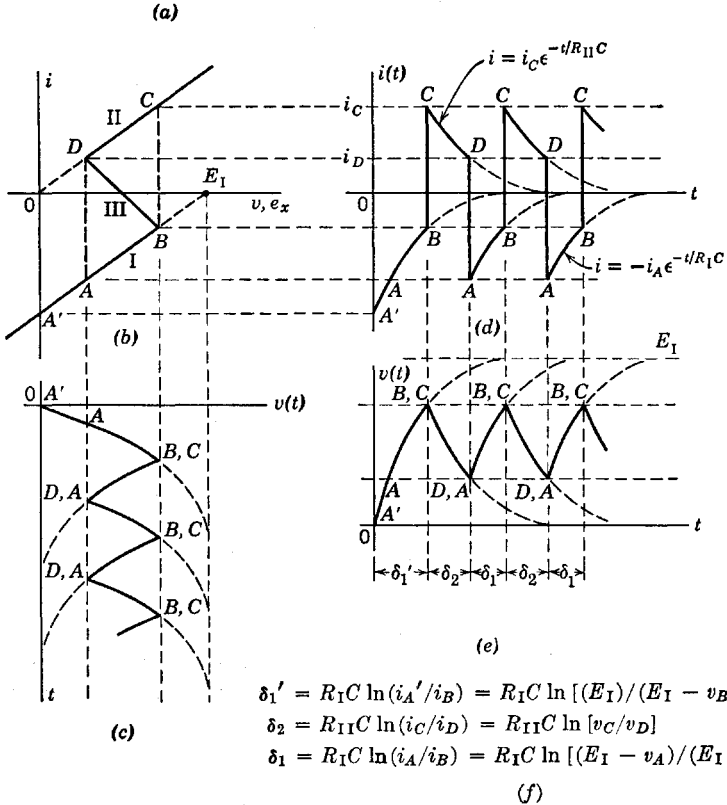
Fig. 9.10. Current-controlled curve with capacitance load.

9.7   A Simple Relaxation Oscillator

Suppose an initially uncharged capacitance is connected to a current-controlled negative resistance at $t = 0$ as indicated in Fig. 9.10(a).

Since the capacitor voltage cannot be changed instantaneously by finite currents, at $t = 0$ we have capacitor voltage e_x and therefore voltage v both equal to zero. This condition can be satisfied only at point A' on the i vs. v plane shown in (b).

At point A', the current i has a large negative value; hence dv/dt is positive, and the capacitor tends to charge. Thus the operating point moves from A' through A toward B. We can see from the plot that, as the operating point moves along the state I line, the negative current diminishes as the voltage increases. This means that dv/dt decreases as v increases. The state I line corresponds to a resistance R_I in series with a battery E_I. With a capacitor C connected, the voltage v will rise exponentially toward E_I as current i proceeds exponentially toward zero with a time constant of $R_I C$. The waveforms of voltage $v(t)$ and current $i(t)$ are shown with the v and i axes aligned with the i vs. v plane [(c) and (d)]. For completeness, $v(t)$ is also shown below $i(t)$ with the time axes aligned. See Fig. 9.10(e).

When the operating point reaches B in the i vs. v plane, the current i is still negative and dv/dt still positive, but any motion beyond point B (in state I) fails to satisfy the relation $i = f(v)$. The relation between capacitor voltage and current cannot be satisfied by moving the operating point along the state III line from point B, since the necessary decrease in v cannot take place with negative current. We conclude, therefore, that the operating point of the resistive circuit must switch instantaneously from B to C, where equilibrium can again be established for both the capacitor and the negative-resistance circuit. The transition from B to C takes place along a vertical line, since the capacitor voltage cannot change instantaneously. In this idealized circuit there is no constraint to prevent the instantaneous change of current from a negative value at B to a positive value at C.

At point C, dv/dt is negative, because the current i is positive. Since the state II line passes through the origin, the circuit model for this state is merely the positive resistance R_{II}. Both v and i approach zero exponentially with a time constant $R_{II} C$. The operating point moves from C to D, where a downward transition to A must occur. The reasoning is similar to that described for the B to C transition. At point A, the capacitor voltage begins to increase again and the operation proceeds continuously around the locus $ABCD$. Thus, except for the first interval, the waveforms $i(t)$ and $v(t)$ are repetitive.

The intervals δ_1 and δ_2, during which the resistive circuit is in state I and state II, respectively, can be calculated from the time constants $R_I C$ and $R_{II} C$ and the initial and final values of the exponentials. The expressions for the intervals are given in (f).

Let us take as a specific example of a relaxation oscillator the series

triode circuit of Fig. 9.5 (a) with capacitor connected to the driving-point terminals, as shown in Fig. 9.11. It is evident that the amplitude dimensions on the waveforms of $v(t)$ and $i(t)$ are obtainable directly
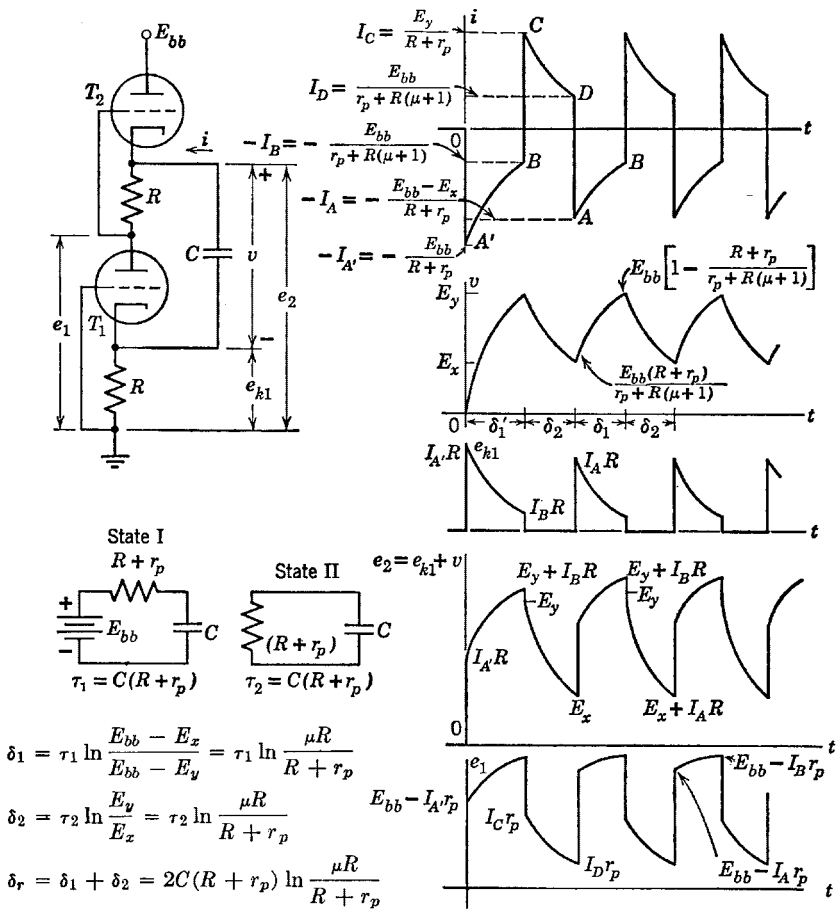
Fig. 9.11. Waveforms in series-triode relaxation oscillator.

from the i vs. v curve given in Fig. 9.5 (e). Dimensions of all other waveforms in the circuit are readily found from these. In a similar manner, the general results obtained in Fig. 9.10 can be applied to other circuits in which a capacitor faces a current-controlled driving-point curve. Using the duality principle, the results can also be applied to circuits with a voltage-controlled curve and inductance as the major energy-storage element.

9.8 Relaxation Oscillator Transition Time

As the operating point moves around the locus $ABCD$ in Fig. 9.10(b), the capacitor voltage is identical with the voltage across the negative resistance circuit during the intervals δ_1 $(A$ to $B)$ and δ_2 $(C$ to $D)$. However, during the instantaneous transitions $(B$ to C and D to $A)$, the capacitor voltage deviates from the voltage across the negative resistance, even though the voltages are measured between the same two points. This irregularity was ignored in the preceding article on the basis that it literally did not exist with instantaneous transitions. The anomaly of two different voltages apparently existing between the same two points is the result of over-idealizing the circuit.

In order to examine the transitions between state I and state II in detail, let us consider the circuit of Fig. 9.12(a). Here a small series inductance representing the inductance of the interconnecting wires is included to absorb the difference between v and e_x when di/dt is large. With inductance L very small, the transitions will still be quite rapid, though not instantaneous. The charge on capacitor C will remain essentially constant during the transition time, hence the capacitor can be replaced by a battery for this interval. We see from the i vs. v plot that the transition occurs in two distinct stages corresponding to states III and II. The two circuits for the transition from B to C are shown in Fig. 9.12(b) and (c). Note that the plots of waveforms $i(t)$ and $e_L(t)$ in this figure are drawn to a time scale which is radically expanded relative to the time scale used for the waveforms of Fig. 9.10 or 9.11. For convenience, assume $t = 0$ corresponds to the time the operating point $(i$ vs. $e_x)$ reaches B on the locus.

We postulate L to be so small that it does not influence the behavior of the circuit during the intervals δ_1 and δ_2. This implies that e_L is negligibly small except at the transitions. The current i has a value $-I_B$ and a rate of change

$$\frac{di}{dt} = \frac{I_B}{R_I C} \tag{9.3}$$

at the break point B, so the actual value of e_L at B is

$$e_L = -\frac{LI_B}{R_I C} \tag{9.4}$$

From these initial conditions, both e_L and i build up along rising exponentials, since the circuit model for $v = f(i)$ involves the negative resistance $(-R_{III})$ in state III, as indicated by (b). The current goes

through zero and positive to I_D, while e_L builds up to a maximum of $(E_y - E_x)$ at break point D. When point P is reached on the locus,

$$v = e_x + e_L$$

(a)

$$e_L = -L\frac{di}{dt} = v - E_y$$
$$= E_{III} - iR_{III} - E_y$$
$$0 < t < t_p$$

(b)

$$e_L = -L\frac{di}{dt} = v - E_y$$
$$= iR_{II} - E_y$$
$$(t > t_p)$$

(c)

For $t > t_p$
$$i = I_C - (I_C - I_D)\epsilon^{-t/\left(\frac{L}{R_{II}}\right)}$$

$$+I_D \quad P \quad \frac{di}{dt} = \frac{E_y - E_x}{L}$$

For $0 < t < t_p$
$$i = -I_B + \frac{I_B L}{R_{III} R_I C}\left[1 - \epsilon^{-t/\left(\frac{L}{R_{III}}\right)}\right]$$

Slope $= \frac{I_B}{R_I C}$

(e)

$(e_L)_0 = -\frac{LI_B}{R_I C}$

$$e_L = -(E_y - E_x)\epsilon^{-t/\left(\frac{L}{R_{II}}\right)}$$
For $t > t_p$

$-(E_y - E_x)$

$$e_L = -\frac{LI_B}{R_I C}\epsilon^{t/\left(\frac{L}{R_{III}}\right)}$$
For $0 < t < t_p$

(f)
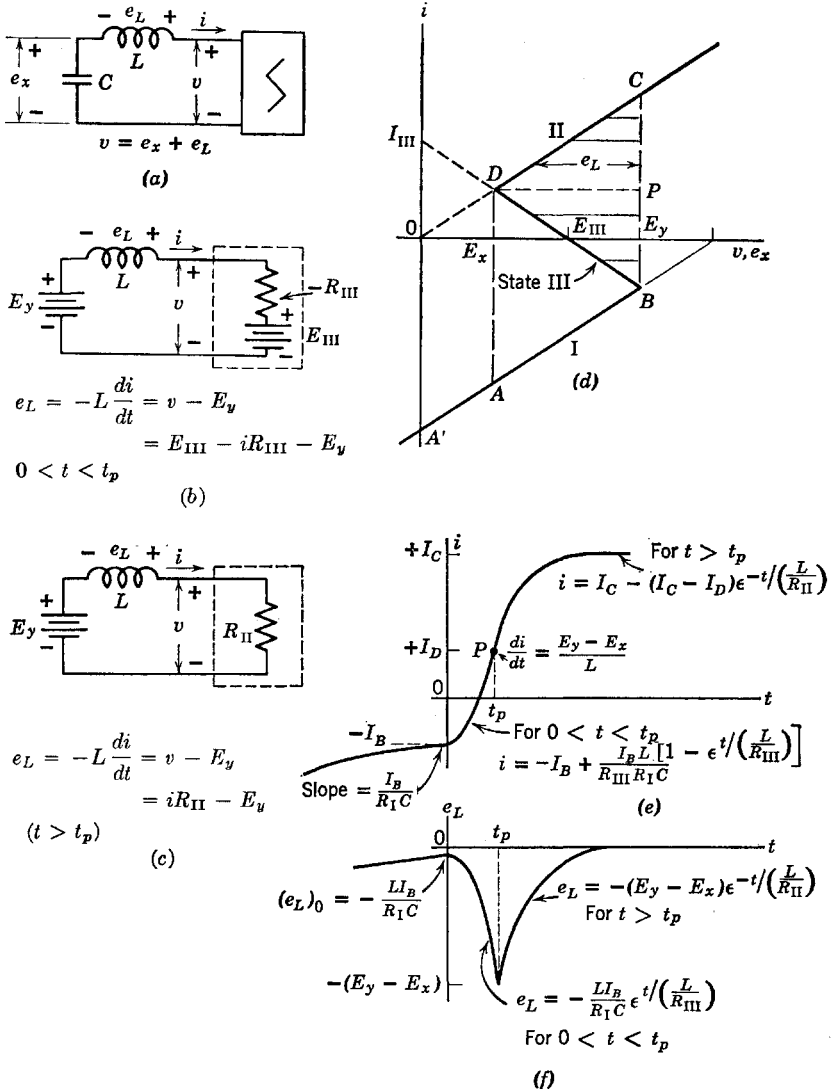
Fig. 9.12. Relaxation oscillator transition time.   ($t = 0$ corresponds to point B.)

the circuit switches from state III to state II. As the transition proceeds from P to C, voltage e_L decays exponentially to zero with time constant

L/R_{II}, while i builds up to I_c with the same time constant. The voltage e_L is continuous at point P; hence di/dt must be continuous. The rate of change of voltage (de_L/dt) has no such restraint, hence the cusp at point P. The time t_p required to reach point P can be calculated as the time required for current i to go from $-I_B$ to $+I_D$, or for e_L to build up from $(LI_B/R_I C)$ to $(E_y - E_x)$. The total transition time can be taken to be approximately $t_p + 5\ L/R_{II}$.

As L decreases, the time scale shrinks on the transition waveforms. As L approaches zero, di/dt increases toward infinity in a manner that makes e_L always increase to $(E_y - E_x)$ and return to zero. Hence, even when L does not appear in the circuit, we can envision the "instantaneous" transition involving a voltage drop e_L which just accounts for the difference between v and e_x.

The transition from D to A is readily obtained by the procedure just outlined for the transition from B to C. In the circuit model of Fig. 9.12(b), we need only change the voltage representing capacitor voltage from E_y to E_x. This change must also be made in the circuit of Fig. 9.12(c); and, in addition, R_{II} must be replaced by R_I.

Separate consideration of the effects of the major energy-storage element C and the stray element L is valid when the time constant involving C is much greater than that involving L. The procedure is similar to that used in analyzing the RC coupling circuit in the preceding chapter, where the effects of coupling capacitance and shunt capacitances were considered separately. In a qualitative way, the stray inductance inserted in the oscillator circuit to make transition time nonzero approximates the effects of interelectrode and wiring capacitances within the circuit that produces the negative driving-point resistance. The analysis of stray effects in terms of a single series inductance is very much simpler than one that considers the energy-storage elements actually associated with the circuit that produces the negative resistance.

9.9 Relaxation Oscillator Locus and Waveforms with Nonzero Transition Time

From the results of Fig. 9.12 we can now see how the locus and waveforms of Fig. 9.10 must be modified when both a large capacitance C and a small inductance L are included in the oscillator circuit. Since the inductance is expected to produce only a minor modification, we begin from the locus and waveforms obtained with the inductance absent. This locus is indicated by the dashed lines in Fig. 9.13, while the solid line shows the qualitative effects of including a small inductance L.

The relations given in Fig. 9.13(a) suffice to trace the modified locus of i vs. e_x shown in (b) and the waveforms shown in (c), (d), and (e).

$$i = f(v)$$
$$i = -C\,de_x/dt$$
$$e_L = -L\,di/dt$$
$$= v - e_x$$
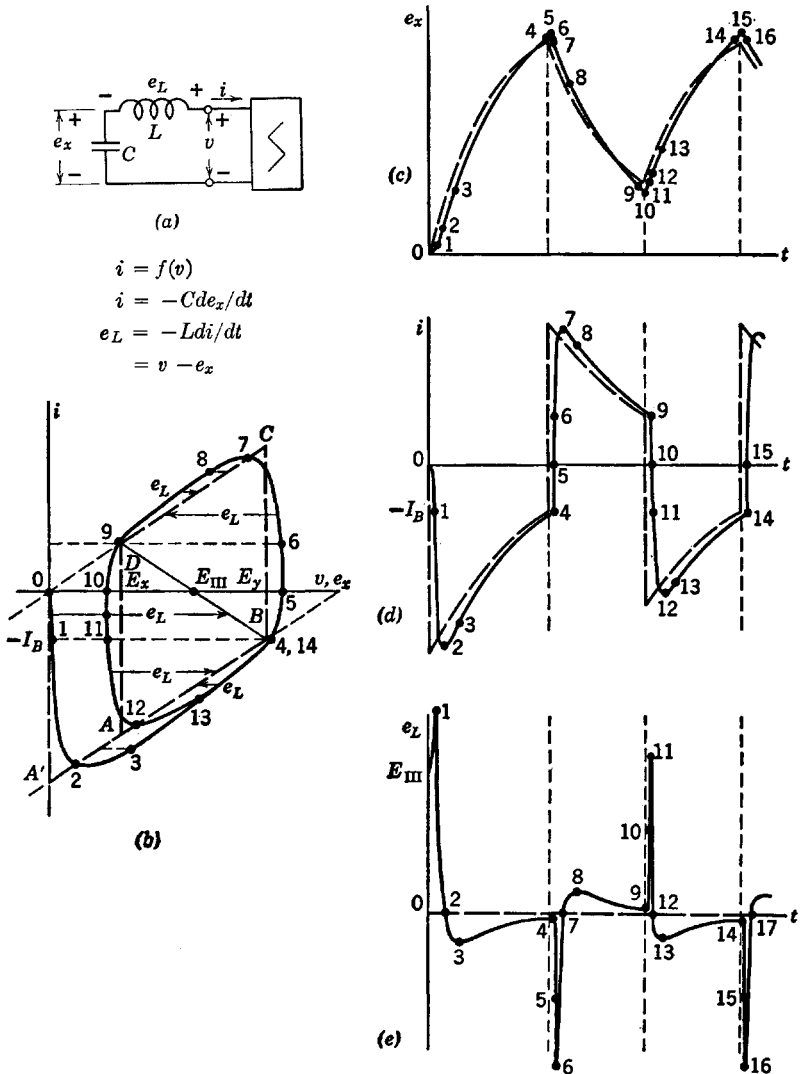
Fig. 9.13. Relaxation oscillator with nonzero transition time.

Beginning at $t = 0$ (point 0) with $i = 0$ and $e_x = 0$, we have $e_L = E_{III}$. Since current $i = 0$, the slope of the capacitance voltage (de_x/dt) is zero at $t = 0$. The slope of the current waveform is $di/dt = -E_{III}/L$,

hence current starts to build up in the negative direction. As it does, v increases faster than e_x, causing e_L to increase until it reaches a maximum nearly equal to E_y at point 1. Here the current waveform will have its maximum negative slope ($di/dt \approx -E_y/L$). During this rapid transition, the capacitor voltage changes only very slightly.

Beyond point 1, e_L diminishes; but since it is still positive, di/dt remains negative and decreases to zero with e_L at point 2. The current i has its maximum negative value here so that de_x/dt is a maximum, and the waveform $e_x(t)$ has its point of maximum slope. With de_x/dt nonzero and di/dt zero at this point, di/de_x is zero. In other words, the locus of i vs. e_x crosses the curve $v = f(i)$ with zero slope. Beyond point 2 the magnitude of the modified current exceeds that of the exponential obtained with only C in the circuit. This fact predicts a crossover of the exponential and the modified waveforms of e_x, for the integral of the modified current will eventually exceed the integral of the exponential current. The crossover will occur where the areas under the two current waveforms become equal.

We can logically say that point 2 marks the end of the rapid transition and after passing this point we have e_L negative and di/dt positive. As the magnitude of the current diminishes, its rate of change must also diminish, since the exponential change of the capacitance dominates the circuit behavior during this portion of the charging interval. Thus we have a maximum of di/dt at point 3, a corresponding negative maximum for e_L and a maximum of d^2e_x/dt^2.

Approaching break-point B, we have di/dt nearly the same for the exponential and the modified current waveforms. This gives us a good approximation for the value of e_L at point 4 which is located at $-I_B$ and marks the beginning of another rapid transition. Beyond point 4, the inductance voltage e_L goes rapidly to a large negative value, causing di/dt to increase rapidly. Prior to reaching point 5, the current has been negative; hence de_x/dt has been positive. At point 5, $i = 0$, $de_x/dt = 0$, hence e_x has reached a maximum. The locus of i vs. e_x must be moving vertically ($di/de_x = \infty$) at point 5, since de_x/dt is zero, whereas di/dt is nonzero.

Point 6 is at the negative peak of e_L corresponding to a maximum of di/dt, and at point 7, e_L and di/dt are zero. At point 8, e_L reaches a maximum (analogous to point 3) and current i therefore has a maximum slope. Point 9, analogous to point 4, marks the beginning of a transition which takes the path of operation through points 10, 11, and 12 in rapid succession. At point 10 (analogous to 5) we have zero current and hence minimum e_x; at point 11, a maximum e_L which produces maximum di/dt. At point 12, di/dt, e_L, and di/de_x are all zero. Sometime after

point 12, the path of operation very nearly merges with the path followed previously and the waveforms and locus become essentially repetitive. The shape of the waveform of e_L is seen to be a series of pulses which approximate impulses. If inductance L is allowed to approach zero, these pulses indeed approach impulses with area just equal to that required to jump the current i from point B to point C or from D to A. The voltage e_L would then be zero between impulses, since L is zero and di/dt is finite.

We have described a general relaxation oscillator based on a current-controlled negative resistance which, together with a capacitance, determines the intervals δ_1 and δ_2. A small series inductance approximates the effects of stray energy-storage elements which limit the rate of transition from one state to another. Comparable results are obtained for the dual of this circuit which is a voltage-controlled negative resistance and an inductance. A small parallel capacitance accounts for the finite transition time.

9.10 Shifting from Astable to Monostable Operation

Minor modification of the circuits used for generating relaxation oscillations (astable circuits) makes them operate in a monostable or bistable mode. A *monostable* circuit is stable in one state only. An external signal must be applied to switch to the unstable state. The circuit then switches itself back to the stable state and remains there until another external signal causes the cycle of operation to repeat. A *bistable* circuit has two stable states, and remains in either state unless an externally applied signal causes a transition.

Relaxation oscillations occur when a circuit has no stable states. Referring to Fig. 9.10 or Fig. 9.13, we see that the charging of capacitor C tends to drive the circuit out of state I, through III, and into II. The capacitor then discharges, driving the circuit out of state II, through III, and into I, and the cycle repeats. To obtain monostable operation, we need only modify the circuit in such a way that the exponential charge or discharge of capacitor C can reach completion ($i = 0$) before a transition occurs.

As indicated in Fig. 9.14(a), the relaxation oscillator circuit is in equilibrium at point P, since we have $e_x = v$, and $i = -C\,dv/dt = 0$. This, however, is a point of unstable equilibrium, since the slightest deviation from P will result in a motion of the operating point away from P. A minute positive current tends to decrease the charge on capacitor C, thus diminishing v, and increasing the current. To produce

a stable equilibrium, the point P for which capacitor current is zero must be shifted to state I or state II. Such a shift requires a direct-current bias source as indicated in Fig. 9.14(b). Here the capacitor current is ($i + i_0$), which will be zero when $i = -i_0$; hence the point P

$$v = f(i)$$
$$i = -C dv/dt$$
(a)

$$v = f(i)$$
$$(i + i_0) = -C \, dv/dt$$
(b)

$$v = f(i_x)$$
$$i_x = i + i_0$$
$$= -C \, dv/dt$$
(c)
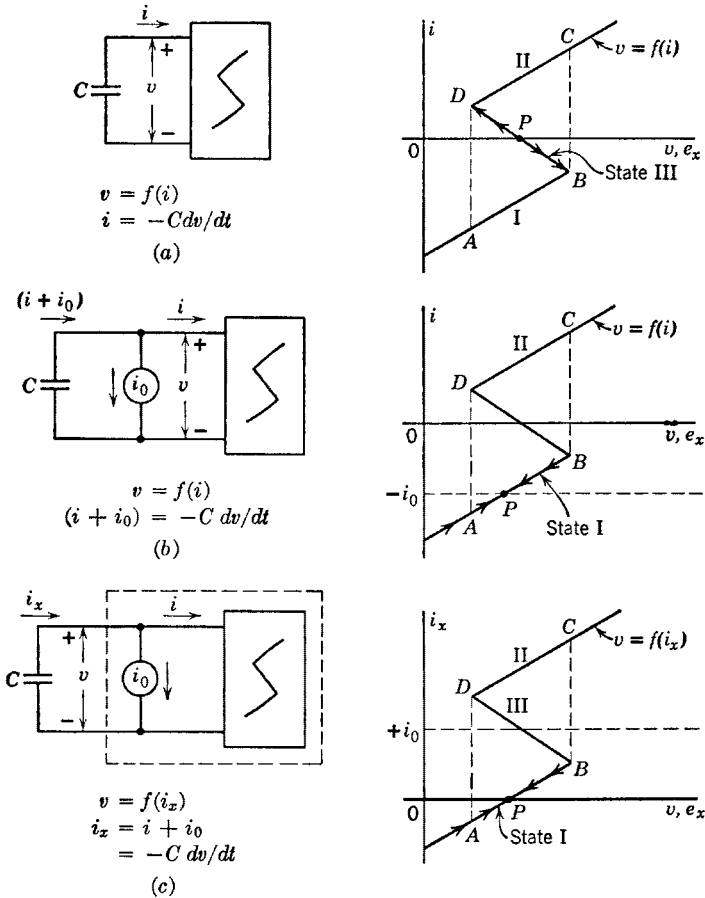
Fig. 9.14. Use of current bias to shift from astable to monostable circuit conditions.

has been shifted to state I. Now any change in current away from the point P results in a polarity for dv/dt that tends to return the operating point to P. Thus the point P is stable.

In the circuit and graphical plot of Fig. 9.14(b), we have assumed the fixed current i_0 and the capacitance C as the load on the resistive circuit specified by $v = f(i)$. The alternative viewpoint, illustrated by

Fig. 9.14(c), considers i_0 as part of the resistive circuit. The result is a modification of the curve of v vs. i. The viewpoint expressed by Fig. 9.14(c) is useful, since all of the resistive circuit elements are lumped together.

Referring to Fig. 9.14(b) or (c), we conclude that the operating point must reach point P and remain there, regardless of its initial location in the i vs. v plane. For example, if by some means the circuit is put into state II, the capacitor voltage will decrease as the operating point moves toward point D. There, a transition to point A in state I takes place. Negative capacitor current then causes the voltage to increase until point P is reached and the circuit is in stable equilibrium.

9.11 Triggering the Monostable Circuit

A monostable circuit is usually "triggered" by a pulse which initiates a transition from the stable state to the unstable state. The time during which the circuit remains in the unstable state is determined by the exponential charge or discharge of the capacitor. During this interval the operation is like that of a free-running relaxation oscillator. At the end of the interval the circuit switches back to the stable state, and remains there until the next triggering pulse initiates the cycle again, as illustrated by Fig. 9.15.

The current-controlled resistance in the circuit of Fig. 9.15(a) is assumed to have a stable equilibrium, as indicated by point P on the curve $v = f(i)$ shown in (b). The waveform of current from the synchronizing source i_s is the train of pulses in Fig. 9.15. The circuit model shown in (c) indicates the circuit conditions just prior to the occurrence of one of the synchronizing pulses; thus all currents are zero, and the circuit is in stable equilibrium.

In order to switch the circuit from state I to state II, the synchronizing pulse must drive the operating point from P to slightly beyond break point B. The operating point will return to point P when the trigger pulse ends, unless the capacitor voltage is increased by the amount

$$\Delta v \geqslant (E_b - E_p) \qquad (9.5)$$

which requires a charge content

$$Q = I_s \delta_s \geqslant C(E_b - E_p) \quad \text{for } R_1 C \gg \delta_s \qquad (9.6)$$

At break point B the circuit switches to state II and follows the operating path $CDAP$. Since the current pulses cannot change the capacitor voltage instantaneously, current i shows no jump at the trailing edge of

the trigger pulse. The negative jump at the end of the trigger pulse appears in i_x, since

$$i_x = i - i_s \tag{9.7}$$
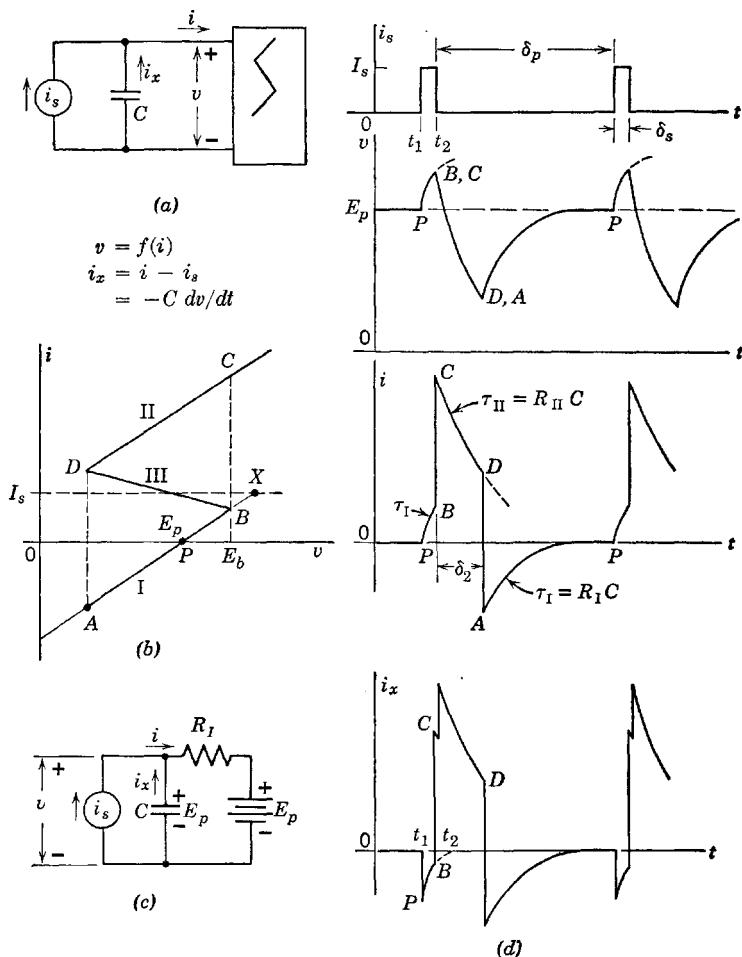
Fig. 9.15. Monostable operation with current pulse triggering.

The interval during which the circuit is in state II can be calculated in the same way as the corresponding interval in the waveforms of the relaxation oscillator. In order to assure return to the stable point P after interval δ_2, the repetition interval of the synchronizing signal should be

$$\delta_r \geqslant (\delta_2 + 5 R_I C) \tag{9.8}$$

Allowing about five time constants for recovery permits the capacitor to recharge to the voltage E_p.

The time required to accomplish triggering will vary inversely with the trigger amplitude I_s, because the larger the trigger amplitude, the

$$v = f(i)$$
$$e_L = v - (e_x + e_s)$$
$$= -L\,di/dt$$

(a)

(b)

(c)
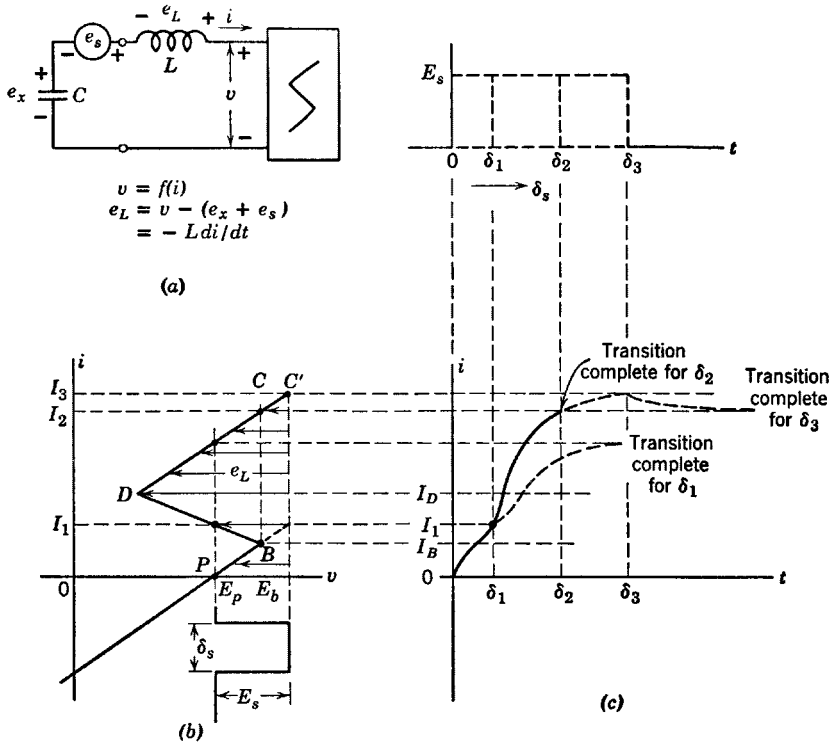
Fig. 9.16. Monostable operation with voltage pulse triggering.

larger the value of dv/dt. It is not proper to assume that instantaneous triggering could be had by letting the pulses approach impulses of the same area, since the effects of the stray parameters would then have to be included in the form of a small series inductance. We have tacitly assumed the pulse duration δ_s to be large compared with the time required to build up current from 0 to I_s in the stray inductance. If inductance is included, the instantaneous changes of current shown in the waveforms of current in Fig. 9.15(d) must be rounded off to approximate exponentials as in Fig. 9.13.

A circuit for obtaining faster triggering is shown in Fig. 9.16(a). Instead of trying to change the capacitor voltage from E_p to E_b, we can bias the resistive curve in the opposite direction. This has the

effect of sliding the entire curve of $v = f(i)$ to the left until break point B passes the capacitor voltage E_p. Then a transition must occur from B to some point C' in state II. This shift of the curve along the voltage axis is readily accomplished by a series voltage source e_s subtracted from v, as indicated in Fig. 9.16(a). Alternatively, we can think of the voltage e_s as being added to capacitor voltage e_x instead of being subtracted from v. The net result will be the same.

If we idealize the circuit by neglecting the series inductance, the triggering transition occurs instantaneously. The only requirement is a series-connected voltage source producing pulses of amplitude $E_s >$ $(E_b - E_p)$. With a nonzero inductance present, we shall see from Fig. 9.16 that a minimum pulse area $E_s\delta_s$ is required. This is the change in flux linkages required to change the current in the inductor sufficiently to trigger the circuit.

Let us assume that the triggering transition takes place so rapidly that the charge on the capacitor remains essentially constant. Then we can use the method of calculation given in Fig. 9.12. For the circuit of Fig. 9.16(a), the values of e_L during the transition are shown on the i vs. v plane in (b). It is apparent that the voltage across the inductance ($e_L = -L\,di/dt$) varies directly with E_s; hence, the larger E_s, the more rapid the transition. Transition waveforms for a specific value of E_s are shown in Fig. 9.16(c) for three values of synchronizing-pulse duration.

The duration of the pulse must exceed δ_1, the value required to change the current from zero to I_1. Beyond that point the transition goes to completion without dependence on the trigger pulse, since the negative value of e_L leaves di/dt positive. If the current has not reached I_1, removal of the synchronizing voltage pulse makes e_L positive, a condition that makes di/dt negative. Therefore, the operating point will be driven downward out of state III to point P in state I.

Maintaining the trigger pulse for the duration δ_2 required for the current to reach the value I_2 at point C results in minimum transition time for a given amplitude of e_s. If the trigger pulse persists for a duration δ_3 the current reaches too large a value (I_3 at point C') and the transition time becomes considerably longer. However, these transition times are all short in comparison with those resulting from current-pulse triggering, particularly if capacitor C is large.

9.12 Bistable Operation

Since a current-controlled curve, $v = f(i)$, is by definition a single-valued function of current, it crosses the axis of zero current only once. Thus there can be only one stable point ($i = 0$) with capacitance as

the energy-storage element. However, if an inductance is used the transient is completed when the voltage across the inductor is zero. If we use a bias voltage to shift the curve along the voltage axis in order to obtain three intercepts on the axis of zero voltage, bistable operation is obtained.

The location of the bias voltage E_0 and the synchronizing voltage e_s is shown in Fig. 9.17(a). The curve $v = f(i)$ is assumed to be like the one used in preceding examples. The sketch in Fig. 9.17(b) indicates E_0 as a load on the negative-resistance device. Under steady-state conditions, with e_s and e_L both zero, we have circuit equilibrium at points P, S_1, and S_2. Thus, for $v = E_0$ we have three points for which e_L is zero, and two of these are stable.

The curve $v' = f(i)$ shown in Fig. 9.17(c) is the curve of (b) with a change of variable ($v' = v - E_0$), which effectively includes the bias voltage E_0 as part of the resistive circuit. The alternative viewpoints, indicated in (b) and (c), are related to specific circuits in (d) and (e). The circuit in (d) shows the necessary voltage translation achieved by an external battery of value $E_0 = E_{bb}/2$. The circuit in (e) makes the necessary internal modification to shift the axis of symmetry from $v = E_{bb}/2$ to $v = 0$. In either case we can indicate the circuit properties as in Fig. 9.17(f).

To switch the circuit from state I to state II and back again requires alternately positive and negative pulses from the synchronizing source. The method of triggering is like that described in Fig. 9.16 for monostable operation. The waveforms drawn in Fig. 9.17(f) assume that the stray inductance is very small. These waveforms can be modified to make the transitions more realistic by using the results obtained in Figs. 9.12 and 9.16.

9.13   Limitations of Negative-Resistance Methods

A nonlinear driving-point resistance combined with one or more energy-storage elements is a very useful model for examining the behavior of many oscillator circuits. This general method of analysis shows directly the conditions for bistable, monostable or astable relaxation oscillations as well as sinusoidal oscillations. Equally important is the fact that a large number of actual circuits can be represented quite accurately by this simple but general model. In principle the concepts described thus far can be applied to any oscillator circuit, but as a practical matter it is inconvenient to try to fit all circuits into the same mold. The negative-resistance concept is not particularly convenient

$v' = v - E_0$
for $e_L = 0$
(a)

(b)   (c)

(d)   (e)

(f)
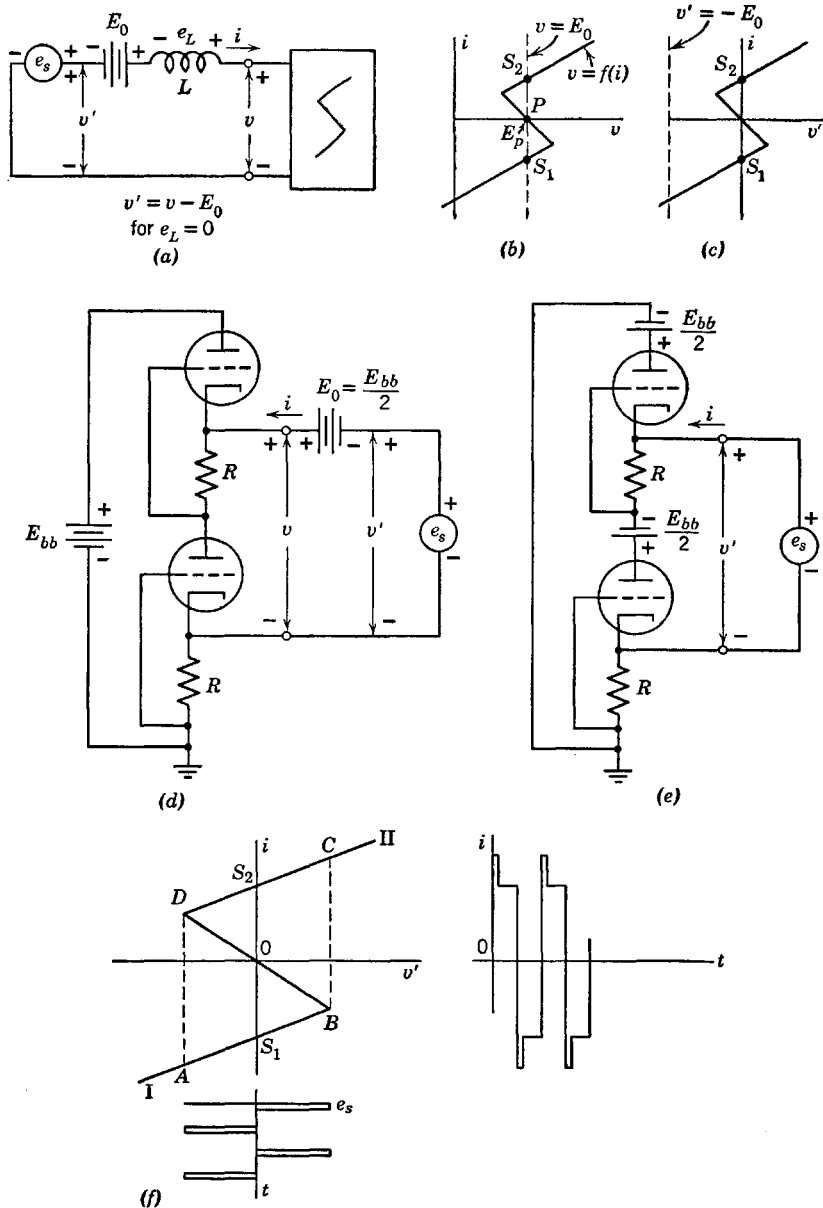
Fig. 9.17. Bistable operation.

$$f_0 \approx \frac{1}{2\pi\sqrt{LC}}$$

(a) (b) (c) (d) (e)
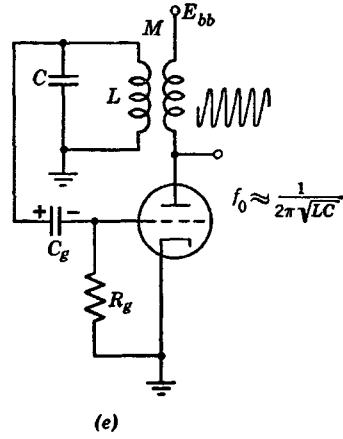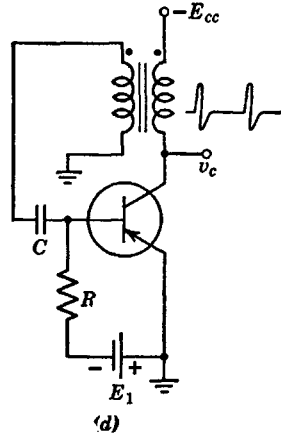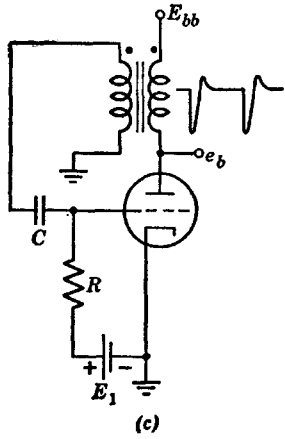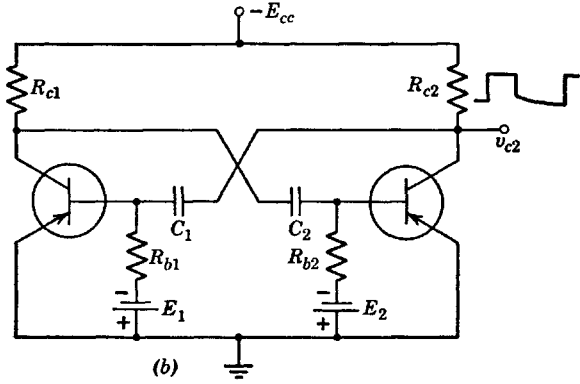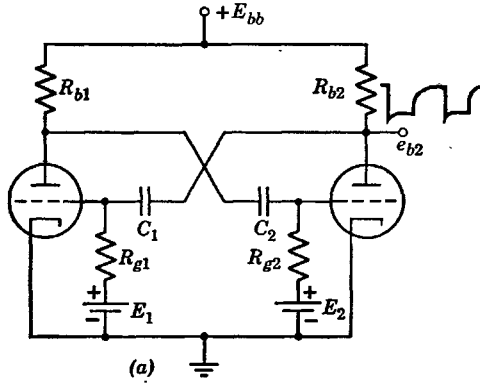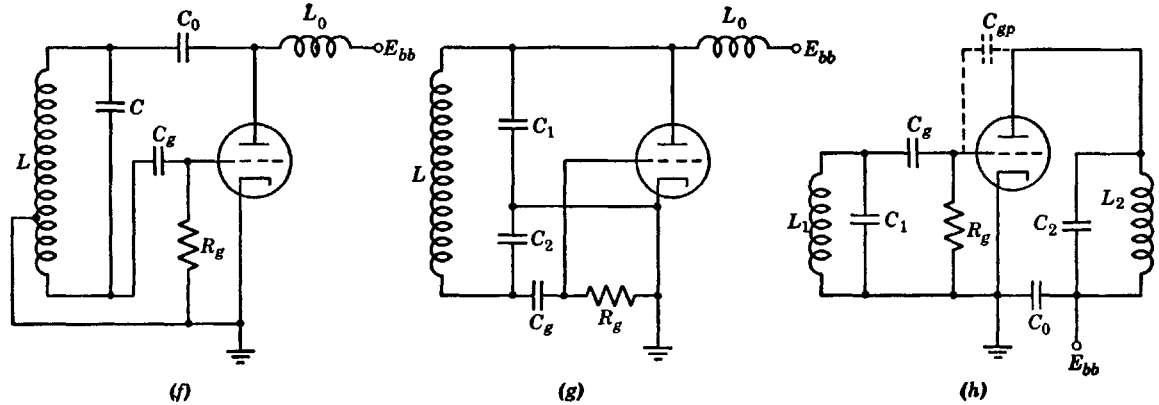
Fig. 9.18. Typical oscillator circuits with major energy-storage elements at two pairs of terminals. (a) Plate-coupled multivibrator; (b) Collector-coupled multivibrator; (c) Triode blocking oscillator; (d) Transistor blocking oscillator; (e) Transformer-coupled sine wave oscillator; (f) Hartley oscillator; (g) Colpitts oscillator; (h) Tuned-plate, tuned-grid oscillator.

for analyzing oscillator circuits in which major energy-storage elements appear at more than one pair of terminals. For example, consider the plate-coupled and collector-coupled multivibrators, shown in Fig. 9.18(a) and (b) respectively. The relaxation oscillations generated by these circuits depend upon two major energy-storage elements, C_1 and C_2. At the terminals of either capacitor, the apparent driving-point relation is a family of negative-resistance curves involving the voltage across the other capacitor as a parameter. Thus, instead of a single resistive curve we have two families of curves, $v_1 = f_1(i_1, v_2)$ and $v_2 = f_2(i_2, v_1)$ describing the terminal conditions faced by the two capacitors. Tracing the corresponding trajectories e_{x1} vs. i_1 and e_{x2} vs. i_2 on the two planes is rather tedious.

These circuits can be analyzed by the methods discussed in Chapter 8, since each is nothing more than a two-stage RC-coupled amplifier with the output of the second stage connected to the input of the first. Analysis of the coupling circuits permits calculation of the waveforms, but the conditions for relaxation oscillations must be established separately. To establish relaxation oscillations, each stage must produce an output of sufficient amplitude (when driven from cutoff to quiescent conduction) to drive the other stage from conduction to cutoff.

Another method for determining whether or not a circuit meets the conditions necessary to establish oscillations is based on an incremental power-gain calculation. For example, in the multivibrator of Fig. 9.18(a) assume both triodes are operating in the linear amplification region. Then apply a signal at one grid and calculate the return signal that would occur with the circuit opened at that point. If the return signal is greater than the signal applied, the circuit is unstable and can generate oscillations.

The transformer-coupled circuits shown in Fig. 9.18(c), (d), and (e) are further examples of circuits with major energy-storage at more than one pair of terminals. The waveforms generated depend both upon the grid-coupling capacitor C_g and upon the properties of the transformer. In the blocking oscillator circuits shown in (c) and (d) the transformer is one designed to transmit brief (one- to ten-microsecond) rectangular pulses. The circuit will generate pulses whose duration depends largely on the pulse transformer magnetizing inductance. The spacing between pulses depends primarily on C and R. In the blocking oscillator, a single control valve provides the necessary power gain, whereas the transformer performs the essential function of polarity inversion. In the multivibrator circuits, the second control valve provides the polarity inversion.

The circuit shown in Fig. 9.18(e) is essentially the same as the blocking oscillator circuit of (c) but the coefficient of coupling M is considerably

less because an air-core coil is used instead of the iron-core coils. The resonant circuit consisting of capacitance C and the equivalent inductance L of the coupled coils produces nearly sinusoidal oscillations at a frequency approximating $f_0 = 1/2\pi\sqrt{LC}$. In this mode of operation, C_g and R_g have relatively little effect on frequency but do influence the amplitude of oscillations. Rectification due to the difference between r_g and R_g produces an average voltage across C_g having the polarity shown. This "grid-leak" bias controls the amplitude of oscillations. If R_g is made too large, the average voltage across C_g builds up during each succeeding cycle of oscillation until eventually plate current is cut off and the oscillations cannot be sustained. Then the capacitor must be discharged through R_g until plate-current conduction begins and oscillations again build up. This "self-pulsed" mode of operation is analogous to blocking oscillator operation except that here each pulse consists of a number of cycles instead of approximately a half cycle.

Three other common sine wave oscillators are shown in Fig. 9.18(f), (g), and (h). In the Hartley circuit the input (grid-to-cathode) voltage is a portion of the output (plate-to-cathode) voltage obtained by tapping the coil. In the Colpitts circuit the same thing is accomplished by a capacitive voltage divider. In the tuned-grid, tuned-plate circuit no coupling is apparent between output and input. The coupling occurs through the grid-to-plate interelectrode capacitance.

9.14   Plate-Coupled Multivibrator

The waveforms generated by the plate-coupled multivibrator circuit can be determined by the procedure used in Chapter 8 for calculating waveforms due to a rectangular-wave input to a resistance-capacitance-coupled amplifier. Assume that the circuit of Fig. 9.19(a) is a relaxation oscillator that remains in state I for an interval δ_1 and in state II for an interval δ_2. The transition from I to II or from II to I via state III occurs very rapidly (instantaneously if stray capacitances are neglected). The rapid transitions are reasonable in view of the high gain available in the circuit and the fact that the circuit is regenerative (output reinforces input). If the circuit is in state III, the "open-loop" voltage gain is $G_1 G_2$; namely, the gain of the two triode amplifier stages. Let the grid of triode 1 be disconnected from R_{g1} and let e_{c1} be changed by a small positive increment Δe_{c1}. Then

$$\Delta e_{b1} = -G_1\,\Delta e_{c1} \qquad (9.9)$$

$(b) \quad \tau_{c2} = C_2[R_{b1} + r_{g2}\|R_{g2}]$

$\tau_{d1} = C_1[R_{g1} + r_{p2}\|R_{b2}]$

$(c) \quad \delta_1 = \tau_{d1} \ln\left[\dfrac{R_{g1}}{R_{g1} + r_{p2}\|R_{b2}}\right]\left[\dfrac{E_1 + E_{bb} - E_{a2}}{E_1 + E_{col}}\right]$

where $E_{col} = E_{bb}/\mu$
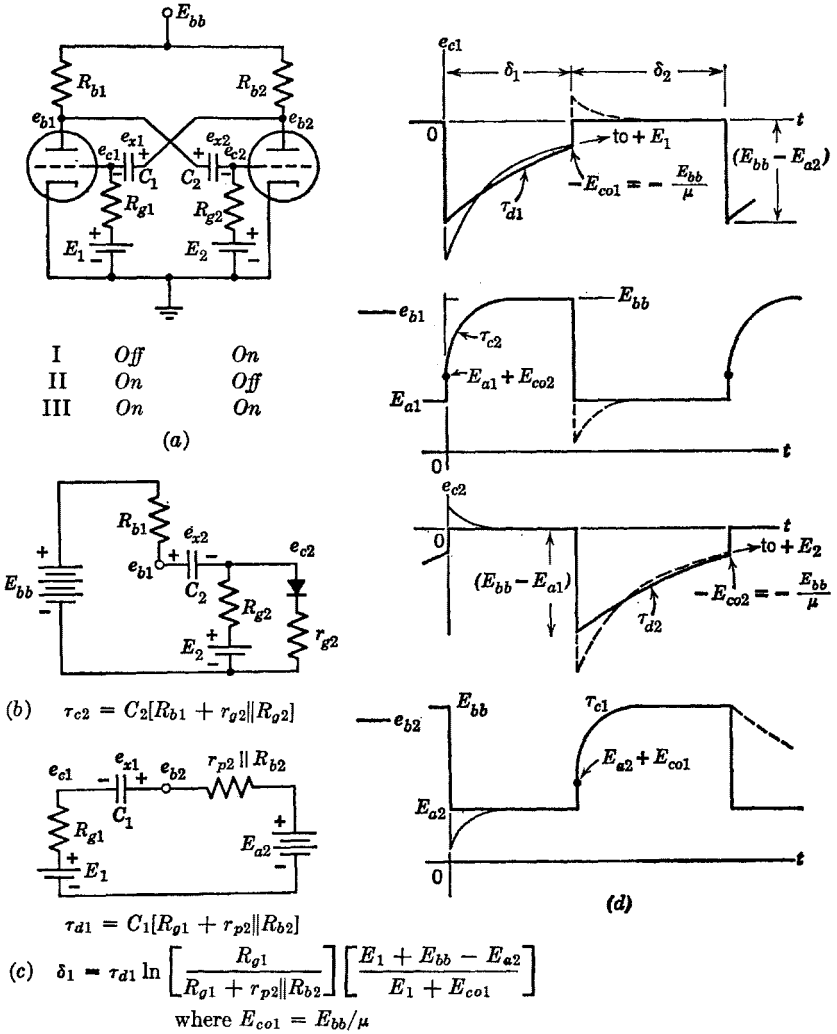
Fig. 9.19. Waveforms for free-running (astable) plate-coupled multivibrator.

This change is coupled to the grid of the second triode by C_2 so that

$$\Delta e_{b2} = -G_2 \Delta e_{c2} = G_1 G_2 \Delta e_{c1} \qquad (9.10)$$

This change in turn is coupled through C_1 and appears across R_{g1}. For typical circuit parameters, $G_1 G_2$ might be 100. If the grid of the first triode is now reconnected to R_{g1} the circuit is unstable in state III since any change in e_{c1} immediately calls for a change in e_{c1} 100 times as

great. Thus the circuit will drive itself out of state III into one of the other states.

If we assume the circuit to be in state I, the two circuit models given in Fig. 9.19(b) and (c) apply. In circuit (b) the capacitor C_2 will tend to charge toward a voltage very nearly equal to E_{bb}. This is true since in most cases $R_{g2} \gg r_{g2}$ and hence the quiescent value of e_{c2} is approximately zero* in state I. Since this leaves triode 2 conducting, the assumption that the circuit is in state I is valid. Now, referring to (c), it is apparent that the voltage e_{x1} must approach the value $(E_{a2} - E_1)$ where E_{a2} is the quiescent open-circuit plate voltage for tube 2 corresponding to the quiescent grid voltage e_{c2} given by the circuit in (b). Circuit models similar to those of (b) and (c) can be obtained for state II by the simple expedient of interchanging all subscripts.

Since state I occurs just after state II, we expect the initial value of e_{x1} to be near E_{bb}. Hence in Fig. 9.19(c) capacitor C_1 is discharging and voltage e_{c1} is rising toward E_1. But e_{c1} must be more negative than E_{bb}/μ to maintain plate-current cutoff in tube 1. This means that the current through R_{g1} and hence the discharge current for C_1 must exceed $(E_1 + E_{bb}/\mu)/R_{g1}$ to maintain state I. When the current falls to this value, the grid voltage e_{c1} rises above the cutoff potential. The resulting onset of plate current i_{b1} causes a drop in e_{b1} and e_{c2}, a rise in e_{b2} and hence a further rise in e_{c1}. Thus a rapid transition to state II occurs. In this state C_1 recharges rapidly to a voltage E_{bb}, while C_2 discharges until the second triode begins to conduct and initiates the transition returning the circuit to state I. This circuit behavior is illustrated by the waveforms in Fig. 9.19(d).

The duration of the interval δ_1 is determined by the exponential expression for e_{c1}. The initial value of e_{x1} is assumed to be E_{bb} and this fixes the initial value of e_{c1} in the interval δ_1. The final value of e_{c1} is $-E_{co1} = -E_{bb}/\mu$. Thus the expression for the interval δ_1 is

$$\delta_1 = \tau_{d1} \ln \left[\frac{R_{g1}}{R_{g1} + (r_{p2} \| R_{b2})} \right] \left[\frac{E_1 + E_{bb} - E_{a2}}{E_1 + E_{co1}} \right] \quad (9.11)$$

A similar expression is obtained for δ_2 by interchanging subscripts. Assuming the initial value of e_{x1} to be E_{bb} amounts to saying that each capacitor charges fully during the interval generated by the discharge of the other capacitor. For this condition to exist, we should have τ_{c2} much less than τ_{d1} and τ_{c1} much less than τ_{d2}. If capacitors C_1 and C_2 have comparable values this requirement reduces to R_{g1} much greater than R_{b2} and R_{g2} much greater than R_{b1}. With these restrictions, the resist-

---

* For simplicity assume $r_{g1} = r_{g2} = 0$ so that e_{c1} and e_{c2} do not exceed zero in states I and II, respectively.

ances $r_{p1}\|R_{b1}$ and $r_{p2}\|R_{b2}$ can be neglected in the circuit models and in the expressions for δ. The waveforms in (d) are drawn subject to this assumption so that $e_{b1} = E_{a1}$ during interval δ_2 and $e_{b2} = E_{a2}$ during δ_1.

Assuming that both r_{g1} and r_{g2} are zero, we obtain perfect limiting of e_{c1} and e_{c2} at zero. With a more realistic value of grid-to-cathode resistance, the recharge of capacitor C_1 causes a positive overshoot on the waveform of e_{c1} at the beginning of interval δ_2 (shown as a dotted line). This variation in e_{c1} causes an amplified replica of the overshoot to appear on the waveform of e_{b1}. The overshoot on e_{b1} is coupled by C_2 to the waveform of e_{c2}. Since this grid voltage is more negative than the cutoff value the overshoot has no effect on the plate voltage e_{b2} and hence proceeds no further in the circuit. In similar fashion an overshoot on e_{c2} at the beginning of interval δ_1 appears on e_{b2} and e_{c1} (shown in light line). We can include the effects of such an overshoot on the calculation of the interval by means of an exponentially time-varying generator $G \Delta e_{c1}$ in series with E_{a2} in the circuit model of Fig. 9.19(c). If we approximate the overshoot by an impulse with the same area, we see that the effect of the overshoot is to remove some charge from the discharging capacitor and thus shorten the interval slightly. In most cases the error in δ owing to neglecting overshoot is very small.

The numerical values of pertinent points on the waveforms of Fig. 9.19(d) are readily determined once the extremes of the capacitor voltages are known. Thus if $\tau_{c2} < \delta_1/5$ and $\tau_{c1} < \delta_2/5$ we know the maximum value of capacitor voltage (e_{x1} or e_{x2}) is E_{bb}. The minimum values are ($E_{a2} + E_{co1}$) and ($E_{a1} + E_{co2}$) respectively. Since $\Delta e_{x1} = \Delta e_{x2} = 0$ at the transitions between states I and II, we have $\Delta e_{b1} = \Delta e_{c2}$ and $\Delta e_{b2} = \Delta e_{c1}$ at these instants.

During interval δ_1, grid voltage e_{c1} heads toward $+E_1$ and during δ_2, the waveform of e_{c2} heads toward $+E_2$. So long as each of these end points lies above $-E_{co}$ the circuit will be astable and hence will generate waveforms of the general character shown. If the polarity of either E_1 or E_2 is reversed and made large enough so that the final value of one grid voltage is more negative than that required for plate-current cutoff, the circuit will be monostable. The triode with the negative grid-return voltage will be normally off and the other one will be conducting. An externally generated trigger pulse can then be used to initiate a transition to the unstable state. After an interval δ_1, the circuit will switch back to the stable state and remain dormant until another trigger pulse is applied. The circuit and typical waveforms for monostable operation are shown in Fig. 9.20 with state II stable and state I unstable. In the waveforms shown, the interval δ_1 is generated as before but δ_2 becomes $\delta_s - \delta_1$, where δ_s is the repetition interval of the synchronizing signal.

For this mode of operation we require $C_2 R_{g2} \gg C_1 R_{g1}$ so that e_{c2} remains near zero during the interval δ_1. Otherwise δ_1 will depend on $\tau_{c2}{}^{(+)}$ and $\tau_{c2}{}^{(-)}$ more heavily than on τ_{d1}.
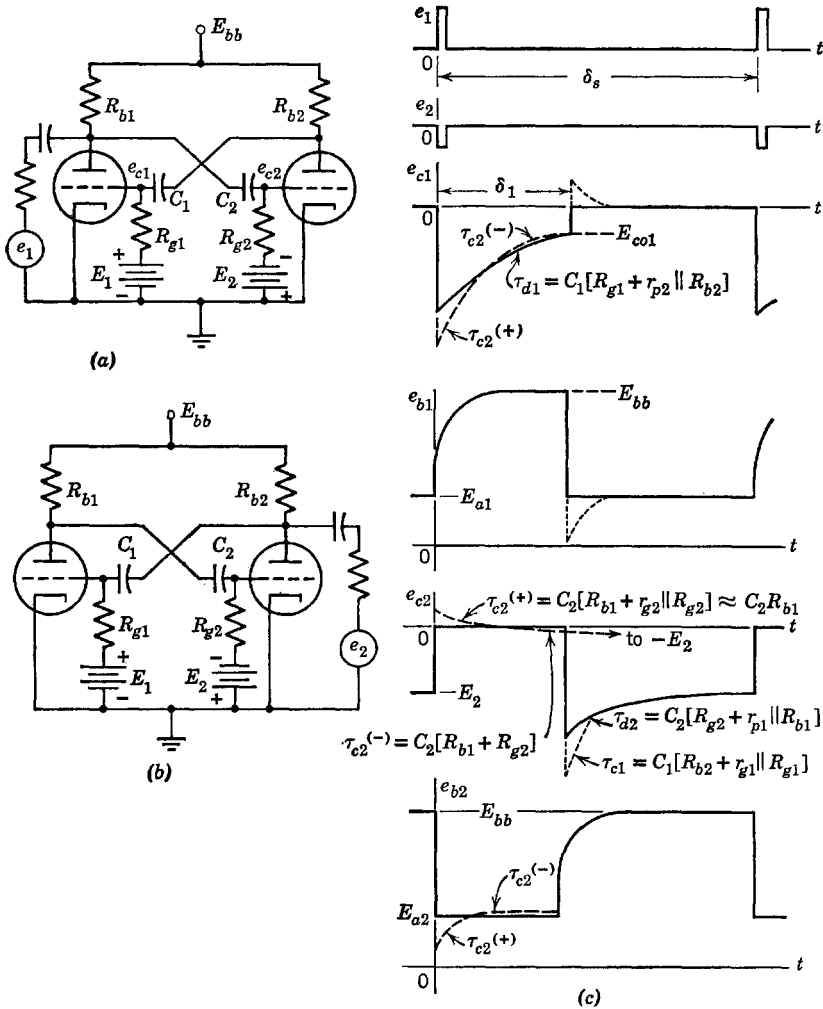
Fig. 9.20. Monostable operation of plate-coupled multivibrator.

The synchronizing signal must be large enough to turn the *off* triode *on* in order to effect triggering. With positive-going pulses applied as shown in Fig. 9.20(a) the amplitude required must exceed the difference between

the grid voltage and the cutoff voltage. Alternatively, as shown in (b), triggering can be accomplished with a negative-going pulse of smaller amplitude applied to the opposite side of the circuit. In this case the trigger pulse is amplified and inverted by the normally *on* triode and applied to the grid of the *off* triode.

9.15   Plate-Coupled Bistable Circuit

If both grids of a symmetrical two-stage plate-coupled circuit are returned to voltages more negative than $-E_{bb}/\mu$, the circuit becomes bistable. One form of the circuit is shown in Fig. 9.21(a). Direct-

Fig. 9.21.   Bistable plate-coupled circuit.

current plate-to-grid couplings are provided by R_1 and R_2. An alternative form of the circuit is shown in (b). Here the negative supply voltage is avoided by operating the cathodes at a positive potential. Let R_k be chosen so that

$$E_k = E_{cc} = I_a R_k \qquad (9.12)$$

where I_a is quiescent plate current for either triode. Then if

$$E_{bb}' = E_{bb} + E_{cc} \qquad (9.13)$$

and all element values correspond, the circuit of (b) is exactly like that of (a), except for the voltage reference level.

The two circuit models shown in (c) correspond to state I (first triode *off*, second triode *on*). Similar circuits with subscripts 1 and 2 interchanged apply to state II. However, since the circuit is symmetrical we need only consider one of the two states.

Transitions from state I to state II or from II to I can occur only when initiated by a trigger source. Two modes of bistable operation result from the two methods of triggering shown with dotted lines in Fig. 9.21 (a) and (b). Either method can be used with either circuit. As shown by the idealized waveforms, for the triggering circuit used in (a), each of the alternatively positive and negative pulses in the synchronizing waveform induces a transition, when e_s is applied at only one plate. When the synchronizing pulses are applied to both plates as in (b), the positive pulses have no effect since they are not passed by the diodes. The *on* tube has a low plate voltage that back-biases the diode connected to its plate, so if $E_{bb} = 200$ and $E_a = 100$, negative synchronizing pulses of amplitude less than 100 volts will be coupled only to the plate of the *off* tube. Thus negative pulses are applied alternately to one plate and then the other so that transitions occur as indicated by the waveforms in (b). In this mode of operation the circuit is a scale-of-two counter since two complete cycles of e_s are required to produce one cycle of e_{b1} or e_{b2}. The compensating capacitor (C_1 or C_2) couples the negative pulse to the grid of the *on* tube where the pulse is amplified and inverted. It is then coupled to the grid of the *off* tube by the other capacitor. If we let $CR = C_{gk}R_g$ in Fig. 9.22 (where C_{gk} is the tube and wiring capacitance between grid and cathode), the coupling from plate to grid is "compensated." We then have a voltage divider with high-frequency attenuation equal to low-frequency attenuation. In a practical circuit, the coupling capacitors are made several times larger than the value thus specified ($C > C_{gk}R_g/R$) in order to enhance circuit gain during the rapid transitions. However, a transition is not complete until one coupling capacitor has charged and the other discharged to the appropriate new values. This condition places a practical upper limit on the

value of C because the recovery time following a transition increases with the value of C. If we assume that $C_{gk} = 10$ μμf, $R = 400$ kilohms, and $R_g = 100$ kilohms, then "compensation" would occur with $C = 2.5$ μμf, whereas the value of C more likely to be used is 25 or 50 μμf.

$$\tau = [C + C_{gk}]\left[\frac{RR_g}{R + R_g}\right]$$

① $C < C_{gk}R_g/R$
② $C = C_{gk}R_g/R$
③ $C > C_{gk}R_g/R$

$\tau_2 = CR = C_{gk}R_g$
$\tau_1 < \tau_2 < \tau_3$

③ $\dfrac{EC}{C + C_g} > \dfrac{ER_g}{R_g + R}$

② $\dfrac{EC}{C + C_g} = \dfrac{ER_g}{R_g + R}$

① $\dfrac{EC}{C + C_g} < \dfrac{ER_g}{R_g + R}$

(a)

$$\tau = \left[C_p + \frac{CC_{gk}}{C + C_{gk}}\right]R_b$$

(b)
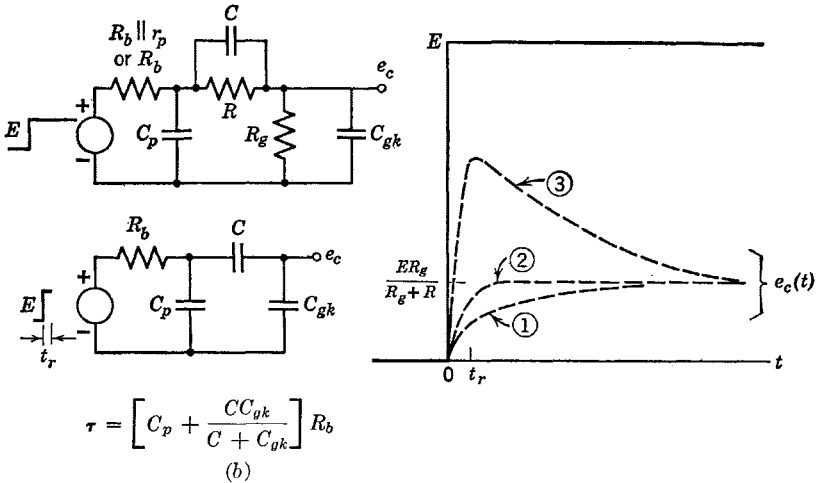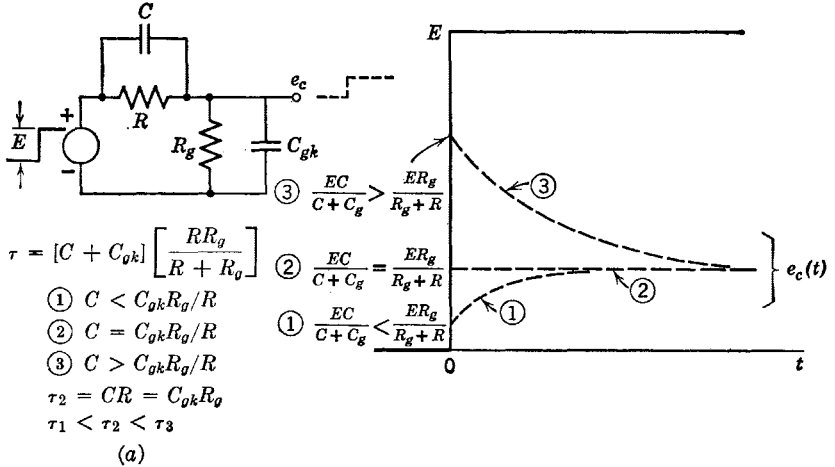
Fig. 9.22. Step response of coupling circuit.

The step response of the coupling circuit (neglecting source resistance R_b or $R_b \| r_p$) is indicated in Fig. 9.22(a) for capacitance C less than, equal to, and greater than the value required for "compensation." If source resistance (and capacitance) are included as in (b), the instan-

taneous change in the voltages are replaced by finite rise or fall times. These can be determined approximately by neglecting the shunting effects of the large resistances R and R_g during the rapid rise. The simplified analysis of the coupling circuit outlined in Fig. 9.22 indicates the considerations involved in making a rapid determination of the approximate transition time of the plate-coupled bistable circuit.

9.16   Blocking Oscillator

Although the blocking oscillator circuit may take a variety of forms, either of the circuits shown in Fig. 9.18(c) or (d) can be used to describe typical circuit behavior. In a physical circuit, drastic nonlinearities can occur either in the active element due to cutoff, etc., or in the iron-core transformer due to saturation, or in both. Let us consider the case in which the transformer can be treated as a linear circuit element for the range of currents and voltages involved in the circuit operation and let the active element be a p-n-p junction transistor.

With the polarity of the bias voltage E_1 as shown in Fig. 9.23(a), the transistor is normally in the *off* state, hence the circuit will be monostable. Triggering is accomplished by making the emitter diode conduct with a negative-going pulse applied between base and ground. This effect can be obtained by applying a positive-going pulse between collector and ground since the transformer will invert the pulse and apply it to the base. Alternatively, a third winding on the transformer can be used to couple the trigger pulse to the base circuit.

The trigger pulse initiates the flow of base current which in turn produces a collector current $a/(1-a)$ larger in the forward-gain state. With the $n:1$ turns ratio shown, the base current will be ideally n times the collector current and the total current gain available is approximately $na/(1-a)$. For nominal values of transformer turns ratio this gain will exceed unity by an appreciable amount. The circuit is therefore unstable in the forward-gain state, so that the transistor is driven into the saturation state (both emitter and collector diodes closed). Referring to the circuit model of Fig. 9.23(b), we see that the current i_d in the collector diode is the excess over that required to sustain the circuit in the saturation state. Since the collector is at ground potential under these conditions, the voltage across the transformer primary is constant at E_{cc}, hence in this idealized case the magnetizing current i_m builds up linearly with time. As it does, there is less excess current available for the collector diode, and when $i_d = 0$ the saturation state can no longer be sustained. Then the transformer voltage v_1 diminishes
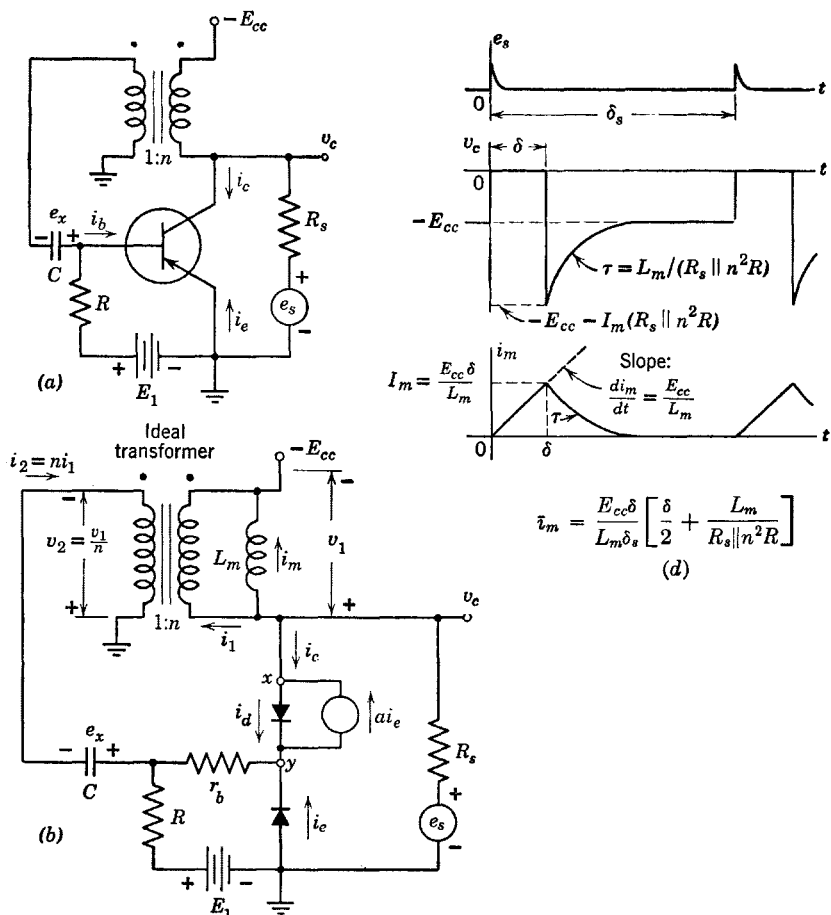
Assume $r_b \ll R$ and R_s; $e_x = E_1$

$$i_2 = (E_{cc}/n - E_1)/r_b$$
$$= n[ai_e - i_d - i_m]$$
$$= i_e(1 - a) + i_d$$

At $t = 0$, $i_m = 0$ and $i_d = \left[\dfrac{E_{cc}/n - E_1}{r_b}\right]\left[a - \dfrac{1-a}{n}\right]$

At $t = \delta$, $i_d = 0$ and $i_m = I_m = \left[\dfrac{E_{cc}/n - E_1}{r_b}\right]\left[\dfrac{a}{1-a} - \dfrac{1}{n}\right]$; $\delta = \dfrac{I_m L_m}{E_{cc}}$

(c)

Fig. 9.23. Monostable blocking oscillator (δ controlled by L_m).

from the value E_{cc} and this change produces a regenerative transition through the forward-gain state to the inactive or *off* state. The magnetizing current then decays exponentially, and, if i_m reaches zero prior to the next trigger, the next cycle of operation will be identical. The assumption that the interval between trigger pulses is very long, compared with the duration of the pulse generated by the circuit, permits i_m to return to zero and e_x to return to the voltage E_1 prior to each trigger pulse.

The duration of the pulse generated may depend on the capacitor, the transformer magnetizing inductance, or both. Which of these cases occurs is determined by the rate at which the capacitor charges relative to the rate at which current builds up in the magnetizing inductance of the transformer. For the case in which the build-up of magnetizing current is largely responsible for controlling the pulse duration, the analysis is simplified if we assume the capacitor voltage remains relatively constant during the pulse and hence never departs appreciably from the value E_1. In this case the capacitor C can be represented by a battery of voltage E_1 in the circuit model of (*b*). The transformer is represented by an ideal transformer of ratio $n:1$ shunted by the magnetizing inductance L_m. The effects of stray parameters such as leakage inductances and shunt capacitances are not included in this simplified analysis. One additional assumption made in deriving the equations of Fig. 9.23 (*c*) is that r_b is much less than R or R_s. The current i_2 immediately following the synchronizing pulse must be

$$i_2 = \frac{E_{cc}/n - e_x}{r_b} = \frac{E_{cc}/n - E_1}{r_b} \qquad (9.14)$$

Since i_2 must be positive to drive the transistor through the forward gain state and into saturation we must satisfy the condition $E_{cc}/n > E_1$. Two other equations for i_2 which express the continuity of current at the nodes x and y respectively are

$$i_2 = n[ai_e - i_d - i_m] \qquad (9.15)$$

$$i_2 = i_e(1 - a) + i_d \qquad (9.16)$$

When a trigger pulse occurs $(t = 0)$, the magnetizing current is zero and the current equations can be solved for i_d. The fact that i_d must be greater than zero imposes another relationship on E_{cc}, E_1, and n. The pulse duration

$$\delta = L_m I_m / E_{cc} \qquad (9.17)$$

is found by noting that the same current equation can be solved for I_m, the value of i_m when $i_d = 0$. After the pulse, the current i_m decreases

exponentially with time constant

$$\tau = L_m/(R_s\|n^2R) \tag{9.18}$$

since the inductor faces resistance R_s in parallel with n^2R reflected through the transformer. The waveforms of i_m and v_c shown in Fig. 9.23 (d) illustrate these points. Note that the average voltage v_1 across the inductor L_m must be zero. Thus the area of the rectangular pulse $(E_{cc}\delta)$ is also the area of the exponential overshoot measured with respect to $-E_{cc}$. This area is the amplitude of the exponential times the time constant, namely

$$\text{Area} = [I_m(R_s\|n^2R)][L_m/(R_s\|n^2R)] = I_mL_m = E_{cc}\delta \tag{9.19}$$

The magnetizing current has an average value readily calculated from the waveform by computing the area and dividing by δ_s.

The current i_2 is constant during the pulse and small between pulses. With i_m known, Eqs. 9.15 and 9.16 can be used to determine i_d and i_e in terms of i_2 and i_m to complete the specification of all pertinent variables in the circuit.

We have calculated a pulse duration based on magnetizing current buildup, subject to the assumption that the coupling-capacitor voltage remains unchanged during the pulse. Let us now consider the other extreme case; namely, a pulse duration determined by the coupling capacitor alone.

For simplicity let the pulse transformer be ideal so that magnetizing current is zero during the pulse. The circuit model of Fig. 9.23 (b) can be used for this calculation if we let the magnetizing inductance L_m be very large.

At the instant a synchronizing pulse occurs, the transistor is driven through the active state to the saturation state. If the interval between synchronizing pulses is large, we can assume the initial capacitor voltage to be E_1. The equilibrium conditions calculated for the previous case at $t = 0$ also apply in this case. The circuit model for calculating the waveforms is shown in Fig. 9.24 (a). The Thevenin equivalent in (b) applies only *during* the pulse interval δ.

As the capacitor charges, the charging current i_2 diminishes exponentially toward zero. The pulse terminates when capacitor voltage is nearly E_{cc}/n and charging current is approximately E_1/R. At these values, the current in the collector diode is zero so that the saturation state can no longer be sustained. When i_d is zero we have $-i_c = i_1 = i_2/n = i_ba/(1 - a)$, but since $anR \gg (1 - a)(R + r_b)$ the base voltage is nearly zero and E_1/R is a valid approximation for i_2. At $t = \delta$ the transistor returns rapidly to the *off* state and the capacitor discharges

slowly through the large resistance R in series with R_s/n^2 to the quiescent value E_1.
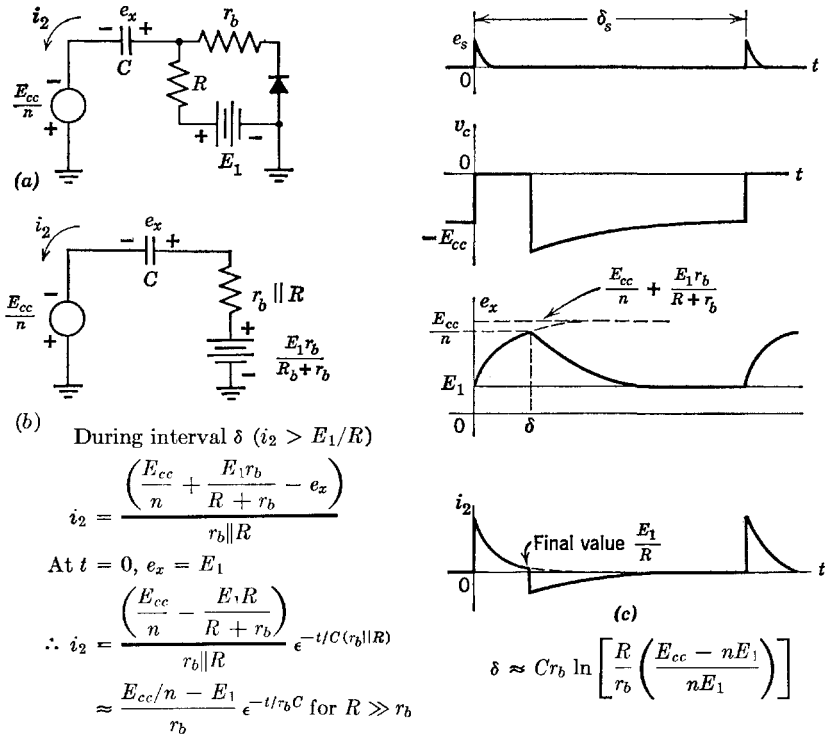
During interval δ $(i_2 > E_1/R)$

$$i_2 = \frac{\left(\dfrac{E_{cc}}{n} + \dfrac{E_1 r_b}{R + r_b} - e_x\right)}{r_b \| R}$$

At $t = 0$, $e_x = E_1$

$$\therefore i_2 = \frac{\left(\dfrac{E_{cc}}{n} - \dfrac{E_1 R}{R + r_b}\right)}{r_b \| R} \epsilon^{-t/C(r_b\|R)}$$

$$\approx \frac{E_{cc}/n - E_1}{r_b} \epsilon^{-t/r_b C} \text{ for } R \gg r_b$$

$$\delta \approx C r_b \ln\left[\frac{R}{r_b}\left(\frac{E_{cc} - nE_1}{nE_1}\right)\right]$$

Fig. 9.24. Monostable blocking oscillator (δ controlled by C).

The pulse duration δ depends on the capacitor charging current. This current is

$$i_2 = \frac{\dfrac{E_{cc}}{n} + \dfrac{E_1 r_b}{R + r_b} - e_x}{r_b \| R} \tag{9.20}$$

Since $e_x = E_1$ at $t = 0$ we have

$$i_2 = \frac{\dfrac{E_{cc}}{n} - \dfrac{E_1 R}{R + r_b}}{r_b \| R} \epsilon^{-t/C(r_b\|R)} \tag{9.21}$$

The collector diode current, which determines the end-point of the

interval, is given by

$$i_d = \frac{\dfrac{E_{cc}}{n} - e_x}{r_b \| R}\left[\left(\frac{R}{R+r_b}\right)\left(\frac{a}{1-a}\right) - \frac{1}{n}\right] - \frac{E_1}{nR} \qquad (9.22)$$

with magnetizing current assumed to be zero. If this expression is compared with the one given in Fig. 9.23(c), we see that the major difference is the replacement of the constant E_1 by the variable e_x. The inclusion of R in parallel with r_b makes relatively little difference numerically since we still assume R to be much greater than r_b. When $i_d = 0$ in Eq. 9.22, the end value of e_x is E_{cc}/n if the factor involving turns ratio and current gain is positive. The restriction imposed on the transformer turns ratio n by this requirement is easily satisfied in practice.

The value of pulse duration calculated from the capacitor voltage waveform is

$$\delta = C(r_b \| R) \ln\left[\frac{\dfrac{E_{cc}}{n} - \dfrac{E_1 R}{R+r_b}}{\dfrac{E_1 r_b}{R+r_b}\left(1 + \dfrac{1}{\left(\dfrac{R}{R+r_b}\right)\left(\dfrac{a}{1-a}\right) - \dfrac{1}{n}}\right)}\right] \qquad (9.23)$$

Since $R \gg r_b$ and $na \gg (1-a)$ we have

$$\delta \approx Cr_b \ln\left[\frac{R}{r_b}\left(\frac{E_{cc} - nE_1}{nE_1}\right)\right] \qquad (9.24)$$

Both the magnetizing-current and the capacitor voltage influence the astable behavior of the blocking oscillator circuit. To obtain continuous oscillations from the circuit of Fig. 9.23(a), we need only reverse the polarity of E_1, so that the transistor is normally biased *on*. The synchronizing source can then be removed as in Fig. 9.25(a). However, the resistance is retained to dissipate the energy in L_m after the pulse.

In most free-running blocking oscillator circuits, the pulse duration is controlled largely by L_m, whereas the interval between pulses is controlled largely by C. When a pulse occurs, the transistor switches to the saturation state and magnetizing current builds up in L_m. If we assume that i_2 remains approximately constant as in the first case considered for monostable operation, the pulse duration δ will be very nearly given by $L_m I_m / E_{cc}$. The constant current i_2 passing through C for the interval will produce a change $\Delta e_x = i_2 \delta / C$. Equations relating variables during the pulse are given in Fig. 9.25(b).

The increment Δe_x must be small compared with E_{cc}/n to avoid reducing i_2 during the pulse, and be slightly positive in order to hold the transistor off while C discharges through $R(r_c = RC)$. If the value

of E_1 is made small, this condition can be met. An approximation for the interval between pulses is given in (c) and pertinent waveforms are shown in (d). In general, the calculation of pulse duration must
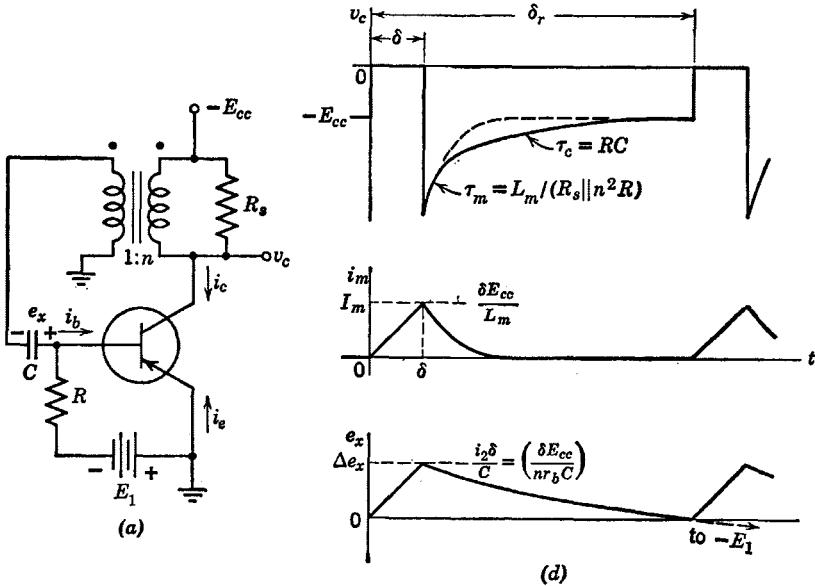
Assume $r_b \ll R$, $n^2 R_s$

Δe_x small

During pulse

$$\frac{di_m}{dt} = \frac{E_{cc}}{L_m}$$

$$I_m = \frac{(E_{cc}/n)}{r_b}\left[\frac{a}{1-a} - \frac{1}{n}\right]$$

$$\delta = L_m I_m / E_{cc}$$

$$i_2 \approx \frac{[E_{cc}/n]}{r_b} = \text{constant}$$

$$\therefore \Delta e_x = \frac{i_2 \delta}{C} = \left(\frac{\delta E_{cc}}{n r_b C}\right)$$

(b)

Interval between pulses

$$(\delta_r - \delta) \approx RC \ln\left(\frac{\Delta e_x + E_1}{E_1}\right)$$

(c)

Fig. 9.25. Astable blocking oscillator.

include the effects of both magnetizing current build-up and coupling capacitor charge.

Practical free-running blocking oscillators readily generate very unsymmetrical rectangular waveforms with large values of the ratio

δ_1/δ_2. Multivibrators, however, are more suitable for generating rectangular waves with small values of δ_1/δ_2 (say, less than 20).

9.17   Sinusoidal Oscillations

If we can predict that nearly sinusoidal oscillations will occur in a particular circuit, a linear incremental analysis can yield a rapid indication of the conditions for oscillation and the approximate frequency. As an illustrative example, consider the tuned-plate circuit shown in Fig. 9.26(a). The circuit is essentially the same as that of the blocking oscillator considered in the preceding article. However, the coefficient of mutual coupling in the transformer is relatively small here, whereas in the pulse transformer used in the blocking oscillator, the coupling coefficient is very close to unity.

In the operation of the tuned-plate circuit the power to the load resistor R plus any incidental circuit losses are supplied by the incremental power gain of the triode. The grid of the triode is driven sufficiently negative to produce plate-current cutoff and sufficiently positive to cause grid current. Waveforms of plate current and grid voltage will be approximately as indicated in Fig. 9.26(b). The pulses of plate current can be thought of as current impulses that charge C and cause the circuit to oscillate at its natural frequency. The time constant $C_g R_g$ is assumed large compared to the period so that e_x remains nearly constant during a cycle. The energy given to the tuned circuit will build up the amplitude of oscillations until the energy dissipated equals the energy supplied. As the amplitude increases, the duration of plate-current pulses diminishes, since the grid-cathode voltage is more negative than $-E_{c0}$ during a larger fraction of a cycle. This means that the "average amplification" of the triode decreases as the amplitude of oscillations increases and hence an equilibrium amplitude is established.

Fourier analysis of the plate current i_b shows the presence of a sinusoidal component i_{b1} at the fundamental frequency. The linear incremental analysis indicated in Fig. 9.26(c) assumes each waveform of current or voltage to be represented by a sine wave at this fundamental frequency. We can therefore consider the circuit to be a single-frequency amplifier with the incremental grid voltage e_g being amplified by the triode and coupled back to the grid circuit. (See Article 8.32.) Self-oscillations occur when the voltage coupled back has the amplitude and phase required to sustain the oscillations. Under these conditions the linear incremental circuit shown in Fig. 9.26(c) sustains any value of e_g and/or e_p that has been established. Since this relation holds for

$$g_m E_g + E_p \left[\frac{1}{r_p} + \frac{1}{R} + \frac{1}{j\omega L} + j\omega C \right] = 0$$

$$E_g = \pm j\omega M \left(\frac{E_p}{j\omega L} \right)$$

$$\pm g_m \frac{M}{L} E_p + E_p \left[\frac{1}{r_p} + \frac{1}{R} + \frac{1}{j\omega L} + j\omega C \right] = 0$$

Imaginary terms   Real terms

$$\frac{1}{j\omega L} + j\omega C = 0 \qquad \pm g_m \frac{M}{L} + \frac{1}{r_p} + \frac{1}{R} = 0$$

$$1 - \omega^2 LC = 0$$

$$\omega = 1/\sqrt{LC} \qquad g_m = \frac{L}{M} \left[\frac{1}{r_p} + \frac{1}{R} \right]$$

(c)
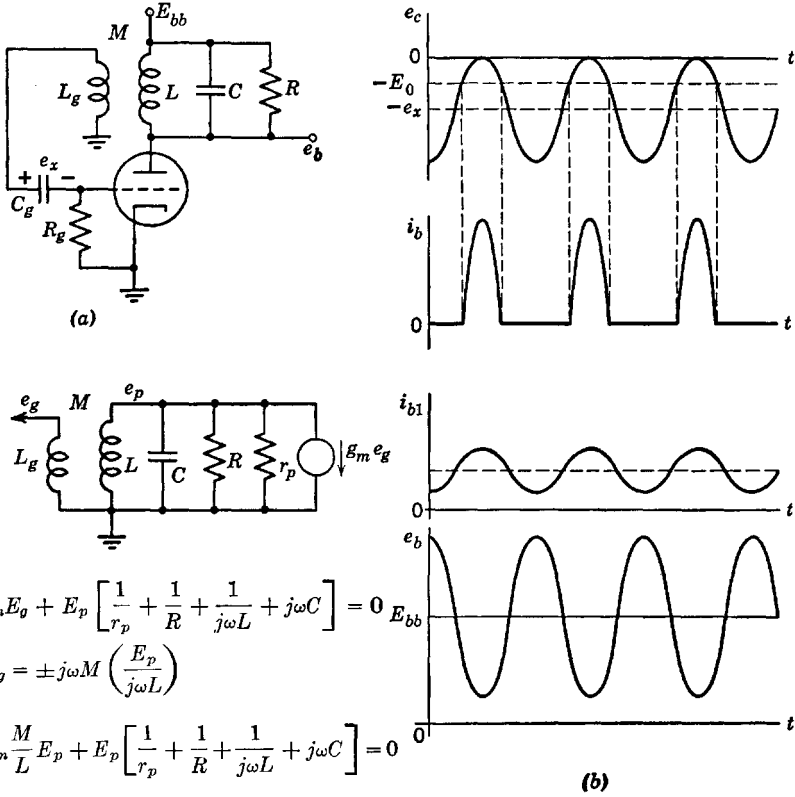
Fig. 9.26. Tuned-plate circuit to generate sinusoidal oscillations.

any (nonzero) value of e_p, the real and imaginary parts must separately equate to zero. These two equations yield the oscillation frequency

$$\omega = \frac{1}{\sqrt{LC}} \tag{9.25}$$

and the condition for oscillations to start,

$$g_m \geqslant \frac{L}{M} \left[\frac{1}{r_p} + \frac{1}{R} \right] \tag{9.26}$$

In an actual nonlinear circuit an appropriately weighted "average" current gain corresponds to the g_m appearing here.

SUPPLEMENTARY READING

A. A. Andronow, C. E. Chaikin, *Theory of Oscillations*, Princeton University Press, Princeton, N.J., 1949.

B. Chance, *Waveforms*, McGraw-Hill, New York, 1948.

William A. Edson, *Vacuum-Tube Oscillators*, John Wiley and Sons, New York, 1953.

L. P. Hunter, editor *Handbook of Semiconductor Electronics*, McGraw-Hill, New York, 1956.

J. Millman and H. Taub, *Pulse and Digital Circuits*, McGraw-Hill, New York, 1956.

J. F. Reintjes and G. T. Coate, *Principles of Radar*, 3rd edition, McGraw-Hill, New York, 1952.

Samuel Seely, *Electron-Tube Circuits*, McGraw-Hill, New York, 1950.

PROBLEMS (See Appendices for device curves)

9.1. Referring to Fig. 9.2, plot the driving-point curve (v_1 vs. i_e) for ($R_b + r_b$) equal to 0.1 kilohm, 1 kilohm, 10 kilohms, and 100 kilohms. Let $V_{cc} = 50$ volts, $R_c = 10$ kilohms, $r_{ef} = r_{cf} = 0$, and $r_{er} = r_{cr} = 30$ kilohms, and assume that $a = 0$ for $i_e < 0$, and $a = 2.5$ for $i_e \geqslant 0$. Determine the numerical value of ($r_b + R_b$) for which $dv/di = 0$ in the forward-gain state. Note that if the base resistance r_b exceeds this value, the circuit has a negative-resistance region even with $R_b = 0$. Alternatively, with $R_b = 0$, determine the range of values of R_c for which a negative-resistance region is obtained.
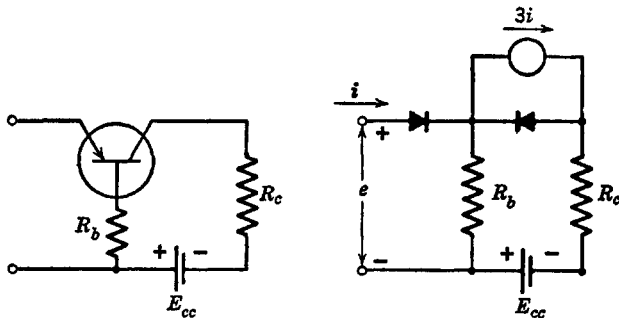
Fig. P9.1

9.2. Sketch and dimension the input e-i characteristics for the circuit shown in Fig. P9.1 assuming the idealized piecewise-linear model as given.

9.3. From Fig. P9.2, sketch and dimension $i(t)$ and $v(t)$ assuming $v = 0$ at $t = 0$.
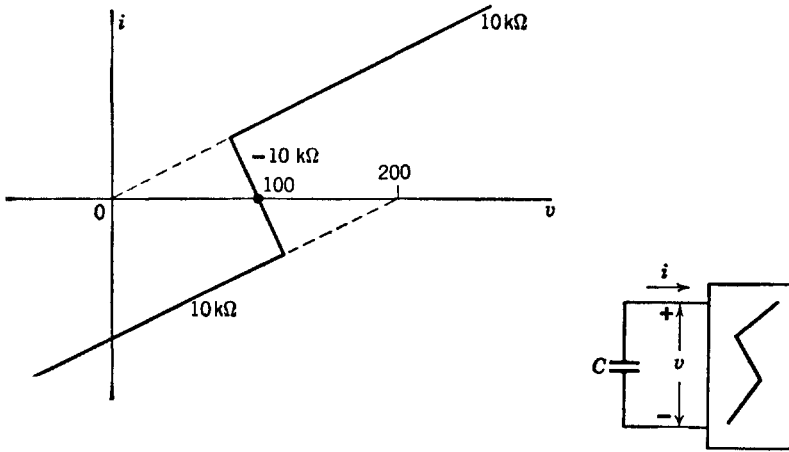
Fig. P9.2

9.4 Sketch and dimension the waveforms $v(t)$ and $i(t)$ for the relaxation oscillator shown in Fig. P9.3.
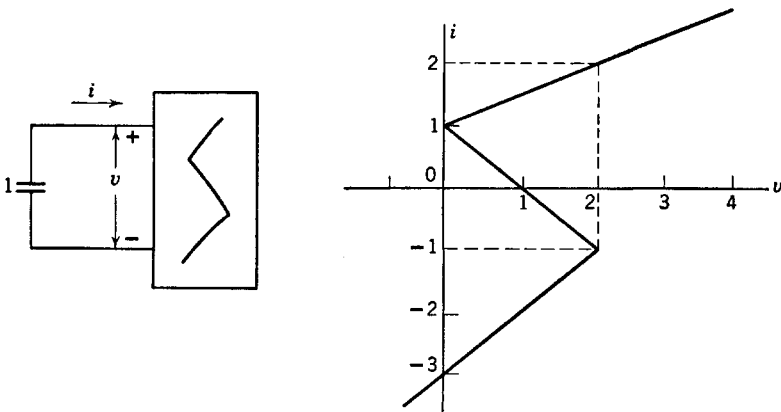
Fig. P9.3

9.5. A vacuum tube circuit exhibits a current-controlled negative-resistance characteristic that may be represented by segments of the following straight lines:

$$v - 50i - 50 = 0$$
$$v + 25i - 200 = 0$$
$$v - 75i - 300 = 0$$

Determine the free-running repetition frequency obtained when a 1000 $\mu\mu f$ capacitor is connected to the terminals of the circuit.

9.6. The driving-point characteristic of an electronic device may be represented by the circuit model shown in Fig. P9.4. Determine the energy-storage
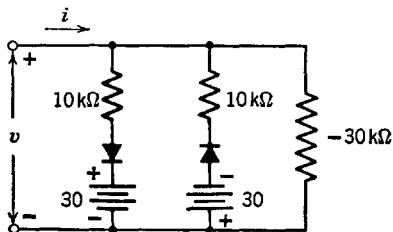
Fig. P9.4

element that should be connected to the terminals in order to generate 50 kcps relaxation oscillations.

9.7. Sketch and dimension the waveforms of v and i for the circuit shown in Fig. P9.5. Let $v = 10$ volts and $i = 1$ ma at $t = 0$.
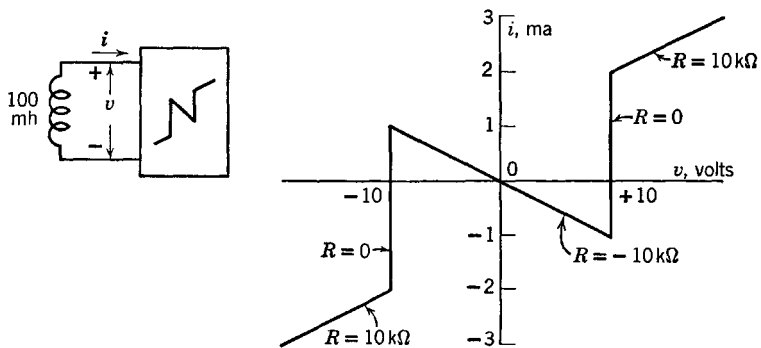
Fig. P9.5

9.8. The driving-point characteristic of a circuit is shown in Fig. P9.6. Sketch and dimension the waveform $v(t)$ that results when a 100 mh inductance (initial current 5 ma) is connected to the terminals at $t = 0$.
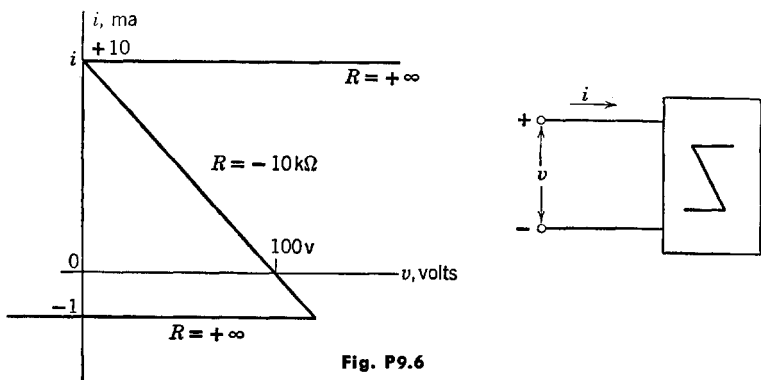
Fig. P9.6
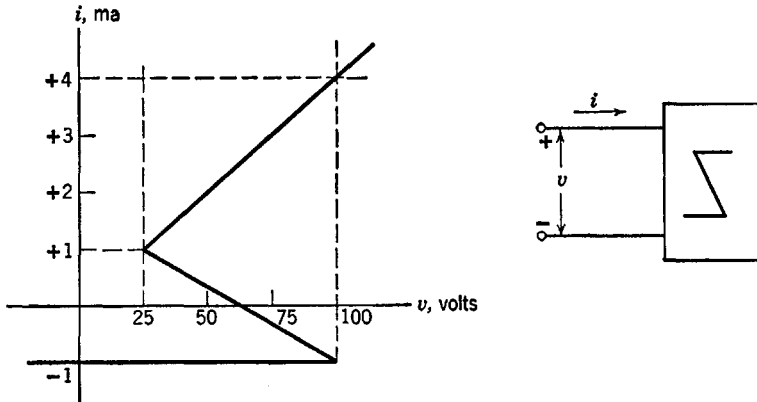
Fig. P9.7

9.9. The driving point characteristic for a circuit is shown in Fig. P9.7. Sketch and dimension the waveform of v obtained if:

(a) An initially uncharged 1000 $\mu\mu$f capacitor is connected to the terminals;

(b) An initially uncharged 1000 $\mu\mu$f capacitor in series with a 10-kilohm resistor is connected to the terminals.
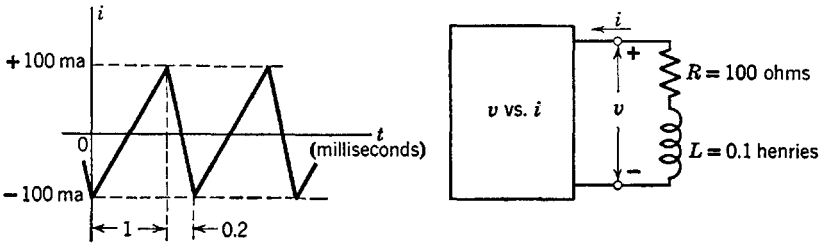
Fig. P9.8

9.10. The magnetic deflection coil for a cathode-ray tube is to be driven by a linearly varying current waveform as indicated in Fig. P9.8. It has been proposed that a two-terminal device with the appropriate negative resistance characteristic could be connected to the coil to produce this required waveform. If the coil has an inductance $L = 0.1$ h and resistance $R = 100$ ohms, sketch and dimension a v–i curve that would produce the desired inductor current waveform.

9.11. Assume that Tubes T_1 and T_2 in Fig. P9.9 are resistive and piecewise-linear with cutoff for $e_{b1} + \mu_1 e_{c1}$ and $e_{b2} + \mu_2 e_{c2}$ negative, respectively.

(a) Sketch and dimension the curve of i vs. v, as determined by the resistive portion of the circuit.

(b) What are the conditions on R_1 and R_2 that result in a negative-resistance region?

(c) Sketch the path of operation in the i vs. v plane when a negative resistance region is present.
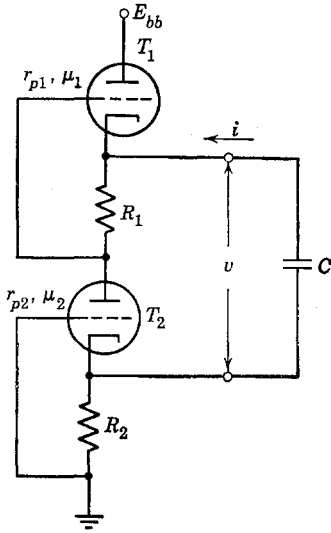
Fig. P9.9
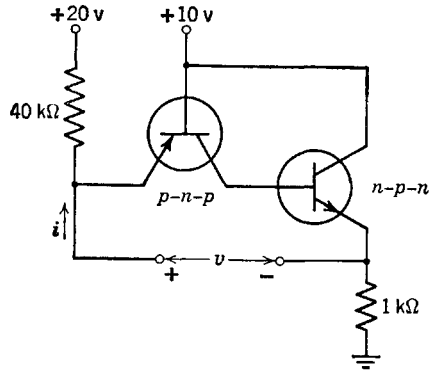
Fig. P9.10

(d) Calculate the period of oscillation and sketch the waveforms of v and i.

(e) As capacitance C is made smaller, the frequency of oscillation increases. Why not, then, use this circuit to generate a frequency of 1000 megacycles, simply by choosing C small enough?

(f) Will the circuit oscillate with a large resistance in series with C?

(g) Will the circuit oscillate with large capacitances added in parallel with R_1 and R_2?

9.12. Using ideal-diode models for the p-n-p and n-p-n transistors in the circuit of Fig. P9.10, sketch and dimension the driving-point curve v vs. i. Assume $a = 0.98$ for both transistors.

9.13. Modify the result of the preceding problem to include the effects of a 100-ohm base resistance in each transistor model.

9.14. Sketch and dimension the waveforms $v(t)$ and $i(t)$ obtained with a 1000 $\mu\mu$f capacitance between the driving-point terminals in Fig. P9.10.
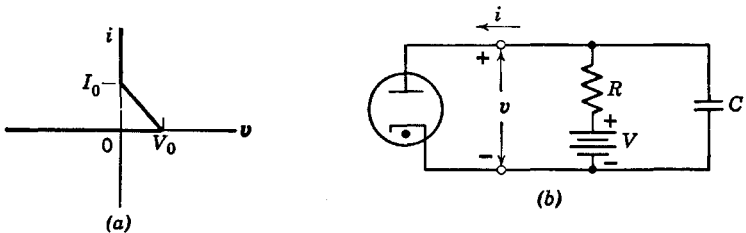
Fig. P9.11

9.15. A gas diode curve is approximated as shown in Fig. P9.11. Sketch and dimension the waveform of $v(t)$ for $V = 2V_0$ and $R = 4V_0/I_0$.

9.16. Each cold-cathode gas diode in the circuit of Fig. P9.12(a) has the odd-symmetric volt-ampere characteristic shown in Fig. P9.12(b). Let us assume the circuit has been at rest for some time with switch K closed. At time $t = 0$, K opens. Sketch and dimension e_1 and e_2 as functions of time.

(a)   (b)
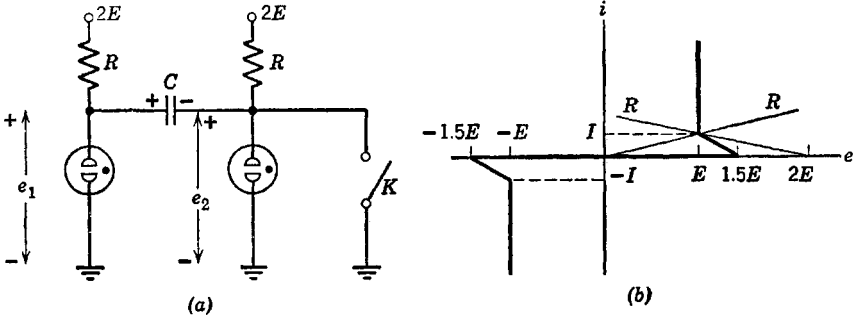
Fig. P9.12

9.17. The e_b vs. i_b characteristic for the tetrode in the circuit of Fig. P9.13 is shown in Fig. P9.14.

(a) Sketch and dimension the i vs. e characteristic.
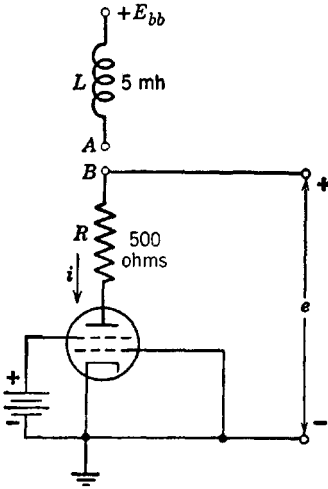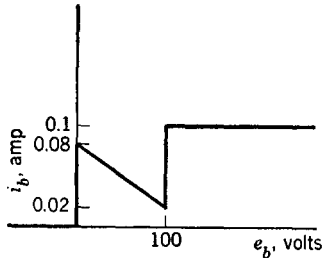
Fig. P9.13   **Fig. P9.14**

(b) Let E_{bb} equal the value of e at the center of the negative resistance region. What is this value?

(c) Show the operating path on the i vs. e characteristic with terminals A and B connected.

(d) Sketch carefully i versus time for two cycles around the operating path.

9.18. Determine the minimum value of current source required to insure monostable operation for the series triode relaxation oscillator in Fig. 9.11. Using a source having twice this magnitude, determine the required charge

content of the trigger pulses. Sketch and dimension the resulting waveforms. (Let $E_{bb} = 200$, $R = r_p = 10$ kilohms, and $\mu = 20$.)

(a)

(b)

(c)

(d)
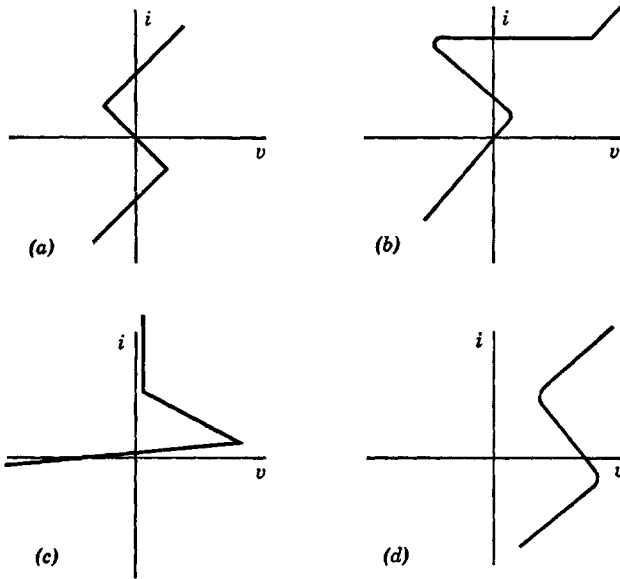
Fig. P9.15

9.19. (a) Which of the v vs. i curves shown in Fig. P9.15 are astable or monostable with parallel capacitance? Why?

(b) Show how the path of operation for the monostable cases always ends at the same point no matter what the capacitor voltage was initially.
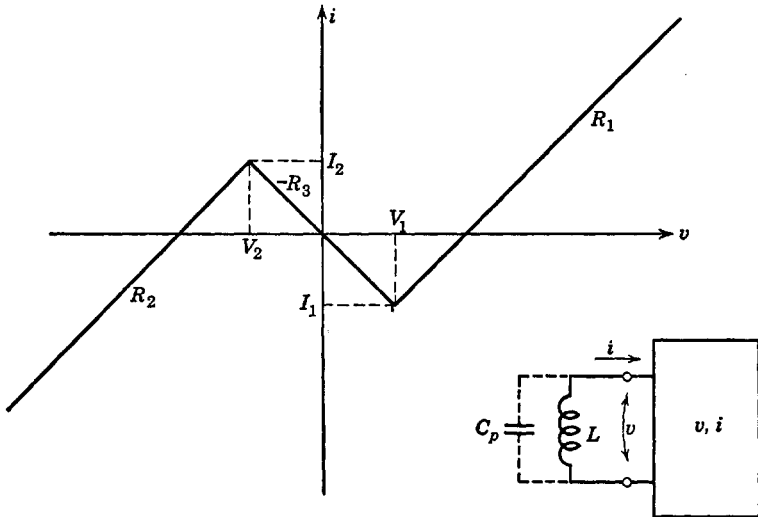
Fig. P9.16

9.20. Apply the results of Article 9.9 to the dual case of a small "parasitic" capacitance in parallel with an inductance in the voltage-controlled case (Fig. P9.16).

9.21. The piecewise-linear representation of an active circuit is shown in Fig. P9.17.
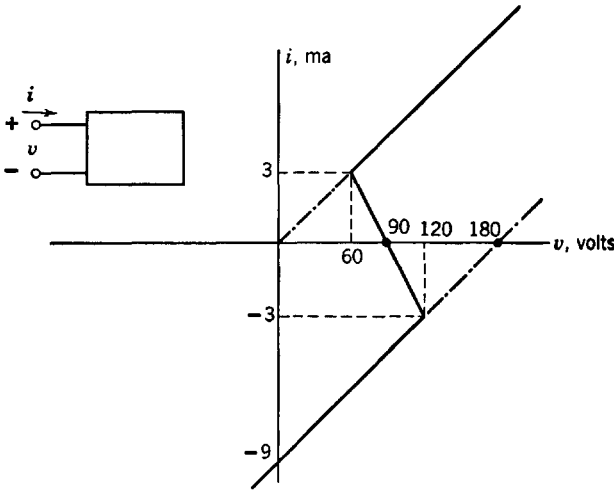
Fig. P9.17

(a) If an inductor with an initial current $I_0 = 6$ ma and an inductance of 1 millihenry is connected at time $t = 0$ to the terminals of the box, sketch and dimension the waveforms of $i(t)$ and $v(t)$. What is the path of operation on the v vs. i plane?

(b) Repeat part (a) for a capacitor of 1 microfarad with zero initial voltage. Again find the path of operation on the v-i plane.
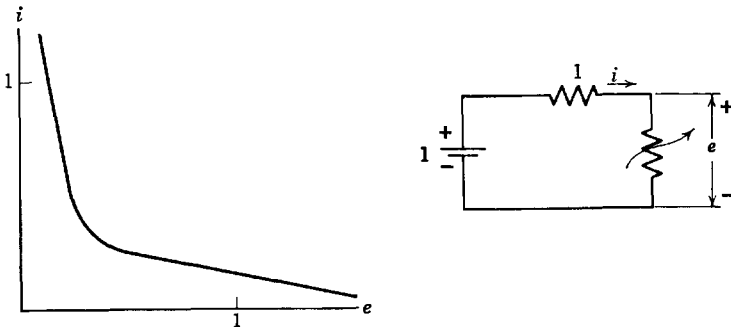
Fig. P9.18

9.22. (a) A nonlinear device having the volt-ampere characteristic shown in Fig. P9.18 is connected in series with a one-volt battery and a one-ohm resistor. Which of the two possible operating points is stable with shunt capacitance across the device terminals? Explain.

(b) The nonlinear device of Fig. P9.18 is replaced by one having the volt-ampere characteristic of Fig. P9.19. Which of the operating points is stable with shunt capacitance across the device terminals? Explain.
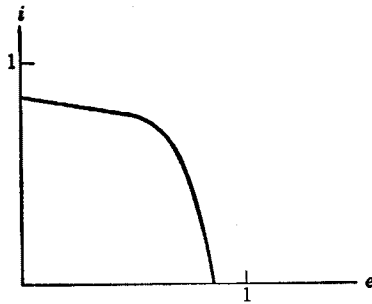
Fig. P9.19

9.23. The triodes in the circuit of Fig. P9.20 can be represented by a piecewise-linear model with $\mu = 20$, $r_p = 10$ kilohms, and $r_g = 1$ kilohm.
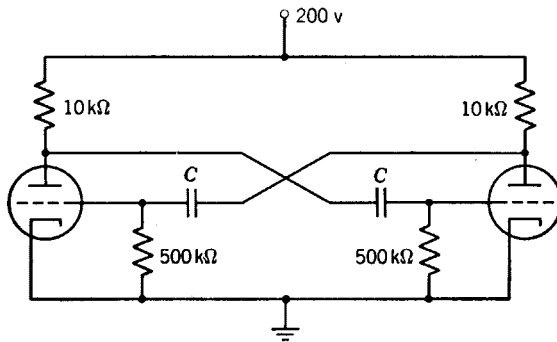
Fig. P9.20

The free-running repetition frequency is to be 1000 cps. Find the value of capacitance required.

9.24. Sketch and dimension the waveforms of pertinent voltages and currents for the transistor multivibrator circuit shown in Fig. P9.21. Assume each junction transistor can be represented by an ideal diode model with $a = 0.98$.

9.25. A multivibrator using a single point-contact transistor is shown in Fig. P9.22 together with the transistor model. Sketch and dimension all pertinent waveforms.

9.26. Calculate the repetition frequency of the free-running multivibrator shown in Fig. P9.23. Assume $r_g = 1$ kilohm for $e_c > 0$. Neglect the effects of stray capacitances.

9.27. Sketch and dimension the waveform of voltage e_0 obtained from the free-running multivibrator shown in Fig. P9.24. Each triode is Type I.
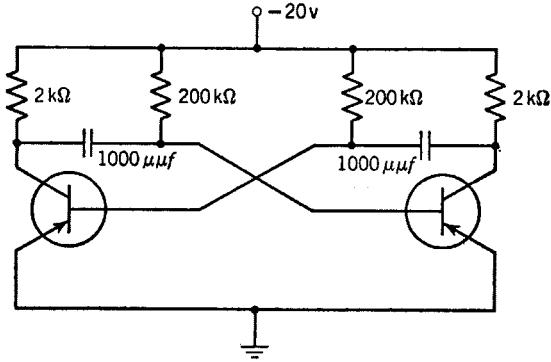
Fig. P9.21
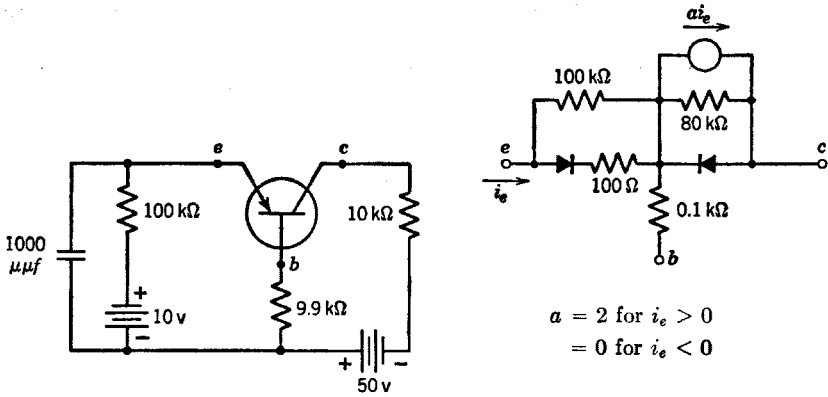
$a = 2$ for $i_e > 0$
$= 0$ for $i_e < 0$
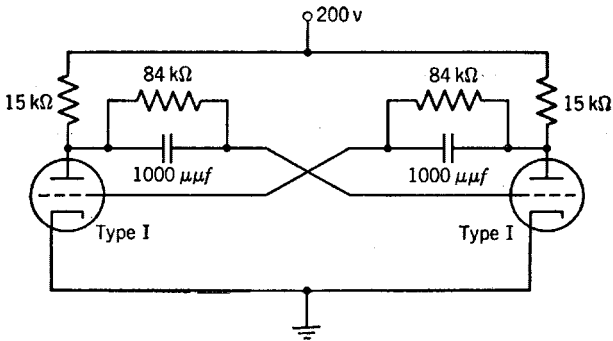
Fig. P9.22
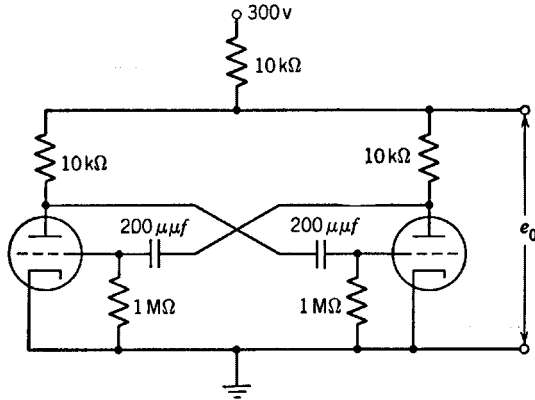
Fig. P9.23

Fig. P9.24

Assume $r_g = 1$ kilohm for $e_c \geqslant 0$ and neglect the effects of tube and wiring capacitances.

9.28. Determine the repetition frequency of the free-running multivibrator shown in Fig. P9.25. Type I triodes are used. The effects of shunt capaci-
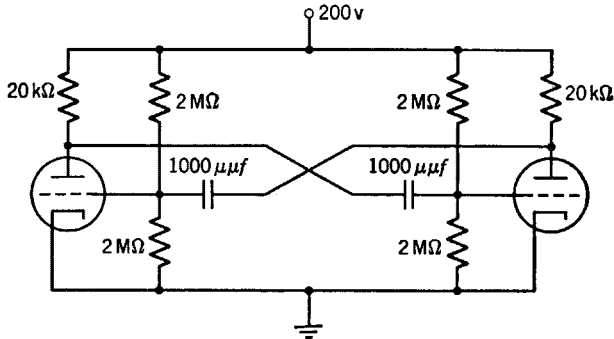
Fig. P9.25

tances and grid-voltage overshoot may be neglected. Assume $r_g = 1$ kilohm for $e_c > 0$.

9.29. The triodes in the circuit of Fig. P9.26 can be represented by a piecewise-linear model with $\mu = 20$, $r_p = 10$ kilohms, and $r_g = 0$. Triodes T_1 and T_2 constitute a monostable multivibrator triggered by $e_1(t)$ applied to triode T_0.

(a) What is the minimum value of input pulse amplitude E_1 that will effect triggering?

(b) What is the maximum value of amplitude E_1 that will result in plate-current cutoff in T_0 as soon as the circuit has been triggered?

(c) What value of C (in $\mu\mu f$) will produce a 100 μsec positive-going pulse at e_2 each time a trigger pulse occurs?

(d) Sketch and dimension $e_2(t)$.
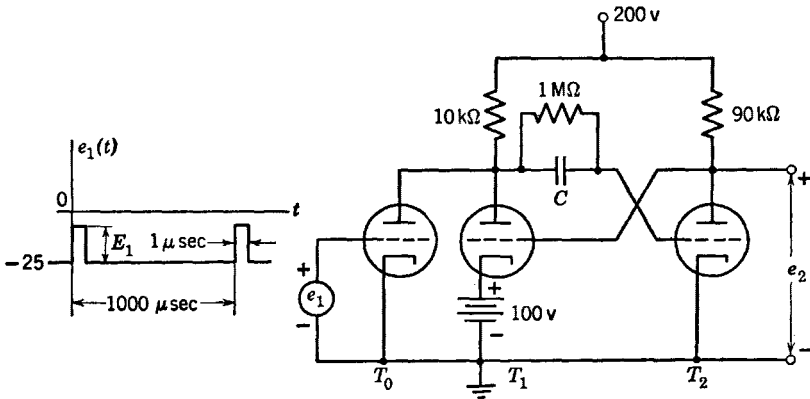
Fig. P9.26

9.30. The monostable multivibrator shown in the diagram of Fig. P9.27 is triggered through diode D_1 by a brief negative voltage pulse. The plate characteristics for the triodes can be represented by piecewise-linear curves with the $e_c = 0$ curve passing through the origin. Let $r_g = 1$ kilohm.
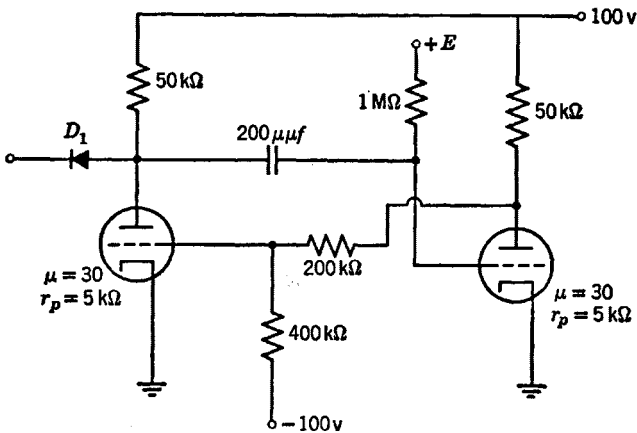
Fig. P9.27

(a) What must be the value of bias potential E to make the duration of the unstable state equal to 100 μsec?

(b) What is the time constant of recovery of the circuit at the end of the unstable state?

9.31. For the free-running multivibrator circuit shown in Fig. P9.28, express the repetition interval δ_r in terms of R_b, R, C, and the triode constants μ and r_p. Assume $r_g = 0$ when grid current flows.

9.32. Type I triodes are used in the monostable circuit shown in Fig. P9.29. The circuit is synchronized by brief pulses having a repetition frequency of
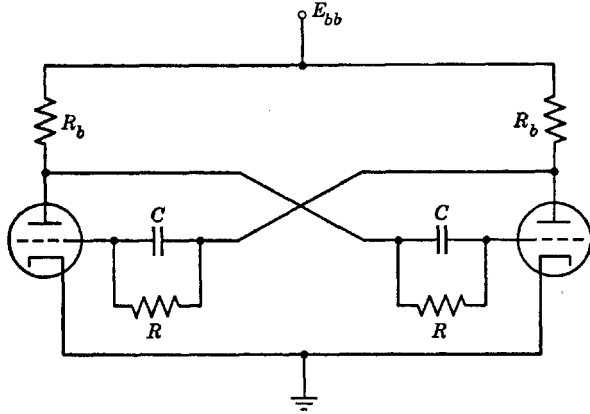
Fig. P9.28
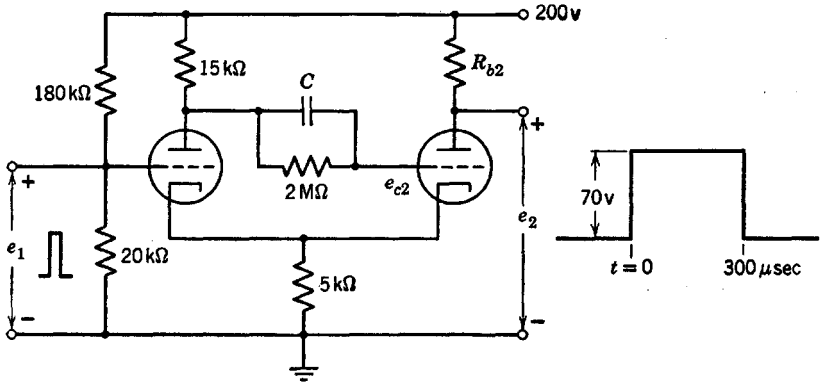
Fig. P9.29
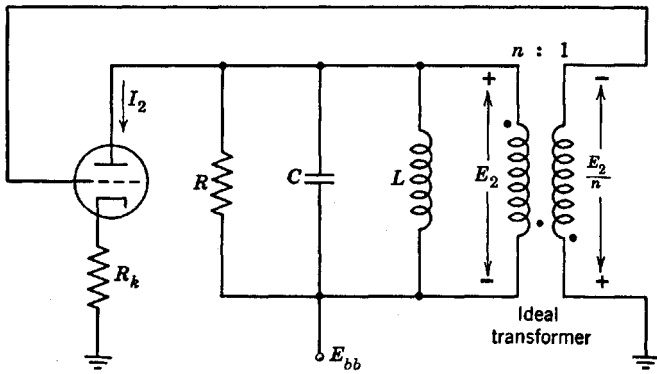
Fig. P9.30

1000 per second. The output pulse resulting from each input pules is required to have a peak-to-peak amplitude of 70 volts and a duration of 300 μsec.

Determine the required values of R_{b2} and C. Assume $e_{c2} = 0$ during conduction of the output triode.

9.33. A tuned-plate oscillator is shown in Fig. P9.30. The transformer is represented by an ideal inductor L and an ideal transformer. The losses of the tuned circuit are represented by R. Find the turns ratio n required for sustained oscillations in the linear region of the tube. What is the frequency of oscillations?

Oscillations in *RLC* Circuits

10.1   Introduction

The *LC* circuit is the basic electrical oscillator. Since the circuit is lossless, the total stored energy remains constant with time. However, the current and voltage oscillate sinusoidally as the stored energy is transferred back and forth from capacitor to inductor. Such ideal oscillations can actually occur at very low temperatures (say 4° Kelvin) where superconductivity exists in certain materials such as lead. However, no power can be supplied to a useful load by the ideal oscillator, because any power output results in a decrease in current and voltage amplitudes.

The linear *RLC* circuit is a more general oscillatory circuit than the *LC* circuit. If all elements are linear, continuous oscillations cannot be sustained. If the resistance is positive, any energy initially stored in the inductance or capacitance is eventually dissipated, hence currents and voltages approach zero. If the resistance is negative, currents and voltages increase without limit. With nonlinear resistance, positive during part of the cycle and negative during part of the cycle, continuous oscillations can occur. The amplitudes of currents and voltages stabilize at the level for which energy given to the *LC* circuit by the negative resistance each cycle equals energy dissipated in the positive resistance.

The relaxation oscillations described in Chapter 9 are special cases in

which only one energy-storage element need be considered at any one time. For this approximation, the circuit operation can be represented by first-order equations. When sinusoidal (or nearly sinusoidal) waveforms are being generated, two energy-storage elements are important throughout the cycle. The equilibrium conditions for the circuit are therefore described by a second-order differential equation. For actual circuits such equations will be nonlinear. The use of a piecewise-linear model for a nonlinear resistance is very helpful in simplifying the analysis.

$$v = E + Ri$$
$$= e_x + e_L$$
$$i = -C \frac{de_x}{dt}$$
$$e_L = -L \frac{di}{dt}$$
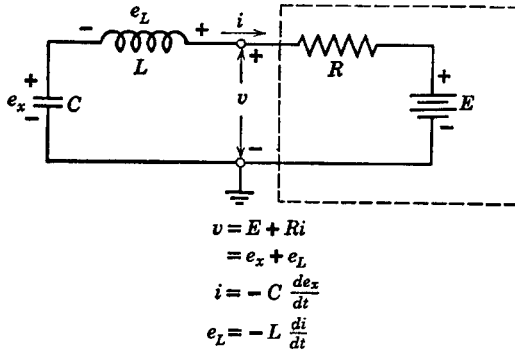
Fig. 10.1. Linear oscillatory circuit.

10.2 Series Oscillatory Circuit

The series circuit shown in Fig. 10.1 is a general linear oscillatory circuit. Since any one state of a piecewise-linear resistive circuit reduces to a resistance and a source, the behavior of this circuit is an aid to understanding the generation of continuous "sinusoidal" oscillations.

From the terminal relations given in Fig. 10.1 we can write the equilibrium equation for the circuit as follows:

$$E + Ri + L \frac{di}{dt} + \frac{1}{C} \int i \, dt = 0 \qquad (10.1)$$

The homogeneous equation, obtained by differentiation, is

$$\frac{d^2 i}{dt^2} + \frac{R}{L} \frac{di}{dt} + \frac{i}{LC} = 0 \qquad (10.2)$$

From this equation, the characteristic equation is

$$s^2 + \frac{R}{L} s + \frac{1}{LC} = 0 \qquad (10.3)$$

which has the roots

$$s_1 = -\frac{R}{2L} - \sqrt{\left(\frac{R}{2L}\right)^2 - \frac{1}{LC}} \qquad (10.4)$$

$$s_2 = -\frac{R}{2L} + \sqrt{\left(\frac{R}{2L}\right)^2 - \frac{1}{LC}} \qquad (10.5)$$

The solution of Eq. 10.2 has the form

$$i = \frac{K}{s_2 - s_1} \left(\epsilon^{s_1 t} - \epsilon^{s_2 t}\right) \qquad (10.6)$$

for the initial condition $i = 0$ at $t = 0$. Any of the voltages can be obtained once the current is known.

The character of the current and voltage waveforms depends on the values of the roots s_1 and s_2.* For comparison with the results obtained for relaxation oscillators in Chapter 9 we shall consider the waveforms $i(t)$ and $e_x(t)$. In addition, a plot of i vs. e_x for corresponding instants of time yields a compact representation of oscillatory circuit behavior. If the waveforms $i(t)$ and $e_x(t)$ have been determined, the phase-plane† trajectory (i vs. e_x) can be plotted point-by-point for successive values of time t. The expressions for $i(t)$ and $e_x(t)$ are thus seen to be a parametric-equation representation of $i(e_x)$.

The functional relation between i and e_x can also be expressed as a differential equation written directly from the equations given in Fig. 10.1.

$$\frac{e_L}{i} = \frac{E + Ri - e_x}{i} = \frac{L}{C}\frac{di}{de_x} \qquad (10.7)$$

Thus the differential equation relating i and e_x is a first-order equation, whereas the equilibrium equation is second-order. In general, the equation for the phase trajectory is one order lower than the equilibrium equation.

Taking the undamped circuit ($R = 0$) as a simple example, we have

$$(E - e_x)\, de_x = \frac{L}{C}\, i\, di \qquad (10.8)$$

If we let $E - e_x = e$, then $de = -de_x$ so that (Eq. 10.7) reduces to

$$\frac{L}{C}\, i\, di + e\, de = 0 \qquad (10.9)$$

---

*Note that the series RLC circuit under discussion is the dual of the parallel RLC circuit described in Article 8.32.

† A plane in which the axes represent a variable and its time derivative is called a phase-plane.

$$\omega_0 = \frac{1}{\sqrt{LC}}, \ A = I\sqrt{\frac{L}{C}}$$
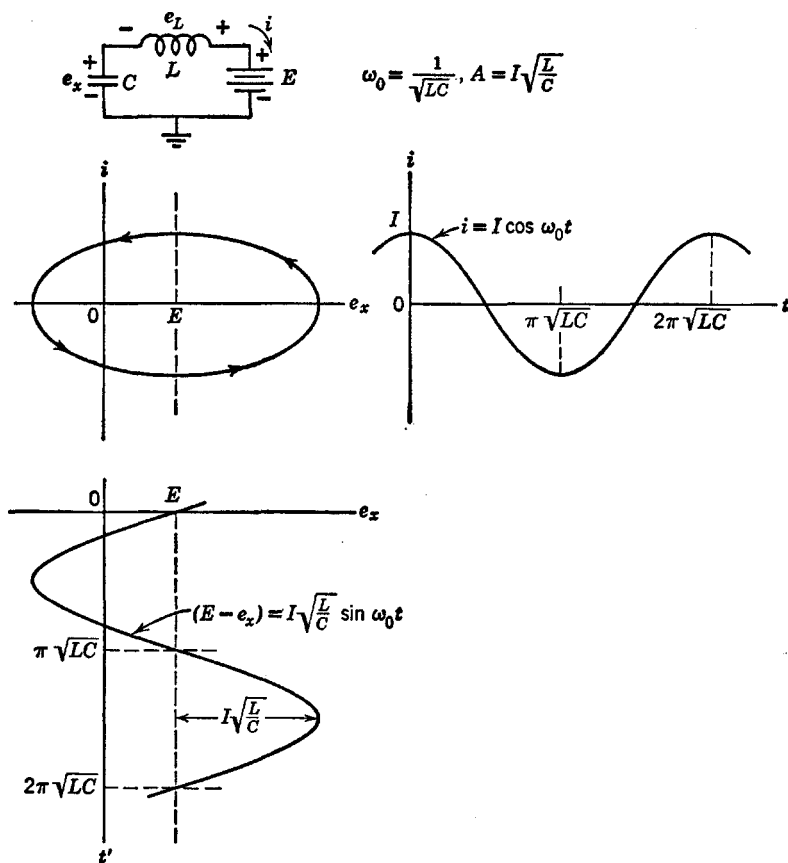
Fig. 10.2.   The undamped oscillatory circuit.

Integrating this equation and substituting for e, we obtain

$$\frac{L}{C} i^2 + (E - e_x)^2 = A^2 \tag{10.10}$$

The corresponding undamped differential equation for the current is

$$\frac{d^2 i}{dt^2} + \frac{i}{LC} = 0 \tag{10.11}$$

which yields

$$i = I \cos (t/\sqrt{LC}) \tag{10.12}$$

and

$$(E - e_x) = I \sqrt{L/C} \sin (t/\sqrt{LC}) \tag{10.13}$$

Comparing Eq. 10.10 with Eqs. 10.12 and 10.13 we can establish $A = I\sqrt{L/C}$. The waveforms of current and voltage as well as the elliptical phase-plane trajectory are shown in Fig. 10.2 for one specific value of A. For different values of A, corresponding to different values of stored energy in the system, the trajectories are concentric ellipses.

10.3 Scale Factors and Normalization

In any analysis, normalization of variables is usually desirable because the results are somewhat more general. When graphical methods are to be used, as in oscillator analysis, either normalization or a suitable choice of scale factors is virtually a necessity. For example, in a circuit like that of Fig. 10.2, the ratio L/C may be of the order of 10^6 or more, so that in terms of volts and amperes the major and minor axes of the ellipse may have a ratio of 10^3 or more. Thus, comparable scales for volts and milliamperes yield reasonable plots. Actually, the most convenient form for the elliptical trajectory is a circle, which can be obtained

TABLE 10.1 NORMALIZATION OF OSCILLATOR EQUATIONS

Let $\quad i' = \sqrt{L}\,i, \quad v' = \sqrt{C}\,v, \text{ and } t' = t/\sqrt{LC}$

Then $\quad e_L' = \sqrt{C}\,e_L, \quad e_x' = \sqrt{C}\,e_x, \quad E' = \sqrt{C}\,E, \text{ etc.}$

and $\quad R' = R/\sqrt{L/C}$

and $\quad \omega_0' = \omega_0\sqrt{LC} = 1$

Original Equation	Normalized Equation
$i = -C\dfrac{de_x}{dt}$	$i' = -\dfrac{de_x'}{dt'}$
$e_L = -L\dfrac{di}{dt}$	$e_L' = -\dfrac{di'}{dt'}$
$\dfrac{d^2i}{dt^2} + \dfrac{R}{L}\dfrac{di}{dt} + \dfrac{i}{LC} = 0$	$\dfrac{d^2i'}{dt'^2} + R'\dfrac{di'}{dt'} + i' = 0$
$s_1 = -\dfrac{R}{2L} - \sqrt{\left(\dfrac{R}{2L}\right)^2 - \dfrac{1}{LC}}$	$s_1' = -\dfrac{R'}{2} - \sqrt{\left(\dfrac{R'}{2}\right)^2 - 1}$
$s_2 = -\dfrac{R}{2L} + \sqrt{\left(\dfrac{R}{2L}\right)^2 - \dfrac{1}{LC}}$	$s_2' = -\dfrac{R'}{2} + \sqrt{\left(\dfrac{R'}{2}\right)^2 - 1}$
$\dfrac{v - e_x}{i} = \dfrac{L}{C}\dfrac{di}{de_x}$	$\dfrac{v' - e_x'}{i'} = \dfrac{di'}{de_x'}$

by expanding the current scale by $\sqrt{L/C}$ or (more symmetrically) by multiplying current by $\sqrt{L}$ and voltage by $\sqrt{C}$. It is also convenient to express time in terms of $\sqrt{LC}$. The effects of normalization on all of the quantities involved in the more general oscillatory circuit of Fig. 10.1 are given in Table 10.1. Note that the normalized variables i' and e' actually have different dimensions than voltage and current.

The normalization in Table 10.1 changes the dimensions of the variables used to describe the oscillations. The same form of equations can be obtained by setting L and C equal to unity, in which case the variables are still current and voltage. Since it is more convenient to think in terms of current and voltage than in terms of the square root of joules ($\sqrt{L}\, i$ or $\sqrt{C}\, e$) we shall assume L and C equal to unity in the following discussions. We know from the normalization procedure that there is no loss of generality in this simplification, since we can always replace current and voltage by the normalized variables in any of the equations.

10.4   Oscillations in Linear Circuits

In the simplified *RLC* circuit shown in Fig. 10.3, the only adjustable parameter is the damping resistance R. We shall select a number of resistance values to illustrate the distinctly different character of current
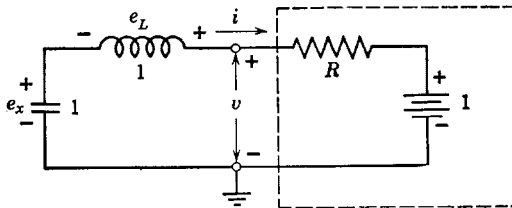
Fig. 10.3. Simplified linear *RLC* circuit.

and voltage waveforms for various amounts of damping. In addition, we shall plot the trajectories in the phase plane (i vs. e_x), since these curves are generally more useful than the waveforms when we examine nonlinear or piecewise-linear oscillatory circuits.

For the circuit of Fig. 10.3, the homogeneous differential equation is

$$\frac{d^2i}{dt^2} + R\,\frac{di}{dt} + i = 0 \qquad (10.14)$$

The roots of the characteristic equation are

$$s_1 = -\frac{R}{2} - \sqrt{(R/2)^2 - 1} \qquad (10.15)$$

$$s_2 = -\frac{R}{2} + \sqrt{(R/2)^2 - 1} \qquad (10.16)$$

Case I: Overdamped ($R > 2$, s_1 and s_2 both **negative real**)
The solution for the current is

$$i = \frac{1}{s_2 - s_1} \left(\epsilon^{s_1 t} - \epsilon^{s_2 t} \right) \qquad (10.17)$$

and the capacitor voltage is

$$e_x = \frac{1}{s_2 - s_1} \left[\frac{1}{s_1} (1 - \epsilon^{s_1 t}) - \frac{1}{s_2} (1 - \epsilon^{s_2 t}) \right] \qquad (10.18)$$
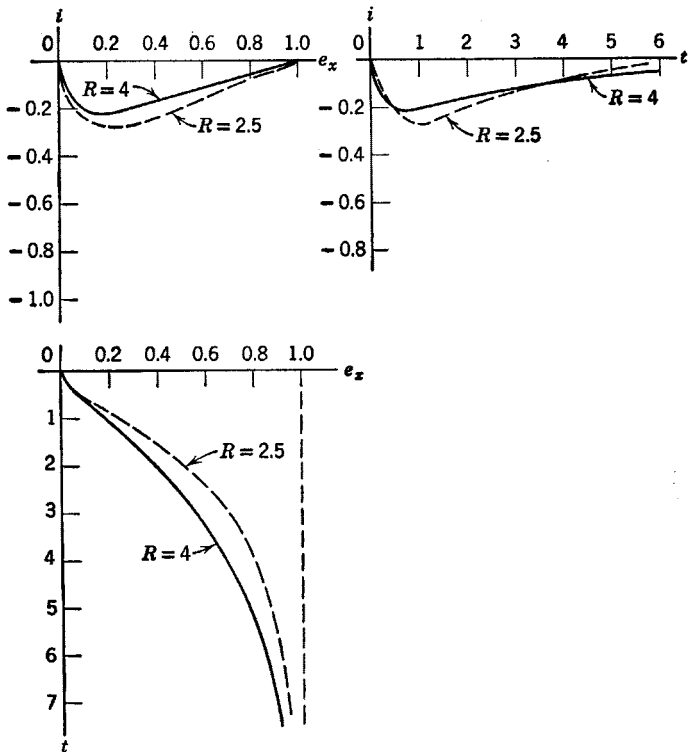
Fig. 10.4. Normalized response of overdamped linear circuit ($R > 2$).

These waveforms are plotted in Fig. 10.4 together with i vs. e_x for $R = 4$ and $R = 2.5$. The plot of i vs. e_x is most readily obtained by taking corresponding values of i and e_x from the waveforms.

Case II: Critically Damped ($R = 2$, $s_1 = s_2 = -1$)

Since in this case s_1 equals s_2, the expression for current given in Eq. 10.17 is indeterminate. Evaluating the limit for the indeterminacy $(0/0)$ as R approaches the critical value, we have

$$i = -t\,\epsilon^{-t} \tag{10.19}$$

and

$$e_x = 1 - \epsilon^{-t}(1 + t) \tag{10.20}$$

Critical damping yields the most rapid return to zero possible for the current waveform without an overshoot or reversal.
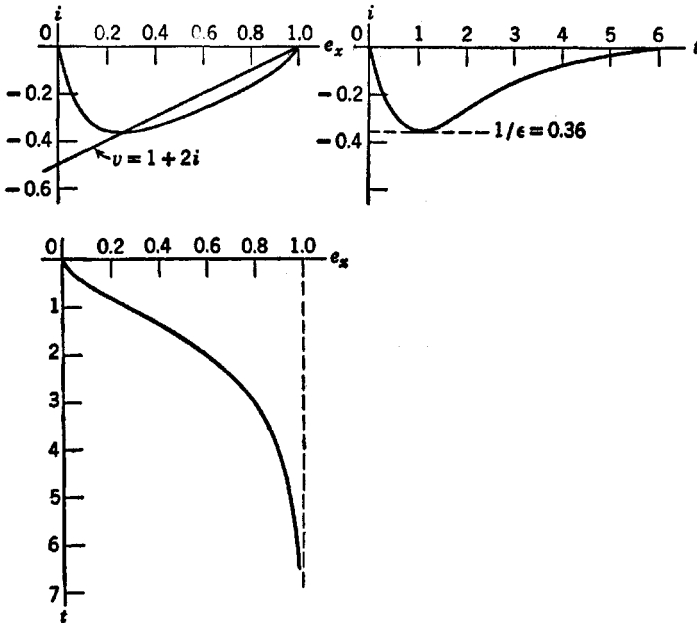
Fig. 10.5. Normalized response of critically-damped linear circuit ($R = 2$).

The waveforms $i(t)$ and $e_x(t)$ are shown in Fig. 10.5 along with the plot of i vs. e_x. The resistive relation $v = 1 + 2i$ is also shown on the current versus voltage plane.

Case III: Underdamped ($0 < R < 2$, s_1 and s_2 are complex conjugates with negative real parts)

For values of R between zero and two, the waveforms are oscillatory. Since s_1 and s_2 are complex, it is convenient to express them as

$$s_1 = -\alpha - j\beta \tag{10.21}$$

$$s_2 = -\alpha + j\beta \tag{10.22}$$
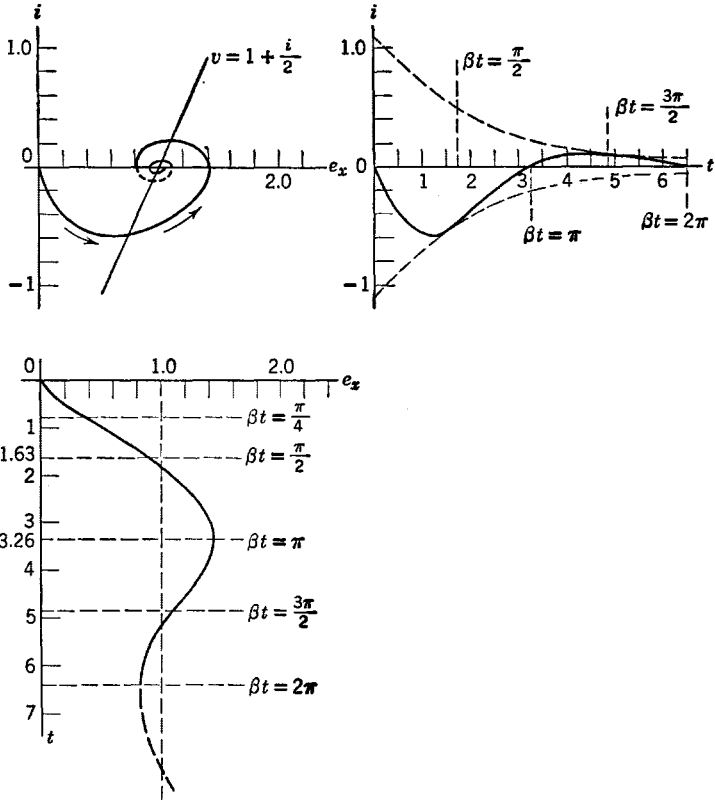
Fig. 10.6. Normalized response of underdamped linear circuit $(R = \frac{1}{2})$.

where $\alpha = R/2$ and $\beta = \sqrt{1 - (R/2)^2}$. In terms of α and β the current is

$$i = \epsilon^{-\alpha t} \frac{\epsilon^{-j\beta t} - \epsilon^{+j\beta t}}{2j\beta} \qquad (10.23)$$

$$= - \frac{\epsilon^{-\alpha t}}{\beta} \sin \beta t \qquad (10.24)$$

The expression for capacitor voltage is

$$e_x = 1 - \frac{\epsilon^{-\alpha t}}{\beta} [\alpha \sin \beta t + \beta \cos \beta t] \qquad (10.25)$$

The curves in Fig. 10.6 are drawn for $R = \frac{1}{2}$, whereas those in Fig. 10.7 are drawn for $R = \frac{1}{5}$. For the latter we have $\alpha = 0.1$ and $\beta = 0.995$ (closely approaching the undamped case, $\alpha = 0, \beta = 1$).

$v = 1 + i/5$

$\epsilon^{-0.1t}$

$i \approx - \epsilon^{-0.1t} \sin t$

$e_x \approx 1 - \epsilon^{-0.1t} [0.1 \sin t + \cos t]$
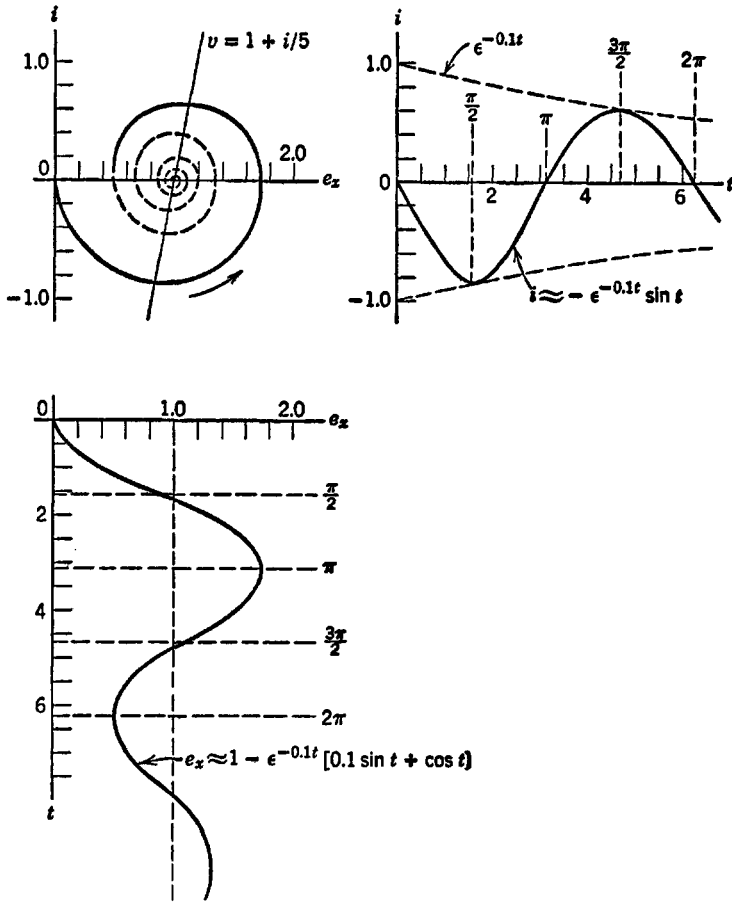
Fig. 10.7. Normalized response of underdamped linear circuit ($R = \frac{1}{5}$).

Case IV: Undamped ($R = 0$, s_1 and s_2 imaginary)

Although we have already discussed the undamped case (see Fig. 10.2) it is desirable to include the normalized representation here for the sake of completeness. Furthermore, it is instructive to obtain the expressions for current and voltage as limiting forms of the underdamped case as R approaches zero. We see from Eqs. 10.21 and 10.22 that α approaches zero and β approaches unity. Therefore, we have

$$i = -\sin t \qquad (10.26)$$

and

$$e_x = 1 - \cos t \qquad (10.27)$$

These waveforms and the circular locus of i vs. e_x are shown in Fig. 10.8.
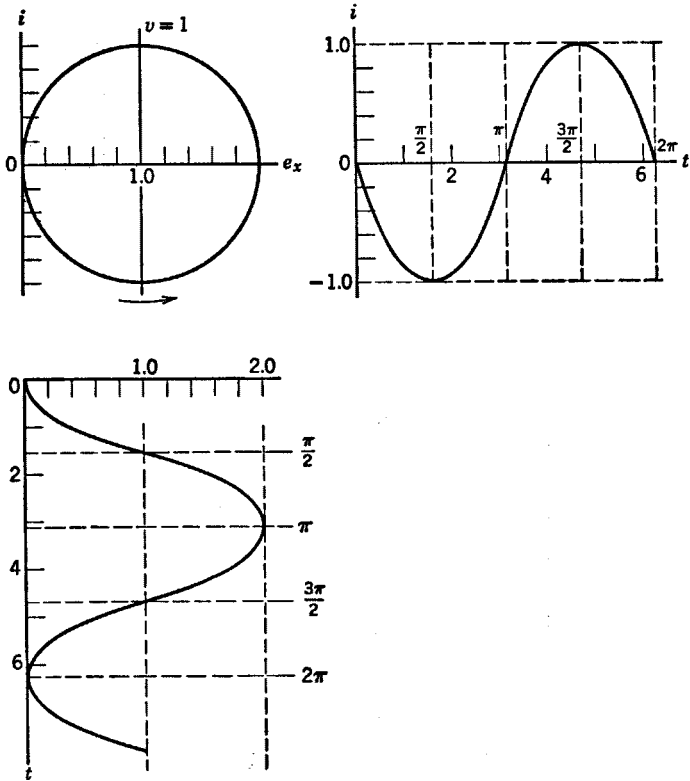
Case V: Negative Underdamped ($0 > R > -2$, s_1 and s_2 complex conjugates with positive real parts)

A negative damping resistance reverses the sign of α but does not change β. Thus, just as in Case III, we have

$$s_1 = -\alpha - j\beta; \qquad s_2 = -\alpha + j\beta \qquad (10.28)$$

but now

$$\alpha < 0 \qquad (10.29)$$

The general equations for current and capacitor voltage are the same as Eqs. 10.23–10.25. In Fig. 10.9 we have chosen $R = -\frac{1}{5}$ for comparison with the corresponding positive underdamped circuit (Case III) illustrated in Fig. 10.7. The exponential envelope for the waveforms of

current and voltage describing the negative-underdamped circuit are like negative extensions of those for the positive-underdamped circuit because a reversal of the sign of α or t has the same effect. However, in both cases the sine wave begins with the negative half cycle.
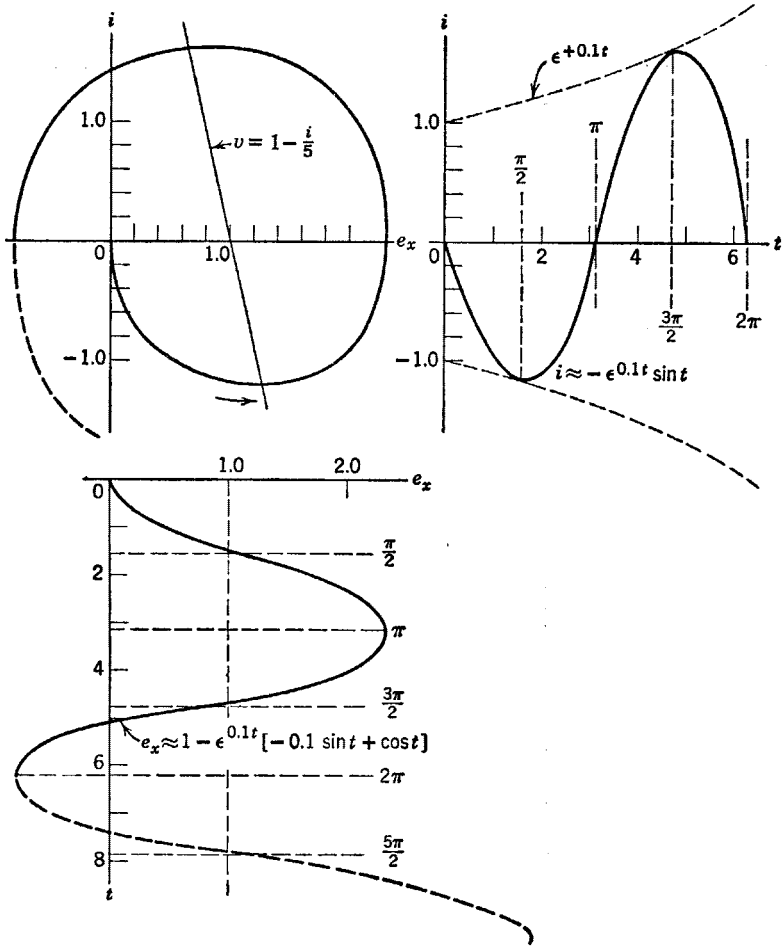
Fig. 10.9. Normalized response of linear circuit with negative damping ($R = -\frac{1}{5}$).

Case VI: Negative Critically Damped ($R = -2$, $s_1 = s_2 = +1$)

The current and capacitor voltage in this case tend to increase indefinitely without ever being oscillatory. The pertinent equations are

$$i = -t\epsilon^t \qquad (10.30)$$

and
$$e_x = 1 + \epsilon^t(t - 1) \tag{10.31}$$

Case VII: Negative Overdamped ($R < -2$, s_1 and s_2 both positive)

As resistance becomes more negative, the current and capacitor voltage merely tend toward infinite values more rapidly than in the previous case. The expressions given in Eqs. 10.17 and 10.18 apply directly, but we note that the values of s_1 and s_2 are now positive.
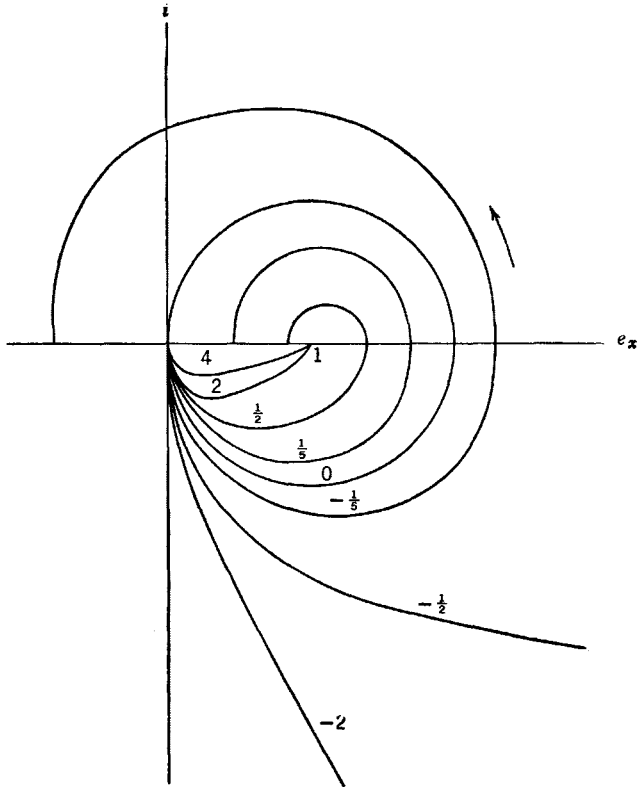
Fig. 10.10. Normalized operating paths for various values of R.

Summary. For ease of comparison, the trajectories of current versus capacitor voltage for both positive and negative damping resistances are shown in Fig. 10.10. For the initial conditions used here ($i = e_x = 0$ at $t = 0$) all trajectories start at the origin. All values of positive damping lead to the final value $i = 0$, $e_x = 1$. With no damping the trajectory is a circle (continuous sinusoidal waveforms) while for negative damping the final values are infinite.

10.5   Some Properties of Phase-Plane Trajectories

In Article 10.4, phase-plane trajectories for a linear series RLC circuit are obtained from the waveforms $i(t)$ and $e_x(t)$ by eliminating the time t. In nonlinear circuit problems, phase-plane trajectories are often easier to determine than waveforms. The phase-plane trajectories determine the character of the solution and can be used to plot waveforms if these are desired.

As an aid to determining trajectories in more complicated systems, let us deduce some basic properties of phase-plane plots. From the differential equation for the trajectories,

$$\frac{v - e_x}{i} = \frac{di}{de_x} \tag{10.32}$$

it is apparent that the slope di/de_x is everywhere unique except at the point $e_x = v$, $i = 0$. At this point the slope is indeterminate since $0/0$ can have any value. The slope of any given trajectory in the vicinity of such a "singular" point must be found by a limiting process as the trajectory approaches or recedes from the point.

From the uniqueness of slope di/de_x at all but singular points we conclude the following:

A. A specific trajectory can never cross itself at a nonsingular point in the finite plane.

B. Two different trajectories cannot cross each other at a nonsingular point in the finite plane.

These points are illustrated by Fig. 10.11(*a*), where three trajectories corresponding to different initial conditions are drawn for $R = \frac{1}{5}$. Different starting points on the same trajectory (like points *B* and *C*) correspond only to a shift of the time reference, but an initial point not on any trajectory shown, must result in a separate curve.

For the undamped case illustrated in Fig. 10.11(*b*), different initial conditions may produce a change in the phase of current and voltage waveforms, or a change of amplitude or both. For initial conditions e_{x0} and i_0 the amplitude of e_x or i for the normalized case will be $\sqrt{(e_{x0} - 1)^2 + i_0{}^2}$, since the trajectories are circles centered at $e_x = 1$, $i = 0$. Initial points O, B and C on the same trajectory merely represent different time origins but the same total energy in the system. The trajectory starting at point A yields waveforms having the same phase relations but smaller amplitudes than those for the circle through point C.
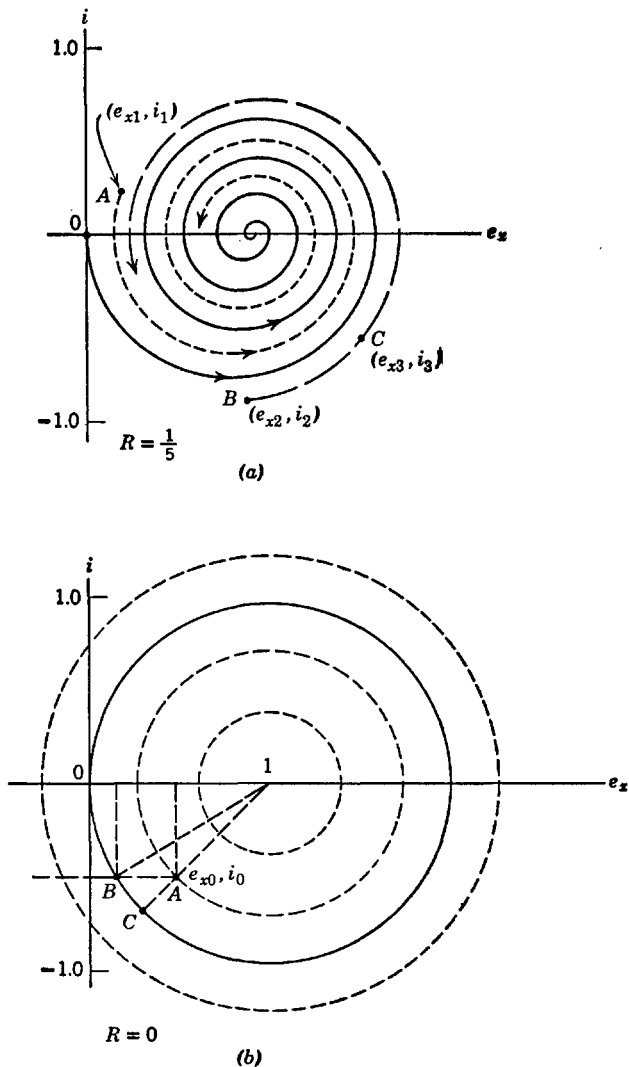
10.6   Circular-Arc Construction of Trajectories

The circular phase-plane trajectories of the undamped circuit [Fig. 10.11(b)] lead to an approximate method of constructing trajectories for damped circuits using only a compass.* Suppose we make a stepwise

* R. D. Thornton (M.I.T. *Research Laboratory of Electronics, Quarterly Progress Report*, July 1955, pp. 95–100).

approximation to a resistive curve as indicated in Fig. 10.12(a). This amounts to replacing the resistance by a number of voltage sources, each of which applies over a small range of current. The circuit is therefore

$$v = E_j$$
for $-\ .2 < i < +.2$ $v = 1.0$
    $-\ .6 < i < -.2$ $v = .92$
    $-1.0 < i < -.6$ $v = .84$
etc.

(a)   (b)

(c)

(d)

c b a
d e f g   All centers
j i h   actually on
k l   x axis
n m
o
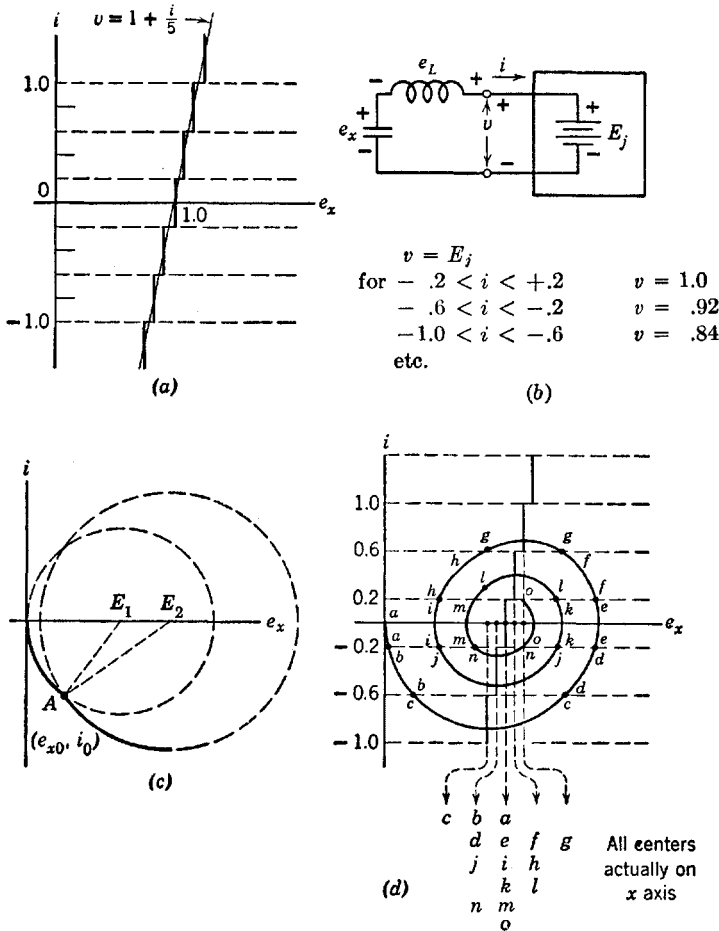
Fig. 10.12. Circular-arc approximation to trajectory resulting from stepwise approximation to resistance.

undamped as indicated by the model shown in (b) and the approximate trajectory will consist of segments of circular arcs.

At a transition from one range of currents to the next, the step change in voltage amounts to nothing more than a corresponding change in the circle center and radius. Compare the change from E_1 to E_2 in Fig. 10.12(c) with the change from initial point A to B in Fig. 10.11(b).

The step change of voltage appears across the inductance which holds current i constant at the instant of change.

The circular-arc method is indicated in Fig. 10.12(d). Beginning with $e_x = i = 0$, the centers used and the extremities of the corresponding circular arcs have been given the same letter designations. Following the letters in sequence leads to the complete trajectory for $R = \frac{1}{5}$. As the trajectory re-enters a current interval that has been encountered previously, the center used before is used again. Thus, if the arcs are

$$v = 1 + \frac{i}{5}$$

$i = I_j$

for $1.0 < v < 1.08,$ $i = 0.2$

$1.08 < v < 1.16,$ $i = 0.6$

etc.

$0.92 < v < 1.0,$ $i = -.2$

etc.

(b)

(a)
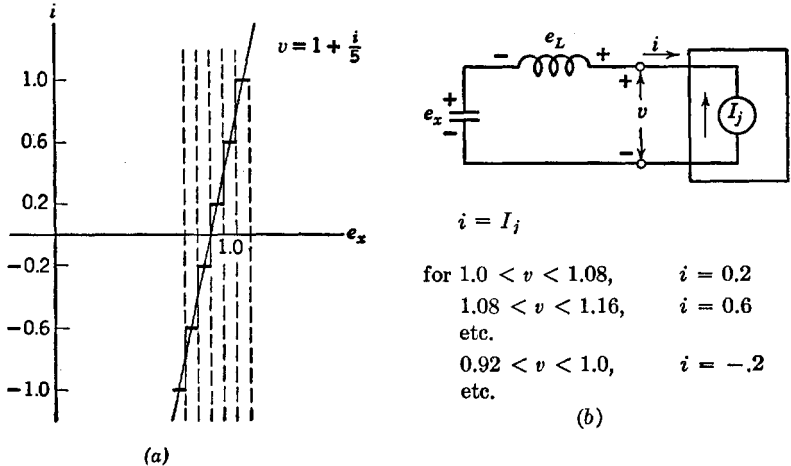
Fig. 10.13. Current-source model for stepwise-resistance approximation.

labelled in sequence, the centers will each have several designations for an oscillatory case such as the one illustrated. Comparing the complete trajectory in (d) with that in Fig. 10.7, we see that a relatively coarse stepwise approximation to the resistance yields a fairly good approximation to the trajectory until the size of the spiral becomes comparable to the size of the steps. Since the steps need not be uniform, more can be used where needed to obtain greater accuracy.

It is apparent that an alternate representation for a stepwise approximation to a resistance is a series of current sources, each of which applies over a small range of voltage. Use of a current-source model like that shown in Fig. 10.13 is not appropriate for the series circuit, since i and e_x will be independent of inductance L. Connecting a circuit element in series with a current source has no effect on the terminal current. With the current source model, the voltage across the inductance will be zero while the current is constant and will have an impulse discontinuity as each step in current occurs. This is obviously not a good approximation

to the actual current waveform in a series *RLC* circuit. The current-source model is the appropriate one to use for a parallel *LC* circuit connected to a stepwise resistance approximation.

10.7 Oscillator Limit Cycles by Circular-Arc Method

With a current-controlled negative resistance connected to a series *LC* circuit, continuous oscillations can be obtained. When repetitive waveforms are thus generated, the corresponding phase-plane trajectory must also repeat itself. The resulting closed path in the phase plane is called a limit cycle. If specific trajectories approach such a limit cycle from both sides, the limit cycle is stable. The free-running relaxation oscillators described in Chapter 9 exhibited closed trajectories (limit cycles) in the phase plane.

The physical argument for the existence of a limit cycle is based on the fact that negative resistance acts as a power source and causes growing transients while positive resistance dissipates power and causes decaying transients. Equilibrium occurs when the amplitude of the oscillations reaches the value for which the energy dissipated during each cycle just equals the energy given to the tuned circuit each cycle. Although simple in physical concept, an analytical solution of this problem is difficult (even for a piecewise-linear negative resistance). However, various techniques permit rapid construction of limit cycles. The circular-arc method is particularly convenient for comparing oscillatory systems. The stepwise approximation can be applied directly to a nonlinear curve or to a piecewise-linear curve. For ease of comparison with results already obtained in Chapter 9, we shall assume a symmetrical current-controlled negative resistance.

The character of the waveforms and phase-plane trajectories of an oscillator circuit depends on the positive and negative damping resistances. Furthermore, since a linear negative resistance causes a growing transient we know that a limit cycle can only exist somewhere outside of the break points, so the trajectory spends some time in the positive-damped regions (states I and II) as well as in the negative-damped region (state III). As a convenience, and without loss of generality, the break points are set at $i = \pm 1$ for all of the following analyses. The construction of the limit cycle proceeds in the same manner as the construction for the linear circuits considered in the preceding article.

The limit cycles shown in Fig. 10.14 correspond to a negative resistance $R_{III} = -\frac{1}{5}$ and positive resistances ranging over the values $\infty, 4, 2$, and $\frac{1}{2}$. This sequence of limit cycles goes from a relaxation oscillation in (*a*)

to a nearly sinusoidal oscillation in (d). Thus with a specific negative resistance, the value of positive resistance alone sets the mode of operation. If states I and II are heavily overdamped we tend to have relaxation oscillations, whereas the underdamped case yields nearly sine
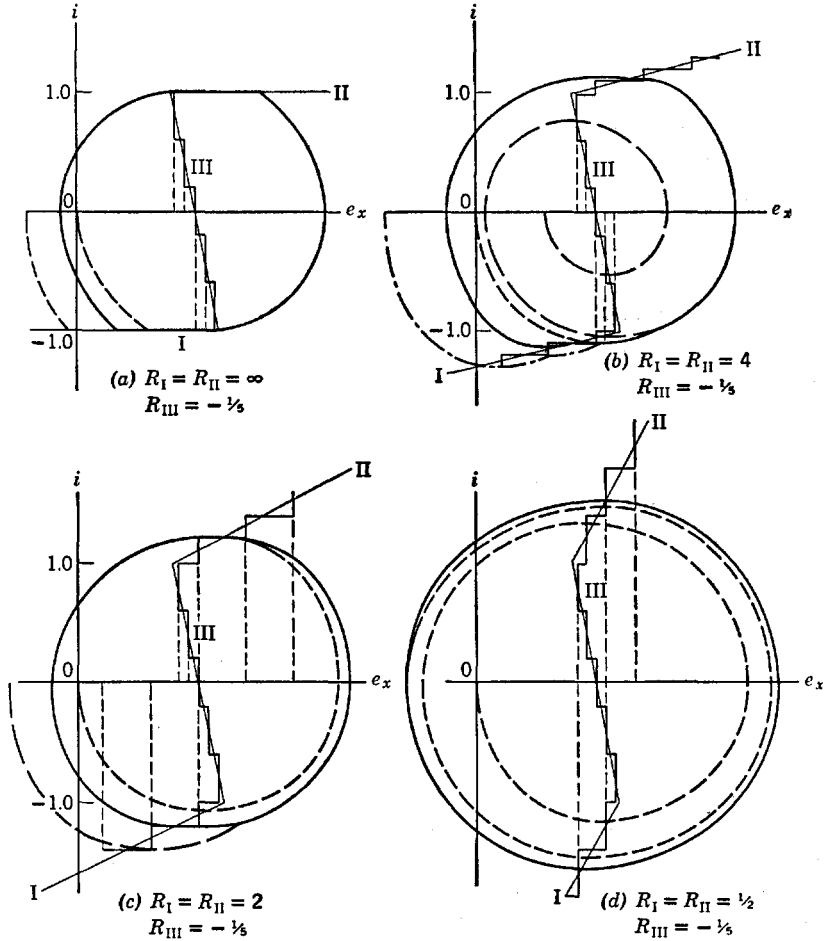
Fig. 10.14. Limit cycles by circular-arc method for $R_{III} = -\frac{1}{5}$.

waves. Note that as the positive damping is decreased the size of the limit cycle increases, as might be expected.

For small damping resistances (both positive and negative) the curve $r = f(i)$ approaches the vertical line that corresponds to no damping.
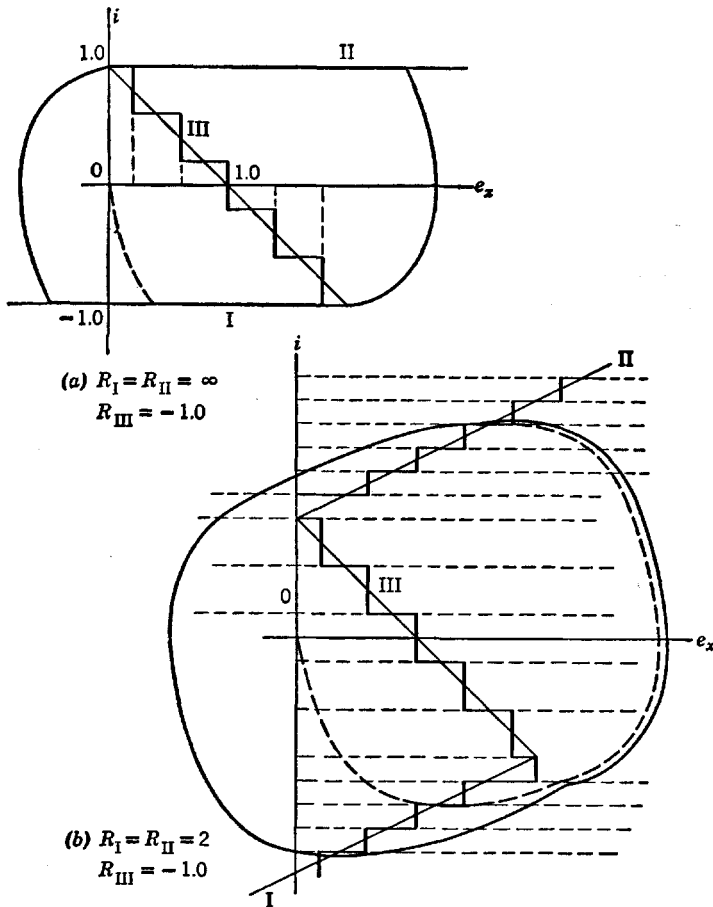
Fig. 10.15. Limit cycles by circular-arc method for $R_{III} = -1.0$.

Note how nearly circular the limit cycle is in Fig. $10.14(d)$. The corresponding waveforms will be nearly sinusoidal. However for such small values of damping, the oscillation amplitude is sensitive to changes in load.

The limit cycles constructed in Fig. 10.15 have a negative resistance $R_{III} = -1$ (a stronger power source than the $R_{III} = -\frac{1}{5}$ of Fig. 10.14). The positive damping is infinite in (a) and critical in (b). Note the tendency of both limit cycles to "elongate." The waveforms will be relaxation oscillations in both cases. If a positive underdamped resistance is used here [say $\frac{1}{2}$ as in Fig. $10.14(d)$], the limit cycle will be very large and will again tend to be nearly circular.

10.8   The Method of Isoclines

An isocline is a line in the phase plane along which the slope of a trajectory must have a fixed value. Any trajectory must intersect the isocline with a unique slope anywhere along the line except at a singular point. Let us apply the method of isoclines to the series RLC circuit, first with linear and then piecewise linear resistance.

The differential equation for RLC circuit trajectories shows slopes explicitly. Thus we merely let di/de_x take on a succession of constant values

$$\frac{v - e_x}{i} = \frac{di}{de_x} = k \qquad (10.33)$$

The equation for an isocline in this case is

$$v - e_x = k\,i \qquad (10.34)$$

We note immediately that for k equal to zero we have e_x equal to v, whereas for infinite k we must have i equal to zero. Other convenient values for k, such as plus and minus one, two, three, etc., are then selected. As indicated in Fig. 10.16 the slope for each isocline is indicated at frequent intervals along the isocline by a short line called a "director." When a number of these have been drawn, a trajectory can be sketched relatively easily beginning from any initial point.

Note that the directors do not specify the direction of rotation of the trajectory for positive time. The direction is readily obtained from the terminal relations such as those specified by Fig. 10.3. In this case, $e_L = -di/dt$ and $i = -de_x/dt$ and thus the rotation is counterclockwise with increasing time.

The values of damping resistance and initial conditions in Fig. 10.16(a) and (b) correspond to those used in Figs. 10.7 and 10.8, whereas (c) is drawn for $R = -1$ and the singular point as the starting place. Each trajectory emerges from the singular point with a different slope.

The method of isoclines is readily extended to nonlinear or piecewise-linear resistance curves. For a piecewise-linear curve we merely treat each state like the linear examples shown in Fig. 10.16, hence the isoclines form a family of radial lines centered at the voltage intercept for that state. As shown in Fig. 10.17, the negative-resistance region (state III) for the current-controlled curve extends between break points. States I and II flank this region. The isoclines in this case form a set of curves like the negative-resistance curve $v = f(i)$ but distorted or "sheared" clockwise.

In Fig. 10.17, $R_{III} = -1$, hence the trajectories within state III will be like those of Fig. 10.16(c). The positive damping is critical, hence the resulting oscillations tend to be fairly nonsinusoidal. The particular value $R_I = R_{II} = 2$ make the isocline for $k = 1$ in states I and II
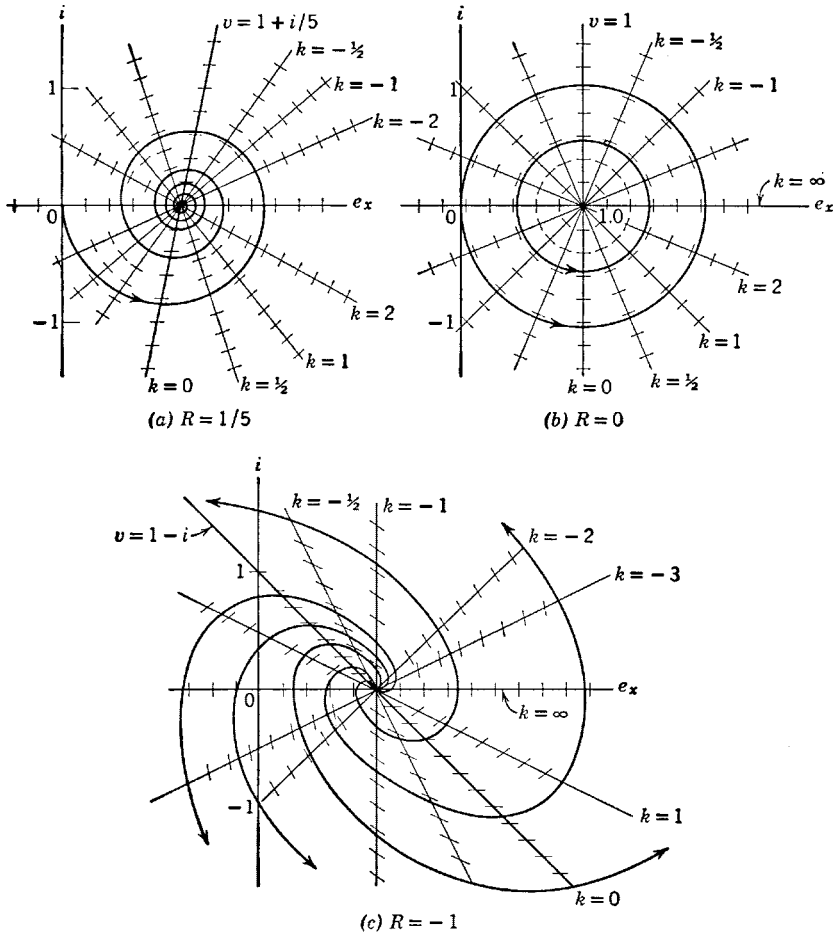
Fig. 10.16. Phase-plane trajectories by the method of isoclines.

different from all the others, because the directors lie along the isocline. This means that a trajectory cannot cross these lines; the operating point can be tangent or move along this isocline until a break point is reached. See, for example, the trajectory originating at point Q.

The trajectories starting at 0 and M in Fig. 10.17 illustrate the

formation of a limit cycle. The trajectory that begins at the origin spirals inward while the trajectory beginning at point M spirals outward. The limit cycle is the repetitive path approached by the two. A third trajectory from an initial point N would fall between the successive
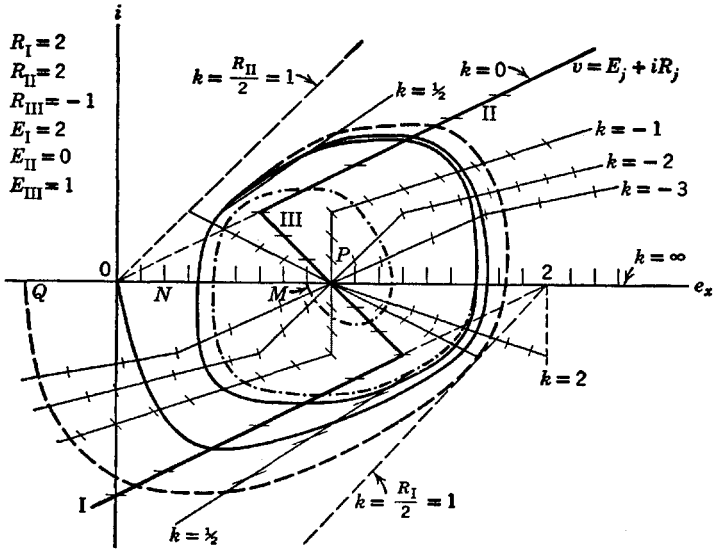
Fig. 10.17. Limit cycle by the method of isoclines.

rotations of the trajectory starting at the origin since two trajectories cannot cross each other.

The limit cycle in Fig. 10.17 resembles that of a relaxation oscillator. If the positive and negative damping is reduced, the trajectory becomes more nearly circular as in Fig. 10.14(d) and the current and voltage waveforms will be nearly sinusoidal.

SUPPLEMENTARY READING

A. A. Andronow and C. E. Chaikin, *Theory of Oscillations,* Princeton University Press, Princeton, N. J., 1949.

William A. Edson, *Vacuum-Tube Oscillators,* John Wiley and Sons, New York, 1953.

PROBLEMS

10.1. A current generator connected across the terminals of capacitance C in Fig. 10.2 applies a brief pulse of current (approximately an impulse) to change the voltage e_x. Assume the area of the pulse $I_0 \delta_0 = Q_0$ changes the

voltage e_x by an amount $\Delta e_x = E_0 = Q_0/C$. Sketch waveforms $i(t)$, $e_x(t)$ and the locus of i vs. e_x for $E_0 = 0$, $E/2$, E, $3E/2$ and $2E$ if the pulse is applied at $t = 0$.

10.2. Plot waveforms $i(t)$, $e_x(t)$ and the locus of i vs. e_x for the circuit of Fig. 10.3 with $R = 1$.

10.3. Referring to Fig. 10.5, indicate the effect of an initial capacitor voltage $e_x(0)$ on the waveforms and locus of i vs. e_x. Let $e_x(0) = -1$, -0.5, 0, 0.5, 1.0, 1.5, and 2.0.
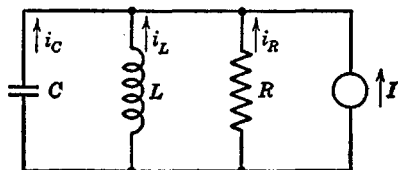
Fig. P10.1

10.4. Write the differential equations relating v and i_L for the linear RLC circuit shown in Fig. P10.1.

10.5. Normalize the equations for the circuit of Fig. P10.1 so that the only parameter appearing in the oscillatory equation is the damping factor.

10.6. By means of duality, relate the results of Problems 10.4 and 10.5 to the corresponding results for the series RLC circuit.

10.7. What values of damping resistance R (or conductance $G = 1/R$) yield the seven possible cases of damping for the parallel RLC circuit?

10.8. Consider a parallel RLC circuit with L and C equal to unity and with an initial unit current in the inductor. Let R assume convenient values such
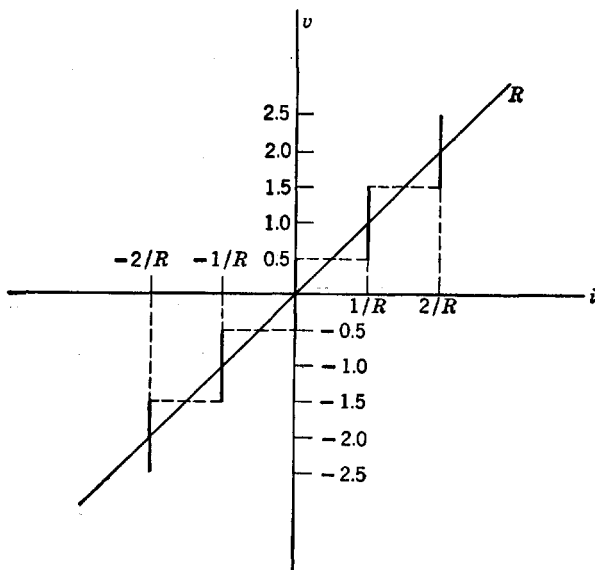
Fig. P10.2

as ±0.25, ±0.5, ±1, and ±2. In each case the resistance is to be approximated by a stepwise-linear curve as indicated in Fig. P10.2. Construct phase-plane trajectories.

10.9. In what region (or regions) of the phase plane is the approximation used in Problem 10.8 inadequate for each value of R considered?
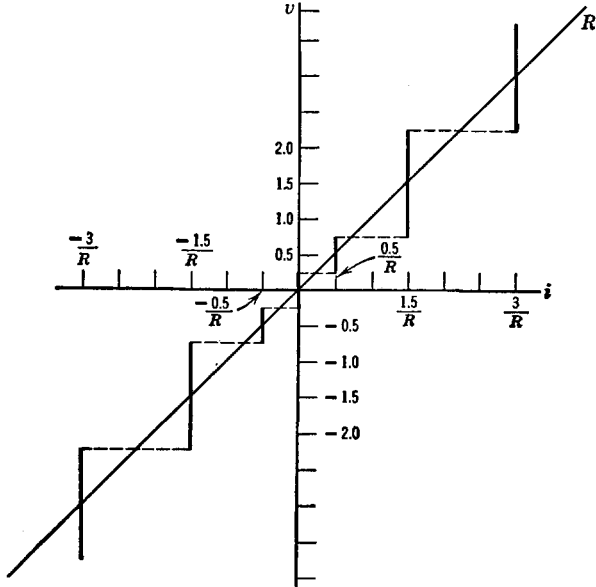
Fig. P10.3

10.10. Compare the results of Problem 10.8 with the results obtained from the approximation shown in Fig. P10.3.
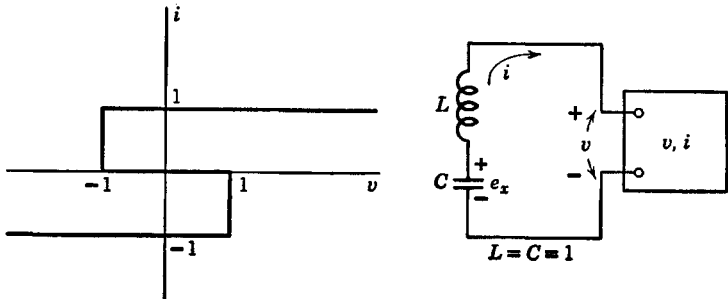
Fig. P10.4

10.11. (a) For the device illustrated in Fig. P10.4, write the differential equations governing v, e_x, and i, and construct a limit cycle.

(b) Sketch and dimension the waveforms of e_x and i vs. t.

(c) Using the information in parts (a) and (b), sketch $v(t)$ and $e_x(t)$.
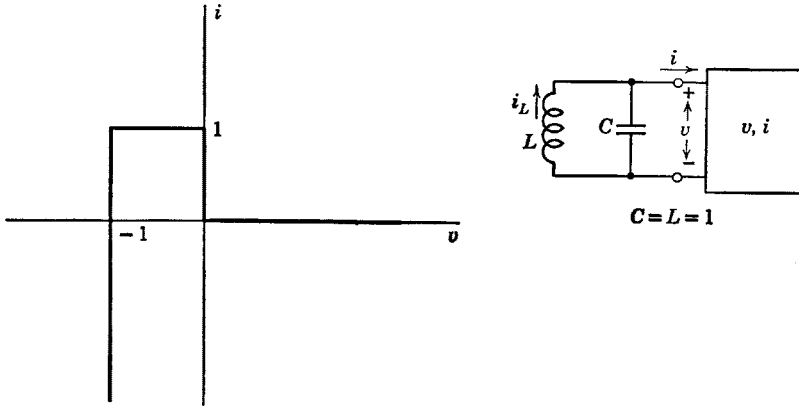
Fig. P10.5

10.12. (*a*) Find the limit cycle for the circuit shown in Fig. P10.5.

(*b*) What is the maximum energy stored in the tuned circuit at any instant of time?

(*c*) By utilizing the differential equations for the circuit and the limit cycle, carefully plot and dimension the waveforms, $v(t)$ and $i(t)$.
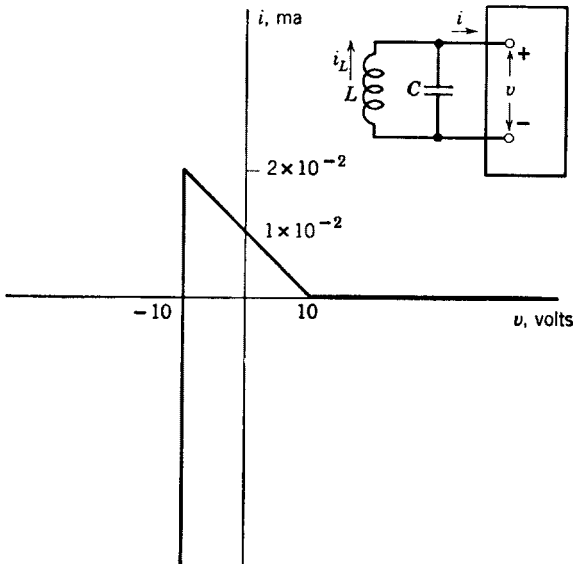
Fig. P10.6

10.13. For $L = 10^{-4}$h, $C = 1$ μf, find from Fig. P10.6 (by using step approximation for the linear slope) the limit cycle of $x = \sqrt{C}v$ and $y = \sqrt{L}i_L$. What is the maximum energy in the circuit?

10.14. (a) From Fig. P10.7, find the differential equations relating i_L with v and i.

(b) Using isoclines, find the phase-plane trajectory for v vs. i_L with an initial charge q_0 on C (giving $v = V_0$ at $t = 0$).
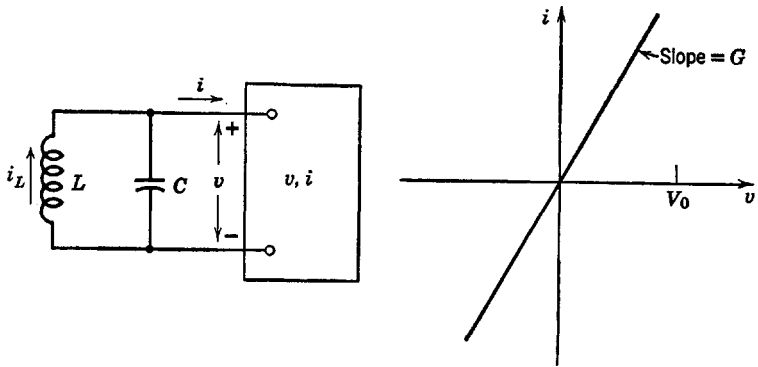
Fig. P10.7

10.15. A number of phase-plane trajectories for various initial conditions constitutes the "phase-plane portrait" of a circuit [Fig. 10.16(c) is an example]. Using the method of isoclines, construct phase-plane portraits for various values of positive damping in a linear parallel RLC circuit with L and C equal to unity. Use v and i_L for the phase-plane variables and distribute initial conditions equally around a unit circle.

10.16. From the results of Problem 10.15 deduce and sketch the phase-plane portraits for the corresponding values of negative damping.

10.17. Using the method of isoclines, find the limit cycle for the oscillator represented by the normalized circuit shown in Fig. P10.8.
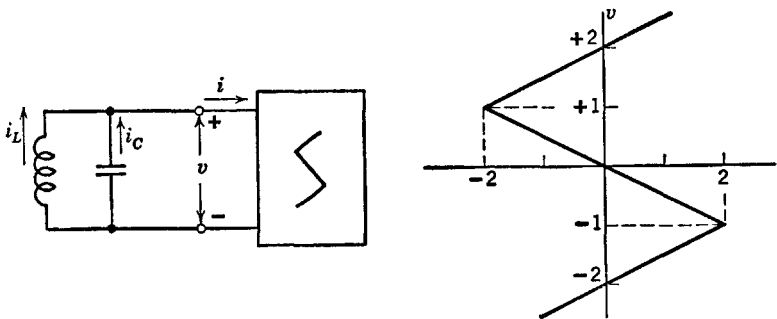
Fig. P10.8

10.18. Modify the resistive curve of Problem 10.17 to include the effects of a 1-ohm load resistance connected in parallel with the tuned circuit. Determine the limit cycle. For what value of load resistance will the circuit cease to sustain oscillations?

10.19. (*a*) Making reasonable approximations, find the piecewise-linear representation of v vs. i (as a function of R) for the circuit of Fig. P10.9(*a*).

(*a*)

(*b*)

Fig. P10.9

(*b*) If the actual resonant circuit consists of a parallel RLC combination with $R_0 = 10$ kilohms as shown in Fig. P10.9(*b*), sketch the phase-plane operation for the ranges of R from nonoscillation, to barely sustained, to over-driven. What value of R will just sustain sinusoidal oscillations?

Symmetry and
Balanced Circuits

11.1 Introduction

The exploitation of symmetry is a basic principle of circuit analysis and design—a revealing principle whose foundations lie in the mathematical theory of transformation groups. In circuit problems the recognition of some form of symmetry is almost certain to lighten the work of analysis. By restricting attention mainly to one or two forms of symmetry that are very common in electronic circuits we can avoid much of the formalism required in a general theory and at the same time illustrate the simplicity, precision, and power of the fundamental idea.

11.2 Symmetry, Symmetrical Components, and Superposition

Figure 11.1 shows an elementary balanced circuit exhibiting "lateral symmetry." Indeed, the left is the same as the right. This is too crude a notion of symmetry, however, and intuition will fail in more complicated examples unless we make the concept of symmetry more precise. Fortunately, a definition of symmetry does exist and it is both precise and simple. In the circuit of Fig. 11.1 the situation is unchanged by

turning the diagram over as one turns the page of a book. A mathematician might say that the geometrical entity is invariant under half-revolution about the y axis, or, for that matter, under any integral number of half revolutions. Inversion about the yz plane, equivalent to reflection in a mirror, would do just as well. These remarks give the key. Let it be understood henceforth that *symmetry* is exhibited by, consists of, and can be defined rigorously only as the *invariance* of a structure
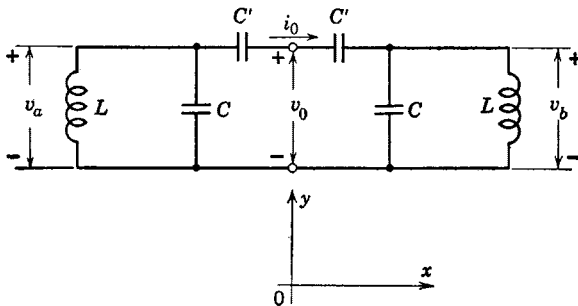
Fig. 11.1. A balanced circuit.

under some *transformation* or group of transformations. In the simple problems to be treated here, the structure is a circuit model and the transformations are rotations, translations, or inversions, all in three-dimensional space.

Having identified certain geometrical symmetries in a given circuit model, our next step is to look for electrical symmetry. *Electrical symmetry* is exhibited by, consists of, and can be defined rigorously only as the *invariance* of an *electrical experiment* under some *transformation* or group of transformations.

In particular, let x_i represent the voltage (or current) in the ith branch of the network. Assume that the transformation not only moves branch i to the position originally occupied by branch j, but that it also *multiplies x_i by some number K*. Thus the transformation can be represented as

$$Kx_i \rightarrow x_j, \quad \text{for various pairs } ij \tag{11.1}$$

If the *electrical experiment* is to be *invariant*, then we must have

$$Kx_i = x_j, \quad \text{for the pairs } ij \tag{11.2}$$

Solution of the Eqs. 11.2 gives, in general, a set of permissible values $K^{(1)}$, $K^{(2)}$, $K^{(3)}$, $\cdots$, $K^{(n)}$, for the parameter K, and each of these values of K defines a related set of voltages (or currents) $x^{(k)}$ which is *sym-*

metric; that is, invariant under the transformation. In particular

$$K^{(k)}x_i^{(k)} = x_j^{(k)} \tag{11.3}$$

In Fig. 11.1, for example, mirror-reflection about the y axis and multiplication of each voltage and current by K gives

$$Kv_a = v_b \tag{11.4}$$

$$Kv_b = v_a \tag{11.5}$$

$$Kv_0 = v_0 \tag{11.6}$$

$$Ki_0 = -i_0 \tag{11.7}$$

Eliminating v_b from Eqs. 11.4 and 11.5, we obtain

$$K^2v_a = v_a \tag{11.8}$$

or

$$K^2 = 1 \tag{11.9}$$

Hence

$$K^{(1)} = 1 \tag{11.10}$$

$$K^{(2)} = -1 \tag{11.11}$$

Taking the first value $K^{(1)}$ we find, from Eq. 11.3,

$$v_a^{(1)} = v_b^{(1)} \tag{11.12}$$

$$i_0^{(1)} = -i_0^{(1)} \tag{11.13}$$

with the result that

$$i_0^{(1)} \equiv 0 \tag{11.14}$$

The second value $K^{(2)}$ gives

$$-v_a^{(2)} = v_b^{(2)} \tag{11.15}$$

$$-v_0^{(2)} = v_0^{(2)} \tag{11.16}$$

so that

$$v_0^{(2)} \equiv 0 \tag{11.17}$$

Figure 11.2 shows the circuit-model interpretation of Eqs. 11.12 through 11.17.

Now back to the general case.   The variables $x^{(k)}$ are called *symmetrical components*. The advantage of symmetrical components lies in the fact that the circuit equations become simpler, as illustrated in Fig. 11.2.   Moreover, since superposition is valid in linear circuits, the solution of any problem can be represented as a superposition of symmetrical components.

Using Fig. 11.1 once more as the example, we see that all possible solutions for v_a and v_b can be represented as

$$v_a = v_a^{(1)} + v_a^{(2)} \qquad (11.18)$$

$$v_b = v_b^{(1)} + v_b^{(2)} \qquad (11.19)$$
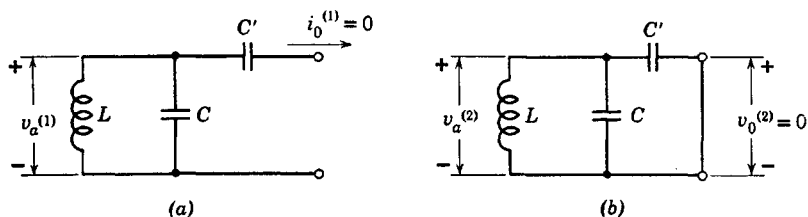
Fig. 11.2. Circuit models for (a) even-symmetric voltages $v^{(1)}$ and (b) odd-symmetric voltages $v^{(2)}$.

where $v_a^{(1)}$ and $v_a^{(2)}$ are solutions of the simpler circuit problems posed in Fig. 11.2. In view of Eqs. 11.12 and 11.15, it follows from Eqs. 11.18 and 11.19 that

$$v_a^{(1)} = \tfrac{1}{2}(v_a + v_b) \qquad (11.20)$$

$$v_a^{(2)} = \tfrac{1}{2}(v_a - v_b) \qquad (11.21)$$

Hence we could have written Eq. 11.19 as

$$v_b = v_a^{(1)} - v_a^{(2)} \qquad (11.22)$$
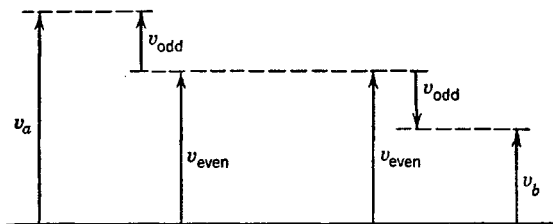
Fig. 11.3. Resolution of an arbitrary pair (v_a, v_b) into even and odd components.

In a laterally symmetric circuit, therefore, the electrical variables are associated in pairs (v_a and v_b in Fig. 11.1) or degenerate pairs (v_0 "and" v_0, or i_0 "and" $-i_0$). An arbitrary pair can always be resolved into even and odd components, as indicated in Fig. 11.3, in terms of which the circuit problem reduces to an independent pair of simpler problems.

11.3   Some Elementary Symmetries

In the preceding article, the general definition of symmetry was introduced and "lateral" symmetry was used as a running example. Lateral symmetry leads to even-symmetric and odd-symmetric com-
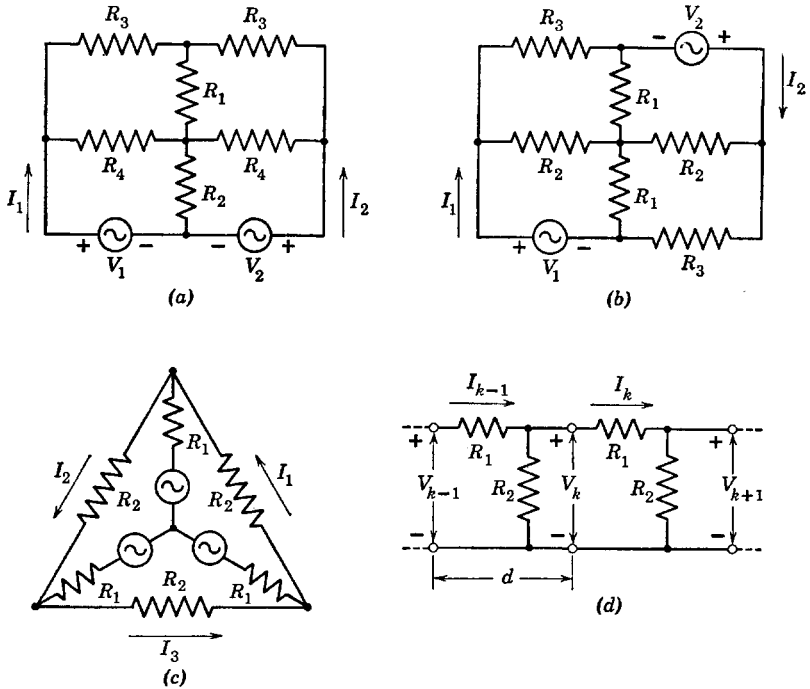
Fig. 11.4.   Various elementary symmetries. (a) Inversion about the vertical center line; (b) π-rotation (or inversion about the center point); (c) $2\pi/3$-rotation (or three different line inversions); (d) Translation (through distance d).

ponents.   Other symmetries yield other components, but *the principle is the same.*   The networks shown in Fig. 11.4 offer a few illustrations.

The experiment shown in Fig. 11.4(b) is invariant under π-rotation (180° rotation about an axis normal to the paper), provided $I_2 = KI_1$ and $I_1 = KI_2$ (and the same for the other current pairs and voltage pairs).   Hence both lateral symmetry, Fig. 11.4(a), and π-rotation symmetry, Fig. 11.4(b), lead to even and odd symmetric components.

In Fig. 11.4(c) we see a slightly different symmetry.   Consider a transformation which rotates the diagram counterclockwise through 120° and at the same time multiplies all currents and voltages in the

circuit by some factor K. If the experiment is to be invariant, then

$$I_2 = KI_1 \tag{11.23}$$

$$I_3 = KI_2 \tag{11.24}$$

$$I_1 = KI_3 \tag{11.25}$$

with similar triple-relationships holding for the other symmetrically disposed current or voltage triplets in the circuit. It follows directly that the permissible values of K are the cube roots of unity; three complex numbers lying 120° apart in the complex plane. Each value of K determines a symmetric-component-set. The three possible symmetric triplets are, therefore,

$$I_1^{(1)} = I_2^{(1)} = I_3^{(1)} \tag{11.26}$$

$$I_1^{(2)} = aI_2^{(2)} = a^2 I_3^{(2)} \tag{11.27}$$

$$I_1^{(3)} = a^2 I_2^{(3)} = a I_3^{(3)} \tag{11.28}$$

where

$$a = \epsilon^{-j2\pi/3} \tag{11.29}$$

Any distribution of voltage and current which satisfies the circuit equations can be represented as a superposition of such components. Symmetric-triplets are a basic tool in the study of unbalanced faults, loads, or interconnections on three-phase power transmission systems.

Figure 11.4(d) offers a final example. The structure is an extended ladder network, only two of whose sections are shown. Assume that all sections are identical. The symmetry test here involves translation of the diagram through a distance d together with multiplication of all currents and voltages by some complex factor K as yet unspecified. The set of symmetrical components is therefore determined by the invariance relations

$$I_k = KI_{k-1} = I_0 K^k \tag{11.30}$$

$$V_k = KV_{k-1} = KZI_{k-1} = ZI_0 K^k \tag{11.31}$$

for any section k, where I_0 and Z are convenient constants to be determined later. A study of Fig. 11.4(d) yields the circuit equations

$$V_{k-1} - V_k = R_1 I_{k-1} \tag{11.32}$$

$$R_2(I_{k-1} - I_k) = V_k \tag{11.33}$$

which become, with the aid of Eqs. 11.30 and 11.31,

$$Z(1 - K) = R_1 \tag{11.34}$$

$$R_2(1 - K) = ZK \tag{11.35}$$

Quantity I_0 has cancelled and thus is relegated to the role of an arbitrary constant. Equations 11.34 and 11.35 lead to a quadratic in K,

$$K^2 - \left(\frac{R_1}{R_2} + 2\right) K + 1 = 0 \qquad (11.36)$$

and the two solutions determine two sets of symmetrical components. Each set has its own value of Z, found from K through Eqs. 11.34 or 11.35. Thus, all possible currents in the extended ladder network are simple superpositions of two basic components of the form

$$I_k^{(1)} = A[K^{(1)}]^k \qquad (11.37)$$

$$I_k^{(2)} = B[K^{(2)}]^k \qquad (11.38)$$

The arbitrary constants A and B can be adjusted to meet the specified boundary conditions at the left and right terminations of the ladder.
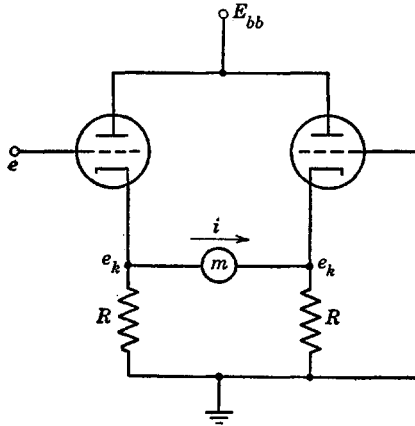
Fig. 11.5. A balanced vacuum-tube voltmeter circuit.

11.4 A Balanced Vacuum-Tube Voltmeter Circuit

The circuit of Fig. 11.5 may be used as a d-c vacuum-tube voltmeter, with the milliammeter m calibrated to indicate values of the applied voltage e. Let us assume that the two tubes are identical and that the meter m has negligible internal resistance. Moreover, we shall consider only small values of e so that operation is linear and superposition is consequently valid.

When e takes the value zero, the circuit is perfectly symmetrical and i, having no preferred direction, must also vanish. The immediate

result of symmetry, therefore, is that the voltmeter transfer curve, $i = f(e)$, must pass through the origin. Let us now see how symmetry can be exploited to calculate the slope of this curve at the origin. First observe that the grid-to-ground voltages shown in Fig. 11.5 can be resolved into the even and odd components indicated in Fig. 11.6. Current i remains zero for even grid excitation [Fig. 11.6(a)]. Furthermore, the common cathode voltage e_k is unchanged by odd grid inputs
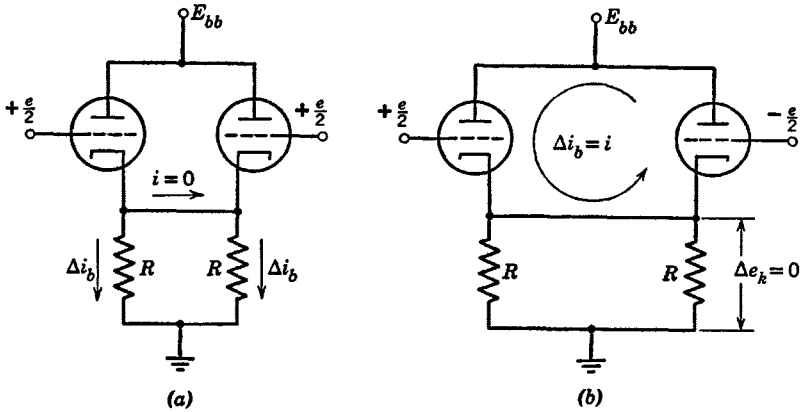
Fig. 11.6. Symmetrical components in the voltmeter circuit. (a) Even components; (b) Odd components.

(b), provided e is small enough so that operation is linear. Since e_k is fixed in (b), the two plate voltages are fixed and the grid-to-cathode voltage increments are just equal to the corresponding applied grid-to-ground voltages, $e/2$ and minus $e/2$. Hence the plate current of the left tube increases by

$$\Delta i_b = g_m \frac{e}{2} \tag{11.39}$$

and the right tube current decreases by the same amount. Since no change can occur in the current carried by either of the cathode resistors R, it follows that the flow of incremental current takes the circulatory pattern shown in (b), down through the left tube, across through the meter and up in the right-hand triode. Hence we have the very simple result,

$$\frac{i}{e} = \frac{g_m}{2} \tag{11.40}$$

giving the slope of the voltmeter transfer curve at the origin. Sym-

metry considerations have been much help here, offering a simple path to the simple answer. Failure to exploit symmetry would have detoured us along a more complicated path to the same answer.

11.5   A Balanced Power Amplifier with Direct-Coupled Load

Figure 11.7(a) shows an amplifier designed for good linearity at a relatively high level of output power to the load R. The three-winding input transformer applies the input signal e_g to the two grids with the
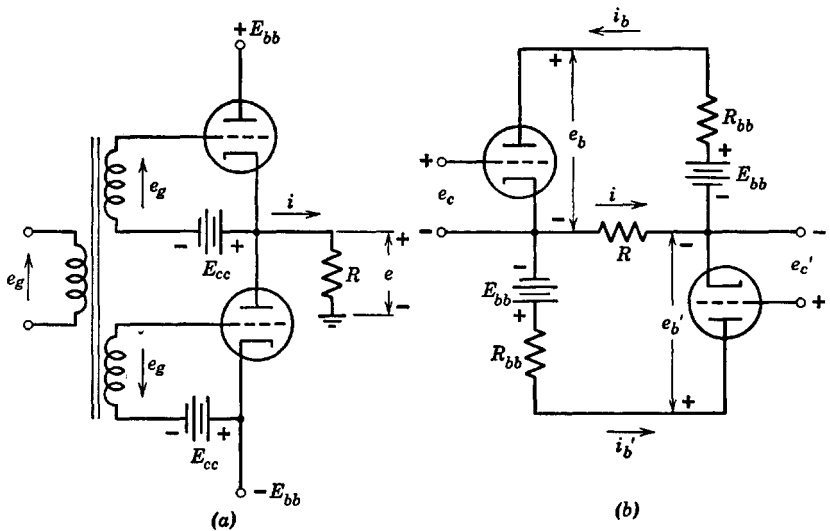
Fig. 11.7. A balanced power amplifier.

polarities indicated in the figure. Recognition of symmetry is facilitated by interchanging the series connection of the lower tube and its plate supply voltage E_{bb}, to obtain the modified circuit shown in (b). This does not affect the operation of the amplifier. The input transformer and grid-biasing batteries are omitted from (b) for simplicity and the effect of internal plate-supply resistance is accounted for by the resistances R_{bb}. Circuit (b) exhibits the same form of symmetry as that shown in Fig. 11.4(b).

For a qualitative description of performance, suppose that the tubes are biased almost to cutoff, so that neither tube conducts much current when e_g is zero. Now apply a large signal e_g varying sinusoidally with time. As e_g swings positive, the lower tube cuts off and the upper tube

conducts, producing a nearly-sinusoidal positive load current i during the first half cycle of the sine wave. In the following half cycle, e_g goes negative, the upper tube is cut off, and the lower tube supplies a nearly sinusoidal negative surge of load current. The load current, therefore, is nearly sinusoidal over the entire cycle, although the operation of each tube is grossly nonlinear. This permits more than twice the amplitude of load current obtainable with one tube alone for a specified amount of distortion.

So much for the qualitative behavior; now to calculate the precise manner in which the load current i varies with the input voltage e_g, taking into account the actual nonlinearity of the tubes. By inspection of Fig. 11.7(b),

$$e_b = E_{bb} - R_{bb}i_b - Ri \tag{11.41}$$

$$e_b{}' = E_{bb} - R_{bb}i_b{}' + Ri \tag{11.42}$$

$$i = i_b - i_b{}' \tag{11.43}$$

Now let us resolve the plate currents and plate voltages into even and odd symmetrical components

$$e_b = e_{be} + e_{bo} \tag{11.44}$$

$$e_b{}' = e_{be} - e_{bo} \tag{11.45}$$

$$i_b = i_{be} + i_{bo} \tag{11.46}$$

$$i_b{}' = i_{be} - i_{bo} \tag{11.47}$$

Substitution into Eqs. 11.41 and 11.42 and elimination of i with the aid of Eq. 11.43 gives a pair of equations which may be added and subtracted to obtain

$$e_{be} = E_{bb} - R_{bb}i_{be} \tag{11.48}$$

$$e_{bo} = -(R_{bb} + 2R)i_{bo} \tag{11.49}$$

The even-symmetric current component may be visualized as a current running counterclockwise around the outside loop of circuit in Fig. 11.7(b), whereas the odd component appears as two loop currents, counterclockwise in the upper loop of the circuit and clockwise in the lower loop. Resistance R is doubled in Eq. 11.49 because both odd currents flow in the same direction through R, thereby doubling the voltage drop. Voltages E_{bb} form an even pair and hence are absent from Eq. 11.49. Even-component current does not flow in the load so that R does not appear in Eq. 11.48. Thus, Eqs. 11.48 and 11.49 could have been deduced directly from the pattern of symmetrical components in the circuit.

Equations 11.48 and 11.49 have the advantage over 11.41 and 11.42 in that they lend themselves to a much simpler interpretation as a graphical construction on the triode plate curves. The construction is shown in Fig. 11.8. For a specified pair of grid-voltage values the upper tube operates at point A and the lower tube at point A'. The resistance in Eq. 11.49 determines the slope of the "odd" load line AA', but does not say where this line should be located. The constants in
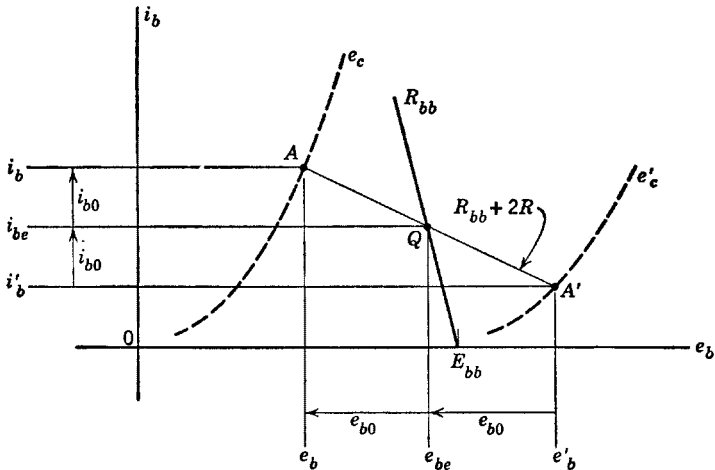
Fig. 11.8. Construction of symmetrical components on the triode plate curves.

Eq. 11.48, on the other hand, completely fix the "even" load line as shown on the diagram, giving its resistance-slope R_{bb} and its intercept E_{bb}. The odd load line intersects the even load line at some point Q. If we now slide the odd load line up and down, maintaining the specified resistance-slope $R_{bb} + 2R$, we will find one position for which point Q lies midway between A and A'. This is the proper location, since Eqs. 11.44 through 11.47 are then satisfied, as indicated by the constructions and dimensions below and to the left in Fig. 11.8.

For a different pair of grid-voltage values point Q would shift to a new position along the even load line and two new operating points A and A' would be found. The actual construction is not as time consuming as it may appear. A pair of draftsman's triangles may be used to shift the odd load line parallel to itself and the proper position can be estimated by eye and checked with a pair of dividers. For a given input signal e_g, the two grid voltages are

$$e_c = -E_{cc} + e_g \qquad (11.50)$$

$$e_c' = -E_{cc} - e_g \qquad (11.51)$$

Choosing a number of different values of e_g, we compute the correspond-
ing grid-voltage pairs and locate A and A' for each pair. A smooth
curve drawn through the various points A (shown in Fig. 11.9 but not
in Fig. 11.8) represents the operating locus of the upper tube as e_g is
varied, and the same for points A' along the operating locus of the
lower tube.
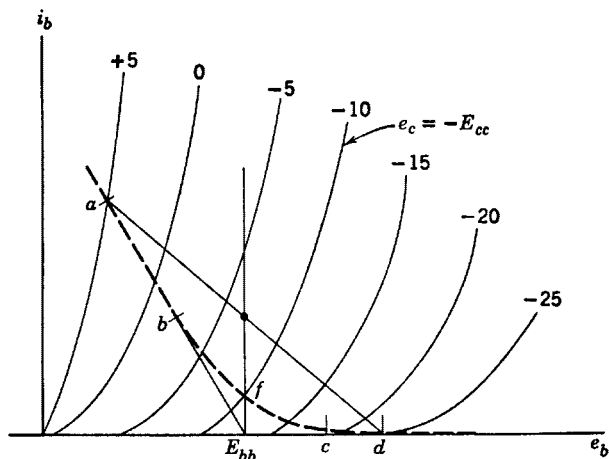
Fig. 11.9. Common operating locus of the two tubes.

It is tacitly assumed that the tubes are identical, although the con-
struction of operating loci could be carried out on the superimposed
plate curves of two different tubes in just the same manner. For identical
tubes, and for the balanced input specified by Eqs. 11.50 and 11.51, the
two operating loci coincide, as shown by the dashed curve in Fig. 11.9.
This is assured by symmetry. When e_g is zero both tubes operate at
point f. As e_g goes positive, the operating point of the upper tube moves
up the locus toward b and that of the lower tube slides down toward
cutoff at c. For a value of e_g greater than about 7.5 volts, the lower tube
is cut off and has no effect upon current flow in the upper loop of the
circuit. Hence the upper tube operates along the straight load line ba,
whose resistance-slope is just $R_{bb} + R$. At 15 volts input the odd load
line has shifted upward as far as position ad. For a sinusoidal input
signal the operating points of the two tubes run back and forth along the
locus in opposite directions, meeting at point f twice during each cycle.

Once the locus is determined, the amplifier transfer curve shown in
Fig. 11.10 may be plotted by taking values of i_b and i_b' directly from
the locus. An odd grid-signal pair e_g produces both even and odd changes
in the plate currents since the circuit of Fig. 11.7 is not linear. Only the

odd component of plate current appears in the load, however, so that the transfer curve is odd-symmetric. The curvature of triode characteristics is such that for proper choice of E_{cc} the transfer curve is actually very nearly linear.
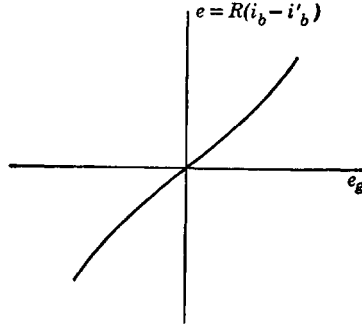
Fig. 11.10. Voltage transfer curve of the balanced amplifier.

The slope of the transfer curve at the origin is

$$G_0 = \frac{2\mu_0 R}{r_{p0} + R_{bb} + 2R} \qquad (11.52)$$

where μ_0 and r_{p0} are the triode amplification factor and plate resistance measured at point f in Fig. 11.9. At the upper extremity of operation, in region ab of the locus, the amplifier voltage gain is

$$G = \frac{\mu R}{r_p + R_{bb} + R} \qquad (11.53)$$

where μ and r_p are measured in region ab. The bias voltage E_{cc} may be increased so long as the approximation

$$G_0 \approx G \qquad (11.54)$$

is valid. A further increase in E_{cc} would then cause the transfer curve to depart more drastically from linearity.

11.6   The Transformer-Coupled Push-Pull Amplifier

No discussion of balanced amplifiers would be complete without mention of the push-pull amplifier, which is widely used as a driver for loudspeakers. The basic circuit is shown in Fig. 11.11, with the output transformer idealized except for resistances R_w, which account for transformer winding resistances. The transformer has n turns in each

half of the primary winding, per turn on the secondary winding, as indicated in the figure. In comparison with the circuit of Fig. 11.7(a), the push-pull amplifier in Fig. 11.11 has its two tubes in parallel (rather than in series) with respect to the plate-voltage supply E_{bb}, and in series (rather than in parallel) with respect to the output load.
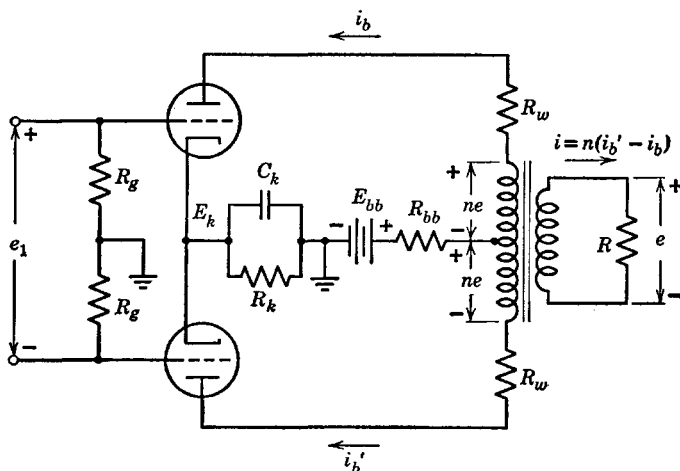
Fig. 11.11. A push-pull amplifier with transformer-coupled load.

The grid-biasing circuit in Fig. 11.11, consisting of R_k and C_k, may be replaced by a battery E_k for purposes of analysis. The bias actually increases somewhat for large input signal amplitudes but we shall assume that R_k has been adjusted to produce the proper value of E_k with a given signal applied.

The equations of the two principal loops of the push-pull circuit are

$$e_b = (E_{bb} - E_k) + ne - R_w i_b - R_{bb}(i_b + i_b') \qquad (11.55)$$

$$e_b' = (E_{bb} - E_k) - ne - R_w i_b' - R_{bb}(i_b + i_b') \qquad (11.56)$$

Addition and subtraction of these, together with substitution of the transformer relation

$$i = n(i_b' - i_b) \qquad (11.57)$$

leads to the symmetrical equations

$$\frac{e_b + e_b'}{2} = (E_{bb} - E_k) - (R_w + 2R_{bb})\left(\frac{i_b + i_b'}{2}\right) \qquad (11.58)$$

$$\frac{e_b - e_b'}{2} = -(R_w + 2n^2 R)\left(\frac{i_b - i_b'}{2}\right) \qquad (11.59)$$

With even and odd subscript notation, these become

$$e_{be} = (E_{bb} - E_k) - (R_w + 2R_{bb})i_{be} \qquad (11.60)$$

$$e_{bo} = -(R_w + 2n^2 R)i_{bo} \qquad (11.61)$$

From here on the analysis picks up the previous trail at Eqs. 11.48 and 11.49 and follows exactly the same path to a transfer curve of output voltage $nR(i_b' - i_b)$ versus input voltage e_1. Notice that in the push-pull circuit the odd component of grid signal is only half of the input voltage, since e_1 splits equally across the two resistors R_g. For the push-pull form of symmetry the even component of plate current flows counterclockwise in the upper loop and clockwise in the lower, producing no magnetization of the transformer core. Hence output voltage arises entirely from the odd component of plate current, which runs downward through both tubes in series and returns upward through the transformer primary winding.

Before leaving the push-pull circuit we should say something about the choice of the transformer turns-ratio n. The optimum n will depend, of course, upon the particular conditions and restrictions laid down, and these are somewhat arbitrary. One possible set of conditions is the following.

1. Load power to be a maximum
2. Grid voltage never positive
3. Plate supply voltage $(E_{bb} - E_k)$ fixed
4. R_{bb} and R_w negligible
5. Input signal adjustable to any desired amplitude
6. Grid bias near cutoff

Under these conditions the operating path has the general shape shown in Fig. 11.9, with the special dimensions indicated in Fig. 11.12. For any given waveform, the output power is proportional to the product of the output current and voltage amplitudes. Examination of Figs. 11.11 and 11.12 shows that

$$x = ne_{max} \qquad (11.62)$$

$$y = \frac{1}{n} i_{max} \qquad (11.63)$$

so that (for a specified signal waveform) output power is proportional to the product

$$p = xy \qquad (11.64)$$

Since grid voltage is restricted to nonpositive values, the greatest power output will occur when the extremity of the operating path lies some-

where along the zero-grid-voltage curve; at point q in Fig. 11.12. Geometrically speaking, our job is to maximize the area of the triangular region xy by varying the resistance-slope n^2R. As point q moves upward along the zero-grid-voltage curve, a position is reached at which y is increasing percentagewise just as fast as x is decreasing percentagewise.
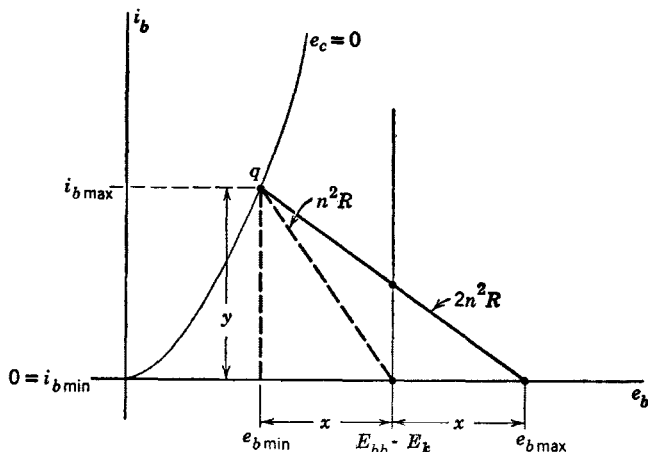
Fig. 11.12. Maximization of push-pull power output.

At this point the product xy is stationary and also evidently a maximum. In short,

$$\frac{dy}{y} = -\frac{dx}{x} \tag{11.65}$$

The relation given in Eq. 11.65 can be obtained formally by setting the total differential of Eq. 11.64 equal to zero. By inspection of Fig. 11.12,

$$r_p = -\frac{dx}{dy} \tag{11.66}$$

$$n^2R = \frac{x}{y} \tag{11.67}$$

and from Eq. 11.65 we see that the condition for maximum output power is

$$r_p = n^2R \tag{11.68}$$

where r_p is measured at point q. Thus the transformer turns-ratio n should be chosen to "match" the load resistance R to the value of the incremental plate resistance r_p at the extremity of the operating swing,

where one tube is cut off and only half of the transformer primary winding is effective. The load resistance reflected into the plate circuit by half of the winding is, of course, just equal to n^2R.

11.7   A Classical Puzzle

The story is told of a classical puzzle put to a famous mathematician. Two players sit at a rectangular table, each with a pile of identical ideal short sticks in his lap. The first player takes his turn by placing one little stick on the table. The second player then places another stick so that it does not overlie or otherwise disturb the first. The turns alternate and the last player able to place a stick legally on the table wins the game. The problem is to give a plan for winning, if possible, and to decide whether or not it is advantageous to have the first move. The famous mathematician, so the story goes, came to the solution in twelve seconds. This problem would be hopeless without help from the elegant and powerful concept of symmetry. [Solution: The first player places his first stick in the center of the table; the orientation is immaterial. Thereafter, he simply matches every move of the second player, maintaining symmetry about the center.] An interesting extension of the problem arises from the question, "Can the first player win without placing his first stick in the center?"

SUPPLEMENTARY READING

Herman Weyl, *Symmetry*, Princeton University Press, Princeton, N. J., 1952.
T. S. Gray, *Applied Electronics*, John Wiley and Sons, New York, 1954.
E. A. Guillemin, *Introductory Circuit Theory*, John Wiley and Sons, New York, 1953.

PROBLEMS

11.1. In the circuit of Fig. P11.1(a), let $v_1 = v_1^{(1)} + v_1^{(2)}$, $v_2 = v_1^{(1)} - v_1^{(2)}$, and similarly for i_1 and i_2. Show that $v_1^{(1)}$, $v_1^{(2)}$, $i_1^{(1)}$, and $i_1^{(2)}$ are related as in the circuits of Fig. P11.1(b) and (c). Using superposition of symmetrical components, find v_1/i_1 when v_2 is zero. Find v_1/i_1 when i_2 is zero.

11.2. Show that the construction in Fig. P11.2(b) determines the currents and voltages in the circuit of Fig. P11.2(a). The conditions are: v and i are the coordinates of point p; v' and i' are the coordinates of point p'; and O bisects pp'. Show how you would find v_1 and v_2 from the construction.
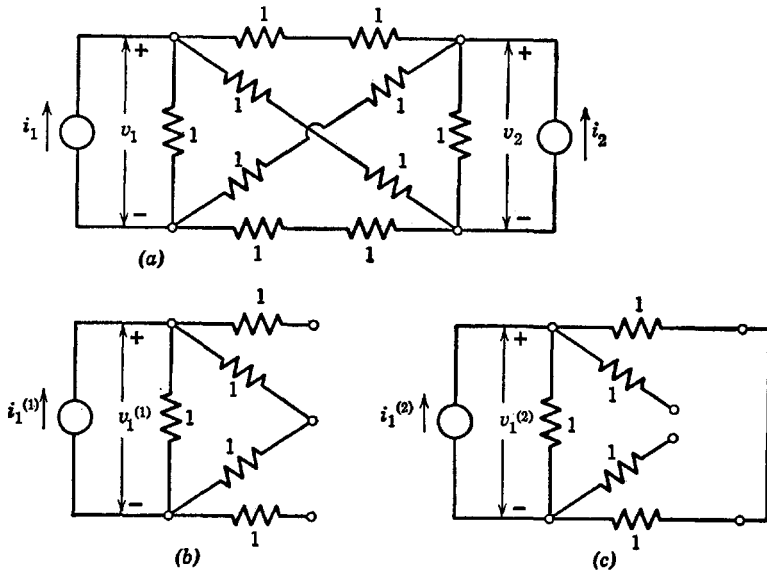
Fig. P11.1
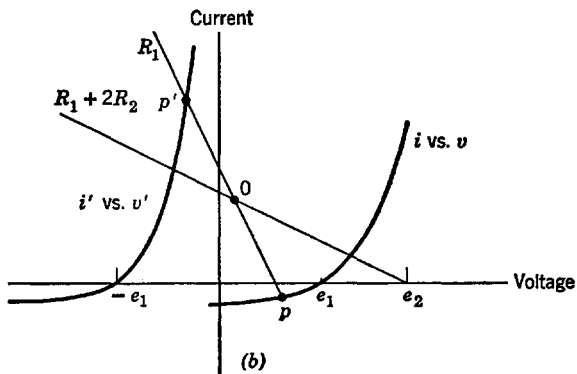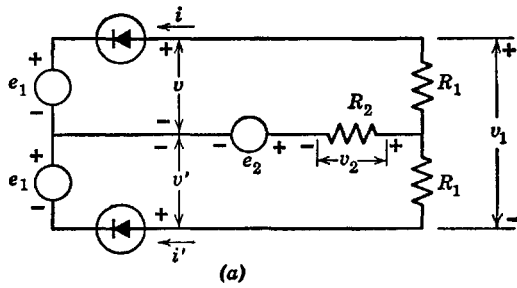
Fig. P11.2

11.3. For the push-pull class A power amplifier circuit shown in Fig. P11.3, determine R and E_{cc} for maximum undistorted power output without exceed-
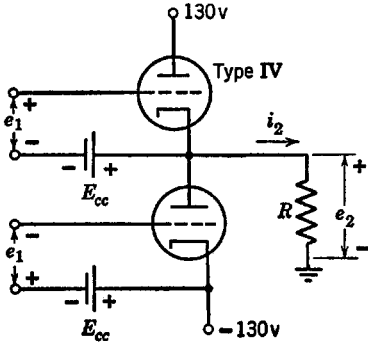
Type IV ratings per triode:

$E_{bb\ max} = 250$ volts
$I_{b\ max} = 120$ ma
$P_{b\ max} = 13$ watts

Fig. P11.3

ing the tube ratings. For this load locate the path of operation in the e_b vs. i_b plane and determine the maximum power output.
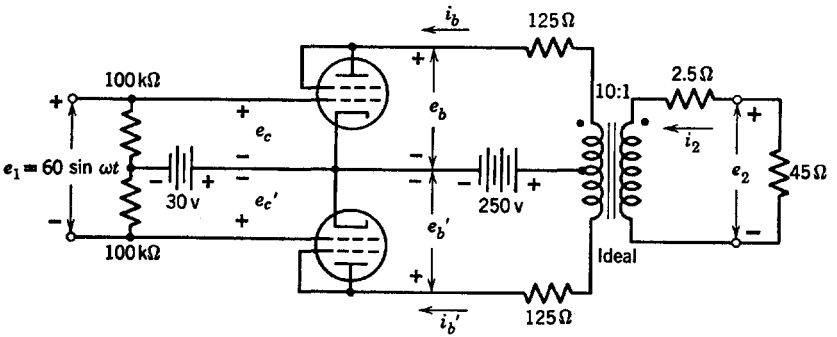
Fig. P11.4

11.4. Two Type V triodes are used in the push-pull amplifier of Fig. P11.4. Under operating conditions with the input signal applied, show that the instantaneous plate voltages and currents are:

$$E_{sum} = \left(\frac{e_b + e_b'}{2}\right) + R_{sum}\left(\frac{i_b + i_b'}{2}\right)$$

$$0 = \left(\frac{e_b - e_b'}{2}\right) + R_{dif}\left(\frac{i_b - i_b'}{2}\right)$$

where $E_{sum} = 250$ volts, $R_{sum} = 125$ ohms, $R_{dif} = 2500$ ohms.

11.5. For the circuit of Fig. P11.4,

(a) Show that point Q must be the mid-point of line segment AA' in Fig. P11.5.

(b) Using the Type V curves compute e_c, e_c', i_b, i_b', and i_b-i_b' for $e_1 = 0$, 15, 30, 45, 60. Notice that the operating path is a straight line for $e_c > -15$.
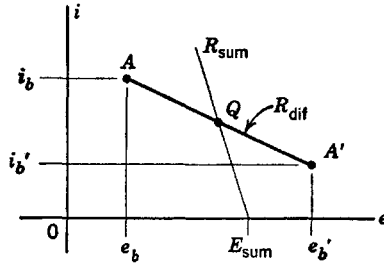
Fig. P11.5

11.6. For the circuit of Fig. P11.4,

(a) Show that $e_2 = 225(i_b' - i_b)$;

(b) Show that the transfer curve $e_2 = f(e_1)$ is an odd function.

(c) Plot $e_2 = f(e_1)$ for $0 < e_1 < 60$. Use roughly 10 volts per inch for e_1 and 2.5 volts per inch for e_2. This curve can be used to find the output waveform $e_2(t)$ due to any symmetrical input wave $e_1(t)$;

(d) Approximate $e_2 = f(e_1)$ by a straight line $e_2 = Ae_1$, choosing A for least magnitude of error in e_2 over the range $0 < e_1 < 60$. Now sketch and dimension the output distortion voltage $e_{2d} = e_2 - Ae_1$ versus t for one-half cycle of the input $e_1 = 60 \sin \omega t$. What distortion frequency is strongly evident? What is the distortion amplitude in per cent of the $e_2(t)$ amplitude? What is the gain A? What is the output power?
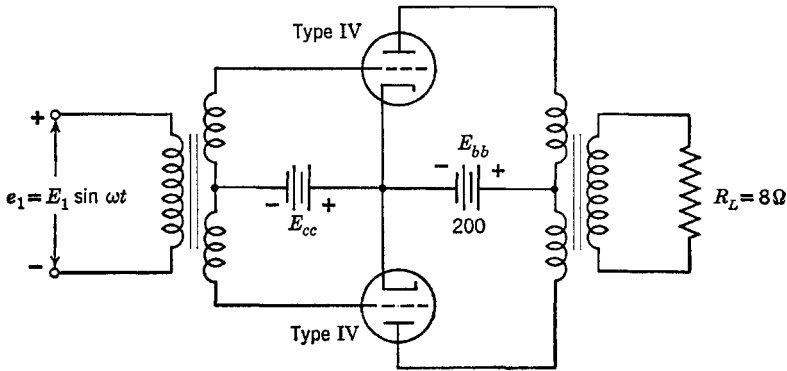
Fig. P11.6

11.7. The basic circuit for a push-pull amplifier is shown in Fig. P11.6. The input transformer has a turns ratio of $1:2$ from primary to each half of the secondary. The turns ratio $n:1$ of the output transformer from either half of the primary to the secondary is to be determined for a load resistance of 8 ohms.

Determine E_1, E_{cc}, and n to yield maximum power output with each triode conducting plate current during approximately one-half cycle, with $e_c < 0$ (class B). What is the power output?

Transistor Curves
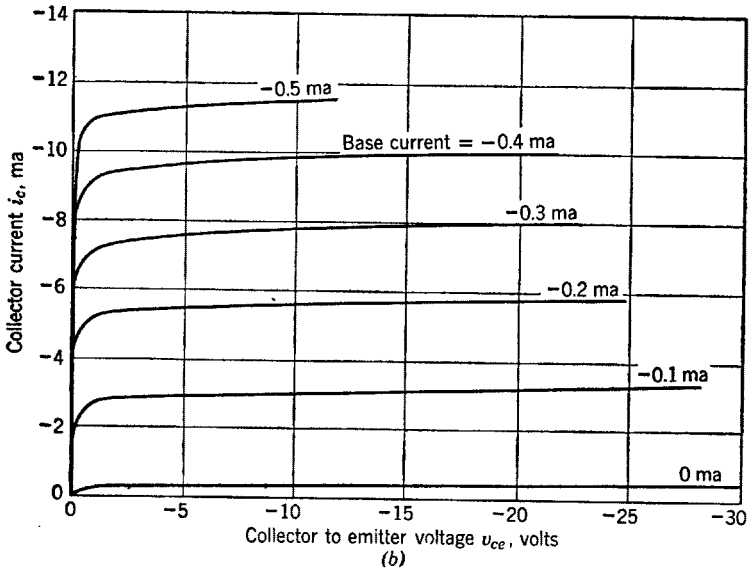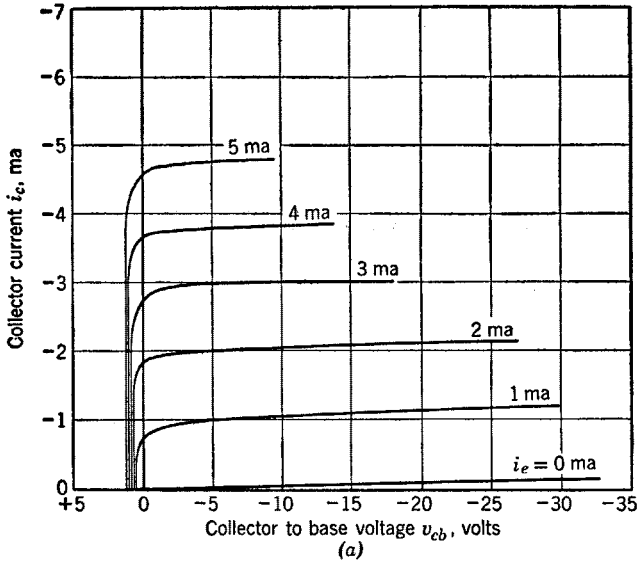
Fig. A.1. Type I transistor (p-n-p junction) collector curves: (a) Common base; (b) Common emitter.
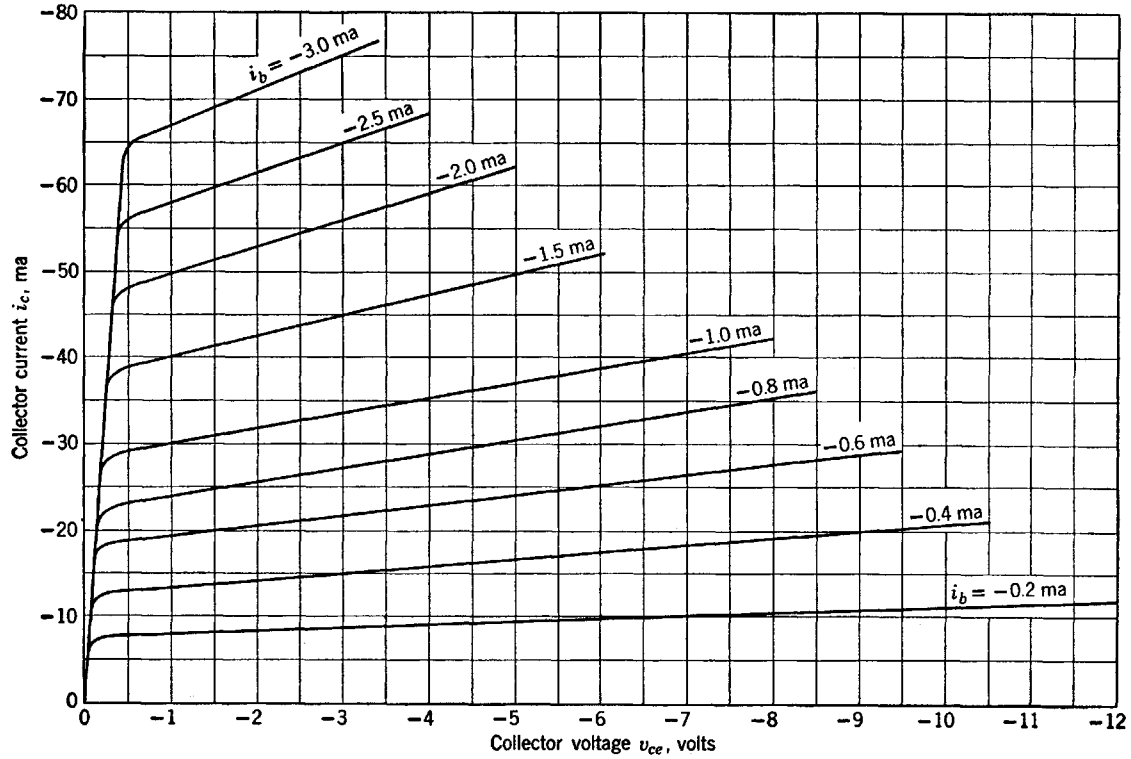
Fig. A.2. Type II transistor (*p-n-p* junction) collector curves: common emitter.
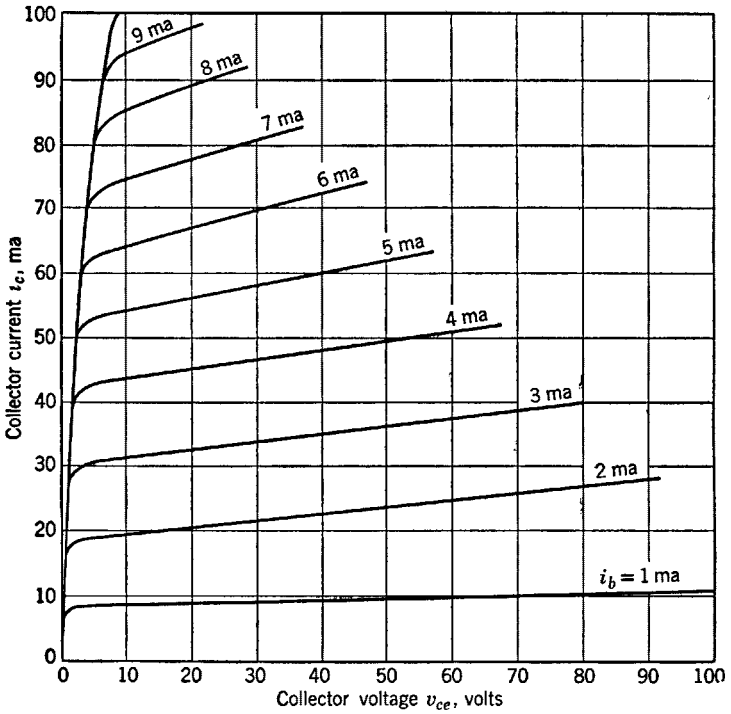
Fig. A.3. Type III transistor (*n-p-n* junction) collector curves: common emitter.
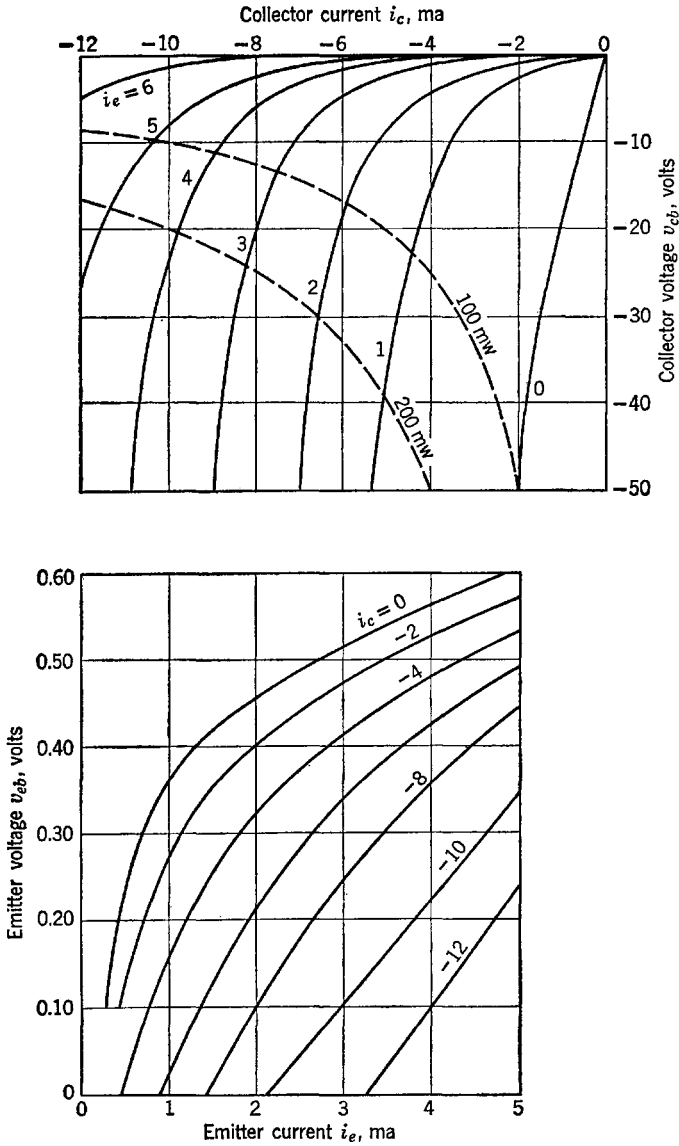
Fig. A.4. Type IV transistor (point-contact) collector and emitter curves: common base.
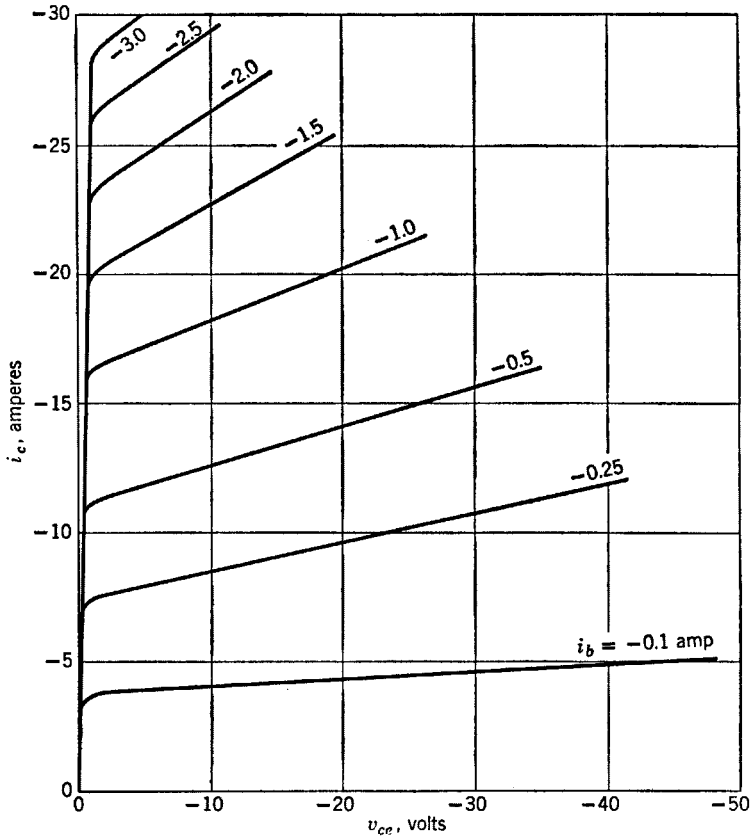
Fig. A.5. Type V transistor (p-n-p junction) collector curves: common emitter.

Triode Curves
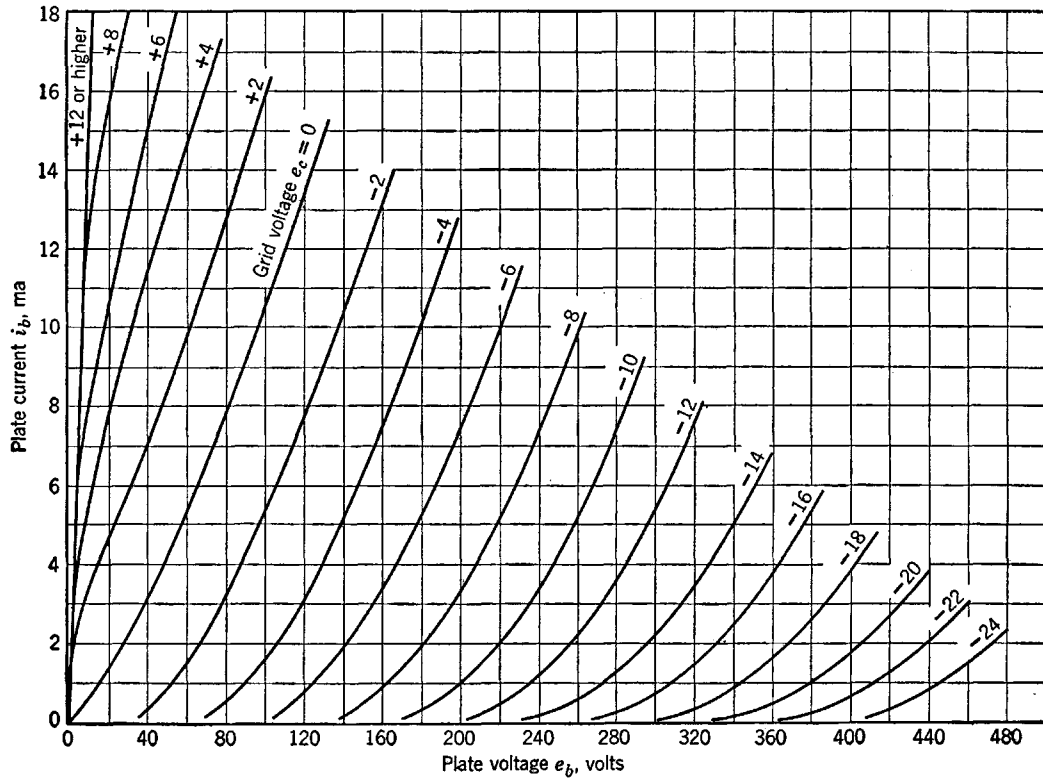
Fig. B.1. Type I triode plate curves.
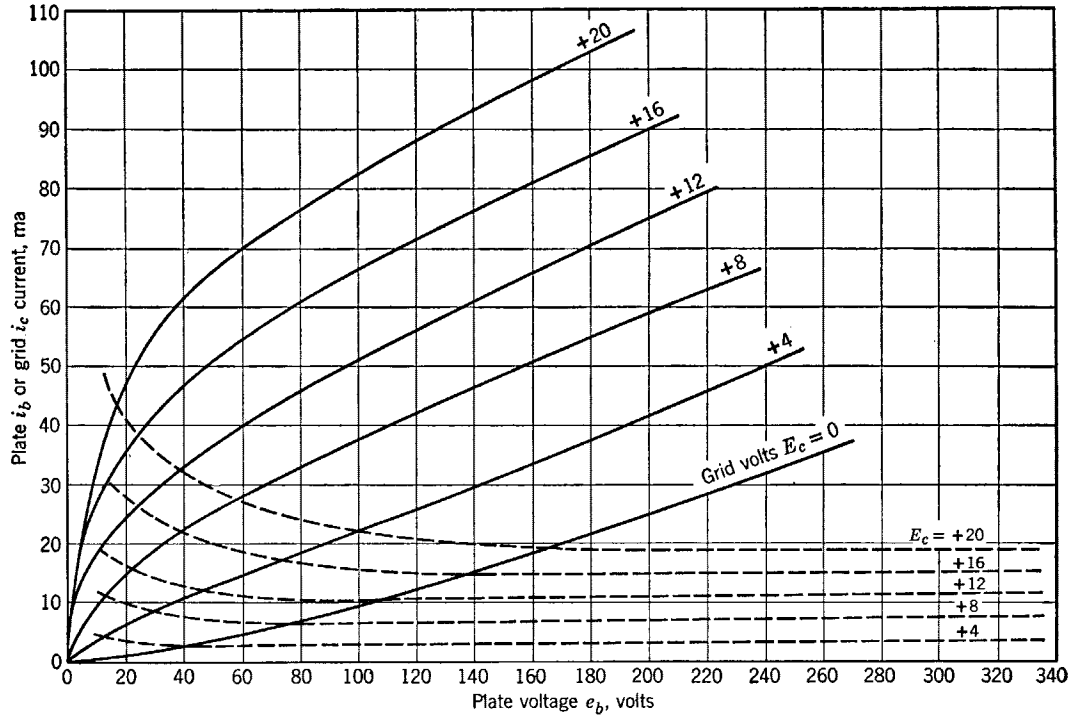
Fig. B.2. Type I triode plate and grid curves for positive grid voltage.
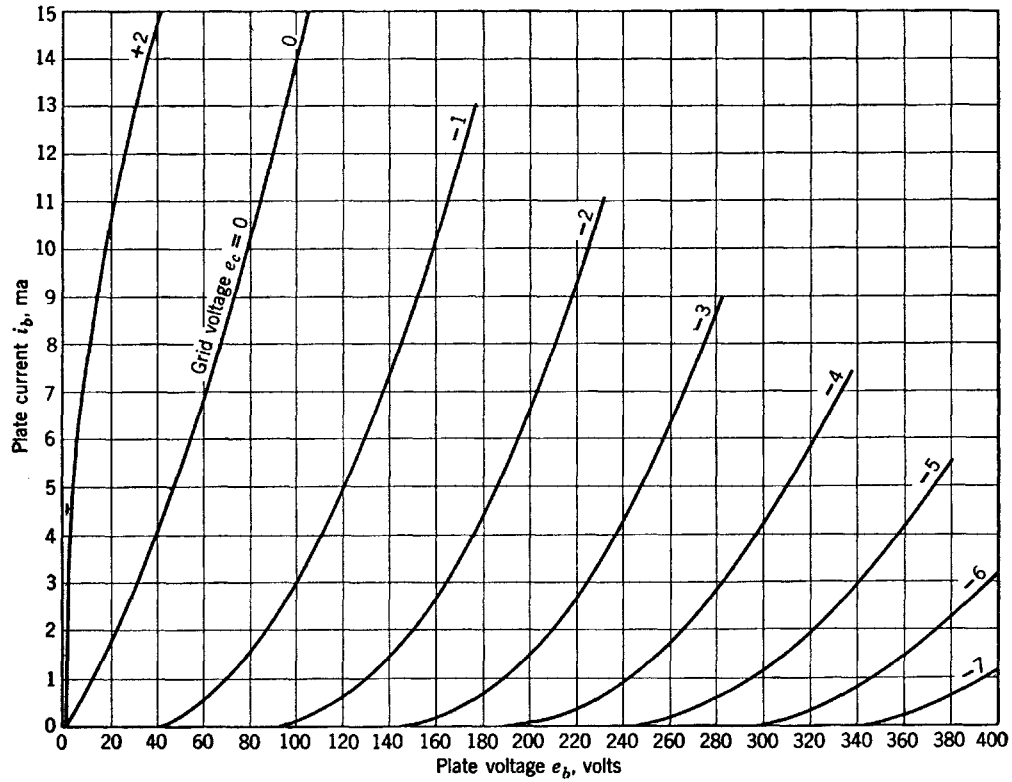
Fig. B.3.   Type II triode plate curves.
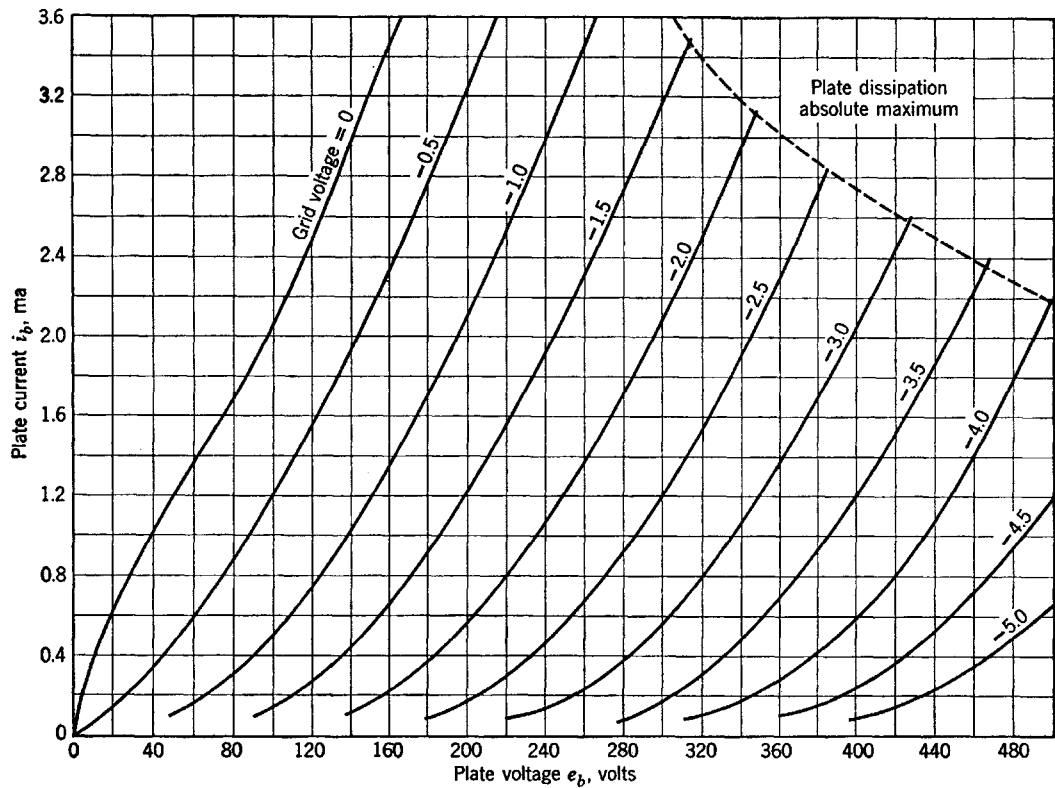
Fig. B.4. Type III triode plate curves.
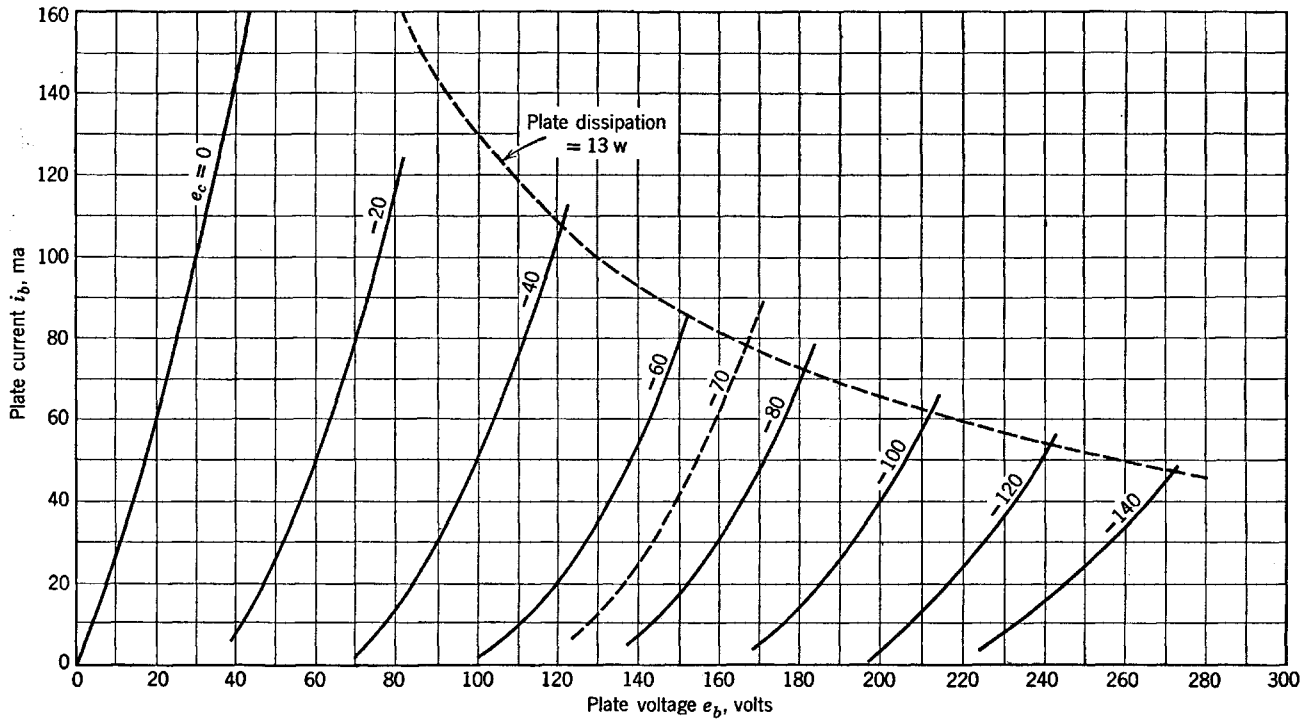
Fig. B.5. Type IV triode plate curves.
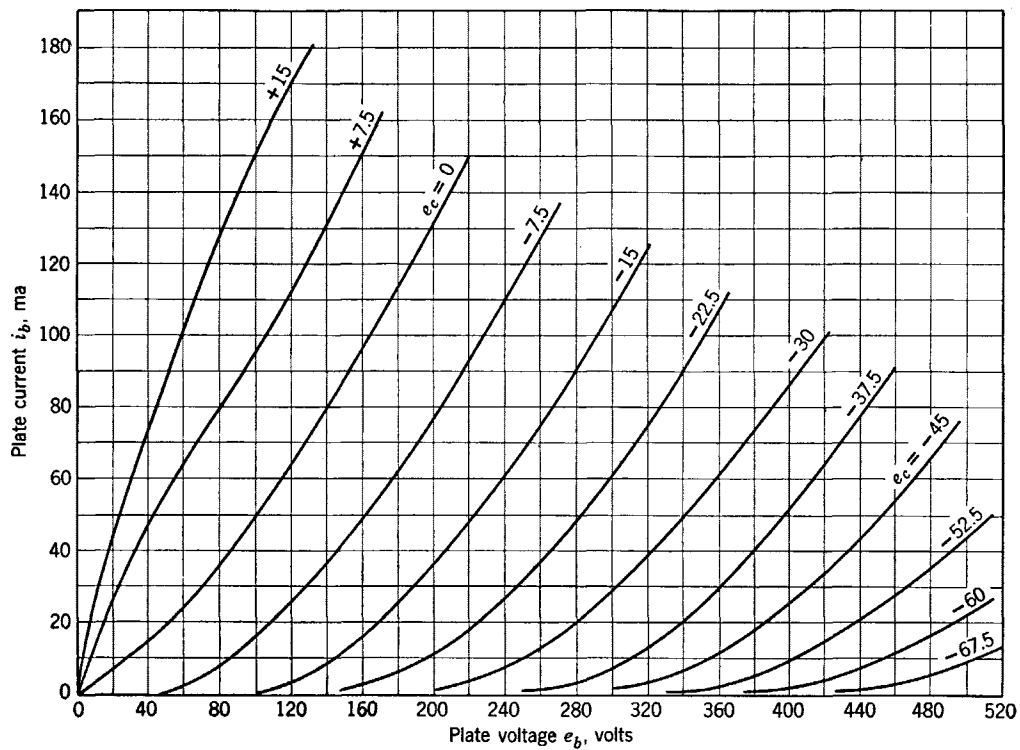
Fig. B.6. Type V triode plate curves.

Pentode Curves
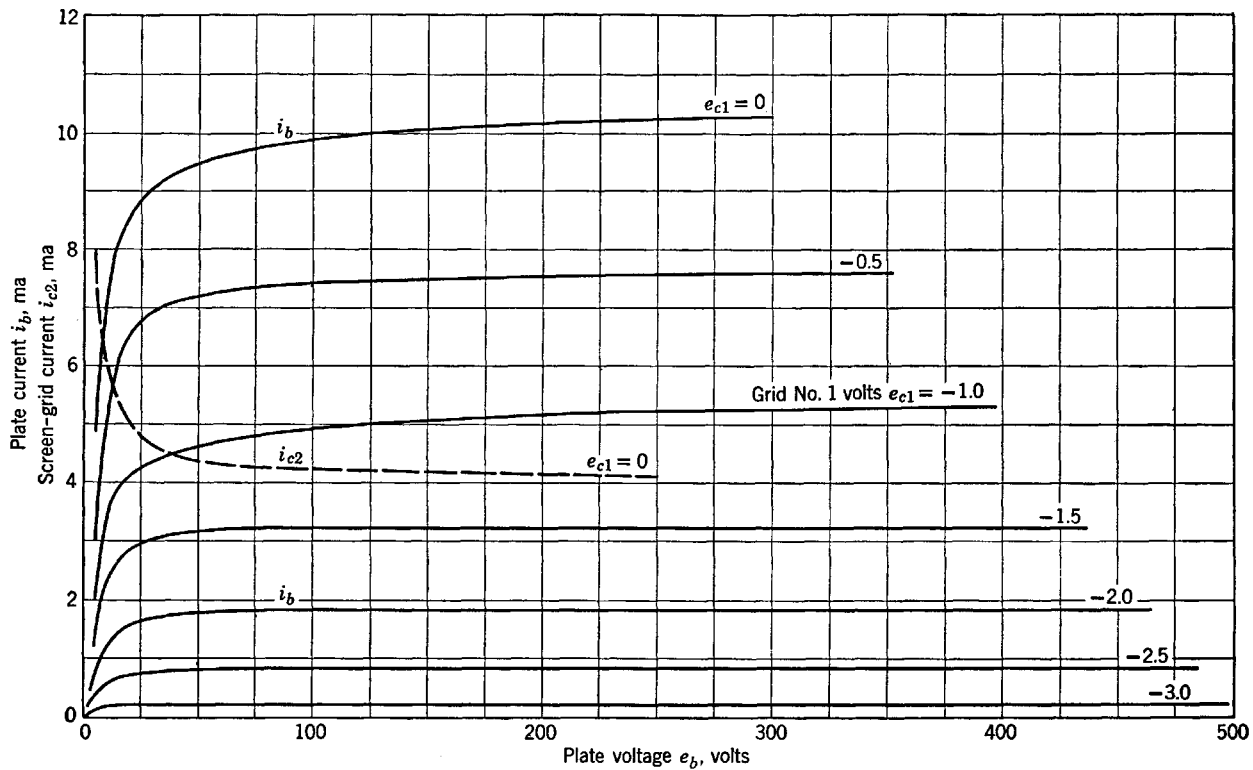
Fig. C.1. Type I pentode plate curves. (Screen-grid voltage $e_{c2} = 100$ volts).
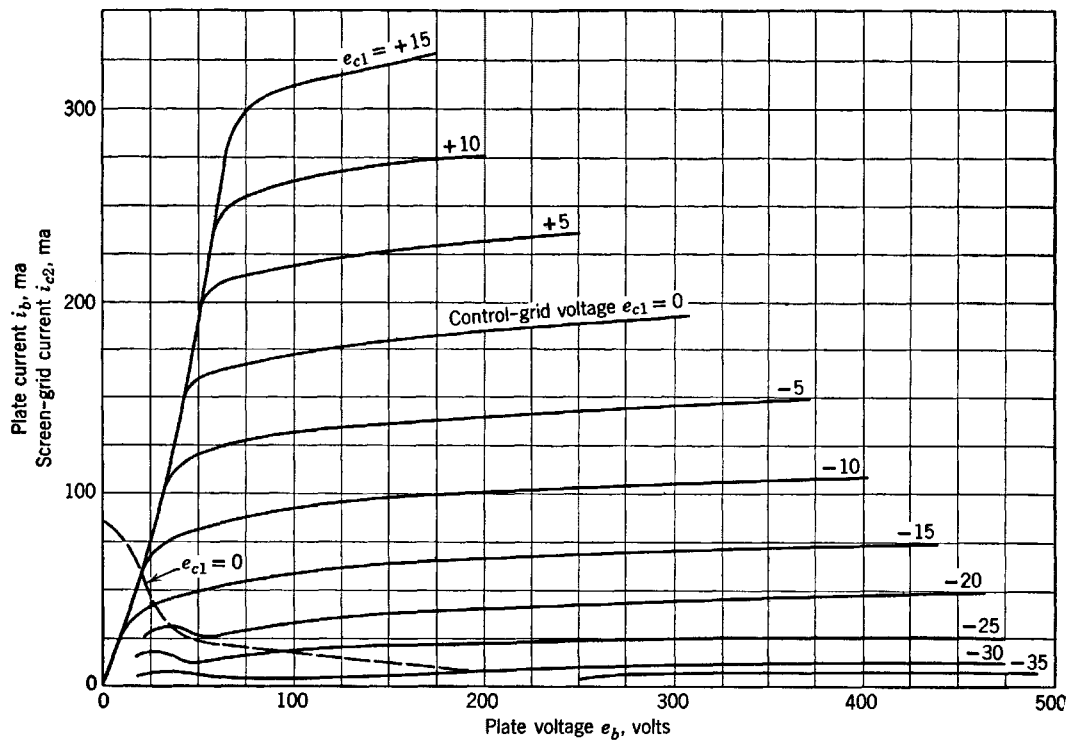
Fig. C.2. Type II pentode plate curves. (Screen-grid voltage $e_{c2} = 250$ volts.)

Exponential Curves
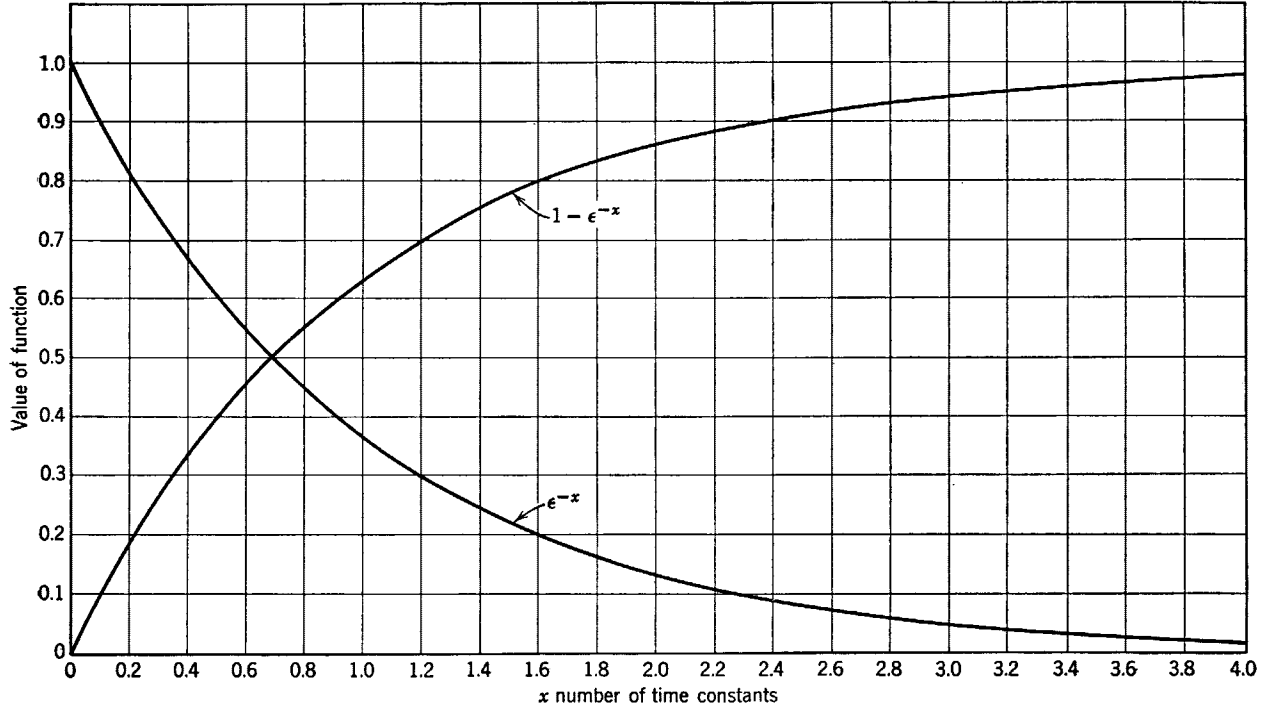
Fig. D.1. Exponential functions ϵ^{-x} and $1 - \epsilon^{-x}$.

$$x = \frac{t}{RC} \text{ in } RC \text{ circuits}$$

$$x = \frac{Rt}{L} \text{ in } RL \text{ circuits}$$

I N D E X