E. J.
ANGELO

Electronic Circuits

McGRAW-HILL

McGRAW-HILL
ELECTRICAL AND
ELECTRONIC
ENGINEERING
SERIES

BROOKLYN
POLYTECHNIC
INSTITUTE
SERIES

ELECTRONIC CIRCUITS

A Unified Treatment of Vacuum Tubes and Transistors

McGRAW-HILL ELECTRICAL AND ELECTRONIC ENGINEERING SERIES

AHRENDT AND SAVANT · Servomechanism Practice
ANGELO · Electronic Circuits
ASELTINE · Transform Method in Linear System Analysis
BAILEY AND GAULT · Alternating-current Machinery
BERANEK · Acoustics
BRENNER AND JAVID · Analysis of Electric Circuits
BRUNS AND SAUNDERS · Analysis of Feedback Control Systems
CAGE · Theory and Application of Industrial Electronics
CAUER · Synthesis of Linear Communication Networks
CHIRLIAN AND ZEMANIAN · Electronics
CLEMENT AND JOHNSON · Electrical Engineering Science
COTE AND OAKES · Linear Vacuum-tube and Transistor Circuits
CUCCIA · Harmonics, Sidebands, and Transients in Communication Engineering
CUNNINGHAM · Introduction to Nonlinear Analysis
EASTMAN · Fundamentals of Vacuum Tubes
EVANS · Control-system Dynamics
FEINSTEIN · Foundations of Information Theory
FITZGERALD AND HIGGINBOTHAM · Basic Electrical Engineering
FITZGERALD AND KINGSLEY · Electric Machinery
FRANK · Electrical Measurement Analysis
FRIEDLAND, WING, AND ASH · Principles of Linear Networks
GEPPERT · Basic Electron Tubes
GLASFORD · Fundamentals of Television Engineering
GREINER · Semiconductor Devices and Applications
HAMMOND · Electrical Engineering
HANCOCK · An Introduction to the Principles of Communication Theory
HAPPELL AND HESSELBERTH · Engineering Electronics
HARMAN · Fundamentals of Electronic Motion
HARRINGTON · Introduction to Electromagnetic Engineering
HARRINGTON · Time-harmonic Electromagnetic Fields
HAYT · Engineering Electromagnetics
HILL · Electronics in Engineering
JOHNSON · Transmission Lines and Networks
KOENIG AND BLACKWELL · Electromechanical System Theory
KRAUS · Antennas
KRAUS · Electromagnetics
KUH AND PEDERSON · Principles of Circuit Synthesis
LEDLEY · Digital Computer and Control Engineering
LEPAGE · Analysis of Alternating-current Circuits
LEPAGE · Complex Variables and the Laplace Transform for Engineers

LePage and Seely · General Network Analysis
Ley, Lutz, and Rehberg · Linear Circuit Analysis
Linvill and Gibbons · Transistors and Active Circuits
Lynch and Truxal · Introductory System Analysis
Millman · Vacuum-tube and Semiconductor Electronics
Millman and Seely · Electronics
Millman and Taub · Pulse and Digital Circuits
Mishkin and Braun · Adaptive Control Systems
Moore · Traveling-wave Engineering
Pettit · Electronic Switching, Timing, and Pulse Circuits
Pettit and McWhorter · Electronic Amplifier Circuits
Pfeiffer · Linear Systems Analysis
Reza · An Introduction to Information Theory
Reza and Seely · Modern Network Analysis
Rogers · Introduction to Electric Fields
Rudenberg · Transient Performance of Electric Power Systems
Ryder · Engineering Electronics
Schwartz · Information Transmission, Modulation, and Noise
Seely · Electrical Engineering
Seely · Electron-tube Circuits
Seely · Introduction to Electromagnetic Fields
Seely · Radio Electronics
Seifert and Steeg · Control Systems Engineering
Siskind · Direct-current Machinery
Skilling · Electronic Transmission Lines
Skilling · Transient Electric Currents
Spangenberg · Fundamentals of Electron Devices
Spangenberg · Vacuum Tubes
Stevenson · Elements of Power System Analysis
Stewart · Fundamentals of Signal Theory
Storer · Passive Network Synthesis
Strauss · Wave Generation and Shaping
Terman · Electronic and Radio Engineering
Terman and Pettit · Electronic Measurements
Thaler · Elements of Servomechanism Theory
Thaler and Brown · Analysis and Design of Feedback Control Systems
Thompson · Alternating-current and Transient Circuit Analysis
Tou · Digital and Sampled-data Control Systems
Truxal · Automatic Feedback Control System Synthesis
Valdes · The Physical Theory of Transistors
Williams and Young · Electrical Engineering Problems

BROOKLYN POLYTECHNIC INSTITUTE SERIES

ANGELO · Electronic Circuits
MISHKIN AND BRAUN · Adaptive Control Systems
SCHWARTZ · Information Transmission, Modulation, and Noise
STRAUSS · Wave Generation and Shaping

ELECTRONIC CIRCUITS

A Unified Treatment of Vacuum Tubes and Transistors

E. J. ANGELO, Jr.

Professor of Electrical Engineering
Polytechnic Institute of Brooklyn

McGRAW-HILL BOOK COMPANY, INC.

New York Toronto London

1958

ELECTRONIC CIRCUITS

 Library of Congress Catalog Card Number 58-8837

V

01790

THE MAPLE PRESS COMPANY, YORK, PA.

FOREWORD

The rate of scientific discoveries and important engineering applications during and following World War II has presented a peculiar challenge to electrical engineering education. Though originally rooted in the physical discoveries of Faraday and the embracing theory of Maxwell, electrical engineering had for many years been absorbed in narrow and specialized areas of this potentially most powerful field of engineering endeavor. The linkage of network theory with the mathematical field of functions of a complex variable, the successful monitoring of electron flow in solids aided by the judicious implantation of impurities, and the perception of the broad principles of controlled feedback for automatic regulation of processes have enlarged the scope of electrical engineering to the point where a fundamental reorientation of the entire undergraduate curriculum appears necessary.

In fairness to students as well as to instructors, it is important to incorporate the practically inescapable emphasis on sound and rigorous engineering science—as contrasted with the basic sciences of physics, chemistry, and mathematics—in measured steps and with careful evaluation of the learning capacity on the one hand and the ability for constructive use on the other hand. We have thus embarked at the Polytechnic Institute of Brooklyn on a comprehensive program of revision of all the courses in the electrical engineering program. Upon discussion with McGraw-Hill representatives, we conceived of a series of basic textbooks portraying the essential concepts of the evolutionary process rather than detailed design procedures still in the process of maturation. This series will carry in the title the name of the Institute and will cover the fields of electronic circuits, communication theory, electromechanical transduction, digital techniques, feedback principles, and others, jointly comprising the essence of the undergraduate course program in electrical engineering.

The present volume by Professor E. J. Angelo, Jr. is the first tangible result of this long-range planning started more than eight years ago. It contains all the elements of the basic approach we have chosen. We hope that it can serve as catalyst as well as a helpful guide for many who have accepted the challenge of the new developments in technology, and in particular the trend toward scientific engineering.

Ernst Weber

PREFACE

The study of electronics is currently considered to include studies of physical electronics, solid-state physics, linear amplifiers, nonlinear amplifiers, pulse circuits, rectifiers, and a variety of other related topics. Any attempt to cover all these topics in detail in a single volume must inevitably result in a large and unwieldy book. Partly for the purpose of avoiding such a result, this book does not undertake a thorough development of all these subjects; it does, however, attempt to establish the fundamental concepts and techniques that are basic to all of them. The subject matter of the book has served for several years as the basis for a one-year first course in electronics for juniors in electrical engineering. The important topics not treated in detail in this course are studied in a subsequent course.

The first half of the book is concerned primarily with the development of linear and piecewise-linear circuit characterizations for tubes and transistors and with examining the behavior of these devices in the basic amplifier configurations; thus it is concerned with the properties of active devices and with circuit representations for such devices. The techniques employed are quite general and are used in a subsequent course to obtain circuit representations for mechanical, electromechanical, and hydraulic devices. The second half of the book is concerned almost solely with linear tube and transistor circuits; thus it is an introduction to active-circuit theory. This study is closely correlated with the study of passive-circuit theory; in fact, it is an extension of passive-circuit theory to include active circuits. The methods employed in characterizing the active devices make it both feasible and desirable to treat tubes and transistors simultaneously. As implied above, these two devices are seen to be special cases of a large class of amplifying devices.

Physical electronics and solid-state physics are presented in just enough detail to give the student some understanding of the properties of the devices and to acquaint him with the principal factors limiting the performance that can be obtained. It is believed that the student is better motivated and better equipped for a detailed study of these subjects after he has studied the applications of the devices and has learned of the annoying limitations on their performance. Similarly, nonlinear electronic circuits are treated only in an introductory manner in this book; thus power amplifiers, pulse circuits, modulators, and related circuits are not presented in detail.

Even with this restricted scope, the book contains more material than can be covered comfortably in a one-year course. Thus it is appropriate to mention certain sections that can be omitted without eliminating material that is prerequisite to later portions of the text. Chapters 11 and 16 can be omitted entirely; however, if Chap. 16 is omitted it may be necessary to present the Nyquist test of Chap. 17 without proof. The following sections can also be omitted: 3-7, 3-8, 13-5, 15-8, 15-9, and 15-10. In addition, various sections in Chap. 2 can be omitted in accordance with the desires of the instructor.

The author is indebted to many people among his colleagues and his students for their contributions to this work. The course for which the book was developed has been taught by more than 20 different instructors, and each of these has aided the development of the subject matter in one way or another. Special acknowledgment is due Professor Athanasios Papoulis for the resonant-peaking circle of Sec. 15-7 and for the central features in the analysis of the double-tuned amplifier.

E. J. Angelo, Jr.

CONTENTS

CHAPTER 1

INTRODUCTION

The manner in which electronic circuits are presented in this book is the result of an effort to unify and systematize the study of such circuits through the use of some of the elementary but powerful techniques of modern network theory. The result is a somewhat unconventional treatment of the subject matter. This introductory chapter presents a qualitative discussion of the principal objectives of the book.

1-1. Basic Premises. The variety of circuits using electronic devices is so great that it is not possible, because of the limited time available, to discuss each of them separately. Moreover, the study of electronics on such a basis is unsatisfactory in many respects, for it tends to become the study of a large collection of more or less unrelated circuits. A more effective basis for the study is provided by certain fundamental concepts and analytical techniques that are applicable to large classes of circuits and that therefore bring unity into the study.

The unifying concepts that are the principal concern of this book apply equally well to transistor and vacuum-tube circuits; hence it is feasible to study these two types of circuits simultaneously. In fact, these concepts are fundamental to linear networks in general; hence they are currently used in the analysis of mechanical, electromechanical, hydraulic, and thermal systems as well as electronic circuits. Of particular importance is the fact that these concepts provide a kind of insight and understanding that is especially useful in the design of electronic circuits.

Encyclopedias of electronic circuits and handbooks of factual information have important uses in engineering work; however, as a rule they do not provide the best guidance for an organized basic study. A textbook should concentrate on the reasoning processes used in analysis and design; when the reasoning processes are mastered, no special tutoring is needed to utilize the factual information contained in handbooks.

The belief that learning proceeds from familiar things and specific cases to new things and general cases dictates that new concepts be introduced in terms of specific and, wherever possible, familiar circuits. This practice has, in addition, the desirable properties of motivating the study and of relating theoretical developments to engineering applications. A body of powerful theory has questionable value to the engineer (in contrast with the mathematician) unless he can interpret it in terms of a physical system. This is not to say that an elegant theory must be defiled by a diluted expression in practical terms; however, the physicist and the engineer must go one step beyond the mathematician and relate their theories to physical phenomena. This, in fact, is the chief function of the physicist. The chief function of the engineer is to use these relations to design systems which technicians then build. Most engineers are part physicist and part technician, with altogether too little true mathematical competence (as distinct from skill in manipulation and computation).

In beginning the study of a new concept, the student is confronted with two distinct difficulties, neither of which is necessarily serious. The first of these is the difficulty of grasping the concept and understanding its significance; the second is associated with the computational effort required in the quantitative application of the concept. It is essential that these two difficulties be separated; in particular, the purely computational difficulties must not be permitted to obscure the concept. As a general rule, the properly prepared student has less trouble with the conceptual difficulties than with those of computational origin, even though the new concept be quite abstract, provided that the concept is clearly related to other concepts that are familiar from previous studies or experiences. It is not necessary that these familiar concepts be practical or physical; they can be completely abstract so long as they are familiar. Thus a competent electrical-engineering student can in a very few weeks become proficient in analyzing the dynamics of linear mechanical systems, provided that he relates the mechanical system, which is completely physical, to its electrical analog, which is quite abstract in its electrical properties. If the new concept is truly fundamental, however, the student is not likely to perceive all its implications at once. Computational difficulties should be avoided whenever possible by judicious choice of the method of analysis; however, in many cases laborious computations cannot be avoided. Here the human propensity for error takes a heavy toll. The principal cause of errors is not a lack of knowledge, but carelessness. Nevertheless, the engineer must learn to face computational difficulties and to carry out lengthy calculations with accuracy. In this connection, a thorough understanding of the concepts involved and a clear picture of the fundamental relations do much to relieve the tedium of computation, to expose computational errors, and to reveal the most efficient computational procedure.

For reasons stated above, problems intended to illustrate a basic principle should be clearly formulated, and the principle should not be obscured by computational difficulties or vague design consideration.

Such problems may, however, be expressed in terms of carefully chosen practical applications for the purpose of motivation and to provide factual information. Having mastered the principle, the student should advance to problems involving more computational effort and including some aspects of design.

1-2. Analysis and Design. The problem of calculating the performance characteristics of a specified circuit is the problem of analysis. In the case of linear circuits, both electronic and otherwise, there are systematic methods for solving any analysis problem. These methods of solution can be learned by rote, and with their aid the solution of analysis problems does not necessarily require any imagination. Nevertheless, as is shown in the following paragraph, analysis is fundamental in the more challenging problem of design, and in this connection it is often desirable to discard the systematic methods of analysis mentioned above.

The problem of specifying in detail a circuit to meet a given set of performance specifications is the problem of design. In general, there is no systematic method for solving this problem, and usually there are many different solutions to any given design problem. (This latter fact means that more conditions could have been included in the specifications.) Accordingly, the solution of the design problem requires imagination and is therefore stimulating to the mind. Systems are usually designed by a cut-and-try procedure. First, a circuit is chosen tentatively on the basis of factual knowledge and past experience. This circuit is then analyzed, and its performance characteristics are compared with the specifications that must be met. The comparison often shows that the original circuit must be modified in some way, or it may even show that the choice of a different basic circuit is necessary. After these changes are made, the circuit is again analyzed, and the results are compared with the specifications. This process is repeated until a final design is selected. (Occasionally it is necessary to terminate the process before the optimum design is reached in order to meet a production deadline.) Thus efficiency in design depends upon a good store of useful factual information to serve as a guide in making the initial choice and upon skill in analysis to permit evaluation of the choice with a minimum of labor. Of special importance is the *kind* of factual information used to characterize various circuits, for this consideration governs to a large extent the effectiveness of the initial choice of a circuit and the facility with which the initial choice can be modified to obtain improved performance.

In view of these facts, analysis problems can be divided into two classes. In the first class, the objective is to determine some specific quantity such as the number of volts appearing at the output terminals of a given circuit. In this case the method of analysis is chosen to provide the answer with the least effort and with the smallest probability of numerical error; elementary systematic methods are usually satisfactory for the solution of this problem. In the second class of analysis problem, the objective is to examine the general properties of the circuit and to obtain factual information about these properties that will subsequently be useful in circuit design. In this case the method of analysis should be chosen to give the greatest insight into the properties of the circuit and to present the factual information in the form that is the most useful to the designer. Elementary methods of analysis are usually not suitable for this purpose, for they usually provide little insight. On the other hand, the network theorems, for example, are particularly valuable tools in this class of analysis. In this type of analysis, the method used for characterizing the properties of the circuit is especially important. In this connection, the logarithmic frequency characteristics and the pole-zero patterns of network functions are particularly effective for characterizing the dynamic properties of circuits.

1-3. Analysis of Physical Systems. It is shown in Sec. 1-2 that the analysis of physical systems is an important part of the design of such systems and that therefore the engineer, who is ultimately concerned with design, should be a skilled analyst. The analysis of a physical system usually consists of three steps: (1) A study of the physics of the system, either theoretical or experimental, leading to a set of relations that describe the behavior of the system. These relations, which may take the form of a set of equations or a set of curves, constitute a mathematical model for the system. In the case of electrical systems, this study is usually accompanied by the formulation of an electric network model corresponding to the mathematical model. (2) The mathematical solution of the model. (3) The interpretation of the solution in terms of the physical behavior of the system.

Approximately the first half of this book is concerned primarily with the application of step 1 of the analysis procedure to vacuum-tube and semiconductor devices of engineering importance. The analysis of certain simple but basic circuits is carried through to completion, however, in order to show their characteristics and limitations. In formulating network models for physical systems it is usually necessary to make certain simplifying approximations; thus the models correspond approximately to the physical system over a restricted range of operating conditions. This is true of the models developed for vacuum tubes and transistors. For example, most of these models are valid only in a restricted range of operating voltages and frequencies. These models, which are electric-circuit representations for tubes and transistors, are purposely called models rather than equivalent circuits to emphasize the fact that they are not exactly equivalent to the physical device but that they represent the device approximately over a certain range of operating conditions.

Passive, linear, bilateral, lumped-parameter circuits are commonly characterized in terms of ideal resistance, inductance, and capacitance. However, these ideal components are not sufficient for the characterization of systems containing transistors and vacuum tubes. The study of these devices shows that in most of their important applications they can be characterized in terms of ideal R, L, and C plus two additional elements, the ideal diode and the ideal controlled source.

The second half of the book is concerned primarily with steps 2 and 3 of the analysis procedure. Most of this study is restricted to the case of small-signal operation in which the circuits are linear and can be described in terms of ideal R, L, C and controlled sources. Thus the study is concerned with the extension of passive-network theory to the analysis of circuits containing controlled sources. The methods of analysis are chosen to give the greatest insight into the important properties of the circuits and therefore to be of the greatest help in circuit design. The means used to characterize the properties of the circuits are also chosen on the basis of utility in design.

In conclusion, it should be said that for the engineer there are no problems of analysis that cannot be solved; he *must* obtain a solution by one means or another. If the mathematics becomes too cumbersome to be useful, he must make acceptable approximations to reduce the mathematical complexity. If circuits are nonlinear, he must use linear or piecewise-linear approximations, or he may be forced to use graphical methods of analysis. In many circumstances he must be prepared to use electronic computers in obtaining solutions to the more difficult problems. In the worst cases he may be forced to use empirical methods. In no case, however, should he abandon his theoretical knowledge; this knowledge provides valuable guidance in obtaining a solution by any means. When complete analytical solutions are not feasible, partial solutions usually serve as valuable complements to experimental studies.

1-4. Summary. The study of electronic circuits presented in this book has as its primary objective the development of certain fundamental, unifying concepts that give useful insight into the properties of electronic circuits and that serve as valuable guides in circuit design. The more important phases of the study can be summarized as follows: (1) an examination of the electrical characteristics of tubes and transistors; (2) the formulation of network models for tubes and transistors and the emergence of the ideal diode and the ideal controlled source as network components; (3) the extension of passive-network theory to circuits containing controlled sources; (4) the characterization of the dynamic properties of circuits in terms of the logarithmic frequency characteristics; (5) the characterization of the dynamic properties of circuits in terms of the poles and zeros of the network functions; (6) an examination of the effect of feedback on circuit properties.

A secondary objective of the study is to provide the student with a store of factual information about the more important electronic circuits. There are several reasons for presenting this information, among which are (1) motivating the study, (2) providing a nucleus of factual information to serve as a basis for circuit design, and (3) strengthening the grasp of basic concepts by expressing them quantitatively in a realistic context. The factual information is kept subordinate to the basic principles, however, and it is presented, for the most part, through the illustrative examples in the text and through the problems included at the end of each chapter.

CHAPTER 2

THE IDEAL DIODE

It is explained in Chap. 1 that the set of ideal elements consisting of resistors, capacitors, inductors, and constant current or voltage sources is not adequate for the representation of the electrical properties of electronic circuits. However, the addition to the set of two new elements, the ideal diode and the ideal controlled source, removes this limitation and makes possible the network representation of most of the important electronic circuits. The objective of this chapter is to introduce the ideal diode as a circuit element and to investigate some of the useful things that can be done with its aid.

The ideal diode is a nonlinear, two-terminal element; it is used to represent the nonlinear properties of electronic devices such as vacuum tubes and transistors. Circuits containing diodes are therefore nonlinear, and for this reason they can perform operations that cannot be performed by circuits containing linear R's, L's, and C's alone. Thus a wide range of new and useful circuits is made possible by the diode. Some of these circuits are analyzed in the pages that follow.

2-1. Characteristics of the Ideal Diode. The ideal diode is represented symbolically in Fig. 2-1*a*; its volt-ampere characteristic is given by the heavy lines in Fig. 2-1*b*. This characteristic shows that the ideal diode behaves in the following way. When the diode current i_d is positive, the voltage drop e_d across the diode is zero, regardless of the magnitude of the current; that is, the diode behaves as a short circuit to current in the forward direction. On the other hand, when the voltage drop across the diode is negative, the current through the diode is zero, regardless of the magnitude of the voltage; that is, the diode behaves as an open circuit to voltage in the reverse direction. The diode changes

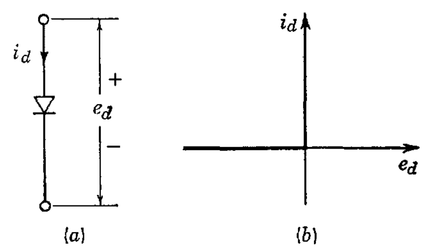


FIG. 2-1. The ideal diode. (*a*) Symbol; (*b*) volt-ampere characteristic.

from the open-circuit to the short-circuit condition at the point where both the current through the diode and the voltage drop across it are zero. Of course, no physical device has exactly these characteristics; however, a number of devices have characteristics that approximate the ideal very closely.

The upper terminal of the diode in Fig. 2-1a is called the anode terminal because positive charge flows into the diode at that terminal. The lower terminal is called the cathode terminal because positive charge flows out of the diode at that terminal.

The most important feature of the ideal diode is the fact that its volt-ampere characteristic is not a straight line; it consists of two straight-line segments joined at right angles. The ideal diode is therefore a nonlinear device, and for this reason it can produce results that cannot be obtained with the linear elements R, L, and C. The importance of the diode lies in this fact.

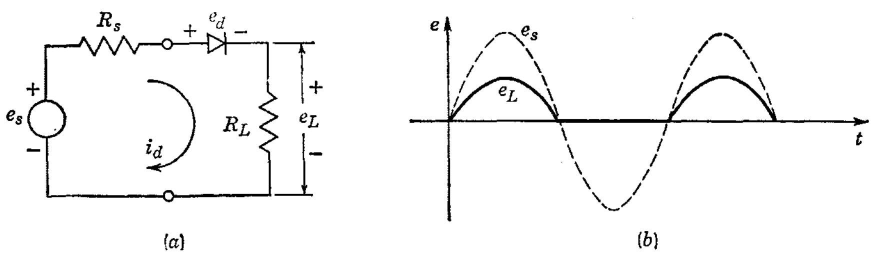


FIG. 2-2. A diode rectifier circuit with supply voltage $e_s = E_s \sin \omega_s t$. ($a$) Circuit; ($b$) waveforms.

Devices having volt-ampere characteristics consisting of straight-line segments are called piecewise-linear devices because they are linear along each separate piece of the characteristic. If the combinations of current and voltage are restricted to values lying entirely on one piece of the characteristic, the device behaves as a linear device. Its nonlinear nature comes into consideration only if the ranges of current and voltage involved extend across a break point in the characteristic.

2-2. The Half-wave Rectifier. A resistive half-wave rectifier circuit is shown in Fig. 2-2a. The voltage source e_s supplies a sinusoidal voltage to the circuit; R_L represents a resistive load being supplied with power, and R_s represents the source resistance. When e_s is positive it produces a current in the positive direction indicated in the diagram. Since this is the forward direction for the diode, it acts as a short circuit, and the magnitude of the current is determined by e_s, R_s, and R_L. When e_s is negative it acts to produce a current in the opposite direction. However, since this is the reverse direction for the diode, it behaves as an open circuit, and no current flows. The waveforms of supply voltage and load voltage, $e_L = R_L i_d$, are shown in Fig. 2-2*b*. Under these conditions the load voltage is

$$e_L = \frac{R_L}{R_s + R_L} e_s \qquad \text{for } e_s \text{ positive} \tag{2-1}$$

$$e_L = 0 \qquad \text{for } e_s \text{ negative} \tag{2-2}$$

The load voltage pictured in Fig. 2-2*b* is periodic; therefore it can be represented by a Fourier series. If the peak instantaneous value of e_L is designated by E_L, the series is

$$e_L = \frac{1}{\pi} E_L + \frac{1}{2} E_L \sin \omega_s t = \frac{2}{3\pi} E_L \cos 2\omega_s t - \frac{2}{15\pi} E_L \cos 4\omega_s t + \cdots \tag{2-3}$$

Thus e_L consists of a d-c component and sinusoidal components at the fundamental radian frequency ω_s and at integer multiples of ω_s. The voltage across the load therefore contains components at frequencies not

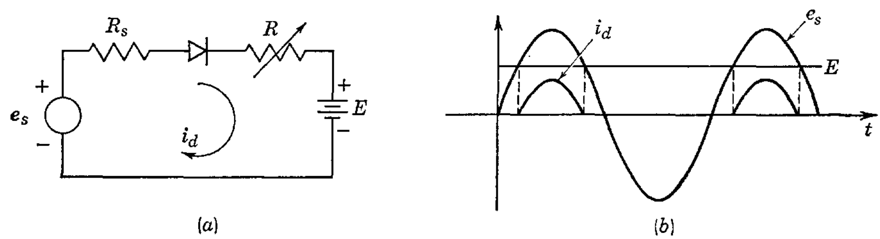


FIG. 2-3. A battery charger with supply voltage $e_s = E_s \sin \omega_s t$. (*a*) Circuit; (*b*) waveforms.

present in the voltage applied to the circuit. The appearance of these new frequencies is a consequence of the nonlinearity of the diode. In circuits consisting entirely of linear elements the only frequencies appearing are those present in the applied voltages and currents. One of the primary uses of the diode is the production of these new frequencies; often the d-c component generated by the diode is the quantity of interest. The action by which a diode generates a direct voltage from an alternating supply voltage is called rectification.

The rectifier circuit shown in Fig. 2-3*a* is a battery charger. In order for charge to be accumulated in the battery it is necessary that i_d have some positive average value; that is, i_d must have a positive d-c component. The action of the diode in generating a d-c component is therefore essential in the charging of a battery from an a-c supply. The action of the circuit can be understood from the waveforms shown in Fig. 2-3*b*. At those instants when the curve of e_s lies below E (e_s less than E) the net voltage around the loop acts to send current in the reverse direction through the diode. At such instants the diode therefore acts as an open circuit, and no current flows. At those instants when the curve of e_s

is above E (e_s greater than E) the net voltage acts to send current in the forward direction through the diode, the diode acts as a short circuit, and current flows in the proper direction to charge the battery. The variable resistance R can be adjusted to give the desired value of charging current. The charging current at each instant is given by

$$i_d = \frac{E_s \sin \omega_s t - E}{R_s + R} \qquad \text{for } e_s > E \tag{2-4}$$

$$i_d = 0 \qquad \text{for } e_s < E \tag{2-5}$$

An ideal diode can conduct any value of current in the forward direction, and it remains an open circuit for all values of inverse voltage. This is not the case, however, with physical diodes. In general, there is a limit to the peak instantaneous current that can be passed without damage, and there is a limit to the peak instantaneous inverse voltage that can be applied. It is therefore necessary to ensure that these two quantities do not exceed the maximum permissible values.

No current flows in the circuit of Fig. 2-2*a* when an inverse voltage exists across the diode; hence there is no voltage drop across the circuit resistances, and the peak inverse voltage is simply E_s. The peak forward current occurs when e_s has its maximum positive value; the magnitude of this current is $(i_d)_{\max} = E_s/(R_s + R_L)$. Similarly, there is no current in the circuit of Fig. 2-3*a* when an inverse voltage exists across the diode; hence the peak inverse voltage occurs when $e_s = -E_s$ (Fig. 2-3*b*), and its value is $E_{pi} = E_s + E$. The peak current occurs when e_s has its maximum positive value, and it is given by $(i_d)_{\max} = (E_s - E)/(R_s + R)$.

It is clear from Fig. 2-3*b* that the diode in the battery-charging rectifier conducts for somewhat less than half the cycle of e_s. If E_s is only slightly larger than E, the diode will conduct for only a very small fraction of the time. Under these conditions the average current, which measures the charging rate, will be much smaller than the peak current, which must be limited to a safe value. This operating condition corresponds to poor utilization of the diode.

The calculation of instantaneous and average power in circuits in which all currents and voltages are d-c or are sinusoids of the same frequency is a relatively simple matter. In circuits such as those discussed above the currents and voltages have complex waveforms, and computations of power must be made with care. In particular, it is usually desirable to base power calculations on the fundamental law

$$p = ei \tag{2-6}$$

If e is the instantaneous potential difference in volts across the terminals of any two-terminal device, and if i is the current in amperes into the positive terminal at the same instant, then p is the power in watts delivered to the pair of terminals at that instant. This relation is always true; it comes directly from the definitions of potential difference and current and is independent of the circuit. If e and i are time-varying quantities, p is a time-varying quantity. The average value of p over any interval T is given by

$$p_{av} = P = \frac{1}{T}\int_0^T ei\,dt \tag{2-7}$$

To find P this integral must be evaluated. If analytic expressions for e and i as functions of time are known, then it may be possible to evaluate the integral analytically. If analytic expressions for e and i are not known, or if they are too complicated for easy integration, the integral can be evaluated by approximate methods such as Simpson's rule or by counting squares on a graphical plot. In any case it must be stated most emphatically that, in general,

$$P \neq e_{av}i_{av} \tag{2-8}$$

In other words, the average of the product of any two time-varying quantities x and y *is not,* in general, the product of the averages of x and y; however, it reduces to the product of the averages in the special case that either x or y or both are constant in the integration.

Example 2-1. In a diode rectifier circuit like that shown in Fig. 2-2 the supply voltage is $e_s = 475 \sin 377t$ volts, the load resistance is $R_L = 1000$ ohms, and the source resistance is $R_s = 500$ ohms. Determine (*a*) the d-c component of the load voltage, (*b*) the peak instantaneous diode current, (*c*) the peak inverse voltage across the diode, and (*d*) the average power absorbed by the load.

Solution. *a.* During the negative half cycle of e_s the diode acts as an open circuit and $e_L = 0$. During the positive half cycle of e_s the diode acts as a short circuit, and the load voltage is

$$e_L = \frac{R_L}{R_s + R_L}\,e_s = \frac{1000}{1500}\,475 \sin\theta = 316 \sin\theta \qquad 0 < \theta < \pi$$

The d-c component of e_L is the average value of the voltage over a full cycle of operation:

$$E_{dc} = \frac{1}{2\pi}\int_0^{2\pi} e_L\,d\theta = \frac{1}{2\pi}\int_0^{\pi} 316 \sin\theta\,d\theta = \frac{316}{\pi} \approx 100 \text{ volts}$$

b. The peak diode current occurs when e_s has its maximum positive value; hence

$$(i_d)_{max} = \frac{(e_s)_{max}}{R_s + R_L} = \frac{475}{1500} = 0.316 \text{ amp}$$

c The peak inverse voltage across the diode occurs when e_s has its maximum negative value; hence

$$E_{pi} = (e_s)_{max} = 475 \text{ volts}$$

d. During the negative half cycle of e_s the load absorbs no power. During the positive half cycle of e_s the power absorbed by the load at each instant is

$$p_L = \frac{e_L^2}{R_L} = \frac{316^2}{1000}\sin^2\theta = 100\sin^2\theta$$

The average power absorbed in one full cycle is

$$P_L = \frac{1}{2\pi}\int_0^{2\pi} p_L\,d\theta = \frac{100}{2\pi}\int_0^{\pi}\sin^2\theta\,d\theta = 50\frac{1}{\pi}\int_0^{\pi}\sin^2\theta\,d\theta$$
$$= 50\,(\sin^2\theta)_{\text{av}} = 50[\tfrac{1}{2}(1 - \cos 2\theta)_{\text{av}}]$$

But the average value of $\cos 2\theta$ in the interval $0 < \theta < \pi$ is zero. Hence

$$P_L = 25 \text{ watts}$$

Example 2-2. The supply voltage in the battery charger of Fig. 2-3 is $e_s = 10 \sin 377t$ volts, the battery voltage is $E = 7$ volts, and the total series resistance is $R + R_s = 1$ ohm. Determine (*a*) the charge delivered to the battery in 1 hr, (*b*) the peak diode current, and (*c*) the peak inverse voltage across the diode.

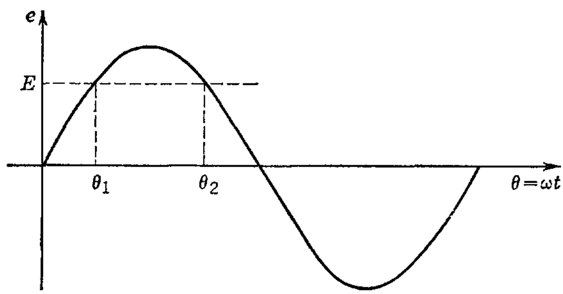


FIG. 2-4. Waveform for the battery charger of Example 2-2.

Solution. *a.* The charge delivered to the battery in 1 hr is

$$Q = 3600(i_d)_{\text{av}}$$

When e_s is less than E, the diode current is zero. When e_s is greater than E (the interval $\theta_1 < \theta < \theta_2$ in Fig. 2-4), the diode current is

$$i_d = \frac{e_s - E}{R_s + R}$$

The average diode current over a complete cycle is thus

$$(i_d)_{\text{av}} = \frac{1}{2\pi}\int_{\theta_1}^{\theta_2}\left[\frac{E_s \sin(\theta) - E}{R_s + R}\right]d\theta$$

Integrating and substituting the numerical value of $R_s + R$ yields

$$(i_d)_{\text{av}} = \frac{1}{2\pi}[-E_s\cos(\theta) - E\theta]_{\theta_1}^{\theta_2}$$

But $\theta_2 = \pi - \theta_1$. Hence

$$(i_d)_{\text{av}} = \frac{1}{2\pi}[-E_s\cos(\pi - \theta_1) - E(\pi - \theta_1) + E_s\cos(\theta_1) + E\theta_1]$$
$$= \frac{1}{2\pi}[2E_s\cos(\theta_1) + 2E\theta_1 - \pi E]$$

When $\theta = \theta_1$, $e_s = E_s \sin \theta_1 = E$; hence $\theta_1 = \sin^{-1} (E/E_s) = \sin^{-1} 0.7$, and $\theta_1 =$ 0.775 radian. Substituting numerical values in the above expression yields

$$(i_d)_{av} = \frac{1}{2\pi} [20 \cos (\theta_1) + 14\theta_1 - 7\pi]$$
$$= 0.50 \text{ amp}$$

The charge delivered to the battery in 1 hr is thus

$$Q = (3600)(0.478) = 1720 \text{ coulombs}$$

b. The peak diode current occurs when e_s has its positive maximum value and is

$$(i_d)_{max} = \frac{E_s - E}{R_s + R} = \frac{10 - 7}{1} = 3 \text{ amp}$$

The ratio of peak to average diode current is thus

$$\frac{(i_d)_{max}}{(i_d)_{av}} = \frac{3}{0.478} = 6.3$$

c. The peak inverse voltage across the diode occurs when e_s has its maximum negative value and is

$$E_{pi} = E_s + E = 10 + 7 = 17 \text{ volts}$$

2-3. The Diode Limiter. The circuit shown in Fig. 2-5*a* is a diode limiter. It has the property that no matter how wide the range over

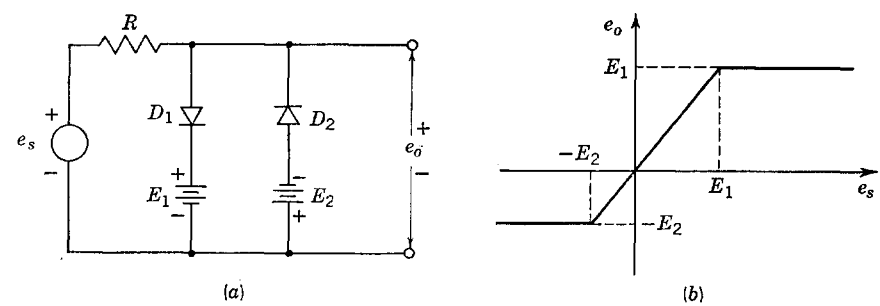
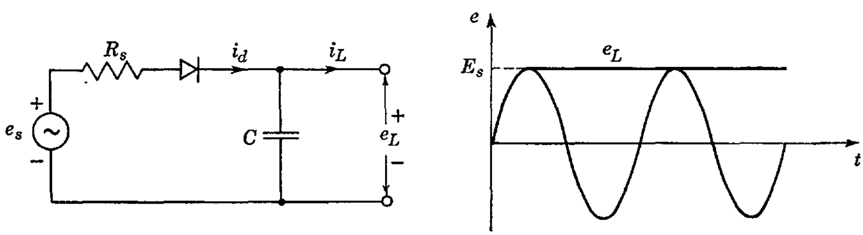


FIG. 2-5. The diode limiter. (*a*) Circuit; (*b*) transfer characteristic.

which e_s varies, the output voltage e_o is restricted to the range between the values of E_1 and $-E_2$. The voltage transfer characteristic of Fig. 2-5*b* shows how the output voltage varies as a function of the input voltage. When the input voltage is greater than E_1, diode D_1 conducts, and $e_o = E_1$. When the input voltage is less than (more negative than) $-E_2$, diode D_2 conducts, and $e_o = -E_2$. For values of e_s between these two limits both diodes are biased in the reverse direction and therefore behave as open circuits; hence in this range the output voltage varies with the input voltage as shown in Fig. 2-5*b*. If no load is connected across the output terminals, $e_o = e_s$ when e_s is in the range between $-E_2$ and E_1. The diode limiter might be used to protect a circuit against the application of excessive voltage, or it might be used to alter the waveform of the input voltage. For example, if $E_1 = E_2$, and if e_s is a sinusoid with an amplitude much larger than E_1, then the output will be very nearly a square wave of voltage.

The battery symbols are used in Fig. 2-5 to emphasize the fact that E_1 and E_2 are direct voltages; they do not imply that batteries are necessarily used to obtain these voltages. This practice will be followed throughout this book. It is to be understood that the battery symbols designate ideal sources of direct voltage.

2-4. The Peak Rectifier. The half-wave rectifier discussed in Sec. 2-2 delivers a pulsating, unidirectional current to the load. Such circuits are satisfactory for many applications such as battery charging and electrolytic processes. There are many other applications, however, where it is desired to obtain a pure direct voltage from the standard a-c



(a) (b)

FIG. 2-6. A peak rectifier with supply voltage $e_s = E_s \sin \omega_s t$. (a) Circuit; (b) waveforms.

power mains. The peak rectifier circuit shown in Fig. 2-6a is often used for this purpose when the load current i_L is small, a few milliamperes or less.

The operation of the circuit[1]* can be understood with the aid of the waveforms in Fig. 2-6b. If C initially has no charge, and if e_s is applied at time $t = 0$, then as e_s increases from zero to its positive maximum, current flows in the forward direction through the diode, and charge is stored in the capacitor. If the source resistance R_s is very small, the voltage drop across it is negligibly small, and e_L, the voltage across C, is essentially equal to e_s at every instant until e_s reaches its maximum positive value. Thus C charges to a voltage equal to the maximum positive value of e_s. Now if $i_L = 0$, there is no way for C to discharge, for the diode does not conduct current in the reverse direction. Thus the charge accumulated by C during the first quarter cycle is trapped and cannot escape. This trapped charge maintains the voltage across the capacitor at the value E_s. The circuit of Fig. 2-6a is called a peak rectifier

* Superior numbers designate references listed at the end of the chapter.

because its output voltage is equal to the positive peak value of the input voltage.

If the load on the peak rectifier consists of a large resistance R_L, as shown in Fig. 2-7a, then the capacitor can discharge slowly through R_L while the diode is not conducting. Under these conditions the load voltage and current will consist of a small ripple component superimposed upon a large d-c component. As R_L is made smaller, the capacitor discharges by a greater amount each cycle, and the ripple becomes greater

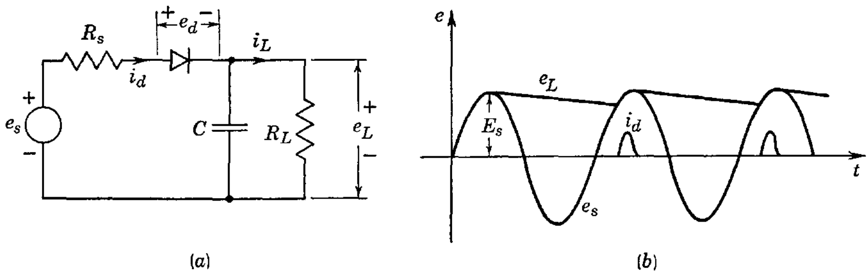


FIG. 2-7. A peak rectifier with resistive load and supply voltage $e_s = E_s \sin \omega_s t$. ($a$) Circuit; ($b$) waveforms.

in magnitude. During a brief interval near the positive peak of e_s in each cycle a pulse of charging current flows through the diode; this pulse restores to the capacitor the charge lost through R_L during the interval when the diode is not conducting.

It follows from the discussion above and from Figs. 2-3b and 2-7b that the peak rectifier acts, in some respects, like the battery charger. The charge trapped on the capacitor makes the capacitor behave somewhat like a battery. As in the case of the battery charger, current flows through the diode in short pulses, and the peak diode current is much greater than the direct component of current delivered to the load. Again, this condition represents ineffective utilization of the diode.

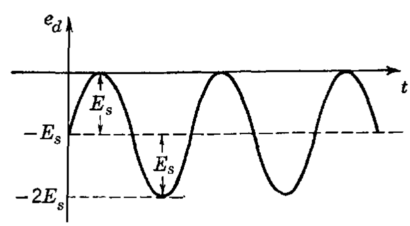


FIG. 2-8. Voltage across the diode in the peak rectifier.

If the voltage drop across R_s is negligibly small, the voltage across the diode is $e_d = e_s - e_L$. The waveform of this voltage is shown in Fig. 2-8 for the case where the ripple component of e_L is very small. It is clear from this waveform that the peak inverse voltage across the diode is $2E_s$.

2-5. The Diode Clamper. The circuit of a diode clamper[2] is shown in Fig. 2-9a. This circuit is identical with that of the peak rectifier shown in Fig. 2-6a except that the positions of the capacitor and the diode are interchanged. Thus the operation of the clamper is like that of the peak rectifier; however, the output voltage is taken from a different pair of terminals. Accordingly, if the voltage drop across R_s is negligibly small, the capacitor charges to the positive peak value of e_s as indicated in Fig. 2-9*a*, and the output voltage, which in this case is the voltage across the diode, is $e_d = e_s - E_s$. Thus the waveform of the output

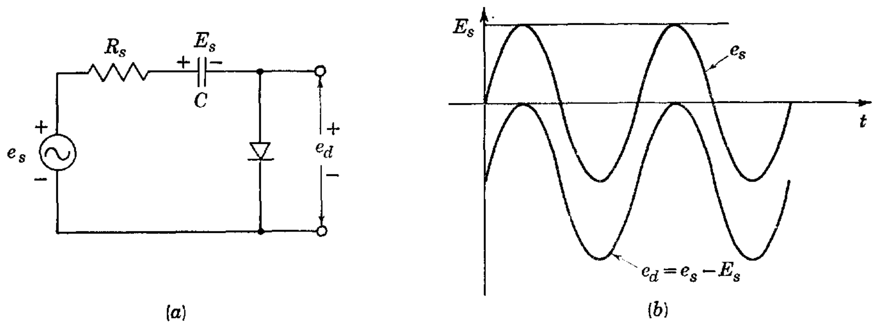


FIG. 2-9. A diode clamper with input voltage $e_s = E_s \sin \omega_s t$. (*a*) Circuit; (*b*) waveforms.

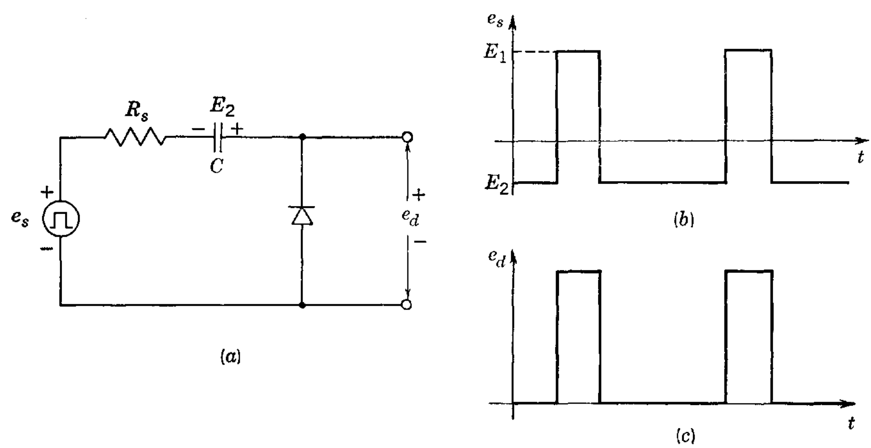


FIG. 2-10. Another diode clamper. (*a*) Circuit; (*b*) input voltage; (*c*) output voltage.

voltage is the same as the waveform of the input voltage, but the wave is shifted down by an amount equal to the positive peak value of the input voltage as illustrated in Fig. 2-9*b*. Since the output voltage rises just to the value zero when e_s has its positive peak value, the circuit is said to clamp the positive peak of e_s at zero volts.

The diode clamper has a number of useful applications. For example, in TV receivers it is necessary that the voltages appearing at certain points in the circuit have fixed peak values. Clamping circuits are used to meet this need. As another example, clamping circuits are often used in a-c vacuum-tube voltmeters; this application is discussed further in Sec. 2-6.

Another form of the diode clamper is shown in Fig. 2-10*a*. This circuit is like that of Fig. 2-9*a* except that the diode is reversed. For variety, a rectangular waveform of input voltage, shown in Fig. 2-10*b*, is assumed. When e_s is negative, the diode conducts and charges the capacitor to the maximum negative value of e_s with the polarity shown. Thus the voltage across the diode in this circuit is $e_d = e_s + E_2$, and, as shown in Fig. 2-10*c*, the waveform of the input voltage is shifted up by an amount equal to its negative peak value. This circuit clamps the negative peak of the signal waveform at zero volts.

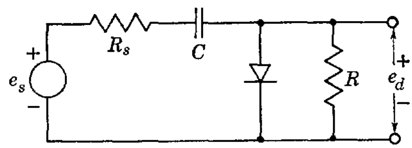


FIG. 2-11. A diode clamper with a resistive load.

It is usually necessary to connect a large resistance across the output terminals of a diode clamper as shown in Fig. 2-11 so that the charge trapped on the capacitor will not be trapped indefinitely. This resistance usually draws negligible current and does not alter the analysis given above.

2-6. The A-C Vacuum-tube Voltmeter. The a-c vacuum-tube voltmeter[2] consists basically of two parts, a diode circuit to rectify the alternating voltage and a d-c voltmeter to measure the rectified voltage. The rectifier often consists of a diode clamper and an *RC* filter forming a circuit like that shown in Fig. 2-12*a*. The resistances R_1 and R_2 are very large and draw a negligible current; hence the diode and C_1 form a diode clamper like that of Fig. 2-9*a*. If R is small, the waveform of e is clamped with its positive peak at zero, as indicated by the waveform of e_d in Fig. 2-12*b*. This voltage consists of a sinusoidal component equal to e and a d-c component equal to the positive peak value of e. The filter capacitor C_2 is chosen so that it acts as a short circuit in comparison with R_2 at the frequency of the voltage e that is being measured; C_2 acts, of course, as an open circuit to the d-c component. Hence the a-c component of e_d does not appear at the output terminals of the filter, and the voltage e_o is equal to the d-c component of e_d, which in turn is equal to the peak value of e. If the voltage e_o is measured with a d-c vacuum-tube voltmeter, the value obtained will be the positive peak value of e. Such voltmeters are called *peak-above-average* voltmeters.

Vacuum-tube voltmeters like the one described above are often calibrated to read the effective, or rms, value for sinusoidal voltages. This result is accomplished by designing the scale to indicate the peak value of the unknown voltage divided by $\sqrt{2}$. When such an instrument is used to measure nonsinusoidal voltages, the reading obtained is not the rms value of the voltage, it is simply the positive peak value of the voltage divided by $\sqrt{2}$. The peak value of the nonsinusoidal voltage, which is usually the quantity of interest, is obtained by multiplying the voltmeter reading by $\sqrt{2}$.

In the above discussion it is assumed that the resistance R in series with the voltage being measured is small. If this resistance is large, as often is the case, the current drawn by the voltmeter may cause an appreciable voltage drop across R so that the voltage at the terminals of the voltmeter is not the desired voltage. The result will be an appreciable error in the measurement unless a correction is made for the loading effect

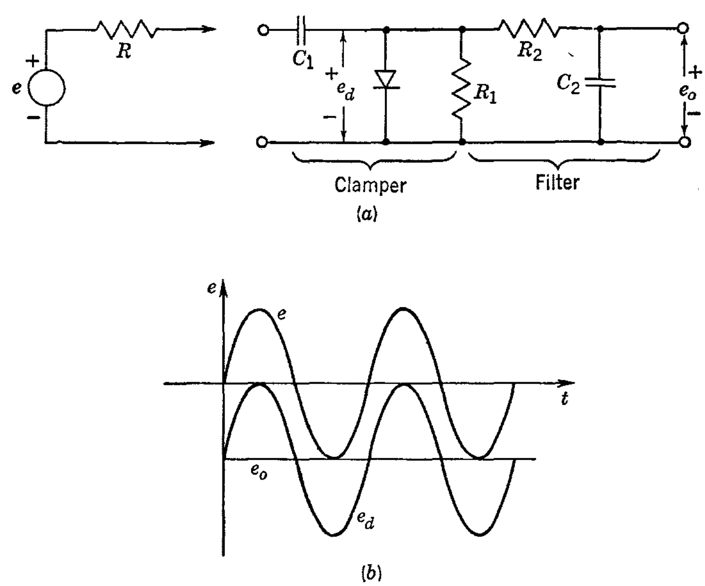


FIG. 2-12. A diode clamper with a filter. (*a*) Circuit; (*b*) waveforms.

of the instrument. For this reason it is important that the voltmeter circuit be designed to draw the least possible current from the circuit in which voltages are being measured.

2-7. The Voltage Doubler. The rectifier shown in Fig. 2-13 has the interesting and useful property that it develops a direct voltage equal to the peak-to-peak value of the input voltage.[1,4] Hence, if the input is sinusoidal, the direct voltage at the output is twice the peak value of the sinusoid. The operation of the circuit can be explained as follows. The peak rectifier consisting of D_2 and C_2 draws negligible current from the diode clamper consisting of D_1 and C_1. Therefore the clamper operates like the one shown in Fig. 2-10*a*. Capacitor C_1 charges to E_s volts with the polarity shown, and the voltage across D_1 is $e_{d1} = e_s + E_s$. The waveform of this voltage is shown in Fig. 2-13*b*; it constitutes the input voltage to the peak rectifier. The peak rectifier then charges C_2 to the peak value of e_{d1}, which is $2E_s$ as indicated in Fig. 2-13*b*. This voltage appears at the output terminals of the circuit.

By extending the basic ideas involved in the voltage doubler, diode circuits can be devised to act as voltage triplers, quadruplers, . . . , n-tuplers.[1,4] A voltage quadrupler, shown in Fig. 2-14, consists of two conventional voltage doublers. If another voltage doubler is added in the same manner, the output voltage will be $6E_s$. Circuits giving odd

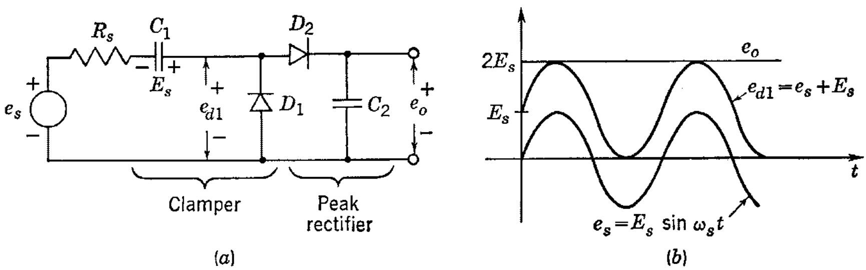


FIG. 2-13. A voltage doubler. (*a*) Circuit; (*b*) waveforms.

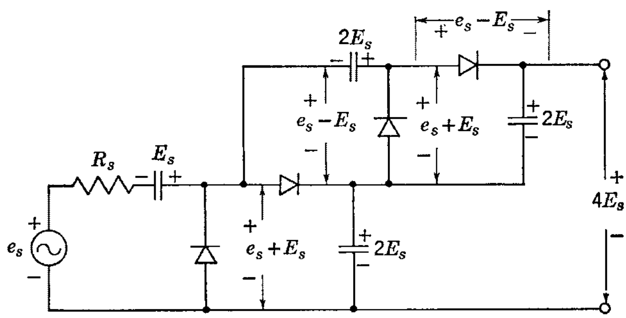


FIG. 2-14. A voltage quadrupler with supply voltage $e_s = E_s \sin \omega_s t$.

multiples of E_s are obtained by removing the capacitor and diode on the left and continuing with voltage doublers as in Fig. 2-14.

2-8. The Full-wave Rectifier. A full-wave rectifier circuit supplying power to a resistive load is shown in Fig. 2-15*a*. The circuit consists basically of two half-wave rectifiers connected to a single load resistor and supplied with sinusoidal input voltages that are equal in magnitude but opposite in phase. During the positive half cycle of e_s, D_1 acts as a short circuit, D_2 acts as an open circuit, and $e_L = e_s$. During the negative half cycle of e_s, D_2 acts as a short circuit, D_1 acts as an open circuit, and $e_L = -e_s$. The waveform of e_L is shown in Fig. 2-15*b*. The advantage of full-wave rectification over half-wave rectification in this application is that for a given peak diode current the average current in the load is twice as great. The full-wave rectifier offers additional advantages in other applications that are discussed later.

The load voltage in the circuit of Fig. 2-15 can be expressed mathematically as

$$e_L = |e_s| \tag{2-9}$$
$$= |E_s \sin \omega_s t| \tag{2-10}$$

Alternatively, since e_L is periodic it can be expressed in the form of a Fourier series:

$$e_L = E_s\left(\frac{2}{\pi} - \frac{4}{3\pi}\cos 2\omega_s t - \frac{4}{15\pi}\cos 4\omega_s t - \cdots\right) \tag{2-11}$$

Thus the load voltage consists of a d-c component of magnitude $(2/\pi)E_s$, which is the average value of each half cycle of e_s, and a set of sinusoidal

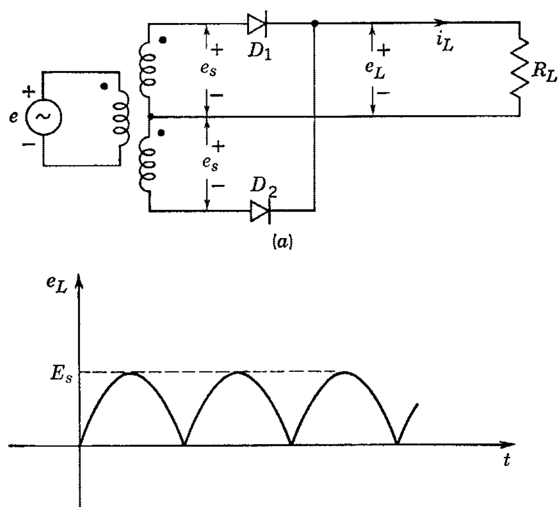


FIG. 2-15. A full-wave rectifier with supply voltage $e_s = E_s \sin \omega_s t$. (a) Circuit; (b) waveforms.

components at frequencies that are even multiples of the frequency of e_s. Thus if e_s is a 60-cycle voltage with an rms value of 115 volts, e_L will contain a d-c component of $(115)(2\sqrt{2}/\pi) = 103.5$ volts plus sinusoidal components at 120, 240, 360, . . . cps.

In the discussion above it is assumed that the center tap on the power transformer in Fig. 2-15 is exactly at the center and that the voltages across the two halves of the transformer are therefore equal. In practice these two voltages are likely to be slightly unbalanced, with the result that the successive peaks in the load-voltage waveform shown in Fig. 2-15b are not exactly equal. In this case the fundamental frequency of the periodic voltage across the load is 60 cps, and e_L contains components at 60 cps and its harmonics.

The peak inverse voltage across the diodes in Fig. 2-15*a* can be determined in the following way. The sum of the voltages across the two diodes is equal to the full secondary voltage of the transformer, $2e_s$. However, at each instant of time one of the diodes acts as a short circuit and the other acts as an open circuit; hence the voltage $2e_s$ appears across the diode that acts as an open circuit, and the maximum inverse voltage that appears across either diode is $2E_s$.

Another full-wave rectifier circuit is shown in Fig. 2-16. This circuit, which is called the bridge rectifier, requires four diodes, but it has the advantage that a center-tapped transformer is not required. During the positive half cycle of e_s current flows through D_1, R_L, and D_3 and during the negative half cycle current flows through D_4, R_L, and D_2. In both cases current flows in the same direction through R_L. During the positive half cycle R_L is connected across the transformer by D_1 and D_3; during the negative half cycle it is connected across the transformer by D_2 and D_4 with the terminals reversed. It follows that the waveform of voltage across R_L is the same as that shown in Fig. 2-15*b* and that e_L is given by Eqs. (2-9) to (2-11).

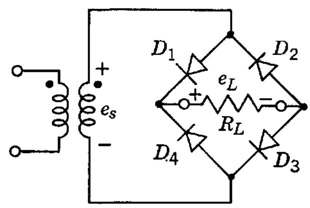
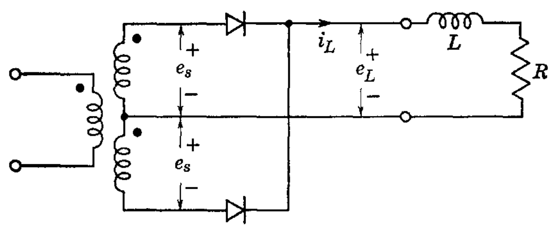


FIG. 2-16. A bridge rectifier with supply voltage $e_s = E_s \sin \omega_s t$.

The peak inverse voltage across the diodes can be determined by noting that the sum of the voltages across D_1 and D_4 equals the supply voltage e_s and that likewise the sum of the voltages across D_2 and D_3 equals e_s. Thus during the positive half cycle of e_s, when D_1 and D_3 act as short circuits, the voltage e_s appears across both D_4 and D_2 in the inverse direction, and the peak inverse voltage is E_s for each of these diodes. During the negative half cycle D_2 and D_4 act as short circuits, the voltage e_s appears across both D_1 and D_3 in the inverse direction, and the peak inverse voltage is E_s for each of these diodes. Thus, for a specified d-c component of voltage at the load, the diodes in the bridge rectifier need withstand only half as much inverse voltage as those in the rectifier of Fig. 2-15. This is an important consideration when semiconductor diodes are used.

2-9. The Full-wave Rectifier with an Inductive Load. Figure 2-17*a* shows a full-wave rectifier supplying power to a load consisting of a series connection of resistance and inductance. Circuits of this type occur, for example, when a full-wave rectifier is used to supply field current to a d-c generator from an a-c source of power. The load current, shown in Fig. 2-17*b*, consists of a d-c component and a time-varying component. The inductive part of the load tends to prevent any change in the load current; hence the time-varying component of load current is smaller than in the case of the full-wave rectifier with a resistive load. In particular, the load current never drops to zero, and either one or the other of the diodes conducts at every instant.

Since one or the other of the diodes conducts at every instant, the load voltage has the form shown in Fig. 2-17*b*, and it can be expressed mathematically by Eq. (2-11). This fact makes the calculation of the load current a relatively simple matter. Each sinusoidal component of the load



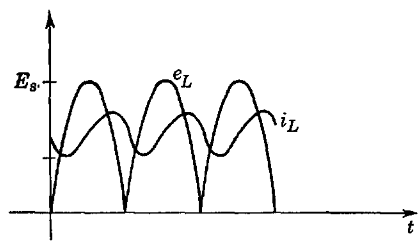


FIG. 2-17. A full-wave rectifier with an inductive load and supply voltage $e_s = E_s \sin \omega_s t$. (*a*) Circuit; (*b*) waveforms.

voltage given by (2-11) can be treated separately by superposition, and the impedance of the load at any frequency is $\mathbf{Z}_L = R + j\omega L$. Hence the d-c component of i_L is

$$I_{dc} = \frac{2E_s}{\pi R} \tag{2-12}$$

If the frequency of the supply voltage is 60 cps, then $\omega_s = 377$ rps, and the load reactance at the fundamental frequency of i_L is $2\omega_s L = 754L$. Thus the amplitude of the fundamental-frequency component of i_L is

$$I_1 = \frac{(4/3\pi)E_s}{[R^2 + (754L)^2]^{1/2}} \tag{2-13}$$

The amplitude of each sinusoidal component of i_L can be evaluated in this manner. It is clear that the inductive reactance increases as the order of the harmonic increases; hence the amplitudes of higher harmonics are relatively small.

Figure 2-18 shows a circuit that is used to control the speed of a small d-c motor such as might be used to drive a lathe, an application in which close speed control is often required. The bridge rectifier R_1 supplies a constant current to the shunt field. Its load consists of a series connection of the resistance and inductance of the shunt field; hence the circuit is a full-wave rectifier with an RL load, and the field current can be determined in the manner outlined in the preceding paragraph.

Transformer T_2 and rectifier R_2 supply an adjustable direct voltage to the armature of the motor; the rotational speed of the armature is roughly proportional to this voltage. The load on rectifier R_2 consists of a series

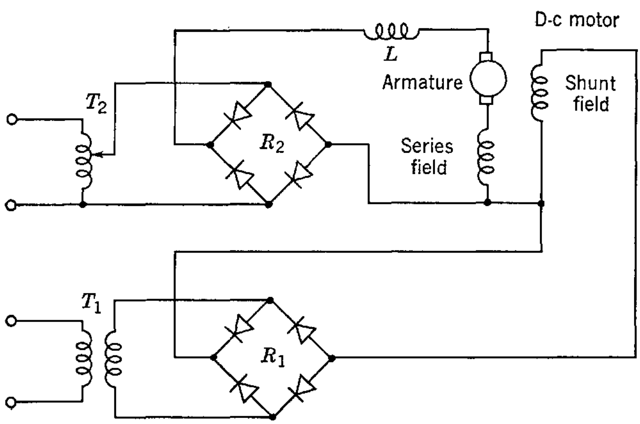


FIG. 2-18. An adjustable-speed d-c-motor circuit.

connection of the armature-circuit resistance, the armature-circuit inductance, the series-field inductance, the added inductance L, and a voltage source accounting for the back emf generated by the rotation of the armature. The inductance L is added to limit the alternating components of armature current to small magnitudes, for these components generate heat in the armature but do not produce any average torque.

The analysis of the armature circuit is complicated by the simultaneous presence of an inductance and a voltage source in the load. If the inductance were present alone, the armature current could be calculated in the same way as the field current; if the voltage source were present alone, the current could be calculated by the method used in the analysis of the battery charger of Example 2-2. The voltage source tends to restrict the conducting interval of the diodes to a fraction of a cycle of the input voltage, as in the case of the battery charger, whereas the inductance tends to maintain a constant flow of current. The inductance prevails when the direct current in the load is large; the voltage source prevails when the direct current in the load is small. This problem is studied in greater detail in Sec. 2-11.

It is pointed out in Sec. 2-4 that there are many applications in which it is desired to obtain a direct voltage (or current) with negligible ripple from the standard a-c power mains. In that section it is shown that when only a small direct current is required, the desired result can be obtained by connecting a capacitor in parallel with the output terminals. It follows from the discussion in the first part of this section that when large direct currents are required the desired results can be obtained by connecting an inductor in series with the output. Thus the circuit in Fig. 2-17*a* can be viewed as a full-wave rectifier delivering a nearly pure direct current to the resistor R. The inductor L serves as a choke, or filter choke, to limit the alternating components of current in the load.

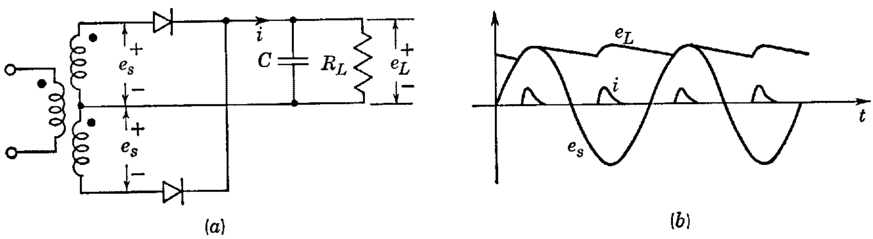


FIG. 2-19. A full-wave rectifier with a capacitor filter and supply voltage $e_s = E_s \sin \omega_s t$. (*a*) Circuit; (*b*) waveforms.

The alternating components cannot be made zero, but if the direct component, which is limited only by R, is large, then the alternating components, which are limited primarily by L, can be made relatively small.

The use of an inductor to reduce the ripple in the load current has the advantage, compared with the use of a capacitor, that each diode conducts for a full half cycle. Hence the ratio of average to peak diode current is much larger than in the case of the capacitor filter, and better use is made of the diodes. However, in many applications the d-c output current required is so small that suitable smoothing cannot be obtained with an inductor of reasonable size. Hence it is often necessary to use a capacitor for the smoothing or, in many cases, to use more elaborate low-pass filters consisting of combinations of L and C.

2-10. The Full-wave Rectifier with a Capacitor Filter. Figure 2-19*a* shows a full-wave rectifier with a smoothing capacitor connected across its output terminals. This circuit is similar to the single-diode peak rectifier discussed in Sec. 2-4; in this case, however, the capacitor receives two charging pulses in each cycle of the input voltage instead of one. Consequently, for the same values of R_L and C, the full-wave peak rectifier has less ripple in its output voltage than does the half-wave rectifier of Sec. 2-4.

When the ripple component of voltage is small compared with the d-c component, the magnitude of the ripple can be estimated easily. The peak value of e_L is E_s; hence e_L can be written as

$$e_L = E_s - e_r \tag{2-14}$$

where e_r is the saw-tooth ripple voltage shown in Fig. 2-19*b*. The ripple voltage results from the capacitor discharging through the load resistor during the interval between charging pulses. During this interval, which is approximately $T_s/2$, half the period of e_s, e_L is nearly constant at the value E_s. Hence the current discharging C is approximately constant at the value

$$i_L = \frac{E_s}{R_L} \tag{2-15}$$

The charge lost by C while the diodes are not conducting is, then,

$$\Delta q = \frac{T_s i_L}{2} \tag{2-16}$$

and the corresponding change in the capacitor voltage is

$$\Delta e_L = E_r = \frac{\Delta q}{C} = \frac{T_s i_L}{2C} = \frac{T_s E_s}{2CR_L} \tag{2-17}$$

where E_r represents the peak-to-peak value of the ripple component of e_L. The quantity CR_L is the time constant τ of the combination of load resistor and filter capacitor; thus (2-17) can be written as

$$E_r = \frac{T_s}{2\tau} E_s \tag{2-18}$$

The ratio of peak-to-peak ripple voltage to d-c output voltage is thus approximately

$$\frac{E_r}{E_s} = \frac{T_s}{2\tau} = \frac{1}{2\tau f_s} = \frac{\pi}{\omega_s CR_L} \tag{2-19}$$

where $f_s = 1/T_s = \omega_s/2\pi$ is the frequency of the a-c supply.

Example 2-3. The half-wave peak rectifier of Fig. 2-20*a* is supplied with the square wave of voltage shown in Fig. 2-20*b*. Determine the approximate values of the d-c component and the ripple component of the load voltage.

Solution. When the capacitor is charging, the diode acts as a short circuit, and the circuit connected to the capacitor can be replaced by a Thevenin equivalent as shown in Fig. 2-20*c*. The time constant of the charging circuit is

$$T_c = (99)(10^{-6}) = 0.000099 \text{ sec}$$
$$\approx 0.1 \text{ msec}$$

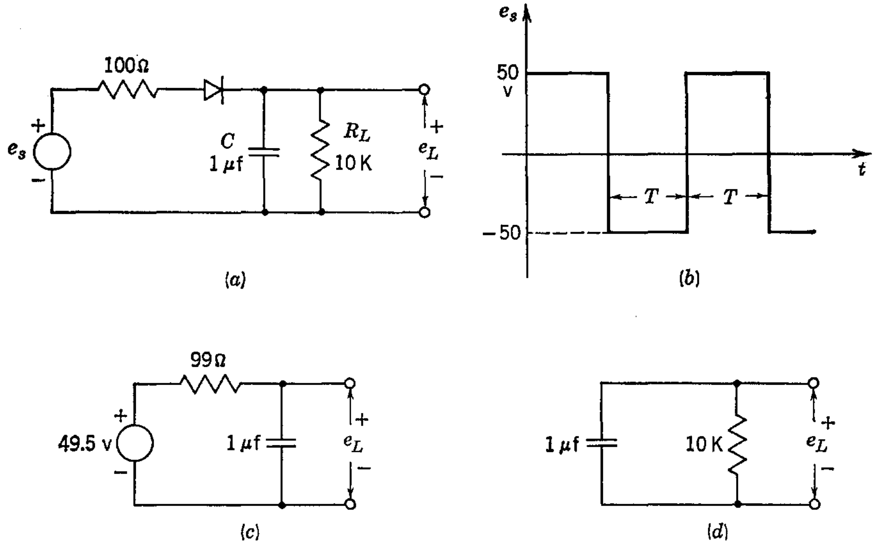


FIG. 2-20. Diode rectifier for Example 2-3. (a) Circuit; (b) waveform, $T = 0.5$ millisecond; (c) charging circuit; (d) discharging circuit.

Since T_c is approximately $T/5$, C charges essentially to 49.5 volts during the positive half cycle of e_s.

When the capacitor is discharging, the diode acts as an open circuit and the load voltage is given by the circuit of Fig. 2-20*d*. The time constant of the discharging circuit is

$$T_d = (10^4)(10^{-6}) = 0.01 \text{ sec} = 10 \text{ msec}$$

Since $T_d = 20T$, C discharges only a small amount, and e_L remains nearly constant during the negative half cycle of e_s.

The amplitude of the ripple voltage is given approximately by

$$E_r = \frac{\Delta q}{C} = \frac{Ti_L}{C} = \frac{Te_L}{10^4 C} = \frac{(5)(10^{-4})(49.5)}{(10^4)(10^{-6})} = 2.48 \text{ volts}$$

As a first approximation, the d-c component of the output voltage can be taken as 49.5 volts. For a closer approximation,

$$E_{dc} \approx \frac{1}{2T}\left(49.5T + 49.5T - \frac{2.48T}{2}\right) \approx 49 \text{ volts}$$

2-11. The Full-wave Rectifier with an *LC* Filter. The parallel-capacitor filter of Sec. 2-10 gives satisfactory smoothing in applications where the direct current required is less than a few milliamperes, and the series-inductor filter of Sec. 2-9 is satisfactory in applications where the direct current required is greater than a few amperes. In the host of applications where the direct current required lies between these two ranges, a more elaborate filter circuit must be used. One filter circuit that is widely used in these applications is shown in Fig. 2-21. This filter, which combines the series-inductor and shunt-capacitor actions, is known as the choke-input filter.[1-3]

The inductor in the filter of Fig. 2-21 tends to maintain the input current to the filter, i_f, constant, and the capacitor tends to maintain the load voltage e_L constant. At no load, with $R_L = \infty$ and $i_L = 0$, the capacitor charges to the peak value of e_s as in the case of the peak rectifier. This fact is indicated in Fig. 2-22a, where the d-c component of load voltage is plotted against the d-c component of load current. As the load current

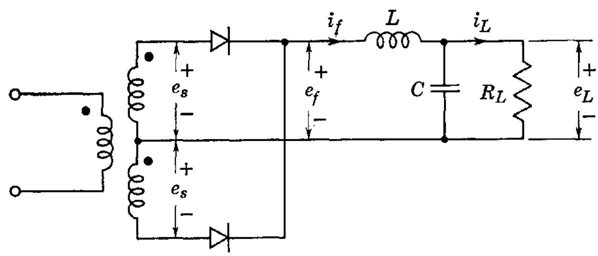


FIG. 2-21. A full-wave rectifier with a choke-input filter and supply voltage $e_s = E_s \sin \omega_s t$.

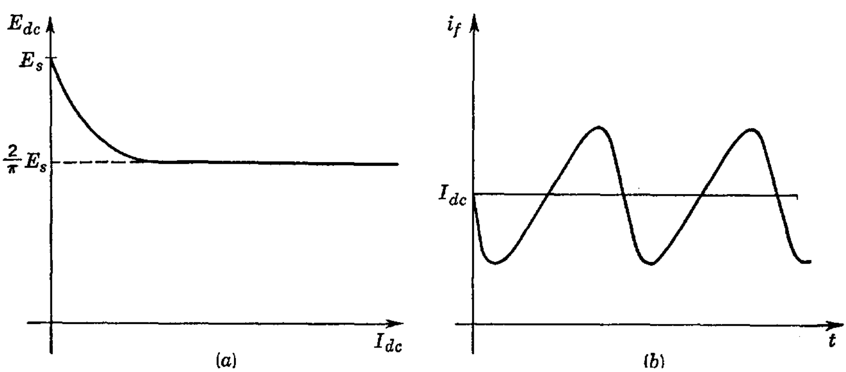


FIG. 2-22. Characteristics of the full-wave rectifier with a choke-input filter. (a) Output characteristic; (b) waveforms.

is increased from zero, the load voltage decreases rather rapidly and one or the other of the diodes conducts during a portion of each half cycle of e_s. When the d-c component of load current exceeds a certain value, the diodes conduct for a full half cycle, and the input to the filter is connected to the transformer at each instant through one diode or the other. Under these conditions e_f, the voltage at the input to the filter, is a full-wave rectified sinusoid like that shown in Fig. 2-15b. These conditions exist for all load currents greater than this value, and the d-c component of load voltage remains constant at the average value of $|e_s|$ as shown in Fig. 2-22a. (In actual rectifier circuits the load voltage decreases slightly with increasing load current because of the series resistance associated with the inductor and the diodes.)

The minimum value of I_{dc} for which the diodes conduct during a full half cycle can be determined rather easily, and the method of analysis is basic in the study of a number of electronic circuits. When the diodes conduct for a full half cycle, the input voltage to the filter e_f is given by the right-hand side of Eq. (2-11). The input current to the filter can be calculated from this expression and the input impedance to the filter by considering each sinusoidal component of applied voltage separately. The input impedance at any frequency is

$$\mathbf{Z}_f = j\omega L + \frac{R_L}{1 + j\omega C R_L} \tag{2-20}$$

As shown in Fig. 2-22*b*, the input current i_f consists of the d-c load current plus a time-varying component. Ordinarily the time-varying component consists primarily of the fundamental-frequency component in the Fourier series. In a well-designed filter, the capacitor C is so large that at the frequency of this component the second term in Eq. (2-20) is negligible compared with the first term; hence the time-varying component of i_f is approximately sinusoidal with an amplitude given by

$$I_1 = \frac{2}{3\pi\omega_s L} E_s \tag{2-21}$$

This amplitude is independent of the load resistance and the load current.

It is clear from Fig. 2-22*b* that for all values of I_{dc} greater than I_1, the input current to the filter never drops to zero, and one or other of the diodes must therefore conduct at all times. If I_{dc} is reduced to a value equal to I_1, i_f drops just to zero at its minimum value. The current i_f cannot have any negative values, for the diodes in Fig. 2-21 cannot conduct in the reverse direction. Hence if I_{dc} is less than I_1, both diodes become nonconducting during certain portions of the cycle, and neither diode conducts for a full half cycle of the input voltage. Thus the minimum value of I_{dc} for which the diodes conduct for a full half cycle corresponds to the condition

$$I_{dc} = I_1 = \frac{2}{3\pi\omega_s L} E_s \tag{2-22}$$

But $I_{dc} = E_{dc}/R_L$, and in the range of load currents where both diodes conduct for a full half cycle, $E_{dc} = 2E_s/\pi$ (Fig. 2-22*a*). Substituting these relations in (2-22) yields

$$I_{dc} = \frac{1}{3\omega_s L} E_{dc} \tag{2-23}$$

and

$$R_L = 3\omega_s L \tag{2-24}$$

Equation (2-24) gives the maximum value of R_L for which both diodes conduct for a full half cycle of the a-c supply voltage. In order to avoid operating on the steeply rising portion of the volt-ampere characteristic of Fig. 2-22*a*, it is common practice to connect a resistance of this value, called a bleeder, across the output terminals of the filter. When the frequency of the a-c supply is 60 cps, Eq. (2-24) becomes

$$R_L = 1130L \tag{2-25}$$

where R_L is in ohms, and L is in henrys.

When the d-c load current is greater than the critical value determined above, both diodes conduct for a full half cycle, and the input voltage to the filter is given by Eq. (2-11). The ripple voltage appearing at the output of the filter can be calculated by considering each Fourier component of the input voltage separately. In a well-designed filter the ripple voltage at the output consists primarily of a sinusoidal component at the fundamental ripple frequency. The fundamental-frequency

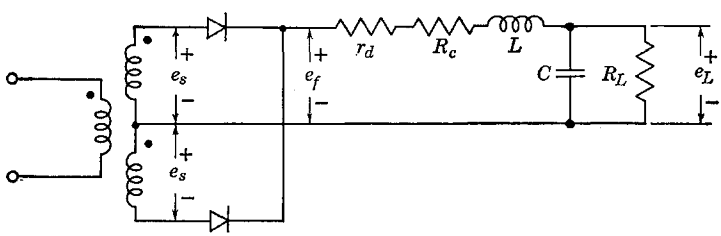


FIG. 2-23. The full-wave rectifier including diode and inductor resistances.

component of voltage at the input to the filter is given by Eq. (2-11) as $4E_s/3\pi$, and its radian frequency is $2\omega_s$. It then follows from the circuit shown in Fig. 2-21 that the complex amplitude of the fundamental-frequency component of output voltage is

$$\mathbf{E}_1 = \frac{R_L/(1 + j2\omega_s CR_L)}{j2\omega_s L + R_L/(1 + j2\omega_s CR_L)} \frac{4E_s}{3\pi} \tag{2-26}$$

$$= \frac{R_L}{R_L(1 - 4\omega_s{}^2LC) + j2\omega_s L} \frac{4E_s}{3\pi} \tag{2-27}$$

For good filtering, L and C should be chosen to make $4\omega_s{}^2LC$ much greater than unity.

If there is a series resistance r_d associated with each diode in the circuit of Fig. 2-21, the input current to the filter must flow through r_d, and the input voltage to the filter is reduced accordingly. Since only one diode conducts at any instant, the effect of the two diode resistances can be accounted for conveniently by a single resistance at the input to the filter as shown in Fig. 2-23. This circuit shows also a resistance R_c that accounts for the resistance of the filter inductor. The d-c component of the load voltage is reduced by the voltage divider action of r_d, R_c, and R_L.

Example 2-4. The supply voltage for the circuit of Fig. 2-23 is $e_s = 550 \sin 377t$ volts, and the circuit parameters are $r_d = 400$ ohms, $R_c = 100$ ohms, $R_L = 3000$ ohms, $L = 10$ henrys, and $C = 16$ μf. Determine the d-c component and the 120-cps component of voltage across the load.

Solution. The d-c load current is limited by the resistance $r_d + R_c + R_L = 3500$ ohms. Since this resistance is less than $1130L = 11{,}300$ ohms, both diodes conduct for a full half cycle, and the voltage at the input to the filter (including r_d) is

$$e_f = |550 \sin 377t|$$

The d-c component of e_f is therefore

$$(e_f)_{dc} = (0.636)(550) = 350 \text{ volts}$$

and the d-c load voltage is

$$E_{dc} = (350)(3000/3500) = 300 \text{ volts}$$

Ignoring the small effect of r_d and R_c on the 120-cps component of voltage, Eq. (2-27) gives

$$\mathbf{E}_1 = \frac{3000}{3000[1 - 4(377)^2(10)(16)(10^{-6})] + j2(3770)} \frac{(4)(550)}{3\pi}$$

$$= \frac{3000}{-(27)(10^4) + j7540}(234)$$

Thus the amplitude of the 120-cps component of load voltage is

$$E_1 \approx 2.6 \text{ volts}$$

2-12. The Full-wave Rectifier with a *CLC* Filter. The full-wave rectifier with a *CLC* filter is shown in Fig. 2-24*a*. This is perhaps the most widely used rectifier-filter combination. In comparison with the *LC* filter, the *CLC* filter gives better filtering and higher output voltage for a given supply voltage; however, the voltage regulation of the *CLC* filter is not as good as that of the *LC* filter. A typical output volt-ampere characteristic for a *CLC*, or capacitor-input, filter is shown in Fig. 2-24*b*.

The action in this circuit is somewhat like the action in the full-wave peak rectifier discussed in Sec. 2-10. A pulse of current flows through one or the other of the diodes when e_s is near its peak value in each half cycle. Because the inductor L tends to hold its own current constant, the pulses of current serve primarily to charge C_1. When the load current is small, C_1 is charged nearly to the peak value of e_s. The voltage across C_1 has a saw-tooth waveform similar to that shown in Fig. 2-19*b*; the combination of L and C_2 then acts like an LC filter to transmit the d-c component of voltage across C_1 to the load.

The quantitative analysis[1,3] of the rectifier with a capacitor-input filter is complicated by the fact that the diodes conduct for less than a

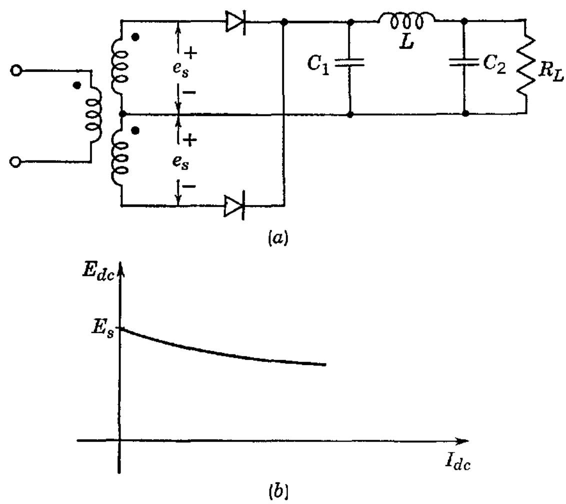


FIG. 2-24. A full-wave rectifier with a capacitor-input filter and supply voltage $e_s = E_s \sin \omega_s t$. (a) Circuit; (b) output characteristic.

half cycle of the supply voltage; thus the voltage at the input to the filter is one of the unknowns that must be determined. An analysis can be made with the aid of certain simplifying approximations; however, it will not be undertaken here. The design of such rectifiers is usually based on families of output characteristics, similar to that of Fig. 2-24b, published by the tube manufacturer for the particular diode to be used.

REFERENCES

1. Gray, T. S.: "Applied Electronics," 2d ed., John Wiley & Sons, Inc., New York, 1954.
2. Terman, F. E.: "Electronic and Radio Engineering," 4th ed., McGraw-Hill Book Company, Inc., New York, 1955.
3. Arguimbau, L.: "Vacuum Tube Circuits," John Wiley & Sons, Inc., New York, 1948.
4. Seely, S.: "Radio Electronics," McGraw-Hill Book Company, Inc., New York, 1956.

PROBLEMS

2-1. A battery charger like that shown in Fig. 2-3 is used to charge a 6.3-volt automobile battery. The supply voltage is $e_s = 20 \sin 377t$.

a. What value of $R + R_s$ is required to limit the peak diode current to 5 amp?

b. With the adjustment of part *a*, what is the direct current delivered to the battery?

c. What is the average power delivered to the battery, the average power dissipated in the resistors, and the average power delivered by the source?

d. What is the peak inverse voltage that appears across the diode?

2-2. The supply voltage for a peak rectifier like that shown in Fig. 2-7 is $e_s = 100 \sin 377t$. The capacitor and the load resistor are large so that there is no appreciable ripple voltage across the load, and e_L is 90 volts. Under these conditions the action of the capacitor is similar to that of a 90-volt battery. What value of R_s is required to limit the peak diode current to 2 ma?

2-3. The periodic voltage shown in Fig. 2-25 is applied at the input of the diode clamper shown in Fig. 2-11. The source resistance R_s is negligibly small, and R and C are large so that the capacitor cannot discharge appreciably through R during one cycle of the signal voltage.

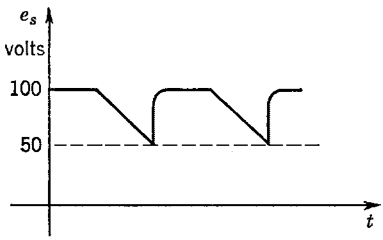


FIG. 2-25. Waveform for Prob. 2-3.

a. When the circuit is in the steady state, what is the value of the voltage across the capacitor?

b. Sketch the waveform of the output voltage e_d. Show on this sketch the values of all significant voltages.

2-4. The field current for a d-c motor is supplied by a bridge rectifier like that shown in Fig. 2-18. The a-c supply voltage at the secondary of T_1 is 115 volts, rms, at 60 cps. The inductance of the field winding is 100 henrys, and its resistance is 75 ohms.

a. Determine the d-c component of the field current.

b. Determine the amplitude of the 120-cps component of the field current.

2-5. The load on a bridge rectifier like that in Fig. 2-16 is a parallel RC combination like that shown in Fig. 2-19. The supply voltage is $e_s = 10 \sin 200t$, the load capacitance is 1 μf, and the load resistance is 3 megohms. Under these conditions the ripple voltage across the load is small.

a. What is the average current in the load resistor?

b. What is the average current in each diode?

c. What is the peak inverse voltage appearing across each diode?

2-6. In a peak rectifier like that shown in Fig. 2-19 the filter capacitance is 8 μf, and the d-c component of the load current is 1 ma. The supply voltage is sinusoidal with a frequency of 60 cps. Under these conditions the ripple voltage across the load is small.

a. What value of supply voltage E_s is required to give a d-c load voltage of 100 volts? Give the rms value.

b. What is the peak-to-peak value of the ripple component in the load voltage?

c. What is the peak inverse voltage appearing across the diodes?

2-7. The choke in a rectifier circuit like that shown in Fig. 2-23 has an inductance of 8 henrys and a resistance of 100 ohms. The filter capacitance is 32 μf, the series resistance associated with each diode is 400 ohms, the load resistance is 4000 ohms, and the supply voltage is sinusoidal with a frequency of 60 cps.

a. What must be the value of the supply voltage E_s to give a d-c load voltage of 250 volts? Give the rms value.

b. Calculate the amplitude of the 120-cps component of the load voltage. (Neglect the resistances of the choke and the diodes in this calculation.)

c. What is the largest value of bleeder resistance that will ensure that the diodes conduct for a full half cycle?

2-8. A certain rectifier is identical with the one analyzed in Example 2-4 except that the diodes are reversed. What is the value of the d-c component of e_L for this circuit?

2-9. The supply voltage for a rectifier like that shown in Fig. 2-17 is $e_s = 200 \sin 800\pi t$. The filter choke has an inductance of 10 henrys and a series resistance of 100 ohms. The load resistance is 1000 ohms. The diode resistance is negligible.

a. What is the d-c component of voltage across the load resistance?

b. What are the amplitude and the frequency (in cycles per second) of the fundamental-frequency component of the ripple voltage at the load?

c. What is the peak inverse voltage appearing across the diodes?

2-10. The voltage applied to the diode limiter of Fig. 2-5 is $e_s = 10 \sin 2000\pi t$. The resistance R is 1000 ohms, and the battery voltages are $E_1 = E_2 = 5$ volts.

a. Sketch two full cycles of the supply voltage e_s, the output voltage e_o, and the voltage drop across R. Mark on this sketch the values of all significant voltages.

b. What is the peak current through the diodes and the peak inverse voltage across the diodes?

2-11. A full-wave rectifier like the one shown in Fig. 2-23 is supplied with a 400-cps sinusoidal voltage. The inductance of the choke is 1 henry, its resistance is 25 ohms, and the series resistance associated with the diodes is 200 ohms. What is the largest value of bleeder resistance that will ensure that the diodes conduct for a full half cycle?

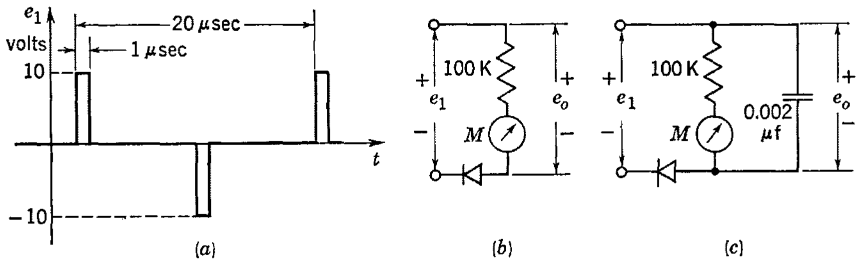


FIG. 2-26. Circuits and waveform for Prob. 2-12. (*a*) Voltage waveform; (*b*) first circuit; (*c*) second circuit.

2-12. The periodic voltage shown in Fig. 2-26*a* is applied to the two circuits shown in Figs. 2-26*b* and *c*. The d-c microammeter M has negligible resistance, and it indicates the average value of current.

a. Determine the steady-state meter reading in each case.

b. Sketch the waveform of e_o for steady-state conditions in each circuit. Mark the values of all significant voltages and time intervals.

2-13. Peak rectifiers like the one shown in Fig. 2-19 are sometimes supplied with alternating voltages having the waveform shown in Fig. 2-27. The filter capacitor required in such cases is much smaller than would be needed with a sinusoidal supply voltage of the same frequency.

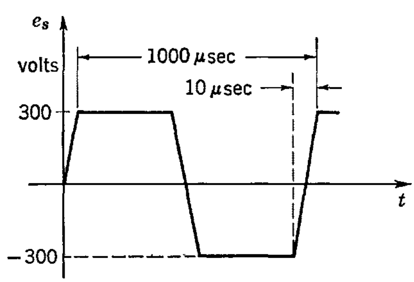


FIG. 2-27. Supply-voltage waveform for the full-wave rectifier of Prob. 2-13.

a. Sketch the waveform of load voltage for $C = 0$.

b. If the d-c load current is 100 ma, and if the peak-to-peak value of ripple voltage is to be less than 1 volt, what is the smallest value of C that can be used? Sketch the waveform of the ripple voltage.

c. Using the fact that the current in a capacitor is $i = C\,de/dt$, determine the peak diode current. (In practice, a resistor in series with each diode would limit the diode current to a much smaller value than this.)

d. Repeat part *b* for the case where e_s is a 1000-cps sinusoid with an amplitude of 300 volts. Assume that the load voltage is approximately equal to the peak value of e_s.

CHAPTER 3

PRACTICAL RECTIFIERS

It is shown in Chap. 2 that the addition of a nonlinear component, the diode, to the collection of components used in electric circuits permits a number of valuable results to be achieved, results that cannot be obtained with the linear elements R, L, and C alone. The discussions of Chap. 2 are based on the assumption that ideal diodes are available. Before the circuits of Chap. 2 can be built, however, it is necessary to devise some physical component that behaves at least approximately like the ideal diode. It is the purpose of this chapter to study several physical devices that can be used in the construction of the circuits discussed in Chap. 2. The physical laws governing these devices are examined so that, among other things, it can be seen how and why the physical diodes do not behave exactly like ideal diodes.

The properties of the ideal diode are given by the volt-ampere characteristic of Fig. 2-1*b*; in fact, this characteristic can be looked upon as the definition of the ideal diode. The extent to which any physical device behaves like an ideal diode can be perceived by comparing its volt-ampere characteristic with that of the ideal diode. Hence the study of physical diodes is concerned largely with the study of their volt-ampere characteristics and the physical laws underlying them.

3-1. The Vacuum Diode. A vacuum diode is shown schematically in Fig. 3-1*a*. It consists of two active electrodes, an anode (or plate) and a cathode, designated p and k, respectively, in Fig. 3-1, and a cathode heater that does not enter directly into the action of the tube. These elements are enclosed in an envelope of mètal or glass from which as much air has been removed as is economically practical, leaving a very high vacuum in the interelectrode space.

When the electrodes are at room temperature the current I_b is extremely small, no matter how large E_b is made and regardless of the polarity of E_b (assuming, of course, that the insulation of the wiring outside the tube does not break down). If, however, sufficient power is applied to the cathode heater to raise the cathode temperature to about 750°C (for typical small diodes), the behavior of the diode changes. At this temperature electrons escape from the cathode in large numbers in somewhat the same way that water molecules evaporate from the surface of a container of hot water. This phenomenon of electron emission is considered in greater detail in Sec. 3-6. The electrons issuing from the cathode emerge with some kinetic energy; hence, if E_b is zero, electrons arrive at the anode at an appreciable rate, and the meter M in Fig. 3-1*a* indicates an appreciable current. This phenomenon, first observed by Edison in his studies of the incandescent lamp, is known as the Edison effect. It is shown, somewhat exaggerated, in the volt-ampere characteristic of Fig. 3-1*b*.

If the polarity of the battery in Fig. 3-1*a* is reversed, the plate of the diode is held at a negative potential relative to the cathode, and there is an electric field in the space between the plate and cathode that opposes

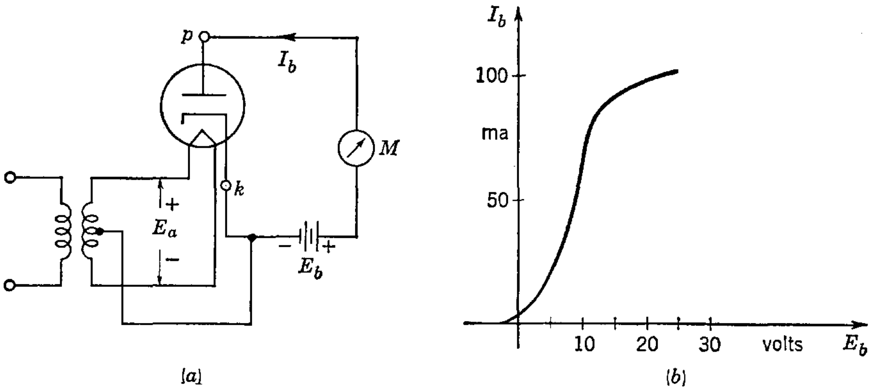


FIG. 3-1. The vacuum diode. (*a*) Diode and supply voltages; (*b*) volt-ampere characteristic.

the flow of electrons from cathode to anode. Moreover, since the plate is at a relatively low temperature, it emits electrons at a negligible rate. Hence there is no flow of electrons from plate to cathode, for there is no source of free electrons at the plate. If the reverse voltage applied to the diode is greater than about ½ volt, the diode current I_b is essentially zero. Thus there can be no reverse current through the diode, and in this respect it behaves as an ideal diode.

If the battery is connected in the circuit with the polarity shown in Fig. 3-1*a*, the plate is held positive relative to the cathode, and the electric field in the interelectrode space accelerates the electrons emitted from the cathode toward the plate. The current I_b therefore increases with increasing E_b as shown by the characteristic curve of Fig. 3-1*b*. When E_b reaches a sufficiently high value, about 10 volts for the conditions pictured in Fig. 3-1, electrons are drawn to the plate as fast as they are normally emitted from the cathode, and the diode current increases relatively slowly with further increases in E_b. The slow increase in I_b in this region results from the fact that increasing E_b increases the emission from the cathode somewhat. In this region the diode current is limited by the cathode emission. Prolonged operation in this region usually results in permanent damage to the cathode; hence vacuum diodes are normally operated on that part of the characteristic curve lying below the knee. It is clear from a comparison of the volt-ampere characteristics of Figs. 2-1*b* and 3-1*b* that the vacuum diode differs from the ideal diode primarily in that there is a voltage drop across the vacuum diode when it conducts in the forward direction; it does not act as a short circuit to forward current.

The representation of the vacuum diode used in Fig. 3-1*a* is a symbolic, or schematic, representation; it does not picture the physical structure of the tube. The cathode of such tubes usually takes the form of a small hollow nickel rod, coated on its outer surface with a thin layer of barium and strontium oxides to increase the electronic emission. The heater is a tungsten wire, like the filament of an ordinary incandescent lamp, placed inside the hollow rod. The plate is usually a metal cylinder surrounding and concentric with the cathode. When the heater power is supplied through a transformer from an a-c source it is common practice to provide the transformer with a center-tap connection. This center tap is usually connected to the negative terminal of the plate power supply, as shown in Fig. 3-1*a*, to fix the potential of the heater relative to the cathode. Sometimes the heater is held 20 or 30 volts positive relative to the cathode to prevent electrons emitted by the heater from reaching the surrounding cathode and interfering with the operation of the tube.

Another type of construction that is occasionally used does not employ a separate filament to heat the cathode. The filament is made of nickel wire or ribbon, and the oxides of barium and strontium are applied directly to the surface of the filament. These are termed directly heated cathodes in contrast to the indirectly heated cathodes described above.

Oxide-coated cathodes are not suitable in certain applications involving voltages greater than about 1000 volts. Tubes built for these applications often employ an ordinary tungsten filament as the cathode. In order to obtain sufficient emission from these filaments, the tungsten must be raised to a much higher temperature than is required by oxide-coated cathodes, usually about 2100°C. The emission from such filaments can be increased by mixing thorium with the tungsten, giving rise to thoriated-tungsten filaments. These more efficient filaments are usually operated at about 1700°C.

3-2. Semiconductor Diodes. A semiconductor diode of the junction type is shown pictorially in Fig. 3-2*a* and schematically in Fig. 3-2*b*. The diode consists of two semiconducting materials having different electrical properties that are joined together along a common boundary known as a junction. The volt-ampere characteristic for a typical junction diode is shown in Fig. 3-2c. There is no obvious reason why the device should have such a characteristic; hence no explanation will be attempted until the physical laws governing the flow of electrons in semiconductors have been examined. Nevertheless, it is clear from Fig. 3-2c that the characteristics of the semiconductor diode are very much like those of the ideal diode provided the inverse voltage does not exceed a certain critical value. The inverse voltage at which the diode breaks down is sometimes referred to as the Zener voltage. The characteristic of the junction diode is so good that, in order to display clearly the imperfections, it is necessary to plot the reverse voltage and current to a different scale from that used for the forward voltage and current. For many applications the characteristics of the junction diode are

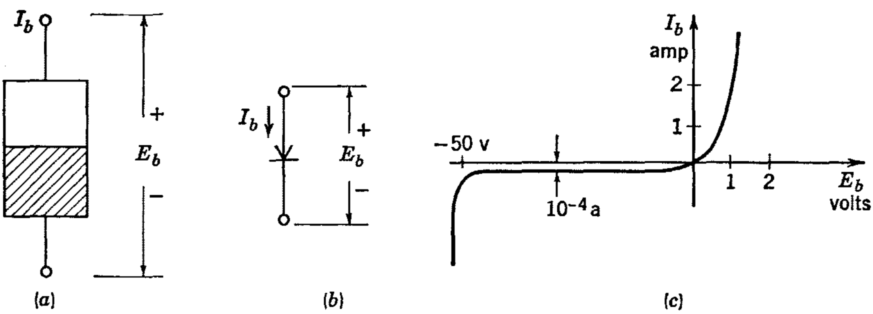


FIG. 3-2. The semiconductor diode. (a) Pictorial representation; (b) schematic representation; (c) volt-ampere characteristic.

superior to those of the vacuum diode. In addition, the junction diode can be made much smaller than the vacuum diode, and it requires no heater power. The principle disadvantage of the junction diode is that its characteristics deteriorate rapidly when the junction temperature rises above about 75°C for diodes made from germanium and about 200°C for diodes made from silicon.

Semiconductor diodes can also be made in forms different from the junction type described above. The point-contact diode consists of a germanium or silicon crystal to which one of the two contacts is made by the tip of a fine wire. The action that takes place at this contact leads to a volt-ampere characteristic that is a good approximation to the ideal diode characteristic, although it is not as good an approximation as is provided by the junction diode. The point-contact diode is superior to other diodes at high frequencies, however, for the parasitic capacitances associated with it are extremely small. Such diodes are used primarily in high-frequency, low-power applications.

Another type of semiconductor diode that is of engineering importance consists of a layer of semiconducting material in contact with a metal. When this contact is formed in a certain way it behaves like the junction of a junction diode and produces a diode characteristic. Diodes of this type are usually referred to as metallic rectifiers. The semiconductors commonly used in metallic rectifiers are selenium and cuprous oxide. Although the metallic rectifier is a much poorer approximation to the ideal diode than is the junction diode, it nevertheless finds wide application in the field of power rectification.

The physical theory underlying semiconductor diodes is rather complex, and little can be said quantitatively about their performance except in the case of the junction diode. Therefore only the junction diode will be treated in the pages that follow. The junction diode is simpler than the other types because it is made of a single crystal; its structure is therefore uniform and simple.

3-3. The Motion of Charged Particles in Electrostatic Fields. In most vacuum-tube and semiconductor electronic devices the flow of electric current is associated with the movement of charged particles under the influence of forces exerted by electrostatic fields. A study of the basic physical laws governing this motion aids in understanding the volt-ampere characteristics of electronic devices. Complete quantitative solutions can be obtained in only a few simple cases; however, a valuable qualitative insight of wide applicability can be obtained from this study.

Figure 3-3 shows a pair of metal electrodes that can be thought of as representing symbolically the electrodes of a vacuum diode. A potential difference is established between these electrodes by a battery as shown in the figure. If a small particle carrying a charge of electricity is placed in the interelectrode space, it experiences a force that results from the action of the battery and the electrodes. Since this force is experienced at every point in the space, a field of force is said to exist in the space; since the force is electric in nature, the field is called an electric field. If the strength of the force at all points is independent of time, the field is a static field.

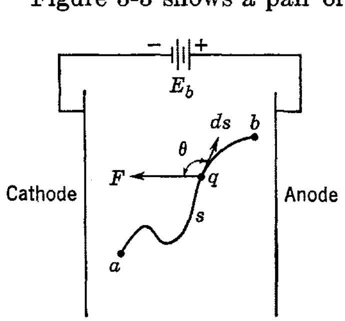


FIG. 3-3. A charged particle moving in an electrostatic field.

Certain important properties of the electrostatic field can be formulated mathematically in terms of the charged particle mentioned above, provided that the particle is so small compared with the dimensions of the electrodes that it can be considered to occupy a point in space.[1] The magnitude of the charge is designated by q, as indicated in Fig. 3-3; it is always a positive number. If the charge is negative electricity, a minus sign is prefixed to the symbol. The force on the particle depends on the position of the particle, the voltage applied between the electrodes, and the magnitude of the charge; the direction of the force reverses when the sign of the charge is reversed. Work must be done on the particle by some external agent to move it against this force. In moving the particle an infinitesimal distance ds the work done by the external agent on the particle is

$$dW = -F(\cos\theta)\, ds \tag{3-1}$$

where, as indicated in Fig. 3-3, F is the magnitude of the force, ds is the magnitude of the displacement, and θ is the angle between the direction of the force and the direction of the displacement. The quantity

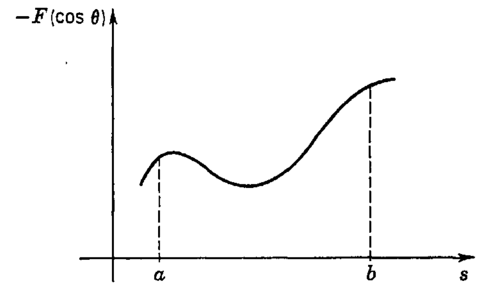


FIG. 3-4. Electrostatic force versus distance along a path.

$F(\cos\theta)$ is the component of the electric force in the direction of ds; $-F(\cos\theta)$ is the force applied by the external agent in the direction of ds. When the particle moves against the electric field, $\cos\theta$ is negative and dW is positive; that is, positive work is done by the external agent.

The total work that must be done by an external agent in moving the charged particle along the path s from point a to point b is found by summing the increments of work required in each increment of distance. Thus

$$W = -\int_s F(\cos\theta)\, ds \tag{3-2}$$

The integral in (3-2) is called a line integral because the values of $F(\cos\theta)$ are taken along the specified path s in Fig. 3-3. This integral can be interpreted in the usual sense with the aid of the diagram shown in Fig. 3-4. Here the values of $-F(\cos\theta)$ have been determined at various points along the specified path and have been plotted as a function of distance along the path. The value of the integral in (3-2) is equal to the area under this curve between the points $s = a$ and $s = b$.

In general, the values of $F(\cos\theta)$ will not be the same along any two paths connecting a and b in Fig. 3-3, and the curve in Fig. 3-4 will have a different shape for each different path. It is a fundamental property of the electrostatic field, however, that the net area under the curve between $s = a$ and $s = b$ is the same for all paths between a and b. Thus the value of the integral in Eq. (3-2) is independent of the path. This remarkable property cannot be deduced from the discussion leading to (3-2), but it can be shown by a further study, based on experimental observations, of the nature of electrostatic forces. Line integrals arising in the study of a number of physical phenomena possess this property. For example, line integrals encountered in the study of the gravitational field possess the property; however, those encountered in the study of magnetic fields do not possess it unless the paths are chosen in a special way. When this property exists it greatly simplifies the solution of field problems.

Two possible paths between points a and b are shown in Fig. 3-5. The same amount of work must be done on a charge to move it from a to b along either path. If the charge is moved in the reverse direction from b to a along one of the paths, the integrand in Eq. (3-2) is the same at each point as when the charge is moved in the forward direction except for a 180° change in θ; thus the work done on the particle in moving it from b to a is the same as that done in moving it from a to b except for a change in sign. Hence if a charged particle moves from a to b and then back to a along s_1 in Fig. 3-5, the net work done on the charge is zero. Moreover, since the work done in traversing paths s_1 and s_2 is the same, the net work done on the charge in moving it from a to b along s_1 and then back to a along s_2 is zero. From these facts it follows that the net work required to move a charged particle through one complete circuit around any closed path in an electrostatic field is zero. Since no net work is required, the particle neither gains nor loses energy in making the circuit; that is, the energy of the particle is conserved. Hence the electrostatic field is called a conservative field. The gravitational field is also a conservative field; the magnetic field is not conservative except in certain restricted regions.

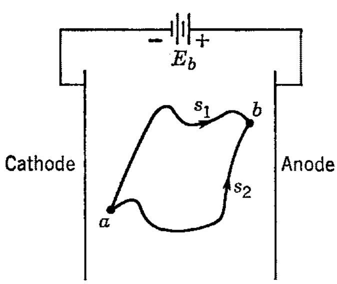


FIG. 3-5. Alternative paths between two points.

If the size of the charge and the magnitude of the voltage applied between the electrodes in Fig. 3-5 are fixed, then the work done in moving the charge from point a to point b depends only on the locations of the two points; it is a property of the space in which the points lie. Moreover, the energy given to the particle in moving it from a to b can always be recovered by returning the particle to a. Therefore the work increases the potential energy of the particle, just as lifting a mass against the force of gravity increases the potential energy of the mass. The quantity W given by Eq. (3-2) is the increase in potential energy of the particle as it moves from a to b.

At any point in an electrostatic field the force on a charged particle is directly proportional to the magnitude of the charge. Hence the force per unit charge, F/q, depends only on the location of the point if the field is fixed. This ratio is the magnitude of the electric-field strength; it is a property given to the point under discussion by the electric field.

From the above discussion it follows that the work done in moving a fixed charge between two fixed points is directly proportional to the magnitude of the charge. Therefore the work per unit charge done in moving the charge between two given points,

$$E_{ab} = \frac{W_{ab}}{q} = \frac{-1}{q} \int_a^b F(\cos \theta)\, ds \tag{3-3}$$

is independent of the magnitude of the charge; it is a property given to the space in which the points lie by the electric field. This quantity E_{ab} is the electric potential of point b with respect to point a. It is also described as the potential difference between a and b, or the rise in potential from a to b, or the fall in potential from b to a. The value of the integral in (3-3) is independent of the path taken between a and b. Hence if point a is a fixed reference point and if b refers to any point in space, then the potential E_{ab} has a unique (single) value for each point in space. The concept of potential would not be useful if this were not the case. In a uniform gravitational field, elevation corresponds to potential. Since the magnetic field is not in general conservative, the integral corresponding to that in Eq. (3-3) does not in general have a unique value, and the notion of potential is not applicable except in restricted regions.

Potential energy and electric potential are quantities relating conditions at one point in space to those at another point. It is usually desirable to pick one point as the reference point and to relate conditions at all other points to conditions at the reference. The reference point is often chosen as some point infinitely remote from the region of interest. In vacuum-tube studies it is convenient, in most cases, to choose a point on the cathode as the reference. Since the cathode (and also the anode) is assumed to be a perfect conductor, no work is required to move a charge from one point to another on its surface, and all points on the cathode are at the same potential. The potential at the cathode is therefore assigned the value zero, and the potential energy of a charge on the cathode is likewise assigned the value zero.

The potential energy gained by a charge q in moving from the cathode in Fig. 3-5 to any point in the interelectrode space, given by Eq. (3-3), is $W = qE$, where E is the potential of the point relative to the cathode. The potential energy gained by q in moving from the cathode to the anode is $W = qE_b$. If the charge is released at the surface of the anode, it is accelerated toward the cathode by the electric force, and it gains kinetic energy as its velocity increases. When the only force acting on a positively charged body at rest is an electrostatic force, the particle must move toward points of lower electric potential, for it can gain kinetic energy only by losing potential energy.

If a charged particle is placed in an electric field, it has a certain amount of potential energy by virtue of its position in the field. If no force other than the electric-field force acts on the particle, it cannot lose any of this energy, although all or part of its potential energy can be converted to kinetic energy and vice versa. The sum of potential and kinetic energies must remain constant. Denoting potential energy by W and kinetic energy by T, this fact is expressed by

$$W + T = C = \text{const} \tag{3-4}$$

If the cathode is taken as the reference for potential and potential energy, then $W = 0$ when the particle is at the cathode, and if the kinetic energy of the particle at the cathode is designated by T_k, then

$$T_k = C$$

and

$$W + T = T_k \tag{3-5}$$

If a positively charged particle is at rest at the anode, it has zero kinetic energy, and its potential energy relative to the cathode is qE_b. Equation (3-5), evaluated for the particle at the anode, gives

$$W = T_k = qE_b \tag{3-6}$$

and, in general,

$$W + T = qE_b \tag{3-7}$$

If the particle is free to move, it moves to the cathode, and its kinetic energy on reaching the cathode is $T = T_k = qE_b$. If the mass of the particle is m and its velocity is v, then (3-7) can be written as

$$qE + \frac{mv^2}{2} = \frac{mv_k^2}{2} = qE_b \tag{3-8}$$

where E is the potential of the point at which the particle happens to be. If q, m, E_b, and E are known, the magnitude of the velocity, v, can be found from this relation; the direction of the velocity remains unknown, however. Equations (3-6) to (3-8) apply only to a particle that starts from rest at the anode.

Suppose that the charged particle under observation is an electron. The mass of the electron is $m_e = 9.11 \times 10^{-31}$ kg, and its charge is $-q_e = -1.60 \times 10^{-19}$ coulomb. The electron is charged with negative electricity; therefore the force exerted on it by an electric field is opposite in direction to that which the field would exert on a positive charge. The equations in the preceding paragraphs apply to the electron when the charge is written as $-q_e$. It follows, therefore, that the electron loses potential energy and gains kinetic energy as it moves from a point of low electric potential to a point of high electric potential. If its potential energy is taken as zero at the cathode, then its potential energy is negative at points of higher electric potential.

If an electron at rest at the cathode in Fig. 3-5 is released, it is accelerated by the electric field toward points of higher potential; by this action it loses potential energy and gains kinetic energy.[2] Since this electron was at rest at the cathode, $T_k = 0$, and

$$W + T = 0 \tag{3-9}$$

holds at all times while the electron moves through the field. If the potential at any point is E and the velocity of the electron at that point is v, then (3-9) yields

$$\frac{m_e v^2}{2} = q_e E \tag{3-10}$$

If the electron goes to the anode, it arrives with a kinetic energy given by

$$\frac{m_e v_a^2}{2} = q_e E_b \tag{3-11}$$

where v_a is the velocity of the electron when it reaches the anode. The kinetic energy of the electron is converted to heat by the impact of the electron on the anode.

Suppose that the path s shown in Fig. 3-6 lies in an electric field that is directed along s (either forward or backward) at every point but varies in strength from point to point along the path. If the potential at point a is taken as zero and the potential at each point along the path is plotted versus distance from a, then a curve such as A or B in Fig. 3-6 might result. This curve shows the distribution of potential along path s; potential-distribution curves are very helpful in studying the motion of charged particles in electronic devices. If the potential along the path to the right of a rises as indicated by either curve A or B, then an electron released at a with no initial kinetic energy is accelerated along the path toward b by the electric field. The ordinate to the curve at each point is proportional to the potential energy lost and the kinetic energy gained by an electron in moving from a to that point. Thus if curve B describes the potential along s, the potential and kinetic energies of the electron at points where curve B crosses the axis have the same values as at point a.

It follows from the definition of electric potential in Eq. (3-3) that the slope of the potential-distribution curve is

$$\frac{dE}{ds} = \frac{1}{q}\frac{dW}{ds} \tag{3-12}$$

Solving (3-1) for dW/ds and substituting the result in (3-12) yields

$$\frac{dE}{ds} = -\frac{1}{q}F\cos\theta \tag{3-13}$$

where $F\cos\theta$ is the component of electric force on the charge q in the positive s direction. If the path is chosen so that the total electric force on the charge is in the positive s direction, then $\cos\theta = 1$, and

$$\frac{F}{q} = -\frac{dE}{ds} \tag{3-14}$$

Thus the slope of the potential-distribution curve along this particular path, which is called the potential gradient, is the negative of the force per unit charge in the positive s direction on a positively charged particle.

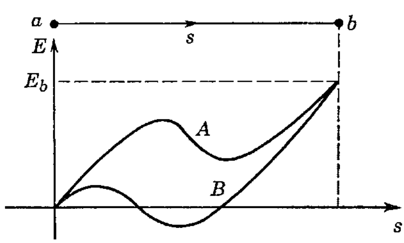


FIG. 3-6. Potential-distribution curves.

The electric field in Fig. 3-6 is specified to be directed along the path s, either in the positive or the negative s direction, at each point. Accordingly, the force on an electron at any point on the path is given by (3-13) as

$$F\cos\theta = q_e\frac{dE}{ds} \tag{3-15}$$

where $\cos\theta$ is either 1 or -1. Since F is the magnitude of the force on the electron, it is a positive quantity. Hence if dE/ds is negative, $\cos\theta$ is negative, and the force on the electron is in the negative s direction. If dE/ds is positive, the force on the electron is in the positive s direction. In short, the magnitude of the force on any electron at any point on path s is proportional to the magnitude of the potential gradient at that point, and it is directed toward the region of higher potential. The force on a positive charge is also proportional to the potential gradient, but it is directed toward the region of lower potential.

If an electron starts from rest at point a in Fig. 3-6, it moves along path s toward point b, and its velocity at each point, given by Eq. (3-10), is

$$v = \sqrt{\frac{2q_eE}{m_e}} = 5.93 \times 10^5 \sqrt{E} \qquad \text{m/sec} \tag{3-16}$$

where E is in volts. Thus the velocity at each point along the path is proportional to the square root of the potential at that point, and it is zero at points where the potential-distribution curve crosses the axis. The velocity is imaginary at points where the curve lies below the axis; an electron starting from rest at point a can never reach points of negative potential. Equation (3-15) shows that the force on the electron is directly proportional to the slope of the potential-distribution curve. The force is to the right when the slope is positive and to the left when the slope is negative; it is zero at points of maximum and minimum potential (except when these occur at the end points). If the potential distribution along s corresponds to curve B in Fig. 3-6, an electron starting from rest at point a is accelerated along the path until it reaches the first point of maximum potential. Beyond this point it slows down and comes to a standstill where the potential-distribution curve crosses the axis. There is a force on the electron at this point directed backward along the path; hence the electron does not remain at this point but moves back to point a. This electron oscillates back and forth indefinitely along this segment of the path.

Suppose that the potential at the minimum point on curve B in Fig. 3-6 is $-E_m$. The kinetic energy of an electron at this point is given by (3-5) as

$$T = T_k - W = T_k - q_eE_m \tag{3-17}$$

where T_k is in this case understood to represent the kinetic energy of the electron at point a. If an electron is to pass the potential minimum and continue to point b, its kinetic energy must always be greater than zero. This is possible only if it leaves point a with an initial kinetic energy T_k that is greater than q_eE_m.

When MKS units are used in the various equations developed above, the unit of energy is the joule. The amounts of energy encountered in the study of the motion of charged particles often are a very small fraction of a joule. For this and other reasons of convenience it is desirable to use a different unit of energy, the electron volt, given by

$$W\ (\text{ev}) = \frac{W\ (\text{joules})}{q'_e} \tag{3-18}$$

where q'_e is a dimensionless number equal to the magnitude of the charge on the electron. Thus the potential energy of a charge q at a point where the potential is E given by

$$W = qE \qquad \text{(joules)} \tag{3-19}$$

$$W = \frac{q}{q'_e}E \qquad \text{(ev)} \tag{3-20}$$

For an electron, Eq. (3-20) becomes

$$W = \frac{-q_e}{q'_e}E = -E \qquad \text{(ev)} \tag{3-21}$$

Thus the potential energy of an electron at a point where the potential is E, when expressed in electron volts, is numerically equal to E. The change in potential energy experienced by an electron in moving through a potential difference of E volts is E electron volts. Thus in moving through a potential rise of 100 volts in going from cathode to anode, an electron loses 100 electron volts of potential energy.

3.4. Electrical Conduction in Crystalline Solids. The quantitative study of the conduction of electricity through solids is a broad field that requires the use of some relatively advanced notions from physics and mathematics. For this reason it is not feasible to discuss the topic in detail here. It is possible, however, to give a semiquantitative picture of the conduction mechanism that permits the important phenomena involved to be understood in a superficial way;[3] that is, the results of detailed theoretical and experimental studies conducted primarily by physicists can be presented and described. In this way a useful insight into the important properties of semiconductor devices can be gained. Such knowledge is helpful to the engineer who is interested in using semiconductor devices in electronic circuits; however, it is not likely to enable him to develop new devices or to improve on those already in existence.

Two crystalline solids that are important in electronics because of their physical properties are germanium and silicon. The discussion that follows is concerned primarily with germanium; however, it applies equally well, in a qualitative sense, to silicon and other crystalline materials.

The normal germanium atom consists of 32 electrons circulating in orbits around a nucleus made up of 32 protons and 41 neutrons; since it contains equal amounts of positive and negative charge it is electrically neutral. Twenty-eight of the electrons are very closely bound to the nucleus; however, the four electrons in the outermost orbits are relatively loosely associated with the atom. These outer four electrons are the valence electrons; they are primarily responsible for the chemical properties of the atom, and it is their presence that permits chemical compounds to be formed. They are required to make the atom electrically neutral, but they represent an excess beyond a preferred chemical state. By way of contrast, the oxygen atom is electrically neutral, but it is deficient by two electrons from a preferred chemical state. If two oxygen atoms and one germanium atom are brought together under favorable conditions, the germanium atom will share two of its excess electrons with each of the oxygen atoms, as illustrated schematically in Fig. 3-7*a*. These three atoms are bound together in a chemical compound called germanium dioxide. This very stable compound, a white powder, is the principal form in which germanium occurs in nature. When joined in this compound and sharing electrons in the manner described, all three of the atoms are closer to a preferred chemical state than when they are separated. The chemical bonds attaching the oxygen atoms to the germanium atom are called electron-pair bonds, or covalent bonds.

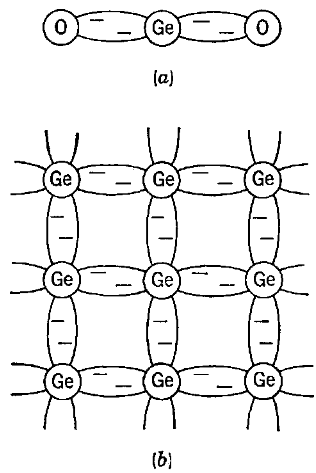


FIG. 3-7. Valence bonds between atoms. (*a*) Germanium dioxide molecule; (*b*) germanium crystal (schematic representation.)

If an aggregation of germanium atoms is brought together under favorable conditions, similar sharing of electrons among the germanium atoms will take place. Each atom will share two electrons in a covalent bond with four of its neighbors in a manner diagramed schematically in Fig. 3-7*b*. Under this condition each atom is closer to a preferred chemical state than when there is no sharing. The covalent bonds hold the atoms fixed in space relative to one another so that the aggregation forms a regular structure, or lattice, in space. This lattice is, of course, a three-dimensional structure; the two-dimensional representation in Fig. 3-7*b* is intended only to represent schematically the relations among the atoms. The entirety of such a regular lattice structure is a crystal. Any particular piece of germanium may consist of a single crystal, or it may consist of many separate crystals oriented in a random manner and joined together at their boundaries. In the latter case the germanium is said to be polycrystalline. The best electronic devices are made from single crystals, although polycrystalline forms can also be used. For example, metallic rectifiers employ polycrystalline forms of the semiconducting material.

If the valence electrons in the crystal of Fig. 3-7*b* are tightly bound in the covalent bonds, there are no free carriers of charge in the crystal and it acts as an insulator. A perfect carbon crystal (diamond) behaves in this manner. If, however, a very strong electric field is applied to such an insulator, the electric forces will tear electrons out of the valence bonds, thereby setting them free and providing mobile carriers of charge; that is, the insulator will break down and become a conductor.

In some materials, such as copper and aluminum, the crystal structure is of such a nature that some of the valence electrons are not bound to any particular location in the crystal. These electrons are free to move through the crystal, and the crystal therefore contains many free charge carriers. Such materials are electrical conductors.

Semiconductors have properties lying between the two extremes described above. At very low temperatures virtually all the valence electrons are bound, and there are essentially no charge carriers present. At room temperature, however, an appreciable number of carriers are created somewhat artificially by thermal energy and similar agents in a manner described below.

In a germanium crystal the valence electrons are not very tightly held in the covalent bonds; only 0.75 electron volt of energy is required to remove an electron from the bond. At room temperatures the particles in the crystal are in constant motion by virtue of thermal energy, and as a result of interactions among the particles, energy is continuously interchanged among them. As a result, many electrons acquire energies in excess of 0.75 electron volt and thereby escape from their bonds. These electrons are free charge carriers. Conditions existing in the crystal when an electron escapes from its bond are pictured in Fig. 3-8. The free electron is represented by the minus sign that is not associated with a covalent bond. An electron is now missing from one of the bonds, and an imperfection, or hole, exists in the regular lattice structure. Associated with this hole, or missing electron, is an excess of positive charge, indicated in Fig. 3-8 by the plus sign in one of the bonds. It is a relatively easy matter for an electron in a nearby valence bond to leave its position and move into the hole left by the thermally ejected electron. When this happens a hole appears in the nearby valence bond just vacated, and in this manner the hole can move from point to point in the crystal, carrying with it a positive charge. Thus there are two carriers of electric charge that are free to move about in the crystal, the free electron with a negative charge and the hole with a positive charge. If there are enough of these carriers present, the crystal may act as a fairly good conductor.

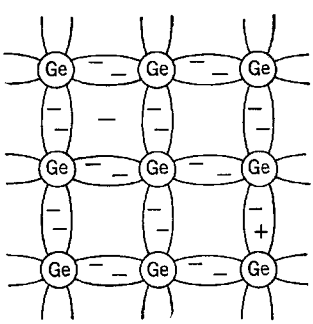


FIG. 3-8. Electron and hole created by thermal agitation.

The free electron and the hole move in a random manner through the crystal as a result of thermal energy and collisions with other particles. The negative electron is not attracted to the positive hole for the following reason. The bound electrons in the vicinity of the hole are attracted by the positive charge of the hole, and the valence bonds are distorted somewhat to permit them to move toward the hole. This negative charge that is displaced toward the hole acts as a screen that masks the presence of the positive hole. In a like manner, bound electrons in the vicinity of the free electron recede somewhat from the free electron, leaving in effect a screen of positive charge that masks the presence of the negative electron. If in their random wandering the hole and electron come close to each other, it is possible that the free electron may move into the empty valence bond represented by the hole. The electron and the hole cancel each other by this recombination process, and charge carriers disappear from the scene in this way.

When the crystal is in thermal equilibrium, the density of the carriers is such that they disappear by recombination just as fast as they are generated by thermal agitation. From a study of the distribution of thermal energy among the particles in the crystal it can be shown that the density of charge carriers (holes plus electrons) is given by

$$n = AT^{3/2} \exp \frac{-q_e E_g}{2kT} \tag{3-22}$$

where the symbolism used is to be read as exp $x = \epsilon^x$, and where T is the absolute temperature, q_e is the magnitude of the electronic charge, E_g is the energy in electron volts required to break the covalent bond (0.75 ev for germanium), k is the Boltzmann constant, and A is a constant that characterizes the material from which the crystal is made. It can be shown that the conductivity of the material is directly proportional to the density of carriers, n; hence it follows from (3-22) that the conductivity of a semiconductor depends strongly on temperature. This fact has been put to useful application in some devices; however, it is principally a source of trouble in semiconductor diodes and transistors because it prevents them from functioning properly at high temperatures.

At room temperature the conductivity of the pure germanium crystal described above is much smaller than the conductivity of an electrical conductor like copper. It is possible, however, to make crystals having a relatively high conductivity at room temperature by building into the crystal lattice a few atoms of another element that differ in a certain way from the basic atoms of germanium. These crystals have an inherent supply of free charge carriers that does not depend on temperature; their conductivity is built in. Crystals that have an inherent supply of free electrons are called *N*-type crystals because conduction occurs through the flow of negative charge. Crystals that have an inherent supply of holes are called *P*-type crystals. These two types of crystals provide the basis for diode and transistor action.

A crystal of *N*-type germanium is illustrated in Fig. 3-9; the *N* character of this crystal results from the presence of a few arsenic atoms in the crystal lattice. In a typical case there is about one arsenic atom for each 10^7 germanium atoms. One of the ways in which the arsenic atom is different from the germanium atom is that it has five rather than four valence electrons. When it enters the lattice structure of the germanium crystal, four of the valence electrons form covalent bonds with adjacent germanium atoms. The fifth valence electron is then only very lightly attached to the parent atom; only 0.01 ev of energy is required to separate it from the atom. Hence at temperatures greater than about 20°K essentially all these excess electrons become free electrons, and the crystal behaves somewhat as a conductor. The conductivity of the crystal depends on the density of carriers, which in turn depends on the density of impurity atoms. Since the normal arsenic atom is electrically neutral, the loss of the free electron leaves a net positive charge associated with the atomic nucleus. This positive charge cannot move through the crystal because the atom is tightly bound in the lattice by the covalent bonds; the circle around the plus sign in Fig. 3-9 signifies that this is a bound charge. The free electron is not attracted to the bound positive charge because of the shielding action described previously. Since the arsenic impurity in the crystal contributes free electrons for conduction, it is called a donor material. Other pentavalent atoms, such as those of antimony and phosphorous, can also be used as donors.

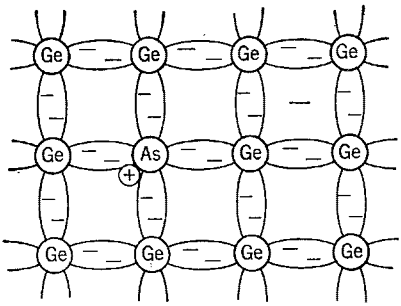


FIG. 3-9. A germanium crystal with arsenic impurity.

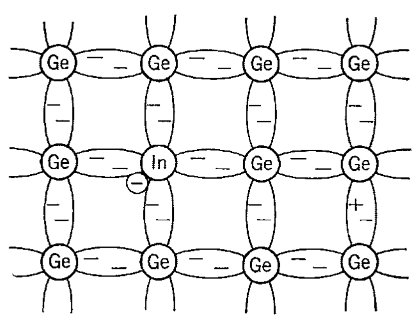


FIG. 3-10. A germanium crystal with indium impurity.

A crystal of *P*-type germanium is shown in Fig. 3-10; the *P* character of this crystal results from the presence of indium atoms in the crystal lattice. One of the ways in which the indium atom differs from the germanium atom is that it has three rather than four valence electrons. When it enters the lattice structure of the germanium crystal, its three valence electrons enter covalent bonds with adjacent germanium atoms; however, there is no electron to form the fourth bond of the normal lattice structure, and therefore a hole exists. This hole is free to move through the crystal in the manner described previously, carrying a positive charge. When the hole moves away from the indium atom an electron has moved into the empty valence bond associated with the indium atom. As a result there is an excess of negative charge associated with this atom; this bound negative charge is indicated in Fig. 3-10 by

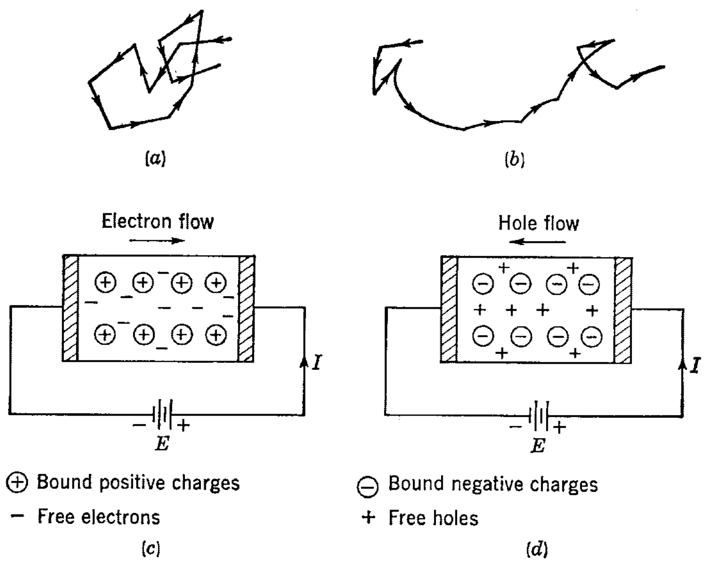


FIG. 3-11. The flow of charged particles in semiconductors. (*a*) Random motion; (*b*) random motion plus drift; (*c*) flow in *N*-type crystal; (*d*) flow in *P*-type crystal.

the encircled minus sign. Since the indium impurity captures electrons from the germanium atoms it is called an acceptor material. Other trivalent atoms, such as those of aluminum, gallium, and boron, can also be used as acceptors.

The type of path described by a free charge carrier in a crystal is illustrated in Fig. 3-11*a*; on the average the particle goes nowhere. The discontinuities in direction correspond to collisions with other particles. If an electric field is applied to the crystal as indicated in Figs. 3-11*c* and *d*, the particle moves under the influence of the field and describes a path such as that shown in Fig. 3-11*b*. The electric field superimposes a drift on the random wanderings of the particle. Figures 3-11*c* and *d* indicate the net motion of charge through *N* and *P* types of germanium under the influence of an applied field.

Holes and electrons are created by thermal agitation in *N*- and *P*-type crystals just as they are in the intrinsic (pure) crystal shown in Fig. 3-8. In an *N*-type crystal, however, there are many free electrons present; hence thermally generated holes quickly recombine with an electron, and as a result there are very few holes present in *N*-type crystals. Electrons are the majority carriers in an *N*-type crystal, and holes are the minority carriers. The converse is true for *P*-type crystals; thermally generated electrons quickly recombine with holes that are present in abundance. Holes are the majority carriers in a *P*-type crystal, and electrons are the minority carriers.

3-5. The *P-N* Junction Diode. The characteristics of semiconductor diodes are described qualitatively in Sec. 3-2. The internal mechanism of the *P-N* junction diode can be explained qualitatively in terms of the properties of semiconductors presented in the preceding section. Two forms of the *P-N* junction are illustrated in Fig. 3-12. In each case the diode is formed from a single crystal of either silicon or germanium, one section of which is *P*-type and one section of which is *N*-type.

The method of making germanium diodes of the two types shown in Fig. 3-12 is first described briefly; similar methods are used in making silicon diodes. The grown junction diode is made by touching a single-crystal seed of germanium to the liquid surface of molten germanium and then slowly withdrawing it. As the seed is withdrawn, the molten germanium crystallizes onto the seed in the form of additional lattice layers, and the crystal grows. By adding an acceptor impurity such as indium to the melt, a *P*-type crystal is obtained. When the *P*-type crystal has attained a suitable size, a donor impurity such as arsenic is added to the melt in quantity sufficient to neutralize and override the acceptor, and the portion of the crystal grown thereafter is *N*-type.

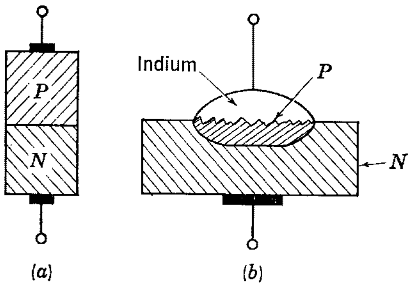


FIG. 3-12. *P-N* junction diodes. (*a*) Grown junction; (*b*) alloy junction.

The alloy junction diode is made[4] by placing a dot of indium on the surface of a wafer of *N*-type germanium and heating the combination to a temperature well above the melting point of indium. Germanium then dissolves into the indium. When the combination is cooled, the dissolved germanium recrystallizes on the original crystal, and the result is again a single crystal of germanium. However, the recrystallized volume is *P*-type germanium because of the presence of indium atoms in the lattice structure.

Figure 3-13*a* shows a junction diode with no voltage applied; the free charge carriers are indicated by the plus and minus signs. Any hole that wanders by diffusion into the N region is quickly canceled by recombination with one of the many free electrons present; hence there are few holes in this region, and the majority carriers are electrons. Similarly, any electron that wanders by diffusion into the P region quickly recombines with one of the many holes present; hence there are few electrons in this region, and the majority carriers are holes. When this system is in equilibrium there is no net flow of charge across the junction.

If a forward voltage is applied to the diode as shown in Fig. 3-13*b*, holes in the P region move across the junction into the N region, and electrons in the N region move across the junction into the P region. These carriers recombine and cancel each other very quickly in the vicinity of the junction so that neither carrier penetrates very deeply into the domain of the other. The view can be taken that electrons flow from the external circuit into the N region and holes flow from the

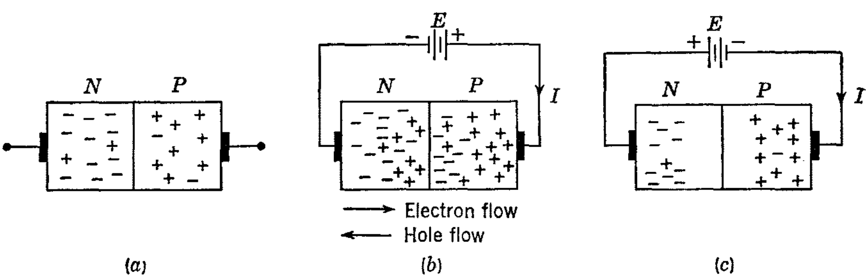


FIG. 3-13. Conduction through a P-N junction diode. (*a*) No voltage applied; (*b*) forward voltage applied; (*c*) reverse voltage applied.

external circuit into the P region; these two carriers move through the crystal to the vicinity of the junction where they cancel each other. This is the condition of easy conduction; it corresponds to the portion of the volt-ampere characteristic in the first quadrant of Fig. 3-2*c*.

If a reverse voltage is applied to the diode as shown in Fig. 3-13*c*, holes in the P region and electrons in the N region both move away from the junction. Minority carriers cannot flow from the external circuit into either the N or the P regions with normal applied voltage, for this action requires the breaking of covalent bonds. Hence ideally there would be no current flow with reverse voltage applied. However, electron-hole pairs are generated continously by thermal agitation in both the N and the P regions. A hole so generated in the N region moves toward the junction, and the associated electron moves into the external circuit. Similarly, a thermally generated electron in the P region moves toward the junction, and the associated hole moves into the external circuit. These thermally generated carriers thereby give rise to a small reverse current. Since this reverse current depends on the rate at which carriers are generated thermally, it is strongly dependent on temperature and is essentially independent of the applied voltage. This current is shown in the third quadrant of the diode characteristic of Fig. 3-2*c*.

The foregoing discussion explains the behavior of an ideal *P-N* junction diode. The characteristics of physical diodes often differ appreciably from those of the ideal *P-N* junction, especially for reverse currents and voltages. These differences result to a large extent from leakage currents on the external surface of the diode and from other more complicated surface phenomena that are not fully understood. Usually the reverse current, called the saturation current, is observed to increase slowly with increasing reverse voltage until the breakdown voltage is reached.

The equation of the volt-ampere characteristic for an ideal junction diode (exclusive of the breakdown region) is simple and can be developed from a consideration of the potential distribution inside the diode.[3] Figure 3-14*a* shows a diode with its external terminals short-circuited; the potential distribution inside the crystal under conditions of thermal equilibrium is shown by the curve beneath the diode. The shape of the potential-distribution curve can be explained in the following way. Holes diffusing from right to left across the junction recombine with free electrons just to the left of the junction, and electrons diffusing from left to right across the junction recombine with holes just to the right of the junction. Thus there are practically no free carriers in a small region on either side of the junction. This region is called the carrier-depletion region. The portion of the depletion region in the *N*-type crystal contains bound positive charges that are not neutralized by negative carriers, and the portion in the *P*-type crystal contains bound negative charges that are not neutralized by positive carriers. These bound charges are said to be uncovered. An electric field extends across the junction between these uncovered charges, and its direction is such as to oppose the diffusion of holes into the *N* region and of electrons into the *P* region. Because of the field across the junction, the potential in the diode has the distribution shown in Fig. 3-14*a*; the potential hill, or potential barrier, in this curve is associated with the field between the uncovered charges. Free positive charges in the crystal tend to move to points of lower potential, and free negative charges tend to move to points of higher potential.

The potential-distribution curves of Fig. 3-14 do not show what happens to the potential at the metallic terminals through which external connections are made with the diode. The nature of these connections is not fully understood. However, negligible current flows in the external circuit when no voltage is applied to the diode; hence there is negligible potential drop across the external circuit. Therefore the potential drops at the metallic contacts with the crystal just compensate for the potential difference across the P-N junction, and the metallic terminals are at equal potentials. When the contacts are properly formed they do not enter into the analysis of the diode behavior; the diode characteristic depends only on the potential barrier at the P-N junction.

The height of the potential hill at the P-N junction is determined by an equilibrium between two factors: (1) the generation of electron-hole pairs on both sides of the junction by thermal agitation, and (2) the diffusion of carriers across the junction against the potential barrier

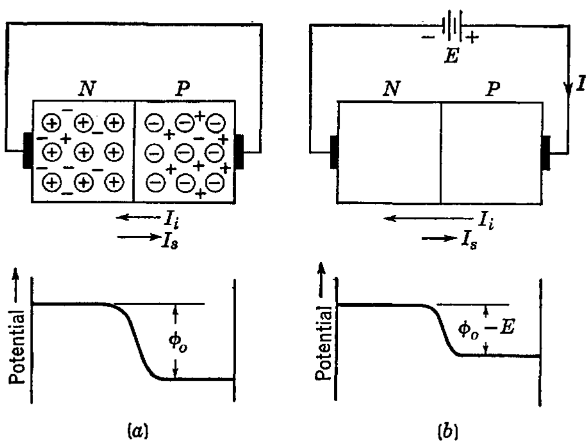


FIG. 3-14. Potential distribution in a junction diode. (*a*) No voltage applied; (*b*) forward voltage applied.

by virtue of thermal energy. Under equilibrium conditions, with no current in the external circuit, these two factors result in equal and opposite currents across the junction as indicated in Fig. 3-14*a*. The thermally generated electron-hole pairs result in a small stream of electrons from the low to the high potential region and a similar small stream of holes from the high to the low potential region. This current, which is also the saturation current, is symbolized by I_s. It depends on the density of thermally generated carriers; hence it depends strongly on temperature and is independent of the voltage applied to the diode.

The second component of current across the junction is associated with those free electrons in the N region and holes in the P region that gain enough kinetic energy from thermal agitation to climb the potential hill. If the height of the potential hill is ϕ_o volts, the potential energy gained by a hole in climbing the hill is, from Eq. (3-3),

$$W_o = q_e\phi_o \tag{3-23}$$

Any hole having more than this amount of kinetic energy directed toward the junction can move from the P region into the N region. A similar condition applies to the electrons in the N region. From a more detailed study of the random motion of such particles it can be shown that if n_o is the total number of carriers starting across the potential barrier per second, then the number capable of getting all the way across in each second is

$$n = n_o \exp \frac{-W_o}{kT} = n_o \exp \frac{-q_e \phi_o}{kT} \tag{3-24}$$

where k is the Boltzmann constant and T is the absolute temperature. Multiplying (3-24) by q_e gives the amount of charge crossing the junction per second as a fraction of the total amount of charge that attempts the crossing in each second; thus

$$I_i = I_o \exp \frac{-q_e \phi_o}{kT} \tag{3-25}$$

When no voltage is applied to the diode, negligible current flows in the external circuit, and the two components of current across the junction are equal in magnitude and opposite in direction. Thus

$$I_s = I_i = I_o \exp \frac{-q_e \phi_o}{kT} \tag{3-26}$$

The value of I_s is fixed by the rate at which electron-hole pairs are generated by thermal agitation. Under thermal equilibrium with no voltage applied, the height of the potential barrier adjusts itself so that the current I_i just equals the saturation current I_s. If ϕ_o is too small, more holes move into the N region and more electrons move into the P region, more bound charges are uncovered, and the height of the barrier is increased. If ϕ_o is too large, the converse occurs.

When a forward voltage is applied to the diode as shown in Fig. 3-14b, the height of the potential hill is reduced by the amount of the applied voltage, E. Many more carriers have enough kinetic energy to cross this reduced barrier, and I_i increases accordingly. Under this condition Eq. (3-25) becomes

$$\begin{aligned} I_i &= I_o \exp \frac{-q_e(\phi_o - E)}{kT} \\ &= I_o \left(\exp \frac{-q_e \phi_o}{kT}\right) \left(\exp \frac{q_e E}{kT}\right) \end{aligned} \tag{3-27}$$

Substituting (3-26) into (3-27) yields

$$I_i = I_s \exp \frac{q_e E}{kT} \tag{3-28}$$

With E a positive number, I_i is greater than I_s, and the current in the external circuit is the difference:

$$I = I_i - I_s = I_s \exp\left(\frac{q_e E}{kT}\right) - I_s$$
$$= I_s\left[\exp\left(\frac{q_e E}{kT}\right) - 1\right] \tag{3-29}$$

When a reverse voltage is applied to the diode, the height of the potential barrier is increased to $\phi_o - E$, where E is now a negative number. The number of carriers capable of crossing the barrier is thereby reduced, I_i becomes less than I_s, and a reverse current flows through the

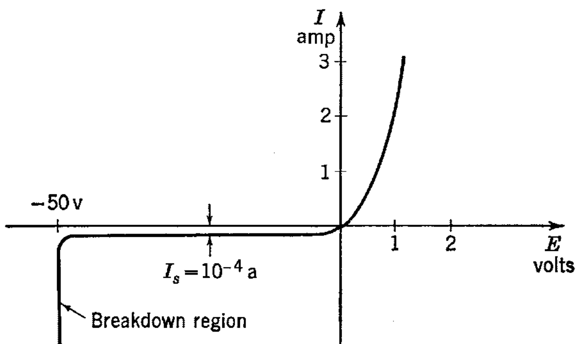


FIG. 3-15. A junction-diode characteristic.

diode. The value of the reverse current is given by (3-29) when the proper negative value is substituted for E.

Equation (3-29) gives the diode terminal current as a function of the applied voltage; it is the equation of the diode volt-ampere characteristic. The typical diode characteristic of Fig. 3-2 is repeated in Fig. 3-15; for clarity the forward voltage and current are not plotted to the same scale as the reverse voltage and current. Substituting numerical values for q_e and k and taking room temperature to be 300°K yields

$$I = I_s(\epsilon^{39E} - 1) \tag{3-30}$$

When E is more positive than about 0.1 volt, (3-30) becomes

$$I \approx I_s \epsilon^{39E} \tag{3-31}$$

This is the equation for the forward-current portion of the volt-ampere characteristic. When E is more negative than about -0.1 volt, (3-30) becomes

$$I \approx -I_s \tag{3-32}$$

This is the equation for the reverse-current portion of the volt-ampere characteristic up to the point of breakdown.

The equations developed above show the dependence of the diode current on temperature. The effect of temperature on the reverse current is of primary importance, for the reverse current represents a departure from the ideal characteristic. The reverse current in a germanium diode approximately doubles for each 10°C rise in temperature; the rectifying properties of these diodes is seriously impaired at temperatures above 75°C. It is in this respect that silicon is far superior to germanium. Although the reverse current in a silicon diode increases with temperature at about the same rate, the value of the reverse current is several orders of magnitude less than that in a germanium diode. Hence the silicon diode retains its rectifying properties up to temperatures as high as 200°C. The ratio of forward to reverse current, which is a measure of the diode quality, is given by Eq. (3-29) as

$$\frac{I}{I_s} = \exp\left(\frac{q_e E}{kT}\right) - 1 \tag{3-33}$$

The mechanism of the breakdown that occurs when excessive inverse voltage is applied to the diode can also be explained in terms of the potential distribution in the diode. With a small inverse voltage applied to the diode, a potential difference exists across the depletion region at the junction; this potential difference is equal to the height of the potential hill shown in Fig. 3-14. Thermally generated holes in the N region and thermally generated electrons in the P region are accelerated across the depletion region, and they gain kinetic energy in the process. These carriers collide with bound particles in the depletion region and near its boundaries, and they exchange energy with the bound particles. If the reverse voltage applied to the diode is increased sufficiently, the thermally generated carriers crossing the depletion region will gain enough kinetic energy from the strong electric field to knock bound electrons out of the covalent bonds. This process is called ionization. The free carriers created by ionization are in turn accelerated by the field and may themselves have ionizing collisions, creating still more free carriers. Thus the originally small saturation current is greatly multiplied by this process, which is described as avalanche breakdown. When the reverse voltage applied to the diode is sufficient to make efficient ionizing agents of the free carriers, further increases in diode current require negligible increases in reverse voltage; in fact, the voltage drop across the diode may decrease slightly with further increases in current. The voltage at which breakdown occurs is often called the Zener voltage, for originally the breakdown was erroneously thought to result from the Zener effect, in which electrons are torn out of covalent bonds by a strong electric field.

When junction diodes are used in the circuits described in Chap. 2, the peak inverse voltage appearing across the diodes must not exceed the breakdown voltage. This fact limits the direct voltage that can be obtained from rectifiers using junction diodes. On the other hand, however, if the current through the diode under breakdown conditions is limited to a value giving a safe power dissipation, the diode is not damaged in any way by operation in the breakdown region. This fact suggests new applications for the diode. For example, a properly made diode can be used to provide a constant voltage drop that is nearly independent of diode current and temperature. The breakdown voltage can be controlled over the range between one or two volts and several hundred volts by controlling the appropriate factors in the manufacturing process.

The abrupt change in the diode characteristic at the breakdown voltage is similar to that occurring at the zero voltage point. Hence the device can be operated as a diode about the breakdown point. When the inverse voltage exceeds the breakdown value, the diode conducts readily; when the inverse voltage is less than the breakdown value the diode behaves essentially as an open circuit. This mode of operation has the advantage that the diode can switch from the conducting to the nonconducting state at the breakdown point much faster than it can at the zero voltage point; thus the high-frequency performance of the diode is improved greatly. It has the disadvantage that for safe power dissipation the diode current must be limited to a relatively small value, for the conducting voltage drop across the diode is relatively large.

When a forward voltage is applied to the junction diode, charges are distributed in the diode as shown in Fig. 3-13*b*. Many holes cross the junction and enter the N region as minority carriers, and many electrons enter the P region as minority carriers. When the voltage applied to the diode is reversed, the steady-state conditions are as pictured in Fig. 3-13*c*. A depletion region exists in the vicinity of the junction, and there is a very small flow of thermally generated carriers across the junction. This new condition is not reached instantaneously, however. When the applied voltage is reversed, a reverse current flows until the minority carriers on each side of the junction either disappear by recombination or return to their natural domain. This reverse current may last for several microseconds. Thus the diode acts very much as a capacitor that must be charged to a new voltage when the applied voltage is reversed. The capacitance is associated with presence of minority carriers on each side of the junction, and it is often referred to as the junction storage capacitance. It is this capacitance that limits the high-frequency performance of the diode when it is operated in the normal mode. When the diode is operated in the breakdown region, the distribution of carriers is like that shown in Fig. 3-13*c* with the addition of many carriers generated by ionization in the depletion region, which is a region of high field strength. When the applied voltage is abruptly reduced below the breakdown value, the carriers in the depletion region are rapidly swept out by the strong field, and there is no reverse current. As a result, the switching action is much faster than when the diode is operated in the normal mode.

The junction diodes described above can be used in any of the circuits presented in Chap. 2 provided the peak inverse voltage does not exceed the breakdown voltage and provided the forward current does not cause excessive heating. In many cases the forward voltage drop across the diode and the reverse current through the diode are so small that they can be neglected. In such cases the diode can be represented as an ideal diode, and the circuits behave in the manner described in Chap. 2.

3-6. Conduction through Vacuum Diodes. The study of the conduction of electric current through vacuum diodes consists of two more or less separate parts, a study of the emission of electrons from solids and a study of the flow of current in a vacuum. Both of these parts have features in common with the study of conduction through semiconductor diodes.

In accordance with the discussion of Sec. 3-1, the action in a vacuum diode depends on the copious emission of electrons from the cathode. The cathode is made of a metallic conductor which, as described in Sec. 3-4, consists of atoms bound in crystal lattices, valence electrons bound in valence bonds, and free electrons that are not bound to any particular place in the metal. Electronic emission is the process by which the free electrons escape from the surface of the metal.

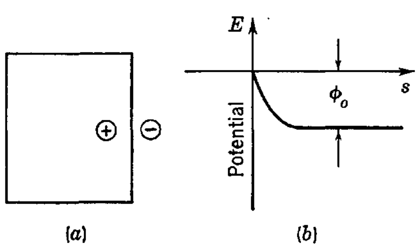


FIG. 3-16. Electron emission from a metal. (*a*) An electron escaping from the surface of a metal; (*b*) potential-distribution diagram for the region near the surface of a metal.

Figure 3-16*a* pictures symbolically a free electron in the process of escaping from a metal. As indicated in the figure, the escaping electron leaves behind an excess of positive charge associated with the atomic nuclei in the metal. Since there is a force of attraction between these charges, work must be done on the electron to move it away from the metal; the electron therefore gains potential energy in leaving the metal. If the potential energy of an electron at any point relative to the cathode is W and the electric potential at that point is E, then $W = -q_eE$. If the magnitude of the potential at points remote from the metal is designated ϕ_o, as indicated in Fig. 3-16*b*, the potential energy gained by an electron in escaping completely from the metal is $W_o = -q_e(-\phi_o) = q_e\phi_o$. The energy required for a complete escape is a property of the metal and is called the work function of the metal; ϕ_o is the work function expressed in electron volts.[2]

The potential barrier encountered by an escaping electron at the surface of a metal is similar to the potential barrier at the junction of a semiconductor diode. As in the case of the junction diode, only those electrons having kinetic energies directed toward the barrier in excess of ϕ_o ev can surmount the barrier. The number of electrons capable of crossing such a barrier in each second is given by Eq. (3-24). Hence the emission current is

$$I_e = I_o \exp \frac{-W_o}{kT} = I_o \exp \frac{-q_e\phi_o}{kT} \tag{3-34}$$

This equation has exactly the same form as (3-25), for the two equations describe the same phenomenon. The factor I_o, which is the emission current that would exist if ϕ_o were zero, can be evaluated from further considerations of the motion of electrons inside the metal. The resulting equation for the emission current per unit of cathode area is

$$I_e = AT^2 \exp \frac{-q_e\phi_o}{kT} \tag{3-35}$$

where A is a constant depending on the metal from which the cathode is made.

The work functions of common materials range from 1 to 5 ev, and the emission coefficients A range from values much less than unity up to about 100 amp/(cm^2)(°K^2). It follows from these facts that emission currents are extremely small at room temperatures; to obtain useful emission currents, vacuum-tube cathodes must be heated to a suitable temperature. The list of materials that can be used as thermionic cathodes is quite restricted, for materials that have small work functions tend to have low melting points, and vice versa. Almost all thermionic cathodes are made of either nickel coated with barium and strontium oxides, thoriated tungsten, or pure tungsten. The work functions of these cathodes are, respectively, 1.0, 2.6, and 4.52 ev. The relatively small differences in the work functions of these cathodes make very great differences in emission current because of the exponential relationship given by (3-35). The normal operating temperatures for these three cathodes are, respectively, 750, 1700, and 2100°C. Since the oxide-coated cathode is the most efficient, it is by far the most widely used. It is not suitable for high-voltage applications, however; thoriated-tungsten and pure-tungsten cathodes are used in such applications. Tubes in which the current flow depends on the emission of electrons from a hot cathode are called thermionic tubes.

A phenomenon related to the foregoing discussion, that of contact potential difference,[2] is illustrated in Fig. 3-17. Figure 3-17*a* shows a vacuum diode with its electrodes connected together by a short circuit. Now suppose that the work function of the cathode is $\phi_{oc} = 1.0$ ev. The potential just outside the cathode is -1.0 volt relative to the cathode, as shown in Fig. 3-17*b*. Suppose further that the work function of the anode is $\phi_{oa} = 4.0$ ev. The potential just outside the anode is -4.0 volts relative to the anode (and cathode), as also shown in Fig. 3-17*b*. It then follows that there is a potential difference of 3.0 volts between a point just outside the cathode and one just outside the anode and that there is an electric field in the interelectrode space. This electric field exerts a force on electrons in the interelectrode space and thereby affects their motion. In the particular case described, the field resulting from the contact potential difference opposes the flow of electrons from cathode to anode.

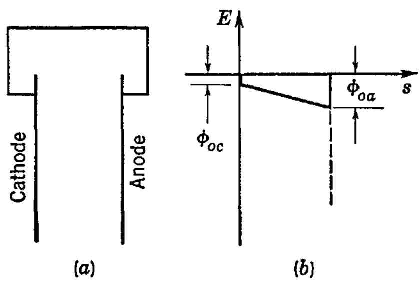


FIG. 3-17. Contact potential difference. (*a*) A vacuum diode with electrodes short circuited; (*b*) potential distribution in the interelectrode space.

If a battery is connected between the electrodes, making the potential of the anode E_b volts relative to the cathode, the apparent anode potential in so far as an electron in the interelectrode space is concerned is

$$E_b' = E_b - \phi_{oa} + \phi_{oc} \tag{3-36}$$

The contact potential is often small compared with the applied voltage, and the usual practice is to ignore it, even in some cases where it is not really negligible.

The factors regulating the flow of current through a diode can be examined with the aid of the diagrams in Fig. 3-18. Figure 3-18*a* represents a diode in which the electrodes are parallel planes of infinite extent. When there is no emission from the cathode in this simple geometry, the electric force on a charged particle has the same strength at every point in the interelectrode space, and its direction at every point is perpendicular to the electrodes. Hence an electron starting from rest at the cathode moves toward the anode along a straight path, such as *s*, that is normal to the electrodes. The equation for the electric potential,

$$E = \frac{W}{q} = \frac{-1}{q} \int F(\cos \theta) \, ds \tag{3-37}$$

takes a simple form along this path because both F and θ are constant, the latter being 180°. Thus

$$E = \frac{F}{q} \int ds = \frac{F}{q} s \tag{3-38}$$

Using the fact that $E = E_b$ when $s = d$ yields

$$E = \frac{E_b}{d} s \tag{3-39}$$

The potential-distribution curve for the path s when there is no emission from the cathode is thus a straight line as shown by curve A in Fig. 3-18c.

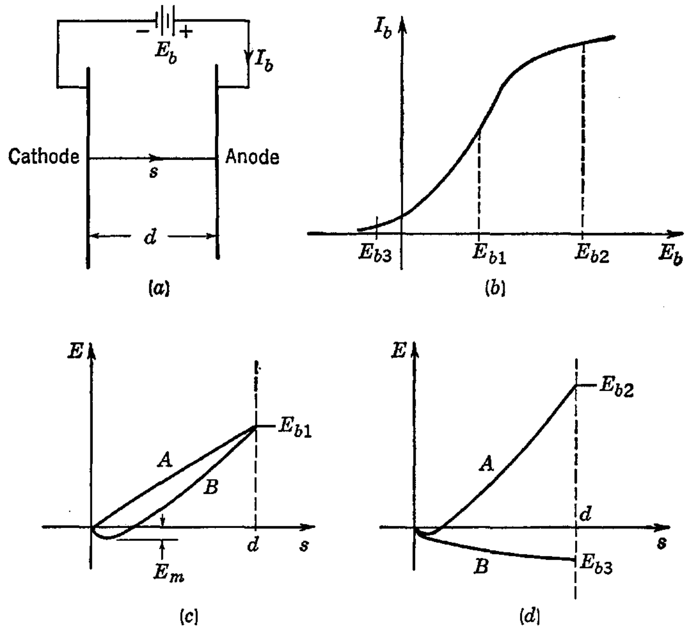


FIG. 3-18. Conduction through a vacuum diode. (a) The diode; (b) the volt-ampere characteristic; (c) potential-distribution curves with and without space charge; (d) potential-distribution curves for other plate voltages.

If the cathode is heated to normal operating temperature, the diode exhibits a volt-ampere characteristic of the form shown in Fig. 3-18b. The shape of this curve can be explained qualitatively in terms of the potential distribution along path s. With the cathode at operating temperature and the anode held positive relative to the cathode, electrons issuing from the cathode move across the interelectrode space in large numbers, and the space contains many electrons in transit to the anode. Since each electron carries a charge, there is a space charge in the interelectrode space; this space charge alters the potential distribution in the space. Consider the work required to move a small positive charge from the cathode toward the anode when appreciable space charge is present. Since the negative space charge exerts an attractive force on the positive charge, less work is required to move the test charge away from the cathode than would be required in the absence of negative space charge. Hence the potential of points in the interelectrode space is reduced by negative space charge, and the potential-distribution curve is depressed as illustrated by curve B in Fig. 3-18c.

Equilibrium is reached under the conditions described above with the potential-distribution curve depressed so that a potential minimum is established a short distance in front of the cathode; the magnitude of this potential minimum is indicated in Fig. 3-18c as E_m. In this way a potential barrier for electrons is established in front of the cathode, and only those electrons having kinetic energies greater than E_m ev directed toward the barrier can pass the potential minimum and continue to the anode; the remainder are stopped by the potential hill and return to the cathode. Again, the number of electrons capable of crossing the barrier is given by Eq. (3-24), and the current through the diode is

$$I_b = I_o \exp \frac{-q_e E_m}{kT} \tag{3-40}$$

If the current tends to increase from this value, the space charge becomes greater, E_m increases, and I_b is restored to the equilibrium value. If the current tends to decrease, the converse action takes place. Since the current is limited by the potential minimum, which in turn is established by the space charge, the current is said to be space-charge-limited under these conditions. The corresponding point on the diode volt-ampere characteristic is shown at E_{b1} in Fig. 3-18b.

If the anode potential is increased, the potential-distribution curve is raised, and a new equilibrium is reached at a larger value of anode current. If the anode potential is made sufficiently great, however, a point will be reached at which the potential minimum is fixed by the work function of the cathode rather than by the space charge. The potential-distribution curve for this condition has the form shown by curve A in Fig. 3-18d; the corresponding point on the diode characteristic is indicated at E_{b2}. Under these conditions the value of the minimum potential is more or less firmly set by the properties of the cathode, and all electrons with energies greater than the work function of the cathode flow to the anode. Accordingly, the anode current is given by Eq. (3-34):

$$I_b = I_e = I_o \exp \frac{-q_e\phi_o}{kT} \tag{3-41}$$

To the extent that the potential minimum is fixed by the work function of the cathode and is not influenced by the anode potential, I_b is independent of E_b. In the case of pure-tungsten cathodes the potential minimum is very firmly fixed, and the diode current does not increase very rapidly with further increases in E_b; a pronounced saturation is observed. As indicated by (3-41), however, I_b increases rapidly with temperature. Under these conditions the current is said to be temperature-limited. In the case of oxide-coated cathodes the potential minimum is not firmly fixed, and no pronounced saturation occurs. However, prolonged operation with the current limited by the cathode rather than by space charge usually results in permanent damage to oxide-coated cathodes. Vacuum tubes are normally operated in the space-charge-limited regions of their characteristics.

As the anode potential is reduced under space-charge-limited conditions, the potential minimum drops lower, and it moves toward the anode. When E_b is made sufficiently negative (a fraction of a volt), the potential minimum occurs at the anode, as illustrated by curve B in Fig. 3-18d. Under this condition, $E_m = -E_b$, since E_m is the magnitude of the potential minimum. It then follows from Eq. (3-40) that

$$I_b = I_o \exp \frac{q_e E_b}{kT} \tag{3-42}$$

A point on the diode characteristic corresponding to this condition is shown at E_{b3} in Fig. 3-18b. In the region where (3-42) applies, E_b is negative. In this range of E_b the plate current is relatively small and is an exponential function of the plate voltage. There are a few useful applications for this exponential volt-ampere relation.

As mentioned above, vacuum tubes are usually operated with space-charge-limited current. Under this condition the plate current is given by Eq. (3-40). Equation (3-40) reveals the mechanism by which space charge limits the flow of current in a vacuum diode, but it is not useful in calculating the diode current, for the value of the potential at the minimum is not known. Moreover, E_m depends in a complicated way on the applied voltage, the cathode temperature, and the current I_b that is to be calculated. However, another approach to the problem which makes use of certain reasonable approximations leads to the three-halves-power law,

$$I_b = KE_b^{3/2} \tag{3-43}$$

for the volt-ampere law of the vacuum diode.[2,5] The perveance of the tube, K, is a constant depending on the geometry of the electrodes. The derivation of (3-43) assumes, among other things, that the potential minimum lies at the surface of the cathode, that the minimum potential is zero, and that all electrons emerge from the cathode with zero initial velocity. It is shown in Sec. 3-9 that this equation is of very limited value in the analysis and design of vacuum-tube circuits because it is a nonlinear relation between I_b and E_b.

If a diode is in operation with E_b volts applied and with I_b amp flowing, it is clear that it absorbs energy at the rate of

$$P_b = E_b I_b \qquad \text{watts} \tag{3-44}$$

The effect of this power absorption may not be so clear; therefore it merits further consideration. According to Eq. (3-11), an electron emerging from the cathode with negligible initial velocity arrives at the anode with a kinetic energy given by

$$\frac{m_e v_a^2}{2} = q_e E_b \tag{3-45}$$

This energy is converted to thermal energy by the impact of the electron on the anode. If the number of electrons reaching the anode each second is n, then the rate at which heat is generated at the anode by electron impacts is

$$P = n q_e E_b \tag{3-46}$$

But nq_e is the amount of charge that arrives at the plate each second; that is, it is the plate current I_b. Hence

$$P = I_b E_b = P_b \tag{3-47}$$

Thus the electrical energy delivered to the diode is all converted to thermal energy at the anode by electron impacts.

The thermal energy generated at the anode is associated with a rise in temperature of the anode. Since the diode is in an evacuated envelope, little heat is lost by conduction or convection. However, with the temperature of the anode higher than that of surrounding objects, heat energy is radiated from the plate to the environment at a rate depending on the temperature difference. An equilibrium is established at that temperature at which heat energy is radiated to the surroundings just as fast as it is generated. (The total heat energy appearing at the anode includes radiant heat from the cathode as well as heat generated by electron impacts.) If heat is generated at too great a rate, the temperature rise will be excessive, and the anode will vaporize or melt. Thus there is a limit on the permissible power dissipation of a tube. For most small tubes the maximum permissible plate dissipation, $P_{b\max}$, is in the range of 1 to 10 watts, average value.

In addition to the limitation on the permissible plate dissipation, there are separate limitations on the maximum permissible plate voltage and current. The maximum instantaneous plate current must be limited to a value that will not damage the cathode if the tube has an oxide-coated cathode, and the maximum instantaneous plate voltage must not be great enough to break down the insulation between the external leads connecting to the electrodes. The maximum permissible plate current, voltage, and power dissipation are specified by the manufacturer.

The foregoing discussion of the vacuum diode is concerned with static conditions in which the currents and voltages do not vary with time. When very-high-frequency voltages are applied to the diode, it is necessary to take into account the fact that the plate and cathode serve as the electrodes of a capacitor and that the total diode current is the sum of the conduction current and the capacitive current. At high frequencies the capacitive current may be comparable with the conduction current, and since it flows equally well in both directions, it impairs the diode characteristics. The nature of the interelectrode capacitance is rather complicated, for it depends on the amount and distribution of the space charge. It is customary to represent the capacitive effect approximately by a fixed capacitance of a few micromicrofarads connected in parallel with the diode.

The vacuum diode is discussed above in terms of the idealized case of infinite, plane, parallel electrodes. Practical diodes cannot use such a simple geometry. The details of electron flow in a practical geometry are considerably more complicated; indeed, it is not possible to make a complete study of the electron flow in any but the simplest geometries. Nevertheless, the basic principles developed in terms of the idealized diode apply equally well to diodes of all geometries.

3-7. Gas-filled Thermionic Diodes. When the anode of a vacuum diode is made positive relative to its cathode, electrons move across the interelectrode space from cathode to anode, and there is a negative space charge in the interelectrode space. This negative space charge depresses the potential-distribution curve in the interelectrode space and thereby limits the current that flows with a given applied voltage. Larger currents can be obtained with a given applied voltage and tube size if the negative space charge can be neutralized in some way. For example, if a grid of fine wires is placed in the region of high negative space charge and is held positive relative to the surrounding space, then the negative space charge is partially neutralized by the positive charge on the wires, the potential distribution curve is raised, and the anode current increases. A few tubes using such a space-charge grid were built in the early days of electronics. However, better results can be obtained in rectifiers intended for low-frequency operation by the method described below.

If a small amount of liquid mercury is included inside the envelope of a thermionic diode such as the one discussed in Sec. 3-6, the space inside the envelope will contain mercury vapor, mercury atoms that have evaporated from the surface of the liquid mercury. If the anode is then made positive relative to the cathode, electrons move from the cathode toward the anode, and many of them collide with mercury atoms on the way. If the colliding electron has enough kinetic energy (10.39 ev for mercury vapor), the collision may separate a valence electron from its parent atom. Thus two new charged particles, a negatively charged electron and an incomplete, positively charged mercury atom called a positive ion, are created by ionization of a mercury atom that was originally neutral. The relatively light electron created by ionization moves

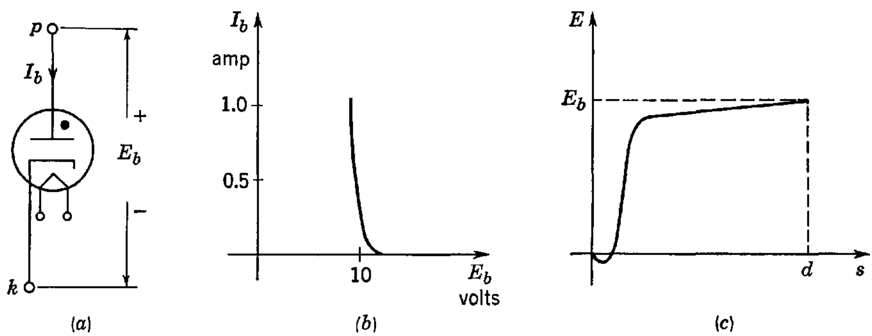


FIG. 3-19. The gas-filled diode. (*a*) The diode; (*b*) the volt-ampere characteristic; (*c*) the potential distribution in the interelectrode space.

rapidly to the anode, while the much heavier positive ion moves slowly toward the cathode. If many positive ions are created in this way, they neutralize the negative space charge of the electrons in most of the interelectrode space, the potential-distribution curve inside the tube is raised, and large currents flow with small impressed voltages.

There is very little recombination of electrons with positive ions in the space inside the tube, for it turns out that it is highly improbable that an electron and a positive ion will meet under conditions that will allow them to recombine and at the same time conserve both momemtum and energy. Recombination takes place primarily at the surfaces of the electrodes and the surrounding envelope; the presence of the third body permits both momentum and energy to be conserved when the recombination takes place.

A gas-filled diode is represented schematically in Fig. 3-19*a*; Fig. 3-19*b* shows a typical volt-ampere characteristic for such a diode, and Fig. 3-19*c* shows the potential distribution along a typical path between the cathode and anode when the tube is conducting a normal current.[2] The volt-ampere characteristic shows that when the tube conducts, the voltage drop across the tube is nearly independent of the plate current; it decreases slightly with increasing plate current, especially at low currents. This characteristic is quite similar to that of the ideal diode except that it is shifted to the right by about 10 volts, the conducting tube drop. The potential distribution curve shows that throughout most of the interelectrode space, from the plate almost to the cathode, the potential gradient, and hence the electric-field strength, is very small. In this region, called the plasma, the electronic and the positive-ion space charges neutralize each other, and the particles move in a random manner, colliding with each other and exchanging energy, in much the same way that holes and electrons move in a semiconductor. Superimposed on this random motion is a drift of electrons toward the anode and of positive ions toward the cathode. For a short distance in front of the cathode, greatly exaggerated in Fig. 3-19*c*, the potential distribution is much like that in a vacuum diode with very small plate-to-cathode spacing; nearly all the tube drop appears across this region. Positive ions flowing across this cathode sheath tend to raise the potential-distribution curve and prevent the potential minimum in front of the cathode from being very large in spite of a large electron current flowing across the sheath into the plasma. Any amount of current, up to the limit of emission from the cathode, can flow across the tube if positive ions are generated in sufficient numbers in the plasma.

The kinetic energy gained by an electron in moving across the cathode sheath into the plasma is almost E_b ev. This energy is approximately equal to the minimum energy that an electron must have in order to ionize an atom of the gas in the tube, although, because of multiple-step ionizations, it is often slightly less than this amount. It corresponds to a tube drop of approximately 10 volts in mercury-vapor tubes and between 10 and 20 volts for other gases that are commonly used. After entering the plasma some of the electrons have ionizing collisions after traveling a short distance, others must travel almost to the plate before having such a collision, and still others reach the plate without having a collision. Thus ions are generated in all parts of the plasma.

At small values of plate current there are few electrons in transit, and the rate of ion production would drop to a low value were it not for the fact that the tube drop rises somewhat, thereby increasing the ionizing efficiency of the electrons and maintaining the rate of ion production at the value needed to sustain the plasma. At very small values of current, however, the plasma cannot be sustained, the ions disappear by surface recombination, and the tube behaves somewhat like a vacuum diode. The current flowing under this condition is too small to show on the volt-ampere characteristic of Fig. 3-19*b*.

Certain precautions that must be observed in the use of gas-filled tubes can be explained in terms of the simple rectifier circuit shown in Fig. 3-20. When the supply voltage e_s is greater than the conducting tube drop E_o, the diode conducts, and as shown by the volt-ampere characteristic of Fig. 3-19*b*, the voltage drop across the tube is approximately constant at the value E_o regardless of the value of the tube current. Thus the circuit of Fig. 3-20*b* is approximately equivalent to the rectifier circuit when e_s is greater than E_o. It is clear from this circuit that when e_s is greater than E_o, the diode current is limited only by the resistance R_L. If this resistance is too small, excessive current flows and the diode is destroyed. It is always necessary to ensure that the current through gas-filled rectifiers is limited to a safe value by an external impedance in series with the tube.

Another precaution that must be observed in the use of gas-filled tubes arises when power is applied to the circuit. When the circuit of Fig. 3-20*a*

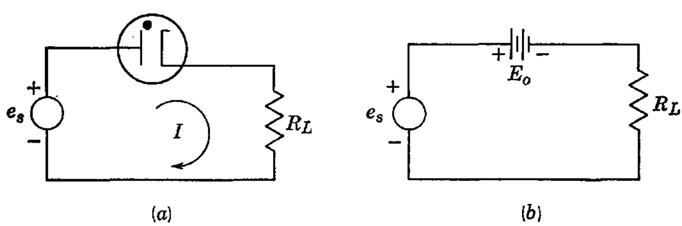


FIG. 3-20. A gas-diode rectifier. (*a*) Circuit; (*b*) an approximately equivalent circuit for $e_s > E_o$.

is in normal operation, the load resistance limits the tube current to a safe value. The voltage drop across the tube when it is conducting is about 10 volts, and the remainder of the applied voltage appears as an IR drop across the load resistance. However, if the a-c supply is applied at the same time that the cathode-heating power is turned on, the ensuing sequence of events may lead to the destruction of the cathode. The cathode temperature rises slowly after the heater power is applied, and there is negligible electron emission until some time has elapsed. During this time there is no current in R_L, and the full voltage of the a-c source appears across the diode. As the cathode approaches its operating temperature a small temperature-limited current begins to flow during the positive half cycles of the supply voltage, there is a small drop across R_L, and the drop across the tube is still relatively large. Electrons moving toward the plate of the tube ionize gas atoms, and the resulting positive ions are accelerated toward the cathode by the electric field in the interelectrode space. Since the drop across the tube is still relatively large, some of these ions move through a large potential difference and strike the cathode with correspondingly large kinetic energies. This positive-ion bombardment is very likely to destroy the delicate oxide coating of the cathode. Experiment shows that the destructiveness of positive-ion bombardment increases rapidly when the conducting tube drop exceeds about 22 volts.[2] In order to avoid cathode damage from this source it is necessary to allow the cathodes of gas-filled rectifiers to reach normal operating temperature, so that normal current can flow and normal voltage drop can appear across the load, before applying power to the plate circuit. The warm-up time required varies from 30 sec to 5 min, depending on the type of tube.

Circuits using gas-filled diodes must be designed so that the maximum permissible instantaneous current and the maximum permissible inverse voltage are not exceeded. As in the case of the vacuum diode, the maximum permissible current is limited by the emission capabilities of the cathode; typical values range from a few hundred milliamperes to a few

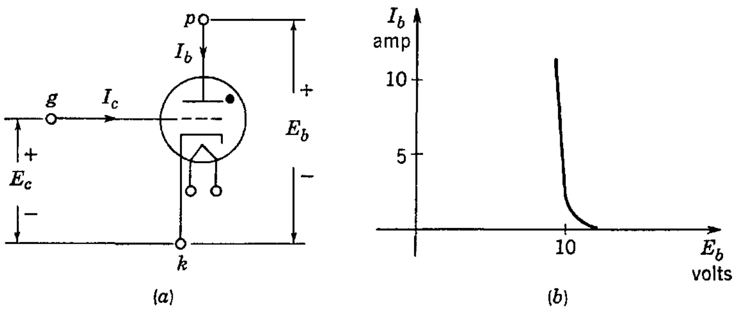


FIG. 3-21. The thyratron. (*a*) The tube; (*b*) the volt-ampere characteristic.

hundred amperes. The maximum permissible inverse voltage is limited by the possibility of an arc-over in the reverse direction through the gas inside the tube; typical values range from a few hundred volts to several thousand volts. Both the maximum current and the maximum inverse voltage that the tube can tolerate depend rather strongly on the gas pressure inside the tube, and hence on the temperature of the tube.

3-8. Gas-filled Thermionic Triodes. The gas-filled triode, or thyratron, is a rectifier that behaves much like the gas-filled diode, but through the action of a third electrode, or grid, it permits simple, efficient control over the current that it conducts.[2] Therefore rectifier circuits using thyratrons can be arranged to give an adjustable output voltage. The thyratron is shown schematically in Fig. 3-21*a*. The anode and cathode are similar to those of the gas-filled diode; the grid is a metal electrode containing one or more holes that is placed between the cathode and anode in the region occupied by the plasma when the tube is conducting. When the tube conducts, electrons flow to the plate and positive ions flow to the cathode through the holes in the grid. The volt-ampere characteristic for the plate-cathode terminals is shown in Fig. 3-21*b*; it is similar to that of the gas diode. The potential distribution in the thyratron when it is conducting is similar to that in the gas diode, shown in Fig. 3-19*c*, and the mechanism of conduction is the same as in the gas diode. The thyratron differs from the gas diode primarily in the action of the grid when the tube is not conducting.

The utility of the grid lies in the fact that the grid-cathode voltage can be used to control the potential that must be applied between the plate and cathode to cause the gas to be ionized and a plasma to be formed. If the grid is made negative relative to the cathode, it lowers the potential-distribution curve in front of the cathode, and virtually no electrons are able to pass the potential minimum and move to the anode, even though the anode be positive relative to the cathode. Under these conditions no current flows in the plate circuit. If the plate is now made more positive, it raises the potential-distribution curve and may make it possible for an appreciable number of electrons to pass the potential minimum in front of the cathode and proceed to the anode. The positive ions generated by these electrons in transit move toward the grid, which is the most negative point in the tube. If positive ions are generated fast enough, the negative charge on the grid is partially shielded by the approaching positive ions, and the potential in front of the cathode rises further. More electron flow to the plate results, more ions are generated, and finally the grid is completely shielded by the great number of positive ions surrounding and moving toward it. The grid then has no further effect on the current, a plasma forms, and the thyratron behaves like a gas diode.

The cumulative action described above that leads to the formation of a plasma is known as ignition, or firing. The minimum plate-to-cathode potential necessary for ignition depends on the grid-to-cathode potential; the more negative the grid the more positive the plate must be for ignition. Figure 3-22 shows the minimum plate voltage required for ignition as a function of the grid potential for a typical thyratron; this curve is the starting, or control, characteristic for the tube. Points lying above this curve correspond to combinations of grid and plate potentials that cause the tube to fire. The time required for the formation of the plasma, usually a few microseconds, is called the ionization time; it limits the speed with which the tube can be turned on.

After the tube fires the grid is surrounded and neutralized by a sheath of positive ions moving toward it. If the grid is made more negative, it merely attracts more ions from the plasma and remains electrically neutralized. Hence after the tube fires the grid has no further effect on the current flowing between plate and cathode; in particular, the plate current cannot be stopped by the application of any reasonable voltage to the grid except in the case of certain small thyratrons conducting currents of a few milliamperes or less. To extinguish the arc and restore the tube to the nonconducting state it is usually necessary to reduce the plate current to a value too low to maintain the plasma. This is the extinction current; it is of the order of microamperes for small thyratrons. If the current remains below this value long enough, usually a few tens or hundreds of microseconds, the positive ions in the tube disappear by surface recombination, the grid regains control, and the tube does not conduct again until the plate voltage exceeds the critical value given by the control characteristic. The time required for the positive ions to disappear is the deionization time for the tube; it places an upper limit on the speed at which the tube can be operated in repetitive cyclic action.

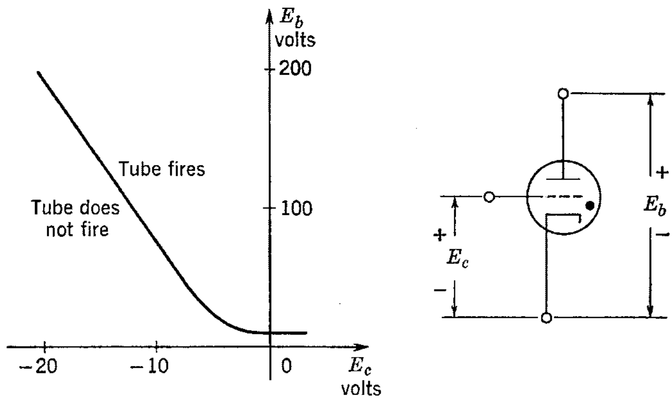


FIG. 3-22. Thyratron control characteristic.

The properties of a thyratron depend on the gas pressure inside the tube. When the gas is mercury vapor obtained by evaporation from liquid mercury in the tube, the gas pressure increases exponentially with the temperature of the coldest spot in the tube; hence the properties of mercury-vapor thyratrons are rather sensitive to the temperature of the tube. In particular, the control characteristic may change appreciably with the changes in temperature encountered in normal operation. This factor must be taken into account in the design of circuits for mercury-vapor thyratrons.

Temperature dependence is often avoided in small, low-power thyratrons through the use of a noble gas such as argon or neon in place of mercury vapor. Since these gases are not derived by evaporation from a liquid source, the density of gas atoms in the tube is independent of temperature, and the tube characteristics are essentially unaffected by changes in temperature. The density of gas atoms in the tube is a function of time, however, for the gas is slowly adsorbed by the electrodes and the inner walls of the envelope. This action is called cleanup. As a consequence of cleanup, the tube characteristics change slowly with time, and eventually the tube is no longer able to perform properly. This fact limits the life of tubes using inert gases. Cleanup is not a factor in mercury-vapor tubes, for additional mercury vapor is always available from the supply of liquid mercury in the tube. Large thyratrons intended for high-voltage, high-power applications almost always use mercury vapor as the gas.

In certain applications, such as in the electronic flash lamps used in high-speed photography and in high-power pulse generators for radar, thyratrons having very short ionization and deionization times are required. This need is met by a group of thyratrons in which hydrogen is the gas used. The relatively light hydrogen ions diffuse through the tube much more rapidly than the heavy mercury ions and give much faster action in the tube.

As in the case of other rectifiers, the maximum instantaneous current and the maximum inverse voltage ratings of the thyratron must be respected. The current and voltage ratings for thyratrons lie in the same ranges as those given for gas-filled diodes.

3-9. The Analysis of Diode Circuits. It is shown in the preceding sections that several physical rectifiers behave very much as ideal diodes. When these physical diodes are used in the circuits presented in Chap. 2, it is often permissible to assume that the physical diodes are ideal. In certain applications, however, especially where vacuum diodes are used, it is necessary to take into account the actual diode characteristics. It is therefore appropriate to examine certain techniques for the detailed analysis of diode circuits. Of even greater importance is the fact that the techniques to be devloped here in terms of relatively simple diode circuits are of fundamental importance in the analysis of vacuum-tube and semiconductor amplifier circuits to be considered later.

Figure 3-23a shows a vacuum diode in the half-wave rectifier circuit of Sec. 2-2. (The cathode heater and its power supply are not shown; in this and all subsequent circuits it is to be understood that all thermionic tubes must have a cathode heater and a source of heating power.) The usable part of the diode volt-ampere characteristic is shown in Fig. 3-23b. The problem is to find the current and the load voltage as functions of time when the value of E_s is given.

The loop equation relating the current to the voltages in the circuit is

$$(R_s + R_L)i_b + e_b = e_s \tag{3-48}$$

In this form it contains two unknowns, e_b and i_b; it can be solved only with the aid of an additional relation between e_b and i_b. The required relation is the volt-ampere law for the diode given by Eq. (3-43); when

(3-43) is substituted in (3-48) the result is

$$(R_s + R_L)i_b + \left(\frac{i_b}{K}\right)^{2/3} = e_s = E_s \sin \omega t \tag{3-49}$$

This equation gives i_b implicitly in terms of the circuit parameters and the applied voltage; however, it is nonlinear and cannot be solved analytically for i_b without considerable difficulty. For this reason graphical methods of solution are almost always used. The graphical solution is based on the measured volt-ampere characteristic of the diode; this characteristic is the counterpart of the three-halves-power law given by Eq. (3-43). The measured characteristic will in general deviate somewhat from (3-43), and it will in general vary considerably for different

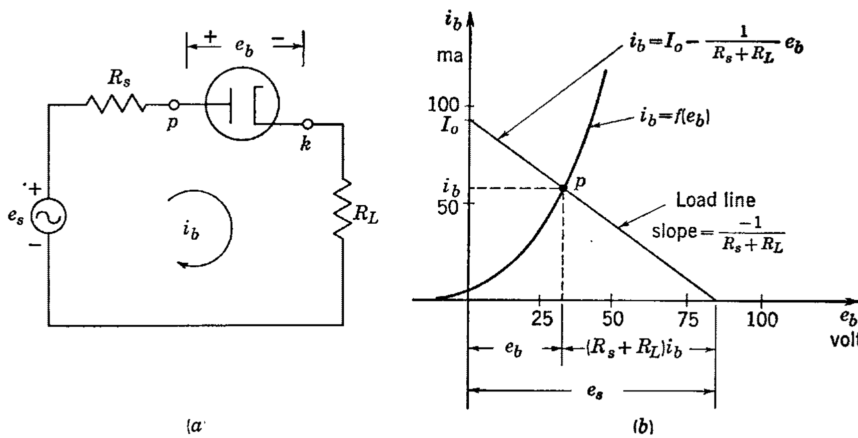


FIG. 3-23. A half-wave rectifier with supply voltage $e_s = E_s \sin \omega_s t$. (a) Circuit; (b) graphical analysis.

tubes of a given type. Therefore the results are approximate unless the characteristic is measured for the particular diode to be used. However, the approximate solution obtained by using the average characteristic for the type of diode employed (obtained from the manufacturer) is satisfactory in practically all cases.

The diode characteristic shown in Fig. 3-23*b* represents a functional relation between i_b and e_b:

$$i_b = f(e_b) \tag{3-50}$$

A second relation between these variables is given by (3-48); this equation can also be written as

$$(R_s + R_L)i_b = e_s - e_b \tag{3-51}$$

and as

$$i_b = \frac{e_s}{R_s + R_L} - \frac{e_b}{R_s + R_L} \tag{3-52}$$

$$= I_o - \frac{1}{R_s + R_L} e_b \tag{3-53}$$

For any fixed value of e_s, I_o is fixed, and (3-53) plots as a straight line on the $e_b - i_b$ coordinates. This line, known as the load line, is shown in Fig. 3-23*b* for one value of e_s. Equation (3-50) is a relation between i_b and e_b imposed by the tube; (3-53) is a relation imposed by the circuit connected to the tube; both relations must be satisfied simultaneously. The diode characteristic gives all combinations of i_b and e_b that satisfy (3-50), and the load line gives all combinations that satisfy (3-53) for one value of e_s. The intersection of these two curves gives the operating point p; both (3-50) and (3-53) are satisfied simultaneously at this point.

As time passes and e_s goes through its cycle of values the load line shifts horizontally in accordance with the value of e_s at each instant, and the operating point p moves along the diode characteristic as the load line shifts. The corresponding values of i_b can be read from the scale

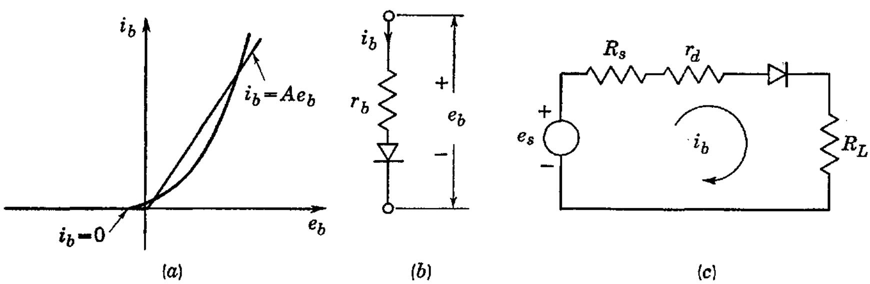


FIG. 3-24. The piecewise-linear approximation for vacuum diodes. (*a*) Diode characteristics; (*b*) model; (*c*) model for rectifier circuit.

of ordinates, and the waveform of i_b can be plotted if desired. The load voltage can be calculated at each instant from the corresponding value of i_b.

The graphical analysis described above is somewhat tedious, and it requires a considerable investment of time and labor, especially when the circuit contains several diodes. As an alternative procedure, the analysis of many diode circuits can be greatly simplified by approximating the diode characteristic with two straight line segments. One such approximation is shown in Fig. 3-24*a* superimposed on the actual diode characteristic. Characteristics consisting of a series of linear segments are called piecewise-linear characteristics; they are used in approximating the characteristics of a wide variety of nonlinear devices. To be sure, an analysis based on this idealized characteristic will not give exactly the results that would be measured experimentally; however, in many applications the results of the idealized analysis are quite close to the experimental results in all important respects. The fact that individual tubes may vary considerably from the published characteristics further justifies the use of idealized characteristics. The value of the idealized characteristics resides in the fact that it greatly simplifies the analysis of many diode circuits and gives results that are quite satisfactory in so far as circuit performance is concerned.

When the behavior of the physical diode is to be represented approximately by the idealized characteristic, it is helpful to have some way of symbolizing the fact in the circuit diagram, and it is desirable to do so by the use of conventional parameters, ideal R, L, C, diodes, and sources. It is readily seen that the volt-ampere characteristic of the circuit in Fig. 3-24*b* corresponds exactly to the idealized characteristic of Fig. 3-24*a* if the resistance r_b has the correct value. For negative values of e_b the ideal diode is an open circuit, and i_b is zero; for positive values of e_b the ideal diode is a short circuit, and the volt-ampere characteristic is a straight line with a slope depending on r_b. If $r_b = 1/A$, the circuit will correspond to the piecewise-linear characteristic of Fig. 3-24*a*. The circuit of Fig. 3-24*b* is a network model for the physical diode that behaves approximately like the physical diode. More specifically, the circuit is a piecewise-linear model for the physical diode. The nonlinear nature of the circuit is concentrated at the break point in the characteristic. The fact that the segments of the characteristic are linear simplifies the analysis problem.

If the piecewise-linear model is substituted for the diode in the half-wave rectifier of Fig. 3-23, the circuit of Fig. 3-24*c* results. It is clear that this circuit can be analyzed by the simple procedure used in Sec. 2-2. Thus the half-wave rectifier using a vacuum diode behaves essentially as if the diode were ideal; the principal difference is that the circuit resistance is increased by the diode resistance r_b. In a like manner, when vacuum diodes are used in any of the circuits presented in Chap. 2, the behavior can be analyzed approximately by using piecewise-linear models for the diodes. If the diode resistances are small, the results of the analysis will be essentially the same as those obtained for ideal diodes.

Figure 3-25 shows volt-ampere characteristics and piecewise-linear approximations for two typical vacuum diodes. The 6AX5-GT is intended primarily for use in power supplies for electronic equipment. The piecewise-linear model corresponding to the piecewise-linear characteristic shown in the figure consists of an ideal diode in series with a resistance of 400 ohms. The 6AL5 is a small tube intended for low-power applications such as AM detectors, diode clampers, and rectifiers to supply small direct currents. The piecewise-linear model implied by the approximating straight-line characteristic consists of an ideal diode in series with a resistance of 200 ohms. Figure 3-26 shows the piecewise-linear model for the 6AL5, including the interelectrode capacitance.

All the preceding discussion regarding the analysis of circuits containing vacuum diodes applies in a qualitative sense to circuits containing semiconductor diodes. If a semiconductor diode is used in the half-wave rectifier of Fig. 3-23, the circuit can be analyzed by the same graphical procedure as that used with the vacuum diode. Figure 3-27*a* shows the

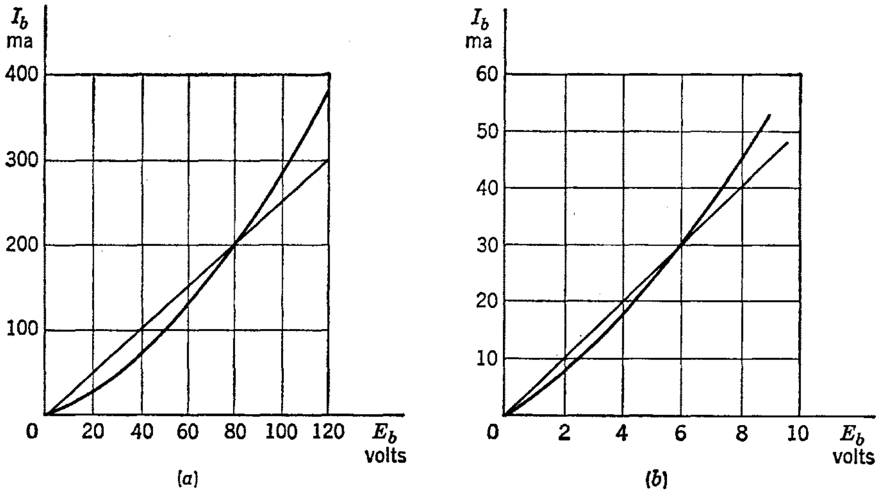


FIG. 3-25. Typical diode characteristics. (*a*) 6AX5-GT twin diode, peak permissible inverse voltage 1250 volts, maximum permissible instantaneous plate current 375 ma; (*b*) 6AL5 twin diode, peak permissible inverse voltage 330 volts, maximum permissible instantaneous plate current 54 ma, plate-to-cathode capacitance, 3 $\mu\mu$f (approximately).

volt-ampere characteristic for a typical silicon *P-N* junction diode and the graphical construction appropriate to the analysis of the half-wave rectifier. The scale of voltages would be suitable if the peak value of e_s were less than about 10 volts. Figure 3-27*b* shows the same construction for the case where the peak value of e_s is of the order of 100 volts. The scale of voltage and current required is such that the diode characteristic cannot be distinguished from the axes; hence the performance of the diode cannot be distinguished from that of an ideal diode in so far as the load voltage and current are concerned.

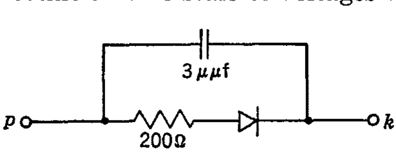


FIG. 3-26. Piecewise-linear model for the 6AL5.

In some circuits such as AM detectors and diode clampers, where small currents and voltages are involved, it may not be desirable to treat the semiconductor diode as ideal. On the other hand, the piecewise-linear representation is often suitable. One approximation to the diode characteristic is shown in Fig. 3-28*a*. The corresponding model is shown in Fig. 3-28*b*. When e_b is less than E_o, a reverse voltage is applied to the ideal diode and it acts as an open circuit. When e_b is greater than E_o, a forward current flows through the ideal diode; its magnitude is $i_b = (e_b - E_o)/r_b$. If r_b and E_o are given suitable values, the model will correspond to the characteristic in Fig. 3-28*a*. Figure 3-28*c* represents the diode approximately at high frequencies; the capacitance C represents the capacitance across the junction of the diode.

It is worthwhile to note that adding the source E_o to the model has the effect of shifting the piecewise-linear characteristic to the right by E_o

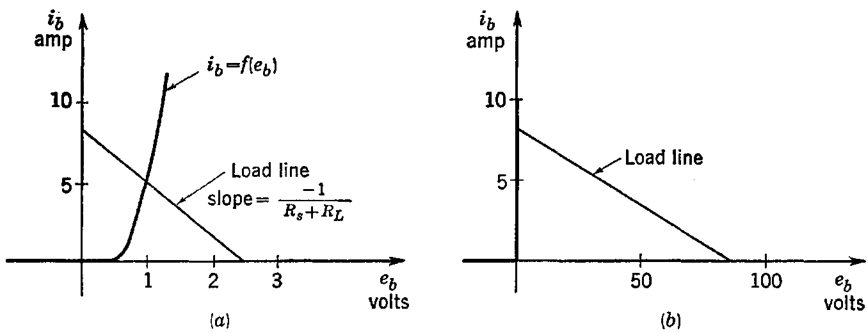


FIG. 3-27. Graphical analysis of a silicon-diode rectifier circuit. (*a*) Low voltage; (*b*) high voltage.

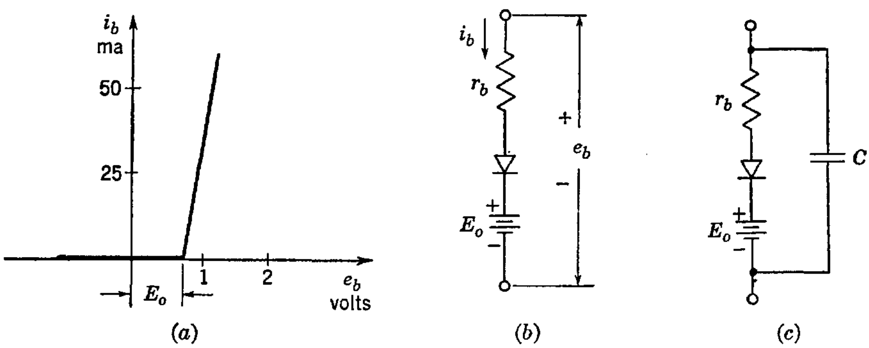


FIG. 3-28. Piecewise-linear models for semiconductor diodes. (*a*) Piecewise-linear characteristic; (*b*) low-frequency model; (*c*) high-frequency model.

volts when the source has the polarity shown in Fig. 3-28*b*. If the polarity of the source is reversed, the characteristic is shifted to the left by E_o volts. These facts indicate that a variety of piecewise-linear characteristics can be generated by various combinations of sources and ideal diodes. For example, volt-ampere characteristics having several linear segments can be generated by using several diodes and associated voltage sources. This feature is useful in the formulation of models for more complicated devices. It is also interesting to note that electronic analog computers often use combinations of physical diodes, resistors, and batteries to generate nonlinear relations between two variables.

The volt-ampere characteristic for a gas-filled diode shown in Fig. 3-29*a* is approximated satisfactorily by the piecewise-linear characteristic of Fig. 3-29*b*. It is clear that the piecewise-linear model of Fig. 3-28*b* corresponds to the characteristic of Fig. 3-29*b* if r_b is made zero. (In some instances it may be desirable to give r_b a small negative value to account for the slight decrease in tube drop with increasing tube current. The apparent negative resistance associated with conduction through gases finds useful applications.)

The thyratron can also be represented by a piecewise-linear model, but the complications needed to account properly for the action of the grid make the model unwieldy. Therefore it is usually simpler to represent the plate circuit of the thyratron by two separate models, an open circuit when the tube is nonconducting and a constant voltage drop (battery) when it is conducting.

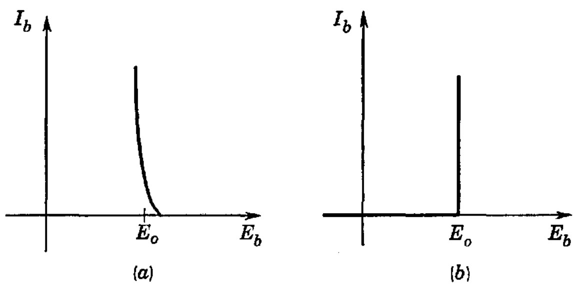


FIG. 3-29. Piecewise-linear approximation for gas-filled diodes. (*a*) Volt-ampere characteristic; (*b*) piecewise-linear approximate characteristic.

The use of grid control to provide an adjustable load voltage in a thyratron rectifier is illustrated by the circuit shown in Fig. 3-30*a*.[2] The a-c supply voltage is e_s, and the control voltage for the grid is the series combination of e_1 and e_2. The voltage e_1 is an adjustable direct voltage by means of which the load voltage is adjusted, and e_2 is a fixed sinusoidal voltage that lags 90° behind e_s. The relations among these voltages are pictured in Fig. 3-30*b*. The load resistance is R_L; as in the case of the gas-filled diode, it must be large enough to limit the plate current to a safe value. In addition, when the grid is positive relative to the cathode by an amount greater than the conducting tube drop, the grid and cathode act as a conducting diode. Therefore it is necessary to include a resistance in series with the grid to limit the grid current to a safe value; this resistance is R_c in Fig. 3-30*a*.

Two simplifying assumptions are made to remove distracting second-order effects from the analysis. They are (1) that the negative grid potential needed to hold the tube nonconducting is negligibly small, and (2) that the conducting tube drop is negligibly small. Thus if the tube is nonconducting with a positive plate potential and a negative grid potential, as during the first portion of the cycle of operation shown in Fig. 3-30*b*, then it cannot fire (begin conduction) as long as the grid potential is negative. With the grid negative relative to the cathode, electrons emitted by the cathode cannot go to the grid, and since there are no positive ions in a nonconducting tube, there is no ion flow to the grid. Hence the grid current is zero, there is no voltage drop across R_c, and the grid voltage is $e_1 + e_2$. At the instant t_f the grid potential, $e_1 + e_2$, changes from negative to positive, the grid is no longer able to hold the tube nonconducting, and the tube fires. Since the conducting tube drop is assumed to be zero, the tube acts as a short circuit and the full supply voltage appears across the load during conduction. Since ordinarily the grid cannot stop the flow of plate current, conduction continues until the plate potential becomes negative, regardless of the

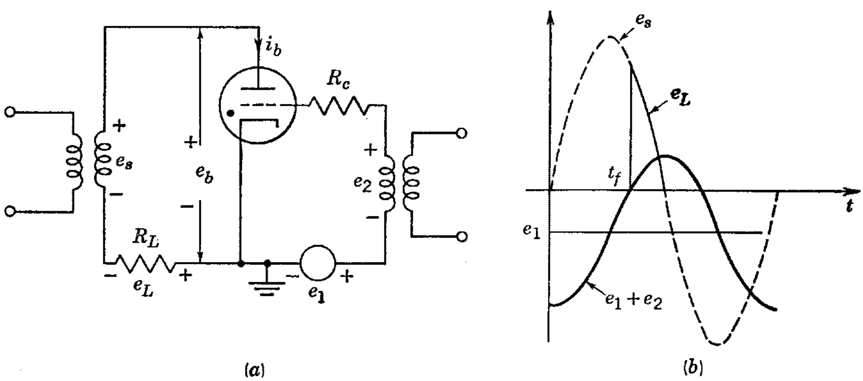


FIG. 3-30. A half-wave thyratron rectifier with d-c plus a-c control. (*a*) Circuit, $e_s = E_s \sin \omega_s t$ and $e_2 = E_2 \sin (\omega_s t - 90°)$; (*b*) waveforms.

value of the grid potential. The corresponding waveform of load voltage is shown in Fig. 3-30*b*.

The duration of the conducting interval, and hence the average and effective values of the load voltage, can be controlled by the adjustable direct voltage e_1. If e_1 is adjusted over a range of positive and negative values, the waveform of $e_1 + e_2$ shifts up or down, and the instant of firing can be made to occur at any point in the positive half cycle of e_s. For example, if e_1 is made zero, the tube fires when e_s has its positive peak value; if e_1 is made equal to E_2, the tube fires at the beginning of the cycle. The tube deionizes and the grid regains control during each negative half cycle of e_s. The average value, or d-c component, of the load voltage is

$$E_{dc} = \frac{1}{T} \int_{t_f}^{T/2} (E_s \sin \omega_s t)\, dt \qquad (3\text{-}54)$$

where T is the period of the supply voltage.

A thyratron rectifier circuit of more practical importance is shown in Fig. 3-31; it is the full-wave counterpart of the half-wave rectifier circuit in Fig. 3-30. The a-c component of the control voltage e_2 lags 90° behind the supply voltage e_s, and the d-c component of the control voltage e_1 is adjustable. The primary of the transformer supplying e_2 is not shown in Fig. 3-31. The firing of each thyratron is controlled separately in the manner described in connection with Fig. 3-30; the two tubes conduct on alternate half cycles.

The load indicated in Fig. 3-31 is the armature of a d-c motor. The voltage applied to the armature, and hence the speed of the motor, can be controlled by controlling e_1. If a magneto tachometer is connected to the shaft of the motor as indicated in Fig. 3-31, it gives an output

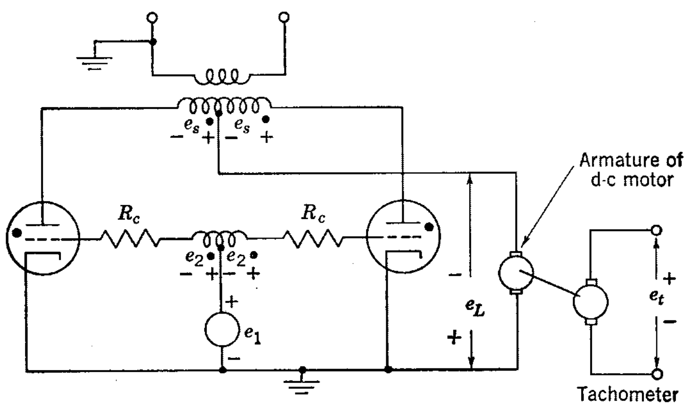


FIG. 3-31. A full-wave thyratron rectifier.

voltage e_t that is directly proportional to the speed of the motor. With the aid of some additional circuitry this tachometer voltage can be used to provide a control voltage e_1 that will automatically adjust the armature voltage so as to keep the motor speed constant in the face of wide variations in load on the motor. Such systems have been used for the control of motors powering machine lathes and for other adjustable-speed applications.

3-10. Summary. Semiconductor, vacuum, and gas-filled rectifiers are satisfactory approximations to the ideal diode, at least over a certain range of voltages and currents. Hence they can be used in the physical realization of the circuits presented in Chap. 2. In the process of examining these physical diodes, certain concepts of fundamental importance emerge. Now that the over-all picture is completed, it is desirable to restate these basic ideas in a brief form.

Current flow in practical rectifiers can be studied in terms of the potential distribution inside the devices. The amount of current flowing in each case depends on the number of carriers that have enough kinetic energy to surmount a potential energy barrier that exists inside the device. When the carriers are in random motion under conditions of thermal equilibrium, the energies of the carriers can be determined from probability theory. From this study the current crossing the potential energy barrier is found to be

$$I = I_o \exp \frac{-qE}{kT}$$

where q = charge on carrier
E = height of barrier, ev
k = Boltzmann constant
T = absolute temperature
I_o = current that would flow if E were zero

In the semiconductor diode, E is a direct function of the applied voltage; hence the volt-ampere law is an exponential function. Under normal operating conditions in the vacuum diode, the value of E depends in a complicated way on the space charge and the applied voltage. It can be shown that the current through the vacuum diode increases approximately as the applied voltage raised to the three-halves power. The conduction process is even more complicated in gas-filled rectifiers.

It follows from the foregoing statements that the volt-ampere laws for practical rectifiers are nonlinear relations between the voltage and current. Indeed, the rectifying property of the diode depends entirely on this nonlinearity; a linear device cannot rectify. However, as a result of this nonlinearity the analysis of diode circuits is complicated, and exact analytical solutions are not often attempted.

Circuits that are not too complicated can be analyzed by simple graphical methods. Alternatively, the analysis can often be simplified by representing the diode approximately by a suitable piecewise-linear model. The approximation of nonlinear devices by piecewise-linear models is a general technique that has been applied to a wide variety of devices in addition to diode rectifiers. The technique does not provide a simple solution to all nonlinear problems, however. If the circuit is very complicated, or if it contains several reactive elements, the piecewise-linear model may be of little help.

REFERENCES

1. Frank, N. H.: "Introduction to Electricity and Optics," 2d ed., McGraw-Hill Book Company, Inc., New York, 1950.
2. Gray, T. S.: "Applied Electronics," 2d ed., John Wiley & Sons, Inc., New York, 1954.
3. Shockley, W.: Transistor Electronics, *Proc. IRE*, vol. 40, no. 11, p. 1289, November, 1952.

4. Herold, E. W.: New Advances in the Junction Transistor, in "Transistors I," p. 27, RCA Laboratories, Princeton, N.J., 1956.
5. Seely, S.: "Radio Electronics," McGraw-Hill Book Company, Inc., New York, 1956.

PROBLEMS

3-1. A certain vacuum diode has parallel-plane electrodes that are large compared with the spacing between them. Hence when there is a negligible charge in the space between the electrodes, the field in the interelectrode space is essentially uniform and perpendicular to the electrodes. The potential-distribution curve for a straight path between cathode and plate is thus a straight line. The total potential difference across the interelectrode space is 100 volts, with the anode positive relative to the cathode.

a. If a single electron starts from rest at the cathode, what is its velocity when it reaches the anode? Give the answer in millions of feet per second. The mass of the electron is 9.11×10^{-31} kg, and the magnitude of its charge is 1.6×10^{-19} coulomb.

b. If the electrodes are separated by 1 cm, how long is the electron in transit? Note that the velocity of the electron is not constant.

3-2. The work functions of the cathode and anode of the diode of Prob. 3-1 are 1 and 2 ev, respectively.

a. What is the contact difference of potential between the electrodes?

b. An electron is emitted from the cathode with 1.5 ev of kinetic energy directed toward the anode. What potential must be applied between the electrodes from an external battery to prevent this electron from reaching the anode?

3-3. Electrons emitted from the cathode of the diode shown in Fig. 3-18*a* move toward the anode along the path s shown in the diagram. The potential distribution along the path is given by curve B in Fig. 3-18*c*, and the value of the potential at the minimum is -1 volt.

a. What is the minimum initial velocity directed along the path s that an electron issuing from the cathode must have if it is to reach the anode?

b The number of electrons leaving the cathode each second with kinetic energies greater than E ev directed along path s is given by $n = 10^{18}\epsilon^{-5E}$. What current flows through the tube?

3-4. Circuits similar to that shown in Fig. 3-32 have been used as multipliers in certain elementary types of electronic analog computers. The three identical silicon diodes are operated so that their volt-ampere relations are given by Eq. (3-31). The input voltages to the amplifier are e_1 and e_2; the input currents are zero. The output voltage from the amplifier is $e_o = e_1 + e_2$. Show that the output current i_3 is proportional to the product of the diode currents i_1 and i_2.

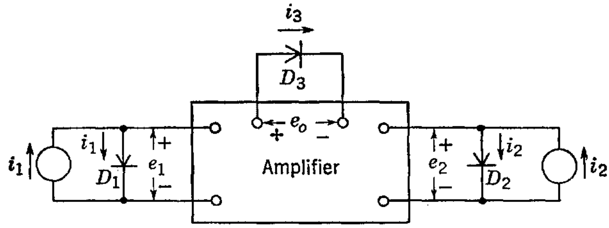


FIG. 3-32. An electronic analog multiplier for Prob. 3-4.

3-5. One diode section of a 6AX5 twin diode is used in a half-wave reactifier like that shown in Fig. 3-23. The supply voltage is $e_s = 120 \sin 377t$ volts, and the circuit resistances are $R_s = 0$ and $R_L = 200$ ohms.

a. Plot the diode characteristic on a suitable sheet of graph paper (Fig. 3-25). Determine graphically the value of i_b for various values of e_s, and plot one full cycle of the waveform of i_b.

b. Replace the tube by a piecewise-linear model consisting of a 400-ohm resistor in series with an ideal diode. Plot one cycle of i_b for this model on the same coordinates as those used for the waveform of part *a*.

c. Would a different value of series resistance in the model give closer agreement between the results of parts *a* and *b*? If so, what value of r_b would you recommend?

3-6. A power supply to deliver 300 volts d-c to a load drawing 100 ma is required. A 6AX5 twin diode is to be used in a full-wave rectifier with a capacitor-input filter (Fig. 2-24). The power source is a 115-volt 60-cps a-c line.

a. If the input capacitance is 10 μf, how much voltage must the transformer supply on each side of the center tap? Refer to the tube manual for the 6AX5 rectifier characteristics, and assume that the filter choke will have a resistance of 100 ohms.

b. In an alternative design, the same tube is to be used in a choke input filter (Fig. 2-21) with $L = 10$ henrys and $C = 16$ μf. Repeat the calculation of part *a* and compare.

c. At what frequency (in cycles per second) is $\omega^2 LC = 1$ for the filter of part *b*? Is this a suitable design? Explain.

3-7. Two silicon diodes are used in a full-wave rectifier to charge 6.3-volt automobile batteries. The diode characteristics are given in Fig. 3-27*a*. Power is obtained from a transformer delivering $e_s = 27 \sin 377t$ volts on each side of the center tap, and a resistor is connected in series with the battery to limit the charging current.

a. Give a circuit diagram for this battery charger showing the polarities of all voltages.

b. The diodes are to be represented by piecewise-linear models like that shown in Fig. 3-28*b*. From the diode characteristics determine suitable values for r_b and E_o.

c. What value must the series resistance have to limit the peak diode current to 20 amp? Would a serious error be made in this calculation by assuming the diodes to be ideal?

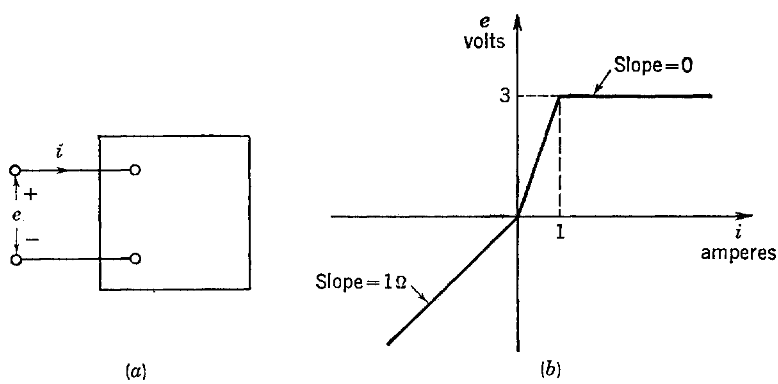


FIG. 3-33. A piecewise-linear device for Prob. 3-8. (*a*) Device; (*b*) volt-ampere characteristic.

3-8. The measured volt-ampere characteristic of an unknown device, represented symbolically in Fig. 3-33*a*, can be approximated closely by the piecewise-linear characteristic shown in Fig. 3-33*b*.

a. Find a piecewise-linear model for the unknown device that has the given piecewise-linear volt-ampere characteristic. The model may contain resistors, ideal diodes, batteries, and d-c current sources.

b. Repeat part *a* when i is replaced by e and e is replaced by i on the volt-ampere characteristic.

3-9. Two gas-filled diodes are used in a full-wave rectifier to charge a 12.6-volt storage battery. Power is obtained from a transformer that delivers $e_s = 35 \sin 377t$ volts on each side of the center tap. The voltage drop across each diode when it is conducting is approximately 10 volts for all values of current. The charger is connected to the battery by a long pair of leads having a resistance of 2 ohms.

a. Give a circuit diagram for this battery charger showing the polarities of all voltages

b. Give a piecewise-linear model for the circuit showing the numerical values of all circuit parameters.

c. What is the peak instantaneous current through each diode?

d. During what fraction of each half cycle does current flow into the battery?

e. If the leads connecting the charges to the battery are shortened to one-quarter of their original length, what is the peak instantaneous current through each diode? The leads are of uniform cross section throughout their length.

3-10. The supply voltage for the thyratron rectifier of Fig. 3-30 is $e_s = 300 \sin \omega_s t$ volts, the supply frequency is $f_s = 400$ cps, and the a-c component of grid voltage is $e_2 = -50 \cos \omega_s t$ volts. For all positive values of plate-to-cathode voltage the tube fires when the grid becomes positive relative to the cathode; that is, the critical grid voltage is zero for all positive values of plate voltage. The tube drop is negligible when the tube is conducting.

a. What value of control voltage e_1 causes the tube to fire when $\omega_s t = 45°$?

b. What is the value of the d-c component of load voltage under the conditions of part *a*?

c. What value of e_1 just reduces the load voltage to zero?

3-11. The load on a full-wave thyratron rectifier like the one shown in Fig 3-31 is the resistance heating element of an oven in which the temperature is controlled automatically. The control voltage e_1 is obtained from a thermocouple and an amplifier in such a way that if the temperature in the oven rises, e_1 decreases, and the rms value of the voltage applied to the heating element decreases. If the temperature in the oven drops, the opposite action takes place. In this way the temperature in the oven can be held nearly constant, or it can be caused to vary in some desired way.

The supply voltage for the rectifier is $e_s = 200 \sin \omega_s t$ volts, and the a-c component of the grid voltage is $e_2 = -50 \cos \omega_s t$ volts. For all positive values of plate voltage, the tube fires when the grid voltage becomes positive; that is, the critical grid voltage is zero for all positive values of plate voltage. The voltage drop across the tube is negligible when the tube is conducting.

a. What value of e_1 will cause each tube to conduct for one-quarter of a cycle of the supply voltage?

b. What is the rms value of the load voltage under the condition of part *a*?

c. An increase in oven temperature of one degree causes e_1 to decrease by 10 volts. If the temperature increases by one degree from the value corresponding to the conditions of part *a*, what is the new rms value of voltage applied to the load?

CHAPTER 4

IDEAL AMPLIFIERS

In Chap. 2 a new circuit component, the ideal diode, is introduced, and it is shown that circuits containing ideal diodes can perform certain useful functions that are not possible when only ideal R, L, and C are employed. The principal objective of this present chapter is to introduce another circuit component, the ideal amplifier, and to examine some of its useful functions. In this connection certain fundamental ideas related to amplification are presented.

The study of amplifiers is initiated with an examination of the properties of ideal amplifiers, for in this way attention can be focused on the process of amplifications without the distractions imposed by physical reality. By concentrating first on the process of amplification, one is conditioned to accept any kind of physical device that may be proposed as an amplifier. Of equal importance, having established the properties of the ideal amplifier, a basis exists for perceiving and evaluating the good and bad features of any physical amplifier and thus, perhaps, for improving the physical amplifier.

The need for power amplification in the field of electric communication is obvious. A telephone instrument delivers a few milliwatts of power to a pair of wires, but only a fraction of this power reaches the receiving end because of losses in the line. For telephone communication over long distances it is necessary to amplify the signal power at more or less regular intervals along the line. In some transcontinental communication circuits the signal power is amplified billions of billions of times in transit from the sending end to the receiving end. A radio transmitter may radiate many kilowatts of power into space, but only a microscopic amount of this power is intercepted by any particular radio receiver. This power must be amplified greatly to give a suitable volume of sound from a loudspeaker. The need for power amplification is equally great in many other fields, as, for example, in electrical instrumentation and in the automatic control machinery.

4-1. The Ideal Voltage Amplifier. One form of ideal electric amplifier is shown enclosed in dotted lines in Fig. 4-1a; it consists simply of a voltage source whose output voltage is at every instant directly proportional to its input voltage. For each value of e_1, the output voltage has the value μe_1 and is independent of the output current. The output volt-ampere characteristic is shown in Fig. 4-1*b*; it consists of a family of vertical straight lines, one for each value of the input voltage. If the values of e_1 corresponding to the curves are separated by 1 volt, then the curves are separated by μ volts on the output-voltage axis. Since this ideal amplifier consists of a source whose output voltage is controlled by the input voltage to the amplifier, it is called a controlled source, or a voltage-controlled source.

If the input voltage e_1 is specified, then the output voltage $e_2 = \mu e_1$ is fixed and is independent of the load current. Hence any amount of power can be drawn from the output terminals by the proper choice of R_L (assuming, of course, that $e_1 \neq 0$). The input current to the amplifier

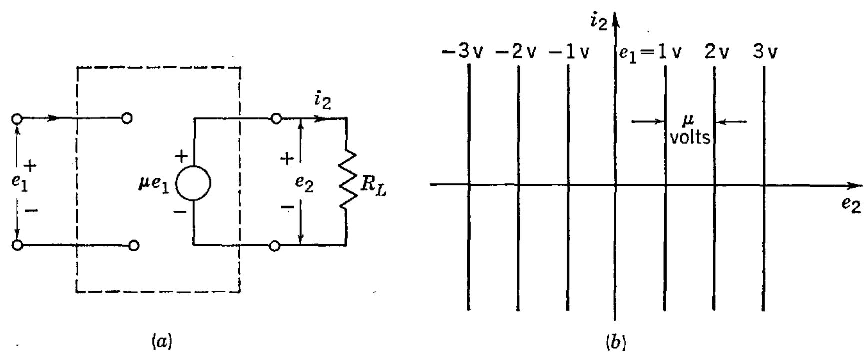


FIG. 4-1. The ideal voltage amplifier. (*a*) Circuit diagram; (*b*) output volt-ampere characteristic.

is zero regardless of the value of e_1; hence the input power is zero under all conditions. From these facts it follows that the power gain, which is the ratio of the output power to the input power, is infinite. The output power from the amplifier at any instant is

$$p = \frac{(\mu e_1)^2}{R_L} \tag{4-1}$$

Equation (4-1) shows that with a fixed load resistance the output power depends on the input voltage e_1; the greater e_1, the greater the output power. With a given value of amplification factor μ, the small voltage received from a telephone line (for example) may not be great enough to give the required power output. This difficulty can be overcome by amplifying the signal voltage before applying it to the power-output amplifier. It is clear that the ideal amplifier of Fig. 4-1*a* can be used as a voltage amplifier provided μ is greater than unity. Figure 4-2 shows a cascade of three amplifier stages for use in this situation. The input and the output power in the first two stages are zero, but the output voltage from each is greater than the input voltage; hence these stages are voltage amplifiers. The third stage gives the required power output; hence it is a power amplifier. The third stage may or may not give voltage amplification also, depending on the value of μ_3. The output voltage is

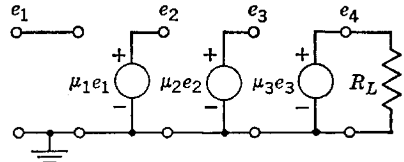


FIG. 4-2. Cascaded voltage amplifiers.

$$e_4 = \mu_1 \mu_2 \mu_3 e_1 \tag{4-2}$$

and the output power is

$$p = \frac{(\mu_1 \mu_2 \mu_3 e_1)^2}{R_L} \tag{4-3}$$

Vacuum-tube amplifiers are conveniently characterized as voltage-controlled amplifiers like those described above, although they fall somewhat

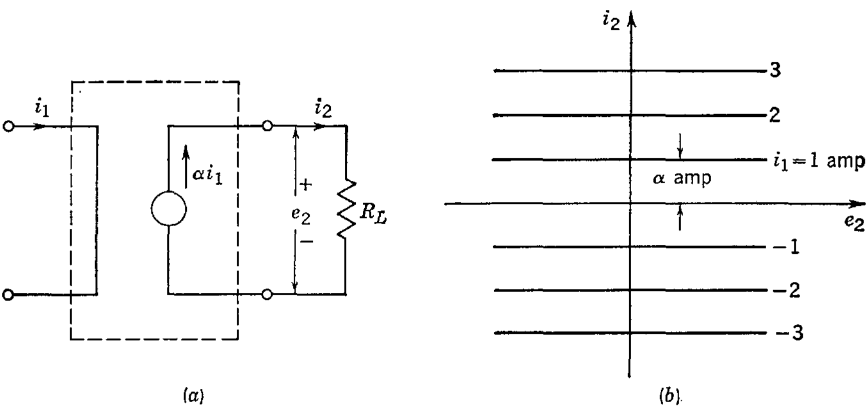


FIG. 4-3. The ideal current amplifier. (*a*) Circuit diagram; (*b*) output volt-ampere characteristic.

short of being ideal. A typical vacuum-tube radio may employ a cascade of five stages similar to those in Fig. 4-2. The first four of these are voltage amplifiers and are designed to give the greatest possible voltage amplification. The fifth stage is a power amplifier and is designed to give the greatest possible power output.

4-2. The Ideal Current Amplifier. Another form of ideal electric amplifier is shown in Fig. 4-3*a*. This amplifier is a current-controlled current source; its output volt-ampere characteristic is shown in Fig. 4-3*b*. For each value of input current i_1 the output current has the value αi_1

and is independent of the value of the load voltage. The ideal current amplifier is thus the dual of the ideal voltage amplifier.

Since the output current is independent of the load voltage, any amount of power can be drawn from the output terminals by the proper choice of R_L, provided $i_1 \neq 0$. Moreover, the input voltage is zero regardless of the value of the input current; hence the input power is zero under all conditions. It follows from these facts that the power gain of the ideal current amplifier is infinite. The power output from the amplifier at any instant is

$$p = (\alpha i_1)^2 R_L \tag{4-4}$$

For a given value of α the input signal current may not be great enough to give the required power in a specified load resistance R_L; thus the need for current amplification arises. The amplifier of Fig. 4-3*a* can be used for current amplification provided α is greater than unity. A cascade of three current-controlled amplifier stages is shown in Fig. 4-4. The first two stages are current amplifiers; the input and output powers in these stages are zero. The third stage provides the required power output; it is a power amplifier. This cascade of current amplifiers is the dual of the cascade of voltage amplifiers shown in Fig. 4-2.

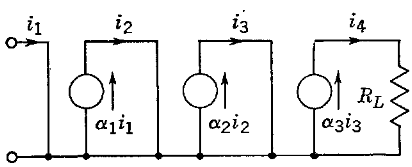


FIG. 4-4. Cascaded current amplifiers.

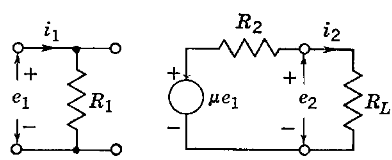


FIG. 4-5. A nonideal amplifier.

Transistor amplifiers are conveniently characterized as current-controlled amplifiers, although they fall somewhat short of being ideal. A typical transistor radio may employ a cascade of five amplifier stages. The first four of these are current amplifiers and are designed to give the greatest possible current amplification. The fifth stage is a power amplifier and is designed to give the greatest possible power output.

4-3. Gain and Amplification in Decibels. Figure 4-5 shows a voltage amplifier that is not ideal; its input resistance is R_1, not infinity, and its output resistance is R_2, not zero. The power input to this amplifier at any instant is $p_1 = e_1^2/R_1$, and the output power at any instant is $p_L = e_2^2/R_L$. The power gain in this case, p_L/p_1, is finite, and it serves as a measure of the effectiveness of the amplifier. Another measure that proves to be more useful in many respects is the power gain in decibels (db), defined by

$$G = 10 \log \frac{p_L}{p_1} \tag{4-5}$$

where *log* is understood to designate the common logarithm. Thus if p_L is 500 watts and p_1 is 5 watts, the power gain is 20 db. The power gain of the circuit in Fig. 4-5 can also be written as

$$G = 10 \log \frac{e_2{}^2 R_1}{R_L e_1{}^2} = 20 \log \frac{e_2}{e_1} + 10 \log \frac{R_1}{R_L} \tag{4-6}$$

If $R_1 = R_L$, as often is the case in communication circuits, $\log R_1/R_L = 0$; in this case, and *only* in this case, the power gain of the amplifier becomes

$$G = 20 \log \frac{e_2}{e_1} \qquad \text{when } R_1 = R_L \tag{4-7}$$

The logarithm of the voltage amplification proves to be useful in several respects, regardless of whether $R_1 = R_L$ or not. For this purpose Eq. (4-7) is appropriated, and the voltage amplification in db is defined as

$$A = 20 \log \frac{e_2}{e_1} \tag{4-8}$$

Thus if $R_1 = R_L$, the voltage amplification in decibels is equal to the power gain in decibels. These two quantities are not equal, however, when $R_1 \neq R_L$; Eq. (4-6) shows that in general the power gain in decibels and the voltage amplification in decibels differ by the amount 10 log R_1/R_L. In order to emphasize the distinction between the quantities given by (4-6) and (4-8), several distinguishing names have been proposed for the logarithmic unit of amplification; among these are *decibels of voltage amplification* (dbv), and *decilog*, although the multiplier of 20 in (4-8) renders the prefix *deci-* somewhat incongruous. Nevertheless, common practice is to use *decibel* as the name for the logarithmic unit of amplification as well as gain. The logarithmic unit of current amplification is defined in a manner similar to that of voltage amplification.

The standard definitions designate the ratio of two powers, or its logarithm, as power gain, and they designate the ratio of two voltages or two currents, or their logarithms, as voltage or current amplification. However, standards to the contrary notwithstanding, it is common practice to use the term *gain* to designate current, voltage, and power ratios and their logarithms. The standard definitions will be followed in this book except in a few special cases where it is desirable to conform to well-established usage. Since according to the standard definitions gain and amplification refer to both a numerical ratio and the logarithm of that ratio, an ambiguity exists unless a clarifying statement is appended. In this book these terms will refer to the numerical ratio unless decibels are specified.

When the output power and voltage of an amplifier are greater than the input power and voltage, the gain and the voltage amplification in decibels are positive numbers. When the output power and voltage are smaller than the input power and voltage, the gain and the voltage amplification in decibels are negative numbers. When the output power and voltage are equal to the input power and voltage, the gain and the voltage amplification in decibels are zero. It is, of course, possible for the gain in decibels to be positive while the amplification in decibels is negative, and vice versa, if R_1 and R_L are not equal.

The decibel is a convenient measure of voltage, current, and power ratios for a number of reasons. One of the most important of these is the fact that the frequency characteristics of electric networks, both electronic and otherwise, have especially simple forms when the response in decibels is plotted against frequency on a logarithmic scale. In addition, the over-all response of a cascade of stages such as those discussed in the preceding section is obtained by a simple addition of the

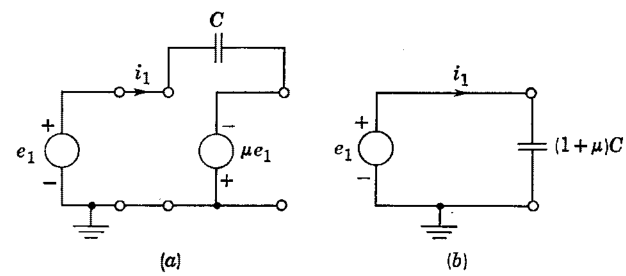


FIG. 4-6. An application of the ideal amplifier. (*a*) Circuit; (*b*) an equivalent circuit.

responses of the individual stages when the response is expressed in decibels.

4-4. Other Applications for Voltage and Current Amplifiers. Ideal amplifiers are capable of a number of interesting and useful operations that seem at first glance to bear little relation to signal amplification. Consider, for example, the combination of capacitor and voltage-controlled voltage source shown in Fig. 4-6*a*. The input current to this circuit is

$$i_1 = C\frac{d}{dt}(e_1 + \mu e_1) = (1 + \mu)C\frac{de_1}{dt} \tag{4-9}$$

Thus the input current is the same as the current that flows into a capacitor of capacitance $(1 + \mu)\ C$ having the voltage e_1 impressed across its terminals; that is, the circuit of Fig. 4-6*b* is equivalent to that of Fig. 4-6*a* in so far as the input terminals are concerned. Circuits of this type are used to obtain the effect of very large capacitors in certain engineering applications.

As another example, consider the combination of resistor and current-controlled current source shown in Fig. 4-7*a*. The current through the resistor is $(1 - \alpha)i_1$, and the voltage drop across the resistor is

$$e_1 = (1 - \alpha)i_1R$$

Since this voltage is also the voltage across the input terminals, the apparent resistance at the input terminals is $(1 - \alpha)R$; that is, the circuit of Fig. 4-7*b* is equivalent to that of Fig. 4-7*a* in so far as the input terminals are concerned. Thus the current source acts as a resistance converter, for it converts a given resistance to a different value. In particular, if α is greater than unity, a positive resistance is converted to a negative resistance. If R is replaced by a complex impedance, the impedance can be converted to a negative impedance. In this case the current-controlled current source is called a negative impedance converter. Negative impedance converters realized with the aid of transistors are important in the design of certain types of electrical filters.

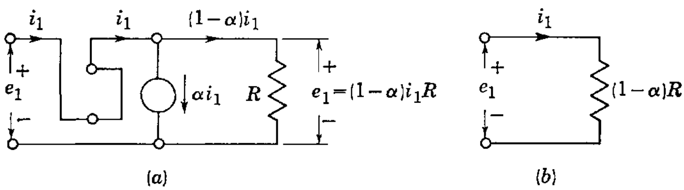


FIG. 4-7. An impedance converter. (*a*) Circuit; (*b*) an equivalent circuit.

It should be noted that the circuit of Fig. 4-6 can also be viewed as an impedance converter.

4-5. Summary. The fundamental notions of voltage, current, and power amplification are presented in the preceding sections, and the concepts of ideal voltage and current amplifiers are established. The basic features of these latter concepts can be summarized as follows.

The ideal voltage amplifier is a voltage-controlled voltage source. Its input impedance is infinite, and its output impedance is zero. Since infinite input impedance permits the greatest possible input voltage to be developed, and since zero output impedance permits the greatest possible output voltage to be developed across any given load, the ideal voltage amplifier provides the greatest possible voltage amplification (for a given amplification factor, μ). The input voltage is independent of the load connected to the output terminals; that is, the input circuit is isolated from the output circuit. Signals can be transmitted through an ideal voltage amplifier in one direction only, from the input to the output.

The ideal current amplifier is the dual of the ideal voltage amplifier. It is a current-controlled current source having zero input impedance and infinite output impedance; these conditions permit the greatest possible current amplification for a given amplification factor, α. As in the case of the voltage amplifier, the input circuit is isolated from the output circuit, and signals can be transmitted in one direction only.

There are many physical devices that are capable of amplifying signals; of special importance among these are the vacuum tube and the transistor. To a certain extent the merits of a physical amplifier can be evaluated by comparing its properties with those of either the ideal voltage amplifier or the ideal current amplifier. Such a comparison may also indicate how the physical device should be modified in order to obtain improved performance.

PROBLEMS

4-1. A three-stage amplifier is shown in Fig. 4-8. The symbol K designates kilohms, and the symbol M designates megohms.

a. Find the voltage amplification and the power gain of each stage in decibels.

b. Find the over-all voltage amplification and power gain in decibels.

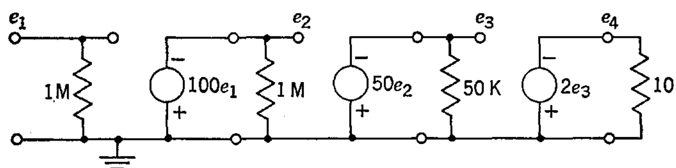


FIG. 4-8. Amplifier for Prob. 4-1.

4-2. A certain phonograph pickup delivers a sinusoidal voltage of 10 mv, rms value, on open circuit, and it has an internal resistance of 500 ohms. This pickup is to deliver a signal to a loudspeaker having a resistance of 10 ohms.

a. If the pickup is connected directly to the loudspeaker, how much power does it deliver to the speaker?

b. A cascade of amplifier stages similar to that shown in Fig. 4-2 is to be used to amplify the signal from the pickup and deliver 10 watts to the loudspeaker. Two types of amplifiers, shown in Fig. 4-9, are available for the cascade connection. These

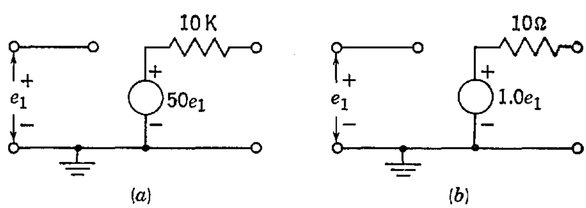


FIG. 4-9. Amplifiers for Prob. 4-2. (*a*) Type 1; (*b*) type 2.

are not ideal voltage amplifiers, for they do not have zero output impedance. Which of these amplifiers can deliver 10 watts to the loudspeaker with the smallest input voltage?

c. Which of the amplifiers in Fig. 4-9 gives the greatest no-load voltage amplification?

d. Give the circuit diagram for a cascade of stages like those in Fig. 4-9 that will meet the specifications of part *b*. Use the smallest possible number of stages, and show the exact number of stages in the diagram. Show how a 1-megohm potentiometer can be used in the first stage as a volume control.

4-3. The capacitor in the circuit of Fig. 4-6 is replaced by a resistor with R = 1000 ohms, and the amplification factor is $\mu = 20$.

a. Determine the input resistance, $R_n = e_1/i_1$.

b. What is the value of R_n if the polarity of the controlled source is reversed? Give the correct algebraic sign.

4-4. A 10-mh inductor is to be simulated with a 1-mh inductor and a voltage-controlled voltage source.

a. Prove that the desired result can be accomplished by a circuit similar to that shown in Fig. 4-6.

b. What value of amplification factor is required?

4-5. In a certain series resonant circuit the parameters are L = 10 mh, C = 0.01 μf, and R = 100 ohms. The resonant Q of the circuit is $\omega_o L/R$, where ω_o is the resonant frequency in radians per second.

a. Determine the resonant Q of the circuit.

b Show how the negative impedance converter of Fig. 4-7 can be used with a resistance R' to increase the resonant Q of the circuit to 100. If the amplification factor of the negative impedance converter is $\alpha = 2$, what value of R' is required?

4-6. The circuit shown in Fig. 4-10 is the idealized form of a type of voltage regulator that is commonly used in electronic power supplies. The function of the regulator s, among other things, to hold the output voltage constant in the face of changes in a-c line voltage and changes in load connected across E_2. Any change in the output voltage is amplified and causes a change in the voltage $6E_5$ that tends to compensate for the initial change in E_2. By taking E_3 from a potentiometer instead of the fixed voltage divider shown in Fig. 4-10, the output voltage can be made adjustable.

a. Let E_1 be 350 volts, d-c. Starting with the relation $E_2 = E_1 + 6E_5$, find the value of E_2. (Be careful with the signs.)

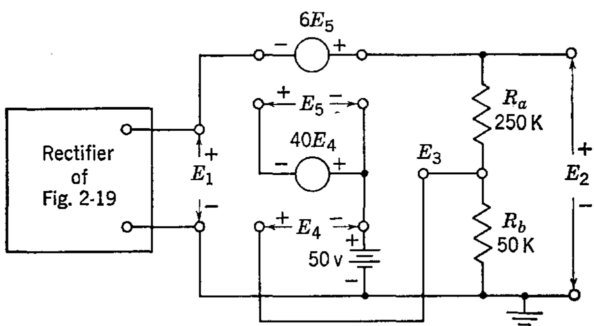


FIG. 4-10. An idealized voltage regulator for Prob. 4-6.

b. If E_1 increases by 50 volts, what is the new value of E_2?

c. What is the value of the ratio $\Delta E_2/\Delta E_1$?

d. If the peak-to-peak ripple in E_1 is 25 volts, what is the peak-to-peak value of the ripple in E_2? The results of the previous parts can be used in this calculation. On the basis of these results, would you say that the rectifier of Fig. 2-19 can use a smaller filter capacitor when it is used with a voltage regulator than when it is used alone?

4-7. Control of the output voltage from the regulator of Prob. 4-6 by means of the output voltage-divider ratio is to be studied. The sum of the voltage-divider resistances is to remain fixed at 300 kilohms, but the tap is to be moved. These conditions correspond to the case in which a potentiometer is used for the voltage divider. The input voltage to the regulator is constant at $E_1 = 400$ volts, d-c.

a. Starting with the relation $6E_5 = E_2 - E_1$, find the values of R_a and R_b required to make $E_2 = 300$ volts. (Be careful with the signs.)

b. Repeat part *a* for $E_2 = 200$ volts.

c. The output voltage has its smallest possible value when $R_a = 0$, $R_b = 300$ kilohms, and $E_3 = E_2$. Starting with the relation $E_2 = E_1 + 6E_5$, determine the smallest value to which E_2 can be adjusted.

CHAPTER 5

THE BASIC VACUUM-TRIODE AMPLIFIER

The concept of the ideal amplifier is presented and some of the useful things that can be done with such a device are discussed in Chap. 4. The next step is to consider what physical devices might be used as amplifiers. There are a great many devices that can amplify, each one deviating in certain ways from the ideal. For example, the d-c generator is capable of amplification; a small amount of power applied to the field winding can control a large amount of power in the armature circuit. Therefore the d-c generator is used as a power amplifier in certain applications. One of the principal limitations on the d-c generator as an amplifier is the slowness with which it responds to input signals; one of its advantages is the relatively large output power that it can deliver. The vacuum tube is a better amplifier for many applications because its response to input signals is extremely fast.

The first objective of this chapter is to examine the vacuum triode as an amplifier and to compare its characteristics with those of the ideal amplifier. The physical laws governing the behavior of the triode are studied qualitatively so that the reasons for departure from the ideal characteristics can be perceived and so that the limitations on the use of the triode can be understood. Finally, an introduction to the analysis and design of elementary triode circuits is presented. In this latter phase of the study, the vacuum tube is the specific amplifier involved. The methods are quite general, however, and apply equally well, with minor modifications, to other amplifiers, both electrical and nonelectrical.

5-1. The Vacuum Triode as an Amplifier. The vacuum triode is shown schematically in a simple amplifier circuit in Fig. 5-1. The cathode and anode in small triodes intended for low-power applications are usually similar to those described in Sec. 3-1 for the vacuum diode. The grid is usually a spiral of wire surrounding the cathode and lying in the space between the cathode and anode. When the plate of the tube is held at a suitable positive potential relative to the cathode, electrons flow from the cathode to the anode through the space between the grid wires. If e_s is always negative so that the grid is negative relative to the cathode by half of a volt or more, very few electrons have enough kinetic energy to reach the grid wires and the grid current is essentially zero. Thus there is negligible power input to the grid-cathode terminal pair, and the first requirement of a good amplifier is met. However, the grid is able to control the relatively large amount of power delivered to the load resistance R_L.

If the grid is made more negative in the circuit of Fig. 5-1, the potential-distribution curve in the space between the grid and cathode is lowered, and as a consequence the plate current, the load current, the load voltage, and the load power are all decreased. The tube acts as a valve for current in the plate circuit. For this reason the British use the term "valve" instead of "vacuum tube" in referring to electron tubes. In a typical circuit a 1-volt change in e_s causes i_b to change by 0.3 ma and causes the load voltage to change by 15 volts. Thus there is both voltage and power amplification.

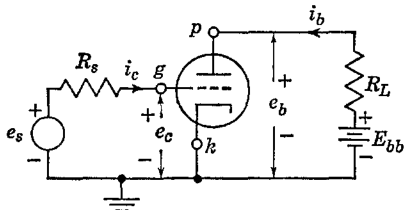


FIG. 5-1. The basic triode amplifier circuit.

The performance of the triode as an amplifier can be examined further by comparing its input and output volt-ampere characteristics with the characteristics of the ideal voltage amplifier presented in Chap. 4. A typical set of experimentally measured triode characteristics is shown in Fig. 5-2; the symbols used on these characteristics are defined in the circuit diagram of Fig. 5-1. If the grid and plate voltages, e_c and e_b, are specified, the input current to the triode, i_c, is fixed. Similarly, if e_c and e_b are specified, the plate current i_b is also fixed; that is,

$$i_c = f_c(e_b,e_c) \qquad \text{and} \qquad i_b = f_b(e_b,e_c)$$

The input and output characteristics of Fig. 5-2 are graphs of these functional relations. The input characteristic is a family of curves of i_c versus e_c for various fixed values of e_b; the output characteristic is a family of curves of i_b versus e_b for various fixed values of e_c. These two families of curves contain complete information as to the relations between the tube voltages and currents.

The input, or grid, characteristic shows that the input current and power are negligibly small provided the grid is kept sufficiently negative relative to the cathode. In this respect the triode behaves like the ideal voltage amplifier. Vacuum-tube amplifiers are operated with negative grid voltage in the great majority of their applications.

The output, or plate, characteristic shows that the plate current depends strongly on the grid voltage when the plate voltage is sufficiently positive; this fact is the basis for amplifier action in the triode. As in the case of the vacuum diode, the plate current can never be negative; hence the plate characteristic shows the plate current different from zero only in the first quadrant. These characteristics are quite similar in certain important respects to the output characteristic of the ideal voltage amplifier shown in Fig. 4-1*b*. The triode characteristics form a useful family only in the first quadrant, and compared with the ideal characteristics, they are rotated clockwise so that they are not vertical. This rotation is evidence of the fact that the triode does not have zero internal resistance. The triode differs from the ideal voltage amplifier further in that its characteristics are not straight lines and they are not exactly uniformly spaced for equal increments in grid voltage. These facts mean that the plate current does not vary linearly with the grid and plate voltages; hence the waveform of variations in input voltage is not

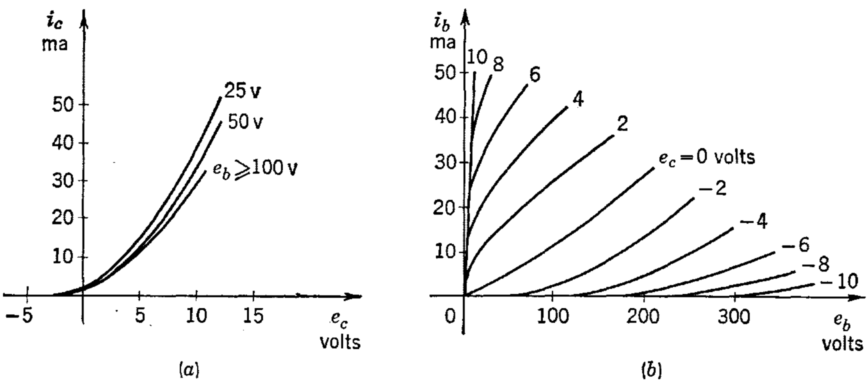


FIG. 5-2. Static triode characteristics. (*a*) Input characteristic, $i_c = f_c(e_b,e_c)$; (*b*) output characteristic, $i_b = f_b(e_b,e_c)$.

reproduced exactly in the plate circuit. This nonlinearity may result in appreciable distortion of the signal waveform. A minor difference between the triode and the ideal amplifier of Fig. 4-1 is that a positive increment of input voltage shifts the triode characteristic to the left rather than to the right.

5-2. Current Flow in Vacuum Triodes. The flow of current in the triode and its control by the grid potential can be studied with the aid of potential-distribution curves for the space between the plate and cathode. Figure 5-3*a* represents a triode whose plate and cathode are infinite parallel planes; the grid is a set of parallel wires located in a plane parallel to and between the plate and cathode. Figure 5-3*b* shows the potential distribution along two paths between the plate and cathode for typical operating conditions with the grid at a small negative potential and the plate at a larger positive potential relative to the cathode. Path *A* passes midway between two grid wires, and path *B* passes through one of the grid wires. Both of these paths are lines of geometric symmetry; hence the electric field, and the force on an electron at all points on each path, is directed along the path either toward the plate or toward the cathode.

Under conditions of space-charge-limited current, the potential along path A has a minimum value just in front of the cathode; those electrons emitted from the cathode with enough kinetic energy pass the potential minimum and proceed to the plate. Along path B the minimum potential occurs at the grid and is equal to the applied grid voltage. Very few electrons leave the cathode with enough kinetic energy to reach the grid; hence the grid current is negligible. If the grid is made more negative, the potential-distribution curve between the grid and the

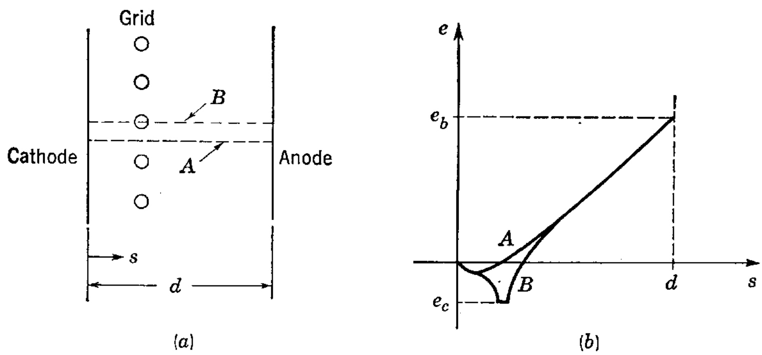


FIG. 5-3. Potential distribution in a triode. (*a*) The triode; (*b*) the potential distribution.

cathode is lowered, fewer electrons are able to pass the potential minimum, and the plate current is decreased. If the grid is made sufficiently negative, practically no electrons are able to pass the potential minimum, and the plate current is essentially zero; this condition is called plate-current cutoff.

With the grid voltage held constant, the plate current varies with the plate voltage in much the same manner as in the vacuum diode; thus the individual characteristic curves in Fig. 5-2*b* are similar in shape to the diode characteristic. As the grid is made more negative, the plate must be made more positive to give the same plate current; hence the curves in Fig. 5-2*b* are shifted to the right for more negative grid voltages. Since the grid is more effective than the plate in controlling the potential in the grid-cathode space, a large change in plate voltage is required to counteract the effect of a small change in grid voltage.

Figure 5-4*a* shows the potential distribution in a triode when the grid is positive, but less positive than the plate. The space-charge-limited current that flows is relatively large under these conditions. Since the grid is positive relative to the cathode, any electron leaving the cathode can reach the grid, and there is an appreciable grid current as indicated in Fig. 5-2*a*. However, most of the electrons pass between the grid wires and continue to the plate. Figure 5-4*b* shows the kind of potential distribution that can exist when the grid is more positive than the plate. Electrons corresponding to a large plate current move relatively slowly across the interelectrode space, and a large space charge exists in the space between the grid and plate. This space charge depresses the potential in the grid-plate space and may cause a minimum to occur at a small negative potential as shown at point *b* in Fig. 5-4*b*. This potential minimum affects the electron flow in exactly the same way as the minimum in front of the cathode at point *a*. Those electrons that leave the cathode with sufficient kinetic energy pass the potential minimum at *b*

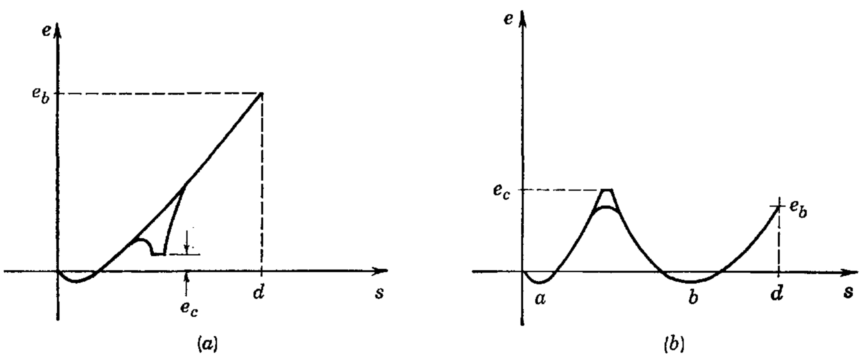


Fig. 5-4. Potential distribution in a triode with positive grid voltage. (*a*) $e_b > e_c > 0$; (*b*) $e_c > e_b > 0$.

and continue to the plate; the remainder are turned back and are ultimately collected by the grid. Thus a virtual cathode exists at point *b*, and the plate current varies with plate voltage much as it would in a diode with its cathode at point *b*. This condition corresponds to that portion of the plate characteristic in Fig. 5-2*b* at low plate voltage where the characteristics for various grid voltages merge into a single curve. This single curve is much like a diode characteristic. The points at which the individual characteristic curves break away from the common curve correspond to plate voltages at which the virtual cathode disappears.

In addition to the input and output characteristics of Fig. 5-2, there are other triode characteristics that display the properties of the tube in different ways. One of these is the constant-current voltage transfer characteristic shown in Fig. 5-5. Each curve in this family shows the combinations of grid and plate voltage that give a certain fixed value of plate current. This family can be plotted from the data contained in the plate characteristic of Fig. 5-2*b*, or it can be measured directly. Since these characteristics are quite linear over a wide range of grid voltage, it follows that a highly linear voltage amplifier could be built by using a current source for the plate-circuit power supply in the amplifier of Fig. 5-1. Unfortunately, practical considerations usually rule this possibility out.

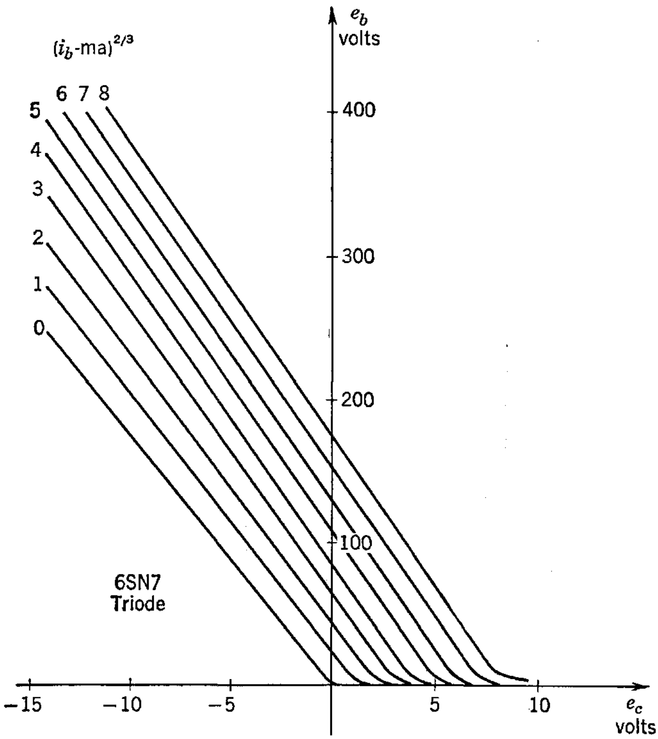


FIG. 5-5. Static triode constant-current characteristic.

The constant-current characteristics are not only linear, but they are also nearly parallel except at low values of plate current, and as shown by Fig. 5-5, they are approximately equally spaced for equal increments of $(i_b)^{2/3}$. Thus the equation for the straight portion of any one of the curves is the equation of a straight line,

$$e_b = A - \mu e_c \tag{5-1}$$

where A is the intercept on the e_b axis and $-\mu$ is the slope of the line. It is shown later that the factor μ is the voltage amplification factor of the triode; it corresponds to the amplification factor of the ideal amplifier in Sec. 4-1. Since the curve for $i_b = 0$ passes through the origin, and since the curves are approximately equally spaced for equal increments of $(i_b)^{2/3}$, the intercept A is proportional to $(i_b)^{2/3}$. Hence

$$A = Bi_b^{2/3} \tag{5-2}$$

and (5-1) takes the form

$$e_b = Bi_b^{2/3} - \mu e_c$$

Solving for i_b gives

$$i_b = K(e_b + \mu e_c)^{3/2} \tag{5-3}$$

where K, the perveance of the triode, depends on the size and shape of the electrodes. This is the three-halves-power law for the triode; it holds reasonably well over the range of voltages corresponding to the linear portion of the constant-current characteristic, including both positive and negative values of grid voltage. However, it is of limited value in the analysis and design of triode circuits because it is a nonlinear relation between the current and the voltages.

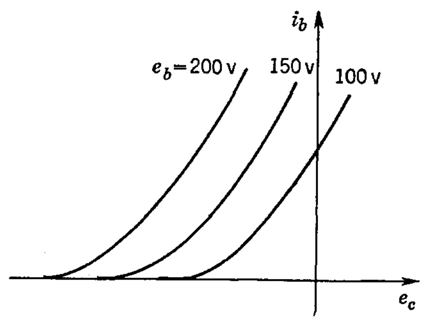


FIG. 5-6. Static triode constant-voltage characteristic.

Another set of triode characteristics, the constant-voltage transfer characteristics, is shown in Fig. 5-6. This set of curves is useful in the study of certain types of triode circuits; it can be constructed from the data contained in the plate characteristics of Fig. 5-2*b*.

The discussion of the triode up to this point is concerned with static conditions of voltage and current. When the tube is used to amplify signals at frequencies greater than a few tens of kilocycles per second, it is found that the capacitances between the electrodes may have an important effect. For example, the capacitance between plate and cathode is in parallel with the output terminals, and at sufficiently high frequencies it tends to short-circuit the output terminals. The interelectrode capacitances are the primary factors limiting the high-frequency performance of the tube. These capacitances are rather complicated in nature, for they depend on the amount and location of the space charge in the tube. It follows from this fact that the capacitance values depend on the currents and voltages in the tube. It is customary to represent the interelectrode capacitances approximately by three fixed capacitances, one between grid and cathode, one between plate and cathode, and one between grid and plate. In typical small triodes these capacitances are roughly equal, and their values usually lie in the range between 1 and

5 $\mu\mu$f. However, the amplifying property of the tube has the effect of magnifying the action of the grid-plate capacitance in the manner discussed in connection with Fig. 4-6; this important matter is considered in more detail in Sec. 14-3.

Another imperfection of the triode amplifier is the noise current that is generated by the tube. The charge arriving at the plate of the tube does not come as the smooth flow of a continuous medium; it is associated with the flow of discrete electrons showering on the plate. Moreover, the emission from the cathode is not exactly the same at every instant of time, and the electrons arrive at the plate in clouds. The result is that the plate current consists of an average value plus a very small time-varying noise component.[3] The time variations of the noise component of current are of a random nature, and they can be represented as the summation of small sinusoidal currents with frequencies distributed continuously over a wide band from zero to very high frequencies. It is customary to account for the noise component of current by assuming the tube to be noiseless and connecting in series with the grid circuit a noise-voltage source to give the noise component of plate current. It is clear that the tube noise places a lower limit on the magnitude of signal that can be amplified by the tube. If the signal magnitude is less than the magnitude of the equivalent noise voltage, it is masked by the noise.

As in the case of the diode, heat is generated at the plate of the triode by the impact of electrons arriving from the cathode, and as in the case of the diode, the rate at which heat energy is generated at any instant is

$$p = e_b i_b \qquad (5\text{-}4)$$

For any tube there is a maximum rate at which heat can be generated at the plate without an excessive temperature rise. This is the maximum permissible average plate dissipation. Under static operating conditions, with e_b and i_b constant at the values E_b and I_b, the average power dissipated at the plate is $P = E_b I_b$. Setting this quantity equal to the maximum permissible plate dissipation gives a relation between E_b and I_b that describes a hyperbola on the plate characteristic. The general form of this maximum-plate-dissipation hyperbola is shown in Fig. 5-7. Under static operating conditions, combinations of E_b and I_b correspond-

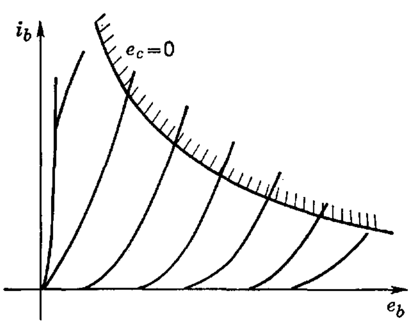


FIG. 5-7. Static plate-dissipation hyperbola.

ing to points lying above this curve give excessive plate dissipation and will damage the tube. Under dynamic operating conditions, the plate dissipation depends on the waveforms of e_b and i_b. Analyses of typical dynamic operating conditions yield operating constraints similar to the static dissipation hyperbola shown in Fig. 5-7. Typical values of maximum permissible plate dissipation for small triodes range roughly from 1 to 15 watts.

The cathodes used in small vacuum triodes are like those used in small diodes, and, as in the case of the diode, there is a limit to the amount of current that can be drawn from the cathode without damage to the oxide coating. The tube manufacturer usually specifies the maximum permissible cathode current. In addition, maximum permissible values of electrode voltages are usually specified.

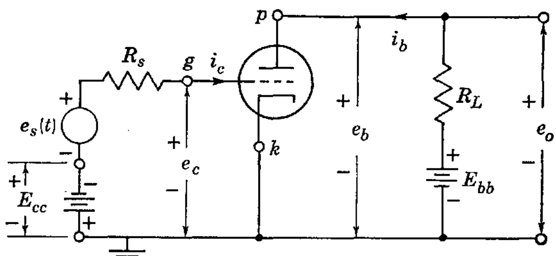


FIG. 5-8. The basic triode amplifier.

5-3. Graphical Analysis. The basic triode amplifier is shown in Fig. 5-8. The signal to be amplified, $e_s(t)$, is in general a time-varying voltage; indeed, there would be no point in amplifying a voltage that remains constant for all time. Furthermore, e_s usually has no d-c component (its average value is zero); if the average value of e_s is not zero in any particular case, the d-c component will be associated with the direct voltage E_{cc}. In the most common mode of operation the grid-bias voltage E_{cc} is negative and of such magnitude that the signal can never make the grid positive relative to the cathode; hence no grid current flows, and no power is drawn from the signal source. The plate supply voltage E_{bb} is chosen so that with no signal applied, the plate voltage e_b and the plate current i_b correspond to a point in the useful region of the plate-characteristic curves; such a point is indicated at Q in Fig. 5-9. Thus the grid-bias and the plate-supply voltages shift the operating point for the tube from the origin of the plate characteristic to a more suitable location. In all four quadrants surrounding this shifted operating point, the tube characteristics are much like the characteristics of the ideal voltage amplifier, except for a rotation.

The circuit of Fig. 5-8 can be used either as a power amplifier or as a voltage amplifier. When it is used as a power amplifier the resistance

R_L is usually the load. When it is used as a voltage amplifier the voltage between plate and ground is usually taken as the output voltage e_o. The variations in e_s are amplified and reproduced approximately as variations in voltage drop across R_L. It is interesting to view R_L and E_{bb} as an approximation to a current source for the plate-circuit power supply; the larger the R_L the better the approximation and the more linear the amplifier, provided i_b is not too small. The resistance R_s in Fig. 5-8 accounts for any resistance associated with the signal source.

The analysis problem consists of finding the power delivered to the load or the voltage delivered to the output terminals when the tube, the circuit resistances, and the applied voltages are specified. The plate current is given by Eq. (5-3) as a nonlinear function of the plate and grid voltages; hence an algebraic solution is difficult. For this reason it is customary to resort to a graphical solution[1,2] as is done in the case of the vacuum diode. In the analysis of simple circuits like that of Fig. 5-8 it is also customary to follow certain conventions regarding the symbols used. These conventions are listed as follows:

E_{cc} = grid-bias voltage
e_c = instantaneous potential of the grid relative to the cathode
e_g = instantaneous value of the time-varying component of e_c
E_c = average value of e_c
i_c = instantaneous value of the grid current
i_g = instantaneous value of the time-varying component of i_c
I_c = average value of i_c
E_{bb} = plate-supply voltage
e_b = instantaneous potential of the plate relative to the cathode
e_p = instantaneous value of the time-varying component of e_b
E_b = average value of e_b
i_b = instantaneous value of the plate current
i_p = instantaneous value of the time-varying component of i_b
I_b = average value of i_b

Certain additional symbols will be introduced as the need for them arises.

The circuit of Fig. 5-8 has two loops; the voltage-law equations for these loops are

$$e_c = E_{cc} + e_s - R_s i_c \tag{5-5}$$

and

$$e_b = E_{bb} - R_L i_b \tag{5-6}$$

These two equations contain four unknowns, e_c, i_c, e_b, and i_b; two additional relations among these variables are needed to obtain a solution. The necessary additional relations are given by the input and output characteristics of the tube; they represent graphically the relations

$$i_c = f_c(e_b,e_c) \qquad \text{and} \qquad i_b = f_b(e_b,e_c) \tag{5-7}$$

These four relations can be solved simultaneously by graphical constructions on the tube characteristics.

Consider first the case in which E_{cc} is chosen so that the grid is always negative relative to the cathode and the grid current is always zero. The grid voltage is then given by (5-5) in terms of the known applied voltages, and only two unknowns, i_b and e_b, remain to be found. The plate-characteristic family gives a relation between these two variables that is imposed by the tube. For any given value of e_c the relation between i_b and e_b is one curve of the family. Equation (5-6) gives another relation between i_b and e_b; this relation is imposed by the external circuit connected to the tube. The fact that these two relations must be satisfied simultaneously fixes i_b and e_b; the graphical solution is shown in Fig. 5-9. This construction can be interpreted by writing (5-6) as

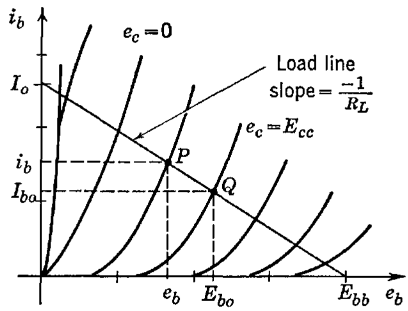


FIG. 5-9. Graphical analysis of the triode amplifier.

$$i_b = \frac{E_{bb} - e_b}{R_L} = I_o - \frac{e_b}{R_L} \tag{5-8}$$

This is the equation of the load line. Only those combinations of i_b and e_b that lie on this line are permitted by the external circuit connected to the tube. When $e_s = 0$, the grid voltage is E_{cc}, and only those combinations of i_b and e_b that lie on the plate characteristic for $e_c = E_{cc}$ are permitted by the tube. Hence the intersection of these two lines at Q gives the simultaneous solution for $e_s = 0$.

The operating point of the tube is at Q when there is no signal input; thus it is called the quiescent operating point, and I_{bo} and E_{bo} are the quiescent plate current and voltage. If a signal is applied so that e_c varies with time, the operating point must move along the load line, occupying at each instant the point at which the load line intersects the appropriate plate-characteristic curve. One such point is designated P in Fig. 5-9. The path covered by the operating point while the input signal varies is called the operating path; for many amplifier circuits it never extends into the positive-grid region and never reaches the $i_b = 0$ point. If e_s is given as a function of time, waveforms of i_b and e_b can be constructed by this graphical process, the load voltage and load power can be evaluated, and the waveform distortion introduced by the nonlinearity of the tube can be studied.

If the combination of e_s and E_{cc} is such that grid current flows during part or all of the time, the analysis is slightly more complicated, for in this case e_c and i_c are also unknowns. However, the grid circuit can be solved graphically for e_c in exactly the same manner as the plate circuit is solved for e_b. The appropriate construction is shown in Fig. 5-10 with the assumption that the input characteristic can be represented with suitable accuracy by a single curve for all values of plate voltage. Having determined e_c by this construction, the plate current and voltage can be determined by the construction shown in Fig. 5-9.

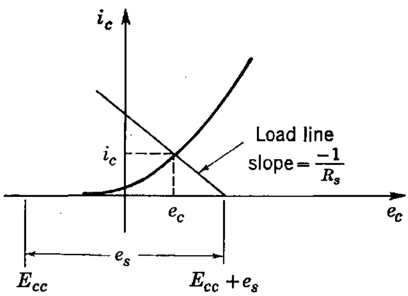


FIG. 5-10. Grid-circuit analysis.

The construction shown in Fig. 5-10 is typical when R_s is a few hundred ohms. If R_s is 10,000 or 100,000 ohms, as often is the case, the grid-circuit load line is almost horizontal, and e_c remains essentially zero for all reasonable positive values of e_s. Thus if e_s has occasional peaks that tend to drive the grid positive, these peaks do not appear in the grid voltage e_c; they are clipped off by the action of the grid. Such clipping is sometimes used deliberately.

Example 5-1. A triode amplifier is shown in Fig. 5-11*a*; the input and output volt-ampere characteristics of the tube are shown in Figs. 5-11*b* and *c*. The maximum permissible plate dissipation is 1.5 watts, and the maximum permissible grid dissipation is 0.5 watt.

a. Find the quiescent tube voltages and currents.

b. Is the maximum permissible plate dissipation exceeded?

c. Is the maximum permissible grid dissipation exceeded?

Solution. a. Under quiescent conditions the grid-circuit load line is given by Eq. (5-5) with $e_s = 0$. This line crosses the grid-voltage axis at $e_c = E_{cc} = 10$ volts, and it crosses the grid-current axis at $i_c = E_{cc}/R_s = {}^{10}\!/_1 = 10$ ma. The corresponding load line is shown in Fig. 5-11*b*. The load line intersects all the grid-characteristic curves at approximately the same point; hence the quiescent grid voltage is approximately 3 volts, and the quiescent grid current is approximately 7 ma.

The plate-circuit load line is given by Eq. (5-6). It crosses the plate-voltage axis at $e_b = E_{bb} = 300$ volts, and it crosses the plate-current axis at $i_b = E_{bb}/R_L = {}^{300}\!/_{10} = 30$ ma. This load line is shown in Fig. 5-11*c*. The grid voltage, determined above, is 3 volts; hence the operating point for the plate circuit is at the intersection of the load line with the plate characteristic for $e_c = 3$ volts. This point is approximately midway between the characteristics for $e_c = 2$ volts and $e_c = 4$ volts; it is shown at Q in Fig. 5-11*c*. Thus the quiescent plate voltage is approximately 60 volts, and the quiescent plate current is approximately 25 ma.

b. The plate dissipation is

$$P_b = E_b I_b = (60)(25)(10)^{-3} = 1.5 \text{ watts}$$

This is exactly the maximum permissible plate dissipation.

c. The grid dissipation is

$$P_c = E_c I_c = (3)(7)(10)^{-3} = 0.021 \text{ watt}$$

Thus the maximum permissible value of 0.5 watt is not exceeded.

The signal transmission properties of the triode amplifier are summarized and conveniently displayed by the voltage transfer characteristic shown in Fig. 5-12. This characteristic shows the output voltage e_o as a function of the input-signal voltage e_s for a particular circuit. It is

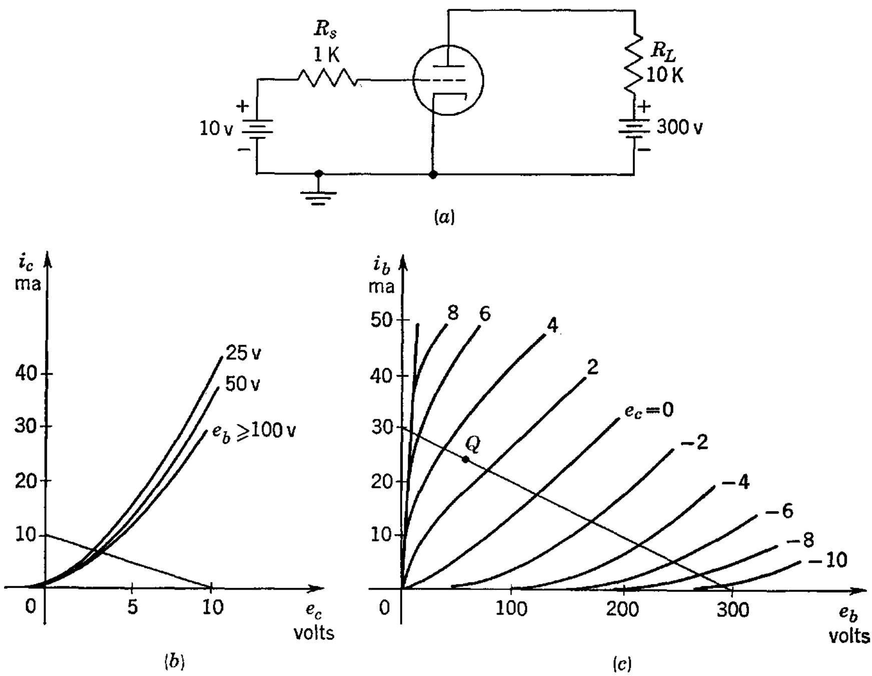


FIG. 5-11. Graphical analysis of a triode amplifier. (*a*) Circuit; (*b*) grid-circuit relations; (*c*) plate-circuit relations.

constructed by reading values of $e_o = e_b$ from the load line on the plate characteristic for values of e_s chosen over a wide range from the positive grid region to plate-current cutoff. In the positive grid region, e_c as well as e_b must be determined graphically. For large negative values of e_s the plate current is cut off, and e_o is constant at the value E_{bb}. For positive values of e_s, grid current flows and e_c remains approximately constant at zero volts. The most common mode of operation is on the steep, linear part of the curve where a relatively small change in e_s gives a large change in e_o. The ratio of the change in e_o to the change in e_s is the voltage amplification of the circuit; thus the small-signal amplification is equal to the slope of the transfer characteristic at the operating point.

It is interesting to compare the voltage transfer characteristic of the amplifier with the constant-current transfer characteristics of the tube given in Fig. 5-5.

If the voltage transfer characteristic were perfectly linear in the operating region, the variations in e_o would be a magnified copy of the variations in e_s, and the amplification would be distortionless. Over the

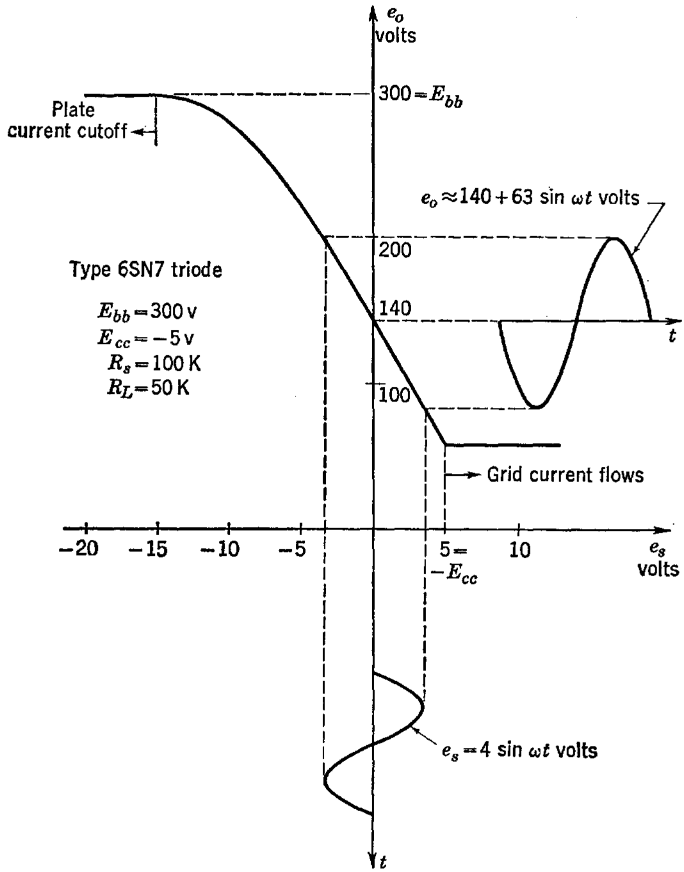


FIG. 5-12. Voltage transfer characteristic for the triode amplifier.

large-signal range, unfortunately, the curve is not linear because the plate characteristics for equal increments of grid voltage do not intersect the load line at equal intervals. Thus the waveform of e_s is not exactly duplicated in e_o. The worst distortion occurs at small plate currents where the plate characteristics are quite crowded; this fact appears in Fig. 5-12 as the rounding of the upper knee of the characteristic. The construction in Fig. 5-12 shows how the voltage transfer characteristic can be used to determine the waveform of the output e_o when the waveform of the input e_s is given. It is clear from the type of curvature in the characteristic that the peaks in the output voltage are somewhat flattened and the valleys are somewhat exaggerated.

When the voltages applied to an amplifier are chosen so that the plate current is never cut off, the mode of operation is designated Class A. Class B designates the mode of operation in which E_{cc} and E_{bb} are chosen so that the plate current is just cut off when no signal is applied. When the tube is biased beyond the Class B condition, the operation is designated Class C. In addition, the term Class AB is used to describe operating conditions that are intermediate between Classes A and B. When no grid current flows, the fact is signified by adding the subscript 1 to the class letter; the subscript 2 indicates that grid current flows. Voltage amplifiers of all types almost always operate in the Class A_1 mode. Power amplifiers for audio frequencies usually operate in either the Class A_1 or the Class AB_2 mode. Power amplifiers for radio frequencies usually operate in the Class C_2 mode.

The distortion resulting from the nonlinear properties of the triode is a matter of basic importance, for it often limits the performance obtainable from the tube. Therefore it is necessary to examine more closely some of the consequences of distortion.[1,3] One aspect of distortion is illustrated in Fig. 5-13. The input signal to the basic amplifier is assumed to be a square wave with an amplitude of 5 volts; the grid bias is assumed to be -5 volts. The resulting current and voltage relations are revealed by the graphical construction on the plate characteristics. With no signal applied, the plate current is constant at the value $I_{bo} = 6.5$ ma. When the signal is applied, the plate current increases by 5.5 ma during the positive half cycle of e_s, and it decreases by 4.5 ma during the negative half cycle. Thus equal positive and negative swings in e_s do not give equal swings in i_b and e_o. As a result, the average current increases from $I_{bo} = 6.5$ ma with no signal to $I_b = 7.0$ ma with signal applied. This change in I_b results from the distortion introduced by the tube; it is often used as an indication that distortion is occurring, and it is occasionally used as a rough measure of the amount of distortion.

Another aspect of signal distortion can be examined by assuming a sinusoidal input signal like that pictured in Fig. 5-12. The output voltage is periodic, but is not sinusoidal; therefore it can be represented by a Fourier series of the form

$$e_o = E_o + E_1 \cos (\omega_1 t + \theta_1) + E_2 \cos (\omega_2 t + \theta_2) + \cdots \qquad (5\text{-}9)$$

The second and higher-order harmonic terms represent distortion, for they are not present in the input. In the usual power amplifier intended for the amplification of sound, the rms value of the distortion components is of the order of 1 to 5 per cent of the rms signal component when the amplifier is delivering rated output with a sinusoidal signal. In voltage amplifiers, where the signal voltages may be quite small, the distortion is much smaller and is ordinarily neglected.

In the practical case where the signals are not sinusoidal, a type of distortion appears that is not contained in Eq. (5-9). Suppose, for

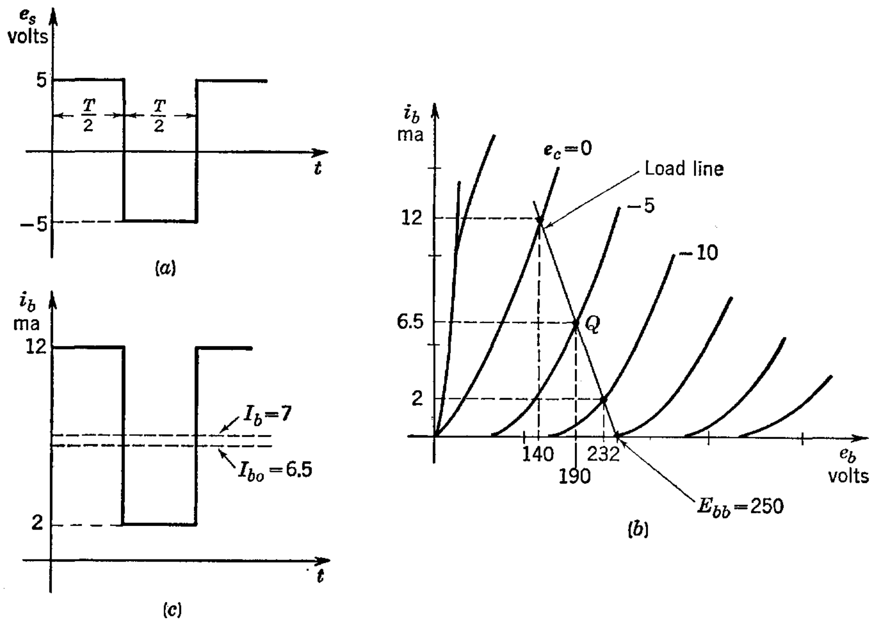


FIG. 5-13. Nonlinearity in triode amplifiers. (*a*) Input signal; (*b*) graphical analysis; (*c*) plate-current waveform.

example, that the signal voltage consists of two sinusoids that are not harmonically related:

$$e_s = E_{s1} \cos (\omega_1 t) + E_{s2} \cos (\omega_2 t) \tag{5-10}$$

The problem is to find the distortion components that appear in the output. Since the voltage transfer characteristic of Fig. 5-12 is not linear, e_o cannot be expressed as a linear function of e_s; it must depend also on higher powers of e_s. In fact, e_o can be expressed as a power series in e_s over a range of e_s; hence

$$e_o = a_o + a_1 e_s + a_2 e_s^2 + \cdots \tag{5-11}$$

These three terms are enough to illustrate the nature of the distortion. The first term is a constant, the second term is an undistorted reproduction of the signal, and the third term is all distortion. From (5-10),

$$e_s^2 = E_{s1}^2 \cos^2 (\omega_1 t) + E_{s2}^2 \cos^2 (\omega_2 t) + 2E_{s1}E_{s2}(\cos \omega_1 t)(\cos \omega_2 t) \qquad (5\text{-}12)$$

But
$$\cos^2 \omega_1 t = \tfrac{1}{2}(1 + \cos 2\omega_1 t) \qquad (5\text{-}13)$$
$$\cos^2 \omega_2 t = \tfrac{1}{2}(1 + \cos 2\omega_2 t) \qquad (5\text{-}14)$$
and
$$(\cos \omega_1 t)(\cos \omega_2 t) = \tfrac{1}{2} \cos (\omega_1 + \omega_2)t + \tfrac{1}{2} \cos (\omega_1 - \omega_2)t \qquad (5\text{-}15)$$

Thus the distortion not only produces harmonics of the input sinusoids, but it also produces components at the sum and difference frequencies. The higher-order terms in (5-11) produce similar distortion components. The mechanism by which the sum and difference frequencies are generated is called intermodulation. Since it results in components that are not harmonically related to the signal, it is one of the most objectionable types of distortion in audio amplifiers. On the other hand, however, it has many useful engineering applications.

5-4. Power Relations in the Triode Amplifier. The plate current in the triode amplifier can be expressed as

$$i_b = I_b + i_p \qquad (5\text{-}16)$$

where I_b is the average value of the current, and i_p is the time-varying component having zero average value. Similarly,

$$e_b = E_b + e_p \qquad (5\text{-}17)$$

where E_b is the average value of the voltage, and e_p is the time-varying component having zero average value. The power drawn from the plate supply in the basic amplifier circuit at each instant is thus

$$p_{bb} = E_{bb}i_b = E_{bb}I_b + E_{bb}i_p \qquad (5\text{-}18)$$

Since E_{bb} is constant and i_p has zero average value, the average value of $E_{bb}i_p$ is zero; hence the average power drawn from the plate supply is

$$P_{bb} = E_{bb}I_b \qquad (5\text{-}19)$$

If the distortion in the amplifier is negligible, and if the input signal has zero average value, then I_b is equal to the quiescent current I_{bo} and is independent of the signal amplitude. Under these conditions the power drawn from the plate supply is also independent of the signal level.

The power absorbed by the load at each instant is

$$p_L = R_L i_b^2 \qquad (5\text{-}20)$$

The average power absorbed by the load is

$$P_L = R_L(i_b^2)_{av} = R_L(I_{b\ rms})^2 = R_L I_b^2 + R_L(I_{p\ rms})^2 \qquad (5\text{-}21)$$

The power dissipated at the plate of the tube at each instant is

$$p_b = e_b i_b = (E_{bb} - i_b R_L)i_b = E_{bb}i_b - R_L i_b^2 \qquad (5\text{-}22)$$

But from (5-18), $p_{bb} = E_{bb}i_b$, and from (5-20), $p_L = R_L i_b^2$; hence

$$p_b = p_{bb} - p_L \tag{5-23}$$

and the average power dissipated at the plate is

$$P_b = P_{bb} - P_L \tag{5-24}$$

Since P_{bb} and P_L are always positive, and since P_L has its minimum value when no signal is applied, it follows from (5-24) that the power dissipated by the tube has its maximum value under no-signal conditions. When signal is applied, P_L increases, and P_b decreases by exactly the same amount; the power drawn from the plate supply remains constant as long as the distortion is negligible.

It follows from these relations that in designing an amplifier to operate with low distortion, the circuit parameters must be chosen so that the quiescent point on the plate characteristic lies on or below the hyperbola of maximum permissible plate dissipation shown in Fig. 5-7. Then, for signals of any magnitude or waveform, it is ensured that the plate dissipation will not exceed the permissible value, provided that distortion does not cause a change in the average value of the current.

The currents and voltages in electronic circuits usually consist of a d-c component plus a time-varying component, and the time-varying components may have a wide variety of waveforms. Therefore computations of average power must be made with care. If e is the potential difference between any pair of terminals, and if i is the current from the high- to the low-potential terminal, then the power delivered to the terminals at any instant is

$$p = ei \tag{5-25}$$

This relation is always true; it results directly from the definitions of voltage and current. The average power delivered to the terminals, which is often the quantity of most interest, is accordingly

$$P = (ei)_{av} \tag{5-26}$$

The average of the product ei must be found by proper means. In general,

$$P \neq e_{av}i_{av} \tag{5-27}$$

The average power equals the product $e_{av}i_{av}$ only in the special cases where either e or i or both do not vary with time. If $e = E + e_s$ and $i = I + i_s$, where E and I are d-c components and e_s and i_s are time-varying components with zero average value, then the instantaneous power is

$$p = (E + e_s)(I + i_s) = EI + Ei_s + e_sI + e_si_s \tag{5-28}$$

Since e_s and i_s have zero average value and E and I do not vary with time, the average power is

$$P = EI + (e_s i_s)_{av} \tag{5-29}$$

Thus the signal power and the d-c power can be calculated separately if it is convenient or otherwise desirable to do so. If the current i flows through a resistance R, the instantaneous power absorbed by the resistor is

$$p = R(I + i_s)^2 = R(I^2 + 2Ii_s + i_s^2) \tag{5-30}$$

and the average power absorbed is

$$P = RI^2 + R(i_s^2)_{av} = RI^2 + R(I_{s\ rms})^2 \tag{5-31}$$

where $I_{s\ rms}$ is the effective value of the signal component of current i_s.

Example 5-2. The plate characteristics for the triode in the amplifier of Fig. 5-14 are shown in Fig. 5-13*b*, and the graphical construction for determining the plate current and voltage is superimposed on the characteristics.

a. Determine the average power drawn from the plate supply, the average plate dissipation, and the average power absorbed by the load under quiescent operating conditions.

b. Determine the quantities listed in part *a* under the condition that the square wave of signal voltage shown in Fig. 5-13*a* is applied at the input.

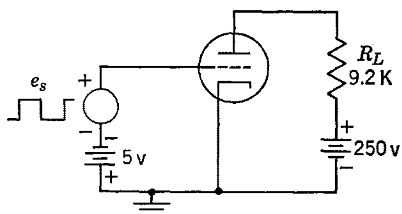


FIG. 5-14. Triode amplifier for Example 5-2.

Solution. a. The quiescent operating point for the amplifier is the point labeled Q in Fig. 5-13*b*. Accordingly, the quiescent average power drawn from the plate supply is

$$P_{bbo} = E_{bb}I_{bo} = (250)(6.5)(10)^{-3} = 1.62 \text{ watts}$$

The quiescent average plate dissipation is

$$P_{bo} = E_{bo}I_{bo} = (190)(6.5)(10)^{-3} = 1.23 \text{ watts}$$

The quiescent average power absorbed by the load is

$$P_{Lo} = P_{bbo} - P_{bo} = 0.39 \text{ watt}$$

b. With the specified signal applied, the plate current is 12 ma during the positive half cycle of e_s and it is 2 ma during the negative half cycle of e_s. Hence the average plate current is

$$I_b = \tfrac{1}{2}(12 + 2) = 7.0 \text{ ma}$$

The average plate current with signal applied is larger than the quiescent plate current because of tube nonlinearities. From Eq. (5-19), the average power drawn from the plate supply with signal applied is

$$P_{bb} = E_{bb}I_b = (250)(7.0)(10)^{-3} = 1.75 \text{ watts}$$

This value is larger than the quiescent power drawn from the plate supply because of the increase in average plate current when the signal is applied.

Because of the fact that the plate current and voltage are square waves, the average plate dissipation with signal applied can be determined in a simple manner. The power dissipated in the tube is constant during each half cycle; during the positive half cycle of e_s it is

$$p_{b1} = (140)(12)(10)^{-3} = 1.68 \text{ watts}$$

and during the negative half cycle it is

$$p_{b2} = (232)(2)(10)^{-3} = 0.464 \text{ watt}$$

The average plate dissipation is the average of these two values,

$$P_b = \tfrac{1}{2}(1.68 + 0.464) = 1.07 \text{ watts}$$

Thus the average plate dissipation with signal applied is less than the quiescent plate dissipation, even though the average plate current increases somewhat when signal is applied.

The average power absorbed by the load with signal applied is

$$P_L = P_{bb} - P_b = 0.68 \text{ watt}$$

Example 5-3. The input signal to the amplifier of Fig. 5-14 is adjusted so that the plate current is $i_b = I_b + i_p = 7 + 5 \cos 2000\pi t$ ma. Determine the average power drawn from the plate supply, the average power absorbed by the load, and the average power dissipated at the plate of the tube.

Solution. The average power drawn from the plate supply is given by Eq. (5-19) as

$$P_{bb} = E_{bb}I_b = (250)(7.0)(10)^{-3} = 1.75 \text{ watts}$$

The average power absorbed by the load is given by Eq. (5-21) as

$$P_L = R_L I_b^2 + R_L(I_p \text{ rms})^2$$

In this case $I_b = 7$ ma, I_p rms $= 5/\sqrt{2}$ ma, and $R_L = 9200$ ohms; thus

$$\begin{aligned} P_L &= (9.2)(49)(10)^{-3} + (9.2)(25/2)(10)^{-3} \\ &= 0.565 \text{ watt} \end{aligned}$$

The average power dissipated in the tube is

$$P_b = P_{bb} - P_L = 1.18 \text{ watts}$$

The maximum and minimum values of plate current in this example have the same values as in Example 5-2. However, the values of the various powers in this example are different from the values of the corresponding powers in Example 5-2 because the waveforms of current and voltage are different.

5-5. Piecewise-linear Analysis. The graphical analysis of triode circuits is likely to prove tedious and time-consuming, especially if there are two or more tubes in the circuit. In many cases the problem of analysis can be greatly simplified by using an approximate representation of the tube that eliminates the need for graphical procedures.

Figure 5-15 shows the input and output characteristics for a typical triode voltage-amplifier tube; superimposed upon these is a set of piecewise-linear characteristics that match the tube characteristics quite well except at small values of plate current or plate voltage. The piecewise-linear grid characteristic is very much like that of the vacuum diode shown in Fig. 3-24*a*. The piecewise-linear plate characteristic consists of a family of straight lines in the first quadrant that are parallel and equally spaced for equal increments of grid voltage.

The piecewise-linear approximation simplifies the analysis of many circuits by making possible the use of algebraic rather than graphical methods. The accuracy of such analyses is quite satisfactory for most needs. In this connection it is pertinent to note that the characteristics of any particular tube may deviate by 10 or 20 per cent from the published

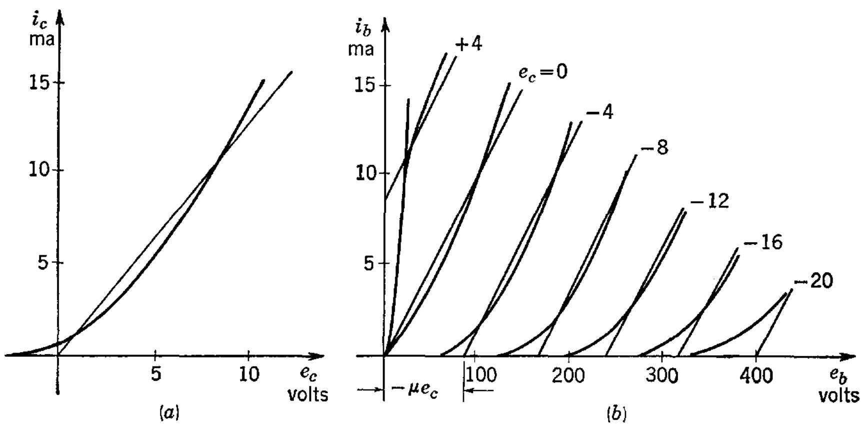


FIG. 5-15. Piecewise-linear approximations for vacuum triodes. (*a*) Grid characteristics; (*b*) plate characteristics.

characteristics for that tube type. Therefore the accuracy of the graphical solution is, in most respects, unlikely to be much better than that of the piecewise-linear solution. The piecewise-linear characteristics, however, give no indication of distortion if the path of operation is restricted to the central portion of the characteristics. In the piecewise-linear representation the nonlinearities of the tube are accounted for only by the breaks in the characteristics.

It is possible to devise a network composed of ideal resistances, diodes, and sources that has a volt-ampere characteristic identical with the piecewise-linear approximate characteristics of the triode. This network is a piecewise-linear model for the triode; it provides additional insight into the behavior of triode amplifiers, and it aids in simplifying the analysis of many triode circuits. The piecewise-linear grid characteristic of Fig. 5-15*a* is like that for the vacuum diode shown in Fig. 3-24*a*. This fact implies that the grid circuit can be represented approximately by an ideal diode in series with an appropriate resistance, as shown in Fig. 5-16*a*.

The piecewise-linear plate characteristic for $e_c = 0$ is also similar to that for the vacuum diode; therefore for $e_c = 0$ the plate circuit can also be represented approximately by an ideal diode in series with a suitable resistance. For other values of e_c the characteristic is shifted to the right or the left, depending on the polarity of e_c, by an amount proportional to e_c. As in the case of the semiconductor diode shown in Fig. 3-28, this shift can be accounted for by adding a voltage source in series with the circuit as shown in Fig. 5-16*b*; the voltage of this source is directly proportional to e_c and of such polarity that the plate characteristic is shifted to the right when e_c is negative.

Since the grid and plate currents both flow out of the tube at the cathode terminal, the circuits of Figs. 5-16*a* and *b* can be combined as

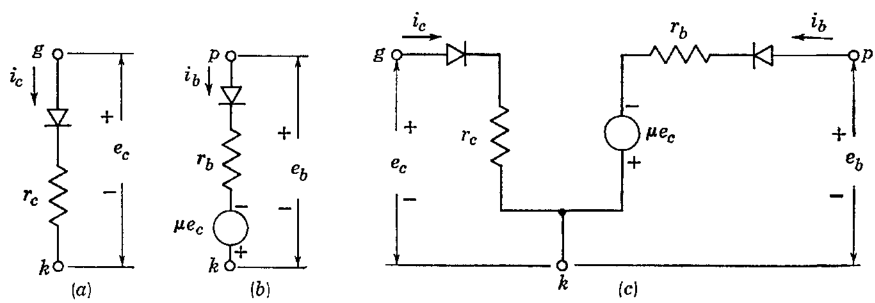


FIG. 5-16. The piecewise-linear model for the vacuum triode, valid for positive e_c provided $e_b > 2e_c$. (*a*) Grid circuit; (*b*) plate circuit; (*c*) combined circuit.

shown in Fig. 5-16*c* to give a complete model for the triode. It is instructive to compare this model with the ideal voltage amplifier shown in Fig. 4-1. When the grid is negative relative to the cathode, the diode in the grid circuit acts as an open circuit; in this respect the model behaves as the ideal amplifier. When the grid is positive relative to the cathode, the grid diode conducts, and the model is quite different from the ideal amplifier. When the plate voltage e_b is greater than $-\mu e_c$, the plate diode conducts, and the model behaves as the ideal amplifier except for the series resistance r_b, which is rarely negligible. The plate resistance is a measure of the amount by which the plate characteristics are rotated from the ideal, vertical position. When e_b is less than $-\mu e_c$, the plate diode is an open circuit, the plate current is cut off, and the model does not behave at all like the ideal amplifier.

As noted in Fig. 5-16, the models in that figure are valid for positive e_c only for values of e_b lying well to the right of the region where the plate characteristics merge into a single diode-like curve. For the characteristics shown in Fig. 5-15*b*, this condition is met if $e_b > 2e_c$.

The parameters μ, r_b, and r_c of the model must be given numerical values that will make the volt-ampere characteristics of the model coincide as closely as possible with those of the triode. One of several possible procedures for determining these values is the following. To determine r_c, choose a suitable straight line on the grid characteristic. The equation of this line is

$$e_c = r_c i_c \tag{5-32}$$

if the line passes through the origin of the coordinates. By substituting values of e_c and i_c for one point on this line (such as the point at which the line crosses the grid characteristic) in (5-32) the value or r_c can be found. It is usually adequate to specify r_c with one or two significant figures.

The horizontal separation between the plate-characteristic curves is usually quite uniform for equal increments of grid voltage except at small values of plate current or plate voltage, a fact which is evident from the characteristics of Fig. 5-15*b*. Thus an average value for the horizontal separation of the curves can be determined easily. The intercepts on the plate-voltage axis μe_c for the piecewise-linear plate characteristics can then be located. These intercepts are, of course, equally spaced for equal increments of e_c. If the piecewise-linear characteristic for a particular grid voltage e_{c1} intersects the plate-voltage axis at the plate voltage e_{b1}, then the amplification factor μ can be calculated from the relation $e_{b1} = -\mu e_{c1}$.

The slope of the family of piecewise-linear plate characteristics is adjusted graphically (with the aid of a straightedge and a triangle) to give the best fit in the region of greatest interest. All curves of the family must have the same slope. The value of the resistance r_b in the model can then be determined from the piecewise-linear characteristic for $e_c = 0$. The equation for this curve,

$$e_b = r_b i_b \tag{5-33}$$

yields the value of r_b when the values of e_b and i_b for one point on the characteristic are inserted.

The nature of the fit obtained by the piecewise-linear approximation can be modified by adjusting the values of μ and r_b; hence, since the procedure outlined above can be completed in a short time, it may be desirable to try several different values for the parameters. It is often possible to adjust the approximation so that the parameter values can be expressed with one or two significant figures.

The parameter values for triodes of various types are spread over a considerable range. For typical small triodes the grid resistance r_c lies in the range between 100 and 1000 ohms, the plate resistance r_b lies in the range between 1000 and 100,000 ohms, and the amplification factor μ lies between 1 and 100. The approximate values of the parameters corresponding to the piecewise-linear characteristics in Fig. 5-15 are $r_c = 0.9$ kilohm, $r_b = 10$ kilohms, and $\mu = 20$.

When the tube is operated in the Class A_1 mode, the grid-circuit diode always acts as an open circuit and the plate-circuit diode always acts as a short circuit. In this case the diodes can be eliminated from the circuit, and the model takes the simple form shown in Fig. 5-17*a*. This model should be compared with the ideal voltage amplifier of Fig. 4-1. At high frequencies the interelectrode capacitances cannot be ignored, and at low

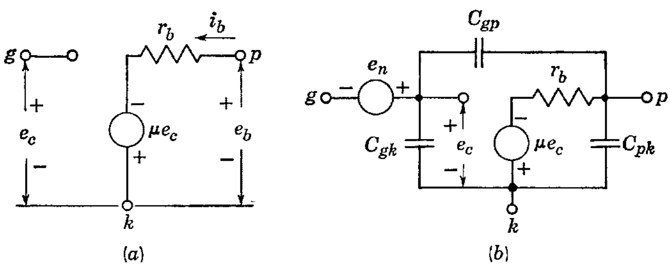


FIG. 5-17. Triode models for Class A_1 operation. (*a*) Basic model; (*b*) model accounting for noise and interelectrode capacitance.

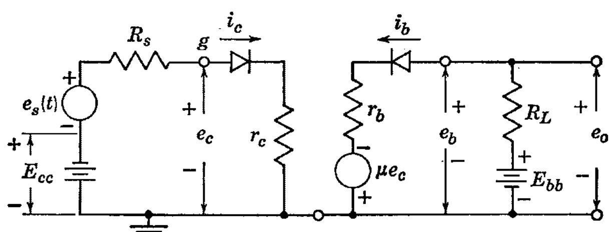


FIG. 5-18. Model for the basic triode amplifier.

signal levels the noise generated by the tube cannot be ignored. The model in Fig. 5-17*b* includes a set of capacitances and a noise-voltage source to account for these effects. It is customary to omit these elements from the model when they have negligible effects on the performance of the circuit.

When the tube in the basic amplifier circuit of Fig. 5-8 is replaced by its piecewise-linear model, the result is the circuit shown in Fig. 5-18. This representation is appropriate when the signal lies in a range of frequencies where the parasitic capacitances are negligible, such as the band of audio frequencies, for example. The analysis of this circuit when the parameters and the applied voltages are known is perfectly straightforward; when the mode of operation is Class A_1, so that the grid and plate diodes act respectively as an open circuit and a short circuit, the analysis can be done by inspection.

In general, the grid voltage is given by

$$e_c = E_{cc} + e_s - R_s i_c \tag{5-34}$$

When the sum $E_{cc} + e_s$ is negative, however, the grid diode acts as an open circuit, and i_c is zero. For this condition

$$e_c = E_{cc} + e_s \tag{5-35}$$

When the sum $E_{cc} + e_s$ is positive, the grid diode acts as a short circuit, and

$$e_c = r_c i_c = r_c \frac{E_{cc} + e_s}{R_s + r_c} \tag{5-36}$$

Having found e_c from (5-36) or (5-35), whichever is appropriate, the plate circuit can be analyzed. The plate voltage is given by

$$e_b = E_{bb} - R_L i_b \tag{5-37}$$

This is also the output voltage e_o. When $-\mu e_c$ is greater than E_{bb}, a reversed voltage is impressed across the plate diode, and it acts as an open circuit. Thus the plate current is cut off and $i_b = 0$. For this condition

$$e_b = E_{bb} \tag{5-38}$$

When the grid is not sufficiently negative to cut the tube off, the plate current is

$$i_b = \frac{E_{bb} + \mu e_c}{R_L + r_b} \tag{5-39}$$

The output voltage can be found by substituting (5-39) into (5-37); it is

$$e_o = e_b = E_{bb} - R_L \frac{E_{bb} + \mu e_c}{R_L + r_b} \tag{5-40}$$

Further understanding of the way in which the piecewise-linear model approximates the triode amplifier can be obtained by comparing the voltage transfer characteristic of the model with that of the amplifier determined by graphical analysis. A typical transfer characteristic obtained by graphical analysis is shown in Fig. 5-12; a typical transfer characteristic for the piecewise-linear model of Fig. 5-18 is shown in Fig. 5-19. The circuit is linear except for the diodes, which are either open or short circuits; therefore the characteristic consists of a set of straight-line segments. The breaks in the characteristic occur where one or the other of the diodes changes from a short circuit to an open circuit. It is a simple matter to calculate the coordinates of the break points; the characteristic can then be completed by adding appropriate straight lines.

The break on the left is the point of plate-current cutoff. At this point the plate diode changes from a short circuit to an open circuit as e_s becomes more negative; hence at this point there is no current through and no voltage across the plate diode. It is clear from the circuit in Fig. 5-18 that with no current through the plate diode, the output voltage is $e_o = E_{bb}$. It also follows from Fig. 5-18 that with no voltage

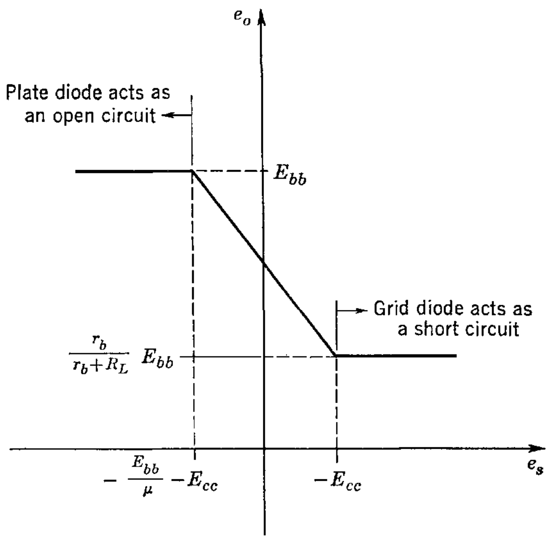


FIG. 5-19. Voltage transfer characteristic for the piecewise-linear model of a triode amplifier.

across the plate diode and $i_b = 0$, $-\mu e_c = E_{bb}$. Since E_{bb} is positive, e_c is negative at this break, and the grid diode acts as an open circuit. Under these conditions $e_c = e_s + E_{cc}$, and the value of e_s at the break is

$$e_s = -\frac{E_{bb}}{\mu} - E_{cc} \tag{5-41}$$

At all values of e_s lying to the left of this point the plate diode acts as an open circuit, and e_o is constant at the value E_{bb}.

The break on the right is the point at which grid current begins to flow. At this point the grid diode changes from an open circuit to a short circuit as e_s becomes more positive; hence at this point there is no current through and no voltage across the grid diode. It is clear from the circuit that under these conditions $e_c = 0$ and $e_c = e_s + E_{cc}$; hence $e_s = -E_{cc}$ at this break point. It also follows from the circuit that when $e_c = 0$, $\mu e_c = 0$, and

$$e_o = \frac{r_b}{r_b + R_L} E_{bb} \tag{5-42}$$

Both break points are now located, and they can be joined by a straight line. For values of e_s greater than $-E_{cc}$, grid current flows. In this region e_c is given by Eq. (5-36); when e_c has been found, e_o is given by (5-40). The right-hand segment of the characteristic can be completed by calculating one point for any convenient e_s greater than $-E_{cc}$. It follows from Fig. 5-18 that if R_s is much greater than r_c, e_c is essentially zero for all reasonable values of e_s greater than $-E_{cc}$. In this case the right-hand segment of the characteristic is essentially horizontal.

Example 5-4. The triode amplifier shown in Fig. 5-20*a* can be represented approximately by the piecewise-linear model shown in Fig. 5-20*b*. A voltage transfer characteristic for the amplifier, determined by graphical analysis, is shown in Fig. 5-12. Determine the voltage transfer characteristic for the model of Fig. 5-20*b*, and compare it with the results obtained by graphical analysis of the amplifier.

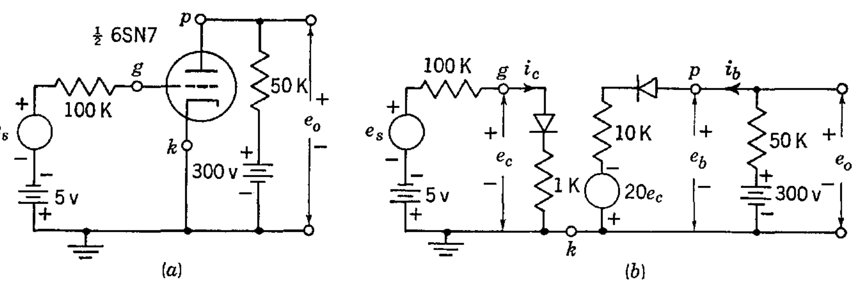


FIG. 5-20. Triode amplifier for Example 5-4. (*a*) Circuit; (*b*) piecewise-linear model.

Solution. The desired characteristic has the general form shown in Fig. 5-19. Hence the first step in the analysis is to determine the coordinates of the break points in the characteristic. At the point of plate-current cutoff, $\mu e_c = -E_{bb}$; thus

$$e_c = -{}^{300}\!/_{20} = -15 \text{ volts}$$

and $$e_s = e_c - E_{cc} = -15 + 5 = -10 \text{ volts}$$

Also, at plate-current cutoff, $i_b = 0$, and

$$e_b = E_{bb} = 300 \text{ volts}$$

Thus the coordinates of the left-hand break are $e_s = -10$ volts, $e_o = 300$ volts.

At the point where grid current begins to flow, $e_c = 0$, $i_c = 0$, and

$$e_s = -E_{cc} = 5 \text{ volts}$$

Also, at this point $\mu e_c = 0$, and

$$e_b = \frac{r_b}{r_b + R_L} E_{bb} = \frac{10}{10 + 50} (300) = 50 \text{ volts}$$

Thus the coordinates of the right-hand break point are $e_s = 5$ volts, $e_o = 50$ volts.

For an additional point in the positive grid region, let $e_s = 10$ volts. Then, from Eq. (5-36),

$$e_c = r_c \frac{E_{cc} + e_s}{R_s + r_c} = (1)\frac{10 - 5}{100 + 1} = \frac{5}{101} \approx 0.05 \text{ volt}$$

Under this condition, $\mu e_c = 1.0$ volt, and from Eq. (5-40),

$$e_b = E_{bb} - R_L \frac{E_{bb} + \mu e_c}{r_b + R_L} = 300 - (50)^{301}\!/\!_{60} \approx 50 \text{ volts}$$

The coordinates of this point in the positive grid region are $e_s = 10$ volts, $e_o = 50$ volts.

The piecewise-linear characteristic corresponding to the points determined above is shown in Fig. 5-21. Superimposed upon this characteristic for comparison is the graphically determined characteristic taken from Fig. 5-12. The correspondence between these two curves is good except in the vicinity of plate-current cutoff.

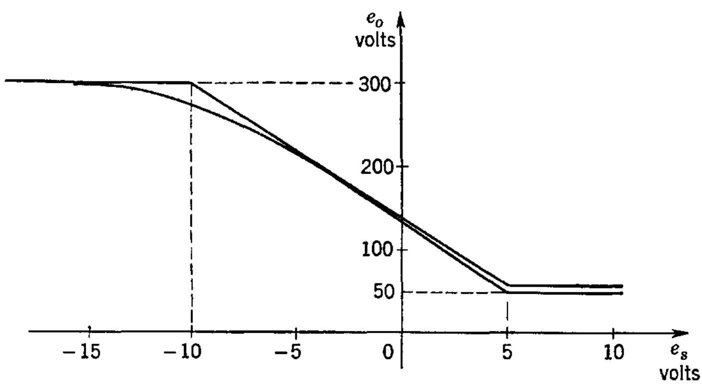


FIG. 5-21. Piecewise-linear and graphical voltage transfer characteristics.

Example 5-5. The circuit shown in Fig. 5-22*a* consists of a full-wave rectifier and an electronic voltage regulator. The problem is to analyze the behavior of the regulator. Any change in output voltage is amplified by triode T_1 and causes the voltage drop across T_2 to change in such a way as to compensate almost entirely for the initial change in E_2. The semiconductor diode in the cathode circuit of T_1 is operated under conditions of avalanche breakdown; thus it provides a constant voltage drop that is essentially independent of the current it carries (Sec. 3-5). Under normal conditions both of the triodes operate in the Class A_1 mode.

A model for the regulator circuit is shown in Fig. 5-22*b* for the condition that both triodes are in Class A_1. The constant voltage drop across the diode is represented by the battery E_o. The parameter values for the circuit are $E_o = 50$ volts, $\mu_1 = 100$, $r_{b1} = 70$ kilohms, $\mu_2 = 4$, $r_{b2} = 0.8$ kilohm, $R_1 = 250$ kilohms, $R_2 = 50$ kilohms, and $R_3 = 200$ kilohms.

a. Determine the output voltage from the regulator when the input voltage is 400 volts, d-c, and the load current is 100 ma.

b. If the input voltage rises to 450 volts when the load current is reduced to zero, what is the no-load output voltage?

c. Making the reasonable assumption that the output voltage from the rectifier, E_1, varies linearly with I_2, give Thevenin equivalent circuits for the rectifier alone and for the rectifier with the voltage regulator.

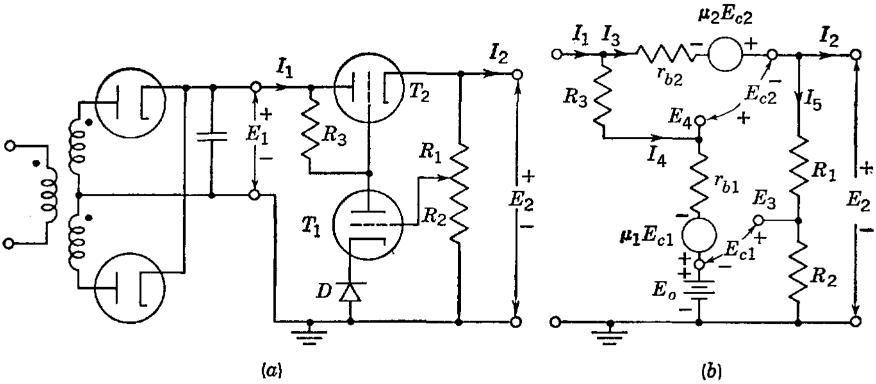


FIG. 5-22. Electronic voltage regulator for Example 5-5. (*a*) Circuit; (*b*) model.

Solution. *a.* The output voltage is given by

$$E_2 = E_1 - I_3 r_{b2} + \mu_2 E_{c2}$$

But $E_{c2} = E_4 - E_2$; hence

$$E_2 = E_1 - I_3 r_{b2} + \mu_2 E_4 - \mu_2 E_2$$

The voltage E_4 is given by

$$E_4 = E_1 - I_4 R_3 = E_1 - R_3 \frac{E_1 + \mu_1 E_{c1} - E_o}{r_{b1} + R_3}$$

and $E_{c1} = E_3 - E_o = E_2/6 - E_o$. Using this last relation and substituting numerical values in the expression for E_4 yields

$$E_4 = 0.259E_1 - 12.35E_2 + 3737$$

Substituting this relation for E_4 into the equation for E_2 yields

$$E_2 = E_1 - I_3 r_{b2} + 1.036E_1 - 53.4E_2 + 14{,}948$$
$$= \frac{2.036E_1 - 800I_3 + 14{,}948}{54.4}$$

when I_3 is in amperes. The terms $2.036E_1$ amd $800I_3$ are small compared to 14,948; hence they have a relatively small effect on the value of E_2 as long as the tubes remain in Class A_1 operation. In addition, I_5 is approximately 1 ma; therefore $I_3 = I_2 + I_5 \approx I_2 = 100$ ma. Thus for $I_2 = 100$ ma and $E_1 = 400$ volts,

$$E_2 = \frac{814.4 - 80 + 14{,}948}{54.4} = 288.3 \text{ volts}$$

b. With $I_2 = 0$, I_3 can be taken as zero. Thus with $I_2 = 0$ and $E_1 = 450$ volts,

$$E_2 = \frac{913.5 + 14{,}948}{54.4} = 291.4 \text{ volts}$$

c. For the rectifier alone, the no-load voltage is 450 volts; this is the voltage of the source in the Thevenin equivalent circuit. With a load current of 100 ma, the terminal voltage is 400 volts; hence the internal resistance is

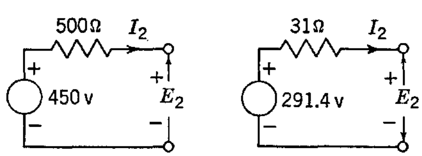
$$R = 50/0.1 = 500 \text{ ohms}$$

The corresponding equivalent circuit is shown in Fig. 5-23*a*.

For the rectifier with the voltage regulator, the no-load voltage is 291.4 volts; this is the voltage of the source in the Thevenin equivalent circuit. With a load current of 100 ma, the terminal voltage is 288.3 volts; hence the internal resistance is

$$R' = 3.1/0.1 = 31 \text{ ohms}$$

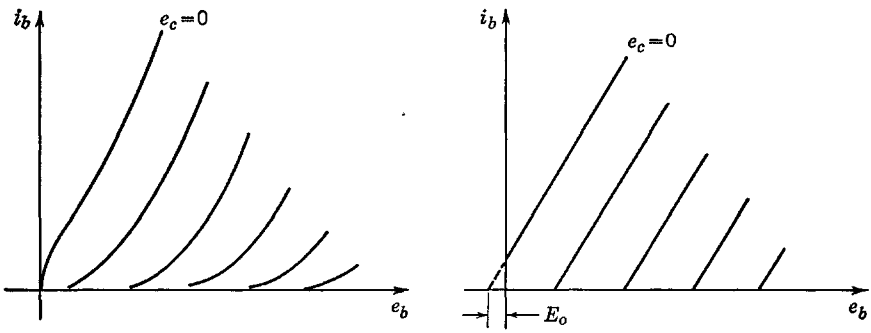
The corresponding equivalent circuit is shown in Fig. 5-23*b*.



(*a*) (*b*)

FIG. 5-23. Thevenin equivalent circuits for the regulated power supply of Fig. 5-22. (*a*) For the rectifier alone; (*b*) for the rectifier with the voltage regulator.

The piecewise-linear triode model of Fig. 5-16*c* gives a good approximation for many types of commonly used triodes. However, certain other triodes have plate characteristics of a slightly different shape so that this model does not give a good approximation. For example,



(*a*) (*b*)

FIG. 5-24. An alternative piecewise-linear approximation for vacuum triodes. (*a*) Characteristics; (*b*) approximation.

certain high-μ triodes have characteristics of the general shape shown in Fig. 5-24*a*. A reasonable piecewise-linear approximation to these characteristics is shown in Fig. 5-24*b*. The important feature of this approximation is that the $e_c = 0$ characteristic does not pass through the origin; when extended, it intersects the plate-voltage axis at a negative voltage, $e_b = -E_o$. Such a characteristic cannot be obtained with the model of Fig. 5-16*c*, for its $e_c = 0$ characteristic must pass through the origin.

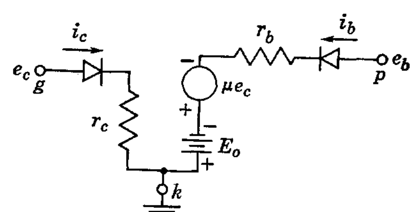


FIG. 5-25. An alternative triode model.

The model shown in Fig. 5-25 gives the characteristics of Fig. 5-24*b* when the parameters are given suitable numerical values. The source E_o shifts the entire family of characteristics to the left by E_o volts, as can be seen by considering the value of e_b that will just produce plate-current cutoff for each value of e_c. This model is more general than the one given in Fig. 5-16*c*, although the latter is quite satisfactory in many applications. In every case it is important to remember that the model is a suitable representation of the triode for positive grid voltages only if the plate voltage is greater than the grid voltage by a sufficient amount. It is also true, however, that there are not many triode applications in which this condition is violated in the normal operation of the circuit.

REFERENCES

1. Gray, T. S.: "Applied Electronics," 2d ed., John Wiley & Sons, Inc., New York, 1954.
2. Seely, S.: "Radio Electronics," McGraw-Hill Book Company, Inc., New York, 1956.
3. Arguimbau, L. B.: "Vacuum Tube Circuits," John Wiley & Sons, Inc., New York, 1948.

PROBLEMS

5-1. One section of a 12AT7 twin triode is used in the basic amplifier circuit shown in Fig. 5-8. The plate supply voltage is $E_{bb} = 300$ volts, and the resistance in the grid circuit is $R_s = 100$ kilohms.

a. What values of R_L and E_{cc} are needed to place the quiescent operating point at $I_{bo} = 5$ ma and $E_{bo} = 200$ volts?

b. Calculate the quiescent plate dissipation and compare it with the maximum permissible value specified by the manufacturer.

5-2. One section of a 12AX7 twin triode is used in the basic amplifier circuit shown in Fig. 5-8. The applied voltages and the circuit parameters are $E_{bb} = 300$ volts, $R_L = 200$ kilohms, $R_s = 100$ kilohms, and $E_{cc} = -1.5$ volts.

a. Trace a set of plate characteristics for the tube. On these curves construct the hyperbola of constant plate dissipation corresponding to the maximum permissible value of plate dissipation.

b. Locate the quiescent operating point for the tube. Is the quiescent plate dissipation less than the maximum permissible value?

5-3. The signal applied to the amplifier of Prob. 5-2 has the waveform shown in Fig. 5-26. The problem is to study the response of the amplifier to this signal.

a. Trace a set of plate characteristics for the tube, and construct the load line for the plate circuit.

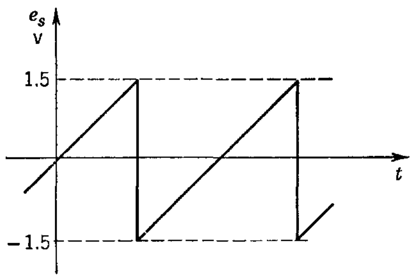


FIG. 5-26. Signal waveform for Prob. 5-3.

b. Construct the voltage transfer characteristic (Fig. 5-12) for e_s between -5 and $+1.5$ volts.

c. Using the curve of part *b*, plot the waveform of e_o when the signal of Fig. 5-26 is applied to the amplifier.

5-4. One section of a 12AT7 twin triode is used in the basic amplifier circuit shown in Fig. 5-8. The applied voltages and the circuit parameters are $E_{bb} = 300$ volts,

$E_{cc} = -2$ volts, $R_L = 15$ kilohms, and $R_s = 100$ kilohms. Calculate the change in output voltage e_o that results from a change in e_s from 0 to 1 volt. By what factor is the change in input voltage amplified?

5-5. One section of a 12AX7 twin triode is used in the basic amplifier circuit of Fig. 5-8. The plate supply voltage is $E_{bb} = 600$ volts, and the load resistance is $R_L = 800$ kilohms. In addition, another 800-kilohm resistor is connected across the output terminals from plate to ground.

a. Give a circuit diagram in which the circuit connected between plate and cathode of the tube is replaced by its Thevenin equivalent.

b. Construct the load line for the tube on its plate-characteristic curves.

c. What value of the grid bias is required to make the quiescent plate voltage be 100 volts?

d. If e_s increases from 0 to 0.5 volt, what is the change in output voltage, e_o? Give the sign as well as the magnitude of the change.

5-6. One section of a 12AX7 is used in the circuit of Fig. 5-8 with $E_{bb} = 300$ volts, $E_{cc} = -1.5$ volts, $R_L = 250$ kilohms, and $R_s = 0$.

a. Find the quiescent power dissipated at the plate of the tube.

b. Sketch and dimension the waveform of e_b versus t when the signal shown in Fig. 5-27 is applied at the input.

c. Calculate the average power dissipated at the plate under the conditions of part *b*. Compare this value with the quiescent dissipation found in part *a*.

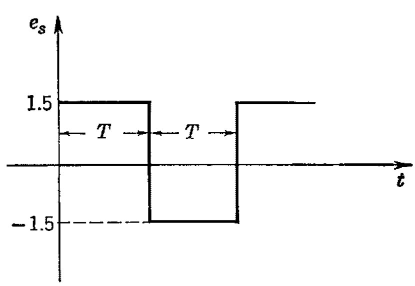


FIG. 5-27. Signal waveform for Prob. 5-6.

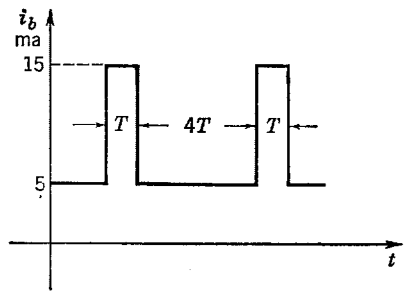


FIG. 5-28. Plate-current waveform for Prob. 5-7.

5-7. A 6J5 triode is used in the basic amplifier circuit with $E_{bb} = 250$ volts and $R_L = 10$ kilohms. The grid bias and the input signal are adjusted so that the plate current has the waveform shown in Fig. 5-28. Find the power delivered by the battery, the power absorbed by R_L, and the power dissipated in the tube. What is the rms value of i_b?

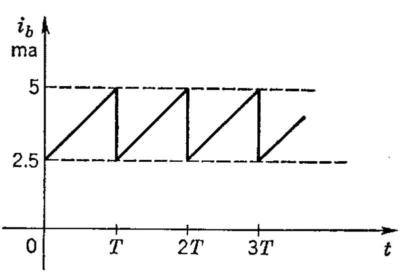


FIG. 5-29. Plate-current waveform for Prob. 5-8.

5-8. One section of a 6SN7 twin triode is used in the circuit of Fig. 5-8 with $E_{bb} = 300$ volts and $R_L = 50$ kilohms. The grid bias and the input signal are adjusted so that the plate current has the waveform shown in Fig. 5-29. The objective of the problem is to study the power relations in the circuit under nonsinusoidal operating conditions.

a. Prove that the average power delivered by the plate-circuit power supply is $P_{bb} = E_{bb}(i_b)_{av}$.

b. Prove that the average power absorbed by R_L is $P_L = R_L(I_{b\,\text{rms}})^2$, where

$$I_{b\,\text{rms}} = \left[\frac{1}{T}\int_0^T i_b^2\,dt\right]^{1/2}$$

c. In the interval $0 < t < T$, i_b varies linearly with time. Hence in this interval i_b is related to t by an equation of the form $i_b = A + Bt$, where A and B are constants. Find the values of A and B in terms of T.

d. Find the average power absorbed by R_L.

5-9. A 6J5 triode is to be represented by the piecewise-linear model shown in Fig. 5-16*c*. The problem is to determine values for the parameters in the model that will give a suitable approximation.

a. When $e_c = 5$ volts, the grid current is 5 ma. Find the value of r_c that will make the model draw the same grid current as the tube when $e_c = 5$ volts.

b. Find the values of μ and r_b that will make the model draw the same plate current as the tube at two points on the plate characteristic: $e_c = 0$, $e_b = 80$ volts and $e_c = -16$ volts, $e_b = 360$ volts.

c. Trace the plate-characteristic curves for the tube with $e_c = 0, -4, -8, -12$, and -16 volts. Superimpose on these characteristics the characteristics for the model with the same set of grid voltages.

5-10. The 2A3 triode is to be represented by a piecewise-linear model of the form shown in Fig. 5-25 in the region of negative grid voltages. Determine values of μ, r_b, and E_o that will give a good approximation. Sketch the model showing the polarities of the voltage sources.

5-11. The piecewise-linear model for a certain amplifier has the form shown in Fig. 5-18. The plate supply voltage is $E_{bb} = 300$ volts, and the circuit parameters are $R_L = 100$ kilohms, $r_b = 10$ kilohms, $r_c = 1$ kilohm, $\mu = 20$, and $R_s = 10$ kilohms. The input voltage is $e_s = E_s \cos \omega t$. The problem is to adjust E_s and E_{cc} to give the largest possible sinusoidal component of voltage at e_o with operation in the Class A_1 mode.

a. What are the required values of E_{cc} and E_s?

b. With the adjustment of part *a*, what is the amplitude of the sinusoidal component of e_o? What is the quiescent plate dissipation?

5-12. A piecewise-linear circuit is shown in Fig. 5-30. For this circuit construct the input volt-ampere characteristic, E_1 versus I_1, and the family of output volt-ampere characteristics, E_2 versus I_2, for E_1 held constant at 1, 0, and -1 volt. Show the coordinates of all break points and the slopes of all lines.

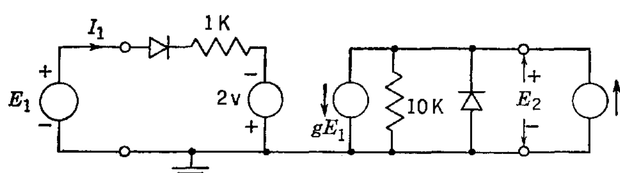


FIG. 5-30. Piecewise-linear circuit for Prob. 5-12. (*Note:* $gE_1 = 5E_1$ ma.)

5-13. The cascode amplifier shown in Fig. 5-31 is used in certain applications as a voltage amplifier. (For example, it is occasionally used in the voltage regulator of Fig. 5-22*a* in place of the one-tube amplifier consisting of T_1 and R_3 to provide more amplification and better regulation.) The grid current in each tube is zero, and the plate circuits can be represented by piecewise-linear models with $\mu = 20$ and $r_b = 10$ kilohms.

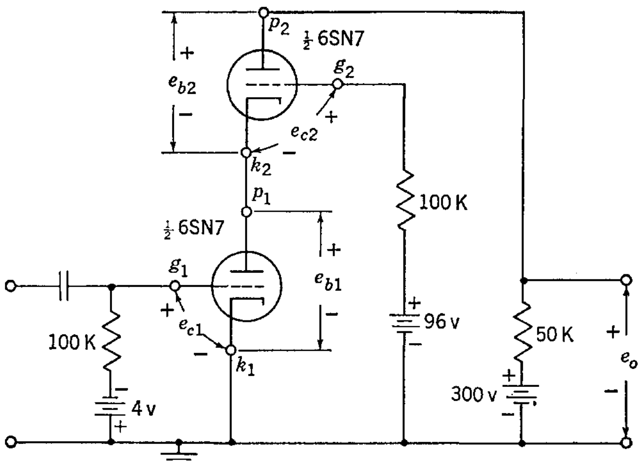


FIG. 5-31. Cascode amplifier for Prob. 5-13.

a. With the aid of the model, determine the quiescent values of i_b, e_{b1}, e_{b2}, and e_o. Note that $e_{c2} = 96 - e_{b1}$ volts.

b. If e_{c1} is given an increment of 0.1 volt, what is the resulting increment in e_o? By what factor is the increment in input voltage amplified?

c. Sketch a diagram showing how the grid voltage for the upper tube can be derived from the plate-supply battery with the aid of a voltage divider. Give the values of resistance required in the voltage divider if the total resistance is 100 kilohms.

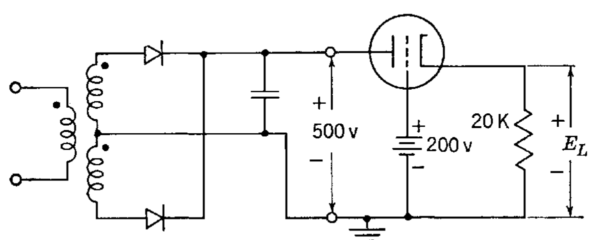


FIG. 5-32. Regulated power supply for Prob. 5-14.

5-14. The circuit of Fig. 5-32 is a full-wave rectifier with a triode used as an elementary voltage regulator. The function of the regulator is, among other things, to hold the load voltage nearly constant in the face of changes in line voltage and to reduce the ripple voltage across the load. As an additional feature, the output voltage can be adjusted over a range of values by adjusting the voltage applied to the grid of the triode. (The performance of this elementary regulator should be compared with that of the more elaborate regulator analyzed in Example 5-5.) The triode can be represented by a piecewise-linear model of the form shown in Fig. 5-16*c* with $\mu = 60$, $r_b = 12$ kilohms, and $r_c = 500$ ohms.

a. Determine the load voltage E_L. *Suggestion:* For the first trial, assume that the grid current is zero, since this is the case in normal operation. Then, when the solution is completed, check the assumption.

b. If the output voltage from the rectifier decreases by 10 per cent, by what percentage does the load voltage change?

c. If the peak-to-peak value of the ripple in the rectifier output is 1 volt, what is the peak-to-peak value of the ripple voltage across the load? The results of the previous parts can be used in this calculation if desired.

5-15. Figure 5-33 shows a circuit that is occasionally used as a voltage amplifier and in which grid bias is developed by grid rectification. Specifically, the grid circuit

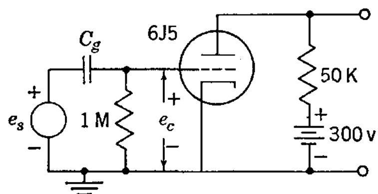


FIG. 5-33. A triode amplifier with grid-leak bias for Prob. 5-15.

acts as a diode clamper. The tube can be represented by a piecewise-linear model of the form shown in Fig. 5-16*c* with $\mu = 20$, $r_b = 10$ kilohms, and $r_c = 0$ (for this application). The input signal is $e_s = 2 \cos \omega_s t$, and C_g is so large that it cannot discharge appreciably through the 1-megohm grid leak in one cycle of the signal.

a. What is the value of the direct voltage across the capacitor?

b. Sketch and dimension the waveform of the grid voltage e_c.

c. What is the value of the grid bias (d-c component of e_c)?

d. A short burst of noise (such as static) makes $e_s = 30$ volts for a brief period. Describe the action of the circuit after the noise has passed. This phenomenon is called grid blocking.

CHAPTER 6

PRACTICAL TRIODE AMPLIFIERS

The basic triode amplifier of Chap. 5 is a perfectly good amplifier; however, certain practical considerations make it desirable to modify the circuit in various ways for different specific applications. For example, it is usually desirable to remove any d-c component in the input signal before applying it to the tube, for the d-c component usually contributes nothing useful to the signal. Another consideration is the fact that it is usually desirable for all the voltage sources used in an amplifier circuit to have one terminal in common; this is not the case with the signal source and the grid-bias source in the basic amplifier circuit of Chap. 5. It is shown below that this latter difficulty can be overcome by a simple circuit modification that provides bias in a convenient and inexpensive way. As still another consideration, the load to which an amplifier is required to deliver power, such as a loudspeaker or a broadcasting antenna, often has a volt-ampere characteristic (impedance, for sinusoidal operation) that does not match the output characteristic of the tube in a suitable manner. In such cases the load may be coupled to the tube through an impedance-matching network; in audio amplifiers an iron-core transformer is commonly used to match the low impedance of the loudspeaker to the tube characteristics.

The first objective of this chapter is to extend the graphical and piecewise-linear methods of analysis introduced in Chap. 5 to more elaborate circuits and to generalize them somewhat. Since the circuits considered contain reactive elements C and L, a second objective is to extend the methods of analysis to include the effects of reactive elements on currents and voltages containing both d-c and time-varying components. The third objective is to examine some of the important properties of certain widely used amplifier circuits.

6-1. Cathode-resistor Bias. Figure 6-1 shows a triode amplifier in which grid bias is provided by the voltage drop across a resistor in the cathode lead. Since conventional current can only flow out of the tube at the cathode, the cathode is always positive with respect to ground. Hence with no signal applied the grid is negative relative to the cathode, negligible grid current flows, there is negligible voltage drop across R_g, and the grid-to-cathode voltage is

$$E_c = -R_k I_b \tag{6-1}$$

If the quiescent operating point is specified, both E_c and I_b are known, and the value of R_k required to give this operating point can be calculated from (6-1).

The signal voltage is usually applied to such amplifiers through a series capacitor, as shown in Fig. 6-1, to remove any d-c component that the signal may contain. Such d-c components are generally of no interest, and they may shift the operating point to an unsuitable location on the plate characteristic. For the present it is assumed that the capacitor is chosen to act as a short circuit to the time-varying component of e_s.

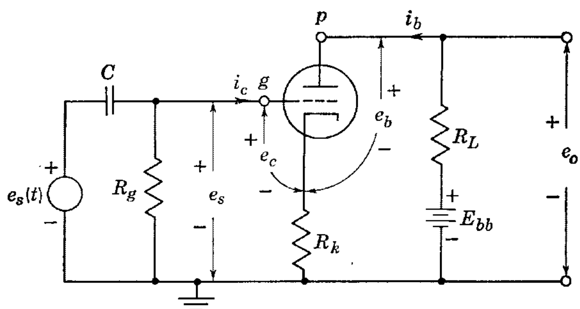


FIG. 6-1. A triode amplifier with cathode-resistor bias.

Under these circumstances the analysis is not altered by assuming that e_s contains no d-c component.

The grid return resistance R_g is necessitated by the fact that in any practical tube there is some flow of charge to the grid, no matter how negative the grid may be. This charge consists in part of high-energy electrons and in part of positive ions created from molecules of residual gas in the tube. If the grid return resistor is omitted, there can be no net flow of charge out of the grid terminal, and the grid assumes a potential such that the flow of positive charge to the grid inside the tube equals the flow of negative charge to the grid. This grid potential usually does not give a satisfactory quiescent operating point. When the grid return resistance is added, a small grid current, usually a fraction of a microampere, flows, and if R_g is not too large, the grid is held at approximately ground potential. For small voltage-amplifier tubes it is usually recommended that R_g not exceed 1 megohm; for larger power-amplifier tubes a maximum value of 50 kilohms may be specified. When proper precautions are taken, however, voltage amplifiers can be operated with values of R_g far in excess of 1 megohm; in some special applications they are operated without any return path at all.

When a time-varying signal to which C acts as a short circuit is applied to the circuit in Fig. 6-1, the signal voltage appears across R_g. If i_c is negligible in comparison with i_b, as normally is the case, the potential of the grid relative to the cathode is

$$e_c = e_s - R_k i_b \tag{6-2}$$

A positive increase in e_s gives a positive increase in e_c, which in turn gives a positive increase in i_b. The increase in i_b results in an increase in the voltage drop across R_k which, as indicated by (6-2), subtracts from the change in e_s. Hence the increase in e_c is not as great as the increase in e_s, and the tube behaves as if the input signal were smaller than e_s. The result is a smaller change in the output voltage and a smaller voltage

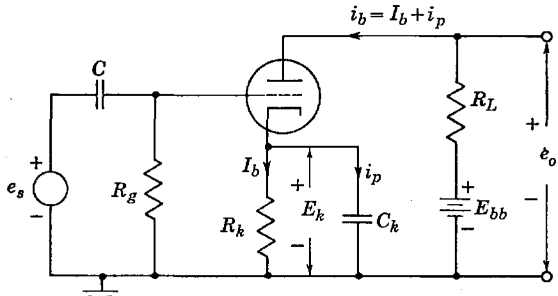


FIG. 6-2. A triode amplifier with a bypassed cathode resistor.

amplification. The cathode resistor provides bias, but it also provides degeneration that reduces the voltage amplification.

The degeneration introduced by the cathode resistor can be removed in so far as the signal is concerned by connecting in parallel with R_k a capacitor that acts as a short circuit to the time-varying component of i_b. This bypass capacitor is shown in Fig. 6-2. In the steady state, the average value of the current through the capacitor must be zero; otherwise the charge on the capacitor and the voltage across the capacitor would increase indefinitely with time. Also, since C_k acts as a short circuit to the time-varying component of i_b, there is no time-varying current through R_k. Hence the current in R_k is I_b, and the current in C_k is i_p. The voltage drop across the parallel combination is $E_k = R_k I_b$; if there is negligible waveform distortion, I_b is independent of the signal amplitude, and $E_k = R_k I_{bo}$.

With the bypass capacitor C_k added to the circuit, its behavior for the d-c components of current and voltage is different from its behavior for the time-varying components. In so far as the d-c components are concerned, the circuit is that shown in Fig. 6-3*a*. For the total currents and voltages (d-c plus the time-varying component) the circuit behaves as the one shown in Fig. 6-3*b*. The battery E_k accounts for the voltage drop across C_k; its voltage is $E_k = R_k I_b$. The circuit of Fig. 6-3*a* can be solved graphically to find the quiescent point, and the circuit of Fig. 6-3*b* can be solved graphically to find the dynamic path of operation and the output voltage.

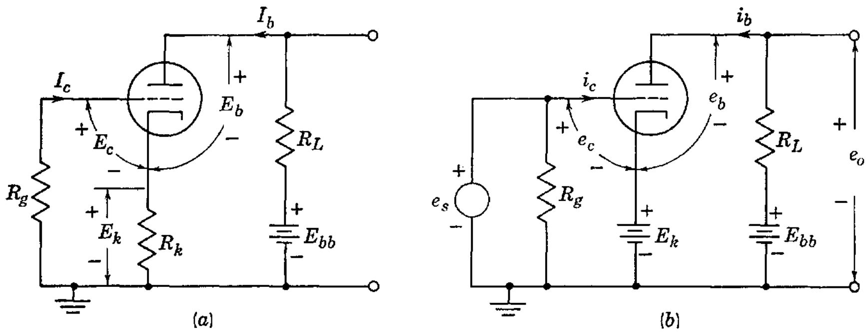


FIG. 6-3. Representation of an amplifier with a bypassed cathode resistor. (*a*) Circuit for d-c components; (*b*) circuit for total currents and voltages.

The graphical construction that gives the quiescent point is shown in Fig. 6-4*a*. The relations that serve as the basis for the construction are obtained by inspection of the circuit in Fig. 6-3*a*:

$$I_c = 0 \qquad (6\text{-}3)$$

$$E_c = -R_k I_b = -E_k \qquad (6\text{-}4)$$

and

$$E_b = E_{bb} - (R_L + R_k)I_b \qquad (6\text{-}5)$$

Equation (6-5) is the equation of the static load line; its intercept with the plate-current axis is at $I_o = E_{bb}/(R_L + R_k)$. Equation (6-4) gives the grid-bias line. This curve is constructed by assuming various convenient values of E_c and computing the corresponding value of I_b from (6-4). These points are plotted on the plate characteristic using the usual scale for plate current and using the grid-voltage scale marked on the individual plate characteristics. The resulting curve is, in most cases, almost a straight line. All points that satisfy (6-4) lie on the grid-bias line, and all points that satisfy (6-5) lie on the static load line. The intersection of these curves satisfies both equations simultaneously and is the quiescent point.

The graphical construction for the operating path with signal applied is shown in Fig. 6-4*b*. The equation of this dynamic operating path is written by inspection of the circuit in Fig. 6-3*b*. It is

$$e_b = E_{bb} - E_k - R_L i_b \qquad (6\text{-}6)$$

By defining $E'_{bb} = E_{bb} - E_k$, (6-6) can be written as

$$e_b = E'_{bb} - R_L i_b \tag{6-7}$$

Thus the dynamic operating path intersects the plate-voltage axis at E'_{bb}, and it intersects the plate-current axis at $I'_0 = E'_{bb}/R_L$. All combinations of e_b and i_b that satisfy Eq. (6-6) lie on this line. Thus, as the signal goes through its sequence of values, the grid voltage, $e_c = e_s - E_k$, goes through a similar sequence of values and the operating point moves along the dynamic operating path specified by Eq. (6-7).

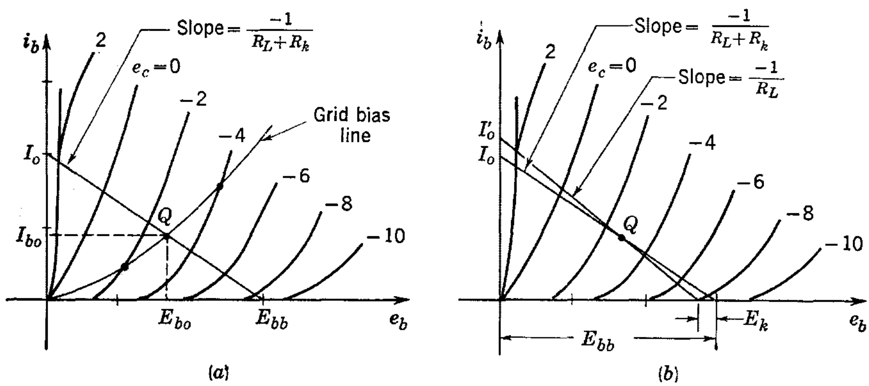


FIG. 6-4. Graphical analysis of an amplifier with a bypassed cathode resistor. (a) D-C components; (b) total currents and voltages.

At any instant when $i_b = I_b$, the plate voltage given by (6-6) is

$$e_b = E_{bb} - E_k - R_L I_b = E_{bb} - (R_L + R_k)I_b \tag{6-8}$$

Substituting (6-5) in (6-8) yields

$$e_b = E_b \tag{6-9}$$

for the value of e_b at any instant when $i_b = I_b$. When there is negligible waveform distortion in the amplifier, $I_b = I_{bo}$ and $E_b = E_{bo}$; hence under this condition the dynamic operating path passes through the quiescent operating point as indicated in Fig. 6-4*b*. Thus when there is negligible distortion, the dynamic operating path can be constructed by passing a line with slope $-1/R_L$ through the quiescent point. This is usually the easiest way to construct the operating path, but it does not give as much insight as the more detailed construction shown in Fig. 6-4*b*. It is significant to note that the slope of the dynamic operating path is greater than that of the static load line. Since the bypass capacitor acts as a short circuit to the time-varying component of i_b, the plate-circuit resistance is less for the time-varying component than for the d-c component.

The foregoing analysis shows that if the average value of the plate current changes as a result of signal distortion, the operating path shifts on the plate characteristic. For example, if $R_k = 1000$ ohms, and if I_b increases by 1 ma as the signal level is increased, then E_k will increase by 1 volt, and the operating path will shift to the left by 1 volt measured on the plate-voltage axis. This shift is ordinarily negligible in Class A_1 operation.

As a useful alternative to the graphical analysis, the circuit of Fig. 6-2 can be analyzed with the aid of the piecewise-linear model for the tube.

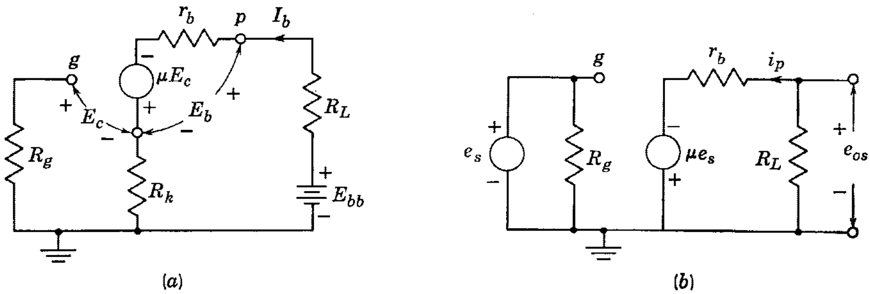


FIG. 6-5. Piecewise-linear models for the triode amplifier with a bypassed cathode resistor. (*a*) Model for d-c components; (*b*) model for signal components.

Figure 6-5*a* shows a model that is valid for the d-c components of current and voltage under the assumption of Class A_1 operation. The quiescent plate current can be found from the two-loop equations for the circuit:

$$E_c = -R_k I_b \tag{6-10}$$

and

$$E_{bb} + \mu E_c = (r_b + R_L + R_k) I_b \tag{6-11}$$

Substituting (6-10) into (6-11) and collecting terms yields

$$E_{bb} = [r_b + R_L + (1 + \mu) R_k] I_b \tag{6-12}$$

After calculating I_b from this relation, the quiescent plate and grid voltages can be calculated easily.

When the mode of operation is Class A_1, the model is a linear circuit, and the applied voltages can be treated separately by superposition. Figure 6-5*b* shows the circuit that results when the signal voltage is considered alone and when the capacitors in the circuit act as short circuits to the signal components of current. The signal component of the output voltage can be calculated by inspection of this circuit when the input signal is given.

For a realistic evaluation of the vacuum triode as a voltage amplifier, the circuit of Fig. 6-5*b* should be compared with the ideal voltage amplifier in Fig. 4-1, for this is about the simplest practical circuit in which the triode can be used. The resistors R_g and R_L appear in addition to r_b as departures from the ideal. As has been explained previously, both of these resistors are needed to ensure proper operation of the tube.

The circuit of Fig. 6-5a contains a controlled source μE_c, whose controlling voltage E_c depends on I_b. But I_b in turn depends on μE_c. Thus the controlling voltage for the controlled source depends on the output from the controlled source. This action, in which the output of a controlled source is transmitted through the circuit in such a way that it affects the controlling quantity, is called feedback. Feedback can have a profound effect on the behavior of an amplifier. Feedback in the circuit of Fig. 6-1 is responsible for the degeneration and consequent reduction in the voltage amplification. The bypass capacitor in the circuit of Fig. 6-2 removes the feedback in so far as the signal components of voltage and current are concerned.

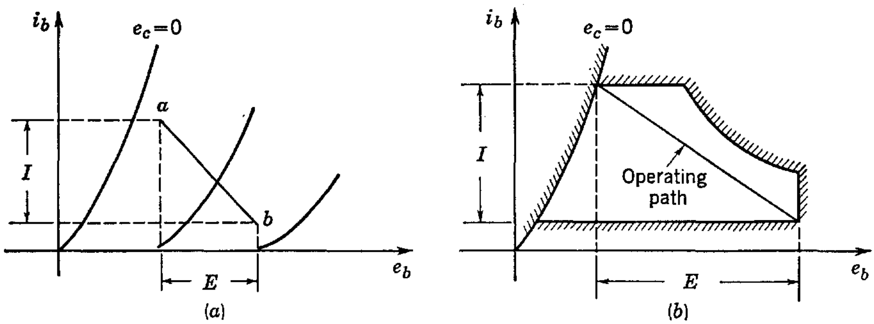


FIG. 6-6. Graphical study of the power output from a triode amplifier. (a) Dynamic operating path; (b) dynamic operating path for maximum power output.

6-2. Transformer-coupled Loads. With a fixed input signal to a given amplifier, the dynamic operating path on the plate characteristics is a line segment such as *ab* in Fig. 6-6a. The peak-to-peak value of the time-varying current through the load is I, and the peak-to-peak value of voltage across the load is E. The power delivered to the load by the time-varying signal is proportional to the product EI. For example, if the signal voltage and current are sinusoidal, the average power delivered to the load by the signal is $P_L = EI/8$. In order to maximize the signal power to the load it is therefore necessary to maximize the product EI. This fact influences the design of amplifiers when the output power, rather than the output voltage or current, is of primary interest.

The product EI, and hence the output power from the amplifier, depends on the slope of the operating path and on the amplitude of the input-signal voltage. In designing an amplifier, these two quantities cannot be chosen at will, however, for the operating path must always be restricted to a certain region on the plate characteristics. A typical set of boundaries for the permissible operating region is shown in Fig.

6-6*b*. In this case, operation is restricted to the negative grid region, the maximum permissible cathode current must not be exceeded, the quiescent point must lie below the hyperbola of maximum permissible plate dissipation, the maximum safe plate voltage must not be exceeded, and the operating path must not extend into the region of low plate current and high distortion. For the conditions pictured in Fig. 6-6*b* it is easy to see that of all the possible operating paths, the one shown in the figure gives the largest value for the product EI, and hence the largest power output. Thus designing an amplifier for maximum power output involves choosing the load resistance to give the optimum slope to the operating path and choosing the plate-supply voltage, the grid bias, and the input-signal amplitude so that each end of the operating path lies at the proper point on the boundary of the permitted region.

In many applications the load resistance is fixed and therefore cannot be adjusted for the optimum value. For example, the nominal impedance of typical loudspeakers lies in the range between 4 and 16 ohms, and this impedance cannot be changed easily. Such a load resistance corresponds to an operating path that is nearly vertical on the plate characteristics of typical vacuum tubes. With the maximum permissible peak-to-peak current through such a load, the peak-to-peak voltage across the load is quite small, and little signal power is delivered to the load. It is customary to remedy this undesirable situation by coupling the loudspeaker to the amplifier through an iron-core transformer. The turn ratio of the transformer is chosen to step up the apparent impedance of the loudspeaker and thereby to give the desired slope to the dynamic operating path.

Another advantage of coupling the load to the amplifier through a transformer is that the d-c component of plate current does not flow through the load. Ordinarily this component serves no useful function in the load, and the power that it delivers to the load is wasted. Moreover, the d-c component may seriously impair the performance of certain types of loads. There are, of course, certain disadvantages associated with the use of transformer coupling. For example, the transformer does not behave well at very low frequencies because the magnetizing inductance is not infinite, and it does not behave well at very high frequencies because the leakage inductance is not zero. In addition, parasitic winding capacitances usually affect the high-frequency performance adversely. Another disadvantage is that the nonlinear nature of the magnetic circuit may cause distortion of the signal waveform.

The circuit diagram of a typical amplifier using transformer coupling to the load is shown in Fig. 6-7. In order to facilitate the analysis of this amplifier it is necessary to devise a linear circuit model to represent the transformer approximately. The model shown in Fig. 6-8 neglects all second-order effects and thereby permits attention to be focused on the first-order effect, namely, the fact that the transformer behaves differently for the d-c and the time-varying components of current and voltage. The model used in Fig. 6-8 represents the transformer as consisting of a magnetizing inductance L and an ideal transformer having a turn ratio N_1/N_2; it ignores winding resistance, leakage inductance, core losses, and distributed capacitance. The inductance L acts as a short circuit to the d-c component of plate current, and it is assumed to act as an open circuit to the time-varying component. In the steady state, the voltage across L, which is given by $L\ di/dt$, must have zero average value; otherwise the current through the inductor would increase indefinitely. It follows from these facts that if C_k acts as a short circuit to time-varying currents, then the voltage across L is e_p, the current through L is I_b, and the current through the primary of the ideal transformer is i_p.

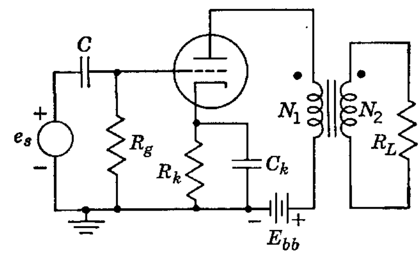


FIG. 6-7. A triode amplifier with a transformer-coupled load.

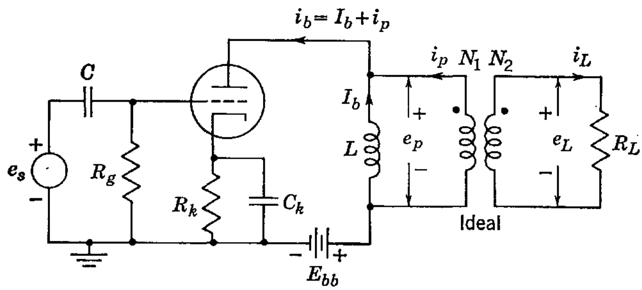


FIG. 6-8. An approximate model for the iron-core transformer.

The ideal transformer is completely governed by two equations:

$$e_L = \frac{N_2}{N_1} e_p \tag{6-13}$$

and

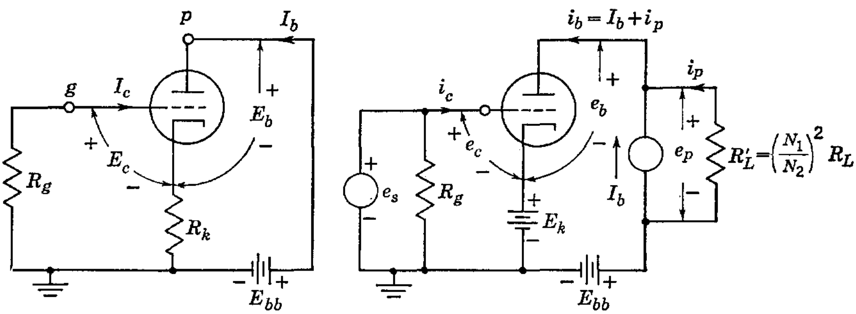
$$i_L = -\frac{N_1}{N_2} i_p \tag{6-14}$$

These relations hold at every instant of time and are independent of frequency and all other considerations.

The circuit of Fig. 6-8 reduces to that shown in Fig. 6-9*a* for the d-c components of current and voltage. The quiescent point for this circuit can be found by the method presented in Sec. 6-1. The equation of the static load line is

$$E_b = E_{bb} - R_k I_b \tag{6-15}$$

This load line is shown in Fig. 6-10; its intersection with the grid-bias line gives the quiescent operating point.



(a) (b)

FIG. 6-9. Representation of an amplifier with a transformer-coupled load. (a) Circuit for d-c components; (b) circuit for total currents and voltages.

For total voltages and currents the circuit can be represented as shown in Fig. 6-9b. The voltage across C_k is represented by the battery $E_k = R_k I_b$, and the current through L is represented by the current source I_b. The resistance R'_L is the load resistance transferred to the primary side of the ideal transformer. From this circuit it follows that

$$e_b = E_{bb} - E_k + e_p = E_{bb} - E_k - R'_L i_p \tag{6-16}$$

Substituting $i_p = i_b - I_b$ yields

$$e_b = E_{bb} - E_k + R'_L I_b - R'_L i_b \tag{6-17}$$

and defining $E'_{bb} = E_{bb} - E_k + R'_L I_b$ leads to

$$e_b = E'_{bb} - R'_L i_b \tag{6-18}$$

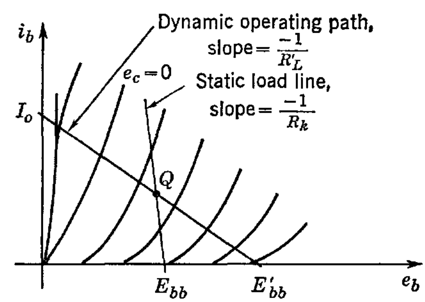


FIG. 6-10. Graphical analysis of an amplifier with a transformer-coupled load.

This is the equation of the dynamic operating path shown in Fig. 6-10. It intersects the plate current axis at $I_o = E'_{bb}/R'_L$, and it intersects the plate voltage axis at E'_{bb}. As the signal voltage e_s takes on various values, the operating point moves along the operating path.

It is clear from the construction in Fig. 6-10 that with signal applied to the amplifier it is possible for the instantaneous value of plate voltage to exceed the plate-supply voltage. This result can be attributed to the fact, shown in Fig. 6-8, that the induced voltage in the magnetizing inductance, $L\ di/dt$, adds in series with E_{bb}. The construction in Fig. 6-10, along with the definition $E'_{bb} = E_{bb} - E_k + R'_L I_b$, also shows that the position of the operating path will shift if I_b changes with signal level as a result of waveform distortion.

At any instant when $i_b = I_b$, the plate voltage given by Eq. (6-17) is

$$e_b = E_{bb} - E_k + R'_L I_b - R'_L I_b = E_{bb} - R_k I_b \qquad (6\text{-}19)$$

Substituting Eq. (6-15) in (6-19) yields

$$e_b = E_b \qquad (6\text{-}20)$$

for the value of e_b at any instant when $i_b = I_b$. When there is negligible waveform distortion in the amplifier, $I_b = I_{bo}$ and $E_b = E_{bo}$; hence under

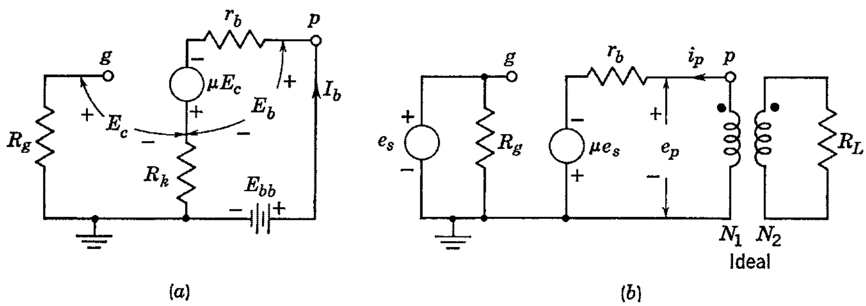


FIG. 6-11. Models for the triode amplifier with a transformer-coupled load. (a) Model for d-c components; (b) model for signal components.

this condition the dynamic operating path passes through the quiescent operating point as indicated in Fig. 6-10. Accordingly, a simple procedure for constructing the dynamic operating path is to pass a line through the quiescent point with a slope $-1/R'_L$.

As an alternative procedure, the amplifier of Fig. 6-7 can be analyzed by replacing the tube with its piecewise-linear model. For Class A_1 operation, the model for the circuit that shows the relations among the d-c components of current and voltage has the form shown in Fig. 6-11a. The quiescent operating point can be calculated from this circuit. In Class A_1 operation the signal source can be treated separately from the direct voltages in the circuit. The model for signal components of current and voltage is shown in Fig. 6-11b; the load voltage and power can be computed from this circuit. The d-c and the signal currents and voltages calculated from the two circuits in Fig. 6-11 can be combined by superposition to obtain the total instantaneous currents and voltages.

Example 6-1. Plate characteristics for the triode in the amplifier circuit of Fig. 6-12a are shown in Fig. 6-12b. The maximum permissible plate dissipation for the triode is 20 watts. The circuit is to be adjusted so that the quiescent operating point is at $i_b = 75$ ma and $e_c = -40$ volts. The problem is to determine suitable values for the circuit parameters and the supply voltage and to study the performance of the amplifier when signal is applied.

a. Determine the values of E_{bb} and R_k that will give the specified quiescent operating point.

b. Determine whether or not the quiescent plate dissipation exceeds the maximum permissible value.

c. The optimum load resistance for the tube is 2250 ohms. What turn ratio should the transformer have?

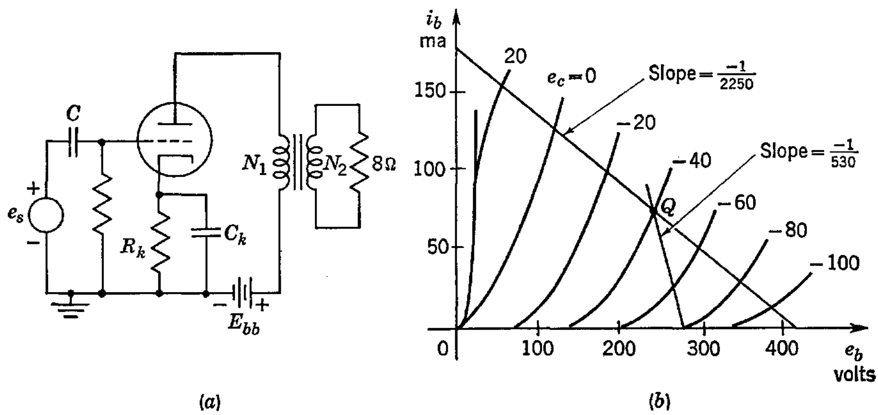


FIG. 6-12. Amplifier with a transformer-coupled load for Example 6-1. (*a*) Circuit; (*b*) graphical construction.

d. A signal voltage, $e_s = 40 \cos 2000\pi t$ volts, is applied at the input. Neglecting the small shift in dynamic operating path caused by distortion, determine the maximum and minimum instantaneous values of e_b.

e. For the conditions of part *d*, determine P_L and P_b.

Solution. *a.* The value of cathode resistance required is fixed by the specified values of e_c and i_b at the quiescent point. Thus

$$R_k = \frac{-E_{co}}{I_{bo}} = \frac{40}{75} = 0.53 \text{ kilohm}$$

The supply voltage required is

$$E_{bb} = E_{bo} + R_k I_{bo} = E_{bo} + 40$$

From the plate characteristics, the quiescent plate voltage is $E_{bo} = 240$ volts. Hence

$$E_{bb} = 240 + 40 = 280 \text{ volts}$$

b. The quiescent plate dissipation is

$$P_{bo} = E_{bo} I_{bo} = (240)(75)(10)^{-3} = 18 \text{ watts}$$

Hence the maximum permissible dissipation is not exceeded.

c. The required turn ratio is given by

$$R'_L = \left(\frac{N_1}{N_2}\right)^2 R_L$$

$$\frac{N_1}{N_2} = \sqrt{R'_L/R_L} = \sqrt{2250/8} = 16.8$$

d. The dynamic operating path is shown in Fig. 6-12*b*. With the specified signal applied, the operating point moves along this path between the plate characteristics for $e_c = 0$ and $e_c = -80$ volts. Thus

For $e_c = 0$: $e_b = e_{b\,\min} = 120$ volts
For $e_c = -80$: $e_b = e_{b\,\max} = 340$ volts

Note that the maximum instantaneous plate voltage exceeds the plate-supply voltage by 60 volts.

e. Under the specified conditions, $e_p \approx 110 \cos 2000\pi t$ volts. Neglecting the losses in the transformer, the power delivered to the primary is also the power delivered to the load. Hence

$$P_L = \left(\frac{110}{\sqrt{2}}\right)^2 \frac{1}{2250} = 2.7 \text{ watts}$$

The plate dissipation is the power delivered by the battery minus the power delivered to the load and the power dissipated in the cathode resistor. Thus

$$P_b = P_{bb} - P_k - P_L$$

But P_{bb} and P_k are independent of the signal as long as the distortion is small, and $P_{bb} - P_k$ is the quiescent plate dissipation. Hence, using the result of part *b*,

$$P_b = P_{bo} - P_L = 18 - 2.7 = 15.3 \text{ watts}$$

It is clear from these results that the efficiency of this amplifier is not very large.

6-3. *RC* Plate-circuit Loads. Another important amplifier circuit is shown in Fig. 6-13. This circuit is commonly used when amplifiers are connected in cascade as indicated in Fig. 4-2; the blocking capacitor C_2 prevents the direct voltage at the plate of the first amplifier from affecting the quiescent conditions in the second stage.

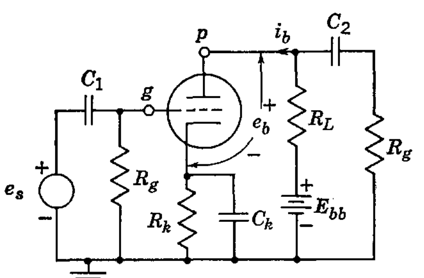


FIG. 6-13. An amplifier with an *RC* plate load.

The static load line and the dynamic operating path can be determined by considering separately the d-c and the time-varying components of current and voltage in the circuit connected to the tube. The capacitors in the circuit act as open circuits to the d-c components; hence these components are related by

$$E_b = E_{bb} - (R_L + R_k)I_b \tag{6-21}$$

The capacitors are assumed to act as short circuits for the time-varying components; hence the time-varying components are related by

$$e_p = -R'_L i_p \tag{6-22}$$

where R'_L is the resistance of R_L and R_g in parallel. Superimposing these two solutions gives

$$E_b + e_p = e_b = E_{bb} - (R_L + R_k)I_b - R'_L i_p \tag{6-23}$$

Substituting $i_p = i_b - I_b$ and collecting terms gives

$$e_b = E_{bb} - (R_L + R_k - R'_L)I_b - R'_L i_b \tag{6-24}$$

Finally, if $E'_{bb} = E_{bb} - (R_L + R_k - R'_L)I_b$, then

$$e_b = E'_{bb} - R'_L i_b \tag{6-25}$$

Equation (6-21) gives the static load line; its intersection with the grid bias line gives the quiescent point, as indicated in Fig. 6-14. Equation (6-25) gives the dynamic operating path; this path is most easily constructed by passing a line through the quiescent point with a slope $-1/R'_L$.

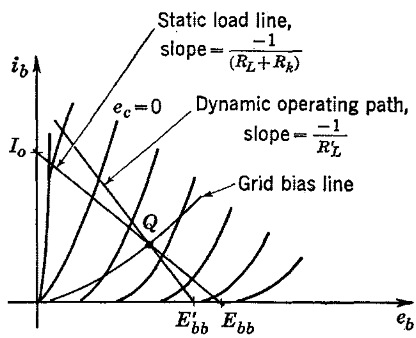


FIG. 6-14. Graphical analysis of an amplifier with an RC plate load.

6-4. Amplifiers with Reactive Loads. In the examples considered in the preceding sections, all capacitors and inductors have been considered to act either as open circuits or short circuits to time-varying currents. Thus the circuits have been in fact resistive circuits. When these conditions are not fulfilled, the relatively simple results obtained above do not hold. Suppose, for example, that i_p in Fig. 6-13 is sinusoidal and of such a frequency that the reactance of C_2 is comparable with the resistance of R_g and R_L. Then e_p is sinusoidal but not in phase with i_p, and if the combinations of i_b and e_b existing at various instants of time are plotted on the plate characteristics, an elliptical operating path (Lissajous pattern) is obtained. In general, the signals to be amplified can be represented as the superposition of sinusoidal components of many different frequencies. If the reactance cannot be neglected in such cases, each sinusoidal component of the signal causes an elliptical movement of the operating point, and the dynamic operating path consists of the superposition of many elliptical motions. Thus the dynamic operating path under these conditions is very complicated indeed.

The dynamic operation of an amplifier with a reactive load can be studied further by considering the circuit shown in Fig. 6-15*b*. The signal applied to the amplifier, which in this simple circuit is also the grid-to-cathode voltage, is shown in Fig. 6-15*a*, and the time constant of the *RL* load is such that the circuit reaches the steady state during each half cycle of the signal voltage. Thus at the end of each half cycle the operating point is on the static load line shown in Fig. 6-15*c*; for example, at the end of the time interval between 0 and *T*, the operating

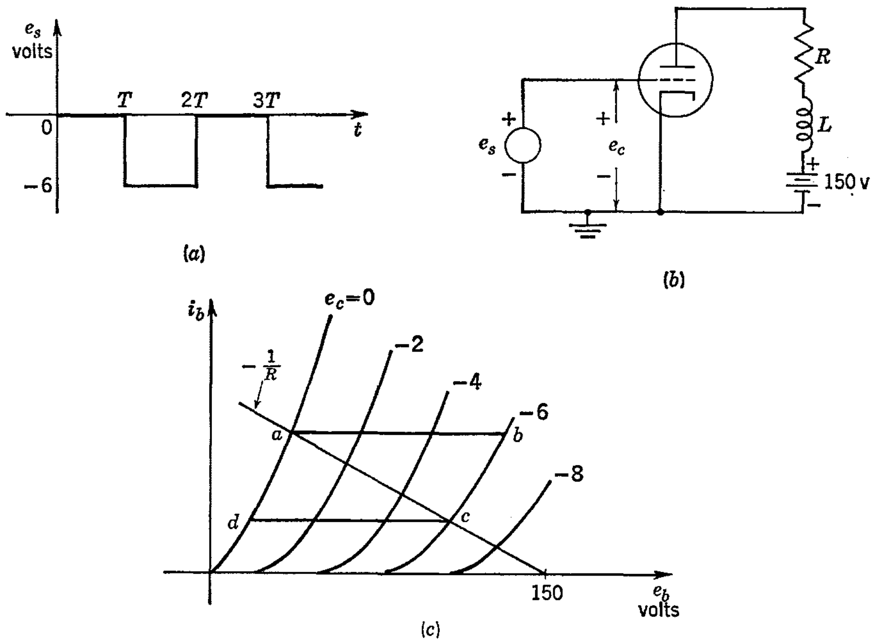


FIG. 6-15. Graphical analysis of a triode amplifier with a reactive load. (*a*) Input signal waveform; (*b*) circuit; (*c*) graphical construction.

point is at *a* in Fig. 6-15*c*. During the interval between *T* and $2T$, the grid voltage is -6 volts, and at the end of the interval the operating point is at *c* in Fig. 6-15*c*; however, it does not move along the static load line from *a* to *c*. The current in the inductor cannot change instantaneously. Therefore, when e_s changes abruptly from 0 to -6 volts at the instant $t = T$, the operating point moves instantaneously along a line of constant i_b to point *b* on the plate characteristic for $e_c = -6$ volts. The plate current then decays to its new steady-state value with the operating point moving along the line of constant $e_c = -6$ volts to point *c*. Similarly, when the grid voltage changes abruptly to 0 volts at $t = 2T$, the operating point moves abruptly along a line of constant i_b to point *d* on the $e_c = 0$ characteristic, and then it moves along the line of constant $e_c = 0$ to point a. Thus the complete dynamic operating path for one full cycle of the signal is the contour $abcd$ shown in Fig. 6-15c.

6-5. The Use of Thevenin's Theorem in Graphical Analyses. The circuits that have been analyzed in the preceding sections have been simple in form. However, any resistive circuit containing only one tube, no matter how complicated, can be reduced to such a simple form by a slight extension of Thevenin's theorem. Figure 6-16a shows a tube connected to a general linear resistive network; Fig. 6-16b shows a network that is equivalent to the one of Fig. 6-16a in so far as the tube is concerned, provided the parameters are given appropriate values. If no grid current flows, this equivalent circuit can be analyzed graphically by the methods developed in the foregoing sections. The graphical procedures must be developed further if the case in which grid current flows is to be considered.

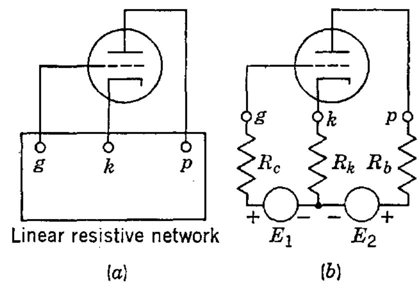


FIG. 6-16. Thevenin's theorem applied to vacuum-tube circuits. (a) Vacuum-tube circuit; (b) an equivalent circuit.

The voltage sources and resistances in the equivalent network can be determined by open-circuit and short-circuit measurements (or calculations) made at the terminals of the linear network. The internal voltages are determined directly from open-circuit measurements. With p-k short-circuited and g-k open-circuited, R_b and R_k are determined by the current in the short circuit and the voltage across the open circuit. Then R_c is determined from the short-circuit current between g and k when p-k is open-circuited.

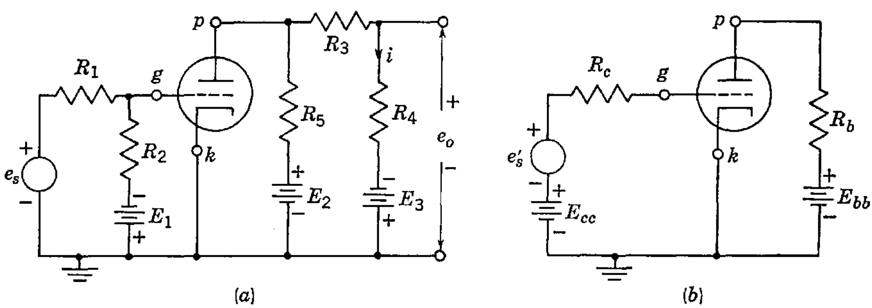


FIG. 6-17. Simplification of an amplifier circuit. (a) Amplifier; (b) an equivalent circuit.

The circuit of Fig. 6-17a is presented as an example. The grid circuit in this amplifier is arranged so that the two sources e_s and E_1 have a common terminal with each other and with the sources in the plate circuit. The voltage source E_3 and the voltage divider R_3-R_4 at the output permit the d-c component of the output voltage to be adjusted to a desired value. For example, it may be desired that $e_o = 0$ when $e_s = 0$.

An equivalent circuit for the amplifier is shown in Fig. 6-17*b*. In this relatively simple circuit there is no coupling between the grid and plate circuits in the network external to the tube; hence the equivalent circuit is obtained by the simple process of applying Thevenin's theorem separately to the grid and plate circuits. The form of this circuit is identical with that of the basic triode amplifier, and it can therefore be solved for i_b and e_b by graphical methods already developed. Then, knowing i_b and e_b, all voltages and currents in the plate circuit can be found. For example, the current i in resistor R_4 is

$$i = \frac{e_b + E_3}{R_3 + R_4} \tag{6-26}$$

and the output voltage is

$$e_o = R_4 i - E_3 \tag{6-27}$$

6-6. Summary. Most practical vacuum-tube circuits contain reactive elements, either L or C or both. The former acts as a short circuit, and the latter acts as an open circuit to the d-c components of voltage and current. In many cases it is permissible to consider the L's to be open circuits and the C's to be short circuits to the signal components of voltage and current. In this case the amplifier can be represented by two different resistive circuits, one for the d-c components and the other for the signal components. When such circuits are analyzed by graphical constructions on the tube characteristics, two different load lines, or operating paths, are required, one for the d-c components and the other for the signal components. These operating paths are straight lines because the corresponding circuits connected to the tube are linear resistive networks. If the networks are not resistive or are not linear, the operating paths will in general not be straight lines.

PROBLEMS

6-1. A 6J5 triode is used in the amplifier of Fig. 6-2. The plate-supply voltage is 300 volts, and the values of the circuit resistances are $R_L = 22$ kilohms, $R_k = 2$ kilohms, and $R_g = 1$ megohm.

a. Find the quiescent plate voltage and plate current by a graphical construction on the plate characteristics.

b. What is the quiescent value of e_o?

c. Determine the quiescent plate dissipation.

d. The signal voltage e_s is a square wave having half cycles of equal duration and having a peak-to-peak value of 16 volts, and the capacitors in the circuit act as short circuits to the time-varying components of voltage and current. Determine the average value of the plate dissipation with this signal applied, and compare it with the quiescent dissipation.

6-2. One section of a 12AX7 twin triode is used in the amplifier of Fig. 6-2. The plate-supply voltage is 300 volts, and the quiescent operating point is to be at I_{bo} = 0.5 ma and E_{bo} = 150 volts. What values of R_L and R_k are required? With this adjustment, what is the greatest peak-to-peak variation that the output voltage e_o can have in Class A_1 operation?

6-3. A 6B4-G power triode is used in the transformer-coupled amplifier shown in Fig. 6-7. (The plate characteristics for this tube are identical with those for the 2A3.) The load resistance R_L is 6 ohms, and the transformer turn ratio is $N_1/N_2 = 20:1$.

a. What values of E_{bb} and R_k are required to place the quiescent operating point at E_{bo} = 200 volts, I_{bo} = 70 ma, and E_{co} = −30 volts?

b. Draw the plate characteristics for E_c = 0, −30, and −60 volts. Construct the static load line and the dynamic operating path on these curves under the assumption that $I_b = I_{bo}$. The transformer can be represented by the model shown in Fig. 6-8.

c. A square-wave signal voltage having half cycles of equal duration and having a peak-to-peak amplitude of 60 volts is applied at e_s. The capacitors in the circuit act as short circuits, and the magnetizing inductance of the transformer acts as an open circuit to the time-varying components of voltage and current. With this signal applied, determine the average plate current I_b and check the assumption of part *b* that $I_b = I_{bo}$.

d. Sketch the waveform of current in the load; show on this sketch the amplitude of the load current.

e. Calculate the average power delivered to the load resistor.

f. Calculate the average plate dissipation, and compare it with the quiescent plate dissipation.

6-4. A 2A3 power triode is used in the amplifier of Fig. 6-7. The input signal is $e_s = 30 \cos 2000\pi t$ volts, and the load resistance is R_L = 500 ohms. The capacitors act as short circuits, and the magnetizing inductance of the transformer acts as an open circuit to the time-varying components of voltage and current.

a. What values of E_{bb} and R_k are required to place the quiescent operating point at I_{bo} = 75 ma and E_{co} = −30 volts?

b. What is the smallest value of turn ratio N_1/N_2 that will give reasonably linear operation with the specified signal applied? The answer to this question is a matter of judgment; some guidance is given in Sec. 6-2 and Fig. 6-6*b*.

c. If the load is removed from the transformer secondary, the dynamic operating path becomes a horizontal line. Under this condition, and with other conditions as in part *b*, what is the peak value of voltage across the transformer primary?

d. What is the average power delivered to the load under the conditions of part *b*? The small effects of waveform distortion can be neglected.

6-5. An amplifier with an *RC* plate load is shown in Fig. 6-18. The operation of the circuit is to be studied with the aid of graphical constructions on the plate characteristics.

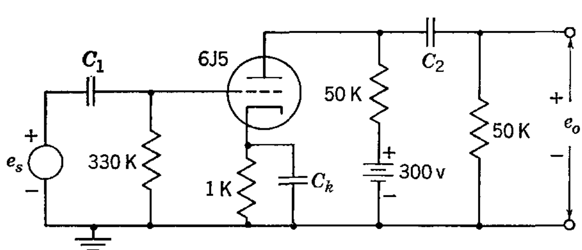


FIG. 6-18. Amplifier for Prob. 6-5.

a. Construct the static load line and locate the quiescent operating point.

b. Construct the dynamic operating path, assuming $I_b = I_{bo}$, so that this path passes through the quiescent point.

c. For linear Class A_1 operation, the dynamic operating point should not go into the positive grid region and it should not go below the $i_b = 1$ ma line. Subject to these restrictions, what is the greatest amplitude that a sinusoidal input signal can have?

d. What is the maximum amplitude of output voltage that can be obtained from the amplifier under the conditions of part *c*?

6-6. In high-gain multistage amplifiers it is usually necessary to use decoupling networks, like that formed by the 100-kilohm resistor and capacitor C in the circuit of Fig. 6-19, to reduce interaction among the stages. The effect of the decoupling network on the procedures of graphical analysis is to be examined.

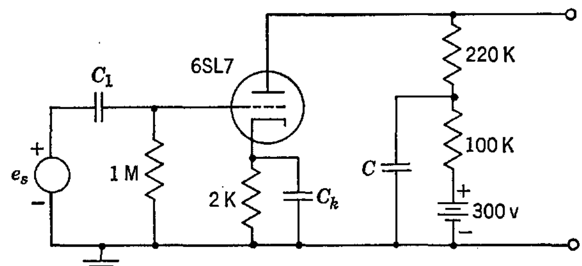


FIG. 6-19. Amplifier with a plate-decoupling network for Prob. 6-6.

a. Sketch the static load line on the e_b, i_b coordinates, and locate the quiescent operating point. The plate-characteristic curves can be omitted from this sketch.

b. Add the dynamic operating path to the sketch of part *a*, making the assumption that $I_b = I_{bo}$. A simple procedure is to pass a line having the proper slope through the quiescent point. Indicate the voltage at which this path intersects the e_b axis.

6-7. The amplifier shown in Fig. 6-20 is identical with the one shown in Fig. 6-18 except that the plate load resistor is replaced with an inductor. With this arrangement the plate-supply voltage required for a given quiescent point is smaller than in the circuit of Fig. 6-18, and slightly higher gain is obtained; on the other hand, the size, weight, and cost are greater, and, as a rule, the performance is poorer at very high and very low frequencies. Certain features of this amplifier are to be examined by graphical analysis.

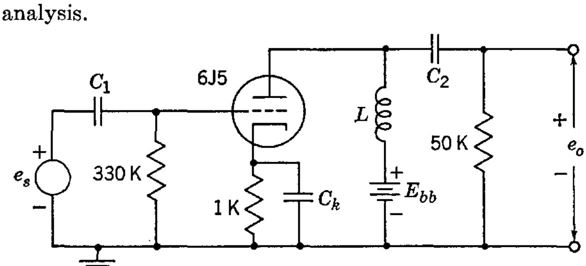


FIG. 6-20. Amplifier for Prob. 6-7.

a. The quiescent operating point for the amplifier in Fig. 6-18 is approximately at $i_b = 4$ ma and $e_c = -4$ volts. Assuming that the inductor has negligible resistance, what value of E_{bb} is required to give this quiescent point in the amplifier of Fig. 6-20? Compare this result with the plate-supply voltage used in the circuit of Fig. 6-18.

b. Sketch the static load line and the dynamic operating path on the e_b, i_b coordinates. Assume that the inductor acts as an open circuit and the capacitors act as short circuits to the time-varying components of current and voltage. The characteristic curves can be omitted from this sketch, but the voltage at which each operating path intersects the e_b axis should be indicated. A simple procedure for constructing the dynamic operating path when $I_b = I_{bo}$ is to pass a line having the proper slope through the quiescent point.

c. If the input signal is sinusoidal, what is the greatest amplitude that it can have in Class A_1 operation? What is the amplitude of the output voltage under this condition?

6-8. The characteristics of the 6SN7 triode used in the amplifier of Fig. 6-21 can be approximated by a piecewise-linear model in which $\mu = 20$ and $r_b = 10$ kilohms. Certain features of the amplifier are to be examined by piecewise-linear analysis.

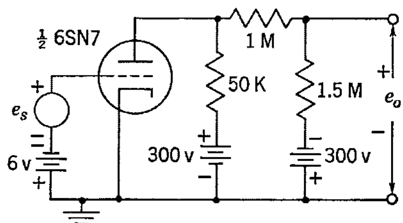


FIG. 6-21. Amplifier for Prob. 6-8.

a. Find the quiescent operating point.

b. Construct the curve of e_o versus e_s. Cover the range of e_s between -20 and 6 volts. Give the coordinates of all break points, and give the slope of each segment of the curve.

c. By what factor does this circuit amplify the input signal when the operation is Class A_1?

6-9. The circuit shown in Fig. 6-22*a* is an elementary d-c vacuum-tube voltmeter. The resistance of the d-c milliammeter *M* is negligibly small. The problem is to determine the meter reading as a function of the input voltage.

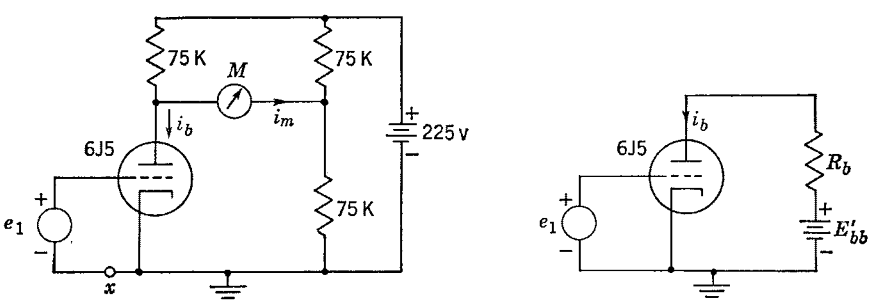


FIG. 6-22. Elementary vacuum-tube voltmeter for Prob. 6-9.

a. Find R_b and E'_{bb} in the Thevenin equivalent circuit of Fig. 6-22*b*.

b. Construct an accurate curve of i_b versus e_1 in the range between plate-current cutoff and $e_1 = 0$. Points on this curve can be determined by graphical analysis with the aid of the equivalent circuit in Fig. 6-22*b*.

c. Since the circuit connected to the tube is linear, i_b and i_m are related by a linear equation of the form $i_m = A + Bi_b$. Find the numerical values of the constants A and B. (Consider whether two special cases, $i_b = 0$ and $i_m = 0$, might permit an easy solution.)

d. Plot i_m versus e_1 on the same coordinates as the curve of part *b*. This curve is a calibration curve for the voltmeter.

e. What value of grid bias must be inserted at x to make $i_m = 0$ when $e_1 = 0$? This voltage provides the zero adjustment for the voltmeter.

6-10. A 6J5 triode is used in the circuit of Fig. 6-15*b* to control the operation of a relay. The plate-supply voltage is 150 volts, and the signal voltage has the waveform shown in Fig. 6-15*a*. The inductance of the relay coil is represented by L, and the resistance of the coil plus an additional current limiting resistor is $R = 10$ kilohms. The behavior of the circuit is to be examined by graphical analysis under the assumption that the steady state is reached during each half cycle.

a. Plot the plate characteristics for $e_c = 0$ and $e_c = -6$ volts. Construct the dynamic operating path, and give the coordinates of each corner of the path.

b. At each corner of the dynamic operating path, determine the voltage drops across the tube, the inductor, and the resistor. Give the polarity of each voltage by assigning a negative sign to voltage rises in the direction of i_b.

c. Compare the maximum instantaneous value of e_b with the plate-supply voltage.

CHAPTER 7

INCREMENTAL LINEAR MODELS FOR THE VACUUM TRIODE

The analysis and design of voltage amplifiers can be separated into two distinct parts. The first part is concerned with the location of the quiescent operating point; this point must lie in a suitable region of the plate characteristic if the tube is to be effective as a voltage amplifier. The second part is concerned with the signal components of voltage and current. The voltage amplification provided by the circuit is usually the feature of primary interest in this second part of the study.

Calculations related to the quiescent operation of the circuit are usually based on either a graphical or a piecewise-linear analysis of the circuit. Calculations related to the signal are practically always based on a suitable linear network model for the circuit, for voltage amplifiers usually operate in the linear Class A_1 mode. In linear circuits such as these it is permissible to treat the signal components of voltage and current separately by superposition; this fact permits a degree of simplification in the models used to represent the circuits. Models that account only for the signal components of voltage and current are called incremental models.

The input signals to voltage amplifiers are in many cases measured in millivolts or microvolts. Such voltages are too small to be used in graphical analyses unless the usual tube characteristics are redrawn on an expanded scale. Of perhaps even more importance, if the circuit contains reactances that cannot be considered to act either as open circuits or short circuits, as often is the case, then the operating path on the plate characteristic is no longer a straight line, and graphical analysis becomes very tedious indeed. For these and other reasons, incremental models for voltage amplifiers provide the most convenient basis for analysis and design.

The objective of this chapter is to present the incremental models that are most commonly used to represent the vacuum triode and to examine their properties in some detail. These models are derived by general methods that are applicable to other physical devices used as components in electrical circuits. In particular, these same methods are used in deriving models for the transistor amplifier.

7-1. Derivation of Incremental Triode Models. An incremental linear model for the triode can be developed from the piecewise-linear model used in Chaps. 5 and 6. Figure 7-1 shows the plate characteristics of a typical triode and a piecewise-linear model for the tube. The parameters of the model are ordinarily chosen to give the best approximate fit to the tube characteristics over the entire usable portion of the characteristics. Under conditions of small-signal operation, however, the excursions of the operating point are restricted to a small region surrounding the quiescent operating point; such a region is enclosed by a dotted line and labeled R in Fig. 7-1*a*. If operation is to be restricted to the region R, then the parameters of the model in Fig. 7-1*b* should be chosen to give

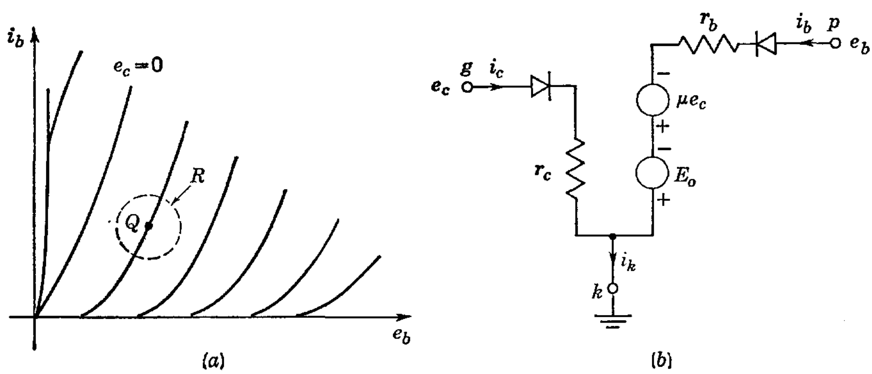


FIG. 7-1. Small-signal operation of a triode. (*a*) Plate characteristic; (*b*) piecewise-linear model.

the best fit inside R. A detailed examination would show that the plate characteristics inside any small region are nearly straight, parallel, and equally spaced for equal increments of grid voltage; hence a very good fit can be obtained inside any such small region. If the quiescent point is moved to a new location, a new set of parameter values is needed to give a good fit inside the new operating region R'.

The tube currents $i_b = I_b + i_p$ and $i_c = I_c + i_g$ resulting from the tube voltages $e_b = E_b + e_p$ and $e_c = E_c + e_g$ can be calculated from the piecewise-linear model when the operation is restricted to the region R and the parameters are chosen accordingly. When the signal components only are of interest, they can be calculated separately by setting the d-c components equal to zero. When the region R corresponds to Class A_1 operation, the corresponding incremental model has the form shown in Fig. 7-2. The plate resistance is symbolized r_p instead of r_b to indicate that it is chosen to give the best representation in the small region R

surrounding the quiescent point; r_p is referred to as the *incremental plate resistance* of the tube.

The incremental linear model for the triode can be derived by an alternative procedure that does not depend on a piecewise-linear model for the starting point. The basic principles used in this derivation have wide application in the physical sciences and therefore merit careful consideration. The input and output characteristics of the triode give the electrode currents as functions of the electrode voltages:

$$i_b = f_b(e_b,e_c) \qquad (7\text{-}1)$$

$$i_c = f_c(e_b,e_c) \qquad (7\text{-}2)$$

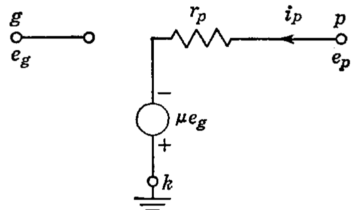


FIG. 7-2. Incremental model for a triode in Class A₁ operation.

Also, the cathode current is given by

$$i_k = i_b + i_c \qquad (7\text{-}3)$$

Thus if the electrode voltages are known, the electrode currents can be found. If e_b is given a small increment, the corresponding increment in i_b can be expressed as

$$\Delta i_b \approx di_b = \frac{\partial i_b}{\partial e_b}\, de_b = g_p\, de_b \qquad (7\text{-}4)$$

where $\partial i_b/\partial e_b$ is the rate of change of i_b with respect to e_b with e_c held constant. It is easily shown from Fig. 7-2 that this partial derivative is $1/r_p$; hence it is designated g_p and called the *incremental plate conductance*.

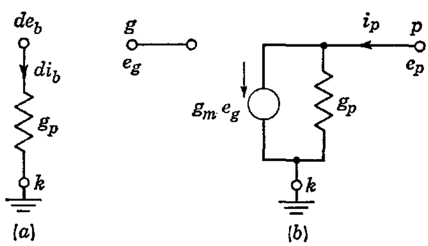


FIG. 7-3. Incremental model for a triode in Class A_1 operation. (*a*) Increment in plate voltage only; (*b*) increments in both grid and plate voltages.

If the mode of operation is Class A_1, i_c is always zero and $di_k = di_b$. Thus Eq. (7-4) can be symbolized by the simple circuit shown in Fig. 7-3*a*.

Suppose now that in Class A_1 operation both e_b and e_c are given small increments; the resulting increment in i_b is

$$\Delta i_b \approx di_b = \frac{\partial i_b}{\partial e_b}\, de_b + \frac{\partial i_b}{\partial e_c}\, de_c \qquad (7\text{-}5)$$

If the partial derivatives, which have the dimensions of conductance, are given suitable symbols, and if the increments in i_b, e_b, and e_c are designated i_p, e_p, and e_g, then Eq. (7-5) becomes

$$i_p = g_p e_p + g_m e_g \qquad (7\text{-}6)$$

where g_p is again the incremental plate conductance, and g_m is the *incre-*

mental grid-to-plate transconductance, or, more briefly, the *mutual conductance.* Again, i_c is identically zero, di_c is zero, and from (7-3)

$$di_k = di_b = i_p \tag{7-7}$$

Equations (7-6) and (7-7) correspond to the circuit shown in Fig. 7-3*b*.

Figures 7-3*b* and 7-2 show two different linear incremental models for the vacuum triode in Class A_1 operation. These circuits give the correct relations among small increments in the currents and voltages around a stated quiescent point. *They cannot in general be used in calculating total voltages and currents,* for they match the tube characteristics over only a

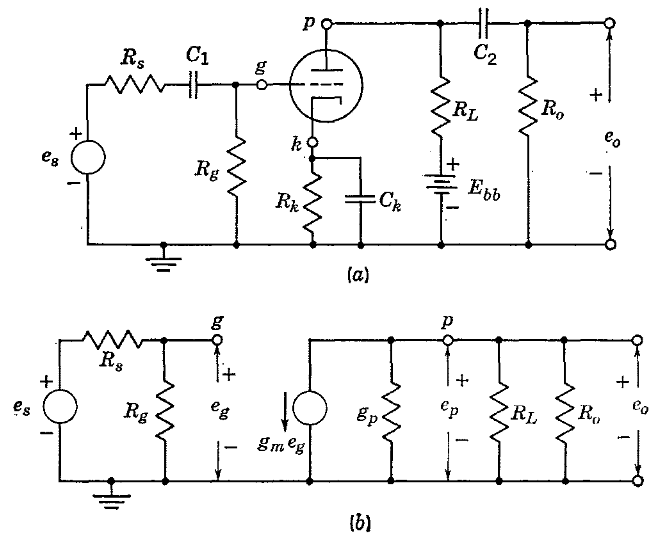


FIG. 7-4. A Class A_1 triode voltage amplifier and its incremental model. (*a*) Circuit; (*b*) incremental model.

small region. The model of Fig. 7-3*b* can be put in the same form as that in Fig. 7-2 by the simple process of a source transformation. The transformation shows that if the two models are to be equivalent, the relation

$$\mu = g_m r_p \tag{7-8}$$

must hold. Both of these models are suitable representations for the triode, and both are widely used. The choice between them is dictated by convenience. A typical triode amplifier and its incremental linear model for the condition that the capacitors act as short circuits to the signal components of current and voltage are shown in Fig. 7-4. The capacitors and the direct voltage source E_{bb} are replaced by short circuits in the model. The output voltage resulting from any specified input signal voltage is easily evaluated from the model in Fig. 7-4*b*.

When the vacuum triode is used at high frequencies, above a few tens of kilocycles per second, the effects of the interelectrode capacitances must be accounted for. Figure 7-5 shows an incremental model for the triode that accounts approximately for these effects. This model is useful in the band of frequencies up to a few tens of megacycles per second. At higher frequencies, lead inductance and the transit time required by the electrons in passing from the cathode to the plate must also be accounted for. It is well to remember that the actual values of the interelectrode capacitances are not constant but depend on the tube voltages and currents. This fact can often be ignored, but in some instances it becomes quite important. For example, it imposes the principal limitation on the frequency stability that can be obtained with many types of vacuum-tube oscillators.

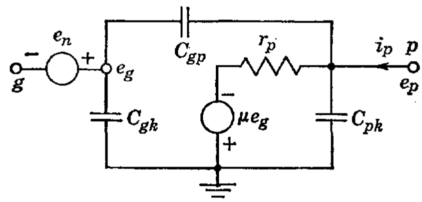


FIG. 7-5. Triode model accounting for tube capacitances and tube noise.

The voltage source e_n in Fig. 7-5 accounts for the noise generated in the tube. The noise source and the interelectrode capacitances are usually not included in the model unless they have a significant effect on the performance of the circuit.

7-2. Incremental Triode Parameters. The parameters μ, g_m, and $r_p = 1/g_p$ are the incremental parameters, or the differential coefficients, of the vacuum tube. They are closely related to the volt-ampere characteristics of the tube, as is brought out in the development that follows. The plate conductance,

$$g_p = \frac{\partial i_b}{\partial e_b} = \frac{di_b}{de_b}\bigg|_{e_c \text{ const}} \tag{7-9}$$

is the slope of the plate characteristic at the quiescent operating point. The mutual conductance,

$$g_m = \frac{\partial i_b}{\partial e_c} = \frac{di_b}{de_c}\bigg|_{e_b \text{ const}} \tag{7-10}$$

is the slope of the constant-voltage transfer characteristic (Fig. 5-6) at the quiescent point. The connection between the amplification factor and the characteristic curves can be found by writing (7-5) in the form

$$di_b = g_p\, de_b + g_m\, de_c \tag{7-11}$$

If e_b and e_c take on increments of such magnitude and sign that i_b remains constant, then $di_b = 0$, and (7-11) yields

$$-\frac{de_b}{de_c}\bigg|_{i_b \text{ const}} = \frac{g_m}{g_p} = g_m r_p = \mu \tag{7-12}$$

Thus the amplification factor is equal to the magnitude of the slope of the constant-current transfer characteristic shown in Fig. 5-5. The minus sign in (7-12) results from the fact that e_b and e_c must take on increments of opposite sign if i_b is to remain constant.

Certain important facts about the incremental parameters of the triode can be deduced from their relations to the slopes of the characteristic curves for the tube. When the slopes of the plate characteristics and the constant-voltage transfer characteristics are positive, g_m and r_p are positive; when the slope of the constant-current characteristic is negative, the amplification factor is positive. These parameters are usually, but not always, positive. The plate characteristics of certain types of vacuum tubes have negative slopes in certain ranges of voltage and current. The incremental plate resistance for such tubes is negative in the vicinity of quiescent points in this region. The slopes of the constant-current characteristics are relatively constant over a wide range of currents and voltages (except at low values of plate current); hence the amplification factor is relatively constant except at low values of plate current. The slopes of the plate characteristics and the constant-voltage characteristics vary widely over the normal operating range, however; hence g_m and r_p depend strongly on the location of the quiescent point. New values of g_m and r_p must be determined for each new quiescent point. Moreover, the quiescent point must be chosen with some care if optimum performance of the tube is to be realized.

It follows from Eqs. (7-9) to (7-12) that the incremental parameters can be evaluated by measuring the slopes of the various characteristics at the quiescent point. It often happens, however, that only the plate characteristics are available. In such cases the tube parameters must be evaluated from the plate characteristics alone; a procedure that yields satisfactory results in this evaluation is presented in Example 7-1.

Example 7-1. The triode in the voltage amplifier of Fig. 7-6*a* has the plate characteristics shown in Fig. 7-6*b*. The problem is to choose the circuit parameters to give a specified quiescent operating point and to study the behavior of the circuit under small-signal operating conditions.

a. Determine the values of R_L and R_k that are required to locate the quiescent point at $i_b = 4$ ma and $e_c = -4$ volts.

b. Determine the incremental parameters for the triode at the quiescent point.

c. The applied signal is a small time-varying voltage, and the capacitors in the circuit behave as short circuits to the signal components of voltage and current. By what factor is the signal voltage amplified?

d. By what factor is the signal of part *c* amplified if the bypass capacitor C_k is removed?

Solution. *a.* The relations between the quiescent operating point and the circuit parameters are illustrated in Fig. 6-4. The voltage drop across R_k must be 4 volts under quiescent conditions; hence,

$$R_k = \frac{-E_{co}}{I_{bo}} = \frac{4}{4} = 1 \text{ kilohm}$$

The dynamic operating path passes through the quiescent point and intersects the plate-voltage axis at $e_b = E_{bb} - E_k = 296$ volts. The graphical construction of Fig. 7-6*b* shows that this path intersects the plate-current axis at

$$I_o = \frac{E_{bb} - E_k}{R_L} = 7 \text{ ma}$$

Thus
$$R_L = \frac{E_{bb} - E_k}{I_o} = \frac{296}{7} = 42.3 \text{ kilohms}$$

b. The amplification factor is nearly constant over most of the area of the plate characteristics shown in Fig. 7-6*b*; hence it can be evaluated with sufficient accuracy

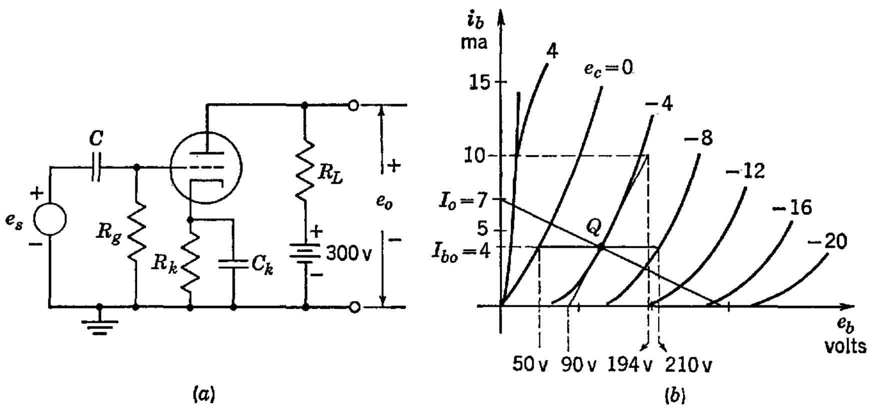


FIG. 7-6. Analysis of a triode amplifier under small-signal conditions. (*a*) Circuit; (*b*) graphical construction.

by substituting relatively large increments of e_b and e_c in Eq. (7-12) provided they are chosen so that i_b is held constant at its quiescent value. In addition, the increments should be chosen so that the quiescent point lies near the center of the range covered. Following this procedure, Fig. 7-6*b* shows that a change of grid voltage from 0 to −8 volts requires a change of plate voltage from 50 to 210 volts to maintain the plate current in the triode constant at its quiescent value. Accordingly, Eq. (7-12) yields

$$\mu \approx -\frac{\Delta e_b}{\Delta e_c} = -\frac{210 - 50}{-8 - 0} = \frac{160}{8} = 20$$

Equation (7-9) shows that the plate conductance is equal to the slope of the plate characteristic at the quiescent point. To facilitate the evaluation of this slope, a tangent to the plate characteristic at the quiescent point is constructed as shown in Fig. 7-6*b*. The slope is thus

$$g_p = \frac{10 - 0}{194 - 90} = \frac{10}{104} \approx 0.1 \text{ millimho}$$

and
$$r_p = 10 \text{ kilohms}$$

Finally, the transconductance is given by Eq. (7-12) as

$$g_m = \frac{\mu}{r_p} = \mu g_p = (20)(0.1) = 2 \text{ millimhos}$$

If an attempt is made to evaluate the transconductance directly from the plate characteristics, it is necessary to take relatively large increments of i_b and e_c along a line of constant e_b. Since the transconductance varies considerably with changes in i_b, this procedure is likely to yield inaccurate results.

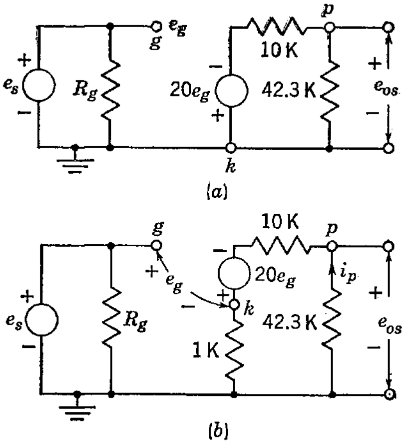


FIG. 7-7. Incremental models for the amplifier of Fig. 7-6. (*a*) With cathode bypass capacitor; (*b*) without cathode bypass capacitor.

c. The incremental model for the amplifier is shown in Fig. 7-7*a*. The output voltage from this circuit, e_{os}, is the signal component of the output voltage e_o shown in Fig. 7-6*a*. Since $e_g = e_s$, the output signal voltage, is

$$e_{os} = -\frac{42.3}{10 + 42.3}(20e_s) = -16.2e_s$$

and $$\frac{e_{os}}{e_s} = A = -16.2$$

d. With the bypass capacitor removed, the incremental model takes the form shown in Fig. 7-7*b*. In this case the output signal voltage is

$$e_{os} = -42.3i_p$$

and the loop equation for the plate circuit is

$$(42.3 + 10 + 1)i_p = 20e_g$$

But $$e_g = e_s - (1)i_p$$

Thus $$(53.3)i_p = 20e_s - 20i_p$$

$$73.3i_p = 20e_s$$

and $$e_{os} = (-42.3)(20/73.3)e_s = -11.5e_s$$

The amplification is thus

$$A = \frac{e_{os}}{e_s} = -11.5$$

Removing the cathode bypass capacitor reduces the voltage amplification by an appreciable amount.

The dependence of the incremental tube parameters on the quiescent operating point can be displayed in a variety of ways; one of these is shown in Fig. 7-8. The dotted lines superimposed on typical triode plate characteristics are lines of constant r_p; that is, the incremental plate resistance is the same at every point lying on any one of these lines. The plate resistance corresponding to each line is marked in kilohms at the end of the line. Since $\mu = g_m r_p$, it follows that in so far as μ is constant over the plate characteristic, the lines of constant r_p are also lines of constant g_m. The values of g_m corresponding to a constant value of 20 for μ are indicated in parentheses on Fig. 7-8. These values of g_m based on the assumption of a constant amplification factor are reasonably accurate except at low values of plate current.

It is well to be realistic about the values of the incremental tube parameters, no matter how they are determined. It has been stated previously that the characteristics of any particular tube of a given type may deviate considerably from the published characteristics. For example, values of g_m may differ by as much as 2:1 in two tubes of the same type. It should be noted also that careless drafting in the preparation

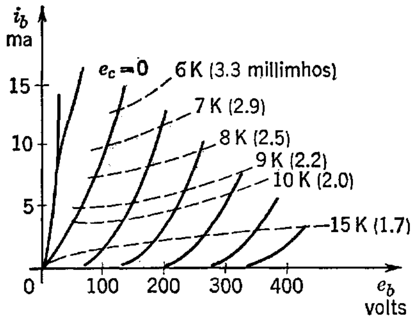


FIG. 7-8. Dependence of tube parameters on the quiescent operating point.

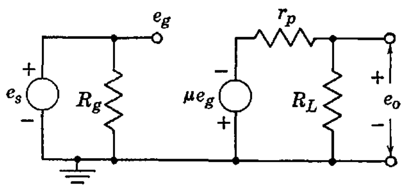


FIG. 7-9. Incremental model for a triode amplifier.

of characteristic curves is sometimes responsible for significant aberrations in these curves.

7-3. Choice of Quiescent Point. It is clear from the preceding sections that the performance of a vacuum tube depends upon the location of the quiescent point. Therefore it is appropriate to summarize the more important factors entering into the choice of the quiescent point. For Class A_1 operation the quiescent point must be chosen so that the signal voltage does not make the grid positive relative to the cathode and so that it does not cut the plate current off. In addition, it is often necessary to choose the quiescent point so that the operating path lies wholly in a region of low distortion.

Further considerations in the choice of the quiescent point follow from Fig. 7-9, which is the incremental model for a common triode amplifier. The voltage amplification of this circuit can be written by inspection. It is

$$A = \frac{e_o}{e_s} = \frac{-\mu R_L}{r_p + R_L} = \frac{-\mu}{1 + r_p/R_L} \tag{7-13}$$

To the extent that μ remains constant, maximum amplification is obtained by making the ratio r_p/R_L as small as possible. But for a fixed plate-supply voltage and a fixed grid bias, the larger the R_L, the smaller the quiescent plate current, and the larger r_p. Hence an optimum compromise among these variables must be sought. In some instances the greatest amplification is obtained with a quiescent point that lies at very low values of plate current and voltage. Such operation is called starved operation.

PROBLEMS

7-1. One section of a 12AX7 twin triode is used in the voltage-amplifier circuit shown in Fig. 7-4*a*. The plate-supply voltage is 250 volts, the source resistance is $R_s = 25$ kilohms, the grid return resistance is $R_g = 330$ kilohms, and the resistance connected across the output terminals is $R_o = 500$ kilohms. The problem is to choose R_L and R_k to give a specified quiescent operating point and to study the behavior of the circuit under small-signal conditions.

a. Determine the values of R_L and R_k that will locate the quiescent operating point at $e_b = 150$ volts and $i_b = 0.5$ ma.

b. Evaluate the three incremental triode parameters at the quiescent point.

c. Give an incremental linear model for the amplifier using the current source representation for the tube. The mode of operation is understood to be Class A_1, the bypass and coupling capacitors act as short circuits to the signal components of current and voltage, and the parasitic capacitances are negligibly small.

d. Determine the voltage amplification, $A = e_o/e_s$.

7-2. One section of a 12AT7 is used in the voltage-amplifier circuit of Fig. 7-4*a*. The circuit parameters and the supply voltage are $R_s = 500$ ohms, $R_g = 500$ kilohms, $R_k = 200$ ohms, $R_L = 10$ kilohms, $R_o = 100$ kilohms, and $E_{bb} = 350$ volts. The mode of operation is Class A_1, parasitic capacitances are negligibly small, and the bypass and coupling capacitors act as short circuits to the signal components of current and voltage. The behavior of the circuit is to be studied under small-signal conditions.

a. Determine the quiescent operating point, and evaluate the incremental tube parameters at this point.

b. The input-signal voltage is a train of pulses having a waveform similar to that shown in Fig. 5-28. The amplitude of each pulse is 50 mv. What is the amplitude of the pulses at the output of the amplifier?

c. How could the circuit be modified to give greater voltage amplification using the same tube?

7-3. There are important applications for each of the circuits shown in Fig. 7-10. The grid current in each tube is zero, and the parasitic capacitances are negligibly small. Give an incremental linear model for each circuit. Show clearly the controlling voltage for each controlled source. Do not assume that the bypass and coupling capacitors are short circuits.

7-4. The amplifiers of Figs. 7-10*a* and *c* are to be compared with respect to voltage amplification under the condition that $R_s = 0$.

a. Give an incremental model for each circuit with $R_s = 0$ and with the coupling and bypass capacitors treated as short circuits for signal components of current and voltage.

b. Derive an expression for the voltage amplification of each circuit in terms of the tube and circuit parameters.

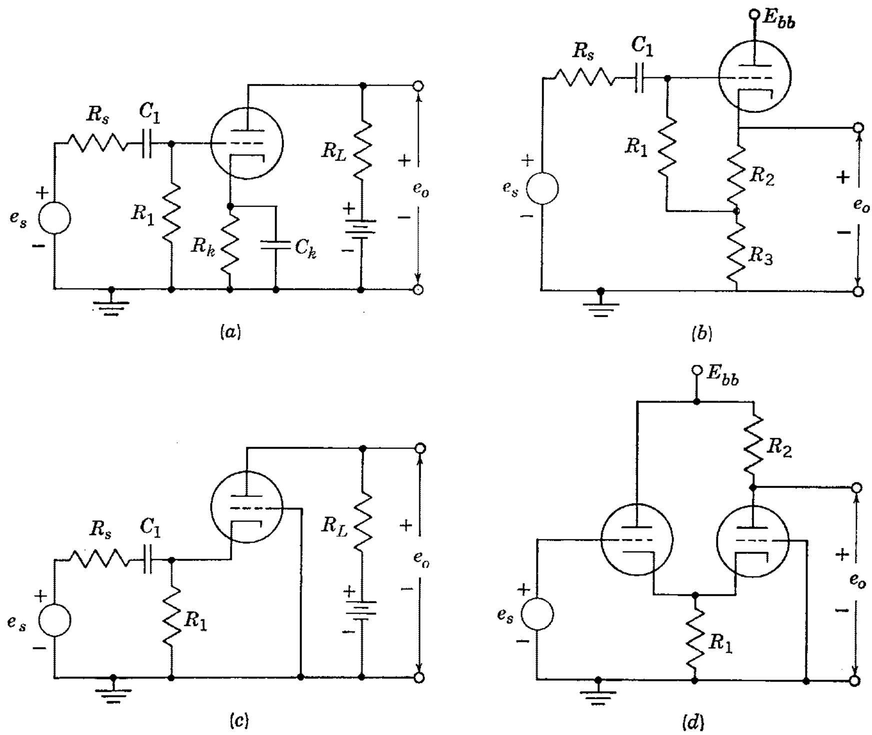


FIG. 7-10. Triode amplifier configurations for Prob. 7-3. (*a*) Grounded-cathode amplifier; (*b*) grounded-plate amplifier; (*c*) grounded-grid amplifier; (*d*) cathode-coupled amplifier.

7-5. One section of a 6SN7 twin triode is used in the amplifier shown in Fig. 6-15*b*. The tube parameters are $\mu = 20$ and $r_p = 10$ kilohms; the circuit parameters are $R = 25$ kilohms and $L = 4$ henrys. The signal voltage is $e_s = -5 + \cos 2000\pi t$.

a. What is the class of operation?

b. Give an incremental model for the amplifier. Parasitic capacitances are negligibly small.

c. Determine the amplitude and phase of the voltage across the inductor, using the sinusoidal component of e_s as the phase reference.

7-6. One section of a 12AX7 is used in the circuit shown in Fig. 7-4*a*. The circuit parameters and supply voltage are $R_s = 0$, $R_g = 1$ megohm, $R_k = 1$ kilohm, $R_o =$

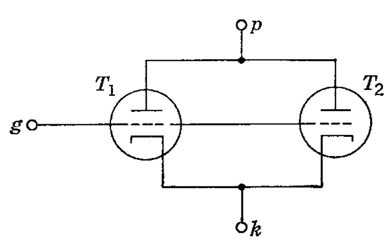


FIG. 7-11. Triodes in parallel for Prob. 7-7.

1 megohm, and $E_{bb} = 300$ volts. The problem is to investigate the dependence of the voltage amplification on the value of R_L.

a. Construct the grid bias line on the plate characteristics, and locate the quiescent operating point for $R_L = 100$, 200, 300, 600, and 1000 kilohms.

b. Evaluate r_p at each of the points located in part *a*.

c. Plot a curve of voltage amplification versus R_L for the range of R_L covered in part *a*. Assume that μ remains constant as R_L is changed.

d. For each value of R_L specified in part *a*, what is the greatest peak value that a sinusoidal input signal can have without making e_c positive?

7-7. Two triodes are connected in parallel as shown in Fig. 7-11. The incremental parameters for T_1 are $g_m = 2$ millimhos and $g_p = 0.1$ millimho; the parameters for T_2 and $\mu = 60$ and $r_p = 15$ kilohms. What are the incremental parameters for the composite tube formed by the parallel connection of T_1 and T_2?

CHAPTER 8

THE BASIC TRANSISTOR AMPLIFIER

The preceding chapters have shown that the vacuum triode is a physical device that behaves approximately like an ideal voltage amplifier. The transistor, which is a semiconductor triode, behaves very much like an ideal current amplifier; hence in many respects it acts like the dual of the vacuum triode. Both vacuum tubes and transistors can be used to amplify electrical signals in general. Vacuum tubes are superior in certain respects, and transistors are superior in others.

Transistors are inherently tiny devices, and they can operate effectively with bias voltages of only one or two volts. In addition, they have no power requirement corresponding to the cathode-heating power required by the vacuum tube. These facts are of great importance in the design of portable equipment such as hearing aids and mobile equipment such as automatic pilots, navigational aids, and communication radios for aircraft and small boats. They are also important in the design of large equipment such as electronic computers, in which amplifiers are used by the thousands. Another factor of importance where many amplifiers are required is the fact that the life of the transistor is unlimited, at least in principle, whereas that of the average vacuum tube is limited to a few thousand hours by deterioration of the cathode.

On the other hand, the transistor, being a semiconductor device, is quite sensitive to changes in temperature, and it cannot be used at all at temperatures above about 75°C in the case of germanium and 200°C in the case of silicon. Since the Zener voltage is ordinarily a few tens of volts, the transistor cannot be used where large signal voltages are required; they can, however, deliver large signal currents and substantial amounts of power.

Transistors are made in both the junction and the point-contact forms. However, since the internal mechanism of the point-contact transistor is not very well understood, and since the junction transistor is superior to the point-contact type in most applications, only the junction transistor is discussed in the following sections.

8-1. The Transistor as an Amplifier. The transistor is shown pictorially in the basic amplifier circuit in Fig. 8-1a; the schematic representation of this amplifier is shown in Fig. 8-1*b*. The physical make-up of the transistor is much like that of the semiconductor diode presented in Sec. 3-2. The *N-P-N* transistor shown in Fig. 8-1 consists of a thin slice of *P*-type semiconductor sandwiched between two pieces of *N*-type semiconductor; thus it has two *P-N* junctions like those discussed in Sec. 3-5. The behavior of the transistor depends primarily on the flow of current across these junctions.

In normal amplifier operation the lower junction in Fig. 8-1*a* is biased in the forward direction; hence electrons in the lower piece of *N*-type material, called the *emitter*, move in the forward direction across this

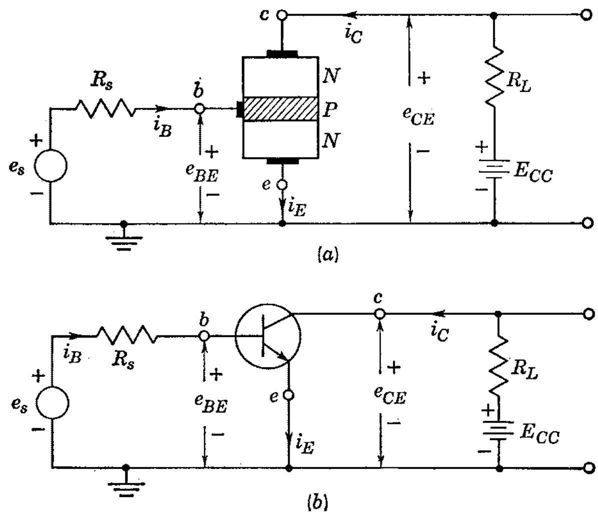


FIG. 8-1. The basic transistor amplifier circuit. (*a*) Pictorial representation; (*b*) schematic representation.

junction into the slice of *P*-type material, called the *base*. These electrons behave much like electrons generated by thermal agitation in the base, and they diffuse through the base in a random manner. In normal amplifier operation the upper junction is biased in the reverse direction; hence the free electrons in the base that wander into the vicinity of the upper junction are swept across the junction into the upper piece of *N*-type material, called the *collector*, creating a reverse current across this junction. Since the width of the base is very small, almost all the electrons injected by the emitter into the base are collected by the collector, and the collector current i_C is approximately equal to the electron current across the emitter junction. The base current i_B is approximately equal to the hole current across the emitter junction. This current can be made much smaller than the electron current by proper adjustment of the impurity concentrations in the emitter and base. Since the ratio of electron current to hole current across the emitter junction remains nearly constant over a wide range of currents, it follows that a small input current, i_B, can control a large output current, i_C, and current amplification is realized. Typical values of current amplification lie in the range between 20 and 100.

The *P-N-P* transistor is made by sandwiching a piece of *N*-type material between two pieces of *P*-type material. The behavior of the *P-N-P* transistor is essentially the same as that of the *N-P-N* transistor. The principal difference in so far as transistor circuits are concerned is that in order to bias the emitter junction in the forward direction and the collector junction in the reverse direction, the voltages applied to the *P-N-P* transistor must be opposite in polarity from those shown in Fig.

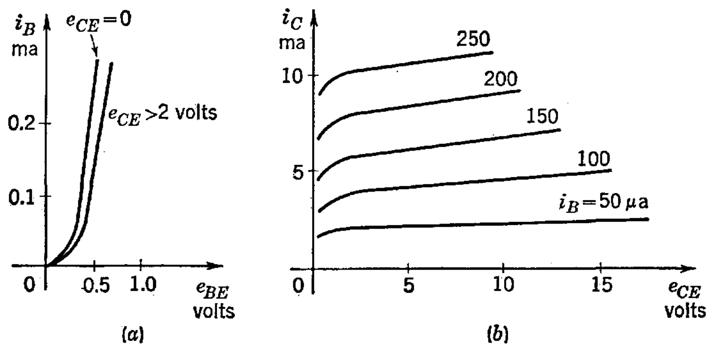


FIG. 8-2. Transistor characteristics. (*a*) Input characteristic, $i_B = f_B(e_{CE},e_{BE})$; (*b*) output characteristic, $i_C = f_C(e_{CE},i_B)$.

8-1. The *N-P-N* transistor is used in most of the discussion that follows for the reason that sign conventions are simpler and more natural than in *P-N-P* transistor circuits. It should be understood that the behavior of the two transistors is essentially identical in so far as signal components of voltage and current are concerned. The arrow at the emitter in the schematic representation of Fig. 8-1*b* indicates the direction of forward current for the emitter junction; hence it indicates the type of transistor.

The performance of the transistor as an amplifier can be examined further by comparing its input and output characteristics with the characteristics of the ideal current amplifier shown in Fig. 4-3. A typical set of experimentally measured transistor characteristics is shown in Fig. 8-2; the symbols used on these characteristics are defined in the circuit diagrams of Fig. 8-1. The input characteristic for any fixed value of positive collector voltage is essentially the same as the volt-ampere characteristic of a semiconductor diode; it deviates from the ideal input characteristic in that the input voltage is not zero. When the base is biased positively, however, the input voltage is relatively small.

The output characteristics in the first quadrant are very much like those of an ideal current amplifier. The fact that they are nearly uniformly spaced for equal increments of base current shows that the output current varies almost linearly with the input current. The fact that the output characteristics are rotated slightly from the horizontal indicates that the transistor has a finite output resistance. It should be noted that the dependence of the collector current on the base current varies considerably among different transistors of the same type. Hence the output characteristic for any particular transistor may be appreciably different from the published characteristics for that transistor type. The published characteristics are the average characteristics for the transistor type.

Transistors can be made in several ways. Grown-junction and alloy-junction transistors are made in the same way as grown-junction and

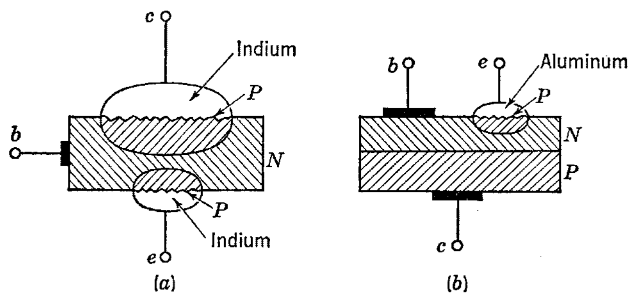


FIG. 8-3. Junction transistor structures. (*a*) Alloy junction; (*b*) diffused base.

alloy-junction diodes;[1] these processes are described briefly in Sec. 3-5. The structure of a grown-junction transistor is illustrated in Fig. 8-1*a*, and that of an alloy-junction transistor is shown in Fig. 8-3*a*.[2] Figure 8-3*b* shows the structure of still another type, known as the diffused-base transistor.[3,4] The base of this latter transistor is formed by diffusing a suitable donor impurity into the surface of a *P*-type wafer from the vapor state to form a thin *N*-type layer. This base layer is separated from the original wafer, which serves as the collector, by a *P-N* junction. The emitter and its *P-N* junction are formed by alloying a film of aluminum with the base layer. Forming the base of the transistor by this diffusion process provides two important features: it permits very thin base layers to be formed (of the order of one ten-thousandth of a centimeter), and it provides a nonuniform distribution of the impurity in the base layer. It will be seen that both of these factors act to enhance greatly the high-frequency performance of the transistor.[5]

8-2. Current Flow in Transistors. The flow of current in junction transistors can be studied with the aid of the potential-distribution diagrams in Fig. 8-4. With no voltages applied to the transistor, an equilibrium is established, as in the case of the junction diode, with a potential barrier across each *P-N* junction; the corresponding potential distribution is shown in Fig. 8-4*b*. If e_{CB} is then made positive, the collector junction is biased in the reverse direction, and the height of the potential barrier at the collector junction increases to $\phi_o + e_{CB}$ volts. A small reverse current flows across this junction as a result of carriers generated by thermal agitation in the collector and base and as a result of electrons in the emitter that have enough energy to pass over the barrier at the emitter junction.

If e_{BE} is now given a small positive value, the potential distribution takes the form shown in Fig. 8-4*c*, and the height of the potential barrier

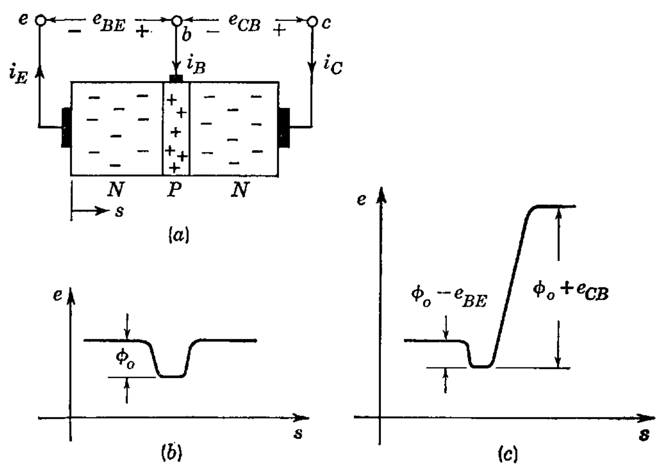


FIG. 8-4. Potential distributions in a transistor. (*a*) Transistor; (*b*) potential distribution for $e_{BE} = e_{CB} = 0$; (*c*) potential distribution for $0 < e_{BE} < e_{CB}$.

at the emitter junction decreases to $\phi_o - e_{BE}$ volts. The electron flow from the emitter into the base therefore increases, and practically all these electrons diffuse across the base and enter the collector; hence the reverse collector current increases. The component of current across the emitter junction associated with the flow of holes from the base into the emitter flows out of the emitter terminal, through the source of e_{BE}, and into the base terminal; the base current is approximately equal to the hole current flowing across the emitter junction.[1] As mentioned in Sec. 8-1, this current is made much smaller than the electron current across the emitter junction by controlling the impurity concentrations in the emitter and base.

The currents that flow in the transistor depend on the number of charge carriers that are able to cross the potential barriers in each second.

The voltage e_{BE} shown in Fig. 8-4a is a forward voltage impressed across the emitter junction. Hence the conventional current that would flow across this junction from the base to the emitter if the collector were removed is given by the volt-ampere law for the P-N junction [Eq. (3-29)]:

$$i_1 = I_1[\exp(\lambda e_{BE}) - 1] \tag{8-1}$$

where $\lambda = q_e/kT$. Similarly, the conventional current that would flow across the collector junction from the base to the collector if the emitter were removed is

$$i_2 = I_2[\exp(-\lambda e_{CB}) - 1] \tag{8-2}$$

The minus sign is required in the exponent because e_{CB} is a reverse voltage applied across the collector junction, whereas Eq. (3-29) gives the current in terms of a voltage applied in the forward direction.

The total current across the collector junction consists of the current i_2 and a portion of i_1. The first of these components is a conventional flow from base to collector. The second component, $\rho_1 i_1$, results from electrons that flow from the emitter into the base and then into the collector. The collector current i_C is the net flow of conventional current across the collector junction from collector to base; hence

$$\begin{aligned} i_C &= \rho_1 i_1 - i_2 \\ &= \rho_1 I_1[\exp(\lambda e_{BE}) - 1] - I_2[\exp(-\lambda e_{CB}) - 1] \end{aligned} \tag{8-3}$$

When e_{CB} is more positive than about 0.1 volt, as is normally the case, the second exponential term in (8-3) is much smaller than unity, and

$$i_C \approx \rho_1 I_1[\exp(\lambda e_{BE}) - 1] + I_2 \tag{8-4}$$

In a similar manner the emitter current is found to be

$$i_E = I_1[\exp(\lambda e_{BE}) - 1] - \rho_2 I_2[\exp(-\lambda e_{CB}) - 1] \tag{8-5}$$

$$\approx I_1[\exp(\lambda e_{BE}) - 1] + \rho_2 I_2 \tag{8-6}$$

where ρ_2 is the fraction of i_2 that flows into the emitter. The current flowing into the base terminal of the transistor is given by

$$\begin{aligned} i_B &= i_E - i_C \\ &= (1 - \rho_1)I_1[\exp(\lambda e_{BE}) - 1] + (1 - \rho_2)I_2[\exp(-\lambda e_{CB}) - 1] \end{aligned} \tag{8-7}$$

The collector current in a transistor results primarily from the diffusion of charge carriers by random motion across the base of the transistor. When a signal current is applied between the base and emitter terminals, the collector current does not respond until some time has elapsed. This fact seriously impairs the performance of the transistor at high frequencies. In addition, the capacitance across the collector junction, discussed in Sec. 3-5, tends to short-circuit the output at high frequencies.

In grown-junction and alloy-junction transistors these effects become appreciable at frequencies just above the audible band. To minimize these effects the base must be made as thin as possible to reduce the transit time of carriers across the base, and the depletion region at the collector junction must be made as wide as possible to reduce the collector capacitance. Both of these ends are accomplished in the diffused-base transistor, with the result that the behavior of the transistor is in some cases independent of frequency up to a few tens of megacycles per second in the basic amplifier circuit.

The superior high-frequency performance of the diffused-base transistor results from the following facts: First, the process by which base is formed permits a high degree of control over the thickness of the base layer; hence very thin base layers are possible. Secondly, the diffusion process results in a much higher concentration of impurity atoms on the emitter side of the base than on the collector side. This nonuniformity results in an electric field in the base that accelerates the charge carriers from the emitter to the collector, and a uniform drift toward the collector is thereby superimposed on the random diffusion of the carriers. For this reason these transistors are also called drift transistors. Finally, the small concentration of impurity atoms in the base near the collector results in a wide depletion layer at the collector junction and a correspondingly small collector capacitance.

The potential distribution in the drift transistor[5] and the mechanism by which it is developed are illustrated in Fig. 8-5. The concentration of acceptor atoms in the base, and hence the concentration of bound negative charges, is much greater at the emitter junction than at the collector junction. The free carriers in the base tend by the random motion of diffusion to assume a uniform concentration, and as a result there is an excess of negative charge near the emitter junction and an excess of positive charge near the collector junction. This charge distribution, indicated in Fig. 8-5a, sets up an electric field that accelerates free electrons across the base from the emitter to the collector. The nature of the potential distribution in the transistor is indicated in Fig. 8-5b. The excess of negative charge near the emitter junction depresses

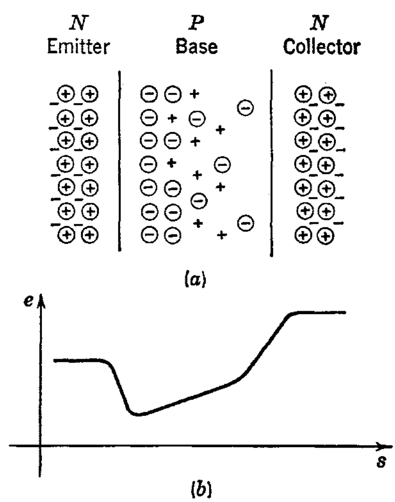


FIG. 8-5. Potential distribution in a drift transistor. (a) Charge distribution; (b) potential distribution.

the potential in that region, and the excess of positive charge near the collector junction raises the potential in that region. Free electrons in the base are accelerated toward the region of high potential.

A carrier-depletion region exists in the vicinity of each junction in the transistor in accordance with the discussion of the *P-N* junction in Sec. 3-5. As a result of the relatively small impurity concentration in the base at the collector junction, the depletion region extends relatively far into the base of the drift transistor. It follows from this fact that the capacitance across the collector junction of the drift transistor is quite small when a suitable reverse voltage (bias) is applied.

The power input to the transistor in the basic amplifier of Fig. 8-1 at the collector-to-emitter terminals is

$$p_C = e_{CE} i_C \tag{8-8}$$

This collector dissipation appears as heat in the transistor. Since i_B and e_{BE} are normally much smaller than i_C and e_{CE}, respectively, the power input at the base-to-emitter terminals is normally negligible in comparison with p_C. Since the transistor is damaged if the temperature becomes excessive, there is a maximum permissible collector dissipation that must not be exceeded. For typical small current-amplifier transistors, the collector dissipation must not exceed a few tens or hundreds of milliwatts. As in the case of the vacuum triode, the maximum permissible collector dissipation fixes a hyperbola on the collector-characteristic family; ordinarily the quiescent operating point for the transistor must lie below this hyperbola.

8-3. Graphical Analysis of the Basic Transistor Amplifier. The graphical analysis of the transistor amplifier is essentially identical with the graphical analysis of the vacuum triode. The circuit and the graphical construction are shown in Fig. 8-6. The base-to-emitter voltage is

$$e_{BE} = E_1 - R_1(i_B - i_s) \tag{8-9}$$

In normal operation e_{BE} is negligibly small in comparison with E_1; hence (8-9) gives i_B approximately as

$$i_B = \frac{E_1}{R_1} + i_s = I_B + i_s \tag{8-10}$$

where $I_B = E_1/R_1$ is the average value of i_B under the assumption that i_s is a time-varying current with zero average value.

The collector-to-emitter voltage is given by

$$e_{CE} = E_2 - R_L i_C \tag{8-11}$$

from which

$$i_C = \frac{E_2}{R_L} - \frac{e_{CE}}{R_L} = I_o - \frac{e_{CE}}{R_L} \tag{8-12}$$

This is the equation of the load line shown in Fig. 8-6*b*. The intersection of the load line with the collector characteristic corresponding to the value of i_B given by (8-10) is the operating point. Under quiescent conditions, $i_s = 0$ and $i_B = I_B = E_1/R_1$; the corresponding quiescent point is indicated in Fig. 8-6*b*.

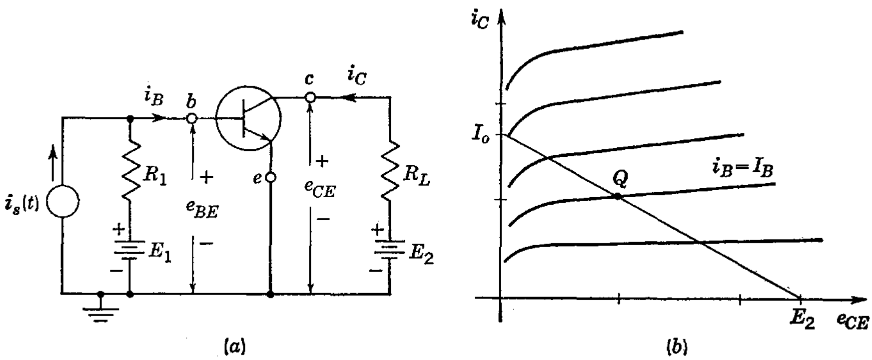


FIG. 8-6. Graphical analysis of the basic transistor amplifier. (*a*) Circuit; (*b*) graphical construction.

The current-transfer characteristic for the amplifier in Fig. 8-6*a* can be constructed by assuming various values of i_s and reading the resulting value of i_C from the load line in Fig. 8-6*b*. The transfer characteristic for a typical amplifier is shown in Fig. 8-7. When i_s is negative and larger than I_B, the emitter junction is biased in the reverse direction, and both i_B and i_C are essentially zero. This condition corresponds to the right-hand end of the load line in Fig. 8-6*b*. When i_s has large positive values, the collector saturates, and further increases in i_s do not produce further increases in i_C. This condition corresponds to the left-hand end of the load line. The central linear portion of the transfer characteristic is the region in which amplifiers normally operate.

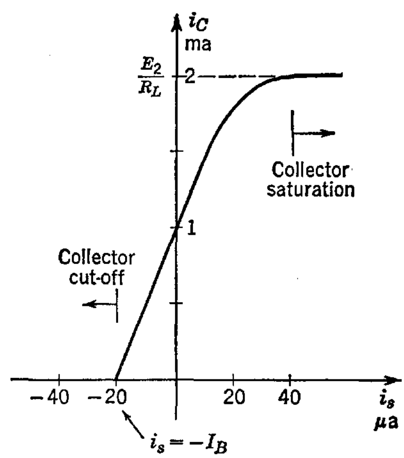


FIG. 8-7. Current transfer characteristic.

It is clear from the transfer characteristic that if the signal applied to the amplifier is too large, there will be waveform distortion as a result of collector cutoff and collector saturation. The discussion of distortion presented in connection with the vacuum triode applies qualitatively to the transistor as well.

Transistors can be operated in Class A, B, and C modes. In Class A the collector current is never cut off; in Class B the transistor is biased approximately at the cutoff point; and in Class C the transistor is biased beyond cutoff.

8-4. Power Relations in the Basic Transistor Amplifier. The power relations in the transistor amplifier are completely analogous to those in the vacuum-triode amplifier. The collector current and the collector-to-emitter voltage can be expressed as

$$i_C = I_C + i_c \qquad \text{and} \qquad e_{CE} = E_{CE} + e_{ce} \tag{8-13}$$

where I_C and E_{CE} are the average values of current and voltage, and i_c and e_{ce} are the time-varying components. The time-varying components have zero average value. The power drawn from the collector supply in the circuit of Fig. 8-6a at each instant is thus

$$p_{CC} = E_2 i_C = E_2 I_C + E_2 i_c \tag{8-14}$$

Since E_2 is constant and i_c has zero average value, the average value of $E_2 i_c$ is zero, and the average power drawn from the collector supply is

$$P_{CC} = E_2 I_C \tag{8-15}$$

If there is negligible waveform distortion, I_C and P_{CC} are both independent of the signal amplitude.

The power absorbed by the load at each instant is

$$\begin{aligned} p_L &= R_L i_C^2 = R_L(I_C + i_c)^2 \\ &= R_L I_C^2 + 2R_L I_C i_c + R_L i_c^2 \end{aligned} \tag{8-16}$$

The quantity $2R_L I_C$ is constant, and the average value of i_c is zero; hence the average value of the second term in (8-16) is zero, and the average power absorbed by the load is

$$P_L = R_L I_C^2 + R_L (i_c^2)_{\text{av}} = R_L I_C^2 + R_L (i_{c\,\text{rms}})^2 \tag{8-17}$$

The power dissipated by the transistor at each instant is

$$\begin{aligned} p_C &= e_{CE} i_C = (E_2 - R_L i_C) i_C = E_2 i_C - R_L i_C^2 \\ &= p_{CC} - p_L \end{aligned} \tag{8-18}$$

The average power dissipated by the transistor is therefore

$$P_C = P_{CC} - P_L \tag{8-19}$$

Thus in distortionless Class A operation, the collector dissipation is maximum under quiescent operating conditions, and it decreases as the signal level is increased.

8-5. Piecewise-linear Analysis. A piecewise-linear model that represents the transistor quite well for many purposes is shown in Fig. 8-8*a*. This model is quite simple, for in certain respects the transistor is a good approximation to the ideal current amplifier. The output characteristics, shown in Fig. 8-2, are nearly straight, horizontal lines, equally spaced for equal increments of base current; this fact is accounted for by the current-controlled current source in the model. Since the emitter and collector junctions act in somewhat the same manner as junction diodes, the base cannot have any appreciable positive potential relative to either the emitter or the collector (in *N-P-N* transistors). The action of these two junctions is accounted for, to a suitable approximation for most purposes, by the two ideal diodes in the model. In normal amplifier operation the emitter diode is biased in the forward

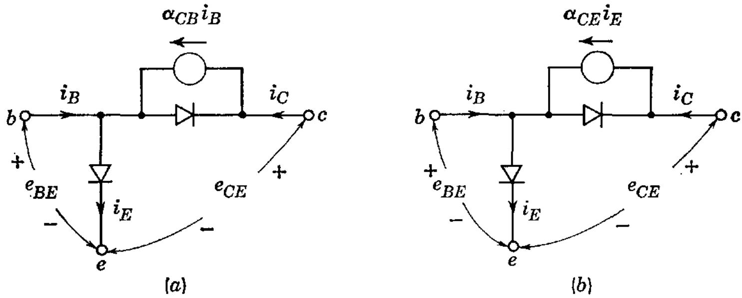


FIG. 8-8. Piecewise-linear models for the *N-P-N* transistor. (*a*) Source controlled by i_B; (*b*) source controlled by i_E.

direction, and the collector diode is biased in the reverse direction; hence in normal operation the former acts as a short circuit and the latter acts as an open circuit.

There is only one parameter, α_{CB}, to be determined in the piecewise-linear model for the transistor. The value of α_{CB} is chosen to make the collector current in the model have the same value as the collector current in the transistor at a suitable point on the collector characteristic. Thus if the base current is i_{Ba} and the collector current is i_{Ca} at the chosen point, then $\alpha_{CB} = i_{Ca}/i_{Ba}$ gives a satisfactory approximation. Typical values for α_{CB} lie in the range between 20 and 100.

The model of Fig. 8-8*a* is usually quite satisfactory for calculating total currents and voltages in transistor circuits. However, for certain purposes it is desirable to have the current of the controlled source expressed as a function of the emitter current i_E rather than the base current i_B. The currents in the model are related by

$$i_E = i_B + i_C \tag{8-20}$$

and

$$i_C = \alpha_{CB} i_B \tag{8-21}$$

Eliminating i_B between these two equations yields

$$i_C = \frac{\alpha_{CB}}{1 + \alpha_{CB}} i_E = \alpha_{CE} i_E \tag{8-22}$$

The quantity

$$\alpha_{CE} = \frac{\alpha_{CB}}{1 + \alpha_{CB}} \tag{8-23}$$

is the emitter-to-collector current-amplification factor, and

$$\alpha_{CB} = \frac{\alpha_{CE}}{1 - \alpha_{CE}} \tag{8-24}$$

is the base-to-collector current-amplification factor. Thus the model shown in Fig. 8-8b is equivalent to the one shown in Fig. 8-8a.

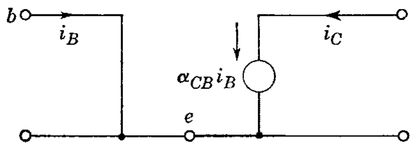


FIG. 8-9. Transistor model for linear Class A operation.

Piecewise-linear models for the *P-N-P* transistor are derived in the same manner. The results are the same except for the fact that the two diodes in the model are reversed. For linear Class A operation, the emitter diode acts as a short circuit and the collector diode acts as an open circuit in both *N-P-N* and *P-N-P* transistors. Hence for this mode

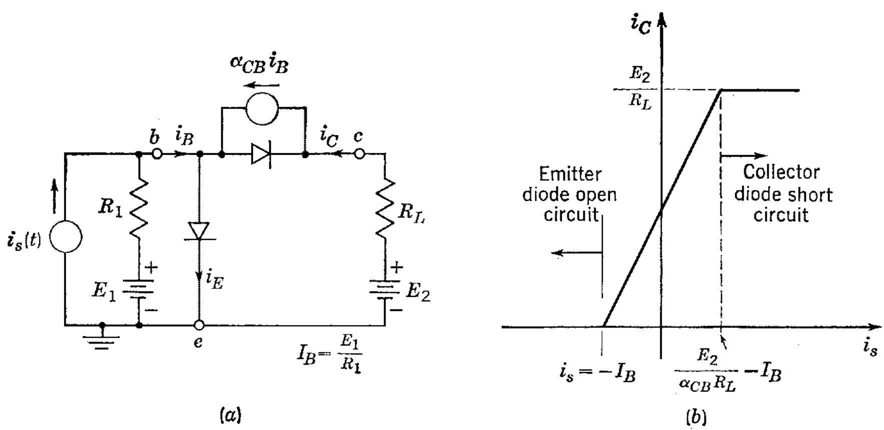


FIG. 8-10. Current transfer characteristic for the piecewise-linear model of a transistor amplifier. (a) Model; (b) transfer characteristic.

of operation the model of Fig. 8-9 represents both types of transistors. This model is identical with the ideal current amplifier; however, it is only an approximate representation for the transistor.

If the transistor in the amplifier of Fig. 8-6a is replaced by its piecewise-linear model, the circuit shown in Fig. 8-10a results. This circuit can be analyzed in a straightforward manner. The current-transfer characteristic for the amplifier can be constructed by determining the points at which the diodes change from the conducting to the nonconducting state; the result is shown in Fig. 8-10*b*. The left-hand break occurs when the voltage across the emitter diode and the current through both diodes are zero. Hence at this point the base terminal is at ground potential, and $i_B = -i_C = -\alpha_{CB}i_B$. The condition $i_B = -\alpha_{CB}i_B$ can be satisfied only if $i_B = 0$. It follows from these facts that at the left-hand break in the transfer characteristic $i_s = -E_1/R_1 = -I_B$.

The right-hand break in the characteristic occurs when the current through the collector diode and the voltage across both diodes are zero.

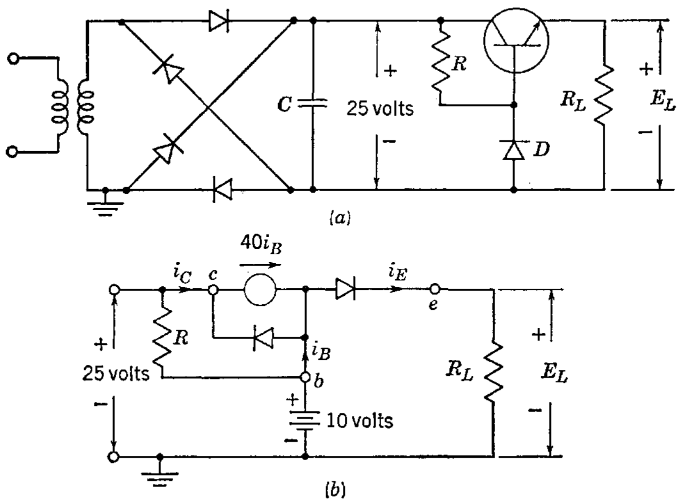


FIG. 8-11. Transistor-regulated power supply for Example 8-1. (*a*) Circuit; (*b*) model for regulator.

Hence at this point both the collector and the emitter are at ground potential, and $\alpha_{CB}i_B = i_C = E_2/R_L$. It follows from these facts that $i_B = E_2/\alpha_{CB}R_L$ and, since the base is at ground potential,

$$i_s = i_B - I_B = E_2/\alpha_{CB}R_L - I_B$$

The piecewise-linear transfer characteristic of Fig. 8-10*b* should be compared with the graphically determined characteristic shown in Fig. 8-7.

Example 8-1. A d-c power supply with a transistor voltage regulator is shown in Fig. 8-11*a*. The rectifier consists of four junction diodes connected in a bridge circuit. The diode *D* in the regulator is operated in the avalanche-breakdown region; it provides a voltage drop of 10 volts that is essentially independent of the current through the diode. Any change in load voltage causes a compensating change in the voltage drop across the transistor from collector to base, and the load voltage is thereby held nearly constant in spite of changes in input voltage or load resistance. Determine the load voltage, the transistor currents, and the voltage drop across the transistor when the load resistance is 500 ohms and when the current amplification factor for the transistor is $\alpha_{CB} = 40$.

Solution. For an approximate solution the transistor can be replaced with its piecewise-linear model as shown in Fig. 8-11*b*. The avalanche diode is represented by a 10-volt battery in this model; it makes the base of the transistor 10 volts positive relative to the ground. Thus if E_L is less than 10 volts, the emitter diode acts as a short circuit and current flows into the base, out of the emitter, and through the load. This current causes a current that is 40 times as great to flow into the collector. Hence the load voltage is held constant at 10 volts, and the voltage across the transistor from collector to base is $25 - 10 = 15$ volts.

The load current is

$$i_E = \frac{E_L}{R_L} = \frac{10}{0.5} = 20 \text{ ma}$$

The base current is given by

$$i_B + 40i_B = 20$$
$$i_B = 20/41 = 0.488 \text{ ma}$$

and the collector current is

$$i_C = i_E - i_B = 20 - 0.488 = 19.5 \text{ ma}$$

The true load voltage in the circuit of Fig. 8-11*a* is less than the voltage across the avalanche diode by the amount of the base-to-emitter voltage drop in the transistor. It is clear from the input characteristic of Fig. 8-2*a*, however, that in normal operation the load voltage is never more than a fraction of a volt less than the avalanche voltage.

8-6. Incremental Linear Models for Transistors. When considering total voltages and currents in the piecewise-linear model for the transistor amplifier, the small voltage drop between base and emitter is usually negligible in comparison with the applied bias voltage. When only the increments of voltage and current are considered, however, such a comparison is usually not possible, and the base-to-emitter voltage drop may have an appreciable effect on the performance of the circuit. Furthermore, when calculating increments of current and voltage it is usually desirable to account for the small effects of variations in collector voltage on both the input and output circuits. For these reasons the incremental model for the transistor is somewhat more complicated than the piecewise-linear approximation.

The base-to-emitter voltage and the collector current can be expressed as functions of the base current and the collector-to-emitter voltage:

$$e_{BE} = f_B(i_B, e_{CE}) \tag{8-25}$$

and

$$i_C = f_C(i_B, e_{CE}) \tag{8-26}$$

If i_B and e_{CE} are given small increments, the resulting increment in e_{BE} can be expressed as

$$\Delta e_{BE} \approx de_{BE} = \frac{\partial e_{BE}}{\partial i_B} di_B + \frac{\partial e_{BE}}{\partial e_{CE}} de_{CE} \tag{8-27}$$

The partial derivative in the first term of (8-27) has the dimensions of resistance, and that in the second term is a dimensionless voltage ratio. Replacing these derivatives by suitable symbols and using lower-case subscripts to denote incremental quantities yields

$$e_{be} = r_n i_b + \mu_{bc} e_{ce} \tag{8-28}$$

In a similar manner, the increment in i_C can be written as

$$\Delta i_C \approx di_C = \frac{\partial i_C}{\partial i_B} di_B + \frac{\partial i_C}{\partial e_{CE}} de_{CE} \tag{8-29}$$

Defining suitable symbols for the quantities in (8-29) and using lower-case subscripts to denote incremental quantities yields

$$i_c = \alpha_{cb} i_b + g_o e_{ce} \tag{8-30}$$

Equations (8-28) and (8-30), along with the relation $i_e = i_b + i_c$, imply the circuit shown in Fig. 8-12*a*. This circuit is an incremental network model for the transistor.

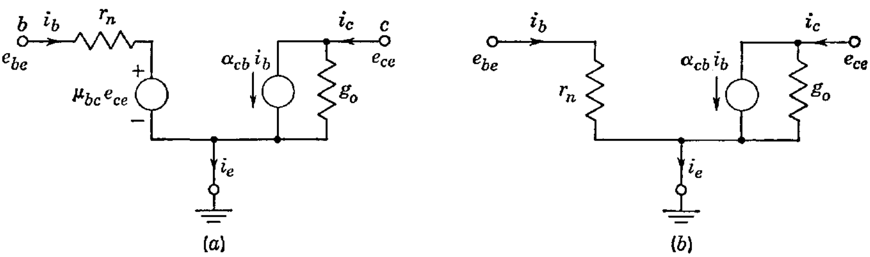


FIG. 8-12. The hybrid incremental model for the transistor. (*a*) Model; (*b*) model when μ_{bc} can be taken as zero.

The model of Fig. 8-12*a* is known as the hybrid model[1,6] for the transistor because a mixed set of voltages and currents is chosen as the set of independent variables in Eqs. (8-25) and (8-26). This model is convenient for several reasons, among which are the facts that the hybrid parameters are easily measured and that they are related in a simple way to the input and output characteristics shown in Fig. 8-2. The input resistance r_n is related to the slope of the input characteristic, and the output conductance g_o is related to the slope of the output characteristic. The reverse voltage amplification factor μ_{bc} is related to the horizontal displacement of the input characteristic resulting from a change in e_{CE}, and the forward current amplification factor α_{cb} is related to the vertical displacement of the output characteristic resulting from a change in i_B. These incremental parameters must, of course, be evaluated at the quiescent operating point; a change in the quiescent point calls for a reevaluation of the incremental parameters.

A further useful feature of the hybrid parameters is the fact that they give a measure of the extent to which the transistor approximates the ideal current amplifier. The parameters r_n, g_o, and μ_{bc} are all zero in the ideal case; to the extent that these parameters for a given transistor are negligibly small in comparison with the parameters of the external circuit, the transistor behaves as an ideal current amplifier. The values of the hybrid parameters for a typical alloy-junction transistor are $r_n = 2.5$ kilohms, $\mu_{bc} = 3 \times 10^{-4}$, $\alpha_{cb} = 50$, and $g_o = \frac{1}{50}$ millimho. For the drift transistor, μ_{bc} and g_o are much smaller (in theory, at least) than

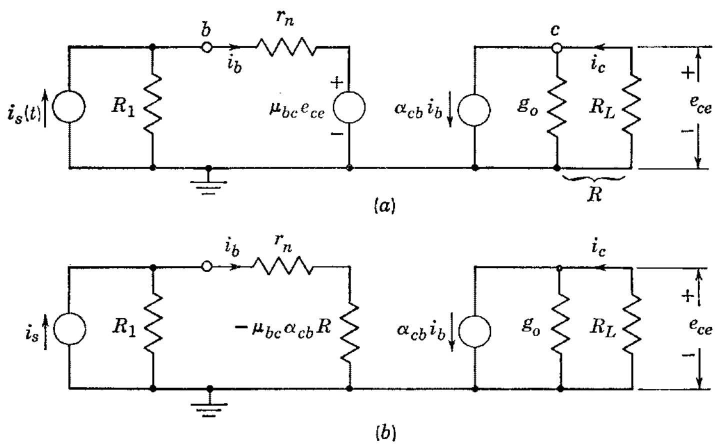


FIG. 8-13. Transistor-amplifier analysis. (*a*) Incremental model; (*b*) equivalent circuit.

for the alloy-junction transistor, while α_{cb} and r_n have about the same values as for the alloy-junction transistor. The reverse voltage-amplification factor is always a small number, and in many cases the voltage source $\mu_{bc}e_{ce}$ can be replaced by a short circuit in the incremental model. In such cases the model reduces to that shown in Fig. 8-12*b*, and the problem of circuit analysis is considerably simplified thereby.

The hybrid model provides a simple means for evaluating the effects of the reverse transmission through the transistor. Consider, for example, the circuit in Fig. 8-13*a*, which is an incremental model for the transistor amplifier of Fig. 8-6*a*. If the resistance R_L in parallel with g_o is designated by R, then the collector voltage is given by

$$e_{ce} = -\alpha_{cb}i_bR \tag{8-31}$$

Hence

$$\mu_{bc}e_{ce} = -\mu_{bc}\alpha_{cb}Ri_b \tag{8-32}$$

It follows that if the voltage source in the input circuit of the transistor model is replaced by a resistance $-\mu_{bc}\alpha_{cb}R$, the current and voltage relations in the input circuit are not changed; thus the input circuit can be represented as shown in Fig. 8-13*b*. It is clear from the Thevenin equivalent of the circuit connected to the negative resistance in Fig. 8-13*b* that if $\mu_{bc}\alpha_{cb}R$ is much smaller than $r_n + R_1$, the negative resistance can be replaced by a short circuit without affecting i_b appreciably. If i_b is unaffected by this change, the currents and voltages in the collector circuit are not affected. In this way a quick estimate can be made of the effect of neglecting the voltage source in the input circuit of the transistor model. When this source is negligible, the current amplification of the circuit is, by inspection,

$$A_c = i_c/i_s = \frac{R_1}{r_n + R_1} \alpha_{cb} \frac{r_o}{r_o + R_L} \tag{8-33}$$

where $r_o = 1/g_o$.

Values for the hybrid parameters at a typical quiescent operating point are often provided by the manufacturer of the transistor. The symbols used in the preceding paragraphs are not commonly employed in transistor specifications, however; the hybrid parameters are often designated[6] by $r_n = h_{11}$, $\mu_{bc} = h_{12}$, $\alpha_{cb} = h_{21}$, and $g_o = h_{22}$. These symbols, along with the additional definitions $e_{be} = e_1$, $i_b = i_1$, $e_{ce} = e_2$, and $i_c = i_2$, result in a symmetrical form for Eqs. (8-28) and (8-30):

$$e_1 = h_{11}i_1 + h_{12}e_2 \tag{8-34}$$
$$i_2 = h_{21}i_1 + h_{22}e_2 \tag{8-35}$$

Alternatively, the hybrid parameters may be designated by $r_n = h_{ie}$, $\mu_{bc} = h_{re}$, $\alpha_{cb} = h_{fe}$, and $g_o = h_{oe}$. The letter *e* in the subscripts indicates that the parameters apply for the transistor operated with its emitter terminal grounded, as in the basic amplifier circuit of Fig. 8-6. A similar set of symbols with the letter *b* replacing the letter *e* designates the hybrid parameters for the transistor operated with its base terminal grounded and the input signal applied to the emitter. The value of the symbols used in Eqs. (8-28) and (8-30) lies in the fact that they show at a glance the dimensions of the parameters when it is understood that μ and α always represent dimensionless amplification factors.

It is often necessary to evaluate the transistor parameters for a particular quiescent operating point, for the parameters may vary appreciably from one quiescent point to another and from one transistor to another of the same type. The parameters can be evaluated from the characteristic curves, or they can be measured experimentally. If the parameters are to be determined experimentally, it is important to understand their significance in terms of open-circuit and short-circuit measurements. If an incremental current i_b is applied between the base and emitter terminals in the model of Fig. 8-12*a* with the collector terminal short-circuited to the emitter, then $e_{ce} = 0$, and the input voltage is $e_{be} = r_n i_b$. Thus r_n is the input resistance with the output short-circuited. Under these same conditions the current in the short circuit between the collector and emitter terminals is $i_c = \alpha_{cb} i_b$. Hence α_{cb} is the forward current amplification factor with the output short-circuited. In a similar manner it can be seen that g_o is the output conductance with the input terminals open-circuited and that μ_{bc} is the reverse voltage amplification factor with the input terminals open-circuited.

Example 8-2. A transistor amplifier is shown in Fig. 8-14*a*. The input signal is i_s, a small time-varying current, and the signal component of current in the load is i_{Ls}.

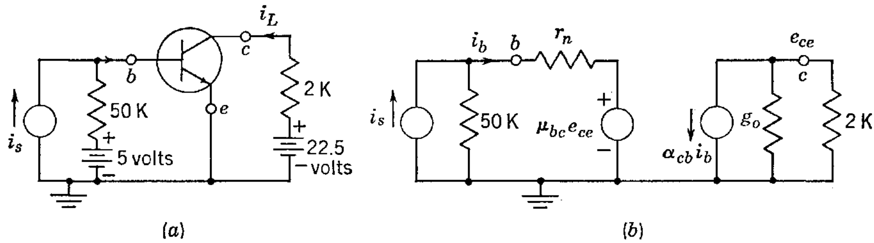


FIG. 8-14. Transistor amplifier for Example 8-2. (*a*) Circuit; (*b*) incremental model.

The transistor parameters are $\alpha_{cb} = 50$, $r_n = 2.5$ kilohms, $\mu_{bc} = 3 \times 10^{-4}$, and $g_o = 0.02$ millimho. Determine the small-signal current amplification i_{Ls}/i_s.

Solution. The incremental model for the amplifier is shown in Fig. 8-14*b*. The effect of the reverse voltage-amplification factor on the behavior of the circuit is to be examined first. The net resistance connected across the output terminals is the resistance of R_L in parallel with g_o:

$$R = \frac{r_o R_L}{r_o + R_L} = \frac{(50)(2)}{50 + 2} = 1.92 \text{ kilohms}$$

The equivalent resistance reflected into the input circuit is

$$-\mu_{bc}\alpha_{cb}R = -(3)(10^{-4})(50)(1.92) = -0.0282 \text{ kilohm}$$

This resistance is much less than 1 per cent of $r_n + R_1 = 2.5 + 50 = 52.5$ kilohms; hence its effect on the behavior of the circuit is entirely negligible.

The signal component of the load current is then

$$i_{Ls} = \frac{50 i_s}{50 + 2.5}(50)\frac{50}{50 + 2} = 45.8 i_s$$

and the current amplification is

$$A_c = \frac{i_{Ls}}{i_s} = 45.8$$

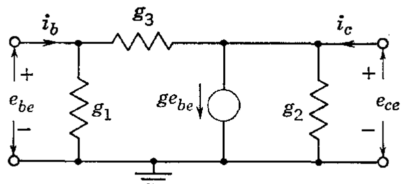


FIG. 8-15. The π model for the transistor.

The hybrid model of Fig. 8-12 is especially convenient when the reverse voltage-amplification factor is negligibly small. When this parameter is not negligible, it is always possible, and often desirable, to replace the hybrid model with an equivalent model having only one controlled source. The π model shown in Fig. 8-15 is one such that is frequently useful; in particular, it is useful in representing the dependence of the transistor behavior on frequency at high frequencies.

The π parameters can be expressed in terms of the hybrid parameters by equating the short-circuit input and transfer conductances for the two models. These relations can often be simplified by the fact that, as a rule, $g_3 \ll g_2 \ll g_1 \ll g$. Thus, with the output terminals short-circuited, $e_{ce} = 0$, and the π model yields

$$\left.\frac{i_b}{e_{be}}\right|_{e_{ce}=0} = g_1 + g_3 \tag{8-36}$$

Under these conditions the hybrid model yields

$$\left.\frac{i_b}{e_{be}}\right|_{e_{ce}=0} = \frac{1}{r_n} \tag{8-37}$$

If the models are to be equivalent, these ratios must be equal, and

$$g_1 + g_3 = \frac{1}{r_n} \tag{8-38}$$

Using the inequalities listed above, (8-38) becomes

$$g_1 \approx \frac{1}{r_n} \tag{8-39}$$

Equating the short-circuit transfer conductances in the forward direction yields

$$\left.\frac{i_c}{e_{be}}\right|_{e_{ce}=0} = g - g_3 = \frac{\alpha_{cb}}{r_n} \tag{8-40}$$

or

$$g \approx \frac{\alpha_{cb}}{r_n} \tag{8-41}$$

Equating the short-circuit transfer conductances in the reverse direction yields

$$\left.\frac{i_b}{e_{ce}}\right|_{e_{be}=0} = -g_3 = \frac{-\mu_{bc}}{r_n} \tag{8-42}$$

or

$$g_3 = \frac{\mu_{bc}}{r_n} \tag{8-43}$$

Finally, equating the short-circuit output conductances of the two circuits yields

$$\left.\frac{i_c}{e_{ce}}\right|_{e_{be}=0} = g_2 + g_3 = g_o - \frac{\alpha_{cb}\mu_{bc}}{r_n} \tag{8-44}$$

or

$$g_2 \approx g_o - \frac{\alpha_{cb}\mu_{bc}}{r_n} \tag{8-45}$$

When the reverse voltage transmission through the transistor is so small that μ_{bc} can be taken as zero, Eq. (8-43) yields $g_3 = 0$. Under these conditions the π and hybrid models become identical except for the controlling quantities for the controlled sources in the models. This is usually the case when the model represents a drift transistor.

The T model shown in Fig. 8-16 is another widely used representation for the transistor. The resistances in this model can be related directly to the base, emitter, and collector of the transistor; historically, it was the first model used to represent the transistor. The parameters of the T model can be expressed in terms of the hybrid parameters by equating the open-circuit input and transfer resistances for the two models. These relations can often be simplified by the fact that, as a rule, $r_e \ll r_b \ll r_d$. Thus, equating the open-circuit output resistances of the two circuits yields

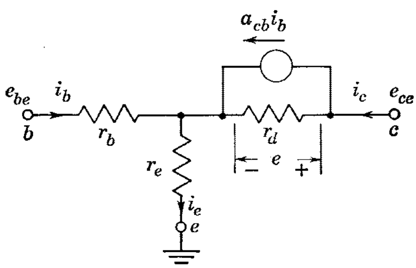


FIG. 8-16. The T model for the transistor.

$$\left.\frac{e_{ce}}{i_c}\right|_{i_b=0} = r_d + r_e = \frac{1}{g_o} \tag{8-46}$$

whence

$$r_d \approx \frac{1}{g_o} \tag{8-47}$$

Equating the open-circuit transfer resistances in the reverse direction gives

$$\left.\frac{e_{be}}{i_c}\right|_{i_b=0} = r_e = \frac{\mu_{bc}}{g_o} \tag{8-48}$$

or

$$\frac{r_e}{r_d} \approx \mu_{bc} \tag{8-49}$$

Equating the open-circuit input resistances yields

$$\left.\frac{e_{be}}{i_b}\right|_{i_c=0} = r_b + r_e = r_n - \frac{\alpha_{cb}\mu_{bc}}{g_o} \tag{8-50}$$

from which

$$r_b \approx r_n - \frac{\alpha_{cb}\mu_{bc}}{g_o} \tag{8-51}$$

Finally, equating the open-circuit transfer resistances in the forward direction yields

$$\left.\frac{e_{ce}}{i_b}\right|_{i_c=0} = r_e - a_{cb}r_d = \frac{-\alpha_{cb}}{g_o} \tag{8-52}$$

Using the inequalities listed above, along with Eq. (8-47), this relation reduces to

$$a_{cb} \approx \frac{\alpha_{cb}}{g_o r_d} \approx \alpha_{cb} \tag{8-53}$$

Thus a_{cb} is nearly equal to the short-circuit current amplification factor α_{cb}; it is customary to ignore the distinction between these two quantities and to use the symbol α_{cb} in both cases.

It is sometimes convenient to have the controlled source in the T model expressed in terms of its dependence on the emitter current i_e rather than the base current i_b. The transistor currents are related by

$$i_e = i_b + i_c \tag{8-54}$$

Also, from Fig. 8-16, the collector current is

$$i_c = \alpha_{cb} i_b + \frac{e}{r_d} \tag{8-55}$$

where the distinction between a_{cb} and α_{cb} is dropped. Eliminating i_b between (8-54) and (8-55) leads to

$$i_c = \frac{\alpha_{cb}}{1 + \alpha_{cb}} i_e + \frac{e}{(1 + \alpha_{cb})r_d} \tag{8-56}$$

Defining two new symbols, $\alpha_{ce} = \alpha_{cb}/(1 + \alpha_{cb})$ and $r_c = (1 + \alpha_{cb})r_d$, gives

$$i_c = \alpha_{ce} i_e + \frac{e}{r_c} \tag{8-57}$$

Thus the current source and shunt resistance in the model of Fig. 8-16 can be replaced by the equivalent current source and shunt resistance shown in Fig. 8-17. The quantity α_{ce} is the emitter-to-collector current amplification factor.

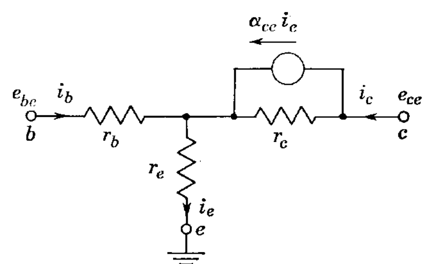


FIG. 8-17. Alternative form for the T model.

8-7. Choice of the Quiescent Operating Point. The incremental parameters of the transistor depend on the location of the quiescent operating point on the transistor characteristics. The dependence of the hybrid parameters of a typical small alloy-junction transistor on the quiescent emitter current is shown by the curves[1] of Fig. 8-18. The best approximation to an ideal current amplifier is obtained when r_n, μ_{bc}, and g_o are as small as possible and α_{cb} is as large as possible. However, since the parameters do not all change in the same way with changes in I_E, conflicting requirements arise, and a compromise must be made. This compromise depends on the nature of the external circuit and the magnitude of the supply voltage available as well as on the transistor itself. As a general rule, the optimum quiescent point for small-signal amplifiers corresponds to an emitter current of 1 or 2 ma. For large-signal operation this statement is not true, of course, for in Class A operation the quiescent emitter current must be at least as large as the peak output

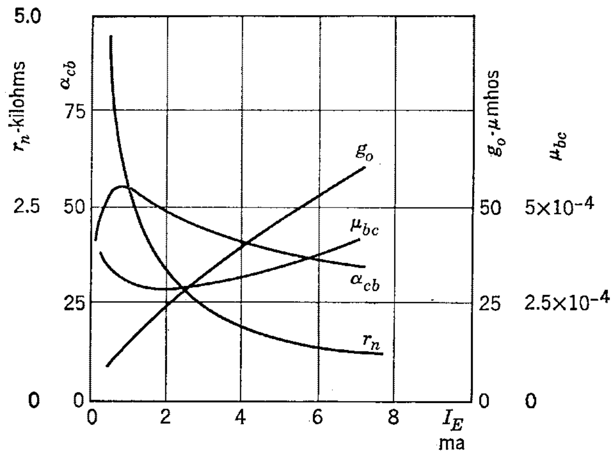


FIG. 8-18. Dependence of the hybrid parameters of an alloy-junction transistor on emitter current.

signal current. Additional factors enter into consideration when the transistor is to be operated at high frequencies; these are discussed in the following section.

8-8. High-frequency Transistor Models.[6] The incremental models presented in Sec. 8-6 do not account for the high-frequency effects in the transistor. Therefore these models are valid only at frequencies less than about 1 megacycle in the case of drift transistors and less than about 10 or 20 kilocycles in the case of alloy-junction transistors. The hybrid model can be modified easily to account for the high-frequency effects in the transistor when $\mu_{bc} = 0$, as usually is the case with the drift transistor. The resulting model, shown in Fig. 8-19*a*, provides a suitable representation of the drift transistor at frequencies up to some tens of megacycles.

The first step in developing the model of Fig. 8-19*a* from the hybrid model of Fig. 8-12*b* is to separate the short-circuit input resistance into two parts, r_b' and r_{be}. The resistance r_b' accounts for that portion of the base resistance that is associated with the base-lead connection and the part of the base that does not lie in the active region between the emitter and collector. The terminal b' in the model is the internal base terminal; it is not available for measurements or for connection to external circuitry. The model is then completed by adding the capacitors C_e and C_c as shown in Fig. 8-19*a*. The capacitance C_e, the emitter storage capacitance, accounts for the diffusion and transit-time effects at the emitter junction; the capacitance C_c is the capacitance across the depletion region at the collector junction. Typical values for these capacitances are $C_c = 2$ μμf and $C_e = 200$ μμf for drift transistors, and $C_c = 30$ μμf and $C_e = 3000$ μμf for alloy-junction transistors.

It is important to note that the controlled source in the circuit of Fig. 8-19a depends on i, the current through r_{be}, rather than on the input current i_b. At low frequencies, where C_e and C_c act as open circuits, $i = i_b$; at high frequencies, however, these two currents are not equal. In particular, if the input current is a sinusoid of constant amplitude, then the amplitude of the current i is a function of the signal frequency. The short-circuit current amplification of the transistor at high frequencies can be evaluated from the circuit of Fig. 8-19a for sinusoidal

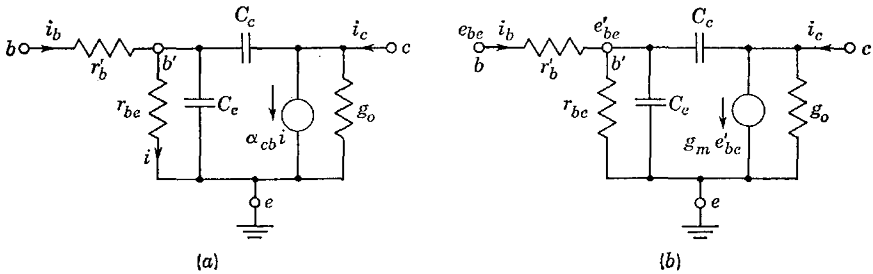


FIG. 8-19. High-frequency models for the drift transistor. (a) Current-controlled source; (b) voltage-controlled source.

operating conditions. For a typical transistor, C_e is about 100 times as big as C_c; therefore when the output terminals are short-circuited, the current in C_c is negligible in comparison with i_b and i_c at frequencies where the model of Fig. 8-19a is valid. Thus if the complex amplitude of a sinusoidal input current is designated $\mathbf{I}_b$, the complex amplitude of the sinusoidal current in r_{be} is, when the output terminals are short-circuited,

$$\mathbf{I} = \frac{1}{1 + j\omega C_e r_{be}} \mathbf{I}_b \tag{8-58}$$

The short-circuit output current is then

$$\mathbf{I}_c = \alpha_{cb}\mathbf{I} = \frac{\alpha_{cb}}{1 + j\omega C_e r_{be}} \mathbf{I}_b \tag{8-59}$$

Thus, if the amplitude of the input current is held constant as the frequency of the current is increased, the output current decreases and tends to zero at very high frequencies. This result is associated with the fact that at high frequencies the signal current tends to flow through C_e rather than through r_{be}.

The quantity $1/C_e r_{be}$ has the dimensions of frequency, and it has a useful interpretation as a frequency. Thus, defining a new symbol, $\omega_{cb} = 1/C_e r_{be}$, yields for the short-circuit output current

$$\mathbf{I}_c = \alpha_{cb}\mathbf{I} = \frac{\alpha_{cb}}{1 + j\omega/\omega_{cb}} \mathbf{I}_b \tag{8-60}$$

The amplitude of this current is

$$I_c = \frac{\alpha_{cb}}{\sqrt{1 + (\omega/\omega_{cb})^2}} I_b \tag{8-61}$$

Thus at the frequency $\omega = \omega_{cb}$, I_c is reduced to $1/\sqrt{2}$ times its low-frequency value; ω_{cb} is known as the cutoff frequency for the short-circuit forward current amplification. This cutoff frequency, which is also referred to as the beta cutoff frequency and symbolized ω_β, provides a useful means for specifying the bandwidth over which the short-circuit current amplification remains reasonably close to its low-frequency value. Typical values of ω_{cb} are of the order of 20 kcps for alloy-junction transistors and of the order of 1 megacycle for drift transistors; the theoretical limit for drift transistors appears to be of the order of several tens of megacycles.[5]

When the output terminals of the transistor are not short-circuited, the effects of the collector capacitance C_c are amplified by the transistor (Sec. 4-4) and may not be negligible; in such a case, the high-frequency performance of the transistor depends on both C_e and C_c. Both of these capacitances decrease with increasing collector voltage; hence, in general, the high-frequency cutoff depends on the choice of the quiescent operating point.

An alternative form for the high-frequency transistor model, shown in Fig. 8-19*b*, is often more convenient in circuit analysis than the model of Fig. 8-19*a*. The only difference between the two circuits is that the source in Fig. 8-19*b* is controlled by the voltage across r_{be} rather than by the current through r_{be}. The current in r_{be} is

$$i = \frac{e'_{be}}{r_{be}} \tag{8-62}$$

and the current of the controlled source in Fig. 8-19*a* is

$$\alpha_{cb} i = \frac{\alpha_{cb}}{r_{be}} e'_{be} \tag{8-63}$$

Thus the controlled source in Fig. 8-19*b* is equivalent to the one in Fig. 8-19*a* if

$$g_m = \frac{\alpha_{cb}}{r_{be}} \tag{8-64}$$

The models of Fig. 8-19 are applicable when the reverse voltage amplification factor μ_{bc} is zero. When this amplification factor is not negligible, as often is the case with alloy-junction transistors, a voltage source $\mu_{bc} e_{ce}$ appears in series with r_{be}. However, this voltage source can be eliminated by the use of a π model, like that in Fig. 8-15, for the *intrinsic* transistor, that part of the transistor exclusive of r'_b. The resulting model is shown in Fig. 8-20.

Equation (8-60) gives the short-circuit output current from the transistor $\mathbf{I}_c$ as a function of the base current $\mathbf{I}_b$ under sinusoidal operating conditions. An alternative expression for $\mathbf{I}_c$ can be obtained from the fact that the collector, base, and emitter currents are related by

$$\mathbf{I}_e = \mathbf{I}_c + \mathbf{I}_b \tag{8-65}$$

Using (8-65) to eliminate I_b in (8-60) and collecting terms yields

$$\mathbf{I}_c = \frac{\alpha_{cb}}{1 + \alpha_{cb} + j\omega/\omega_{cb}} \mathbf{I}_e \tag{8-66}$$

$$= \frac{\alpha_{cb}}{1 + \alpha_{cb}} \frac{1}{1 + j\omega/\omega_{cb}(1 + \alpha_{cb})} \mathbf{I}_e \tag{8-67}$$

Defining a new quantity, $\omega_{ce} = (1 + \alpha_{cb})\omega_{cb}$, and recalling that

$$\alpha_{ce} = \frac{\alpha_{cb}}{(1 + \alpha_{cb})}$$

leads to

$$\mathbf{I}_c = \frac{\alpha_{ce}}{1 + j\omega/\omega_{ce}} \mathbf{I}_e \tag{8-68}$$

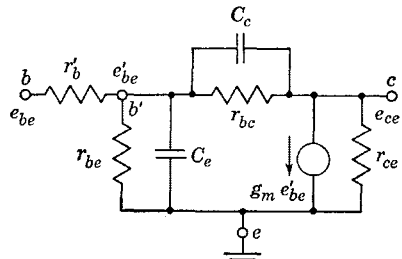


FIG. 8-20. High-frequency model for alloy-junction transistors.

Thus the quantity $\omega_{ce} = (1 + \alpha_{cb})\omega_{cb}$ is the cutoff frequency for the transistor when the amplitude of the emitter current is held constant; it is also referred to as the alpha cutoff frequency and symbolized ω_α. Since α_{cb} usually lies in the range between 20 and 100, ω_{ce} is much larger than ω_{cb}. The value of ω_{ce} is often given in transistor specifications as an indication of the high-frequency performance of the device.

A complete set of parameter values for the high-frequency π model of a typical small alloy-junction transistor is $r_b' = 250$ ohms, $r_{be} = 2600$ ohms, $r_{ce} = 220$ kilohms, $r_{bc} = 5$ megohms, $C_e = 4500$ $\mu\mu$f, $C_c = 17$ $\mu\mu$f, and $g_m = 21$ millimhos. The corresponding set of parameter values for a typical small drift transistor is $r_b' = 40$ ohms, $r_{be} = 1550$ ohms, $r_{ce} = \infty$, $r_{bc} = \infty$, $C_e = 200$ $\mu\mu$f, $C_c = 1.7$ $\mu\mu$f, and $g_m = 37$ millimhos.

8-9. Summary. The transistor is a physical amplifier that behaves somewhat as the dual of the vacuum triode. The emitter injects charge carriers into the base of the transistor, and the magnitude of the current flowing from the emitter into the base is controlled by the height of the potential barrier across the emitter junction. In the alloy-junction transistor the carriers injected into the base by the emitter move by diffusion to the collector and give rise to a reverse collector current that is controlled by the height of the potential barrier at the emitter junction. In the drift transistor a built-in electric field in the base region accelerates the carriers across the base toward the collector, thereby reducing the transit time of the charges crossing the base and improving the high-frequency response of the transistor. The transistor can provide current amplification because the collector current is much larger than the base current and is almost directly proportional to it. The transistor also provides voltage and power amplification.

The basic transistor amplifier has the same general form as the basic vacuum-triode amplifier, and it can be analyzed by graphical methods that are fundamentally the same as those used with the vacuum-triode amplifier. Alternatively, total voltages and currents in the transistor amplifier can be calculated approximately with the aid of a simple piecewise-linear model for the transistor.

When the transistor is operated in the linear Class A mode, the signal components of voltage and current can be calculated with the aid of an incremental linear model for the transistor. A variety of incremental models for the transistor can be formulated. Three of these, the hybrid model, the T model, and the π model, are the most useful.

Only alloy-junction and drift transistors are discussed specifically in the preceding sections. Transistors are made in a variety of other types,[1] among which are grown-junction, surface-barrier, and four-terminal, or tetrode, transistors. The basic principles and the methods of analysis are essentially the same for all these transistors. In so far as circuit analysis is concerned, the principal differences among the various transistor types lie in the parameter values to be used in the models representing them.

REFERENCES

1. Hunter, L. P.: "Handbook of Semiconductor Electronics," McGraw-Hill Book Co., Inc., New York, 1956.
2. Herold, E. W.: New Advances in the Junction Transistor, in "Transistors I," p. 27, RCA Laboratories, Princeton, N.J., 1956.
3. Lee, C. A.: A High Frequency Diffused Base Germanium Transistor, *Bell System Tech. J.*, vol. 35, p. 23, January, 1956.
4. Thomas, D. E., and G. C. Dacey: Application Aspects of the Germanium Diffused Base Transistor, *Trans. IRE*, vol. CT-3, p. 22, March, 1956.
5. Kroemer, H.: The Drift Transistor, in "Transistors I," p. 202, RCA Laboratories, Princeton, N.J., 1956.
6. Lo, A. W., R. O. Endres, J. Zawals, F. D. Waldhauer, and C. Cheng: "Transistor Electronics," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1955.

PROBLEMS

8-1. A 2N170 *N-P-N* transistor is used in the basic amplifier circuit shown in Fig. 8-6. The supply voltages are $E_1 = E_2 = 12.6$ volts, and the circuit resistances are $R_1 = 252$ kilohms and $R_L = 4.2$ kilohms.

a. The output volt-ampere characteristic curves for the transistor are approximately horizontal lines with $i_C = 30i_B$. Sketch a family of these curves for $i_B = 20$, 40, 60, and 80μa. Construct the load line for the amplifier on these characteristics.

b. Determine the quiescent collector current and collector-to-emitter voltage. Is the maximum permissible collector dissipation of 25 mw exceeded?

c. Determine graphically and plot i_C versus i_s for $-75 < i_s < 75$ μa. *Note:* See Eq. (8-10).

d. Repeat part *c* for the condition that the battery voltages are both reduced to 6.3 volts. Plot this curve on the same coordinates with the curve of part *c*.

8-2. A 2N104 *P-N-P* transistor is used in the basic amplifier circuit shown in Fig. 8-6. The supply voltages are $E_1 = E_2 = -22.5$ volts. The circuit resistances are to be chosen to give a specified quiescent operating point.

a. The output volt-ampere characteristic curves for the transistor are approximately horizontal lines with $i_C = 45i_B$. Sketch a family of these curves for $i_B = -25, -50, -75, -100$, and -125 μa.

b. Determine the values of R_1 and R_L required to locate the quiescent operating point at $i_C = -2$ ma and $e_{CE} = -10$ volts. Is the maximum permissible collector dissipation of 35 mw exceeded at the quiescent point?

c. If the input signal is a sinusoidal current, approximately what is the greatest amplitude that the signal component of i_C can have without serious waveform distortion? Is the limit set by collector cutoff or collector saturation?

8-3. An amplifier having the form shown in Fig. 8-6 uses a 2N301 *P-N-P* transistor. The supply voltages are $E_1 = E_2 = -12.6$ volts, and the circuit resistances are $R_1 = 1$ kilohm and $R_L = 8$ ohms. The output volt-ampere characteristics for the transistor are approximately horizontal lines with $i_C = 70i_B$. The input signal is the square wave of current shown in Fig. 8-21.

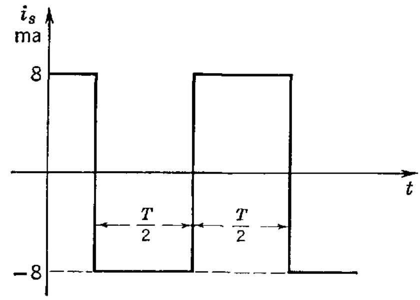


FIG. 8-21. Signal waveform for Prob. 8-3.

a. Sketch a family of output characteristics for 5-ma steps in i_B between 0 and -30 ma. Construct the load line on these characteristics, and locate the quiescent operating point.

b. Determine the power dissipated at the collector, the power absorbed by the load, and the power drawn from the battery E_2 under quiescent conditions. Is the maximum permissible collector dissipation of 5.5 watts exceeded?

c. Repeat the calculations of part *b* with the specified signal applied. Compare these values with those found in part *b*. *Note:* See Eq. (8-10).

8-4. The transistor in the amplifier of Prob. 8-1 is to be represented approximately by the piecewise-linear model shown in Fig. 8-8*a*.

a. Determine the value of α_{CB} to be used in the model.

b. Give the piecewise-linear model for the amplifier; show the numerical values of all parameters and applied voltages.

c. Construct the current transfer characteristic for the model, and give the coordinates of each break point.

8-5. Give the circuit diagram for a piecewise-linear model to represent the *P-N-P* transistor of Prob. 8-3. Show the value of α_{CB} on this diagram.

8-6. A 2N247 *P-N-P* drift transistor is used in the amplifier shown in Fig. 8-6. The supply voltages are $E_1 = -2$ volts and $E_2 = -12.6$ volts; the resistances are $R_1 = 60$ kilohms and $R_L = 2$ kilohms. The transistor can be represented by a piecewise-linear model with $\alpha_{CB} = 60$.

a. Find the quiescent collector current and collector-to-emitter voltage. Is the maximum permissible collector dissipation of 35 mw exceeded under quiescent conditions?

b. If the input signal is $i_s = 10 \cos \omega_s t$ μa, what is the peak value of the signal component of the collector current?

8-7. Four useful transistor circuits are shown in Fig. 8-22. Give an incremental linear model for each of these circuits that is valid in the frequency range where the behavior of the transistor is independent of frequency. Use the hybrid model for the transistors, and assume $\mu_{bc} = 0$. Do not assume that the capacitors shown act as short circuits. Show clearly the controlling current for each controlled source.

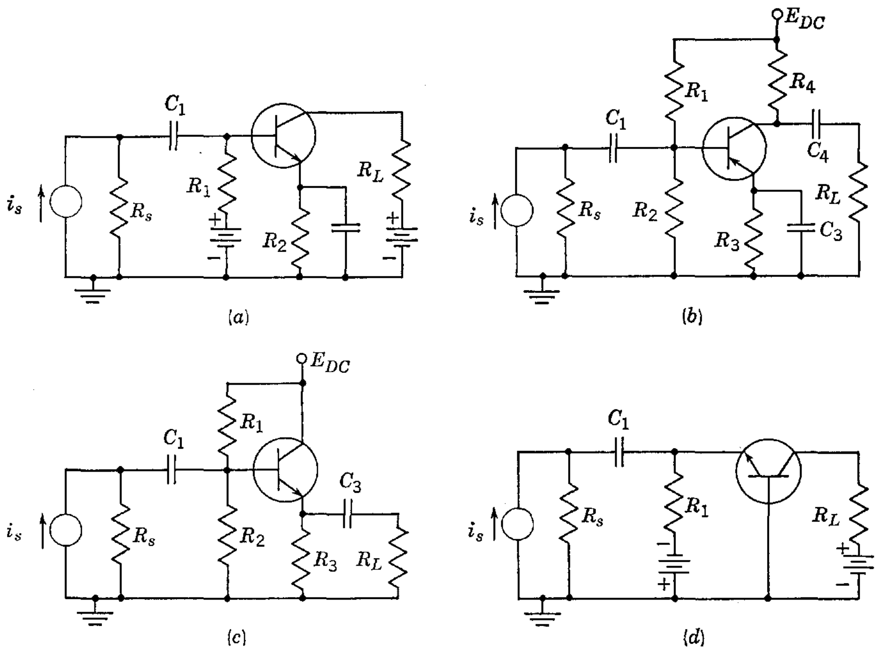


FIG. 8-22. Transistor amplifiers for Prob. 8-7. (*a*) Current amplifier; (*b*) current amplifier; (*c*) grounded-collector amplifier; (*d*) grounded-base amplifier.

8-8. The hybrid parameters for a 2N105 *P-N-P* transistor are $r_n = 2880$ ohms, $\mu_{bc} = 5.5 \times 10^{-4}$, $\alpha_{cb} = 55$, and $g_o = 16.3$ micromhos.

a. Determine the parameters of the π model shown in Fig. 8-15.

b. Determine the parameters of the T model shown in Fig. 8-16.

8-9. The T parameters for a 2N77 *P-N-P* transistor are $r_e = 23$ ohms, $r_b = 1430$ ohms, $r_d = 70$ kilohms, and $a_{cb} = 55$. Determine the hybrid parameters.

8-10. A 2N105 *P-N-P* transistor is used in the amplifier circuit shown in Fig. 8-6. The supply voltages are $E_1 = E_2 = -12.6$ volts.

a. Determine the values of R_1 and R_L that will place the quiescent operating point at $e_{CE} = -4$ volts and $i_C = -0.7$ ma.

b. Give an incremental model for the amplifier using the hybrid representation for the transistor. The parameters have the values given in Prob. 8-8 for the operating point specified in part *a*.

c. Determine the incremental current amplification of the circuit, i_{cs}/i_s, where i_{cs} is the signal component of i_C.

CHAPTER 9

PRACTICAL TRANSISTOR AMPLIFIERS

The basic amplifier circuit presented in Chap. 8 is the prototype transistor amplifier. Practical amplifiers usually consist of this circuit with certain modifications and additions that simplify the physical realization and improve the performance of the amplifier. For example, it is possible and desirable to derive the bias voltages for both the collector and the base from a single d-c supply, and, furthermore, it is often necessary to incorporate in the circuit some mechanism to reduce the effect of temperature on the current in the collector circuit. As in the case of the triode amplifier, when the transistor is required to deliver large amounts of power to a load there is an optimum dynamic path of operation on the collector characteristic; hence, when the load resistance is fixed, it may be desirable to use transformer coupling to transform the load impedance to the optimum value. The objective this chapter is to examine the modifications of the basic amplifier circuit that are often required for the satisfactory operation of practical transistor circuits.

9-1. Transistor Amplifier with a Single Battery. When a transistor is used in normal amplifier operation, the collector junction is biased in the reverse direction, and the emitter junction is biased in the forward direction. Thus for *N-P-N* transistors, both the collector and the base are at positive potentials relative to the emitter, and both bias voltages can be obtained from a single d-c supply in the manner illustrated in Fig. 9-1*a*. The voltage E_{DC} is a positive direct voltage; its value is usually of the order of 10 or 20 volts. The circuit of Fig. 9-1*a* is equally useful with *P-N-P* transistors, the only difference being that in this case the voltage E_{DC} is negative.

Although only one physical battery is used in the actual circuit, it is clear that the circuit in Fig. 9-1*b* is entirely equivalent to the actual circuit. This equivalent circuit has the same form as the basic amplifier circuit presented in Chap. 8, and the procedures for graphical, piecewise-linear, and incremental analyses are the same as those developed in Chap. 8.

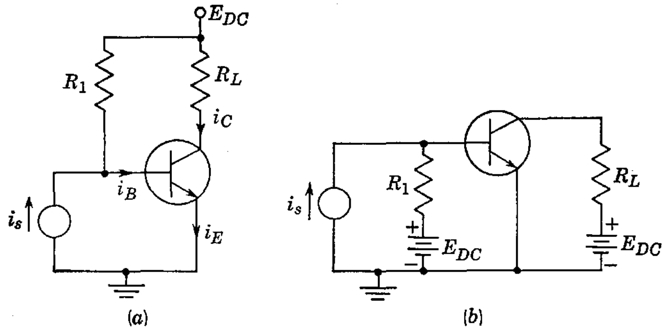


FIG. 9-1. Transistor amplifier with a single battery. (a) Circuit; (b) equivalent circuit.

9-2. Temperature Effects in Transistor Amplifiers.[1,2] In normal amplifier operation, the collector junction of the transistor is biased in the reverse direction, and ideally there would be no collector current when the base current is zero. This is the condition implied by the piecewise-linear model of Fig. 8-8a. In reality, however, a small reverse current flows across the collector junction when the base current is zero, just as a small reverse current flows in the P-N junction diode when a reverse voltage is applied. This current is much larger than that in the junction diode, for in addition to the current resulting from thermally generated carriers in the collector and base, it consists of a relatively large current resulting from free carriers in the emitter that have enough thermal energy to diffuse across the emitter junction into the base. The collector current that flows in typical small transistors with zero base current is of the order of 100 μa; however, the value of this current may vary considerably from one transistor to another.

To account for the collector current that flows when the base current is zero, the piecewise-linear model for the transistor must be modified as shown in Fig. 9-2a. The current I_{CE0} is the current that flows from collector to emitter with $i_B = 0$; the total collector current is

$$i_C = I_{CE0} + \alpha_{CB} i_B$$

The current I_{CE0} is itself no cause for concern, but the fact that it increases more or less exponentially with the temperature of the transistor introduces a problem. The solid lines in Fig. 9-2b are the collector characteristics for a transistor at one temperature. If the temperature of the transistor increases, I_{CE0} increases, and the entire family of collector characteristics is shifted upward by the amount of the increase in I_{CE0}. As the characteristic curves move upward, the quiescent operating point moves along the static load line, as shown in Fig. 9-2b, from its original position at Q to a new position at Q'. The new location of the quiescent point may be unsatisfactory.

The increase in transistor temperature may be caused either by an increase in ambient temperature or by self-heating in the transistor. When power is applied to the amplifier, the quiescent collector dissipation, which goes into heat, is given by

$$P_C = E_{CE}I_C = E_{DC}I_C - R_L I_C^2 \tag{9-1}$$

This quantity has a maximum value when $E_{CE} = E_{DC}/2$. When R_L is large, and especially when $E_{CE} < E_{DC}/2$, the self-heating in the transistor

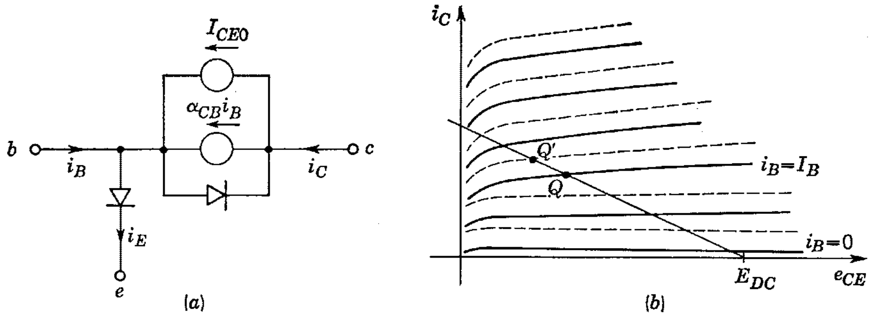


FIG. 9-2. Effect of temperature on collector current. (a) Piecewise-linear model; (b) collector characteristic.

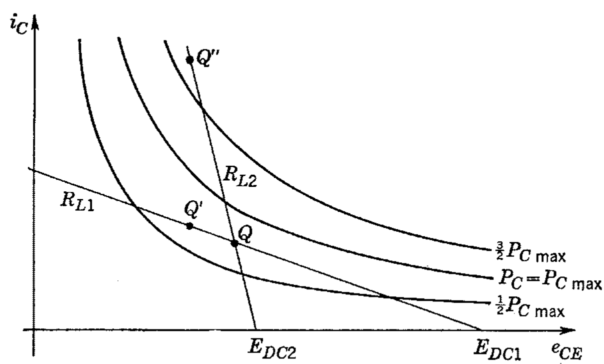


FIG. 9-3. Effect of temperature on collector current. The curves are contours of constant P_C, and $P_{C\max}$ = maximum permissible value of P_C.

will cause only a small shift in the quiescent point, as shown in Fig. 9-3 for $R_L = R_{L1}$. This small shift can often be tolerated. However, when R_L is small and E_{CE} is greater than $E_{DC}/2$, as shown in Fig. 9-3 for $R_L = R_{L2}$, the shift in the quiescent point may be considerable, and the increase in collector current further increases the collector dissipation, which in turn increases the transistor temperature further. This cumulative action usually leads to excessive collector dissipation. In the worst cases, no thermal equilibrium is possible at any safe collector current, and destruction of the transistor is certain. Thus it is often necessary to incorporate in the amplifier some mechanism to stabilize the quiescent point against the effects of self-heating and changes in ambient temperature.

9-3. Stabilization of the Quiescent Operating Point. The circuit arrangement most commonly used[1] to stabilize the quiescent point is shown in Fig. 9-4*a*. Stabilization results primarily from the presence of R_e, a resistor connected in series with the emitter lead. The base bias is provided through the voltage divider R_a-R_b, rather than through a series resistor as in Fig. 9-1*a*, for the reason that a relatively low-resistance path from base to ground is required. A model for the circuit with $i_s = 0$ is shown in Fig. 9-4*b*. The emitter diode is represented as a short circuit, and the collector diode is represented as an open circuit in this

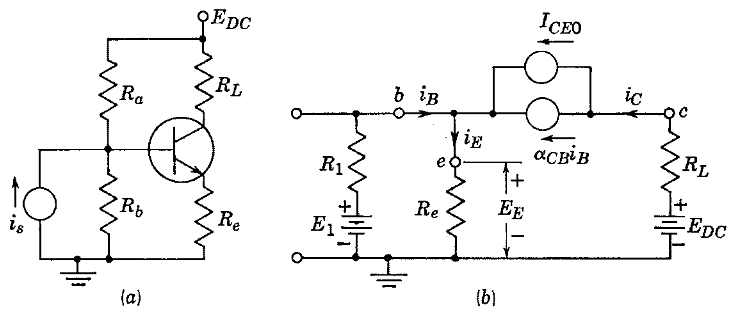


FIG. 9-4. Quiescent-point stabilization with emitter-lead resistance. (*a*) Circuit; (*b*) model.

model; the source E_1 and the resistance R_1 constitute a Thevenin equivalent for E_{DC} and the voltage divider R_a-R_b. It is clear from this circuit that a positive increment in I_{CE0} divides between R_1 and R_e; hence it causes a negative increment in i_B and a corresponding negative increment in $\alpha_{CB}i_B$. This negative increment in $\alpha_{CB}i_B$ tends to compensate for the original increment in I_{CE0}, and the net change in i_C is smaller than the increment in I_{CE0}. The shift in the quiescent operating point with changes in temperature is thereby reduced.

The addition of the resistance R_e to the amplifier introduces feedback, according to the definition presented in Sec. 6-1, into the circuit. This circuit provides an example of the use of feedback to reduce undesirable effects associated with imperfect circuit components.

Further insight into the action of R_e in stabilizing the quiescent operating point can be gained by an alternative interpretation. If the voltage drop across R_1 in Fig. 9-4*b* caused by the small current i_B is negligible in comparison with E_1, then

$$E_E \approx E_1 \tag{9-2}$$

and

$$I_E R_e \approx I_C R_e \approx E_1 \tag{9-3}$$

Thus if R_e and E_1 are constant, I_C remains approximately constant irrespective of changes in temperature. In addition, if it is necessary to change transistors, the new quiescent point will be at approximately the same location as the old, even if the characteristics of the new transistor are appreciably different from those of the old transistor.

The effectiveness of R_e in stabilizing i_C, and hence the quiescent point, can be formulated quantitatively with the aid of the model in Fig. 9-4*b*. The collector current is

$$i_C = \alpha_{CB} i_B + I_{CE0} \tag{9-4}$$

and the loop equation for the base circuit is

$$R_1 i_B + R_e(i_B + i_C) = E_1 \tag{9-5}$$

Solving (9-5) for i_B yields

$$i_B = \frac{E_1}{R_e + R_1} - \frac{R_e i_C}{R_e + R_1} \tag{9-6}$$

and substituting this expression in (9-4) gives

$$i_C = \frac{\alpha_{CB} E_1}{R_e + R_1} - \frac{\alpha_{CB} R_e i_C}{R_e + R_1} + I_{CE0} \tag{9-7}$$

$$\left(1 + \frac{\alpha_{CB} R_e}{R_e + R_1}\right) i_C = \frac{\alpha_{CB}}{R_e + R_1} E_1 + I_{CE0} \tag{9-8}$$

The coefficient on i_C in (9-8) is closely related to the feedback in the circuit, and as is shown in the chapter on feedback amplifiers, it is widely used in the study of feedback systems. If this quantity is designated by

$$F = 1 + \frac{\alpha_{CB} R_e}{R_e + R_1} = 1 + \frac{\alpha_{CB}}{1 + R_1/R_e} \tag{9-9}$$

then (9-8) can be put in the form

$$i_C = \frac{\alpha_{CB}}{(R_e + R_1)F} E_1 + \frac{I_{CE0}}{F} \tag{9-10}$$

Thus the effect of changes in I_{CE0} on i_C is reduced by the factor F. The effect of changes in α_{CB} on i_C, such as might be associated with replacing the transistor, is also reduced by this factor. It is clear from (9-9) and (9-10) that small R_1/R_e gives large F and good stabilization; for other reasons, however, it is usually not practical to make R_1/R_e less than about 5. Under this condition, and with $\alpha_{CB} = 50$, $F = 1 + {}^{50}\!/_{6} = 9.33$.

The degenerative feedback that reduces the effect of changes in I_{CE0} also reduces the signal amplification by the factor F. This fact is evident from Eq. (9-10), for an input signal is equivalent to a variation in E_1. This reduction in amplification can be eliminated for time-varying currents in the same way that cathode degeneration is eliminated in the triode amplifier, namely, by connecting a bypass capacitor in parallel with R_e. A typical circuit with an emitter bypass capacitor and an input coupling capacitor is shown in Fig. 9-5*a*.

The capacitors in the amplifier of Fig. 9-5*a* are normally chosen to act as short circuits to the time-varying components of current and voltage. Accordingly, the dynamic operating path on the collector characteristic

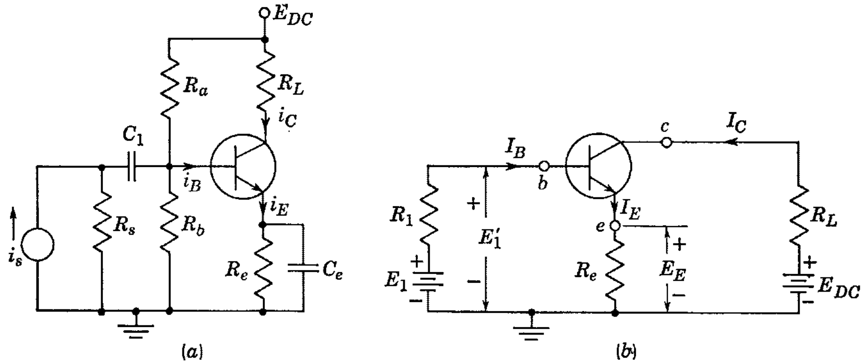


FIG. 9-5. A typical transistor amplifier. (*a*) Complete circuit; (*b*) circuit for d-c components of voltage and current.

is different from the static load line. For the d-c components of current and voltage, the circuit reduces to that shown in Fig. 9-5*b*; the base biasing circuit is replaced by its Thevenin equivalent, E_1 and R_1, in this representation. The loop equation for the collector circuit yields

$$E_{CE} = E_{DC} - R_L I_C - R_e I_E \tag{9-11}$$

But normally $I_E \approx I_C$; thus

$$E_{CE} \approx E_{DC} - (R_L + R_e) I_C \tag{9-12}$$

and

$$I_C \approx \frac{E_{DC}}{R_L + R_e} - \frac{E_{CE}}{R_L + R_e} \tag{9-13}$$

$$\approx I_o - \frac{E_{CE}}{R_L + R_e} \tag{9-14}$$

This is the equation of the static load line shown in Fig. 9-6.

The quiescent operating point can be located by determining the collector current. If the base-to-emitter voltage drop E_{BE} and the drop in R_1 caused by the small base current I_B are small compared with E_1, then as a first approximation

$$E_E = R_e I_C = E_1 \tag{9-15}$$

and

$$I_C = \frac{E_1}{R_e} \tag{9-16}$$

This relation gives a first approximation to the quiescent collector current. If a second, closer approximation is desired, the base current can be read from the characteristics for the approximate quiescent point, and the base-to-ground voltage can be determined approximately from

$$E_1' = E_1 - R_1 I_B \tag{9-17}$$

Again neglecting the small voltage E_{BE}, the second approximation to the quiescent collector current is

$$I_C = \frac{E_1'}{R_e} \tag{9-18}$$

The process can be repeated to obtain a still closer approximation; however, it is seldom necessary to go beyond the second approximation.

Locating the quiescent point by successive approximations as described above gives a useful insight into the behavior of the circuit. Successive approximations can be avoided if desired, however, by the solution of two simultaneous equations. Neglecting the small base-to-emitter voltage, the loop equation for the base circuit is

$$E_1 = R_1 I_B + R_e(I_B + I_C) \tag{9-19}$$

and the collector characteristics are approximated by

$$I_C = \alpha_{CB} I_B \tag{9-20}$$

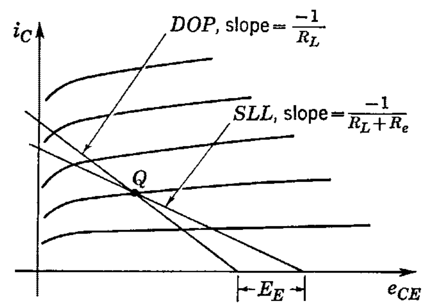


FIG. 9-6. Graphical analysis of a transistor amplifier showing the static load line (SLL) and the dynamic operating path (DOP).

Eliminating I_B between these two equations and solving for I_C yields

$$I_C = \frac{\alpha_{CB} E_1}{R_1 + (1 + \alpha_{CB}) R_e} \tag{9-21}$$

Since α_{CB} occurs in both the numerator and the denominator of this expression, and since $(1 + \alpha_{CB})R_e$ is usually considerably larger than R_1, the inaccuracy involved in the approximation of (9-20) tends to cancel in (9-21).

The capacitors C_1 and C_e in the circuit of Fig. 9-5*a* act as short circuits to the time-varying components of current and voltage. Thus the voltage across R_e is a nonvarying voltage of magnitude $E_E = R_e I_E$, where I_E is the average value of the emitter current; if there is negligible distortion, I_E is the quiescent emitter current. The loop equation for the collector circuit is thus

$$e_{CE} = E_{DC} - E_E - R_L i_C \tag{9-22}$$

This is the equation of the dynamic operating path; it is a line with slope $-1/R_L$ that passes through the quiescent point if there is negligible distortion. The dynamic operating path is shown on the collector characteristic of Fig. 9-6.

Example 9-1. The transistor in the amplifier of Fig. 9-7*a* has the collector characteristics shown in Fig. 9-7*b*. Construct the static load line on the collector characteristic, locate the quiescent operating point, and construct the dynamic operating path.

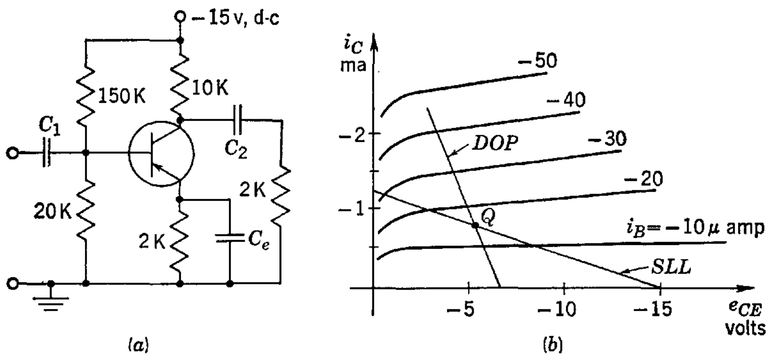


FIG. 9-7. Transistor amplifier for Example 9-1. (*a*) Circuit; (*b*) characteristic curves.

Solution. The static load line has a slope of $-1/2$ ma/volt; it is shown (*SLL*) in Fig. 9-7*b*. The base biasing circuit is equivalent to a battery $E_1 = -1.77$ volts in series with a resistance of 17.7 kilohms.

First approximation to the quiescent collector current: Assuming that $E_E = E_1$,

$$I_C = \frac{E_1}{R_e} = \frac{-1.77}{2} = -0.89 \text{ ma}$$

Second approximation: The point on *SLL* at $I_C = -0.89$ ma corresponds to $I_B \approx -0.018$ ma. Thus

$$E_1' = E_1 - R_1 I_B = -1.77 + (17.7)(0.018) = -1.45 \text{ volts}$$

Assuming that $E_E = E_1'$,

$$I_C = \frac{E_1'}{R_e} = \frac{-1.45}{2} = -0.73 \text{ ma}$$

Third approximation: The point on *SLL* at $I_C = -0.73$ ma corresponds to $I_B \approx -0.014$ ma. Thus

$$E_1' = E_1 - R_1 I_B = -1.77 + (17.7)(0.014) = -1.52 \text{ volts}$$

and

$$I_C = \frac{E_1'}{R_e} = -0.76 \text{ ma}$$

This value is not very different from the value obtained on the second approximation; further approximations yield substantially the same value for I_C.

Alternatively, I_C can be determined directly from Eq. (9-21). The collector characteristics indicate that α_{CB} is approximately 50 for this transistor. Hence

$$I_C = \frac{-(50)(1.77)}{17.7 + (51)(2)} = -0.74 \text{ ma}$$

The quiescent point corresponding to this collector current is shown in Fig. 9-7*b*.

The capacitors C_e and C_2 act as short circuits to the signal components of current and voltage; hence the collector-circuit resistance for signal currents is the resistance of the 10-kilohm and the 2-kilohm resistors in parallel. This parallel resistance is 1.67 kilohms. The dynamic operating path, shown in Fig. 9-7*b* (*DOP*), has a slope of $-1/1.67$ ma/volt, and it passes through the quiescent point when there is negligible distortion.

9-4. Design for a Specified Quiescent Operating Point. If the amplifier of Fig. 9-5*a* is to be designed to give a specified operating point for the transistor, the resistors R_a, R_b, and R_e must be chosen to give a suitable compromise among conflicting factors. The emitter circuit resistance R_e should be made as large as possible to give good stabilization of the quiescent point; however, the larger the R_e, the larger the voltage drop E_E, and the larger the required power-supply voltage. Thus R_e should be made as large as is permitted by the available power-supply voltage.

In the interest of high amplification, the resistances R_a and R_b should be made as large as possible so that the input signal current will flow into the base of the transistor rather than through R_a and R_b. In the interest of good stabilization of the quiescent point, however, R_a and R_b should be made as small as possible so that any change in I_{CE0} will cause a substantial change in i_B. Thus R_a and R_b must be chosen to give a suitable compromise between these two requirements and to provide the desired bias voltage at the base of the transistor. Since the input resistance to typical small transistors is of the order of 2 kilohms, R_1, the resistance of R_a in parallel with R_b, should be at least 10 or 20 kilohms if possible.

Example 9-2. A typical transistor current amplifier is shown in Fig. 9-8. The d-c supply voltage and the circuit resistances are to be chosen so that the quiescent operating point is at approximately $I_C = 1$ ma and $E_{CE} = 5$ volts. At this quiescent point the incremental parameters for the transistor are $\alpha_{cb} = 50$, $r_n = 2.5$ kilohms, and $g_o = 0.02$ millimho; the reverse voltage amplification factor can be taken as zero. The coupling and bypass capacitors act as short circuits to the signal components of current. Determine the necessary parameter values, and calculate the current amplification i_L/i_s.

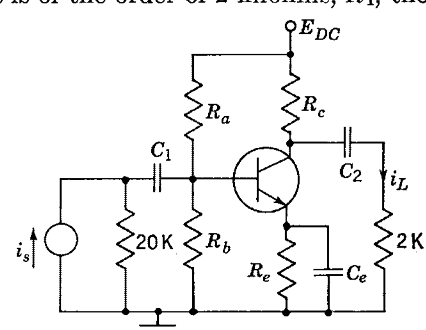


FIG. 9-8. Transistor amplifier for Example 9-2.

Solution. In the interest of current amplification, R_c should be much larger than the 2-kilohm load resistance. Therefore choose $R_c = 20$ kilohms. For the same reason R_1, the resistance of R_a and R_b in parallel, should be much larger than r_n = 2.5 kilohms. Therefore choose R_1 = 20 kilohms. If R_e is made 4 kilohms, R_1/R_e = 5, and the stabilization factor is $F = 1 + {}^{50}\!/_{6} = 9.3$; this value is considered satisfactory.

The voltage drop across R_c is thus R_cI_C = 20 volts, the drop across R_e is approximately 4 volts, and the drop across the transistor from collector to emitter is 5 volts. Thus the d-c supply voltage must be 29 volts.

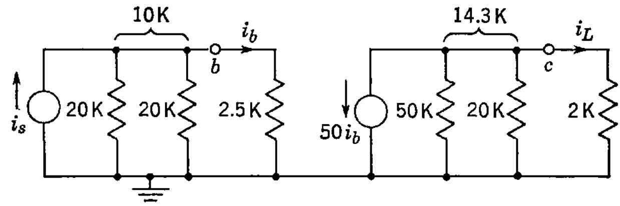


FIG. 9-9. Incremental model for the amplifier of Fig. 9-8.

The equivalent base supply voltage E_1 (Fig. 9-5) is

$$E_1 = E_E + R_1I_B = 4 + (20)(1/50) = 4.4 \text{ volts}$$

Thus

$$E_1 = \frac{R_b}{R_a + R_b} 29 = 4.4$$

$$24.6R_b = 4.4R_a$$

$$R_a = 5.6R_b$$

Also

$$R_1 = 20 = \frac{R_aR_b}{R_a + R_b} = \frac{R_a}{1 + R_a/R_b} = \frac{R_a}{6.6}$$

Thus

$$R_a = (6.6)(20) = 132 \text{ kilohms}$$

$$R_b = 132/5.6 = 23.6 \text{ kilohms}$$

An incremental model for the amplifier is shown in Fig. 9-9. The current amplification is, by inspection of the circuit,

$$A_c = \frac{i_L}{i_s} = -\frac{10}{10 + 2.5}(50)\frac{14.3}{14.3 + 2}$$

$$= -35$$

Any design procedure involving stabilization of the quiescent point is necessarily a rule-of-thumb procedure. The temperature rise caused by any given power dissipation in a transistor depends strongly on the means provided for conducting heat away from the transistor; thus it depends on the environment of the transistor, and in particular on the way in which the transistor is mounted in its environment. When transistors are required to handle large amounts of current and power, it is usually necessary to mount them in close thermal contact with a large metal plate, termed a *heat sink*, to expedite the removal of heat from the transistor. Clearly, no detailed design can be made without taking these factors into account. Fortunately, rule-of-thumb procedures are satisfactory for all but the most exacting designs.

9-5. Choice of Dynamic Operating Path for Maximum Output Power. The dynamic operating path for a transistor must be restricted to a certain permitted region of the collector characteristics. The limitations are similar to those encountered with the vacuum triode and discussed in Sec. 6-2. The permitted operating region for a particular transistor is illustrated in Fig. 9-10*a*. The collector voltage cannot exceed the indicated maximum value without the occurrence of avalanche breakdown, and the operating path cannot extend into the region of very small collector voltages without excessive distortion. The collector characteristics are somewhat crowded together at large values of collector current; hence the operating path cannot extend into the region of very large collector currents without excessive distortion. Similarly, there is a limit on the minimum collector current if excessive distortion

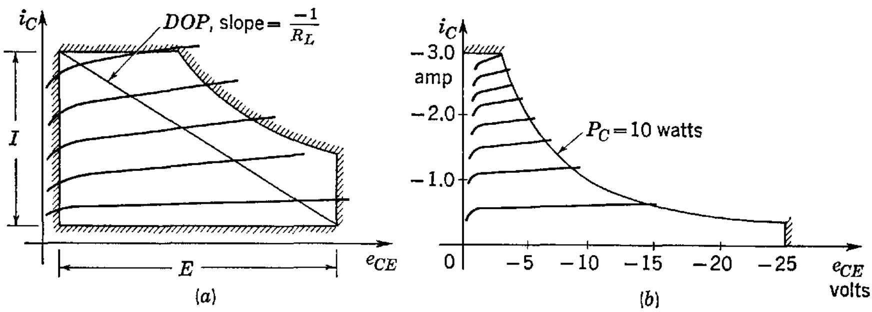


FIG. 9-10. Transistor characteristics. (*a*) Permitted operating region; (*b*) characteristics of a power transistor.

is to be avoided. In addition to these limitations, the quiescent point must lie below the hyperbola of maximum permissible collector dissipation.

If the waveform distortion is negligible, the signal power delivered to the load is proportional to the product of I, the peak-to-peak signal current in the load, and E, the peak-to-peak signal voltage across the load. The signal power delivered to the load is a maximum when the product EI is a maximum. It is clear that the dynamic operating path shown in Fig. 9-10*a* corresponds to maximum signal power in the load under the specified limitations.

The maximum permissible values of collector current, collector voltage, and collector dissipation for many transistors are such that the operating path extending from the upper left-hand corner of the permitted region to the lower right-hand corner results in excessive quiescent collector dissipation. Such a case is illustrated by the family of collector characteristics shown in Fig. 9-10*b*. Further investigation is required to determine the operating path for this transistor that gives the maximum signal power in the load.

The permitted operating region on the collector characteristic of a typical, high-power transistor is shown in Fig. 9-11. It is assumed, for simplicity, that the region is bounded by a maximum collector current, a maximum collector voltage, a maximum collector dissipation, and the coordinate axes. It is further assumed that the operating point makes equal excursions on either side of the quiescent point; hence the quiescent point is the mid-point of the dynamic operating path. Under these conditions, if the quiescent point is fixed with the coordinates E_{CE} and I_C, the dynamic operating path must lie inside a rectangle whose base is $2E_{CE}$ and whose altitude is $2I_C$. It is clear that the operating path giving the maximum EI product, and hence the maximum signal power in the load, is the diagonal of the rectangle. This fact, together with a special property of the equilateral hyperbola P, yields at once a solution to the problem of choosing an optimum path of operation. If the line segment IE terminating on the coordinate axes is tangent to P at any point, then the point of tangency is the mid-point of the line segment. This statement follows directly from the fact that the slope of the hyperbola at the point E_{CE}, I_C is $-I_C/E_{CE}$. Then for all operating paths tangent to P, the quiescent point is at the point of tangency, and the EI product is constant at the value

$$EI = 4E_{CE}I_C = 4P_{C\,\max} \tag{9-23}$$

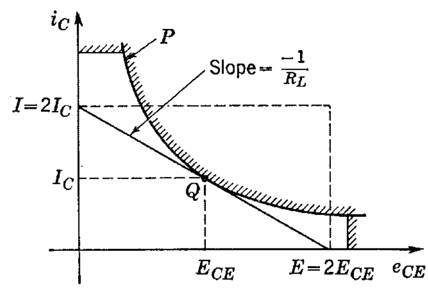


FIG. 9-11. Graphical construction related to maximum power output.

where $P_{C\,\max}$ is the collector dissipation at points on the hyperbola P. There is no operating path with a quiescent point on or below P that gives a greater EI product, for all operating paths must lie inside a rectangle with dimensions $2E_{CE}$ by $2I_C$. Hence the dynamic operating path that yields maximum signal power in the load must be tangent to the hyperbola of maximum permissible collector dissipation; however, the point of tangency may lie anywhere on the hyperbola provided the operating path does not intersect the maximum current limit or the maximum voltage limit.

Since the slope of the optimum operating path is the same as that of the hyperbola at the point of tangency, $-I_C/E_{CE}$, the optimum value for the load resistance is

$$R_L = \frac{E_{CE}}{I_C} \tag{9-24}$$

The collector dissipation has its greatest value under quiescent conditions, and this value is $P_{C\,\max}$, the maximum permissible value. If the signal is sinusoidal, the average signal power delivered to the load under maximum signal conditions is

$$P_L = \frac{E_{CE}I_C}{2} = \frac{P_{C\,\max}}{2} \tag{9-25}$$

The power drawn from the d-c supply is a constant, independent of the signal amplitude, and if there is no d-c voltage drop across the load, as in the case of a transformer-coupled load, then

$$P_{CC} = P_{C\,\max} \tag{9-26}$$

and the collector-circuit efficiency is

$$\frac{100\,P_L}{P_{CC}} = 50 \text{ per cent} \tag{9-27}$$

If the signal is a square wave, the average signal power delivered to the load under maximum signal conditions is

$$P_L = E_{CE}I_C = P_{C\,\max} \tag{9-28}$$

and the collector-circuit efficiency is 100 per cent.

All operating paths chosen according to the conditions stated above yield the same power to the load when there is negligible waveform distortion; hence a variety of designs yield the same maximum signal power to the load. If the power required by the load under sinusoidal operating conditions is specified, the first step in the design is to select a transistor having a collector dissipation at least twice the required P_L. The next step is to choose a suitable quiescent operating point on the collector-dissipation hyperbola. Equation (9-24) then yields the optimum value for the load resistance, and the d-c supply voltage required can be determined from the details of the circuit. If the quiescent point is chosen at a small value of E_{CE}, a small d-c supply voltage is required. With this choice, however, the operating path extends into the region of large collector current, and excessive waveform distortion may result. If the quiescent point is chosen at a large value of E_{CE}, a large d-c supply voltage is required. However, the waveform distortion is less, and the input signal power required is somewhat less.

9-6. Transformer-coupled Loads. Transistors are often required to deliver power to load resistances that do not give the optimum operating path discussed in Sec. 9-5. Iron-core transformers are often used in such cases to transform the load resistance to an optimum value. The circuit diagram of a transistor amplifier with a transformer-coupled load is shown in Fig. 9-12*a*. The graphical analysis of this amplifier runs exactly parallel to the graphical analysis of the transformer-coupled triode amplifier presented in Sec. 6-2. The primary of the transformer acts as a short circuit to the d-c component of collector current; hence the static load line is given by Eq. (9-14) with $R_L = 0$. The quiescent operating point on this line corresponds to the value of I_C given by Eq.

(9-21). The capacitors C_1 and C_e are understood to act as short circuits to the time-varying components of current; hence the dynamic operating path is a line of slope $-1/R'_L = -(N_2/N_1)^2/R_L$ passing through the quiescent point. The graphical construction is shown in Fig. 9-12*b*.

In designing a transistor amplifier of the type shown in Fig. 9-12, the transformer turn ratio is chosen to give a suitable slope to the dynamic operating path. The collector supply voltage E_{DC}, the stabilizing emitter resistance R_e, and the bias resistances R_a and R_b are chosen to give a

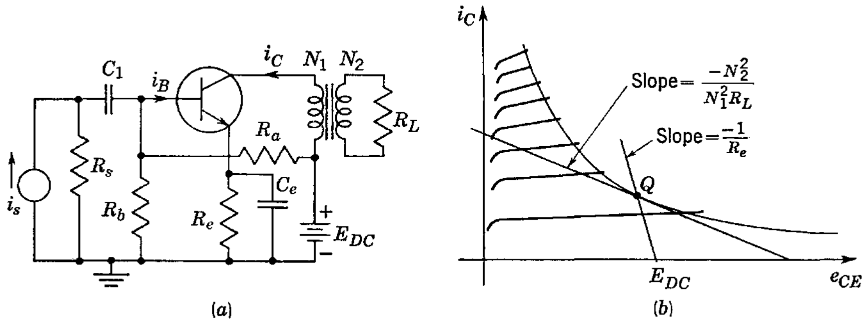


FIG. 9-12. Graphical analysis of a transistor amplifier with a transformer-coupled load. (*a*) Circuit; (*b*) graphical construction.

suitable quiescent point. The optimum quiescent point is usually located on the hyperbola of maximum permissible collector dissipation if maximum output power is required. For maximum output power, the amplitude of the input signal is adjusted to give the maximum permissible excursion of the operating point along the dynamic operating path.

9-7. Summary. Practical transistor amplifiers are usually somewhat different in form from the basic amplifier circuit, partly for reasons of convenience and partly for reasons of necessity. Practical circuits usually contain capacitors, and occasionally, as in the case of the transformer-coupled load, they may contain inductors. The basic principles underlying the analysis and design of such circuits are the same as those developed for the triode amplifier, although the details of the calculations are different. Thevenin's theorem can be used in the manner described in Sec. 6-4 to reduce transistor amplifiers of various configurations to the standard form shown in Fig. 6-16*b*. This standard form is identical with the circuit shown in Fig. 9-5*b* and analyzed in Sec. 9-3.

REFERENCES

1. Lo, A. W., R. O. Endres, J. Zawals, F. D. Waldhauer, and C. Cheng, "Transistor Electronics," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1955.
2. Hunter, L. P., "Handbook of Semiconductor Electronics," McGraw-Hill Book Company, Inc., New York, 1956.

PROBLEMS

9-1. A 2N170 *N-P-N* transistor is used in the amplifier circuit shown in Fig. 9-1*a*. The transistor can be represented by a piecewise-linear model with $\alpha_{CB} = 30$.

a. If the d-c supply voltage is 12 volts, what values of R_L and R_1 are required to locate the quiescent operating point at $I_C = 1$ ma and $E_{CE} = 4$ volts?

b. If the input signal is sinusoidal, what is the maximum amplitude that the sinusoidal component of current in R_L can have without excessive distortion?

9-2. If the d-c supply voltage in the amplifier of Fig. 9-7*a* is reduced to -10 volts, what is the new quiescent operating point? Does this quiescent point appear to be satisfactory for small-signal operation?

9-3. An *N-P-N* transistor is used in the amplifier circuit shown in Fig. 9-5*a*. The transistor can be represented by a piecewise-linear model with $\alpha_{CB} = 40$. The d-c supply voltage is 22.5 volts, and R_b and R_e are 10 and 2 kilohms, respectively.

a. Sketch a set of approximate collector characteristics based on the piecewise-linear model. Show curves for $i_B = 50$, 100, 150, 200, and 250 μa.

b. Determine the values of R_a and R_L that will locate the quiescent operating point at $I_C = 4$ ma and $E_{CE} = 7.5$ volts.

c. The coupling and bypass capacitors act as short circuits to the time-varying components of current. Sketch and dimension the static load line and the dynamic operating path on the collector characteristic constructed for part *a*.

9-4. The effect of the emitter-circuit resistance in stabilizing the quiescent operating point for the transistor in the amplifier of Fig. 9-7*a* is to be examined.

a. Determine a suitable value for α_{CB}.

b. If I_{CE0} increases by 200 μa as the result of an increase in temperature, by what amount does the quiescent collector current increase?

9-5. The transistor described in Prob. 9-3 is used in the circuit shown in Fig. 9-5*a* with $R_L = 1.0$ kilohm, $R_b = 20$ kilohms, and $R_e = 0.5$ kilohm. The circuit is required to deliver a signal current to R_L having a peak-to-peak value of 4 ma when a sinusoidal signal is applied at the input. The coupling and bypass capacitors act as short circuits to the time-varying components of current.

a. Sketch the set of collector characteristics specified in part *a* of Prob. 9-3.

b. What is the smallest value that E_{DC} can have if the amplifier is to meet the design specifications?

c. Determine approximately the required value of R_a.

9-6. A 2N109 *P-N-P* transistor is used in the power amplifier of Fig. 9-12*a* with $E_{DC} = -9$ volts, $R_e = 220$ ohms, $R_b = 5.6$ kilohms, $R_L = 10$ ohms, and $N_1/N_2 = {}^{707}\!/\!_{100}$. The coupling and bypass capacitors act as short circuits to the time-varying components of current, and the transformer acts as an ideal transformer to the time-varying currents.

a. What value of R_a is required to make the quiescent collector current 10 ma? The current amplification factor for the 2N109 is $\alpha_{CB} = 70$.

b. Sketch and label the static load line and the dynamic operating path on the i_C-e_{CE} coordinates. It is not necessary to show the collector characteristics on this sketch.

c. If the signal is sinusoidal, what is the greatest amplitude that the current in R_L can have without marked waveform distortion? Assume that the operating path can extend to the i_C and the e_{CE} axes.

9-7. A transistor having the collector characteristics shown in Fig. 9-10*b* is used in the amplifier of Fig. 9-12*a* with $R_L = 4$ ohms and $R_e = 1$ ohm. The circuit is to be designed for maximum power delivered to the load. The maximum permissible collector current, voltage, and power dissipations are indicated on the characteristic curves.

a. The dynamic operating path is to permit maximum power to the load, and one end is to terminate on the collector-voltage axis at $e_{CE} = -25$ volts. What transformer turn ratio is required?

b. Sketch the collector characteristics, and locate the quiescent operating point that will permit maximum power output for a sinusoidal signal. What value of E_{DC} is required for this quiescent point?

c. Assuming that the operating path can extend to the axes without producing waveform distortion, what is the maximum output power under sinusoidal conditions? *Note:* The actual output power obtainable with small distortion may be appreciably smaller than this value.

CHAPTER 10

MULTIGRID VACUUM TUBES

One of the most troublesome defects of the vacuum triode is the parasitic capacitance between the grid and the plate. Although this capacitance is only a few micromicrofarads, its effects become important at frequencies greater than a few tens of kilocycles per second. It reduces the amplification that can be obtained with the tube, and the feedback that it provides from the plate circuit to the grid circuit may cause self-sustaining oscillations to occur in radio-frequency amplifiers unless special circuit arrangements are employed. Self-sustaining oscillations interfere with the proper operation of an amplifier.

The addition of a second grid in the space between the control grid and the plate provides an electrostatic shield that reduces the grid-to-plate capacitance to a small value and eliminates the faults described above. The plate characteristics of this tetrode vacuum tube are quite irregular, however, and the region corresponding to linear amplification is restricted. The irregularities in the plate characteristics are removed by the insertion of a third grid in the space between the second grid and the plate. Thus the pentode vacuum tube evolves. These additional grids increase the amplification factor and the plate resistance of the tube by a large factor; as a result, greater voltage amplification can be obtained with pentodes than with triodes. The circuitry is slightly more complicated, however. In certain applications, such as the voltage amplifiers in radio receivers, pentodes are used almost exclusively.

The class of multigrid tubes includes other tubes with four or five grids. These are special-purpose tubes intended to perform other functions instead of or in addition to amplification. Such tubes can be studied more effectively when their special functions are under consideration; hence they are not discussed here.

The objective of this chapter is to correlate the characteristics of tetrodes, including beam power tubes, and pentodes with their internal physics, to present typical circuits in which these tubes are used, and to develop network models for the tubes.

10-1. Vacuum Tetrodes and Pentodes. For reasons set forth above, it became necessary early in the history of vacuum tubes to seek some method of reducing the grid-to-plate capacitance inherent in the triode. A solution to this problem was found by inserting an additional grid in the interelectrode space between the control grid and the plate as indicated schematically in Fig. 10-1*a*. If the second grid is held at a fixed potential, it acts as an electrostatic shield, or screen, between the control grid and plate; hence it is called a *screen grid*. The grid-plate capacitance can be reduced by a factor of 100 or more by the addition of a screen grid.

Typical plate characteristics for a tetrode with a fixed screen-to-cathode voltage are shown in Fig. 10-1*b*. It is clear from a comparison of these

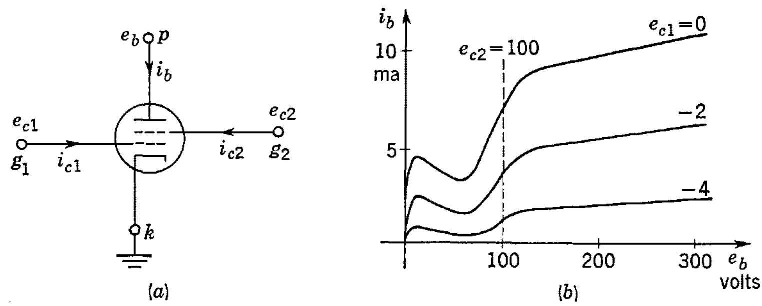


FIG. 10-1. Tetrode characteristics. (*a*) Tetrode vacuum tube; (*b*) static tetrode plate characteristics.

curves with those for a triode that the addition of a screen grid has a pronounced effect on the plate characteristics. Particularly distressing is the fact that the characteristics are quite nonlinear for values of plate voltage smaller than the screen voltage, for waveform distortion will be severe if the operating path extends into this region. The potential distribution diagram for the tetrode reveals the reason for the irregularities in the plate characteristics and indicates how they can be eliminated.[1,2]

An idealized tetrode with infinite, parallel, plane electrodes is pictured in Fig. 10-2*a*; the potential-distribution diagrams for two paths between the cathode and plate are shown in Fig. 10-2*b*. The potential in the region between the cathode and control grid is determined primarily by the control-grid and screen-grid potentials and is much less dependent on the plate potential. Thus the potential minimum in front of the cathode and the current leaving the cathode depend primarily on the control-grid and screen-grid potentials, and the cathode, control grid, and screen grid act much like a triode with holes in its plate. A fraction of the current leaving the cathode is intercepted by the screen grid; the remainder passes through the space between the screen-grid wires and is collected by the plate. The screen grid must be held at a positive potential to give a suitable value of plate current; a typical value for the screen potential is $E_{c2} = 100$ volts.

When the plate voltage is less than the screen-grid voltage, such as E_b' in Fig. 10-2*b*, the following action takes place. Electrons from the cathode strike the plate with sufficient kinetic energy to release one or more electrons from the plate. These secondary electrons find themselves in an electric field that accelerates them toward the screen grid; they are eventually collected by the screen grid because it is the most positive electrode in the tube. The loss of secondary electrons by the plate is manifest as a reduction in plate current, and it is the cause of the irregularities in the tetrode plate characteristics.

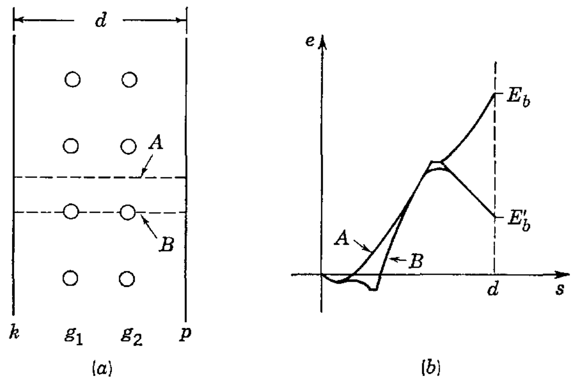


FIG. 10-2. Conduction through a tetrode. (*a*) Idealized tetrode; (*b*) potential distribution.

Secondary emission takes place at the plate of the triode just as it does at the plate of the tetrode. In the triode case, however, the plate is normally the most positive electrode in the tube, and the secondary electrons all return to the plate. In the extreme case where the grid of the triode is more positive than the plate, the secondary electrons are collected by the grid, thereby increasing the grid current and reducing the plate current.

If the potential-distribution curve in the space between the plate and screen grid were pulled down to some low level as illustrated in Fig. 10-3*a*, secondary electrons emitted from the plate would find themselves in an electric field accelerating them back toward the plate, and they would return to the plate just as they do in the triode. The cause of the irregularities in the tetrode characteristic would thereby be removed. This desirable potential distribution can be realized by adding still another grid, called the *suppressor grid*, to the tube. The result is the pentode vacuum tube represented schematically in Fig. 10-3*b*.

The pentode is normally operated with the suppressor grid connected to the cathode, the screen grid held at some fixed positive potential as in the case of the tetrode, and the control grid biased at a suitable negative potential as in the case of the triode. Typical plate characteristics for a pentode with the suppressor grid connected to the cathode and with the screen held at a fixed positive potential are shown in Fig. 10-4. The action inside the tube leading to the sharp drop in plate current and the sharp rise in screen current at low values of plate voltage can be explained in terms of the potential-distribution diagram of Fig. 10-3*a*. Electrons in transit

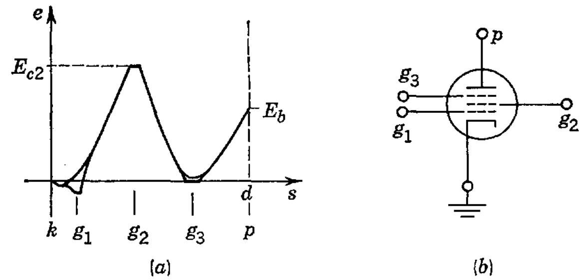


FIG. 10-3. Evolution of the pentode vacuum tube. (*a*) Potential distribution; (*b*) the pentode vacuum tube.

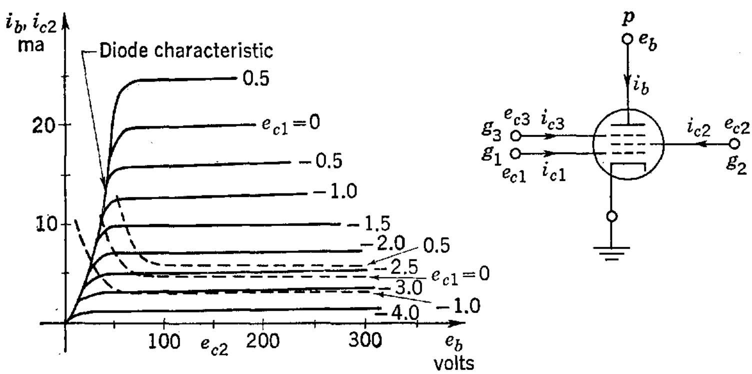


FIG. 10-4. Static pentode plate characteristics; e_{c2} = 150 volts, e_{c3} = 0 volts, i_b given by solid curves, and i_{c2} given by dotted curves.

from the cathode to the plate move rather slowly in the vicinity of the suppressor grid because of the low potential in this region. At low plate voltages a large negative space charge accumulates in the vicinity of the suppressor, the potential is depressed to negative values, and a virtual cathode forms in the space between the suppressor-grid wires. The faster electrons arriving at the virtual cathode pass on to the plate; the slower ones are turned back and are eventually collected by the screen grid. The plate and the virtual cathode act very much like a diode, giving rise to the diode-like portion of the characteristic at low plate voltages. The point at which each separate characteristic breaks away from the diode characteristic is the point at which the virtual cathode disappears for the corresponding value of e_{c1}.

The addition of the suppressor grid has some further effects on the properties of the tube. For one thing, the additional grid gives additional shielding between the plate and the control grid, thereby reducing further the capacitance between these electrodes. The capacitance may be reduced by as much as 1000 by the shielding action of the screen and suppressor grids. Another result is that the plate current in the pentode is almost independent of the plate voltage when there is no virtual cathode at the suppressor grid. It follows from this fact that the amplification

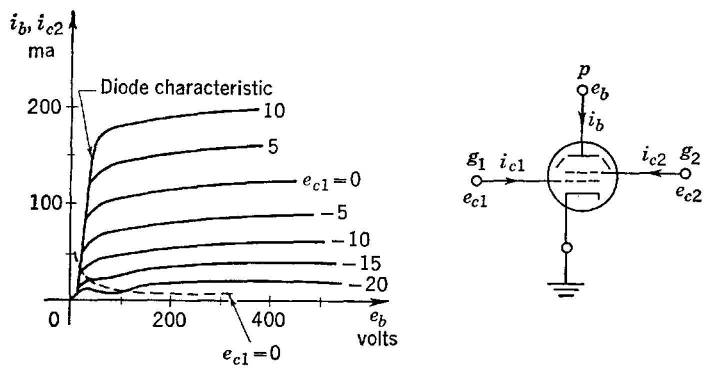


FIG. 10-5. Static beam-power-tube plate characteristics; e_{c2} = 250 volts, i_b given by solid curves, and i_{c2} given by dotted curve.

factor and the plate resistance of the pentode are much greater than the corresponding parameters of the triode; as a result, greater voltage amplification is possible with the pentode than with the triode.

The addition of a suppressor grid is one way of producing the desired minimum shown in the potential-distribution curve of Fig. 10-3*a*. An alternative method, illustrated schematically in Fig. 10-5, is to add electron-deflecting plates to a tetrode arranged so that the electron stream flowing to the plate is concentrated into a dense beam in the space between the screen grid and the plate. The dense negative space charge associated with this concentration of electrons depresses the potential-distribution curve and develops the desired potential minimum. Since the action of the space charge is more or less uniform in space, the plate characteristics of such tubes have sharper knees than comparable pentodes, and the operating path can extend farther to the left than in the case of the pentode. This fact is important in power amplifiers where large voltage swings are necessary for large power output; hence the beam tetrode is used primarily in power amplifiers and is known as a beam power tube.[1,2]

The plate current in multigrid tubes contains a small, randomly varying component like that described in connection with the triode. The noise component of current in tetrodes and pentodes is usually several times that in triodes, however. The greater noise in multigrid tubes is associated with the division of current between the plate and the screen grid. This division is not constant, but varies randomly with time, giving rise to random variations in plate current. Noise originating in this manner is called *partition noise*. Since all the current in the triode is collected by the plate, there is no partition noise. For this reason triodes are often used in preference to pentodes when very feeble signals are to be amplified, even though the triode has a much smaller amplification factor.

10-2. Pentode Voltage Amplifiers. A practical pentode voltage amplifier is shown in Fig. 10-6. All bias voltages are derived from a single d-c source, E_{bb}. The control-grid bias is provided by the voltage drop across R_k, and the screen-grid bias is provided through the series resistor R_s which is chosen to give the desired potential at the screen grid. The bypass capacitors C_k and C_s are chosen large enough to act as short circuits to the signal components of current; hence the voltages across them are direct voltages. It follows from these facts that only the control-grid and plate voltages vary with the signal. The screen and suppressor grids affect the shape of the tube characteristics, but they play no direct part in amplification of the signal; thus the pentode behaves like a triode having the characteristics shown in Fig. 10-4.

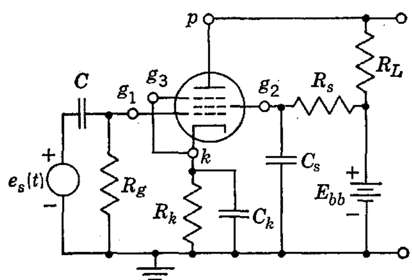


FIG. 10-6. Pentode voltage amplifier.

The design of a pentode amplifier for a specified quiescent operating point is straightforward. The cathode resistor is chosen to give the desired grid bias:

$$e_{c1} = -R_k(i_b + i_{c2}) = -E_k \tag{10-1}$$

The screen-grid resistor is chosen to give the specified screen-grid voltage:

$$e_{c2} = E_{bb} - R_s i_{c2} - E_k \tag{10-2}$$

and the plate load resistor is chosen to give the specified plate voltage:

$$e_b = E_{bb} - R_L i_b - E_k \tag{10-3}$$

The procedure for choosing the size of the bypass capacitors is developed in detail in Chap. 14.

The construction of the static load line on the plate characteristics and the determination of the quiescent operating point when the circuit is given is not a simple problem, for e_{c1}, e_{c2}, e_b, i_{c2}, and i_b are all unknown. To make matters worse, a given set of plate characteristics applies for only one value of the unknown e_{c2}. Fortunately, however, it is seldom necessary to solve this problem. The construction of the dynamic operating path, which is of considerable importance when the signal is large, is often a much simpler problem, for under normal dynamic operating conditions E_k and e_{c2} are constant and i_{c2} is of no interest. When the quiescent operating point is known and when the screen-grid voltage has the value for which the plate characteristics are known, the dynamic operating path can be constructed by the same procedure as that used with the triode. The equation of the dynamic operating path is

$$e_b = E_{bb} - R_L i_b - E_k \tag{10-4}$$

This is the equation of a straight line with a slope $-1/R_L$; it passes through the quiescent point if there is negligible waveform distortion.

All the calculations discussed above require a knowledge of the quiescent screen-grid current. The tube manufacturer often supplies information about this current in the form of a family of screen-grid characteristics superimposed on the plate characteristics as illustrated in Fig. 10-4. In many cases, however, only a single screen-grid characteristic for one fixed value of control-grid voltage is given. In such cases the screen-grid current for other values of control-grid voltage can be estimated with sufficient accuracy on the following basis. In the normal operating region, on the right of the knee of the plate characteristics, the fraction of the total current intercepted by the screen grid is nearly constant and independent of the electrode voltages;[2] that is,

$$i_{c2} = \rho i_b \tag{10-5}$$

The constant of proportionality, ρ, is usually in the range between 0.3 and 0.5 for typical small pentodes. Thus the screen-grid current for any combination of electrode voltages in the normal operating region of the plate characteristics can be determined if ρ and i_b are known. If one screen-grid characteristic is shown on the plate characteristics, it suffices for the determination of ρ.

Tube characteristics are usually supplied for one value of screen-grid voltage only. These curves cannot be used directly in the design of an amplifier to be operated with any other value of screen voltage. However, the given characteristics can be converted rather easily to correspond to any other value of screen voltage in a fairly wide range. When the suppressor grid is connected to the cathode, the plate and screen currents can be expressed as functions of the electrode voltages:

$$i_{c2} = f_{c2}(e_{c1},e_{c2},e_b) \quad (10\text{-}6)$$

and

$$i_b = f_b(e_{c1},e_{c2},e_b) \quad (10\text{-}7)$$

These functions are represented graphically by the tube characteristics. With the aid of certain simplifying assumptions it can be shown that in the normal operating range all electrode currents vary approximately as the three-halves power of the electrode voltages when the voltages are all changed in the same proportion; that is, if all electrode voltages are multiplied by a given factor k, then all electrode currents are multiplied by $k^{3/2}$. Thus (10-6) and (10-7) can be written as

$$k^{3/2}i_{c2} = f_{c2}(ke_{c1},ke_{c2},ke_b) \quad (10\text{-}8)$$

and

$$k^{3/2}i_b = f_b(ke_{c1},ke_{c2},ke_b) \quad (10\text{-}9)$$

in which k is the variable. These relations hold reasonably well in the normal operating region of the characteristics; they do not hold when a virtual cathode exists at the suppressor or at low plate currents where current from parts of the cathode may be cut off.

Suppose now that a set of pentode characteristics corresponding to a given screen voltage E_{c2} is to be converted to a new screen voltage, $E'_{c2} = kE_{c2}$. It follows from (10-8) and (10-9) that the original characteristics will correspond to the new screen voltage if the plate voltage scale and the control-grid voltage scale are multiplied by k and if the plate current scale and the screen current scale are multiplied by $k^{3/2}$. Thus the tube characteristics can be converted to a different screen voltage by the simple expedient of changing the scales of current and voltage. The conversion is illustrated in Fig. 10-7.

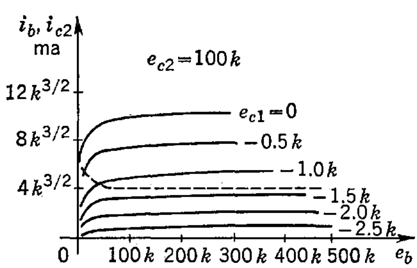


FIG. 10-7. Conversion of pentode characteristics to a new screen-grid voltage.

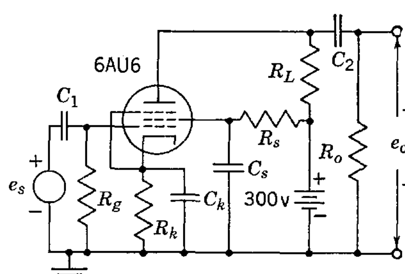


FIG. 10-8. Pentode voltage amplifier for Example 10-1.

Example 10-1. A typical pentode voltage amplifier is shown in Fig. 10-8. The resistances in the circuit are to be chosen to locate the quiescent operating point at $E_{bo} = 100$ volts and $E_{c1} = -2$ volts when $E_{c2} = 100$ volts.

Solution. The plate current given by the tube characteristics at the quiescent point is approximately 1.8 ma. The plate load resistance required is then

$$I_{bo}R_L = E_{bb} - e_b - E_k = 300 - 100 - 2 = 198 \text{ volts}$$
$$R_L = 198/1.8 = 110 \text{ kilohms}$$

The tube characteristics give only one curve of I_{c2}; this curve is for $E_{c1} = 0$. Thus the screen-grid current at the quiescent point must be determined with the aid of Eq. (10-5). For $E_{c1} = 0$ and $E_b = 100$ volts, $I_b = 9.8$ ma and $I_{c2} = 4.2$ ma; hence

$$\rho = \frac{I_{c2}}{I_b} = \frac{4.2}{9.8} = 0.43$$

Thus for $E_{c1} = -2$ volts and $E_b = 100$ volts, $I_b = 1.8$ ma, and

$$I_{c2} = \rho I_b = (0.43)(1.8) = 0.77 \text{ ma}$$

The screen dropping resistance required is then

$$I_{c2}R_s = E_{bb} - E_{c2} - E_k = 300 - 100 - 2 = 198 \text{ volts}$$
$$R_s = 198/0.77 = 257 \text{ kilohms}$$

The cathode bias resistor required is given by

$$E_k = -E_{c1} = R_k(I_{bo} + I_{c2}) = 2 \text{ volts}$$
$$R_k = \frac{2}{1.8 + 0.77} = 0.78 \text{ kilohm} = 780 \text{ ohms}$$

10-3. Beam Power Amplifiers. Beam power tubes are designed especially for applications in which large output power is required. Figure

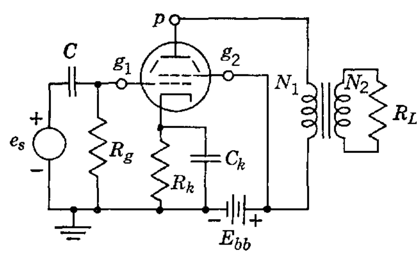


FIG. 10-9. Beam power amplifier for Example 10-2.

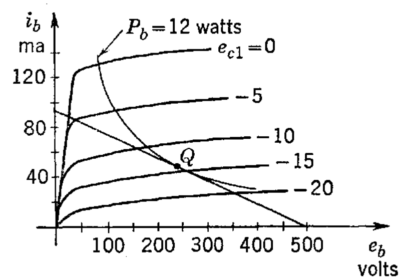


FIG. 10-10. Beam-power-tube characteristics.

10-9 shows a transformer-coupled power amplifier using a beam power tube. As is often the case with such tubes, the full plate-supply voltage is applied to the screen grid; cathode bias is provided in the usual manner.

A set of plate-characteristic curves for the tube is shown in Fig. 10-10. For Class A_1 operation, the dynamic operating path must be restricted to a region bounded by the $e_{c1} = 0$ line, the maximum permissible plate voltage, and the plate-voltage axis. In addition, the quiescent operating point must lie on or below the hyperbola of maximum permissible plate dissipation. Thus the design of a beam power amplifier is concerned largely with determining the operating path inside the permitted region that gives maximum signal power to the load. The permitted region shown in Fig. 10-10 is quite similar to the permitted region on the transistor characteristic shown in Fig. 9-11. The principal difference is that the dynamic operating path for the beam power tube must terminate on the diode-like portion of the characteristics rather than on the plate-current axis. Nevertheless, the results deduced in connection with Fig. 9-11 apply to the geometry in Fig. 10-10 within a reasonable approximation, and they indicate directly the optimum design for the beam power amplifier.

According to the foregoing discussion and the results obtained in Sec. 9-5, the optimum dynamic operating path for the beam power tube is tangent to the hyperbola of maximum permissible plate dissipation. If the operating point makes equal excursions on either side of the quiescent point, then the quiescent point should lie on the hyperbola. The output power is nearly independent of the point of tangency provided the operating path terminates on the diode-like part of the plate characteristics. The optimum load resistance referred to the primary of the output transformer is $R'_L = E_{bo}/I_{bo}$. If the signal is sinusoidal, then the maximum-signal output power is somewhat less than half of the maximum permissible plate dissipation, and the plate-circuit efficiency is somewhat less than 50 per cent.

Example 10-2. A 6CZ5 beam power tube is used in the power amplifier of Fig. 10-9. The plate characteristics and the hyperbola of maximum permissible plate dissipation for this tube are shown in Fig. 10-10. The cathode resistor, the transformer turn ratio, and the plate-supply voltage are to be determined for an optimum design.

Solution. The quiescent operating point is chosen at $E_{bo} = 250$ volts as a compromise between waveform distortion and the requirement of a large plate-supply voltage. The quiescent plate current is thus

$$I_{bo} = \frac{P_{b\max}}{E_{bo}} = \frac{12}{250} = 0.048 \text{ amp} = 48 \text{ ma}$$

The optimum load resistance referred to the primary of the output transformer is

$$R'_L = \frac{E_{bo}}{I_{bo}} = \frac{250}{48} = 5.2 \text{ kilohms}$$

and the required transformer turn ratio is

$$\frac{N_1}{N_2} = \sqrt{\frac{5200}{R_L}}$$

when R_L is expressed in ohms.

The plate characteristics indicate that for the chosen quiescent point, the grid bias must be approximately 14 volts; hence

$$R_k = \frac{-E_{c1}}{I_{bo} + I_{c2}} = \frac{14}{52} = 0.27 \text{ kilohm} = 270 \text{ ohms}$$

The plate-supply voltage must then be

$$E_{bb} = 250 + 14 = 264 \text{ volts}$$

neglecting the small d-c drop across the resistance of the transformer winding. This design corresponds with the recommendations of the tube manufacturer.

For maximum power output with small distortion, the dynamic operating path should terminate approximately on the plate characteristic for $e_{c1} = -5$ volts; thus the input signal should have a peak instantaneous value

$$E_{s\,\text{peak}} = 14 - 5 = 9 \text{ volts}$$

With this signal applied, the peak-to-peak voltage across the transformer primary is $E \approx 460$ volts, and the peak-to-peak current in the transformer primary is $I \approx 74$ ma. If the signal waveform is sinusoidal, the power delivered to the load is

$$P_L = \frac{EI}{8} = \frac{(460)(0.074)}{8} = 4.25 \text{ watts}$$

This is somewhat less than $P_{b\max}/2 = 6$ watts because the path of operation does not extend to the coordinate axes. Greater power output can be obtained by applying a larger signal at the input; however, the waveform distortion is correspondingly greater.

10-4. Pentode Models. The plate characteristics for a pentode with a fixed screen-grid voltage can be represented approximately by a set of

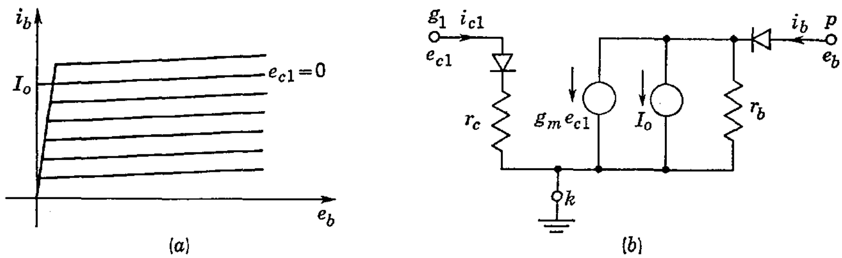


FIG. 10-11. Piecewise-linear model for a pentode with fixed screen-grid voltage. (a) Characteristics; (b) model.

piecewise-linear characteristic like those shown in Fig. 10-11a. These characteristics are straight parallel lines that are equally spaced for equal increments of grid voltage. Corresponding to these characteristics there is a piecewise-linear model for the pentode; the model is shown in Fig. 10-11b. The extended plate characteristic for $e_{c1} = 0$ is a straight line that intersects the plate-current axis at $i_b = I_o$; hence for $e_{c1} = 0$, the plate circuit can be represented by a resistance in parallel with a current source I_o. When grid voltage is applied, the plate characteristic is shifted vertically by a distance that is proportional to e_{c1}. Since this shift is equivalent to a change in the value of I_o, it can be accounted for by a current source $g_m e_{c1}$ connected in parallel with the source I_o. The proportionality constant g_m is the transconductance of the model. The volt-ampere characteristic for the grid circuit of the pentode is similar to that of a triode; hence the model for the grid circuit is like that used for the triode. The diodes in the model symbolize the fact that the grid and plate currents can never be negative.

There are certain limitations on the applicability of the piecewise-linear model in Fig. 10-11b. First, it is valid only for values of e_b lying to the right of the knee of the plate characteristic; second, it is valid for one value of screen-grid voltage only; and finally, it does not account for the component of cathode current that flows in the screen-grid circuit. Because of the last of these facts, the model cannot be used to determine the quiescent point when the cathode-resistor bias is used.

The incremental model for the pentode is useful in analyzing circuits in which the tube is used to amplify small signals. In the most common type of pentode voltage amplifier, the suppressor grid is connected to the cathode, the screen grid is held at a fixed potential relative to the cathode, and the mode of operation is Class A_1. Under these conditions the screen and suppressor grids have no direct effect on the increments of voltage and current in the circuit, and the tube operates in the same general manner as a triode. Consequently, the incremental model for the pentode, shown in Fig. 10-12, has the same form as that developed in Chap. 7 and shown in Fig. 7-3 for the triode. The parameters in this model must be evaluated at the proper operating point; the transconductance especially will vary considerably from one operating point to another. The plate resistance can be determined from the slope of the plate characteristic at the operating point, and the transconductance can be determined from the change in plate current resulting from a small change in grid voltage when the plate voltage is held constant. The plate characteristics used in these calculations must correspond to the quiescent screen-grid voltage.

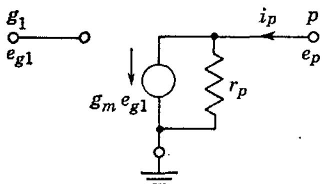


FIG. 10-12. An incremental model for the pentode.

There are some applications for the pentode in which the screen and suppressor grids are not held at constant potentials relative to the cathode, and in such cases it is desirable to have an incremental model that accounts for the effects of these electrodes on the currents flowing in the tube. A complete model for the pentode can be developed by a simple extension of the method employed in Sec. 7-1. This general model can then be specialized to fit the operating conditions that arise in any particular case.

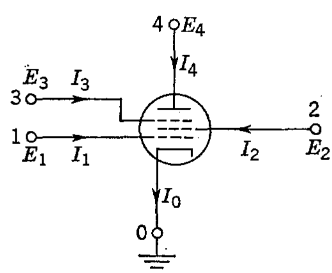


FIG. 10-13. Currents and voltages in a pentode vacuum tube.

To systematize the derivation, the electrodes are numbered and the currents and voltages are symbolized as shown in Fig. 10-13. Then, since only the cathode emits electrons, it follows that

$$I_0 = I_1 + I_2 + I_3 + I_4 \tag{10-10}$$

The plate current is a function of the four electrode voltages:

$$I_4 = f_4(E_1,E_2,E_3,E_4) \tag{10-11}$$

and the increment in I_4 resulting from increments in voltage applied to each of the electrodes can be expressed as

$$\Delta I_4 \approx dI_4 = \frac{\partial I_4}{\partial E_1} dE_1 + \frac{\partial I_4}{\partial E_2} dE_2 + \frac{\partial I_4}{\partial E_3} dE_3 + \frac{\partial I_4}{\partial E_4} dE_4 \tag{10-12}$$

The partial derivatives in (10-12) have the dimensions of conductance; hence giving them appropriate symbols and using lower-case letters for increments of current and voltage yields

$$i_4 = g_{41}e_1 + g_{42}e_2 + g_{43}e_3 + g_{44}e_4 \tag{10-13}$$

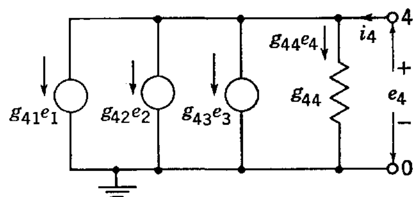


FIG. 10-14. Incremental model for the plate circuit of a pentode.

The circuit shown in Fig. 10-14 corresponds to Eq. (10-13). The conductance g_{44} is the incremental plate conductance; g_{41}, g_{42}, and g_{43} are the incremental transconductances from grids 1, 2, and 3, respectively, to the plate. Each of the controlled sources in Fig. 10-14 accounts for the effect of one of the grids on the plate current; the conductance g_{44} accounts for the effect of the plate voltage on the plate current.

Following this same procedure, models similar to that of Fig. 10-14 can be developed for each of the three grids in the pentode. When these models are combined in accordance with Eq. (10-10) a complete model for the pentode, shown in Fig. 10-15*a*, results. An equivalent model using voltage sources is shown in Fig. 10-15*b*; it is obtained from the circuit in Fig. 10-15*a* by the conventional process of source transformation. The relations among the parameters in the two models, which are given by the transformation process, can be summed up in two statements:

$$r_{kk} = \frac{1}{g_{kk}} \qquad \text{and} \qquad \mu_{jk} = \frac{g_{jk}}{g_{jj}} = g_{jk}r_{jj} \tag{10-14}$$

When the pentode is used at high frequencies, the interelectrode capacitances must be accounted for. In general, a capacitance is required between each pair of electrodes, including the cathode; however, some of these, such as the control-grid-to-plate capacitance, are usually negligible.

The value of the models in Fig. 10-15 lies principally in the fact that they can be specialized to represent any particular operating condition. For example, Fig. 10-16 shows a pentode amplifier in which the screen voltage is obtained from an unbypassed resistor. The mode of operation is Class A_1, interelectrode capacitances are negligible at the signal frequency, and both C and C_k act as short circuits to the signal components of current. An incremental model for this circuit is shown in Fig. 10-17. Since no current flows in the control grid, the portion of the

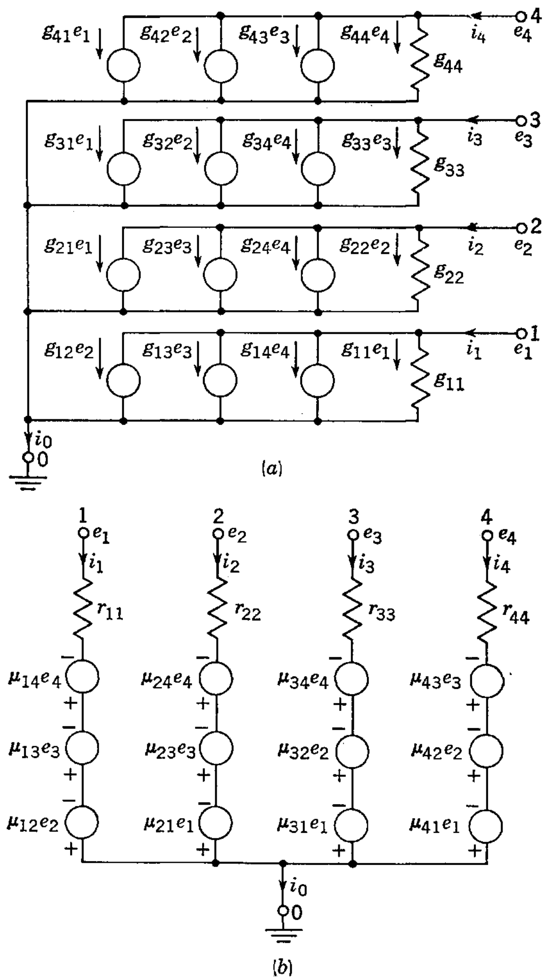


FIG. 10-15. The complete pentode model. (*a*) Current-source representation; (*b*) voltage-source representation.

model representing the control grid is simply an open circuit. The suppressor grid is connected to the cathode; hence $e_3 = 0$, and all sources controlled by e_3 are omitted from the model. Moreover, since $e_3 = 0$, the portion of the model representing the suppressor grid is of no interest and is omitted in Fig. 10-17. An analysis of this model shows that the voltage amplification is materially reduced by the action of the screen grid on the plate current.

The model in Fig. 10-17 can be simplified further by using the relation, presented in Sec. 10-2, between the plate and screen-grid currents. This relation is

$$I_2 = \rho I_4 \qquad (10\text{-}15)$$

Since ρ is a constant, it follows that the increments of plate and screen-grid current are related by

$$i_2 = \rho i_4 \qquad (10\text{-}16)$$

This relation can be symbolized as shown in the circuit of Fig. 10-18. Calculations based on this simplified model are easy to make, and they usually compare quite satisfactorily with experimental results.

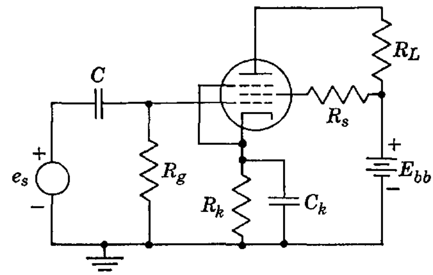


FIG. 10-16. A pentode amplifier with no screen-grid bypass capacitor.

It is clear from the circuits appearing in the last few illustrations that the circuit models for transistor and vacuum-tube amplifiers may contain a number of controlled sources and that these are controlled by

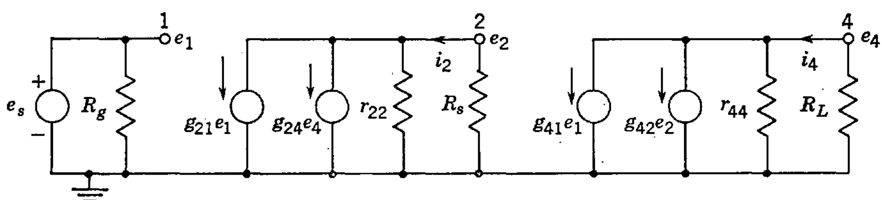


FIG. 10-17. Incremental model for the amplifier of Fig. 10-16.

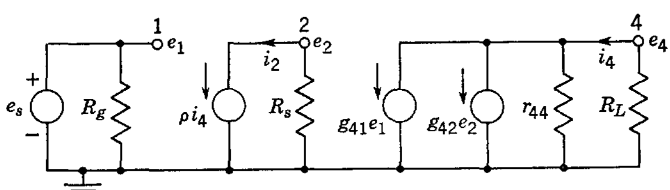


FIG. 10-18. Simplified model for the amplifier of Fig. 10-16.

currents or voltages appearing at various places in the circuit. It is appropriate at this point to call attention to the fact that when using circuit models for transistors or vacuum tubes, it is essential that the controlling current or voltage for each controlled source be clearly marked on the circuit diagram. Only in this way can errors and confusion be avoided. In this connection it should be noted that in the pentode models of Fig. 10-15 the controlling voltages are the increments in the potentials of the various electrodes *relative to the cathode.*

REFERENCES

1. Gray, T. S.: "Applied Electronics," 2d ed., John Wiley & Sons, Inc., New York, 1954.
2. Spangenberg, K. R.: "Vacuum Tubes," McGraw-Hill Book Company, Inc., New York, 1948.

PROBLEMS

10-1. A 6AU6 pentode is used in the voltage amplifier circuit shown in Fig. 10-6. The d-c supply voltage is 300 volts, and the grid return resistance R_g is 1 megohm. The tube is to be biased so that under quiescent conditions $E_b = E_{c2} = 100$ volts and $I_b = 5$ ma. Under these conditions, $I_{c2} = 2.1$ ma. Determine the required values of R_k, R_s, and R_L. Do not overlook the fact that the screen-grid current flows through the cathode resistor.

10-2. The amplifier of Prob. 10-1 is to be biased so that under quiescent conditions $E_b = E_{c2} = 100$ volts and $I_b = 3$ ma.

a. Determine the screen-grid current at this quiescent point. *Note:* See Eq. (10-5).

b. Determine the required values of R_k, R_s, and R_L.

c. Sketch the static load line and the dynamic operating path on the plate characteristics under the assumption that the coupling and bypass capacitors act as short circuits to the time-varying components of current. Estimate the voltage amplification provided by the circuit in linear Class A_1 operation.

10-3. A 6AQ5 beam power tube is used in the power amplifier shown in Fig. 10-9. The circuit is adjusted so that under quiescent conditions $E_b = E_{c2} = 200$ volts and $E_{c1} = -8$ volts. Determine the quiescent plate current. Note that the published tube characteristics must be converted to the value of screen-grid voltage given above.

10-4. A 6AQ5 beam power tube is used in the power amplifier shown in Fig. 10-9. The circuit is to be designed to deliver the greatest possible power to an 8-ohm loudspeaker in Class A_1 operation.

a. Plot the plate characteristic for $E_{c1} = 0$, $E_{c2} = 250$ volts, and construct the hyperbola of maximum permissible plate dissipation on this plot.

b. Under quiescent conditions, E_b and E_{c2} are to be 250 volts. Construct the optimum dynamic operating path on the plot of part *a.* Determine the grid bias, the value of R_k, and the transformer turn ratio required for this operating path.

c. If the input signal voltage is sinusoidal, what is the greatest peak value that it can have in Class A_1 operation? With this signal applied, what is the power delivered to the loudspeaker? The small effects of waveform distortion can be neglected, and the transformer can be considered ideal in so far as signal currents are concerned.

10-5. A 6AK5 pentode is used in the voltage amplifier of Fig. 10-6. The supply voltage is 250 volts, and the circuit elements are chosen to locate the quiescent operating point at $E_{c1} = -3$ volts, $E_{c2} = 120$ volts, and $E_b = 80$ volts.

a. Determine the values of R_k, R_s, and R_L.

b. Determine the incremental tube parameters at the quiescent operating point.

c. Interelectrode capacitances are negligible, and the coupling and bypass capacitors act as short circuits to the signal components of current. Give an incremental model for the circuit, and determine the voltage amplification.

10-6. Figure 10-19 shows two useful pentode circuits in which the electrodes are used in unconventional ways. Incremental models are required for these circuits.

In each case the interelectrode capacitances are negligible. The control grid and the suppressor grid can be omitted from the model for the circuit in Fig. 10-19*a*; the control grid and the plate can be omitted from the model for the circuit in Fig. 10-19*b*. Use the current-source representation for the tubes in both cases, and label the controlling voltages for the controlled sources.

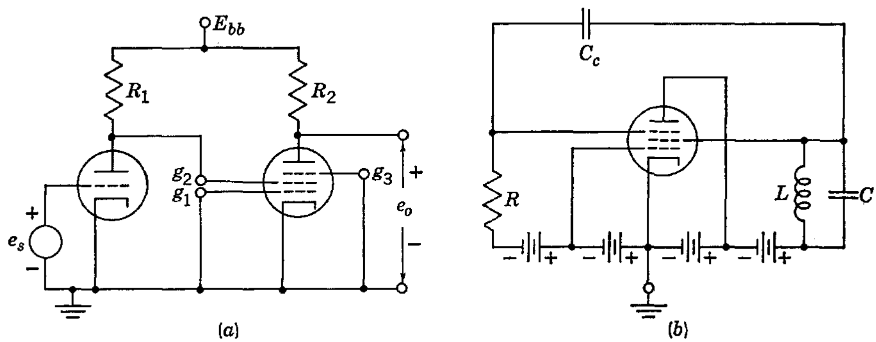


FIG. 10-19. Pentode circuits for Prob. 10-6. (*a*) Directly coupled amplifier; (*b*) transitron oscillator.

10-7. A pentode amplifier is designed in Example 10-1. The performance of this amplifier under small-signal operating conditions is to be studied, and the effect of omitting the screen-grid bypass capacitor is to be determined.

a. Evaluate g_m and g_p at the quiescent operating point.

b. Give an incremental model for the circuit using the circuit-parameter values determined in Example 10-1 and using the value R_o = 500 kilohms. The coupling and bypass capacitors are to be treated as short circuits to the time-varying components of voltage and current.

c. Determine the voltage amplification of the circuit.

d. Give a model for the circuit for the condition that the screen-grid bypass capacitor is removed (Fig. 10-18.) Assume g_{42} = 0.06 millimho and ρ = 0.4.

e. Determine the voltage amplification of the circuit with the screen-grid bypass capacitor removed.

CHAPTER 11

ANALYSIS OF PIECEWISE-LINEAR CIRCUITS

The electrical properties of tubes and transistors are presented in some detail in Chaps. 5 to 10, and the possibility of using these devices as voltage and current amplifiers is examined. Tubes and transistors are also used in a variety of important applications not directly concerned with amplification, and it is desirable to examine a few of these applications before proceding with a more detailed study of linear amplifiers. Of particular importance in this connection are the methods of analysis that can be used to obtain quantitative information regarding circuit performance.

It has been shown that in general the volt-ampere characteristics for tubes and transistors are nonlinear, although they are very nearly linear for small increments of current and voltage. These nonlinear relations are of such a nature that analysis by purely algebraic processes is impractical; hence graphical methods of analysis are often used. However, when a circuit contains more than one tube or transistor, the graphical methods become quite cumbersome. In such cases it becomes desirable to seek more efficient methods. In many applications the piecewise-linear approximations for the tubes or transistors provide effective bases for analysis. The piecewise-linear models for tubes and transistors are developed by approximating the volt-ampere characteristics of the devices by suitable families of straight line segments. The parameters of the models are chosen to give the best approximation over the range of voltage and current encountered in the application under study. Although the representation is approximate, it is quite satisfactory for many applications, and it often permits a great saving of time and effort.

The nonlinear properties of tubes and transistors are undesirable in many applications, for they may lead to distortion of the signal waveform. In other circumstances, however, these nonlinear properties are used to achieve results that cannot be obtained by any means with purely linear components. A number of such applications are presented in Chap. 2 in connection with the ideal diode; similar applications for triodes and transistors are discussed in the sections that follow.

11-1. Quiescent-point Calculations. The determination of the quiescent operating point is often an important step in the analysis of tube and transistor circuits. Graphical methods for determining the quiescent point in the basic amplifier circuits are presented in the preceding chapters. When the circuits become more complicated, however, graphical techniques become quite cumbersome. For example, the cathode-coupled amplifier of Fig. 11-1a can be analyzed by graphical means, but

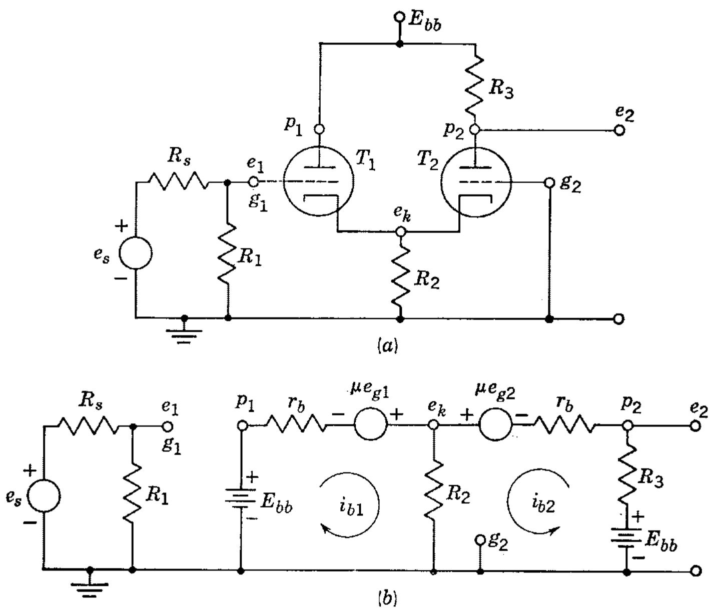


FIG. 11-1. Cathode-coupled amplifier. (a) Circuit; (b) piecewise-linear model. $e_{g1} = e_1 - e_k$ and $e_{g2} = -e_k$.

the procedure is time-consuming. On the other hand, the quiescent conditions can be determined easily by elementary circuit analysis if the tubes are replaced by suitable piecewise-linear models of the form shown in Fig. 5-16.

A piecewise-linear model for the circuit of Fig. 11-1a is shown in Fig. 11-1b. The plate-supply battery is shown in two places to simplify the form of the circuit. It is concluded from an inspection of the original circuit that both grids are biased negatively; hence both grid-circuit diodes are shown as open circuits in the model. It is further concluded from the original circuit that plate current flows in both tubes; hence the plate-circuit diodes are shown as short circuits in the model. It is necessary that these assumptions be verified when the calculations are completed. It is also assumed that the two triodes are identical.

When the circuit parameters and the applied voltages are known, the plate currents in the tubes can be determined from two loop equations. Under quiescent conditions $e_1 = 0$, and the loop equations are

$$(r_b + R_2)i_{b1} + R_2 i_{b2} = E_{bb} + \mu e_{g1} = E_{bb} - \mu e_k \qquad (11\text{-}1)$$

and

$$R_2 i_{b1} + (r_b + R_2 + R_3)i_{b2} = E_{bb} + \mu e_{g2} = E_{bb} - \mu e_k \qquad (11\text{-}2)$$

But

$$e_k = R_2(i_{b1} + i_{b2}) \qquad (11\text{-}3)$$

Substituting this value for e_k into (11-1) and (11-2) and collecting terms yields

$$[r_b + (1 + \mu)R_2]i_{b1} + (1 + \mu)R_2 i_{b2} = E_{bb} \qquad (11\text{-}4)$$

and

$$(1 + \mu)R_2 i_{b1} + [r_b + (1 + \mu)R_2 + R_3]i_{b2} = E_{bb} \qquad (11\text{-}5)$$

Given the numerical values of E_{bb} and the circuit parameters, these equations can be solved for the values of i_{b1} and i_{b2}.

It should be mentioned at this point that the solution of this problem, and others like it, can be further simplified by the use of certain of the network theorems presented in Chap. 13.

Example 11-1. The applied voltage and the parameter values in a circuit having the form shown in Fig. 11-1a are E_{bb} = 300 volts, R_s = 1 kilohm, R_1 = 1 megohm, R_2 = 2 kilohms, R_3 = 20 kilohms, μ = 20, and r_b = 10 kilohms. Determine the quiescent plate current and plate voltage for each tube.

Solution. Substituting numerical values in (11-4) and (11-5) yields

$$52i_{b1} + 42i_{b2} = 300$$

and

$$42i_{b1} + 72i_{b2} = 300$$

in which the currents are expressed in milliamperes. Solving these equations yields i_{b1} = 4.55 ma and i_{b2} = 1.51 ma.

Equation (11-3) gives the voltage at the cathodes as

$$e_k = (2)(6.06) = 12.1 \text{ volts}$$

Thus the plate voltage for T_1 is

$$e_{b1} = E_{bb} - e_k = 300 - 12.1 \approx 288 \text{ volts}$$

The output voltage is

$$e_2 = E_{bb} - R_3 i_{b2} = 300 - (20)(1.51) \approx 270 \text{ volts}$$

and the plate voltage for T_2 is

$$e_{b2} = e_2 - e_k = 270 - 12.1 \approx 258 \text{ volts}$$

11-2. The Triode Limiter. A limiter circuit employing two diodes is presented and discussed in Sec. 2-3. A triode circuit for accomplishing the same result and providing a larger output voltage is shown in Fig. 11-2a; a piecewise-linear model for the circuit is shown in Fig. 11-2b. The two diodes in this model are responsible for the limiting action. The voltage-transfer characteristic for a circuit similar to the one shown in Fig. 11-2b is presented in Example 5-4. When e_c is made more negative than a few volts, the plate current is cut off, the diode in the plate circuit acts as an open circuit, and the output voltage is limited to a maximum value equal to the plate-supply voltage. When the signal voltage goes positive, grid current flows, the diode in the grid circuit acts as a short circuit, and the voltage divider consisting of r_c and R acts to keep e_c much smaller than e_s. Thus e_c never rises much above zero for any reasonable value of e_s, and the output voltage is limited to a minimum value corresponding approximately to $e_c = 0$.

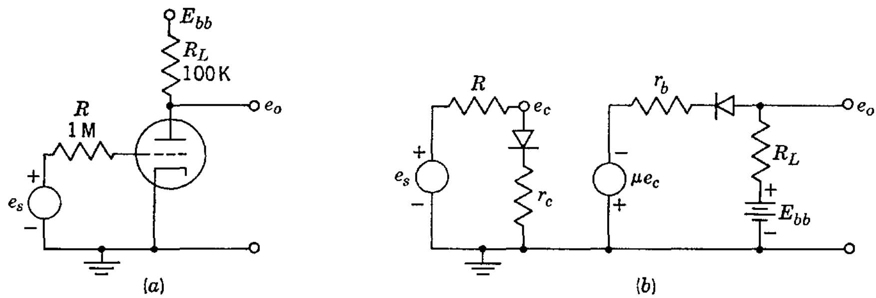


FIG. 11-2. Triode limiter. (a) Circuit; (b) piecewise-linear model. $\mu = 100$, $r_b = 75$ kilohms, and $r_c = 1$ kilohm.

When e_s is positive, the grid voltage is

$$e_c = \frac{r_c}{r_c + R} e_s \approx \frac{r_c}{R} e_s \tag{11-6}$$

where the approximation is good if $R \gg r_c$. For the parameter values specified in Fig. 11-2, $e_c = 0.001e_s$, and even when $e_s = 100$ volts, e_c is only 0.1 volt; hence e_o is only slightly affected by e_s when e_s is positive. The value of e_o corresponding to $e_c = 0$ is easily determined, for under this condition the source μe_c is zero. Accordingly,

$$e_o = \frac{r_b}{r_b + R_L} E_{bb} \qquad e_c = 0 \tag{11-7}$$

This is approximately the lower limit for e_o. For the values given in Fig. 11-2, $e_o = 64.3$ volts for $e_c = 0$.

At the point of plate-current cutoff there is no voltage drop across the plate-circuit diode; hence at this point

$$\mu e_c = -E_{bb} \tag{11-8}$$

and, since $i_b = 0$ at this point,

$$e_o = E_{bb} \qquad i_b = 0 \tag{11-9}$$

This is the upper limit on e_o. For the values given in Fig. 11-2, the upper limit on e_o is 150 volts, and the value of e_c that gives plate-current cutoff is $e_c = -E_{bb}/\mu = -1.5$ volts. Since no grid current flows at plate-current cutoff, $e_s = e_c$, and $e_s = -1.5$ volts is required for cutoff. Thus the circuit of Fig. 11-2 limits, or clips, any signal that exceeds zero volts positively or -1.5 volts negatively.

The triode limiter can be used to block the transmission of excessively large signals, and it is frequently used in waveform shaping circuits. For example, a sinusoidal signal of 15 volts amplitude applied to the limiter of Fig. 11-2 produces a reasonably good square wave at the output with a peak-to-peak amplitude of $150 - 64.3 = 85.7$ volts.

11-3. The Triode with Grid-leak Bias. The use of a resistor in the cathode circuit to provide grid bias for triode amplifiers is treated in

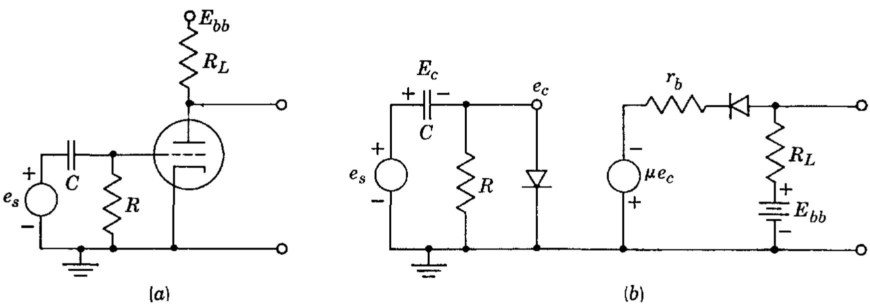


FIG. 11-3. A triode amplifier with grid-leak bias. (*a*) Circuit; (*b*) piecewise-linear model.

Chap. 6. This is by far the most common method of obtaining bias in voltage amplifiers. In some applications, however, the cathode resistor and its bypass capacitor are omitted, as illustrated in Fig. 11-3*a*, and grid bias is provided by action that takes place in the grid circuit. Bias obtained in this way is called *grid-leak* bias.

The action of the circuit of Fig. 11-3*a* can be analyzed with the aid of the piecewise-linear model shown in Fig. 11-3*b*. For the purpose of this analysis the small grid resistance of the triode, r_c, is neglected. It is seen by referring to Fig. 2-11 that the grid circuit in Fig. 11-3*b* acts like the diode clamper analyzed in Sec. 2-5. Accordingly, if a periodic signal is applied to the amplifier, the capacitor charges up to the peak value of e_s with the polarity shown in Fig. 11-3*b*, and if R_1 is large so that C cannot discharge appreciably in one cycle of e_s, the voltage across the capacitor remains essentially constant at the peak value of e_s. The grid voltage is then

$$e_c = e_s - E_c \tag{11-10}$$

and the capacitor voltage E_c serves as grid bias for the tube. Waveforms of e_s and e_c are shown in Fig. 11-4 for sinusoidal operation. If the amplitude of e_s changes, the value of E_c changes accordingly, and the positive peak of the waveform of e_s remains clamped at zero volts.

If a very large signal is applied to the circuit at e_s, the capacitor charges to a correspondingly large voltage. This voltage may be large enough to cut off the plate current when the excessive signal is removed. Such action as this is called *grid blocking*. The large voltage across C decays slowly as C discharges through the large resistance R; the rate of decay is governed by the time constant RC. If the time constant is very large, the plate current in the triode may be cut off for several seconds. Signal transmission through the amplifier is blocked during this time.

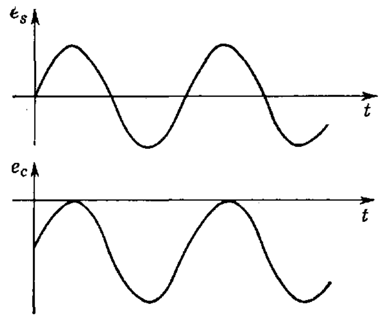


FIG. 11-4. Waveforms of voltage in an amplifier with grid-leak bias.

It is important to note that grid blocking and the clamping of the input signal by grid action are not phenomena restricted to amplifiers with grid-leak bias. Both phenomena may occur in amplifiers with cathode-resistor bias if an input coupling capacitor is used and if the input signal is sufficiently large. These actions may seriously interfere with the operation of the circuits in which they occur.

11-4. The Cathode-coupled Limiter. A simple triode limiter is presented in Sec. 11-2. In many applications, however, it is necessary to couple the signal source to the limiter through a capacitor in order to remove a d-c component of voltage from the signal. In such cases a clamping action takes place in the grid circuit in the manner described in Sec. 11-3, and as a result the limiter may not be able to perform the desired function. For example, under this circumstance it is not possible for the limiter to generate a square wave from a sinusoidal input signal.

The origin of the difficulty described above is the fact that the triode limiter depends on the flow of grid current for part of its limiting action; the clamping action would not occur if there were no grid current. The cathode-coupled limiter, which is the same circuit as the cathode-coupled amplifier shown in Fig. 11-1, provides this desirable feature. Limiting occurs at one signal level in this circuit because of plate-current cutoff in T_1, and it occurs at another signal level because of plate-current cutoff in T_2. Grid current never flows in T_2, and it flows in T_1 only when very large input signals are applied.

The voltage-transfer characteristic for the cathode-coupled limiter is shown in Fig. 11-5. It is evident from the circuit of Fig. 11-1 that T_1 is cut off if the input signal is made sufficiently negative. Under this condition e_k has its smallest value, the plate current in T_2 has its largest value, and e_2 has its smallest value as indicated in Fig. 11-5. With T_1 cut off, e_k and e_2 are unaffected by changes in e_1. As e_1 is increased above the value giving plate-current cutoff in T_1, the current in T_1 increases, and e_k increases. Thus the grid of T_2 becomes more negative with respect to its cathode, the current in T_2 decreases, and the output voltage rises. As this action continues, e_k eventually becomes so great that the plate current in T_2 is cut off. Under this condition e_2 has its maximum value, E_{bb}, and is unaffected by further increases in e_1.

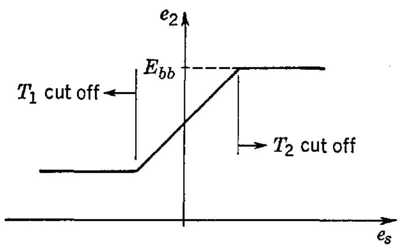


FIG. 11-5. Voltage transfer characteristic for a cathode-coupled limiter.

The analysis of the cathode-coupled limiter is greatly facilitated by the use of the piecewise-linear representation for the circuit shown in Fig. 11-1b. In fact, to construct the piecewise-linear voltage transfer characteristic it is sufficient to determine the coordinates of the two break points. Each break corresponds to one tube reaching the point of plate-current cutoff. For further simplification the circuit is shown in Fig. 11-6a for the condition that T_2 is at the cutoff point, and it is shown in

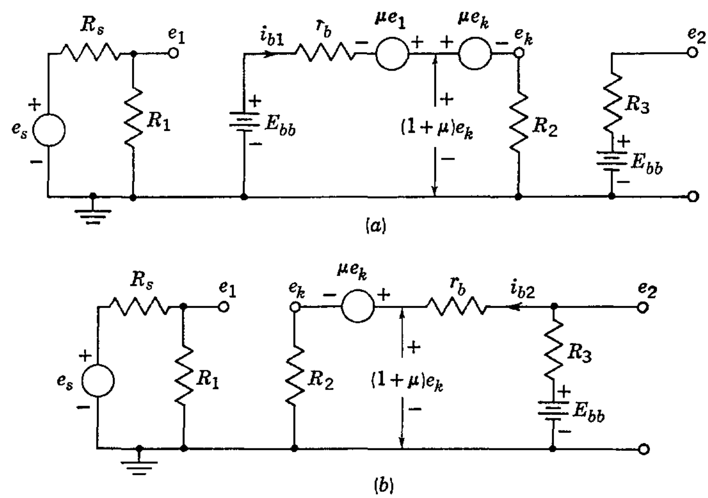


FIG. 11-6. Piecewise-linear models for the cathode-coupled limiter. (a) T_2 cut off; (b) T_1 cut off.

Fig. 11-6b for the condition that T_1 is at the cutoff point. The relations $e_{g1} = e_1 - e_k$ and $e_{g2} = -e_k$ are used in these representations, and in each case it is assumed that there is no grid current in either tube. The validity of the assumption must be checked before the analysis is complete.

When T_2 is at the point of plate-current cutoff, $i'_{b2} = 0$ and, from

Fig. 11-6a,

$$e_2' = E_{bb} \tag{11-11}$$

where the primes are used to denote that T_2 is at plate-current cutoff. Under this condition Fig. 11-6b indicates that there is no voltage drop across r_b and R_3; accordingly

$$(1 + \mu)e_k' = E_{bb} \tag{11-12}$$

and

$$e_k' = \frac{E_{bb}}{1 + \mu} \tag{11-13}$$

This is the value of e_k required to produce plate-current cutoff in T_2. The plate current in T_1 required to develop this value of e_k is

$$i_{b1}' = \frac{e_k'}{R_2} = \frac{E_{bb}}{(1 + \mu)R_2} \tag{11-14}$$

Also, a loop equation for the circuit in Fig. 11-6a yields

$$r_b i_{b1}' + (1 + \mu)e_k' = E_{bb} + \mu e_1' \tag{11-15}$$

Substituting (11-12) into (11-15) yields the simple relation

$$r_b i_{b1}' = \mu e_1'$$

and hence, using (11-14),

$$e_1' = \frac{r_b}{\mu} i_{b1}' = \frac{r_b E_{bb}}{\mu(1 + \mu)R_2} \tag{11-16}$$

Thus the coordinates of one break in the transfer characteristic, e_1' and e_2', are established.

When T_1 is at the point of plate-current cutoff, $i_{b1}'' = 0$, and since there is no voltage drop across r_b in Fig. 11-6a,

$$(1 + \mu)e_k'' = E_{bb} + \mu e_1'' \tag{11-17}$$

A loop equation for the circuit in Fig. 11-6b yields

$$(r_b + R_3)i_{b2}'' + (1 + \mu)e_k'' = E_{bb} \tag{11-18}$$

Substituting (11-17) into (11-18) yields the simple relation

$$(r_b + R_3)i_{b2}'' + \mu e_1'' = 0$$

and hence

$$e_1'' = -\frac{r_b + R_3}{\mu} i_{b2}'' \tag{11-19}$$

But with $i_{b1}'' = 0$, $e_k'' = R_2 i_{b2}''$. Substituting this relation into (11-18) yields

$$(r_b + R_3)i_{b2}'' + (1 + \mu)R_2 i_{b2}'' = E_{bb}$$

and

$$i_{b2}'' = \frac{E_{bb}}{r_b + (1 + \mu)R_2 + R_3} \tag{11-20}$$

Knowing i''_{b2}, e''_1 and e''_2 can be obtained at once. From (11-19)

$$e''_1 = -\frac{r_b + R_3}{\mu} \frac{E_{bb}}{r_b + (1 + \mu)R_2 + R_3} \tag{11-21}$$

and from Fig. 11-6*b*

$$e''_2 = E_{bb} - R_3 i''_{b2} = \frac{r_b + (1 + \mu)R_2}{r_b + (1 + \mu)R_2 + R_3} E_{bb} \tag{11-22}$$

Thus the coordinates of the second break in the transfer characteristic are established, and the characteristic can be completed by constructing the appropriate straight line segments.

Another point of interest is the value of e_1 at which grid current begins to flow in T_1. Grid current begins to flow when $e_{g1} = 0$, and under this condition in a properly designed limiter $i_{b2} = 0$. Thus the circuit in Fig. 11-6*a* applies for this condition, and from the fact that $e_{g1} = 0$ it follows that

$$e_1 = e_k = \frac{R_2}{r_b + R_2} E_{bb} \tag{11-23}$$

when grid current begins to flow in T_1.

It should be noted in passing that certain of the network theorems presented in Chap. 13 can be used to further simplify the analysis of the cathode-coupled limiter.

Example 11-2. The grid of triode T_2 in the limiter of Fig. 11-7*a* is provided with a bias $e_3 = 20$ volts relative to ground. This arrangement permits a larger resistance to be used for R_2, which in turn permits larger values of e_1 without grid current in T_1. The coordinates of the breaks in the voltage-transfer characteristic, e_2 versus e_1, are to be determined, and the value of e_1 at which grid current begins to flow is to be calculated.

Solution. A piecewise-linear model for the circuit is shown in Fig. 11-7*b*. It is assumed that grid current does not flow in either tube, and the plate-circuit diodes are treated as short circuits. The voltage divider providing e_3 is omitted from the model, for it plays no part in the operation of the circuit other than fixing the value of e_3 at 20 volts. The transfer characteristic has the same form as the one shown in Fig. 11-5, and the analysis follows the procedure outlined above with minor modifications to account for the source μe_3.

When T_2 is at the point of plate-current cutoff, $i'_{b2} = 0$, and

$$e'_2 = E_{bb} = 300 \text{ volts}$$

Under this condition there is no voltage drop across R_3 and the plate resistance of T_2; hence

$$(1 + \mu)e'_k - \mu e_3 = E_{bb}$$

giving

$$e'_k = \frac{300 + 400}{21} = 33.3 \text{ volts}$$

The current in R_2 under this condition is i'_{b1}; thus

$$i'_{b1} = \frac{e'_k}{R_2} = \frac{33.3}{5} = 6.66 \text{ ma}$$

The equation for the left-hand loop in Fig. 11-7*b* can be written as

$$r_b i'_{b1} + (1 + \mu)e'_k = E_{bb} + \mu e'_1$$

Substituting the relation $(1 + \mu)e'_k = E_{bb} + \mu e_3$ yields

$$\begin{aligned} r_b i'_{b1} + \mu e_3 &= \mu e'_1 \\ (10)(6.66) + (20)(20) &= 20e'_1 \\ e'_1 &= 23.3 \text{ volts} \end{aligned}$$

This solution yields $e'_k = 33.3$ volts, $e'_1 = 23.3$ volts, and $e_3 = 20$ volts. Thus both grids are negative relative to the cathodes, and the initial assumption of no grid

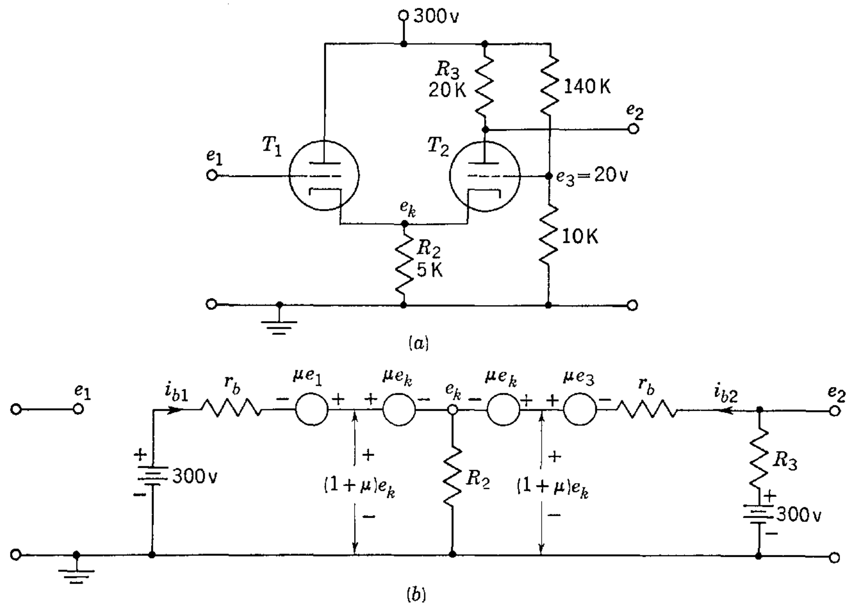


FIG. 11-7. Limiter for Example 11-2. (*a*) Circuit; (*b*) model. $r_b = 10$ kilohms, $\mu = 20$, $e_{g1} = e_1 - e_k$, and $e_{g2} = e_3 - e_k$.

current is valid. The right-hand break in the transfer characteristic therefore occurs at $e'_1 = 23.3$ volts and $e'_2 = 300$ volts.

When T_1 is at the point of plate-current cutoff, $i''_{b1} = 0$, and under this condition the equation for the right-hand loop in Fig. 11-7*b* can be written as

$$(r_b + R_3)i''_{b2} + (1 + \mu)e''_k = E_{bb} + \mu e_3$$

But with $i''_{b1} = 0$, $e''_k = R_2 i''_{b2}$. Substituting this relation in the loop equation yields

$$\begin{aligned} 30i''_{b2} + (21)(5)i''_{b2} &= 300 + 400 \\ i''_{b2} &= {}^{700}\!/_{135} = 5.19 \text{ ma} \end{aligned}$$

Thus

$$\begin{aligned} e''_2 &= E_{bb} - R_3 i''_{b2} = 300 - (20)(5.19) \\ &= 196 \text{ volts} \end{aligned}$$

Also, with $i''_{b1} = 0$, there is no voltage drop across the plate resistance of T_1; hence

$$(1 + \mu)e''_k = E_{bb} + \mu e''_1$$

But $$e_k'' = R_2 i_{b2}'' = (5)(5.19) = 26 \text{ volts}$$

Hence $$e_1'' = \frac{(21)(26) - 300}{20} = 12.3 \text{ volts}$$

This solution yields $e_k'' = 26$ volts, $e_1'' = 12.3$ volts, and $e_3 = 20$ volts. Thus both grids are negative relative to the cathodes, and the initial assumption of no grid current is valid. The left-hand break in the transfer characteristic therefore occurs at $e_1'' = 12.3$ volts and $e_2'' = 196$ volts.

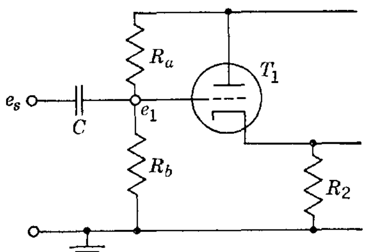


FIG. 11-8. Modified form of the limiter of Example 11-2.

Grid current begins to flow in T_1 when $e_{g1} = 0$; under this condition $i_{b2} = 0$, and

$$e_1 = e_k = \frac{R_2}{r_b + R_2} E_{bb} = (5/15)(300) = 100 \text{ volts}$$

It is often necessary to have the limiting action take place at equal increments of the signal voltage above and below zero. This result can be achieved by providing a bias voltage for the grid of T_1 in the manner shown in Fig. 11-8. The coupling capacitor C prevents the source of signals from altering the bias voltage at e_1. In the circuit under study, the voltage divider should be designed to make the quiescent value of e_1 be

$$e_{1o} = \frac{e_1' + e_1''}{2} = 17.8 \text{ volts}$$

If the amplitude of e_s in this circuit is such that e_1 exceeds 100 volts, grid current flows, and the d-c component of e_1 changes as a result of the clamping action of the grid.

11-5. Saw-tooth Generators. There are many kinds of electronic equipment that require a voltage having a saw-tooth waveform. For example, such a voltage is used in the timing circuits of certain types of

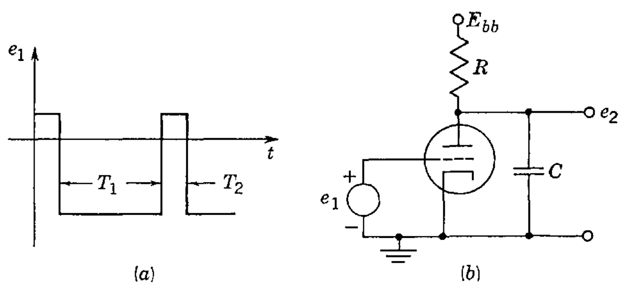


FIG. 11-9. Triode saw-tooth generator. (a) Grid voltage; (b) circuit.

radar receivers, and saw-tooth voltages are generated by the sweep circuits used in cathode-ray oscilloscopes. Because of the wide need for saw-tooth voltages, many kinds of saw-tooth generators have been developed and find use in various applications. One of the simplest of these generators is shown in Fig. 11-9. With the indicated voltage e_1 applied between grid and cathode, the circuit acts in the following way. During the interval T_1, e_1 is sufficiently negative to cut off the plate current in the triode. Thus during this interval the capacitor charges relatively slowly through the resistance R. During the interval T_2, e_1 has a value that permits a large plate current in the triode. Thus during this interval the capacitor discharges rapidly through the triode. The resulting output voltage, which is also the voltage across the capacitor, may have a waveform like the one pictured in Fig. 11-10*a*.

The triode saw-tooth generator can be analyzed quantitatively with the aid of the diagrams shown in Fig. 11-10. During the interval that

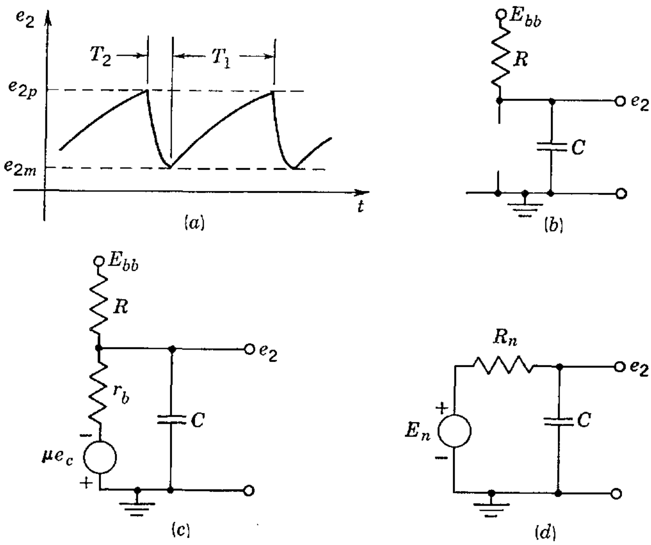


FIG. 11-10. Analysis of the triode saw-tooth generator. (*a*) Output voltage; (*b*) tube cut off; (*c*) tube conducting; (*d*) Thevenin equivalent circuit.

the tube is cut off the circuit reduces to the one shown in Fig. 11-10*b*; thus during this interval C charges through R toward a final voltage equal to E_{bb}. Choosing $t = 0$ at the beginning of the charging interval, the equation for e_2 during this interval has the form

$$e_2 = A + B\epsilon^{-t/RC} \tag{11-24}$$

If C is permitted to charge indefinitely, e_2 approaches E_{bb}; hence, letting t tend to infinity in (11-24) yields

$$A = E_{bb} \tag{11-25}$$

When $t = 0$, e_2 has its minimum value, e_{2m}, as indicated in Fig. 11-10*a*; hence, letting $t = 0$ in (11-24),

$$A + B = e_{2m}$$
$$B = e_{2m} - E_{bb} \tag{11-26}$$

Thus while the tube is cut off

$$e_2 = E_{bb} - (E_{bb} - e_{2m})\epsilon^{-t/RC} \tag{11-27}$$

The exponential in (11-27) can be expanded in a power series to give

$$e_2 = E_{bb} - (E_{bb} - e_{2m})\left[1 - \frac{t}{RC} + \frac{1}{2}\left(\frac{t}{RC}\right)^2 - \cdots\right] \tag{11-28}$$

Thus it is clear that if e_2 is to rise linearly with time, as is usually desired, the circuit must be designed so that the square and higher-degree terms in t are negligibly small; that is, R and C must be chosen so that $t \ll RC$ for all values of t in the charging interval. If this is done, Eq. (11-28) reduces to

$$e_2 \approx e_{2m} + (E_{bb} - e_{2m})\frac{t}{RC} \tag{11-29}$$

The peak value of e_2 is then

$$e_{2p} \approx e_{2m} + (E_{bb} - e_{2m})\frac{T_1}{RC} \tag{11-30}$$

During the time the tube is conducting it can be represented by its piecewise-linear model, as shown in Fig. 11-10*c*, provided the plate voltage always remains somewhat greater than the grid voltage. The evaluation of e_2 as a function of time while the tube is conducting is facilitated by replacing the circuit connected to C by its Thevenin equivalent as shown in Fig. 11-10*d*. The parameters of this circuit are

$$R_n = \frac{r_b R}{r_b + R} \tag{11-31}$$

and

$$E_n = \frac{r_b E_{bb} - R\mu e_c}{r_b + R} \tag{11-32}$$

It follows from Fig. 11-10*d* that if $t = 0$ when the tube begins conducting, then while the tube is conducting e_2 has the form

$$e_2 = A + B\epsilon^{-t/R_nC} \tag{11-33}$$

If the discharge continues indefinitely, e_2 approaches the value E_n; hence, letting t tend to infinity in (11-33) yields

$$A = E_n \tag{11-34}$$

With $t = 0$, e_2 has its peak value, e_{2p}, as indicated in Fig. 11-10*a*; hence, letting $t = 0$ in (11-33),

$$A + B = e_{2p}$$
$$B = e_{2p} - E_n \tag{11-35}$$

Thus while the tube is conducting

$$e_2 = E_n + (e_{2p} - E_n)\epsilon^{-t/R_nC} \tag{11-36}$$

The minimum value of e_2 is, accordingly,

$$e_{2m} = E_n + (e_{2p} - E_n)\epsilon^{-T_2/R_nC} \tag{11-37}$$

If the discharge period is relatively long so that $T_2 \gg R_nC$, then

$$e_{2m} \approx E_n \tag{11-38}$$

In most applications it is desirable that the charging portion of the saw tooth be as nearly linear as possible and that the discharging portion be completed as rapidly as possible. These objectives would be met completely if the capacitor were charged from an ideal current source and if it were discharged through a short circuit. The transistor saw-tooth generator shown in Fig. 11-11 provides an approximation to the ideal conditions. With the indicated voltage e_1 applied to the circuit, the following action takes place. During the interval T_1 the transistor operates in the normal manner with a collector current that is nearly

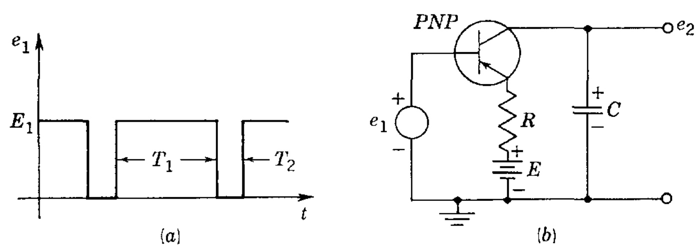


FIG. 11-11. Transistor saw-tooth generator. (a) Controlling voltage; (b) circuit.

constant and independent of the collector voltage. This current charges the capacitor in a nearly linear manner with the polarity shown. During the interval T_2 the voltage on the capacitor biases the collector in the *forward* direction, the direction of easy conduction, and the capacitor discharges rapidly through this junction and the base terminal.

The transistor saw-tooth generator can be analyzed quantitatively with the aid of the diagrams shown in Fig. 11-12. The circuit in Fig. 11-12*a* is a piecewise-linear representation for the circuit during the interval T_1. During this interval the controlling voltage E_1 is less than the d-c supply voltage E, and the emitter current is negative. As long as the output voltage e_2 is less than E_1, the collector-circuit diode is biased in the reverse direction, and $i_C = \alpha_{CE}i_E$ is negative and charges the capacitor with the polarity shown. Thus during this interval the charging current is

$$i_C = \alpha_{CE}i_E = \alpha_{CE}\frac{E_1 - E}{R} \tag{11-39}$$

With E_1 less than E, this current is negative and constant. The output voltage is therefore

$$e_2 = \frac{q}{C} = \frac{i_C t}{C} = \frac{\alpha_{CE}}{RC}(E_1 - E)t \qquad (11\text{-}40)$$

If this voltage remains less than E_1 during the interval T_1, its peak value is

$$e_{2p} = \frac{\alpha_{CE}}{RC}(E_1 - E)T_1 \qquad (11\text{-}41)$$

During the interval T_2 the controlling voltage is zero; thus during this interval the voltage on the capacitor biases the collector-circuit diode in the forward direction and the diode acts as a short circuit. The corresponding circuit representation is shown in Fig. 11-12*b*; it follows from this circuit that the capacitor discharges instantaneously when the control voltage drops to zero. The waveform of the output voltage, shown in

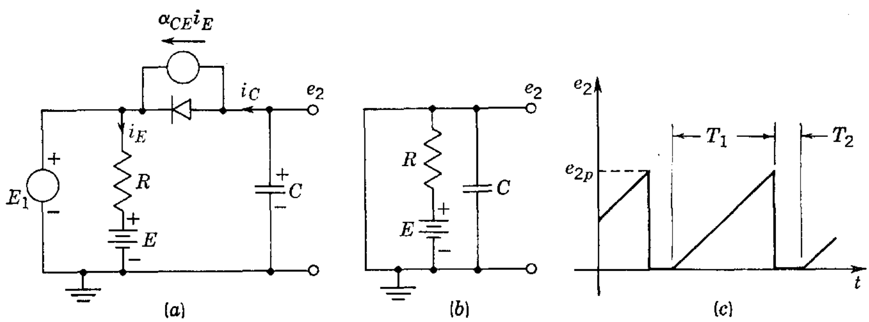


FIG. 11-12. Analysis of the transistor saw-tooth generator. (*a*) Circuit during charging interval; (*b*) circuit during discharging interval; (*c*) output voltage.

Fig. 11-12*c*, is an ideal saw tooth with a perfectly linear rise during the interval T_1 and an instantaneous drop at the end of the interval. In practice this perfect waveform connot be obtained, of course, because of parasitic elements not included in the analysis, although the ideal may be approached rather closely. During the charging interval the transistor does not behave as a perfect current source, and, moreover, any leakage resistance shunting the capacitor and any load resistance connected to the circuit act to prevent the circuit connected to C from behaving as an ideal current source. In addition, the capacitor cannot in fact discharge instantaneously, for there is some resistance associated with the transistor and with the source supplying the control voltage e_1. If the circuit is designed carefully, however, very good performance can be obtained.

The two saw-tooth generators described above require a voltage of special waveform to control the action of the triode and the transistor. The thyratron saw-tooth generator shown in Fig. 11-13 has the ability to generate a periodic saw-tooth voltage without the need for a special controlling voltage. The repetition rate for the saw tooth is determined by the circuit itself. Circuits of this type are often called *free-running* saw-tooth generators.

If the thyratron is nonconducting when the voltage across C is small, the negative grid-to-cathode potential holds the tube cut off. The tube remains nonconducting while the capacitor charges through the resistance R until the capacitor voltage, which is also the plate-to-cathode voltage, becomes great enough to cause the tube to ignite. The voltage at which ignition occurs depends on the grid voltage E, in accordance with the control characteristic for the thyratron. A typical control characteristic is shown in Fig. 3-22.

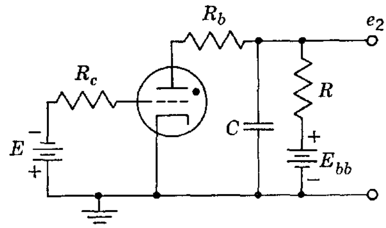


FIG. 11-13. Thyratron saw-tooth generator.

When the tube ignites and becomes conducting, the capacitor discharges through the tube at a rate depending on the size of the current-limiting resistor, R_b. Near the end of the discharge the current becomes too small to maintain a plasma in the tube, the grid regains control of the tube, and the capacitor begins to charge again while the grid holds the tube cut off. This cycle of events is repeated periodically.

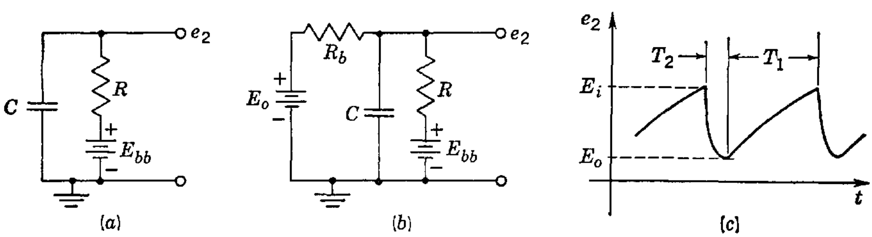


FIG. 11-14. Analysis of the thyratron saw-tooth generator. (*a*) Charging circuit; (*b*) discharging circuit; (*c*) output voltage.

The thyratron saw-tooth generator can be analyzed quantitatively with the aid of the diagrams shown in Fig. 11-14. During the charging interval T_1 the tube is nonconducting, and the output voltage can be determined from the circuit of Fig. 11-14*a*. This circuit has the same form as the one in Fig. 11-10*b* representing the triode saw-tooth generator; hence the results of the triode analysis can be adapted to the present case yielding

$$e_2 \approx E_o + (E_{bb} - E_o)\frac{t}{RC} \tag{11-42}$$

provided $t \ll RC$. The peak value of the saw-tooth wave, which is also the ignition voltage for the thyratron, is

$$E_i \approx E_o + (E_{bb} - E_o)\frac{T_1}{RC} \tag{11-43}$$

The duration of the charging interval is thus

$$T_1 \approx RC\frac{E_i - E_o}{E_{bb} - E_o} \tag{11-44}$$

If the discharge interval T_2 is negligible in comparison with T_1, the frequency of the saw-tooth wave is

$$f \approx \frac{1}{T_1} \approx \frac{1}{RC}\frac{E_{bb} - E_o}{E_i - E_o} \tag{11-45}$$

During the discharging interval the voltage drop across the tube is nearly constant and independent of the tube current (see Fig. 3-29 and the related discussion); hence the tube can be represented by a voltage source with a terminal voltage of E_o volts as shown in Fig. 11-14*b*.

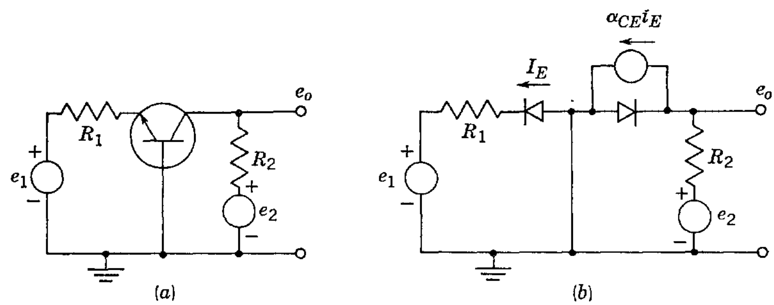


FIG. 11-15. Pulse-amplitude modulator. (*a*) Circuit; (*b*) piecewise-linear model.

Typical values for E_o lie between 10 and 15 volts. The current through the thyratron during the discharge interval is limited to a safe value by the resistance R_b. In the usual case R_b is much smaller than R; thus C discharges toward a voltage that is approximately equal to E_o, and the tube deionizes and ceases conducting when the discharge current drops below the minimum value required to maintain ionization in the tube. In order for the tube to deionize it is necessary that the resistance of R in series with R_b limit the tube current to a value small enough to permit deionization. The waveform of the output voltage is shown in Fig. 11-14*c*.

11-6. A Transistor Pulse-amplitude Modulator. A pulse-amplitude modulator using a transistor is shown in Fig. 11-15*a*; a piecewise-linear model for the circuit is shown in Fig. 11-15*b*. The function of this circuit is to cause the amplitude of a train of pulses, shown as e_2 in Fig. 11-16, to vary in accordance with a signal voltage, shown as e_1 in Fig.

11-16. The result is an amplitude-modulated pulse train, shown as e_o in Fig. 11-16, having an *envelope* that varies with time in the same manner as the signal voltage e_1. In certain respects it is easier to work with the signal in the form of an amplitude-modulated wave than in its original form. The original signal can be recovered from the amplitude-modulated wave by the process of demodulation, or detection, discussed in Sec. 11-7.

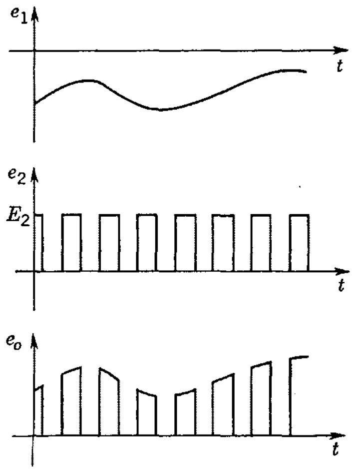


FIG. 11-16. Voltage waveforms in the pulse-amplitude modulator.

The operation of the modulator can be analyzed with the aid of the model shown in Fig. 11-15*b*. The signal is applied at the input of the circuit, and the collector, instead of being supplied with a direct voltage as in an amplifier, is supplied with a train of voltage pulses of uniform amplitude. During the intervals in which e_2 is zero, there is no reverse bias across the collector-circuit diode, the diode acts as a short circuit to the source $\alpha_{CE}i_E$, and the output voltage is zero. During the intervals in which e_2 has the value E_2, the transistor acts as a normal amplifier and gives an output voltage that is proportional to the input voltage. Specifically, when $e_2 = E_2$,

$$\begin{aligned} e_o' &= E_2 - \alpha_{CE} i_E R_2 \\ &= E_2 + \frac{\alpha_{CE} R_2}{R_1} e_1 \end{aligned} \tag{11-46}$$

This is the equation of the envelope of the modulated pulse train. If

$$e_1 = E + E_1 \cos \omega_1 t \tag{11-47}$$

the envelope is given by

$$\begin{aligned} e_o' &= E_2 + \frac{\alpha_{CE} R_2}{R_1} (E + E_1 \cos \omega_1 t) \\ &= E_o(1 + m \cos \omega_1 t) \end{aligned} \tag{11-48}$$

The quantity

$$m = \frac{E_1 \alpha_{CE} R_2 / R_1}{E_2 + E \alpha_{CE} R_2 / R_1} \tag{11-49}$$

is the *modulation index;* its value is normally less than unity so that the envelope never drops to zero.

It is important to obtain and examine a complete expression for the output voltage, not just an expression for the envelope. This can be done easily by noting from Fig. 11-16 that the output voltage can be obtained from the envelope by multiplying the envelope by a train of rectangular pulses in which the amplitude is alternately unity and zero. This pulse train can be expressed in terms of a Fourier series. If $t = 0$ is chosen at the mid-point of a pulse, this series has the form

$$f(t) = a_0 + a_1 \cos \omega_2 t + a_3 \cos 3\omega_2 t + \cdots \tag{11-50}$$

where ω_2 is the fundamental frequency of e_2, the train of voltage pulses applied to the collector circuit. It then follows that the output voltage is

$$e_o = [e_o'(t)][f(t)] \tag{11-51}$$

If the envelope has the form given by (11-48), then

$$\begin{aligned} e_o &= [E_o(1 + m \cos \omega_1 t)][f(t)] \\ &= a_0 E_o(1 + m \cos \omega_1 t) \\ &\quad + a_1 E_o(1 + m \cos \omega_1 t)(\cos \omega_2 t) \\ &\quad + a_3 E_o(1 + m \cos \omega_1 t)(\cos 3\omega_2 t) \\ &\quad + \cdots \end{aligned} \tag{11-52}$$

This expression for e_o is not very useful, for it would be difficult indeed to analyze the response of a circuit to a signal of this form. The difficulty can be avoided easily, however, for it is a simple matter to modify (11-52) so that e_o is expressed as a sum of sinusoidal components. The response of linear circuits to this signal can then be analyzed by conventional means through treating each sinusoidal component separately and superposing the results. Thus the first term in (11-52) can be written as

$$e_{o1} = a_0 E_o + m a_0 E_o \cos \omega_1 t \tag{11-53}$$

Multiplying out the second term and using the trigonometric identity for cos (A) cos (B) yields

$$\begin{aligned} e_{o2} = a_1 E_o \cos \omega_2 t &+ \frac{m}{2} a_1 E_o \cos (\omega_2 + \omega_1)t \\ &+ \frac{m}{2} a_1 E_o \cos (\omega_2 - \omega_1)t \end{aligned} \tag{11-54}$$

Similarly, the third term becomes

$$\begin{aligned} e_{o3} = a_3 E_o \cos 3\omega_2 t &+ \frac{m}{2} a_3 E_o \cos (3\omega_2 + \omega_1)t \\ &+ \frac{m}{2} a_3 E_o \cos (3\omega_2 - \omega_1)t \end{aligned} \tag{11-55}$$

The higher-order terms in the expression for e_o take similar forms. The diagram shown in Fig. 11-17 indicates the frequencies and relative amplitudes of the sinusoidal components associated with the first three terms of (11-52) under typical conditions. This pattern is repeated, with diminishing amplitude, at all odd multiples of ω_2.

Modulators are often used in conjunction with a bandpass filter as illustrated in Fig. 11-18*a*. The filter is a circuit designed to pass signals having frequencies lying in a band centered at ω_2 and to reject signals at all other frequencies, including ω_1, $3\omega_2$, and so forth. Thus the output from the filter consists solely of the second term in (11-52), and

$$e_f = a_1E_o(1 + m \cos \omega_1 t)(\cos \omega_2 t) \tag{11-56}$$

The waveform of this voltage is shown in Fig. 11-18*b*. This signal consists of a sinusoidal voltage at the frequency ω_2, called the *carrier* voltage, and sinusoidal components at the frequencies $\omega_2 + \omega_1$ and $\omega_2 - \omega_1$, called the *sideband* components.

If the signal voltage e_1 consists of the sum of many sinusoidal components, the analysis can be carried through in the same manner as that

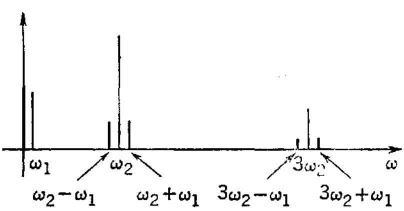


FIG. 11-17. Spectrum of a pulse-amplitude modulated wave.

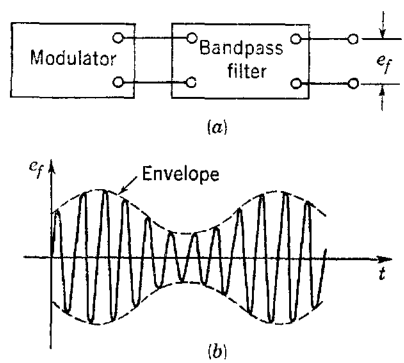


FIG. 11-18. Modulator with bandpass filter. (*a*) Block diagram; (*b*) filter output.

employed above; the only difference is that there are more terms involved. The result shows that all the components in the original signal appear in the upper sideband with relative amplitudes preserved, and they appear again in the lower sideband. The envelope of the modulated wave has the same waveform as the original signal. Thus it is often useful to view modulation as a frequency translation.[1,2] A signal in the audio-frequency band, for example, can be translated by modulation to the radio-frequency band for efficient transmission by radio waves. One of the first operations performed on a signal by a superheterodyne radio receiver is to translate the signal from its original frequency to a band centered at 455 kcps, the intermediate frequency, for efficient amplification. When interest is centered primarily on frequency translation, the term *heterodyne* is often used in place of *modulate*. As a further point of terminology, the difference frequencies produced by modulation are often called *beat frequencies*.

An important feature of modulation is that fact that the output of the modulator contains frequencies that are not present in the input. Modulation is thus a nonlinear operation, and it cannot be performed by linear circuits. Equation (11-56), for example, shows that modulation involves the multiplication of two time functions. Thus an ideal multiplier is an ideal modulator. Unfortunately, ideal multipliers are not easy to contrive; hence modulators with appropriate filters have been used as multipliers in some analog computers.

11-7. The Diode Detector. In most applications where modulated signals are employed it is necessary ultimately to recover the information, or original signal, carried by the modulated wave. The demodulator, or detector, is a device to recover the original signal. It is clear that the output of the demodulator must contain frequencies not contained in the input; hence it must be a nonlinear device. The information contained in the modulated wave cannot be recovered with a linear circuit. On the other hand, almost any nonlinearity produces some kind of demodulation.

The most commonly used demodulator is the peak rectifier shown in Fig. 11-19. This circuit is discussed in Sec. 2-4, and the action of the

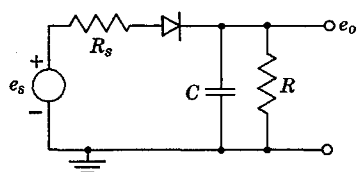


FIG. 11-19. The diode detector.

filter capacitor is analyzed in Sec. 2-10. Ideally, the output voltage from the peak rectifier is equal to the peak value of the input voltage. Thus, ideally, e_o should follow the envelope of the modulated wave exactly, and e_o should then be a duplicate of the original signal. The actual behavior of a well-designed detector is usually a good approximation to this ideal.

Any realistic analysis of the peak detector[1,2] becomes rather complicated; hence an analysis is not undertaken here. It is possible, however, to point out one of the important considerations in the design of these circuits. In order to minimize the ripple in the output of the detector, it is desirable to make R and C as large as possible. However, with R and C large, e_o cannot decrease rapidly when the envelope of the modulated wave decreases. If R and C are very large, and if the envelope decreases rapidly, then e_o is unable to follow the envelope, and the output voltage is no longer a duplicate of the original signal. This form of signal distortion is known as *diagonal clipping*. The values of R and C must therefore be chosen on the basis of a reasonable compromise between good filtering and the danger of diagonal clipping.

11-8. Summary. Tubes and transistors find important applications as linear amplifiers. They also find many important uses that are not related to amplification. A few of these applications are presented in the preceding sections; they include limiters, clampers, waveform generators, modulators, and detectors.

In each of these applications the nonlinear nature of the tube or transistor is an essential feature of the circuit; hence the methods of linear circuit analysis cannot be employed. These circuits can be analyzed and designed by graphical methods, but the process is likely to be laborious and tedious. On the other hand, piecewise-linear approximations are often permissible. These approximations usually provide added insight into the properties of the circuits, and in many cases they simplify the procedures of analysis and design.

REFERENCES

1. Terman, F. E.: "Electronic and Radio Engineering," McGraw-Hill Book Company, Inc., New York, 1955.
2. Gray, T. S.: "Applied Electronics," 2d ed., John Wiley & Sons, Inc., New York, 1954.

PROBLEMS

11-1. The quiescent operating conditions in the circuit of Fig. 11-20 are to be determined. For this purpose the piecewise-linear representation for the circuit can be used. The tubes are identical with $\mu = 70$ and $r_b = 50$ kilohms. Find the quiescent values of the output voltage, e_3, the plate currents, and the plate-to-cathode voltages.

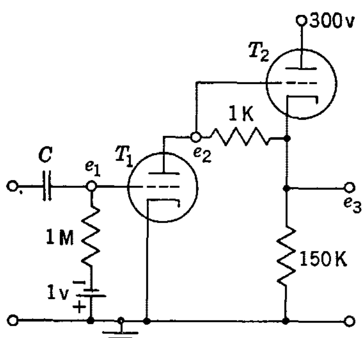


FIG. 11-20. Circuit for Prob. 11-1.

11-2. The plate-supply voltage in an amplifier like the one shown in Fig. 11-3 is 200 volts, and the load resistance is $R_L = 200$ kilohms. The input signal is $e_s = 2 \cos \omega_s t$ volts, and R and C are so large that C cannot discharge appreciably through R during one cycle of the signal. The behavior of the circuit is to be studied with the aid of the piecewise-linear representation with $\mu = 100$ and $r_b = 70$ kilohms.

a. Sketch and dimension the waveform of the grid voltage, e_c.

b. Sketch and dimension the waveform of the output voltage, e_b.

c. What is the greatest amplitude that the sinusoidal input signal can have without waveform distortion?

11-3. A certain cathode-coupled limiter has the form shown in Fig. 11-7*a* with $R_2 = 5$ kilohms, $R_3 = 15$ kilohms, and $E_{bb} = 300$ volts. The circuit is further adjusted so that $e_3 = 46.5$ volts. The characteristics of this circuit are to be determined with the aid of the piecewise-linear model. The tubes are identical with $\mu = 40$ and $r_b = 7$ kilohms. The voltage transfer characteristic has the form shown in Fig. 11-5. Determine the coordinates of the breaks in this characteristic. Sketch the characteristic approximately to scale. At what value of e_1 does grid current begin to flow in T_1?

11-4. The behavior of the thyratron saw-tooth generator shown in Fig. 11-13 is to be examined. When the tube is conducting the plate-to-cathode voltage drop is 15 volts, and the tube ceases to conduct when the plate current becomes less than 0.5 ma. The plate-supply voltage is $E_{bb} = 300$ volts, and the control characteristic for the thyratron is a straight line having the equation $e_b = -10e_c$.

a. If the peak value of the output voltage is to be 75 volts, what must be the voltage of the grid-bias battery, E?

b. If $R = 1$ megohm, what value of C is required to make the charging interval last for 1 msec? This adjustment gives a repetition rate for the saw-tooth wave of about 1000 cps.

c. Describe briefly three ways in which the repetition rate can be made adjustable. Which of these affect the amplitude of the saw tooth?

11-5. The saw-tooth generator of Prob. 11-4 is adjusted in accordance with parts *a* and *b* of that problem. However, as the capacitor ages, its insulation becomes leaky, and after a period of time it behaves as an ideal capacitor in parallel with a resistance of 200 kilohms. The operation of the circuit is seriously affected by this change.

Show by a sketch the form of the output voltage as a function of time under this condition, and explain the action of the circuit. The Thevenin equivalent for the circuit connected to the ideal capacitor during the charging interval may be helpful in this analysis. Note that Eq. (11-42) does not describe the action of the circuit in this case, but an equation of the form of (11-24) is applicable if it should be needed.

11-6. The signal voltage applied to a modulator of the form shown in Fig. 11-15 has the form

$$e_1 = E + E_a \cos (200\pi t) + E_b \cos (300\pi t)$$

The voltage e_2 has the form shown in Fig. 11-16, and its repetition rate is 10,000 cps. List all the frequencies less than 20,000 cps that appear at the output. Give these in cycles per second.

11-7. An AM radio signal at 1-megacycle/sec picked up by a radio receiver is to be translated to the intermediate frequency of 455 kcps. The operation is to be performed by a modulator similar to the one shown in Fig. 11-15 used in conjunction with a filter as shown in Fig. 11-18. The voltage e_2 required by the modulator is generated by a local oscillator in the receiver. What must be the repetition rate of the local oscillator if the desired result is to be achieved? There are several possible answers; the frequency commonly used is the smallest one of these that is greater than the frequency of the radio signal.

CHAPTER 12

SYSTEMATIC ANALYSIS OF LINEAR ELECTRONIC CIRCUITS

The problem of circuit analysis is that of determining the currents and voltages existing in a specified circuit. The first step in the analysis of electronic circuits is to formulate a suitable network model representing the physical system; Chaps. 4 to 10 are largely concerned with the physical behavior of tubes and transistors and with the network models that can be used to represent them. The models for linear electronic circuits in general contain R, L, C, independent current and voltage sources, and controlled current and voltage sources. Since these circuits are linear and obey Kirchhoff's laws, the loop and node methods of analysis provide general systematic procedures for analyzing any circuit, no matter how complicated. One of the objectives of this chapter is to review these systematic methods of analysis and to examine the way in which the presence of controlled sources affects the forms of the equations and the procedures of analysis.

The problem of electronic circuit design is that of specifying in detail a circuit to perform some stated function. It usually involves choosing a type of circuit on the basis of previous experience and then using the techniques of analysis to determine suitable values for the circuit components. Thus the engineer should know as much as possible about the characteristics of various circuits so that he may make the best initial choice of circuit type, and he should be skilled in circuit analysis to facilitate the detailed computations. Both of these ends are usually served best by the analytical procedure that gives the most insight into the behavior of the circuit. Hence effort invested in developing insight is considered well spent.

12-1. Source Transformations. Network models for electronic circuits contain two basically different kinds of sources, independent sources and controlled sources. An independent source is one whose output voltage or current is entirely independent of the circuit in which the source is connected; the output of such a source is one of the independent variables of the network. A controlled source is one whose output voltage or current depends on a controlling voltage or current, which in turn depends on the network in which the source is connected; the output of such a source is one of the dependent variables of the network. Four kinds of controlled sources are encountered in electronic circuits. They are voltage-controlled voltage sources, voltage-controlled current sources, current-controlled current sources, and current-controlled voltage sources. These sources can be transformed from one kind to another just as an independent current source and shunt resistance can be transformed to an equivalent voltage source and series resistance. A simple example illustrates a few of these transformations.

The triode voltage amplifier of Fig. 12-1*a* can be represented, in so far as increments of current and voltage are concerned, by the incremental

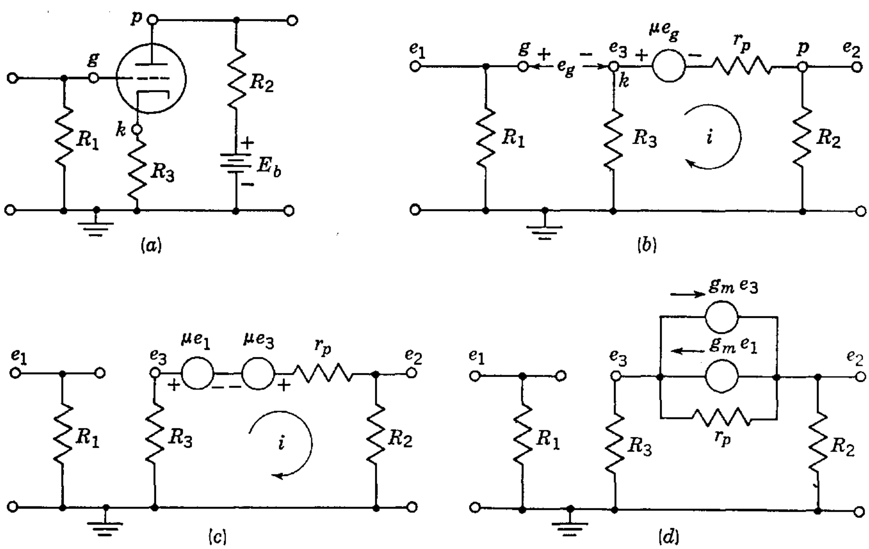


FIG. 12-1. Equivalent representations for a voltage amplifier. (*a*) Circuit; (*b*) model; (*c*) equivalent circuit; (*d*) equivalent circuit.

linear model of Fig. 12-1*b*. The voltage e_g is the incremental potential of the grid relative to the cathode. It follows from the circuit of Fig. 12-1*b* that

$$\mu e_g = \mu(e_1 - e_3) = \mu e_1 - \mu e_3 \tag{12-1}$$

Hence the voltage source μe_g in Fig. 12-1*b* can be replaced by the equivalent series connection of two voltage sources shown in Fig. 12-1*c*. Either one or both of the sources in Fig. 12-1*c* can be associated with r_p and converted to equivalent voltage-controlled current sources; one possibility is shown in Fig. 12-1*d*.

It is clear from Fig. 12-1*c* that

$$e_3 = -R_3 i \qquad \text{and} \qquad \mu e_3 = -\mu R_3 i \tag{12-2}$$

Thus the μe_3 source in Fig. 12-1c can be replaced by an equivalent source having the terminal voltage $\mu R_3 i$. The result, shown in Fig. 12-2a, is that a voltage-controlled voltage source is converted to a current-controlled voltage source. If this source is now associated with the resistance r_p and converted to an equivalent current source, the circuit shown in Fig. 12-2b, which contains a current-controlled current source, results.

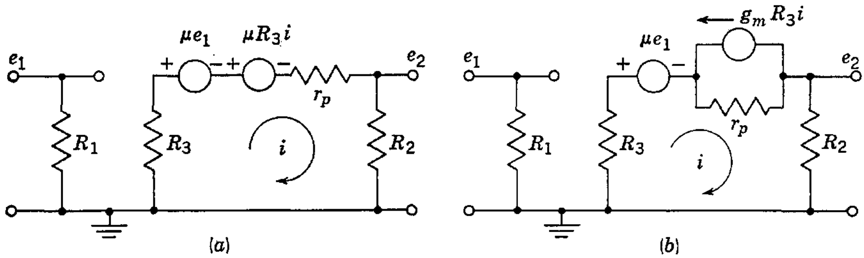


FIG. 12-2. Circuits equivalent to the one in Fig. 12-1b. (a) With a current-controlled voltage source; (b) with a current-controlled current source.

Another circuit that is equivalent to the one in Fig. 12-1b is shown in Fig. 12-3a; this circuit is obtained by using the current-source representation for the tube. A transformation that is often useful in circuit analysis yields the equivalent circuit of Fig. 12-3b. One of the current sources in Fig. 12-3b delivers a current $g_m e_g$ to the ground node, and the other current source takes the same current away from that node.

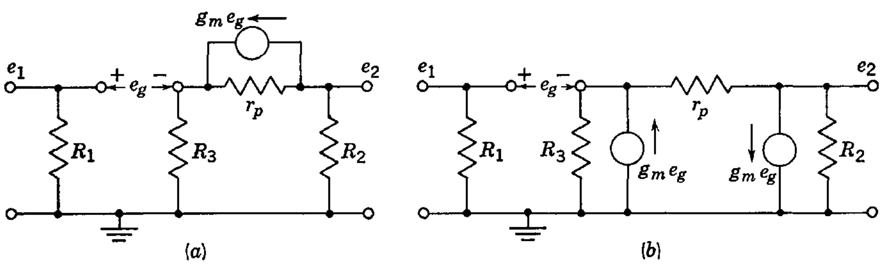


FIG. 12-3. Circuits equivalent to the one in Fig. 12-1b. (a) Current-source form; (b) equivalent current-source form.

Hence the currents flowing into the nodes in Fig. 12-3b are the same as the currents flowing into the corresponding nodes in Fig. 12-3a, the corresponding voltages and currents in the two circuits are the same, and the two circuits are therefore equivalent.

Each of the transformations described above illustrates a mutual equivalence and can be applied in either direction. Each transformation corresponds to an algebraic rearrangement of the equations relating the currents and voltages in the circuit, and equivalent results can be obtained in each case by rearranging the equations rather than the network. However, operating on the network usually provides more insight into the properties of the circuit than manipulating the equations, and it is usually easier to discern helpful transformations by inspection of the circuit than by inspection of the equations.

12-2. Analysis on the Node Basis. The analysis of simple circuits such as those presented in Chaps. 6 and 9 is a simple matter that can often be done by inspection of the circuit. When circuits become more

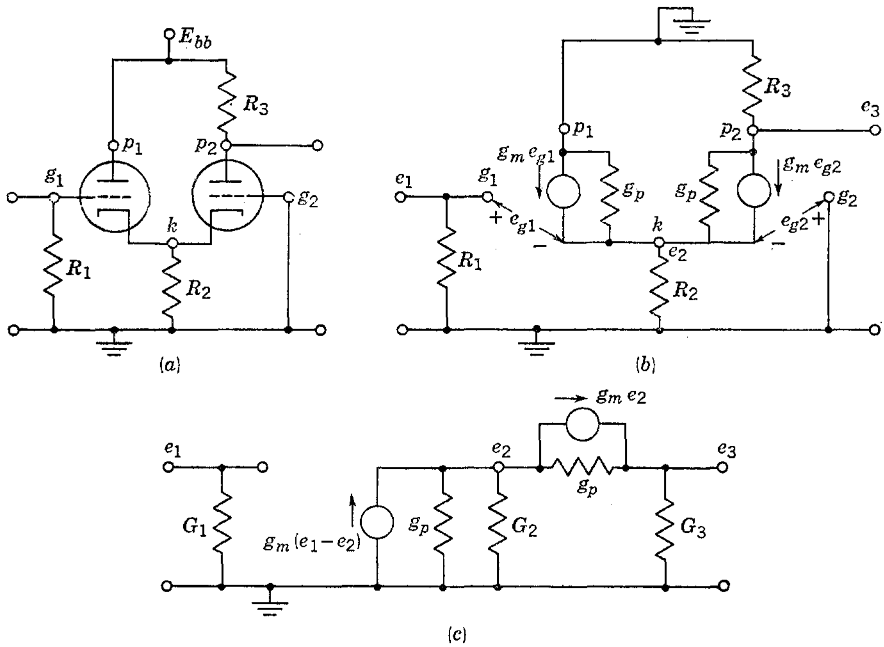


FIG. 12-4. Analysis of the cathode-coupled amplifier. (*a*) Circuit; (*b*) incremental model; (*c*) rearranged model.

complex, however, and especially when they contain feedback, a systematic method of analysis is highly desirable. Consider, for example, the cathode-coupled voltage amplifier shown in Fig. 12-4*a*. An incremental model for this circuit, shown in Fig. 12-4*b*, is obtained by replacing each triode with its current-source model and replacing the d-c supply with a short circuit. This model is redrawn in Fig. 12-4*c* in a more orderly form, and the relations $e_{g1} = e_1 - e_2$ and $e_{g2} = -e_2$ are used to eliminate e_{g1} and e_{g2}.

The node method of analysis[1–3] provides a straightforward solution of the circuit in Fig. 12-4*c* for the unknown voltages and currents as functions of the input signal e_1. The circuit contains four nodes, designated e_1, e_2, e_3, and ground, and it contains two separate parts. The separate parts of a network are parts between which there can be no flow of current; hence the part containing node e_1 in Fig. 12-4c is separate from the rest of the network. Network models for electronic circuits often contain several separate parts, each of which can be analyzed separately from the rest of the network.

The solution of the left-hand part of the network in Fig. 12-4c is trivial. The right-hand part contains three nodes, e_2, e_3, and ground. If the potentials of any two of these relative to the third are known, then the potential across each branch can be found from Kirchhoff's voltage law, and the current in each branch can be computed. Systematic application of the node method of analysis begins with the choice of one node as the reference, or datum, node. Any node can be chosen, and it is marked with the ground symbol. A current-law equation is then written, *in terms of the node-to-datum voltages*, for each node at which the potential is unknown. The primary unknowns in a systematic application of the node method of analysis are therefore the node voltages. The unknown node voltages are found by simultaneous solution of the set of current-law equations.

A systematic procedure for writing the current-law equations is to equate the sum of the currents away from a node in passive branches to the sum of the currents toward the node in source branches. Thus for the nodes e_2 and e_3 in Fig. 12-4c,

$$(g_p + G_2)e_2 + g_p(e_2 - e_3) = g_m(e_1 - e_2) - g_m e_2 \tag{12-3}$$

and

$$G_3 e_3 + g_p(e_3 - e_2) = g_m e_2 \tag{12-4}$$

Collecting terms separately on each side of these equations yields

$$(2g_p + G_2)e_2 - g_p e_3 = g_m e_1 - 2g_m e_2 \tag{12-5}$$

and

$$-g_p e_2 + (g_p + G_3)e_3 = g_m e_2 \tag{12-6}$$

The terms on the left of these equations account for the passive branches in the circuit; the matrix (array) of the coefficients of these terms is symmetrical with respect to the principal diagonal, as is always the case for circuits composed of linear, bilateral, passive elements. The terms on the right account for the action of the sources in the circuit; the terms involving e_2 are unknowns, however. Collecting all terms in (12-5) and (12-6) gives

$$(2g_m + 2g_p + G_2)e_2 - g_p e_3 = g_m e_1 \tag{12-7}$$

and

$$-(g_m + g_p)e_2 + (g_p + G_3)e_3 = 0 \tag{12-8}$$

It is to be noted that transferring the unknown source terms to the left-hand side of these equations destroys the symmetry of the matrix of coefficients. In general, controlled sources introduce unidirectional coupling between various parts of the network, and they thereby destroy the symmetry of the matrix. As a result of this fact, the reciprocity theorem does not apply in general to circuits containing controlled sources.

Equations (12-7) and (12-8) can be solved for the unknown voltages in terms of the input signal. Solving for e_3 with the aid of determinants[1–3] yields

$$e_3 = \frac{g_m(g_m + g_p)e_1}{(g_p + G_3)(2g_m + 2g_p + G_2) - g_p(g_m + g_p)} \tag{12-9}$$

and after simplification,

$$e_3 = \frac{g_m e_1}{g_p + 2G_3 + \dfrac{G_2(g_p + G_3)}{g_m + g_p}} \tag{12-10}$$

Now consider any network in general consisting of a single part, or of one separate part of a larger network, and containing R, L, C, independent current sources, and voltage-controlled current sources. This network has N_t total nodes, and when one of these is designated the reference node, $N = N_t - 1$ unknown node voltages remain. But exactly N independent current-law equations can be written, providing a set of equations that can be solved for the N unknown voltages. If the currents and voltages in the network are all sinusoidal and of the same frequency, then the node equations have the following form when terms are collected separately on each side of the equations:

$$\begin{aligned}
\mathbf{y}_{11}\mathbf{E}_1 + \mathbf{y}_{12}\mathbf{E}_2 + \cdots + \mathbf{y}_{1N}\mathbf{E}_N &= \mathbf{I}_1 + g_{11}\mathbf{E}_1 + \cdots + g_{1N}\mathbf{E}_N \\
\mathbf{y}_{21}\mathbf{E}_1 + \mathbf{y}_{22}\mathbf{E}_2 + \cdots + \mathbf{y}_{2N}\mathbf{E}_N &= \mathbf{I}_2 + g_{21}\mathbf{E}_1 + \cdots + g_{2N}\mathbf{E}_N \\
\cdots\cdots\cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\
\mathbf{y}_{N1}\mathbf{E}_1 + \mathbf{y}_{N2}\mathbf{E}_2 + \cdots + \mathbf{y}_{NN}\mathbf{E}_N &= \mathbf{I}_N + g_{N1}\mathbf{E}_1 + \cdots + g_{NN}\mathbf{E}_N
\end{aligned} \tag{12-11}$$

The y's are combinations of the admittances of various R, L, and C branches, the **I**'s are the complex amplitudes of the total currents from independent sources into the various nodes, and the g's are combinations of the transconductances of controlled current sources. In any particular case, many of the y's, **I**'s, and g's may be zero. Equations (12-11) should be compared with Eqs. (12-5) and (12-6). The matrix of the y coefficients in (12-11) is symmetrical around the principal diagonal; the matrix of the g coefficients is, in general, not symmetrical.

To solve Eqs. (12-11) for the unknown voltages, all terms are collected, yielding

$$\begin{aligned}
\mathbf{y}'_{11}\mathbf{E}_1 + \mathbf{y}'_{12}\mathbf{E}_2 + \cdots + \mathbf{y}'_{1N}\mathbf{E}_N &= \mathbf{I}_1 \\
\mathbf{y}'_{21}\mathbf{E}_1 + \mathbf{y}'_{22}\mathbf{E}_2 + \cdots + \mathbf{y}'_{2N}\mathbf{E}_N &= \mathbf{I}_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots & \cdots \\
\mathbf{y}'_{N1}\mathbf{E}_1 + \mathbf{y}'_{N2}\mathbf{E}_2 + \cdots + \mathbf{y}'_{NN}\mathbf{E}_N &= \mathbf{I}_N
\end{aligned} \tag{12-12}$$

The solution of these equations for any unknown, say $\mathbf{E}_j$, is given by Cramer's rule[1,3] as

$$\mathbf{E}_j = \frac{\begin{vmatrix} \mathbf{y}'_{11} & \mathbf{y}'_{12} & \cdots & \mathbf{I}_1 & \cdots & \mathbf{y}'_{1N} \\ \mathbf{y}'_{21} & \mathbf{y}'_{22} & \cdots & \mathbf{I}_2 & \cdots & \mathbf{y}'_{2N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{y}'_{N1} & \mathbf{y}'_{N2} & \cdots & \mathbf{I}_N & \cdots & \mathbf{y}'_{NN} \end{vmatrix}}{\mathbf{\Delta}} \tag{12-13}$$

where $\mathbf{\Delta}$ is the determinant of coefficients in Eqs. (12-12). Expanding the numerator determinant on the column of I's and designating the cofactor on the element in the kth row and jth column by $\mathbf{\Delta}_{kj}$ gives

$$\mathbf{E}_j = \frac{\mathbf{\Delta}_{1j}}{\mathbf{\Delta}}\mathbf{I}_1 + \frac{\mathbf{\Delta}_{2j}}{\mathbf{\Delta}}\mathbf{I}_2 + \cdots + \frac{\mathbf{\Delta}_{Nj}}{\mathbf{\Delta}}\mathbf{I}_N \tag{12-14}$$

The coefficients in this equation, $\mathbf{\Delta}_{kj}/\mathbf{\Delta}$, have the dimensions of impedance; hence defining new symbols, (12-14) can be written as

$$\mathbf{E}_j = \mathbf{Z}_{j1}\mathbf{I}_1 + \mathbf{Z}_{j2}\mathbf{I}_2 + \cdots + \mathbf{Z}_{jN}\mathbf{I}_N \tag{12-15}$$

where

$$\mathbf{Z}_{jk} = \frac{\mathbf{\Delta}_{kj}}{\mathbf{\Delta}} \tag{12-16}$$

Solving (12-12) in this manner for each of the unknown voltages yields a set of equations expressing each unknown voltage in terms of the known applied currents. These equations are

$$\begin{aligned} \mathbf{Z}_{11}\mathbf{I}_1 + \mathbf{Z}_{12}\mathbf{I}_2 + \cdots + \mathbf{Z}_{1N}\mathbf{I}_N &= \mathbf{E}_1 \\ \mathbf{Z}_{21}\mathbf{I}_1 + \mathbf{Z}_{22}\mathbf{I}_2 + \cdots + \mathbf{Z}_{2N}\mathbf{I}_N &= \mathbf{E}_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathbf{Z}_{N1}\mathbf{I}_1 + \mathbf{Z}_{N2}\mathbf{I}_2 + \cdots + \mathbf{Z}_{NN}\mathbf{I}_N &= \mathbf{E}_N \end{aligned} \tag{12-17}$$

This set of equations represents the solution of the network for the unknown node voltages. The **Z** coefficients are the open-circuit driving-point and transfer impedances[1,2] of the network; they can be evaluated either from the y′ determinant by means of Eq. (12-16) or from the network itself. As an illustration of this latter procedure, it is clear from (12-17) that $\mathbf{Z}_{12}$ is the ratio of $\mathbf{E}_1$ to $\mathbf{I}_2$ when all the I's except $\mathbf{I}_2$ are zero.

Suppose that the general network considered above contains one or more current-controlled sources. Equations (12-11) then contain the N unknown voltages, and, in addition, the current-controlled sources introduce terms on the right-hand side containing unknown currents. Thus there are N equations relating more than N unknowns. This difficulty can be met by eliminating the unknown currents with the help of auxiliary branch equations such as (12-1) and (12-2). Or, alternatively, the current-controlled sources can be transformed to equivalent voltage-controlled sources, as outlined in Sec. 12-1, before the equations are written.

12-3. Analysis on the Loop Basis. Analysis on the loop basis[1-3] is an alternative procedure that is the dual of analysis on the node basis. The two methods serve the same end and, in general, are of equal importance in network analysis. The configurations most commonly used in electronic circuits are such that the node basis usually provides the simplest solution; however, there are cases in which the loop basis is preferable.

A transistor amplifier with an unbypassed resistor in series with the emitter lead is shown in Fig. 12-5*a*, and an incremental model for the

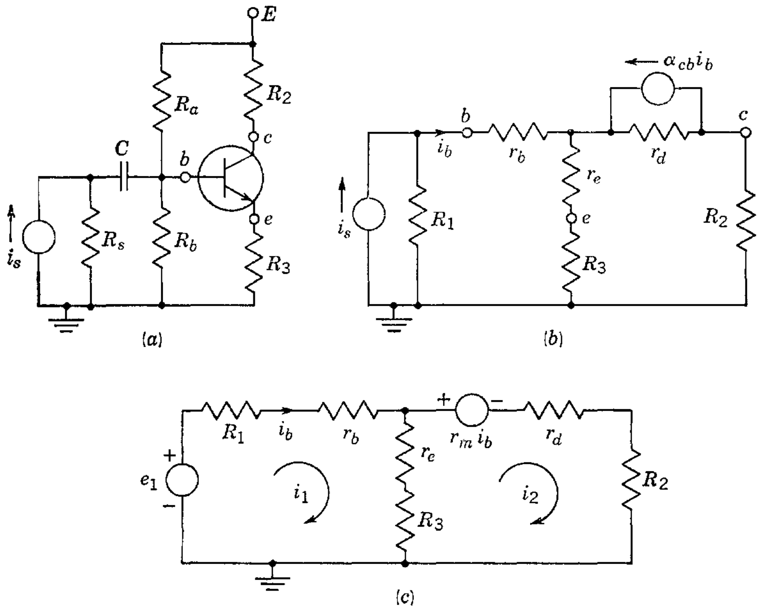


FIG. 12-5. Analysis of a transistor amplifier. (*a*) Circuit; (*b*) incremental model; (*c*) alternative model.

circuit that is valid in the band of frequencies where the coupling capacitor C acts as a short circuit, and where the transistor is independent of frequency, is shown in Fig. 12-5*b*. The resistance R_1 in this circuit represents R_s, R_a, and R_b in parallel. This circuit is put in a more natural form for loop analysis by converting the two current sources to equivalent voltage sources as shown in Fig. 12-5*c*. If the two loop currents in this circuit are known, the current in each branch can be found from Kirchhoff's current law, and the voltage drop across each branch can be computed. The loop method of analysis provides a solution for these loop currents.

Systematic application of the loop method of analysis begins with the choice of a set of independent loop currents. A voltage-law equation is then written, *in terms of the loop currents*, for each loop in which the current is unknown. The primary unknowns in a systematic application of the loop method of analysis are therefore the loop currents. The unknown loop currents are found by simultaneous solution of the set of voltage-law equations.

A systematic procedure for writing the voltage-law equations is to equate the sum of the voltage drops across passive branches in the arrow direction around each loop to the sum of the voltage rises across source branches in the arrow direction around the loop. Thus, when terms are collected separately on each side of the equations for the two loops in Fig. 12-5*c*,

$$(r_b + r_e + R_1 + R_3)i_1 - (r_e + R_3)i_2 = e_1 \qquad (12\text{-}18)$$

and

$$-(r_e + R_3)i_1 + (r_e + r_d + R_2 + R_3)i_2 = -r_m i_1 \qquad (12\text{-}19)$$

The terms on the left of these equations account for the passive branches in the circuit; the matrix (array) of the coefficients of these terms is symmetrical with respect to the principal diagonal. The terms on the right account for the action of the sources in the circuit; the term involving i_1 is unknown, however. Collecting all terms in these equations yields

$$(r_b + r_e + R_1 + R_3)i_1 - (r_e + R_3)i_2 = e_1 \qquad (12\text{-}20)$$

and

$$-(r_e + R_3 - r_m)i_1 + (r_e + r_d + R_2 + R_3)i_2 = 0 \qquad (12\text{-}21)$$

Transferring the source term in (12-19) to the left-hand side destroys the symmetry of the matrix of coefficients. Equations (12-20) and (12-21) can be solved simultaneously for the unknown currents i_1 and i_2.

In the node method of analysis there is no problem in identifying the unknown node voltages and in writing the proper number of independent equations. In the loop method, however, matters are not so clear. First, the number of unknown loop currents required to characterize the network is not always obvious. A more detailed study shows that if the network contains only one part having a total of B branches, N_t nodes, and S_c current sources, then the number of loop currents required is

$$L = B - N_t - S_c + 1 \qquad (12\text{-}22)$$

For comparison, the number of unknown node voltages required to characterize a single-part network is

$$N = N_t - 1 - S_v \qquad (12\text{-}23)$$

where S_v is the number of voltage sources. Second, although the loops can be chosen in a variety of ways, they cannot be chosen completely at random, for some choices lead to a set of L equations that are not independent and that therefore cannot be solved for the unknown currents. The independence of these equations can be ensured by choosing the loops so that every branch is included in at least one loop and so that each loop contains at least one branch that is not contained in any other loop.

Now consider any network consisting of a single part, or of one separate part of a larger network, and containing R, L, C, independent voltage sources, and current-controlled voltage sources. This network has L independent loops, and exactly L independent voltage-law equations can be written providing a set of equations that can be solved for the L unknown currents. If the currents and voltages in the network are sinusoidal and of the same frequency, then the loop equations have the following form when terms are collected separately on each side of the equations:

$$\begin{aligned} \mathbf{z}_{11}\mathbf{I}_1 + \mathbf{z}_{12}\mathbf{I}_2 + \cdots + \mathbf{z}_{1L}\mathbf{I}_L &= \mathbf{E}_1 + r_{11}\mathbf{I}_1 + \cdots + r_{1L}\mathbf{I}_L \\ \mathbf{z}_{21}\mathbf{I}_1 + \mathbf{z}_{22}\mathbf{I}_2 + \cdots + \mathbf{z}_{2L}\mathbf{I}_L &= \mathbf{E}_2 + r_{21}\mathbf{I}_1 + \cdots + r_{2L}\mathbf{I}_L \\ \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots \\ \mathbf{z}_{L1}\mathbf{I}_1 + \mathbf{z}_{L2}\mathbf{I}_2 + \cdots + \mathbf{z}_{LL}\mathbf{I}_L &= \mathbf{E}_L + r_{L1}\mathbf{I}_1 + \cdots + r_{LL}\mathbf{I}_L \end{aligned} \qquad (12\text{-}24)$$

The **z**'s are combinations of the impedances of the various R, L, and C branches, the **E**'s are the complex amplitudes of the total voltages from independent sources summed around the various loops, and the r's are combinations of the transfer resistances of the controlled sources. In any particular case, many of the **z**'s, **E**'s, and r's may be zero. Equations (12-24) should be compared with Eqs. (12-18) and (12-19). The matrix of the **z** coefficients in (12-24) is symmetrical around the principal diagonal; the matrix of the r coefficients is in general not symmetrical.

Collecting all terms in (12-24) yields

$$\begin{aligned} \mathbf{z}'_{11}\mathbf{I}_1 + \mathbf{z}'_{12}\mathbf{I}_2 + \cdots + \mathbf{z}'_{1L}\mathbf{I}_L &= \mathbf{E}_1 \\ \mathbf{z}'_{21}\mathbf{I}_1 + \mathbf{z}'_{22}\mathbf{I}_2 + \cdots + \mathbf{z}'_{2L}\mathbf{I}_L &= \mathbf{E}_2 \\ \cdots\cdots\cdots\cdots & \cdots \\ \mathbf{z}'_{L1}\mathbf{I}_1 + \mathbf{z}'_{L2}\mathbf{I}_2 + \cdots + \mathbf{z}'_{LL}\mathbf{I}_L &= \mathbf{E}_L \end{aligned} \qquad (12\text{-}25)$$

Solving these equations by determinants as the node equations in Sec. 12-2 were solved, and defining

$$\mathbf{Y}_{jk} = \frac{\mathbf{\Delta}_{kj}}{\mathbf{\Delta}} \qquad (12\text{-}26)$$

where $\mathbf{\Delta}$ is the determinant of the coefficients in (12-25) and $\mathbf{\Delta}_{kj}$ is the cofactor of $\mathbf{z}'_{kj}$, yields a set of equations for the unknown loop currents in terms of the applied voltages:

$$\begin{aligned} \mathbf{Y}_{11}\mathbf{E}_1 + \mathbf{Y}_{12}\mathbf{E}_2 + \cdots + \mathbf{Y}_{1L}\mathbf{E}_L &= \mathbf{I}_1 \\ \mathbf{Y}_{21}\mathbf{E}_1 + \mathbf{Y}_{22}\mathbf{E}_2 + \cdots + \mathbf{Y}_{2L}\mathbf{E}_L &= \mathbf{I}_2 \\ \cdots\cdots\cdots\cdots & \cdots \\ \mathbf{Y}_{L1}\mathbf{E}_1 + \mathbf{Y}_{L2}\mathbf{E}_2 + \cdots + \mathbf{Y}_{LL}\mathbf{E}_L &= \mathbf{I}_L \end{aligned} \qquad (12\text{-}27)$$

The **Y** coefficients of (12-27) are the short-circuit driving-point and transfer admittances of the network; they can be evaluated either from the **z**′ determinant by means of Eq. (12-26) or from the network itself. As an illustration of this latter procedure, it is clear from (12-27) that $\mathbf{Y}_{12}$ is the ratio of $\mathbf{I}_1$ to $\mathbf{E}_2$ when all the **E**'s except $\mathbf{E}_2$ are zero.

If the general network contains one or more voltage-controlled sources, Eqs. (12-24) contain the L unknown loop currents, and, in addition, the voltage-controlled sources introduce terms on the right-hand side containing unknown voltages. Thus there are L equations relating more than L unknowns. This difficulty can be met by eliminating the unknown voltages with the help of auxiliary branch equations such as (12-1) and (12-2). Or, alternatively, the voltage-controlled sources can be transformed to equivalent current-controlled sources, as outlined in Sec. 12-1, before the equations are written.

12-4. Two-terminal-pair Networks. In the study of circuits arising in electronic and communication systems, attention is often focused primarily on two pairs of terminals that are designated as input and output terminal pairs. Such a network, known as a two-terminal-pair,[1,3] or two-port, network, is represented in Fig. 12-6. In many cases the details of the currents and voltages inside such networks are of secondary interest; primary interest is centered on the external behavior, as, for example, on the output voltage under specified conditions of load and input voltage. More generally speaking, attention is centered on the relations among $\mathbf{E}_1$, $\mathbf{I}_1$, $\mathbf{E}_2$, and $\mathbf{I}_2$.

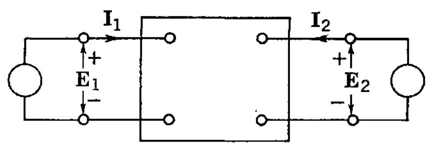


FIG. 12-6. A two-terminal-pair network.

Suppose that $\mathbf{E}_1$ and $\mathbf{E}_2$ are known and that $\mathbf{I}_1$ and $\mathbf{I}_2$ are to be found. If the network contains no internal independent sources, so that $\mathbf{I}_1$ and $\mathbf{I}_2$ depend only on the network parameters and the values of $\mathbf{E}_1$ and $\mathbf{E}_2$, then $\mathbf{I}_1$ and $\mathbf{I}_2$ are given by a set of equations like (12-27). Specifically, they are given by

$$\begin{aligned} \mathbf{I}_1 &= \mathbf{Y}_{11}\mathbf{E}_1 + \mathbf{Y}_{12}\mathbf{E}_2 \\ \mathbf{I}_2 &= \mathbf{Y}_{21}\mathbf{E}_1 + \mathbf{Y}_{22}\mathbf{E}_2 \end{aligned} \qquad (12\text{-}28)$$

These equations completely specify the external behavior of the network; when $\mathbf{E}_1$ and $\mathbf{E}_2$ are known, $\mathbf{I}_1$ and $\mathbf{I}_2$ can be found.

Equations (12-28) can be derived in another manner that is in some respects more convenient. Since the network under consideration is understood to be linear, the principle of superposition can be applied. It follows from this fact that the unknown current $\mathbf{I}_1$ can be evaluated by summing the component resulting from $\mathbf{E}_1$ when $\mathbf{E}_2$ is zero and the component resulting from $\mathbf{E}_2$ when $\mathbf{E}_1$ is zero. The first equation in set

(12-28) is just a statement of this fact. The second equation in set (12-28) can be derived in the same manner.

Now suppose that $\mathbf{I}_1$ and $\mathbf{I}_2$ are known and that $\mathbf{E}_1$ and $\mathbf{E}_2$ are to be found. The unknown voltages can be computed from a knowledge of the circuit parameters and the values of $\mathbf{I}_1$ and $\mathbf{I}_2$. The voltage $\mathbf{E}_1$ is the superposition of two components, one resulting from $\mathbf{I}_1$ when $\mathbf{I}_2$ is zero and the other resulting from $\mathbf{I}_2$ when $\mathbf{I}_1$ is zero; that is,

$$\mathbf{E}_1 = \mathbf{Z}_{11}\mathbf{I}_1 + \mathbf{Z}_{12}\mathbf{I}_2$$

Similarly,

$$\mathbf{E}_2 = \mathbf{Z}_{21}\mathbf{I}_1 + \mathbf{Z}_{22}\mathbf{I}_2 \qquad (12\text{-}29)$$

These equations can also be obtained from (12-17).

Equations (12-28) characterize the external behavior of the network in terms of the short-circuit driving-point and transfer admittances; Eqs. (12-29) characterize the behavior in terms of the open-circuit impedances. For circuits made up of ideal R, L, C, and independent sources,

$$\mathbf{Z}_{12} = \mathbf{Z}_{21} \qquad \text{and} \qquad \mathbf{Y}_{12} = \mathbf{Y}_{21} \qquad (12\text{-}30)$$

The principle of reciprocity[1,3] applies to such circuits. In the usual case with circuits containing controlled sources, however,

$$\mathbf{Z}_{12} \neq \mathbf{Z}_{21} \qquad \text{and} \qquad \mathbf{Y}_{12} \neq \mathbf{Y}_{21} \qquad (12\text{-}31)$$

The principle of reciprocity does not apply in such cases.

12-5. The Hybrid Voltage-amplifier and Current-amplifier Coefficients. In the study of vacuum-tube circuits a third manner of characterizing the external behavior of two-terminal-pair networks is of considerable utility, especially in providing further insight into the properties of voltage amplifiers. Suppose that $\mathbf{E}_1$ and $\mathbf{I}_2$ in Fig. 12-6 are known and that $\mathbf{I}_1$ and $\mathbf{E}_2$ are to be found. The current $\mathbf{I}_1$ can be expressed as the superposition of two components, one resulting from $\mathbf{E}_1$ when $\mathbf{I}_2$ is zero and the other resulting from $\mathbf{I}_2$ when $\mathbf{E}_1$ is zero; that is,

$$\mathbf{I}_1 = \mathbf{Y}_{no}\mathbf{E}_1 + \mathbf{B}_{cs}\mathbf{I}_2 \qquad (12\text{-}32)$$

In a similar manner $\mathbf{E}_2$ can be expressed as

$$\mathbf{E}_2 = \mathbf{A}_{vo}\mathbf{E}_1 + \mathbf{Z}_{os}\mathbf{I}_2 \qquad (12\text{-}33)$$

The coefficients in these equations are the hybrid voltage-amplifier coefficients. The first of these,

$$\mathbf{Y}_{no} = \left.\frac{\mathbf{I}_1}{\mathbf{E}_1}\right|_{\mathbf{I}_2=0} \qquad (12\text{-}34)$$

is the input admittance with the output open-circuited, or, more briefly, it is the open-circuit input admittance. Similarly,

$$\mathbf{B}_{cs} = \left.\frac{\mathbf{I}_1}{\mathbf{I}_2}\right|_{\mathbf{E}_1=0} \qquad (12\text{-}35)$$

is the reverse current transmittance with the input short-circuited, or simply the short-circuit reverse current transmittance,

$$\mathbf{A}_{vo} = \frac{\mathbf{E}_2}{\mathbf{E}_1}\bigg|_{\mathbf{I}_2=0} \tag{12-36}$$

is the open-circuit forward voltage transmittance, and

$$\mathbf{Z}_{os} = \frac{\mathbf{E}_2}{\mathbf{I}_2}\bigg|_{\mathbf{E}_1=0} \tag{12-37}$$

is the short-circuit output impedance. These coefficients provide a convenient means for describing the properties of voltage amplifiers and for comparing them with the ideal voltage amplifier. In the ideal case, shown in Fig. 4-1, $\mathbf{Y}_{no}$, $\mathbf{B}_{cs}$, and $\mathbf{Z}_{os}$ are all zero.

The hybrid coefficients can be evaluated directly from the network with the aid of Eqs. (12-34) to (12-37). As an illustration, consider the

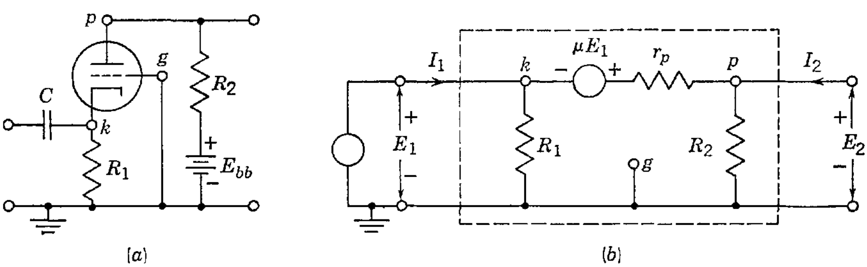


FIG. 12-7. The grounded-grid amplifier. (*a*) Circuit; (*b*) incremental model.

grounded-grid, or cathode-driven, amplifier shown in Fig. 12-7*a*. This circuit, with certain refinements, is widely used as the first amplifier stage in FM and television receivers because it generates relatively little noise. An incremental model that is valid over the band of frequencies in which the coupling capacitor acts as a short circuit and in which the parasitic capacitances act as open circuits is shown in Fig. 12-7*b*. The fact that $E_g = -E_1$ is used in this model.

The hybrid coefficients for the model in Fig. 12-7*b* are to be evaluated. According to (12-34), Y_{no} is the ratio I_1/E_1 with $I_2 = 0$. Under the condition that $I_2 = 0$,

$$I_1 = \frac{E_1}{R_1} + \frac{E_1 + \mu E_1}{r_p + R_2} \tag{12-38}$$

from which

$$Y_{no} = \frac{1}{R_1} + \frac{1 + \mu}{r_p + R_2} \tag{12-39}$$

Also, with $I_2 = 0$,

$$E_2 = \frac{R_2}{r_p + R_2}(E_1 + \mu E_1) \tag{12-40}$$

which, in accordance with (12-36), yields

$$A_{vo} = \frac{(1 + \mu)R_2}{r_p + R_2} \tag{12-41}$$

With $E_1 = 0$ and a current applied at I_2,

$$I_1 = -\frac{R_2}{r_p + R_2} I_2 \tag{12-42}$$

and

$$E_2 = \frac{r_p R_2}{r_p + R_2} I_2 \tag{12-43}$$

These two equations, along with (12-35) and (12-37), yield

$$B_{cs} = -\frac{R_2}{r_p + R_2} \quad \text{and} \quad Z_{os} = \frac{r_p R_2}{r_p + R_2} \tag{12-44}$$

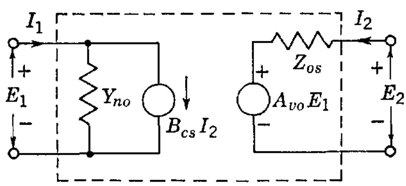


FIG. 12-8. Equivalent circuit for the network of Fig. 12-6.

Equations (12-32) and (12-33) imply that the external behavior of any two-terminal-pair network can be represented by the equivalent network shown in Fig. 12-8. This equivalent network provides a com-

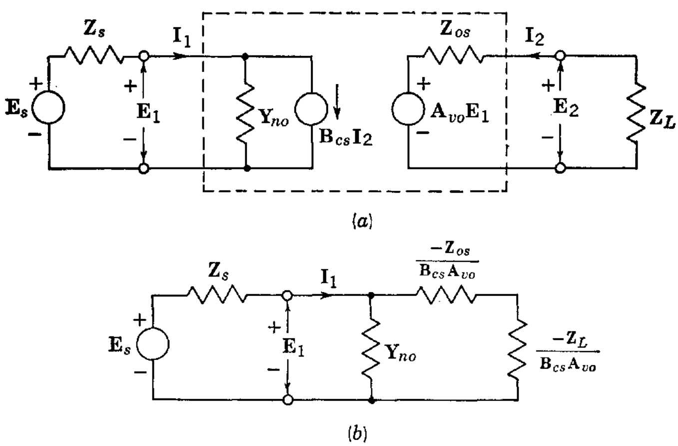


FIG. 12-9. Voltage-amplifier equivalent circuits. (*a*) Network; (*b*) equivalent network for input relations.

pact means of representing the behavior of any network at the external terminal pairs.

When the output terminals in the circuit of Fig. 12-8 are open-circuited, $\mathbf{I}_2 = 0$, and the input admittance to the circuit is $\mathbf{Y}_{no}$. When the circuit is loaded with an impedance $\mathbf{Z}_L$ as illustrated in Fig. 12-9*a*, however, the input admittance is affected through the source $\mathbf{B}_{cs}\mathbf{I}_2$ by the load impedance and in general is not $\mathbf{Y}_{no}$. Similarly, the output impedance is $\mathbf{Z}_{os}$ only if the input signal is supplied by a voltage source with no series impedance. If the load and source impedances are known in advance, they can be accounted for by including them in the network as part of the circuit. When it is desired to study the effects of the terminating impedances on the behavior of the network, however, the alternative procedure described in the following paragraphs provides more insight.

The load current in the circuit of Fig. 12-9*a* is

$$\mathbf{I}_2 = -\frac{\mathbf{A}_{vo}\mathbf{E}_1}{\mathbf{Z}_{os} + \mathbf{Z}_L} \tag{12-45}$$

Therefore

$$\mathbf{B}_{cs}\mathbf{I}_2 = -\frac{\mathbf{B}_{cs}\mathbf{A}_{vo}}{\mathbf{Z}_{os} + \mathbf{Z}_L}\mathbf{E}_1 \tag{12-46}$$

The current source $\mathbf{B}_{cs}\mathbf{I}_2$ in Fig. 12-9*a* is connected across the voltage $\mathbf{E}_1$; therefore it follows from (12-46) that if this source is replaced by a series connection of two impedances, $-\mathbf{Z}_{os}/\mathbf{B}_{cs}\mathbf{A}_{vo}$ and $-\mathbf{Z}_L/\mathbf{B}_{cs}\mathbf{A}_{vo}$, then the current $\mathbf{I}_1$ is not changed. Hence the circuit in Fig. 12-9*b* is equivalent to that in Fig. 12-9*a* in so far as the input terminals are concerned. The input admittance to the circuit when it is loaded with an impedance $\mathbf{Z}_L$ is, therefore,

$$\mathbf{Y}_n = \mathbf{Y}_{no} - \frac{\mathbf{B}_{cs}\mathbf{A}_{vo}}{\mathbf{Z}_{os} + \mathbf{Z}_L} \tag{12-47}$$

The advantage of this representation is that it permits the effect of $\mathbf{Z}_L$ on $\mathbf{Y}_n$ to be evaluated easily.

The equivalent circuit of Fig. 12-9*b* shows that impedances can be transferred from the output to the input side of the network in much the same way as they are transferred across an ideal transformer; the impedance-transformation ratio in this case is $-1/\mathbf{B}_{cs}\mathbf{A}_{vo}$. In a similar manner, the circuit on the input side of the network can be transferred to the output side with the result shown in Fig. 12-10. In this case, impedances are multiplied by $-\mathbf{B}_{cs}\mathbf{A}_{vo}$ and voltages are multiplied by $\mathbf{A}_{vo}$ in being transferred; it follows that admittances are multiplied by $-1/\mathbf{B}_{cs}\mathbf{A}_{vo}$ and currents are multiplied by $-1/\mathbf{B}_{cs}$. The circuit of Fig. 12-10 is equivalent to that of Fig. 12-9*a* in so far as the output terminals are concerned. The output impedance is, therefore,

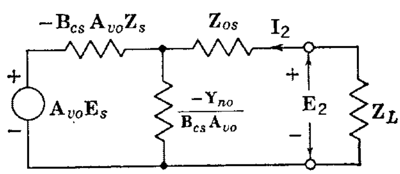


FIG. 12-10. The effect of source impedance.

$$\mathbf{Z}_o = \frac{\mathbf{E}_2}{\mathbf{I}_2}\bigg|_{\mathbf{E}_s=0} = \mathbf{Z}_{os} - \frac{\mathbf{B}_{cs}\mathbf{A}_{vo}}{\mathbf{Y}_{no} + \mathbf{Y}_s} \tag{12-48}$$

where $\mathbf{Y}_s = 1/\mathbf{Z}_s$.

Example 12-1. The circuit of a grounded-grid amplifier is shown in Fig. 12-11. The incremental parameters for the triode at the quiescent operating point are $\mu = 60$ and $r_p = 15$ kilohms. The coupling capacitors can be treated as short circuits and the parasitic tube capacitances can be treated as open circuits in the band of frequencies of interest.

a. Determine the hybrid voltage-amplifier coefficients for the circuit.

b. Determine the forward voltage transmittance, $A_v = E_L/E_s$, when $R_s = 50$ ohms and $R_L = 100$ kilohms.

Solution. *a.* Figure 12-7*b* shows an incremental model for the circuit, and Eqs. (12-39) to (12-44) give the required coefficients in terms of the circuit parameters. Thus

$$Y_{no} = \frac{1}{0.3} + \frac{61}{15 + 33} = 3.33 + 1.27 = 4.6 \text{ millimhos}$$

$$A_{vo} = \frac{(61)(33)}{15 + 33} = 42$$

$$B_{cs} = -\frac{33}{15 + 33} = -0.69$$

$$Z_{os} = \frac{(15)(33)}{15 + 33} = 10.3 \text{ kilohms}$$

b. The impedance reflected into the input circuit by the reverse current transmittance (Fig. 12-9) is

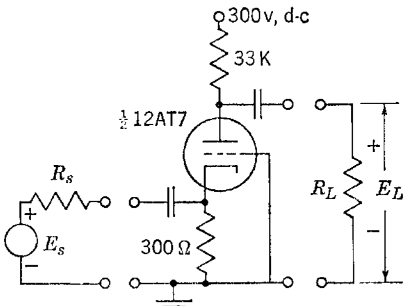


FIG. 12-11. Grounded-grid amplifier for Example 12-1.

$$-\frac{Z_{os} + Z_L}{B_{cs}A_{vo}} = \frac{10.3 + 100}{(0.69)(42)} = 3.81 \text{ kilohms}$$

But the impedance of Z_s in parallel with

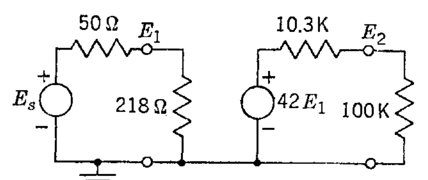


FIG. 12-12. Equivalent circuit for the grounded-grid amplifier.

Y_{no} is less than $R_s = 50$ ohms. Hence when R_s is 50 ohms or less and R_L is 100 kilohms or more, the reflected impedance has a negligible effect on the behavior of the circuit and can be treated as an open circuit. Under these conditions the amplifier can be represented by the circuit shown in Fig. 12-12, and the voltage transmittance is, by inspection of this circuit,

$$A_v = \frac{E_L}{E_s} = \frac{218}{50 + 218}(42)\frac{100}{100 + 10.3}$$
$$= 31$$

The terminal characteristics of any two-port network can be described by Eqs. (12-29):

$$\mathbf{E}_1 = \mathbf{Z}_{11}\mathbf{I}_1 + \mathbf{Z}_{12}\mathbf{I}_2 \quad (12\text{-}49)$$

and

$$\mathbf{E}_2 = \mathbf{Z}_{21}\mathbf{I}_1 + \mathbf{Z}_{22}\mathbf{I}_2 \quad (12\text{-}50)$$

If the network obeys the principle of reciprocity, then

$$\mathbf{Z}_{12} = \mathbf{Z}_{21} \quad (12\text{-}51)$$

and only three of the open-circuit impedances need be determined. It is therefore of interest to determine what relation exists among the hybrid voltage-amplifier coefficients for reciprocal networks.

For the condition that $\mathbf{I}_2 = 0$, dividing Eqs. (12-49) and (12-50) yields

$$\left.\frac{\mathbf{E}_2}{\mathbf{E}_1}\right|_{\mathbf{I}_2=0} = \mathbf{A}_{vo} = \frac{\mathbf{Z}_{21}}{\mathbf{Z}_{11}} \tag{12-52}$$

For the condition that $\mathbf{E}_1 = 0$, Eq. (12-49) yields

$$\left.\frac{\mathbf{I}_1}{\mathbf{I}_2}\right|_{\mathbf{E}_1=0} = \mathbf{B}_{cs} = -\frac{\mathbf{Z}_{12}}{\mathbf{Z}_{11}} \tag{12-53}$$

But $\mathbf{Z}_{12} = \mathbf{Z}_{21}$ for reciprocal networks; hence the condition among the hybrid voltage-amplifier coefficients for reciprocal networks is

$$\mathbf{A}_{vo} = -\mathbf{B}_{cs} \tag{12-54}$$

That is, the forward voltage transmittance is the negative of the reverse current transmittance for reciprocal networks.

An alternative set of hybrid coefficients that is useful in describing current amplifiers is obtained by taking $\mathbf{I}_1$ and $\mathbf{E}_2$ as the known variables in the general network of Fig. 12-6. The two unknown variables, $\mathbf{E}_1$ and $\mathbf{I}_2$, are then given by

$$\mathbf{E}_1 = \mathbf{Z}_{ns}\mathbf{I}_1 + \mathbf{B}_{vo}\mathbf{E}_2 \tag{12-55}$$

and

$$\mathbf{I}_2 = \mathbf{A}_{cs}\mathbf{I}_1 + \mathbf{Y}_{oo}\mathbf{E}_2 \tag{12-56}$$

The coefficients in these equations are the hybrid current-amplifier coefficients. The first of these, $\mathbf{Z}_{ns}$, is the short-circuit input impedance; similarly, $\mathbf{B}_{vo}$ is the open-circuit reverse voltage transmittance, $\mathbf{A}_{cs}$ is the short-circuit forward current transmittance, and $\mathbf{Y}_{oo}$ is the open-circuit output admittance. It is clear that the hybrid parameters used in describing the terminal volt-ampere relations for the transistor are the current-amplifier coefficients for the transistor. Equations (12-55) and (12-56) are the duals of (12-32) and (12-33); the coefficients in (12-55) and (12-56) can be evaluated in terms of the network parameters from the appropriate short-circuit and open-circuit relations in the network. In the ideal current amplifier, $\mathbf{Z}_{ns}$, $\mathbf{B}_{vo}$, and $\mathbf{Y}_{oo}$ are all zero.

Equations (12-55) and (12-56) imply that the general network of Fig. 12-6 can be represented by the equivalent circuit inside the box shown in Fig. 12-13*a*. This circuit is the dual of the voltage-amplifier equivalent circuit of Fig. 12-8, and it is the same as the circuit in Fig. 12-8 with the input and output terminals interchanged. Thus all the results derived from the circuit of Fig. 12-8 apply to the circuit of Fig. 12-13*a*; it is merely necessary to interchange the subscripts 1 and 2 on the voltages and currents, to interchange the subscripts *n* (for input), and *o* (for output) on the coefficients, and to interchange *A* and *B*. For example, reflecting the output-circuit admittance into the input circuit of the current amplifier, shown in Fig. 12-13*b*, corresponds to reflecting the source admittance (with $E_s = 0$) into the output circuit of the voltage amplifier, shown in Fig. 12-10.

The value of the hybrid coefficients is largely conceptual; they provide useful insights into the properties of electronic circuits and into the interaction between the circuits and the external sources and loads connected to them. They provide, in addition, a set of clearly defined coefficients that are sufficient to describe completely the external behavior of two-port networks and that permit such networks to be compared with ideal voltage and current amplifiers. They are particularly useful when it is necessary to determine the performance of a circuit under a variety of terminal conditions, just as Thevenin's theorem is useful in calculating

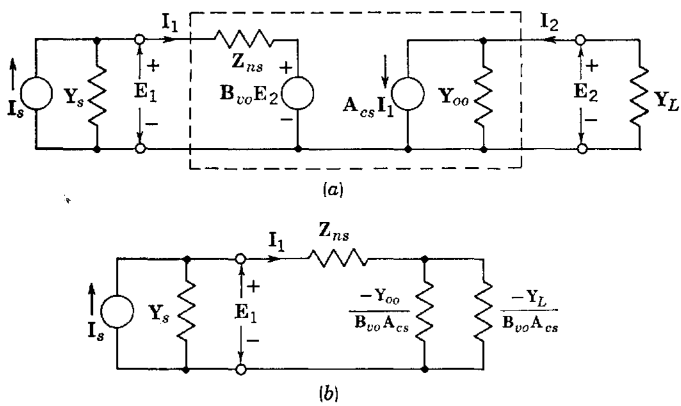


FIG. 12-13. Current-amplifier equivalent circuits. (*a*) Network; (*b*) equivalent network for input relations.

the terminal conditions in a one-port network under various load conditions. But, as is the case with Thevenin's theorem, the hybrid coefficients usually do not simplify the analysis of circuits for a single set of terminal conditions.

12-6. Summary. The currents and voltages in elementary electronic circuits can often be evaluated by inspection of the circuit. When the circuits are more elaborate, however, a systematic method of analysis is essential for orderly, efficient solutions of circuit problems. Both loop and node methods fill this need. The controlled sources that occur in the models for electronic circuits affect the form of the loop and node equations in certain significant ways, and in certain cases they require that auxiliary equations be written to augment the normal set of loop or node equations.

In the study of electronic circuits, attention is often focused on the voltage and current relations at two terminal pairs that are usually designated as input and output terminal pairs. In such cases it is often convenient to characterize the external behavior of electronic circuits in terms of the hybrid voltage-amplifier coefficients or the hybrid current-amplifier coefficients, for these characterizations provide useful insight into the properties of electronic amplifiers and permit practical amplifiers to be compared with the ideal.

REFERENCES

1. Guillemin, E. A.: "Introductory Circuit Theory," John Wiley & Sons, Inc., New York, 1953.
2. Van Valkenberg, M. E.: "Network Analysis," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1955.
3. LePage, W. R., and S. Seely: "General Network Analysis," McGraw-Hill Book Company, Inc., New York, 1952.

PROBLEMS

12-1. The circuit shown in Fig. 12-14 is sometimes used in the measurement of very small currents. The tube and circuit parameters are $g_m = 1.0$ millimho, $g_p = 0.1$ millimho, and $G = 0.1$ millimho. Grid currents and parasitic capacitances are negligible for the purposes of this problem.

a. Give an incremental model for the circuit using the current-source representation for the tubes.

b. Write a set of node equations for the circuit, and determine the numerical values of E_2/E_1 and E_3/E_1.

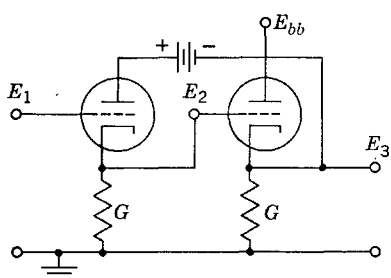


FIG. 12-14. Circuit for Prob. 12-1.

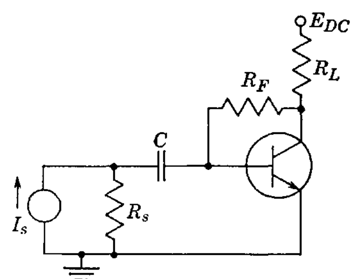


FIG. 12-15. Transistor amplifier for Prob. 12-2.

12-2. The transistor amplifier shown in Fig. 12-15 is used occasionally. The resistance R_F provides the bias current for the base, and it provides feedback that stabilizes the quiescent operating point and reduces the amplification somewhat. The coupling capacitor acts as a short circuit, and the parasitic capacitances of the transistor are negligible at all frequencies of interest. The circuit and transistor parameters are $R_s = 20$ kilohms, $R_L = 10$ kilohms, $R_F = 150$ kilohms, $r_n = 2.0$ kilohms, $\alpha_{cb} = 50$, and $\mu_{bc} = g_o = 0$.

a. Give an incremental model for the circuit using the hybrid representation for the transistor.

b. Write a set of node equations for the circuit, and give the necessary auxiliary equation relating the base current to the unknown node voltages.

c. Determine the numerical value of the current amplification $A_c = I_L/I_s$, where I_L is the signal component of current in R_L.

12-3. The circuit shown in Fig. 12-16 is used as a voltage amplifier in some cathode-ray oscilloscopes. It has the property that changes which occur identically in the two tubes tend to cancel in so far as the output is concerned. (Note that the circuit is a bridge.) Parasitic capacitances and grid currents are negligible for the purposes of this problem.

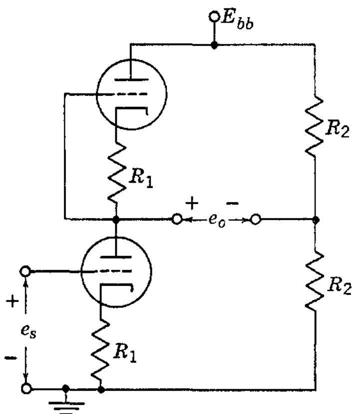


FIG. 12-16. Series-balanced amplifier for Prob. 12-3.

a. Give an incremental model for the circuit using the voltage-source representation for the tubes.

b. Write the loop equation for the circuit with no load connected to the output terminals. Give the necessary auxiliary equations relating the grid-to-cathode voltages for the tubes to the unknown loop current.

c. Determine the voltage amplification, $A = e_o/e_s$, in terms of the tube and circuit parameters. Assume that the two tubes are identical.

12-4. The effect of the unbypassed resistor in the emitter lead of the amplifier shown in Fig. 12-5*a* is to be evaluated. The coupling capacitor acts as a short circuit, and the parasitic transistor capacitances are negligible at all frequencies of interest. The circuit and transistor parameters are $R_s = 20$ kilohms, $R_a = 150$ kilohms, $R_b = 20$ kilohms, $R_2 = 10$ kilohms, $R_3 = 2$ kilohms, $r_n = 2$ kilohms, $\mu_{bc} = 0$, $\alpha_{cb} = 60$, and $r_o = 1/g_o = 70$ kilohms. The input signal is sinusoidal with an amplitude of 1 μa.

a. Give an incremental model for the circuit using the hybrid representation for the transistor. Represent the parallel combination of R_s, R_a, and R_b with a single resistance, and convert all current sources to voltage sources to obtain a two-loop circuit.

b. Write the loop equations for the circuit, and determine the amplitude of the sinusoidal current in R_2.

c. What is the amplitude of the sinusoidal current in R_2 if R_3 is reduced to zero with all other circuit parameters remaining at their original values? What is the ratio of the current amplification with $R_3 = 2$ kilohms to its value with $R_3 = 0$?

12-5. A triode is used in a voltage-amplifier circuit with an unbypassed cathode resistor (Fig. 6-1). The input coupling capacitor acts as a short circuit, and the parasitic capacitances are negligible. The tube and circuit parameters are $\mu = 100$, $r_p = 60$ kilohms, $R_g = 330$ kilohms, $R_L = 220$ kilohms, and $R_k = 1$ kilohm.

a. Give an incremental model for the circuit.

b. Evaluate the hybrid voltage-amplifier coefficients in terms of the tube and circuit parameters by solving the appropriate loop or node equations.

c. Determine the numerical values of the coefficients evaluated in part *b*.

d. Repeat part *c* for $R_k = 0$, assuming that all other circuit parameters retain their original values. What is the ratio of the voltage amplifications for $R_k = 1$ kilohm and $R_k = 0$?

12-6. A certain transistor current amplifier has the form shown in Fig. 12-5*a* with a bypass capacitor added in parallel with R_3. The amplifier is used in the frequency range in which the coupling and bypass capacitors act as short circuits and in which the parasitic capacitances of the transistor are negligible. The circuit and transistor parameters are $R_a = 150$ kilohms, $R_b = 20$ kilohms, $R_2 = 10$ kilohms $R_3 = 2$ kilohms, $r_n = 2$ kilohms, $\mu_{bc} = 0.0005$, $\alpha_{cb} = 50$, and $r_o = 70$ kilohms.

a. Give an incremental model for the amplifier using the hybrid representation for the transistor.

b. Evaluate the hybrid current-amplifier coefficients for the circuit, not including the resistance of the signal source, R_s.

c. What is the input resistance to the amplifier, exclusive of R_s, when a 2-kilohm load resistance is connected from collector to ground in series with a coupling capacitor that acts as a short circuit to signal currents?

12-7. A certain pentode amplifier has the form shown in Fig. 10-16. The circuit parameters are $R_g = 500$ kilohms, $R_s = 30$ kilohms, and $R_L = 50$ kilohms. The cathode bypass capacitor acts as a short circuit at all frequencies of interest. The pentode can be represented by the incremental model shown in Fig. 10-18 with $g_{41} =$ 4.6 millimhos, $g_{42} = 0.16$ millimho, $r_{44} = 1$ megohm, and $\rho = 0.3$.

a. Determine the open-circuit voltage transmittance A_{vo}.

b. If a bypass capacitor that acts as a short circuit to signal currents is connected from the screen grid to ground, what is the value of A_{vo}? What is the ratio of these two values of A_{vo}?

12-8. A 6J5 triode is used in a conventional voltage amplifier (Fig. 6-2). The plate-supply voltage is obtained from a rectifier, and it consists of a direct voltage plus a ripple voltage of 1 mv peak-to-peak amplitude. The power supply can thus be represented as a battery, E_{bb}, in series with a voltage source to account for the ripple. The coupling and bypass capacitors act as short circuits in the frequency range of interest, and the parasitic capacitances are negligible. The tube parameters are $\mu = 20$, $r_p = 10$ kilohms, and $R_L = 50$ kilohms.

It is required that the peak-to-peak signal voltage at the output of the amplifier be at least ten times the peak-to-peak ripple voltage at the output. If this requirement is satisfied, what is the smallest peak-to-peak amplitude that the input signal voltage e_s can have? Note that the ripple voltage sets a lower limit on the amplitude of input signals that can be amplified satisfactorily.

12-9. A 12AU7 twin triode is used in the cathode-coupled amplifier of Fig. 12-4. The two triode sections are identical, and the parasitic capacitances are negligible. Determine the hybrid voltage-amplifier coefficients in terms of the tube and circuit parameters by solving the appropriate node equations.

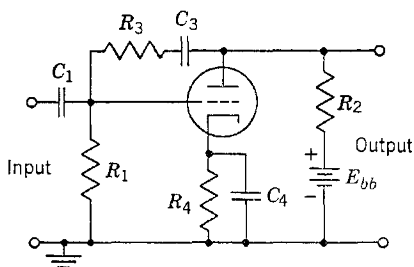


FIG. 12-17. Amplifier for Prob. 12-10.

12-10. Figure 12-17 shows a triode amplifier in which feedback is introduced through R_3 to improve the performance of the circuit in certain respects. The coupling and bypass capacitors act as short circuits in the frequency range of interest, and the parasitic capacitances are negligible.

a. Give an incremental model for the circuit using the current-source representation for the tube. It is instructive to compare this model with the π model for the transistor shown in Fig. 8-15.

b. Determine the voltage-amplifier coefficients in terms of the tube and circuit parameters by solving the appropriate set of node equations. Let $G_1 = 1/R_1$, $G_2 = 1/R_2$, $G_3 = 1/R_3$.

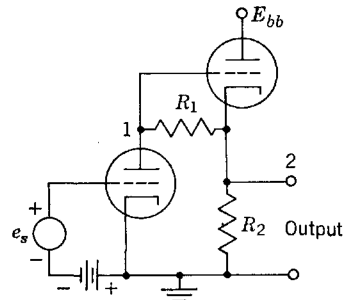


FIG. 12-18. Circuit for Prob. 12-11.

12-11. The circuit shown in Fig. 12-18 has certain useful engineering applications. The two triodes are identical, there is no grid current, and parasitic capacitances are negligible. Determine the forward voltage transmittance A_{vo} in terms of the tube and circuit parameters.

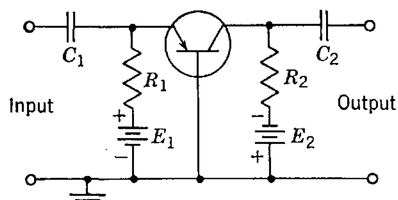


FIG. 12-19. Circuit for Prob. 12-12.

12-12. A transistor amplifier using the grounded-base connection of the transistor is shown in Fig. 12-19. The coupling capacitors act as short circuits to the signal components of current, and the parasitic capacitances are negligible.

a. Give an incremental model for the circuit using the hybrid representation for the transistor. Assume $\mu_{bc} = 0$.

b. Determine the current-amplifier coefficients for the circuit in terms of the transistor and circuit parameters.

c. Can the forward current transmittance A_{cs} ever exceed unity?

d. Can the circuit give voltage amplification in the forward direction?

CHAPTER 13

NETWORK THEOREMS

The loop and node methods of analysis presented in Chap. 12 provide completely general procedures for the analysis of any linear network, no matter how complicated. These general methods of analysis are of fundamental importance to network theory; however, being systematic and mechanistic, they give a limited amount of insight into the properties of specific networks, and they do not necessarily provide the simplest and most direct solution in any specific case. Electrical networks possess certain special properties, expressed in the form of network theorems, that can often be employed to gain useful insight and to obtain direct solutions of particular problems.

A number of valuable theorems have been developed for networks made up of R, L, C, and independent sources. The objective of this chapter is to reexamine these theorems and to determine their applicability to networks containing controlled sources. It is shown in Chap. 12 that the reciprocity theorem does not apply in general to networks containing controlled sources. It is shown in the following sections that certain theorems take on a new significance when controlled sources are present, and a new theorem pertaining directly to controlled sources is presented.

13-1. The Superposition Principle. The node equations for electronic circuits under sinusoidal operating conditions can be put in either of the two general forms expressed by Eqs. (12-11) and (12-12). In these equations the E's are unknown node voltages and the I's are known currents applied to the network. It follows from the fact that these equations are linear that the principle of superposition is applicable. Specifically, let $E'_1, E'_2, \ldots, E'_N$ be solutions of (12-11) and (12-12) for the applied currents $I'_1, I'_2, \ldots, I'_N$, and let $E''_1, E''_2, \ldots, E''_N$ be solutions for the applied currents $I''_1, I''_2, \ldots, I''_N$. Then $(E'_1 + E''_1), (E'_2 + E''_2), \ldots, (E'_N + E''_N)$ are the solutions for the applied currents $(I'_1 + I''_1), (I'_2 + I''_2), \ldots, (I'_N + I''_N)$. The proof of this statement is obtained directly by substitution in either (12-11) or (12-12). A similar argument can be based on the loop equations (12-24) and (12-25).

As a special case of superposition, any unknown voltage, say E_1, can be found by summing the components of E_1 resulting from each of the I's acting one at a time with all other I's made zero.

13-2. The Substitution Theorem. A reexamination of the substitution, or compensation, theorem reveals it to be especially useful in the simplification of circuits containing controlled sources. The circumstances in which it can be used are illustrated in Figs. 13-1*a* and *b*.

The network N in Fig. 13-1*a* can be any network having the terminal current $i(t)$, an arbitrary function of time, and $Ai(t)$ is the voltage of a current-controlled voltage source. The constant A can be any real number. If the network and the source are connected so that the current

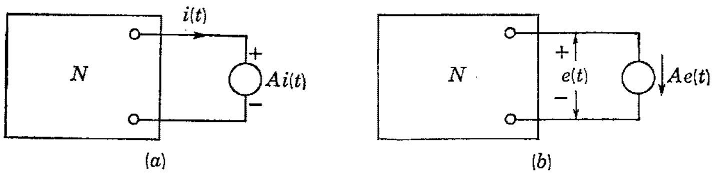


FIG. 13-1. Configurations in which the substitution theorem can be used. (*a*) Current-controlled voltage source; (*b*) voltage-controlled current source.

i flows through the controlled source in the direction of the fall in potential, then the currents and voltages in the network remain unchanged when the controlled source is replaced by a resistance of A ohms. The proof of this theorem follows at once from the general loop equations (12-24).

Figure 13-1*b* illustrates conditions under which the dual form of the theorem can be applied. In this case the voltage-controlled current source $Ae(t)$, which is connected across the voltage $e(t)$, can be replaced by a conductance of A mhos without changing the currents and voltages

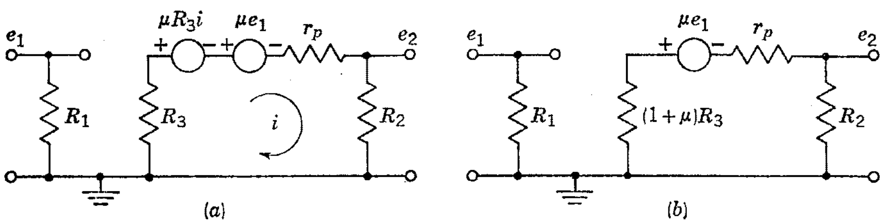


FIG. 13-2. An application of the substitution theorem. (*a*) Circuit; (*b*) equivalent circuit.

in the network. The proof in this case follows directly from the general node equations (12-11).

As a specific example, consider the amplifier of Fig. 12-1*a*. By using the fact that $\mu e_g = \mu e_1 + \mu R_3 i$, the model shown in Fig. 12-2*a* and repeated in Fig. 13-2*a* is obtained. The currents and voltages in this circuit are not changed if the controlled source $\mu R_3 i$ is replaced by a resistance μR_3; the resulting circuit is shown in Fig. 13-2*b*. It is significant that the application of the substitution theorem converts a feedback circuit to a nonfeedback circuit. For this reason the circuit of

Fig. 13-2*b* can be analyzed by inspection, whereas that of Fig. 13-2*a* cannot. It is important to note that the substitution theorem can be applied in this case only if the current through the controlled source is the same as the current through R_3. Thus, for example, it cannot be applied if a resistor is connected in parallel with the tube between plate and cathode.

The circuit shown in Fig. 13-3*a*, which is the prototype for the cathode follower, provides another illustration of the substitution theorem. This important circuit merits a few preliminary remarks. The input voltage is applied between grid and ground, and the output voltage is taken

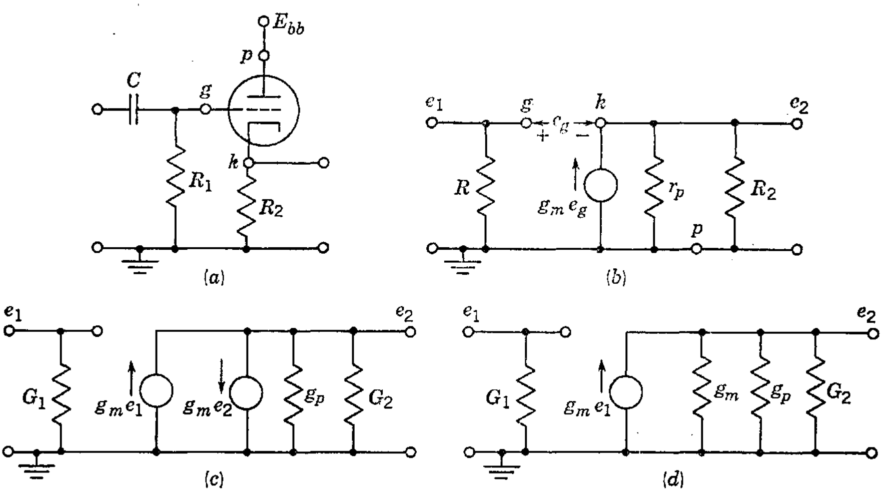


FIG. 13-3. Another application of the substitution theorem. (*a*) Circuit; (*b*) model; (*c*) equivalent circuit; (*d*) simplified equivalent circuit.

between cathode and ground. The plate of the tube is at ground potential in so far as increments of voltage are concerned. A positive increment of input voltage causes a positive increment of plate current and a positive increment of output voltage; hence the cathode potential tends to follow changes in the grid potential. However, the increment of input voltage is always larger than the increment of cathode voltage; it is this fact that permits an increment of plate current. It follows that the voltage amplification of this circuit is always less than unity. The cathode follower has a small input admittance, zero reverse transmittance, and a small output impedance; it therefore serves an important function as an isolating, or buffer, stage between high-impedance sources and low-impedance loads.

A model for the cathode follower that is valid at frequencies where the coupling capacitor acts as a short circuit and where the parasitic capacitances act as open circuits is shown in Fig. 13-3*b*. It is clear from this circuit that $e_g = e_1 - e_2$; thus the controlled current source can be replaced by the equivalent pair of sources shown in Fig. 13-3c. The substitution theorem then permits the current source $g_m e_2$ to be replaced by a conductance g_m, leading to the circuit of Fig. 13-3d. Again feedback is eliminated, and the properties of the cathode follower can be perceived by inspection of the circuit in Fig. 13-3d.

Example 13-1. A practical cathode follower is shown in Fig. 13-4. This circuit arrangement permits a somewhat better circuit design than is possible with the prototype of Fig. 13-3. The resistance $R_2 + R_3$ is chosen to give a suitable operating path on the plate characteristics, and the individual values of R_2 and R_3 are chosen to bias the tube for a suitable quiescent operating point. Since no direct current flows in R_1, the grid bias is the d-c drop across R_3.

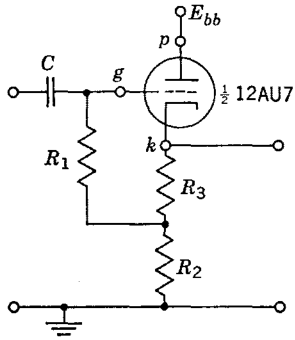


FIG. 13-4. A practical cathode follower for Example 13-1.

The tube used in the particular circuit to be examined is one section of a 12AU7 twin triode; the tube and circuit parameters are $r_p = 10$ kilohms, $g_m = 2.0$ millimhos, $R_1 = 1$ megohm, $R_2 = 50$ kilohms, and $R_3 = 1$ kilohm. The plate-supply voltage is 300 volts. The problem is to evaluate the hybrid voltage-amplifier coefficients with the aid of the substitution theorem.

Solution. A model for the circuit that is valid in the range of frequencies in which C acts as a short circuit and in which the parasitic capacitances are negligible is shown in Fig. 13-5a. After applying the substitution theorem in the manner illustrated in Fig. 13-3, the equivalent circuit of Fig. 13-5 is obtained; application of the theorem removes the feedback from the circuit.

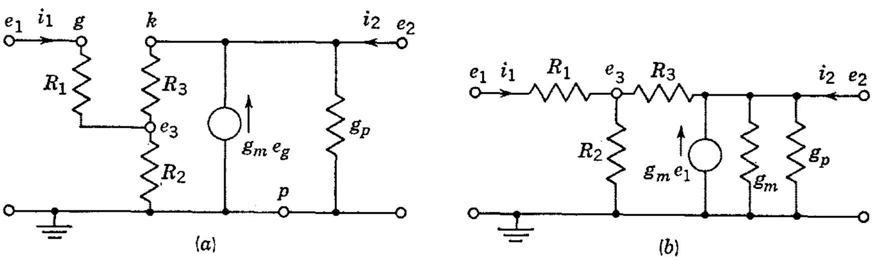


FIG. 13-5. Incremental models for the cathode follower of Fig. 13-4. (a) Model for frequencies at which C acts as a short circuit, $e_g = e_1 - e_2$; (b) model after applying the substitution theorem.

The current in R_1 is much smaller than the current in R_2 and R_3 for two reasons: First, R_1 is much larger than $R_2 + R_3$; and second, an increment of voltage at the grid end of R_1 is accompanied by an almost equal increment at the cathode end, with the result that the increment of voltage across R_1 is much smaller than the increment across $R_2 + R_3$. Thus with $i_2 = 0$, and neglecting the current in R_1, the node equation at the output terminal is

$$\left(g_m + g_p + \frac{1}{R_2 + R_3}\right)e_2 = g_m e_1$$

and
$$A_{vo} = \left.\frac{e_2}{e_1}\right|_{i_2=0} = \frac{g_m}{g_m + g_p + \dfrac{1}{R_2 + R_3}} = \frac{2}{2 + 0.1 + 0.0196} = 0.945$$

This value is somewhat larger than can be obtained with the prototype circuit when the cathode resistor is chosen to give a suitable quiescent point.

With $i_2 = 0$, the input current is

$$i_1 = \frac{e_1 - e_3}{R_1}$$

But, neglecting the small current in R_1 in comparison with the current in R_2,

$$e_3 = \frac{R_2}{R_2 + R_3} e_2 = 0.98e_2 = 0.98A_{vo}e_1$$

Thus
$$i_1 = \frac{e_1 - 0.98A_{vo}e_1}{R_1}$$

and
$$Y_{no} = \left.\frac{i_1}{e_1}\right|_{i_2=0} = \frac{1 - 0.98A_{vo}}{R_1} = \frac{0.074}{R_1} = 7.4 \times 10^{-8} \text{ mho}$$

The corresponding resistance is

$$R_{no} = \frac{1}{Y_{no}} = 13.5 \text{ megohms}$$

Since $R_{no} = R_1$ in the prototype circuit, it is always much smaller than the value obtained with the circuit of Fig. 13-4. Large values for R_{no} are usually desirable.

With $e_1 = 0$, and again neglecting the current in R_1, a current applied at i_2 causes a voltage at e_2 given by

$$e_2 = \frac{1}{g_m + g_p + \dfrac{1}{R_2 + R_3}} i_2 = \frac{A_{vo}}{g_m} i_2$$

Hence the short-circuit output impedance is

$$Z_{os} = \left.\frac{e_2}{i_2}\right|_{e_1=0} = \frac{A_{vo}}{g_m} = \frac{0.945}{2.0} = 0.473 \text{ kilohm}$$

Roughly this same value is realized with the prototype circuit.

The resistance of R_1 in parallel with R_2 is approximately R_2; hence, with $e_1 = 0$, a current applied at i_2 causes a current at i_1 given by

$$i_1 \approx -\left(\frac{e_2}{R_2 + R_3}\right)\left(\frac{R_2}{R_1 + R_2}\right) = -\left(\frac{Z_{os}i_2}{R_2 + R_3}\right)\left(\frac{R_2}{R_1 + R_2}\right)$$

Thus the reverse current transmittance is

$$B_{cs} = \left.\frac{i_1}{i_2}\right|_{e_1=0} = -\frac{Z_{os}R_2}{(R_2 + R_3)(R_1 + R_2)} = -\frac{(0.473)(50)}{(51)(1050)} = -4.41 \times 10^{-4}$$

The value of B_{cs} for the prototype circuit is zero.

Since the reverse current transmittance affects i_1, and hence the load imposed upon the source of signals, it is important to examine its effect more carefully. Figure 13-6 shows the two-port representation of the cathode follower with a load of 600 ohms. The admittance reflected from the output circuit into the input circuit (Fig. 12-9) is

$$Y' = -\frac{B_{cs}A_{vo}}{Z_{os} + R_L} = -\frac{(0.945)(-4.41)(10^{-4})}{1073}$$
$$= 3.88 \times 10^{-7} \text{ mho}$$

Thus the input admittance to the circuit of Fig. 13-6 is

$$Y_n = Y_{no} + Y' = (7.4 + 38.8)(10^{-8})$$
$$= 46.2 \times 10^{-8} \text{ mho} = 0.462 \text{ micromho}$$

The input resistance is

$$R_n = \frac{1}{Y_n} = 2.16 \text{ megohms}$$

Thus the input admittance is determined principally by B_{cs}, rather than Y_{no}, when the load resistance is 600 ohms. The effect of B_{cs} becomes smaller as the load resistance is made larger; however, in many cases the load resistance is fixed at a relatively small value and cannot be changed.

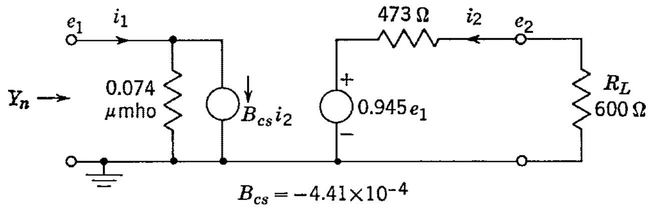


FIG. 13-6. Two-port representation for the cathode follower of Fig. 13-4.

An alternative procedure for examining the effect of the load resistance on the input admittance is to combine R_L in parallel with g_p in the circuit of Fig. 13-5b. The voltage amplification obtained in this case is 0.528, and the input admittance is given by

$$Y_n = \frac{1 - 0.98A_{vo}}{R_1} = \frac{0.482}{R_1} = 0.482 \text{ micromho}$$

The discrepancy between this value and the one obtained previously is largely the result of the approximation made in calculating B_{cs}; accordingly, the last value is the more nearly correct.

13-3. Thevenin's Theorem. The circuit in Fig. 13-7a is an incremental linear model for the triode amplifier with an unbypassed cathode resistor shown in Fig. 12-1. The fact that it is a linear circuit ensures that it can be represented, in so far as the output terminals are concerned, by the Thevenin equivalent circuit[1,2] shown in Fig. 13-7b. The circuit in Fig. 12-8 and the related discussion show this statement to be true for two-terminal-pair networks; the extension of the proof to the general case is not difficult.

Special considerations arise, however, in the evaluation of the internal voltage and resistance of the Thevenin equivalents for circuits containing controlled sources. These considerations can be illustrated by an examination of the circuits in Fig. 13-7. The open-circuit voltage in the circuit of Fig. 13-7*a* is

$$e_{oc} = R_2 i \tag{13-1}$$

An expression for the current i as a function of the circuit parameters and the input voltage e_1 is obtained from the facts that

$$-\mu e_g = (r_p + R_2 + R_3)i \tag{13-2}$$

and

$$\mu e_g = \mu(e_1 + R_3 i) \tag{13-3}$$

Substituting (13-3) and (13-2) and collecting terms gives

$$-\mu e_1 = [r_p + R_2 + (1 + \mu)R_3]i \tag{13-4}$$

Solving this expression for i and substituting in (13-1) yields

$$e_{oc} = \frac{-R_2\mu e_1}{r_p + R_2 + (1 + \mu)R_3} = A_{vo}e_1 \tag{13-5}$$

This is the voltage of the source in the equivalent circuit of Fig. 13-7*b*.

A procedure that is commonly used in evaluating the internal resistance for Thevenin equivalent circuits is to determine the resistance between

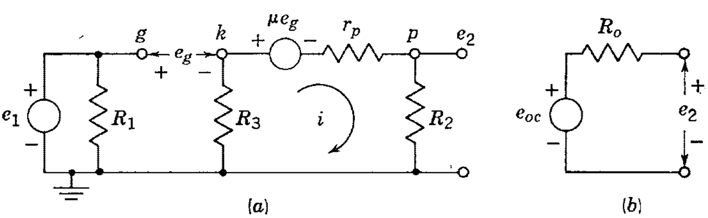


FIG. 13-7. The Thevenin equivalent for a vacuum-tube circuit. (*a*) Circuit; (*b*) equivalent circuit.

the output terminals with all sources made zero; for the circuit in Fig. 13-4*a* this resistance is

$$R_o = \frac{R_2(r_p + R_3)}{r_p + R_2 + R_3} \tag{13-6}$$

Another commonly used procedure is to evaluate R_o from the ratio of open-circuit voltage to short-circuit current. The current in a short circuit connected across the output terminals is given by Eq. (13-4) with $R_2 = 0$; hence

$$i_{sc} = \frac{-\mu e_1}{r_p + (1 + \mu)R_3} \tag{13-7}$$

and

$$R_o = \frac{e_{oc}}{i_{sc}} = \frac{R_2[r_p + (1 + \mu)R_3]}{r_p + R_2 + (1 + \mu)R_3} \tag{13-8}$$

Equations (13-6) and (13-8) do not give the same value for R_o. Thus two conventional procedures for evaluating the internal resistance of Thevenin equivalent circuits, which always give identical results when applied to circuits composed of R, L, C, and independent sources, yield different results when applied to a circuit containing controlled sources. Two important questions must therefore be answered: First, does either of the values of R_o obtained above give a correct representation of the circuit; and second, if one of the values obtained above is correct, why is the other incorrect?

Evaluating R_o from the ratio of open-circuit voltage to short-circuit current involves only measurements made at the output terminals of the circuit; hence this procedure gives a correct measure of the output-terminal characteristics of the circuit, and the value of R_o obtained in this manner yields a correct representation, in so far as the output terminals are concerned, when used in the equivalent circuit of Fig. 13-7*b*. When the resistance between the output terminals is evaluated with all sources in the circuit made zero, the circuit is modified internally, and the value of resistance so obtained does not give a correct representation when used in the circuit of Fig. 13-7*b*. The effect of setting the voltage of the μe_g source in Fig. 13-7*a* equal to zero is illustrated clearly by the two circuits of Fig. 13-2; these circuits are entirely equivalent to that of Fig. 13-7*a* in so far as the output terminals are concerned. It is clear from these circuits that the μe_g source behaves as a voltage source μe_1 in series with a resistance μR_3. Thus, replacing the μe_g source with a short circuit is equivalent to short-circuiting a voltage source μe_1 and a resistance μR_3. The constant voltage source in the Thevenin equivalent circuit, e_{oc}, accounts for the effects of the constant voltage source μe_1, but it cannot account for the effects of the resistance μR_3. These considerations lead to the following general rule: The Thevenin equivalent for a circuit containing controlled sources consists of a source of voltage equal to the open-circuit voltage appearing at the output terminals of the circuit, acting in series with the impedance appearing between the output terminals, with all *independent* sources in the circuit adjusted for zero volts or amperes. The controlled sources in the circuit must not be removed when the impedance of the Thevenin equivalent circuit is evaluated.

The impedance to be used in the Thevenin equivalent for a circuit containing controlled sources can be evaluated correctly in a variety of ways; however, the determination based on the open-circuit voltage and the short-circuit current is often the simplest, especially since the open-circuit voltage must be determined in any event.

13-4. A Reduction Theorem for Controlled Sources. The reduction theorem presented in the following paragraphs is in effect an extension of the substitution theorem. The substitution theorem permits the elimination of current-controlled voltage sources and voltage-controlled current sources in certain configurations; the reduction theorem permits the elimination of voltage-controlled voltage sources and current-controlled current sources in certain other configurations. The reduction theorem, like the substitution theorem, eliminates feedback and often makes circuit analysis possible by inspection.

The reduction theorem applies to the network configurations shown in Fig. 13-8. The network N_1 in Fig. 13-8*a* is any linear two-terminal network having the terminal voltage $e_1(t)$, an arbitrary function of time, and N_2 is any other linear two-terminal network. The constant A of the voltage-controlled voltage source $Ae_1(t)$ is any real number, positive or negative. The two networks and the voltage source are connected in

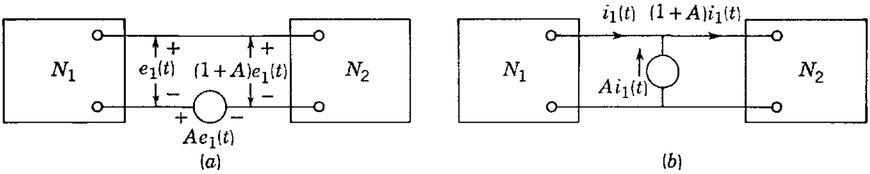


FIG. 13-8. Circuits that can be simplified by the reduction theorem. (*a*) Voltage-controlled voltage source; (*b*) current-controlled current source.

series so that the voltages e_1 and Ae_1 are additive with respect to the loop that they form. The theorem states that all currents in N_1 and N_2 remain unchanged if the source Ae_1 is replaced with a short circuit and if:

(*a*) Each resistance, inductance, elastance (elastance is $S = 1/C$, the reciprocal of capacitance), and voltage source in N_1 is multiplied by $1 + A$

or

(*b*) Each resistance, inductance, elastance, and voltage source in N_2 is divided by $1 + A$

Proof of the theorem is based on the form of the loop equations [see Eqs. (12-24) for the case of sinusoidal operation]. It is clear that if all resistances, inductances, elastances, and voltage sources are multiplied by the same arbitrary factor, then every term in the set of loop equations is multiplied by the same factor, and all the loop currents remain unchanged. Thus if all the parameters and voltage sources in N_1 are multiplied by $1 + A$ as specified in part *a* of the theorem, and if Ae_1 is replaced with a short circuit, then the voltage applied to N_2 is still $(1 + A)e_1$, and all currents in N_1 and N_2 remain unchanged. All voltages in N_1 are multiplied by the factor $1 + A$, but the voltages in N_2 are not changed. Any current sources in N_1 and N_2 are left unchanged in the reduction process.

By the same reasoning process, after the transformation described above is completed, all resistances, inductances, elastances, and voltage sources in the composite network N_1-N_2 can be divided by the factor $1 + A$ without changing any of the currents in N_1-N_2. Thus part *b* of the theorem is proved.

When all currents and voltages in the network are sinusoidal and of the same frequency, the constant of the controlled source may be expressed as a complex number; the reduction theorem holds in this case also.

The dual form of the reduction theorem is applicable to the network configuration shown in Fig. 13-8*b*. It states that all voltages in the network of Fig. 13-8*b* remain unchanged if the current source Ai_1 is replaced with an open circuit and if:

(*a*) All conductances, capacitances, reciprocal inductances, and current sources in N_1 are multiplied by $1 + A$

or

(*b*) All conductances, capacitances, reciprocal inductances, and current sources in N_2 are divided by $1 + A$

Proof of the dual form of the theorem follows from the node equations (12-11) and is dual to the proof outlined above.

In applying the reduction theorem it is essential that the two networks N_1 and N_2 be properly identified at the outset. The two necessary requirements are: (1) the terminal voltage (or current) of the network identified as N_1 must be the controlling quantity for the source to be eliminated; and (2) no current may enter or leave either network through any terminal other than the two terminals by which the networks and controlled source are joined in series (or parallel). It should be noted that connections joining separate parts of a network carry no current and can therefore be ignored in respect to item 2.

A useful application of the reduction theorem is provided by the cascode amplifier shown in Fig. 13-9*a*. This circuit employs two identical triodes, T_1 and T_2; the resistors R_4 and R_5 provide a suitable grid bias for T_2. An incremental model for the amplifier is shown in Fig. 13-9*b*. Since no signal current flows in R_4 and R_5, these resistors play no active part in the operation of the circuit, and they are omitted from the model.

The incremental grid voltage for T_1 is $e_{g1} = e_1 + R_3i$; hence the circuit can be simplified at the outset by applying the substitution theorem to obtain the equivalent circuit shown in Fig. 13-10*a*. This circuit has been separated into two networks, N_1 and N_2, connected in series with the controlled source μe_3, in preparation for the reduction theorem. All the conditions of the reduction theorem are satisfied by this circuit; hence invoking part *a* of the voltage-source form of the theorem yields the reduced network shown in Fig. 13-10*b*. All feedback is eliminated from the circuit, and the reduced network can be analyzed by inspection. It

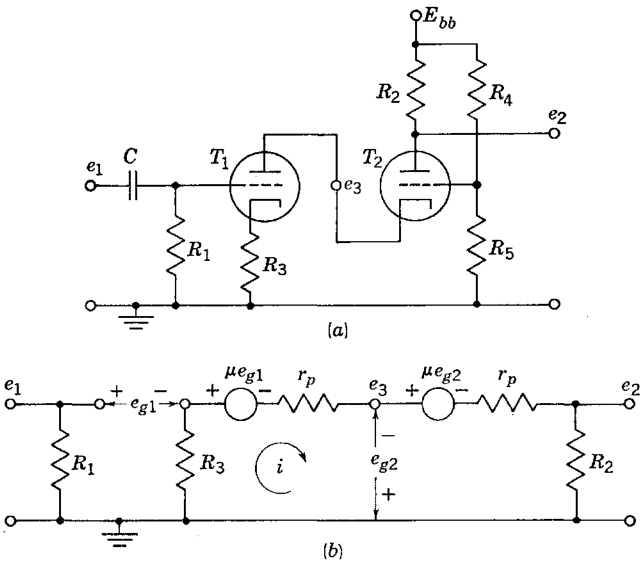


FIG. 13-9. The cascode amplifier. (a) Circuit; (b) model.

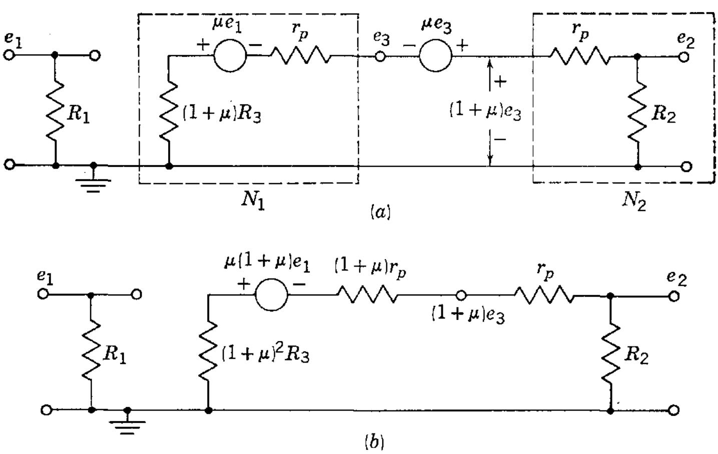


FIG. 13-10. Simplification of the cascode amplifier circuit by application of the reduction theorem. (a) Network prepared for the theorem; (b) reduced network.

is clear from the reduced network that if μ is much larger than unity, the two triodes in cascode act very much as a pentode with an amplification factor μ^2 and a plate resistance μr_p.

The dual form of the reduction theorem can be illustrated with the amplifier shown in Fig. 13-11a. This circuit uses a transistor in the grounded-collector connection. Figure 13-11b shows an incremental model for the amplifier, using the T representation for the transistor, in which it is assumed that the coupling capacitors are short circuits and that the transistor is independent of frequency. The resistance R_1 represents the parallel combination of R_s, R_a, and R_b, and the network is separated into two parts in preparation for the reduction theorem.

Invoking part *a* of the current-source form of the theorem yields the reduced network shown in Fig. 13-11*c*; for compactness the factor $1 + \alpha_{cb}$ is symbolized by k in this diagram. Dividing the resistances R_1 and r_b by k is equivalent to multiplying the corresponding conductances by this factor. Application of the reduction theorem removes all feedback from the circuit; accordingly, the properties of the amplifier can be

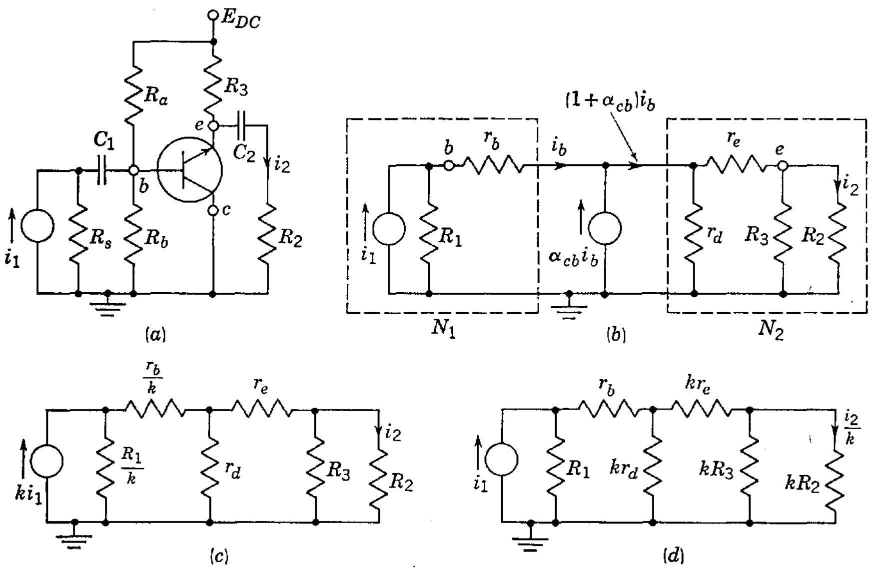


FIG. 13-11. Another application of the reduction theorem. (*a*) Circuit; (*b*) model; (*c*) reduced network, $k = 1 + \alpha_{cb}$; (*d*) alternative reduced network, $k = 1 + \alpha_{cb}$.

seen more clearly from the reduced network than from the original model. The current amplification of the circuit is accounted for by the fact that the input signal current is multiplied by the factor k in the reduced network.

If part *b* of the theorem is applied instead of part *a*, the circuit of Fig. 13-11*d* results. Again all feedback is removed, and a simpler representation results.

The circuit of Fig. 13-11*b* consists of a ladder network having as one of its shunt branches a current source controlled by the current in an adjacent series branch. Applying the reduction theorem to this configuration shows that the controlled current source has the effect of transforming all impedances and currents on one side of the source with respect to the impedances and currents on the other side. Hence current-controlled current sources, realized with the aid of transistors, find useful applications as impedance converters.

It is to be expected that the dual of the configuration shown in Fig. 13-11*b* performs a similar function of impedance conversion. The dual configuration consists of a ladder network having as one of its series branches a voltage source controlled by the voltage across an adjacent shunt branch; it is realized with the grounded-grid triode shown in Fig. 13-12*a*. An incremental model for the circuit is shown in Fig. 13-12*b*. Applying the reduction theorem to the model yields the reduced

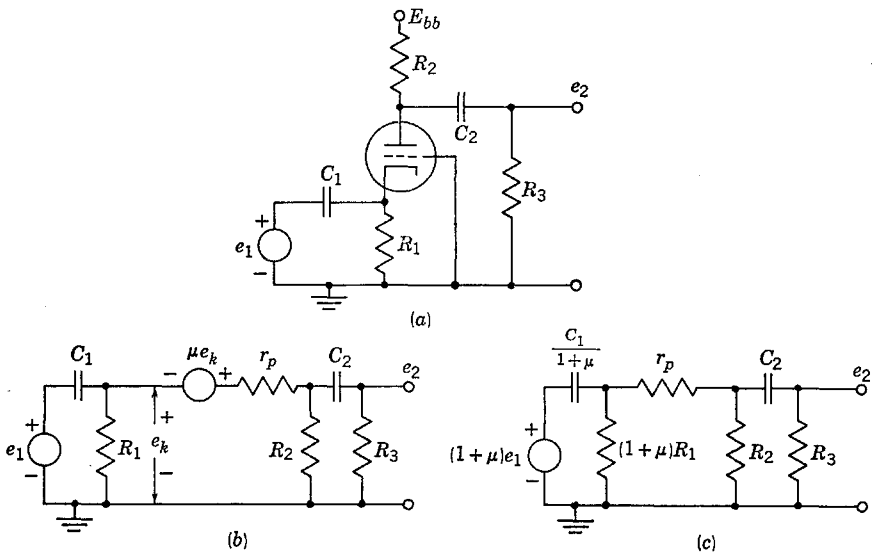


FIG. 13-12. The grounded-grid amplifier as an impedance converter. (*a*) Circuit; (*b*) model; (*c*) reduced network.

network shown in Fig. 13-12*c*; the impedance-transforming property of the grounded-grid triode is exhibited clearly by this circuit.

13-5. The Bisection Theorem. Figure 13-13 illustrates symbolically a network that can be separated into two parts which are mirror images about an axis of symmetry. Many electronic circuits are given this form deliberately, for certain advantages result from the symmetry. The bisection theorem[2] provides a means for using the symmetry of the circuit to reduce the labor required in analysis. It provides, in addition, a point of view that gives further useful insight into the properties of symmetrical networks. Use of the bisection theorem requires that the network contain no internal independent sources; there may, however, be any number of links connecting homologous points in the two parts of the network. The distinguishing feature of such networks, apart from the excitations e_1 and e_1', is the fact that a rotation of 180° about the axis of symmetry produces no change in the network configuration. A circuit meeting these conditions is shown in Fig. 13-14*b*.

The bisection theorem is concerned with the behavior of symmetrical networks with symmetrical and antisymmetrical excitations. Specifically, it states that when the network is excited in the common mode with $e_1 = e_1' = e_c$, the currents and voltages throughout the network are not disturbed if the links joining the two parts of the network are cut, and when the network is excited in the differential mode with $e_1 = -e_1' = e_d$, the currents and voltages are not disturbed if the links are cut and joined together separately in the two parts.

Proof of the theorem is obtained by applying superposition to the two input voltages e_1 and e_1'. With a certain voltage, say e_a, applied at e_1 with $e_1' = 0$, there is a certain current in each link and a certain potential difference between each pair of links. Let the current in link k be designated i_k, and let the potential of link k with respect to link j be

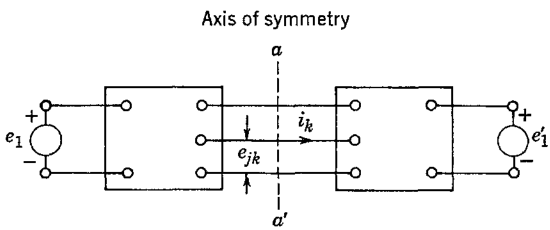


FIG. 13-13. A symmetrical network.

designated e_{kj}. If the voltage e_a is now applied at e_1' with $e_1 = 0$, the current in link k is $-i_k$, and the potential of link k with respect to link j is again e_{kj}. This statement follows from the symmetry of the network, for shifting the excitation e_a from the e_1 terminals to the e_1' terminals is equivalent to rotating the network 180° about the axis of symmetry. When the network is excited in the common mode with $e_1 = e_1' = e_a$, the voltages and currents in the network are obtained by superposing the contributions from the two excitations; hence the total current in link k with common-mode excitation is

$$i_{ko} = i_k - i_k = 0 \tag{13-9}$$

and the total potential of link k relative to link j is

$$e_{kjo} = e_{kj} + e_{kj} = 2e_{kj} \tag{13-10}$$

Equation (13-9) shows that there is no current in the links with common-mode excitation; therefore under this condition the links can be cut without disturbing the currents and voltages.

If a voltage $-e_a$ is applied at the e_1' terminals with $e_1 = 0$, the current in link k is i_k, and the potential of link k relative to link j is $-e_{kj}$. This statement follows from the symmetry of the network and from the fact that reversing the excitation to any network reverses the currents and voltages throughout the network. When the network is excited in the differential mode with $e_1 = -e_1' = e_a$, the voltages and currents in the network are obtained by superposing the contributions from the two excitations; hence the total current in link k with differential-mode excitation is

$$i_{ko} = i_k + i_k + 2i_k \tag{13-11}$$

and the total potential of link k relative to link j is

$$e_{kjo} = e_{kj} - e_{kj} = 0 \tag{13-12}$$

Equation (13-12) shows that there is no potential difference between the links with differential-mode excitation; therefore under this condition the links can be joined together on each side of the axis of symmetry and then cut on the axis without disturbing the currents and voltages in the network.

Any symmetrical network with either common-mode or differential-mode excitation can be bisected, and the analysis of the circuit is concerned with only half of the original network. The analysis of the bisected network is further simplified by the fact that either the link currents or the link-to-link voltages are zero. When the excitation voltages are arbitrary and bear no special relation to one another, the bisection theorem can still be used if the network is symmetrical, for any pair of input voltages can be represented as the superposition of a common-mode pair and a differential-mode pair. Thus any pair of input voltages can be expressed as

$$e_1 = e_c + e_d \qquad \text{and} \qquad e_1' = e_c - e_d \tag{13-13}$$

where e_c and e_d are, respectively, the common-mode and the differential-mode components. When the input voltages e_1 and e_1' are known, the common-mode and differential-mode components can be found by first adding and then subtracting Eqs. (13-13) to obtain

$$e_c = \frac{e_1 + e_1'}{2} \qquad \text{and} \qquad e_d = \frac{e_1 - e_1'}{2} \tag{13-14}$$

If the network is linear, it can be analyzed separately for each of these symmetrical components of excitation with the aid of the bisection theorem, and the results can be combined by superposition to obtain total quantities.

The application of the bisection theorem can be illustrated with the circuit shown in Fig. 13-14a. This circuit is the prototype for a d-c vacuum-tube voltmeter; the deflection of the ammeter M is proportional to the potential difference $e_1 - e_1'$. An incremental model for the circuit is shown in Fig. 13-14*b*. In this model the resistance of the ammeter is represented by R_m, and the circuit is prepared for bisection by separating R_m into an equivalent pair of series resistors and by separating R_3 into an equivalent pair of parallel resistors.

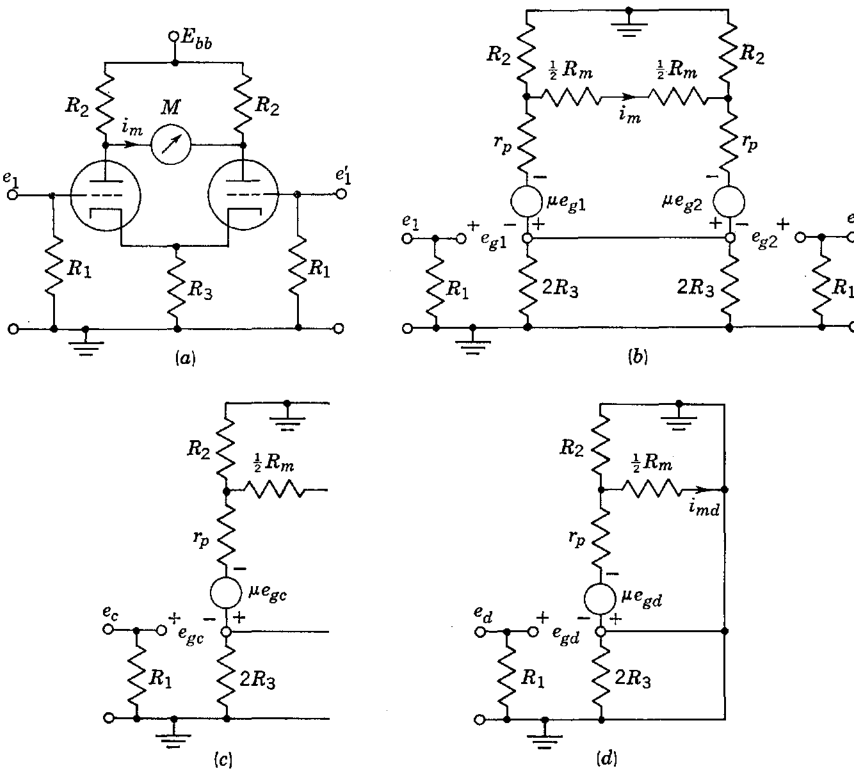


FIG. 13-14. An application of the bisection theorem. (*a*) Circuit; (*b*) model; (*c*) model for the common mode; (*d*) model for the differential mode.

For the common-mode components of current and voltage, the model can be simplified by the bisection theorem to that shown in Fig. 13-14*c*. The currents and voltages in the right-hand half of the circuit are mirror images of those in the left-hand half; hence they need not be calculated separately. The analysis can be further simplified by using the substitution theorem to eliminate the cathode degeneration. It is clear that the common-mode component of the meter current is zero in all events.

For the differential-mode components of current and voltage, the model can be simplified by the bisection theorem to that shown in Fig. 13-14*d*, In this case also it is sufficient to analyze only half of the circuit. There is no cathode degeneration in the model for the differential-mode components, for the cathode resistor in this model, $2R_3$, is short-circuited. The differential-mode component of the meter current, which is also the total meter current, can be evaluated by inspection of the circuit in Fig. 13-14*d*. Since this current depends only on the differential-mode component of excitation, $e_d = (e_1 - e_1')/2$, the circuit of Fig. 13-14*a* is often referred to as a difference amplifier.

The unknown voltage to be measured by the vacuum-tube voltmeter of Fig. 13-14*a* is usually applied at e_1. In this case e_1' is zero, and the differential-mode component of the input voltages is $e_d = e_1/2$. The common-mode component in this case is $e_c = e_1/2$ also; it contributes nothing to the meter current, however.

13-6. Summary. Most of the theorems that are useful in the study of nonelectronic circuits can also be applied to linear circuits containing controlled sources. A notable exception is the reciprocity theorem. When applying Thevenin's theorem to circuits with controlled sources, care must be taken to ensure that the elements of the equivalent circuit are evaluated by a proper method. Certain methods that are useful in the analysis of nonelectronic circuits are not valid when controlled sources are present.

The substitution theorem is especially useful in the study of circuits with controlled sources. Occasions for using this theorem arise frequently, and its application always leads to a simpler network representation by removing part or all of the feedback in the circuit. The reduction theorem is similar in many respects to the substitution theorem, and its value is even greater, for there are more opportunities to use it.

The bisection theorem, which is based upon symmetrical components of current and voltage in symmetrical circuits, permits advantage to be taken of the symmetry that exists in certain types of electronic circuits. The amount of labor involved in the analysis of these circuits may be greatly reduced by the use of the bisection theorem.

The network theorems can be applied only to circuits meeting the conditions specified by the theorem. In using these theorems it is necessary to ascertain that each condition required by the theorem is satisfied. In some cases a rearrangement of the circuit is required before a theorem can be applied, and in other cases it may not be possible to apply a given theorem at all. The principle of superposition and Thevenin's theorem require only that the circuit be linear; however, the substitution, reduction, and bisection theorems can be applied only to circuits having particular configurations.

The application of any network theorem is equivalent to an algebraic rearrangement of the equations relating the currents and voltages in the network, and equivalent results can be obtained in each case by rearranging the equations rather than the network. However, operating on the network usually provides more insight into the properties of the circuit than manipulating the equations, and it is usually easier to discern helpful manipulations by inspection of the network than by inspection of the equations.

REFERENCES

1. Guillemin, E. A.: "Introductory Circuit Theory," John Wiley & Sons, Inc., New York, 1953.
2. LePage, W. R., and S. Seely: "General Network Analysis," McGraw-Hill Book Company, Inc., New York, 1952.

PROBLEMS

13-1. Two identical triodes are used in the circuit of Fig. 12-18. The tube and circuit parameters are $\mu = 20$, $r_p = 10$ kilohms, $R_1 = 5$ kilohms, and $R_2 = 50$ kilohms. Grid currents and parasitic capacitances are negligible. Give an incremental model for the circuit, and determine the internal voltage (in terms of e_s) and the impedance of the Thevenin equivalent circuit with respect to the indicated output terminals. *Note:* This problem can be solved in a straightforward way by the use of loop or node equations; alternatively, the theorems of Chap. 13 can be used in a variety of ways to obtain a solution.

13-2. The transistor amplifier shown in Fig. 12-15 is used with signals lying in the band of frequencies in which the coupling capacitor acts as a short circuit and in which the parasitic transistor capacitances are negligible. The circuit parameters have the values given in Prob. 12-2. Give an incremental model for the amplifier, and determine the internal current (in terms of I_s) and the impedance of the Norton equivalent of the circuit connected to R_L.

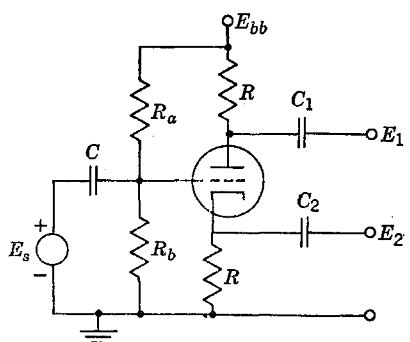


FIG. 13-15. Single-tube phase inverter for Prob. 13-3.

13-3. The circuit shown in Fig. 13-15 is a single-tube phase inverter. Its function is to convert a single signal voltage into two signal voltages that are equal in amplitude but opposite in sign. In normal operation, the same current flows in each of the resistors designated R; hence if these resistors are identical, $|E_1| = |E_2|$. The resistors R_a and R_b provide a suitable grid bias for the tube. The coupling capacitors act as short circuits at all frequencies of interest, the parasitic capacitances are negligible, and there is no grid current.

a. Give an incremental model for the circuit using the voltage-source representation for the tube.

b. Noting that the incremental grid voltage is $E_g = E_s - E_2$, apply the reduction theorem to obtain an equivalent circuit in which the voltage E_1 remains unchanged (part *a* of the theorem).

c. Use the reduction theorem to obtain an equivalent circuit in which E_2 remains unchanged.

d. Evaluate the voltage ratios E_1/E_s and E_2/E_s in terms of the tube and circuit parameters.

13-4. The grounded-collector transistor amplifier shown in Fig. 13-11*a* is used in the range of frequencies in which the coupling capacitors act as short circuits and in which the parasitic capacitances are negligible.

a. Give an incremental model for the amplifier using the hybrid representation for the transistor with $\mu_{bc} \neq 0$. Designate the parallel combination of R_a, R_b, and R_s by R_1.

b. By successive applications of the reduction theorem, obtain an equivalent circuit having no controlled sources. Let $1 + \alpha_{cb} = k_1$ and $1 - \mu_{bc} = k_2$.

13-5. Problem 5-13 is concerned with determining the quiescent operating conditions in a cascode amplifier. Use the reduction theorem to simplify the piecewise-linear model for this circuit, and solve parts *a* and *b* of Prob. 5-13.

13-6. The circuit shown in Fig. 13-16 is the prototype for a d-c vacuum-tube voltmeter. The tubes are identical, there is no grid current, and parasitic capacitances are negligible.

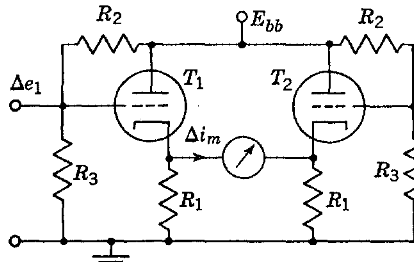


FIG. 13-16. Vacuum-tube voltmeter for Prob. 13-6.

a. Give an incremental model for the circuit using the current-source representation for the tubes.

b. Apply the substitution theorem to eliminate all controlled sources except one controlled by Δe_1.

c. Assuming that the ammeter has zero resistance, evaluate the transfer conductance, $G_t = \Delta i_m/\Delta e_1$. If the tube is a 6SN7 twin triode, roughly how many volts are required at Δe_1 to make $\Delta i_m = 1$ ma?

13-7. The circuit shown in Fig. 13-17 is a sensitive d-c microammeter. If a resistance is connected in series with the input and adjusted to a suitable value, the circuit serves as a high-resistance voltmeter that can substitute for a vacuum-tube voltmeter.

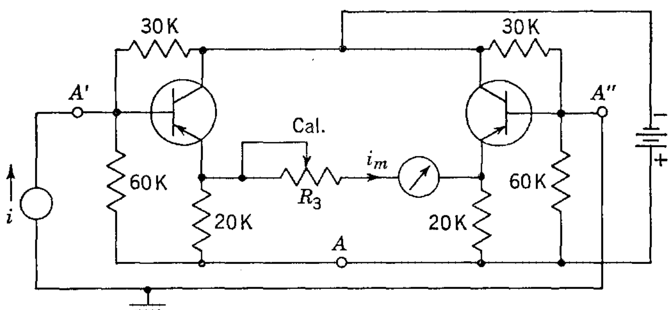


FIG. 13-17. Transistorized microammeter for Prob. 13-7.

The calibrating resistor R_3 is so located in the circuit that it does not affect the quiescent operating point. The problem is to examine the properties of the circuit.

a. Give an incremental model for the circuit that is valid at low frequencies. Use the hybrid representation for the transistors with $\mu_{bc} = 0$, $r_n = 2.5$ kilohms, $\alpha_{cb} = 50$, and $r_o = 1/g_o = 70$ kilohms.

b. Replace the current source i with an equivalent pair of sources, one between terminals A'-A and the other between terminals A''-A (Fig. 12-3). Eliminate all controlled sources from the circuit by successive applications of the reduction theorem to the portion of the circuit on the right of the controlled source (Fig. 13-11).

c. If 250 μa is required for full-scale deflection of the meter, and if R_3 and the meter resistance are negligible, what input current is required for full-scale deflection?

13-8. A 12AT7 twin triode is used in a cathode-coupled amplifier similar to the one shown in Fig. 12-4; d-c sources are added in the grid circuits to provide suitable bias. The triodes are identical, there is no grid current, and parasitic capacitances are negligible. The tube and circuit parameters are $\mu = 60$, $r_p = 15$ kilohms, $R_1 =$ 500 kilohms, $R_2 = 5$ kilohms, and $R_3 = 50$ kilohms.

a. Give an incremental model for the amplifier using the current-source representation for the input tube and the voltage-source representation for the output tube.

b. Apply the substitution and reduction theorems to obtain an equivalent circuit having a single controlled source depending only on the input voltage e_1.

c. Determine the numerical values of the hybrid voltage-amplifier coefficients.

13-9. The high-impedance source in Fig. 13-18 is required to deliver a signal voltage to the relatively low-impedance load. The purpose of the problem is to study, with the aid of the substitution theorem, the use of a cathode follower to isolate the load from the source.

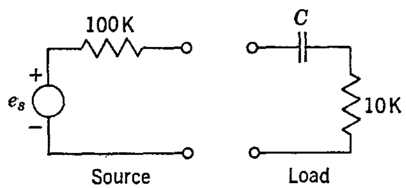


FIG. 13-18. A high-impedance source and a low-impedance load for Prob. 13-9. $e_s = \cos \omega_s t$.

a. If the load is connected directly to the source, and if the coupling capacitor acts as a short circuit, what is the amplitude of the signal voltage across the load?

b. The load is to be isolated from the source by a cathode follower of the form shown in Fig. 13-3; the tube is to be a 6J5 triode, and the grid return resistance R_1 is to be 1 megohm. Determine the values of R_2 and E_{bb} required for a quiescent operating point at $I_{bo} = 4$ ma and $E_{co} = -8$ volts. The techniques presented in Chap. 6 can be adapted to this problem.

c. Give an incremental model for the cathode follower, treating the coupling capacitor as a short circuit and neglecting the parasitic capacitances. Determine the voltage-amplifier coefficients for this circuit. The tube parameters can be approximated by $\mu = 20$ and $r_p = 10$ kilohms.

d. Using the results of part *c*, determine the signal voltage at the input to the cathode follower and the signal voltage across the load when the cathode follower is connected between the source and the load in Fig. 13-18. Explain briefly the function performed by the cathode follower in this application.

13-10. The cathode follower of Example 13-1 is modified by the addition of a bypass capacitor in parallel with the bias resistor R_3. This capacitor acts as a short circuit in the frequency range of interest, and parasitic capacitances are negligible. Evaluate the voltage-amplifier coefficients for the circuit, and compare them with the values obtained in Example 13-1.

13-11. The current source and 5-megohm shunt resistance shown in Fig. 13-19 represent the electrical properties of a photoelectric tube used in measuring light intensity. The signal from the photoelectric tube is transmitted to a cathode follower through a shielded cable that is used to minimize the introduction of stray voltages into the circuit by electrostatic induction. The coupling capacitor acts as a short circuit at the signal frequency, and the parasitic capacitances are negligible. The insulation resistance between the conductor and the shield of the cable is 5 megohms.

a. Find the values of R_1 and R_2 required for a quiescent operating point at $I_{bo} =$ 6 ma and $E_{bo} = 150$ volts.

b. If $I = 1$ μa, what is the value of E_1 when the shield of the cable is grounded as shown in Fig. 13-19? Note that the insulation resistance of the cable cannot be neglected. The tube parameters can be approximated by $\mu = 20$ and $r_p = 10$ kilohms.

c. Repeat part *b* for the case where the shield is connected to the junction of R_1 and R_2. Comment on the relative merits of these two connections.

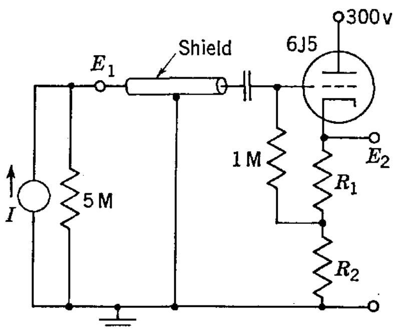


FIG. 13-19. Cathode-follower circuit for Prob. 13-11.

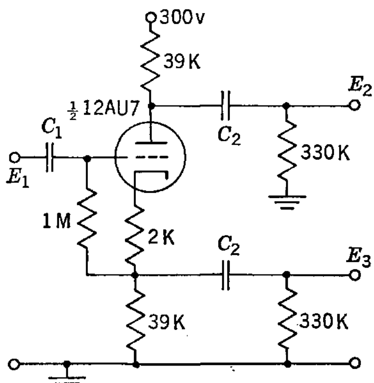


FIG. 13-20. Single-tube phase inverter for Prob. 13-12.

13-12. The circuit shown in Fig. 13-20 is a modified form of the single-tube phase inverter discussed in Prob. 13-3. The tube parameters are $\mu = 17$, $r_p = 20$ kilohms; the coupling capacitors can be treated as short circuits for the signal components of current, and the parasitic capacitances can be neglected. The objective of the problem is to study the properties of the circuit with the aid of the reduction theorem.

a. Determine the amplitudes of E_2 and E_3 when E_1 is a sinusoid with an amplitude of 10 volts. Note that as in the case of the cathode follower analyzed in Example 13-1, the current in the 1-megohm grid return resistor can be neglected. Part *a* of the reduction theorem is useful in evaluating E_2; part *b* is useful in evaluating E_3.

b. The output voltages E_2 and E_3 are measured one at a time with a voltmeter having a resistance of 100 kilohms. What is the reading of the meter in each case?

c. Why are the two voltmeter readings of part *b* unequal? *Suggestion:* Consider the Thevenin equivalent circuit for each of the two output terminal pairs.

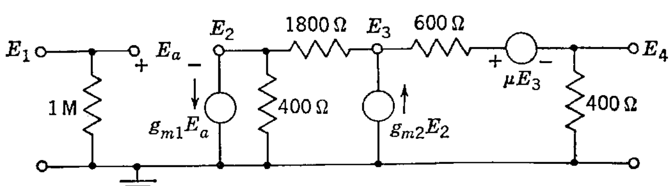


FIG. 13-21. Circuit employing controlled sources for Prob. 13-13. $E_a = E_1 - E_2$, $g_{m1} = 2.5$ millimhos, $g_{m2} = 0.5$ millimho, and $\mu = 0.5$.

13-13. Simplify the circuit of Fig. 13-21 as much as possible by the use of network theorems, and evaluate $A_{vo} = E_4/E_1$.

13-14. The transistorized microammeter of Prob. 13-7 is a symmetrical circuit that can be simplified by the bisection theorem.

a. Give an incremental model for the circuit, and replace the current source i with an equivalent pair of sources as specified in part *b* of Prob. 13-7.

b. Give a simplified circuit for the differential-mode components of current and voltage.

c. If the sum of R_3 and the ammeter resistance is 500 ohms, what is the ratio of i_m to i? Use the circuit and transistor parameters given in Prob. 13-7.

13-15. The prototype vacuum-tube voltmeter of Prob. 13-6 is a symmetrical circuit that can be simplified by the bisection theorem. The input voltage can be represented as the superposition of a common-mode pair and a differential-mode pair applied to the two grids (refer to the discussion related to Fig. 13-14).

a. Give simplified incremental models of the circuit for the differential-mode and common-mode components of current and voltage. Designate the meter resistance by R_m.

b. Determine the ratio of Δi_m to Δe_1 in terms of the tube and circuit parameters.

13-16. The difference amplifier shown in Fig. 13-22 is to be used in an analog computer to perform the operation of subtraction. The two triodes are identical, there are no grid currents, and the parasitic capacitances are negligible.

a. Give incremental models for the differential-mode and the common-mode components of current and voltage.

b. Show that the output voltage is given by an expression of the form $e_o = A(e_1 - e_1')$. Determine the value of A in terms of the tube and circuit parameters.

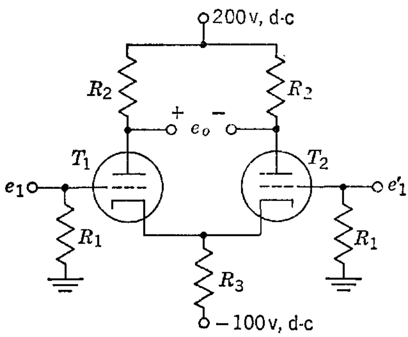


FIG. 13-22. Difference amplifier for Prob. 13-16.

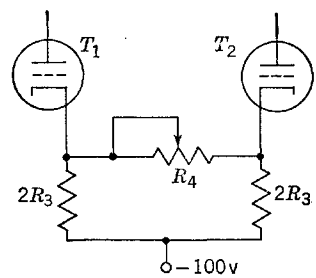


FIG. 13-23. Gain control for the difference amplifier of Fig. 13-22.

13-17. The cathode circuit in the difference amplifier of Fig. 13-22 is rearranged as shown in Fig. 13-23. With the aid of the bisection theorem, show that R_4 can be used as a gain control without disturbing the quiescent operating conditions. Give the maximum and minimum values of amplification, $A = e_o/(e_1 - e_1')$, that can be obtained by adjustment of R_4. Give the result in terms of the tube and circuit parameters.

CHAPTER 14

FREQUENCY DEPENDENCE OF SINGLE-STAGE AMPLIFIERS

Electronic circuits are considered in the preceding chapters only under conditions in which their behavior is independent of frequency in so far as the signal is concerned. However, frequency dependence enters at high frequencies where the parasitic capacitances associated with the tube or transistor and the circuit wiring become important, and it enters at low frequencies where bypass and coupling capacitors do not act as short circuits. An understanding of the basic principles of this dependence is an essential prerequisite for the effective design and utilization of most electronic circuits. This fact is especially important in connection with radar, television, and related fields where high-frequency operation and uniform performance over wide-frequency bands are required.

The primary objectives of this chapter are to present certain basic concepts and techniques that are quite generally applicable in frequency analyses and to develop the concept of the logarithmic amplitude and phase characteristics as a simple and effective means of presenting the results of frequency analysis. A secondary objective is to develop certain factual information about the properties of important vacuum-tube and transistor circuits; for this reason the basic notions are presented in terms of specific circuits.

The term *frequency analysis* as used here means a study of the response of a circuit to sinusoidal signals of various frequencies; hence sinusoidal operating conditions are to be understood in the sections that follow. The response of a circuit to periodic but nonsinusoidal signals can be deduced from the frequency characteristics of the circuit by considering separately each sinusoidal component in the Fourier series for the signal. Similarly, the response of a circuit to a nonperiodic signal can be deduced from the Fourier transform of the signal.

14-1. The Pentode Voltage Amplifier at High Frequencies. The circuit of a typical pentode voltage amplifier is shown in Fig. 14-1*a*. In the usual case the bypass capacitors C_0 and C_2 can be considered short circuits at frequencies greater than a few hundred cycles per second;

hence at frequencies above this limit the amplifier can be represented by the incremental model shown in Fig. 14-1*b*. This model is valid at frequencies up to several tens of megacycles per second. The capacitance C_1 accounts for the total capacitance between grid and ground, including the wiring capacitance as well as the tube capacitances, and C_4 accounts for the total capacitance between plate and ground. The grid-to-plate capacitance of the pentode is assumed to be negligibly small. The behavior of the amplifier at medium and high frequencies can be calculated from this model; the low-frequency behavior is treated in Sec. 14-5. This technique of treating the low and high frequencies separately reduces the amount of labor required in the analysis and, by simplifying the problem, permits a clearer understanding of the circuit behavior.

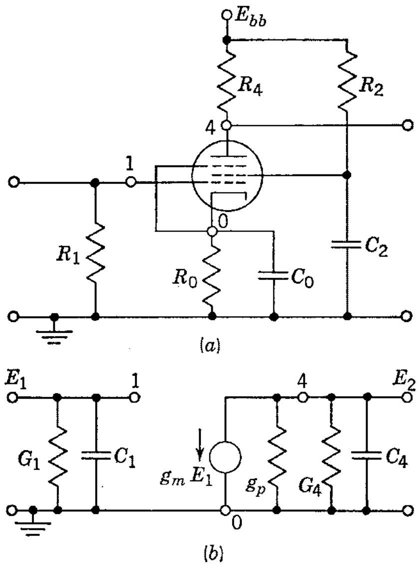


FIG. 14-1. The pentode amplifier at high frequencies. (*a*) Circuit; (*b*) high-frequency model.

The open-circuit input admittance Y_{no} of the pentode amplifier is the admittance of G_1 in parallel with C_1; the short-circuit output impedance Z_{os} is the impedance of g_p, G_4, and C_4 in parallel; and the short-circuit reverse current transmittance B_{cs} is zero. Since

$$B_{cs} = 0, \qquad Y_n = Y_{no}$$

for all values of load impedance, and $Z_o = Z_{os}$ for all values of source impedance. The forward voltage transmittance of the circuit is ordinarily of much greater interest than the other three coefficients; by inspection of the model it is

$$A_{vo}(\omega) = \frac{E_2}{E_1} = \frac{-g_m}{g_p + G_4 + j\omega C_4} \tag{14-1}$$

$$= \frac{-g_m}{g_p + G_4} \frac{1}{1 + j\omega \dfrac{C_4}{g_p + G_4}} \tag{14-2}$$

where E_1 and E_2 are the complex amplitudes of the sinusoidal input and output voltages. At medium frequencies ω becomes relatively small, and the second factor in (14-2) is approximately unity; hence the first factor in (14-2) is the voltage transmittance at medium frequencies. Furthermore, the quantity $(g_p + G_4)/C_4$ has the dimensions of frequency, and it has an important interpretation in terms of frequency. Hence it is desirable to define two new symbols:

$$A_m = \frac{g_m}{g_p + G_4} \qquad \text{and} \qquad \omega_4 = \frac{g_p + G_4}{C_4} \tag{14-3}$$

Substituting these expressions in (14-2) leads to a compact and convenient expression for the voltage transmittance:

$$A_{vo}(\omega) = -A_m \frac{1}{1 + j\omega/\omega_4} = \frac{-A_m}{\sqrt{1 + (\omega/\omega_4)^2}} \epsilon^{j\theta} = -|A_{vo}|\epsilon^{j\theta} \tag{14-4}$$

The magnitude of A_{vo} is the ratio of the magnitudes of E_1 and E_2; it is therefore the voltage amplification of the circuit. The angle θ is the phase angle between E_1 and E_2, except for the sign reversal associated with the minus sign; thus θ is the phase shift of the amplifier apart from the sign reversal. Both $|A_{vo}|$ and θ are functions of ω. The minus sign is carried explicitly in (14-4) in order to avoid adding a constant phase shift of 180° to the angle θ.

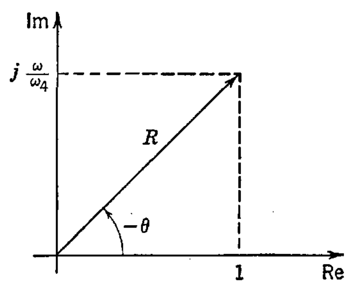


FIG. 14-2. Diagram for the complex number $1 + j\omega/\omega_4$. $R = \sqrt{1 + (\omega/\omega_4)^2}$ and $-\theta = \tan^{-1}(\omega/\omega_4)$.

The dependence of the amplification and the phase shift on frequency arises entirely from the factor $1 + j\omega/\omega_4$ in the denominator of the expression for A_{vo}. This factor is a complex number as well as a function of frequency; the way in which its magnitude and angle vary with frequency is illustrated by the simple geometric diagram of Fig. 14-2. Thus when $\omega = \omega_4$, $\theta = -45°$, $R = \sqrt{2}$, and $|A_{vo}| = A_m/\sqrt{2}$; it follows from these relations that ω_4 is the half-power frequency for the amplifier.

The manner in which A_{vo} varies with frequency can be displayed by calculating A_{vo} for a number of frequencies and plotting the frequency characteristics, $|A_{vo}|$ and phase shift versus frequency. A more useful and far simpler procedure, however, is to plot the voltage amplification in decibels,

$$A_{db}(\omega) = 20 \log |A_{vo}(\omega)| \tag{14-5}$$

and the phase shift as a function of log ω. These logarithmic characteristics possess certain especially simple asymptotic properties that make it possible to construct the complete frequency characteristics without plotting any points in the usual sense. These properties can be developed by substituting (14-4) in (14-5) to obtain

$$A_{\text{db}}(\omega) = 20 \log \frac{A_m}{\sqrt{1 + (\omega/\omega_4)^2}} \tag{14-6}$$

Defining $A_{\text{db}m} = 20 \log A_m$ gives

$$A_{\text{db}}(\omega) = A_{\text{db}m} - 20 \log \sqrt{1 + (\omega/\omega_4)^2} \tag{14-7}$$

$$= A_{\text{db}m} - 20 \log \sqrt{1 + (f/f_4)^2} \tag{14-8}$$

where $\omega = 2\pi f$.

The asymptotic behavior of A_{db} can be examined by considering very small and very large values of ω. Thus for $\omega \ll \omega_4$,

$$A_{\text{db}}(\omega) \approx A_{\text{db}m} - 20 \log 1 = A_{\text{db}m} \tag{14-9}$$

and the low-frequency asymptote of the amplitude characteristic is a constant, $A_{\text{db}m}$, as shown in Fig. 14-3. At high frequencies $\omega \gg \omega_4$, and

$$A_{\text{db}}(\omega) \approx A_{\text{db}m} - 20 \log \frac{\omega}{\omega_4} \tag{14-10}$$

Thus the high-frequency asymptote is a linear function of $\log(\omega/\omega_4)$; it is a straight line when plotted as a function of that variable or when plotted as a function of ω using a logarithmic scale. This asymptote is shown in Fig. 14-3.

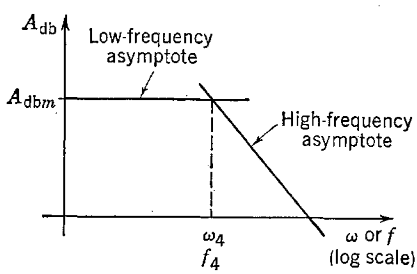


FIG. 14-3. Asymptotic behavior of $A_{\text{db}}(\omega)$.

The equations of the two asymptotes are (14-9) and (14-10); the intersection of the asymptotes occurs at the value of ω that gives the same value of A_{db} in these two expressions. Equating (14-9) and (14-10) gives

$$A_{\text{db}m} = A_{\text{db}m} - 20 \log \frac{\omega}{\omega_4}$$

The value of ω that satisfies this relation is $\omega = \omega_4$; hence the intersection, or break point, of the asymptotes occurs at the half-power frequency for the circuit. When $\omega = \omega_4$, the value of the high-frequency asymptote [Eq. (14-10)] is $A_{\text{db}} = A_{\text{db}m}$; when $\omega = 10\omega_4$, the value of the high-frequency asymptote is $A_{\text{db}} = A_{\text{db}m} - 20$; hence the slope of the high-frequency asymptote is -20 db/decade of frequency, where a decade of frequency is any interval on the frequency scale covering a 10:1 frequency ratio. Frequencies that stand in the ratio 2:1 are separated by one octave; it follows directly from (14-10) that the slope of the high-frequency asymptote is $-6.02 \approx -6$ db/octave.

The transition of the amplitude characteristic from one asymptote to the other is very simple in form. From Eq. (14-7), when $\omega = \omega_4$,

$$A_{db}(\omega) = A_{dbm} - 20 \log \sqrt{2} \approx A_{dbm} - 3 \tag{14-11}$$

Thus at the break frequency the amplitude characteristic is 3 db below the low-frequency asymptote. When $\omega = 2\omega_4$,

$$\begin{aligned} A_{db}(\omega) &= A_{dbm} - 20 \log \sqrt{1+4} \approx A_{dbm} - 10 \log 5 \\ &\approx A_{dbm} - 7 \end{aligned} \tag{14-12}$$

Therefore one octave above the break frequency the amplitude characteristic lies 7 db below the low-frequency asymptote. Since the high-frequency asymptote has a slope of -6 db/octave, the characteristic lies 1 db below it at $\omega = 2\omega_4$. And finally, when $\omega = \omega_4/2$,

$$\begin{aligned} A_{db}(\omega) &= A_{dbm} - 20 \log \sqrt{1 + \tfrac{1}{4}} = A_{dbm} - 10 \log (\tfrac{5}{4}) \\ &\approx A_{dbm} - 7 + 6 = A_{dbm} - 1 \end{aligned} \tag{14-13}$$

Thus one octave below the break frequency the amplitude characteristic lies 1 db below the low-frequency asymptote.

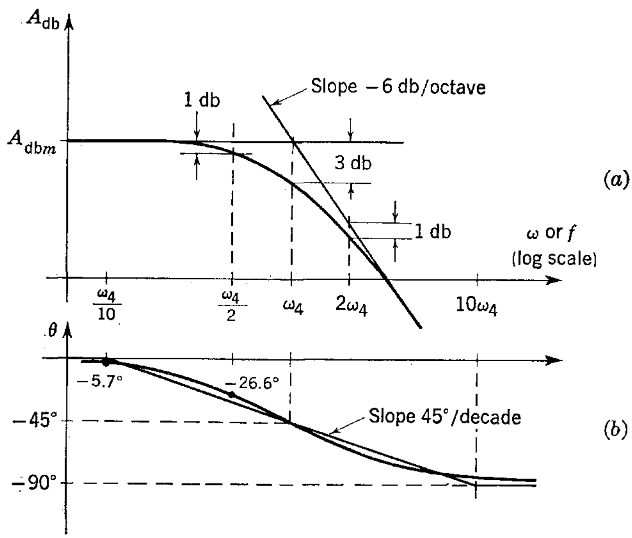


FIG. 14-4. Logarithmic amplitude and phase characteristics for the pentode amplifier at high frequencies. (*a*) Amplitude characteristic; (*b*) phase characteristic.

The amplitude characteristic and its relation to its asymptotes are shown in Fig. 14-4*a*. With the aid of these simple relations the amplitude characteristic can be constructed with negligible effort. For any particular amplifier it is necessary to calculate only the values of A_{dbm} and ω_4; the asymptotes for the amplitude characteristic can be constructed with these two numbers, and the characteristic can be constructed by applying the corrections shown in Fig. 14-4*a*. The corrections are the same for all amplitude characteristics of this form. It is also true that for many needs it is sufficient to construct only the asymptotes of the characteristic.

The phase-shift characteristic, shown in Fig. 14-4*b*, has equally simple properties. As indicated by the diagram in Fig. 14-2, the phase shift is

$$\theta(\omega) = -\tan^{-1}\frac{\omega}{\omega_4} \tag{14-14}$$

The low- and high-frequency asymptotes of the phase shift are 0 and $-90°$, respectively, and the phase shift at the break frequency, ω_4, is $-45°$. The phase characteristic is symmetrical about the point at $\omega = \omega_4$.

Figure 14-4*b* shows that the transition of the phase characteristic from the low- to the high-frequency asymptote can be approximated closely by a straight line beginning one decade below the break frequency at the low-frequency asymptote and ending one decade above the break frequency at the high-frequency asymptote. The exact phase shifts for $\omega = \omega_4/2$ and $\omega = \omega_4/10$ are shown in Fig. 14-4*b*; the phase shifts at $\omega = 2\omega_4$ and $\omega = 10\omega_4$ are, respectively, the complements of these angles. It is clear from Fig. 14-4 that the transition of the phase characteristic from one asymptote to the other occupies a much wider band of frequencies than the transition of the amplitude characteristic. In certain applications this fact is important and requires that careful attention be paid to the phase characteristic.

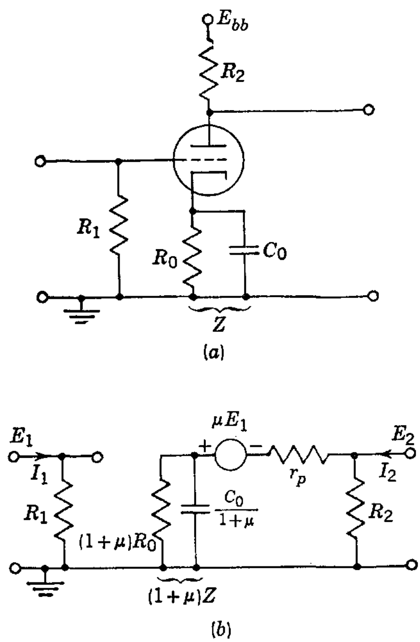


FIG. 14-5. The triode amplifier at low frequencies. (*a*) Circuit; (*b*) low-frequency model.

14-2. The Triode Voltage Amplifier at Low Frequencies. A triode voltage amplifier with a bypassed cathode resistor is shown in Fig. 14-5*a*. The incremental model shown in Fig. 14-5*b* represents the amplifier for increments of current and voltage in the band of frequencies where the parasitic capacitances are negligible; that is, from zero to a few tens of kilocycles per second. This model is obtained by using the reduction theorem in the manner illustrated in Fig. 13-10.

The reverse current transmittance of the model in Fig. 14-5*b*, B_{cs}, is zero; the input admittance Y_{no} is $1/R_1$; and the output impedance Z_{os} is the impedance seen at the output terminals with $\mu E_1 = 0$. The forward voltage transmittance, which is usually of much greater interest, is

$$A_{vo} = \frac{-\mu R_2}{r_p + R_2 + (1 + \mu)Z} = \frac{-\mu R_2}{r_p + R_2 + \dfrac{(1 + \mu)R_0}{1 + j\omega R_0 C_0}} \qquad (14\text{-}15)$$

This expression can be put in a form similar to Eq. (14-4) by an algebraic rearrangement. Such a rearrangement permits the amplitude and phase characteristics to be constructed by the simple methods developed in Sec. 14-1; hence it is well worth the small effort required.

Factoring $r_p + R_2$ out of the denominator of (14-15) and multiplying numerator and denominator by $1 + j\omega R_0 C_0$ yields

$$A_{vo} = \frac{-\mu R_2}{r_p + R_2} \; \frac{1 + j\omega R_0 C_0}{1 + \dfrac{(1 + \mu)R_0}{r_p + R_2} + j\omega R_0 C_0} \qquad (14\text{-}16)$$

In the medium-frequency range, ω becomes relatively large, and the second factor in (14-16) tends to unity; hence the first factor is the voltage transmittance at medium frequencies. Furthermore, the quantity $1/R_0C_0$ has the dimensions of frequencies. Therefore, in the interest of compactness, it is desirable to define the following new symbols:

$$A_m = \frac{\mu R_2}{r_p + R_2}, \qquad \omega_0 = \frac{1}{R_0 C_0}, \qquad \text{and} \qquad k_0 = \frac{(1 + \mu)R_0}{r_p + R_2} \qquad (14\text{-}17)$$

Substituting these relations in (14-16) and factoring $1 + k_0$ out of the denominator yields

$$A_{vo} = \frac{-A_m}{1 + k_0} \; \frac{1 + j\omega/\omega_0}{1 + j\omega/(1 + k_0)\omega_0} = -|A_{vo}|\epsilon^{j\theta} \qquad (14\text{-}18)$$

where $|A_{vo}|$ is the voltage amplification, and θ is the phase shift of the amplifier (excluding the sign reversal). The form of Eq. (14-18) is similar to that of Eq. (14-4).

Both the amplification and the phase shift of the amplifier depend on frequency; this dependence is associated with the factors $1 + j\omega/\omega_0$ and $1 + j\omega/(1 + k_0)\omega_0$, respectively, in the numerator and the denominator of Eq. (14-18). These factors, which are complex numbers, are diagramed in Fig. 14-6. The amplification varies in proportion to the ratio R_n/R_d; for small values of ω this ratio approaches unity, and for large ω it approaches $1 + k_0$. The phase shift of the amplifier is $\theta = \theta_n - \theta_d$; it approaches zero for both very small and very large ω. As in the case of the pentode amplifier of Sec. 14-1, however, the frequency characteristics of the circuit can be presented more clearly by means of the logarithmic amplitude and phase characteristics.

The voltage amplification in decibels is

$$A_{\text{db}} = 20 \log |A_{vo}| = 20 \log \frac{A_m}{1 + k_0} + 20 \log \sqrt{1 + (\omega/\omega_0)^2} - 20 \log \sqrt{1 + [\omega/(1 + k_0)\omega_0]^2} \quad (14\text{-}19)$$

The first term in (14-19) is a constant, and the third term has the same form as the second term in Eq. (14-7); hence the asymptotes of the third term plotted as a function of log ω (or log f) have the form shown by curve A in Fig. 14-7*a*. The low-frequency asymptote is constant at zero db, and the high-frequency asymptote crosses the low-frequency asymptote at $\omega = (1 + k_0)\omega_0$ with a slope of -6 db/octave. The second

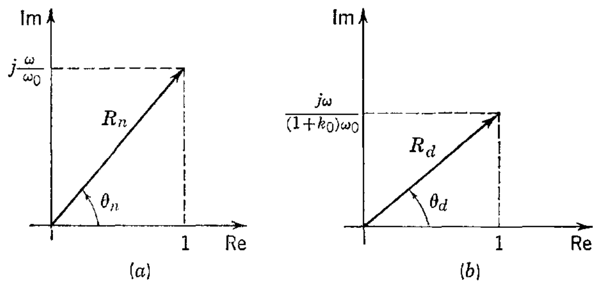


FIG. 14-6. Diagrams of the complex numbers in Eq. (14-18). (*a*) Numerator, $R_n = \sqrt{1 + (\omega/\omega_0)^2}$ and $\theta_n = \tan^{-1} (\omega/\omega_0)$; (*b*) denominator,

$$R_d = \sqrt{1 + [\omega/(1 + k_0)\omega_0]^2}$$

and $\theta_d = \tan^{-1} [\omega/(1 + k_0)\omega_0]$.

term in (14-19) has the same form as the third except for the minus sign; hence its asymptotes have the same form as those of the third term except that the slope of the high-frequency asymptote is $+6$ db/octave. The asymptotes for this term are shown by curve B in Fig. 14-7*a*. In the interest of clarity, only the asymptotes are shown in Fig. 14-7*a*; the amplitude characteristic for each term can be constructed by applying the corrections shown in Fig. 14-4*a*.

The complete amplitude characteristic is given by the sum of the characteristics of the three terms in Eq. (14-19). The asymptotic approximation to the complete characteristic is shown in Fig. 14-7*b*. The amplitude characteristic follows the general shape of the asymptotic approximation except for a rounding of the corners at the break frequencies; it crosses the approximate curve at a point midway between the break frequencies. The exact amplitude characteristic can be constructed either by constructing the exact characteristics in Fig. 14-7*a* and adding them, or by adding the appropriate corrections to the approximate curve in Fig. 14-7*b*.

The phase shift of the amplifier, apart from a sign reversal, is given by

$$\begin{aligned}\theta &= \theta_n - \theta_d \\ &= \tan^{-1}\frac{\omega}{\omega_0} - \tan^{-1}\frac{\omega}{(1 + k_0)\omega_0}\end{aligned} \qquad (14\text{-}20)$$

Each of these terms has the same form as Eq. (14-14) except for a difference of sign in the first term; hence each term has a characteristic like that in Fig. 14-4*b* except for the difference in sign. The construction of a straight-line approximation to the phase characteristic is illustrated in Fig. 14-8, where curves A and B correspond, respectively, to the second and first terms in (14-20). The exact phase characteristic is obtained by adding the appropriate corrections to this approximate curve.

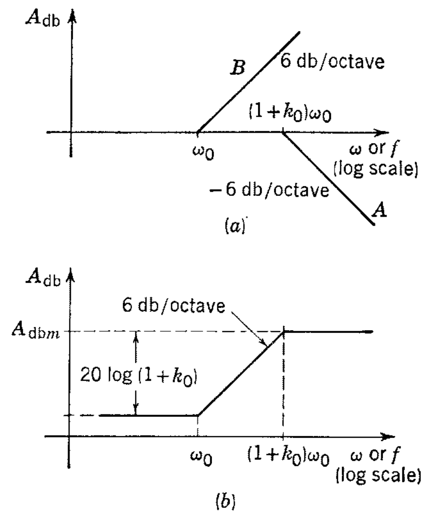


FIG. 14-7. Logarithmic amplitude characteristic for the triode with cathode degeneration. (*a*) Individual terms; (*b*) complete characteristic.

The foregoing analysis provides a basis for choosing the size of bypass capacitor to be used in the amplifier of Fig. 14-5*a*. In the usual case it is desired that the amplifier provide a large uniform amplification in the band of frequencies occupied by the signal; hence it should be designed so that the ramp in the amplitude characteristic of Fig. 14-7*b* occurs at frequencies below the band occupied by the signal. This

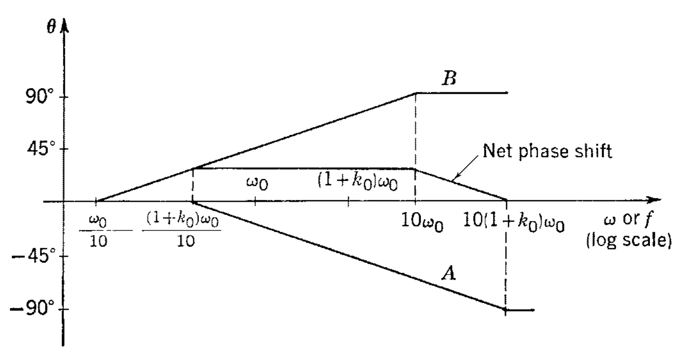


FIG. 14-8. Phase characteristic for the triode with cathode degeneration.

result is accomplished by choosing the circuit parameters so that the break at $\omega = (1 + k_0)\omega_0$ lies below the signal band.

If the tube, the supply voltage, and the quiescent point are specified, then R_2 and R_0 in the circuit of Fig. 14-5*a* must be chosen to give the required quiescent point. This choice fixes k_0 according to Eq. (14-17). The bypass capacitor must then be chosen to give a suitable value of ω_0 according to (14-17). Changes in C_0 affect only the value of ω_0; hence they shift the ramp in the amplitude characteristic parallel to the frequency axis without changing its shape.

Example 14-1. One section of a 12AU7 twin triode is used in the amplifier circuit shown in Fig. 14-5. The circuit parameters are $R_1 = 220$ kilohms, $R_2 = 47$ kilohms, $R_0 = 1.5$ kilohms, and $C_0 = 2.4$ μf. Construct the logarithmic amplitude characteristic for the amplifier in the low- and medium-frequency range.

Solution. A graphical analysis on the plate characteristics shows that the quiescent operating point is at approximately $E_{co} = -5$ volts and $I_{bo} = 3.3$ ma. The incremental tube parameters at this point are $\mu = 17$, $r_p = 11$ kilohms.

With these tube and circuit parameters, the amplification at medium frequencies is

$$A_m = \frac{\mu R_2}{r_p + R_2} = \frac{(17)(47)}{58} = 13.8$$

which in decibels is

$$A_{\text{db}m} = 20 \log 13.8 = (20)(1.14) = 22.8$$

The break frequency for the numerator of $A_{vo}(\omega)$ is

$$\omega_0 = \frac{1}{R_0 C_0} = \frac{10^6}{(1.5)(10^3)(2.4)} = 278 \text{ rps}$$

$$f_0 = \frac{\omega_0}{2\pi} = 44.3 \text{ cps}$$

The factor k_0 is

$$k_0 = \frac{(1 + \mu)R_0}{r_p + R_2} = \frac{(18)(1.5)}{58} = 0.465$$

and the break frequency for the denominator of $A_{vo}(\omega)$ is

$$(1 + k_0)\omega_0 = (1.465)(278) = 407 \text{ rps}$$
$$(1 + k_0)f_0 = (1.465)(44.3) = 65 \text{ cps}$$

The amount by which the amplification at low frequencies is less than its value at medium frequencies is

$$20 \log (1 + k_0) = (20)(0.166) = 3.32 \text{ db}$$

The logarithmic amplitude characteristic and its asymptotes are shown in Fig. 14-9. Note the expanded amplification scale.

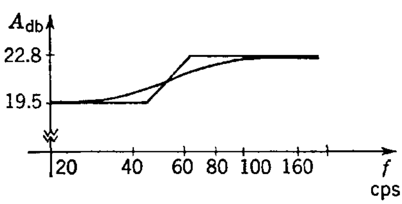


Fig. 14-9. Logarithmic amplitude characteristic for the amplifier of Example 14-1.

14-3. The Triode Voltage Amplifier at High Frequencies. An incremental model for the triode amplifier of Fig. 14-5 that applies in the range of high frequencies where the cathode bypass capacitor is a short circuit is shown in Fig. 14-10. The capacitor C_1 accounts for the total capacitance between grid and ground, including the interelectrode capacitance of the tube and the stray capacitance associated with the circuit wiring; C_2 accounts for the capacitance between plate and ground, and C_{12} accounts for the capacitance between the grid and plate circuits. Since these capacitances are usually a few micromicrofarads, the stray wiring capacitance may have an important effect unless the physical components of the circuit are arranged with care.

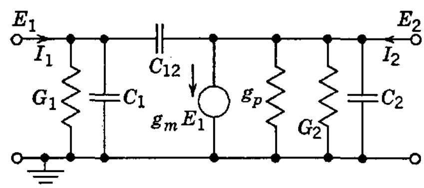


FIG. 14-10. The triode amplifier at high frequencies.

The forward voltage transmittance of the circuit in Fig. 14-10 can be obtained from a single node equation for node E_2. With $I_2 = 0$,

$$(g_p + G_2 + j\omega C_2)E_2 + j\omega C_{12}(E_2 - E_1) = -g_m E_1 \tag{14-21}$$

and

$$\left.\frac{E_2}{E_1}\right|_{I_2=0} = A_{vo} = \frac{-g_m + j\omega C_{12}}{g_p + G_2 + j\omega(C_{12} + C_2)} \tag{14-22}$$

This equation can be rearranged to obtain

$$A_{vo} = \frac{-g_m}{g_p + G_2}\,\frac{1 - j\omega C_{12}/g_m}{1 + j\omega(C_{12} + C_2)/(g_p + G_2)} \tag{14-23}$$

Equation (14-23) can be put in a compact form similar to (14-18) by defining new symbols as follows:

$$A_m = \frac{g_m}{g_p + G_2}; \qquad \omega_{12} = \frac{g_m}{C_{12}}; \qquad \omega_2 = \frac{g_p + G_2}{C_{12} + C_2} \tag{14-24}$$

The quantity A_m is again the voltage transmittance at medium frequencies. Substituting these relations in (14-23) yields

$$A_{vo} = -A_m\,\frac{1 - j\omega/\omega_{12}}{1 + j\omega/\omega_2} = -|A_{vo}|e^{j\theta} \tag{14-25}$$

Equation (14-25) has the same form as Eq. (14-18) except for the sign in the numerator; diagrams for the numerator and denominator factors are shown in Fig. 14-11. The amplification varies with frequency in proportion to the ratio R_n/R_d; at very low frequencies this ratio approaches unity, and at very high frequencies it approaches ω_2/ω_{12}. The phase shift is $\theta = \theta_n - \theta_d$.

The logarithmic amplitude and phase characteristics associated with Eq. (14-25) are constructed in the same manner as the characteristics of Figs. 14-7 and 14-8. In the usual case the break in the amplitude characteristic at ω_{12} lies one or two decades above the break at ω_2; hence the amplitude characteristic for a typical triode amplifier at high fre-

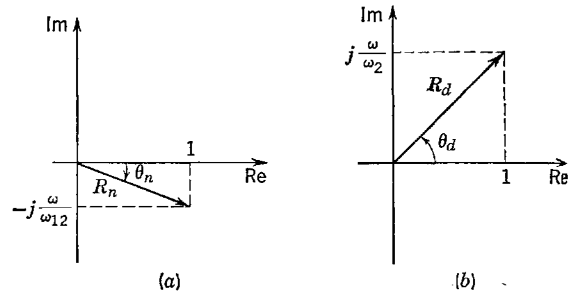


FIG. 14-11. Diagrams for the complex numbers in Eq. (14-25). (a) Numerator; (b) denominator.

quencies has the asymptotic form shown in Fig. 14-12*a*. The break frequency ω_{12}, which usually corresponds to a frequency of the order of 100 megacycles/sec, lies well above the usual frequency range for the amplifier; indeed, at such frequencies the model of Fig. 14-10 is likely to be a poor representation for the circuit because of additional parasitic effects not accounted for. For these reasons the break at ω_{12} is often ignored, and the numerator factor in (14-25) is taken as unity. Such an approximation is valid for all values of ω that are much smaller than ω_{12}.

The minus sign in the numerator of (14-25) affects only the phase characteristic. The phase shift is

$$\theta = \theta_n - \theta_d = \tan^{-1}\frac{-\omega}{\omega_{12}} - \tan^{-1}\frac{\omega}{\omega_2} \tag{14-26}$$

The straight-line approximation to this phase characteristic is shown in Fig. 14-12*b* for a typical triode amplifier.

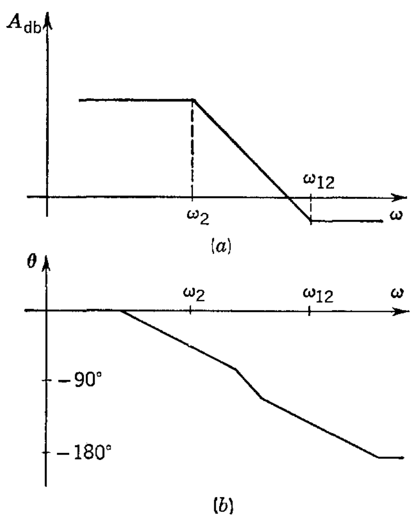


FIG. 14-12. Logarithmic amplitude and phase characteristics for the triode amplifier at high frequency. (a) Amplitude; (b) phase.

The input admittance of the triode amplifier at high frequencies is important, for it places a serious limitation on the performance of the circuit. In addition, the way in which the amplifying action of the tube increases the input admittance at high frequencies is of interest. The input admittance is evaluated by summing the currents at node E_1 with $I_2 = 0$; the result is

$$I_1 = (G_1 + j\omega C_1)E_1 + j\omega C_{12}(E_1 - E_2) \tag{14-27}$$

But $E_2 = A_{vo}E_1$

Hence $I_1 = (G_1 + j\omega C_1)E_1 + j\omega C_{12}(1 - A_{vo})E_1$ (14-28)

Thus if the voltage amplification is much larger than unity, the voltage across C_{12} and the current through it are relatively large. The input admittance is obtained by dividing (14-28) by E_1; the result is

$$Y_{no} = G_1 + j\omega C_1 + j\omega C_{12}(1 - A_{vo}) \tag{14-29}$$

In general, A_{vo} is a function of frequency given by Eq. (14-25); in the useful frequency range of the amplifier, however, $\omega/\omega_{12} \ll 1$, and (14-25) reduces to

$$A_{vo} = -A_m \frac{1}{1 + j\omega/\omega_2} \tag{14-30}$$

Substituting this expression in (14-29) yields

$$Y_{no} = G_1 + j\omega C_1 + j\omega C_{12} + \frac{j\omega C_{12}A_m}{1 + j\omega/\omega_2} \tag{14-31}$$

The last term in (14-31) is an admittance corresponding to an impedance

$$Z = \frac{1 + j\omega/\omega_2}{j\omega C_{12}A_m} = \frac{1}{\omega_2 C_{12}A_m} + \frac{1}{j\omega C_{12}A_m} \tag{14-32}$$

$$= R + \frac{1}{j\omega C} \tag{14-33}$$

Thus Z is the impedance of the series connection of a resistance $R = 1/\omega_2 C_{12}A_m$ and a capacitance $C = A_m C_{12}$. Substituting the values of ω_2 and A_m given by (14-24) into the expression for R and simplifying yields

$$R = \frac{1 + C_2/C_{12}}{g_m} \tag{14-34}$$

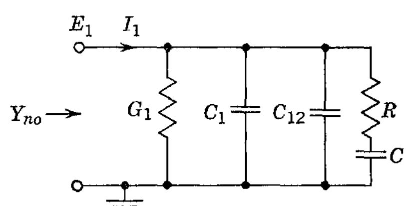


FIG. 14-13. Equivalent input circuit for the triode at high frequencies. $R = 1/\omega_2 C_{12}A_m = (1 + C_2/C_{12})/g_m$ and $C = A_m C_{12}$.

It follows from these facts and from (14-31) that the input admittance to the amplifier is the admittance of the circuit shown in Fig. 14-13; this is an equivalent circuit for the amplifier of Fig. 14-10 in so far as the input terminals are concerned.

For typical triode amplifiers, the value of R in the circuit of Fig. 14-13 is of the order of 1000 ohms; it is much smaller than the reactance of C in the useful frequency range of the amplifier. It is common practice, therefore, to neglect this resistance, leaving only the capacitance $C = A_m C_{12}$ in the right-hand branch of Fig. 14-13. It follows from this simplified representation that, in so far as the input terminals are concerned, the grid-to-plate capacitance C_{12} acts as if it were multiplied by the factor $1 + A_m$ and connected beween grid and ground. This result should be compared with the similar result obtained for the ideal voltage amplifier shown in Fig. 4-6. The action by which the amplifying property of the tube increases the apparent value of the grid-to-plate capacitance is known as the Miller effect. As a result of the Miller effect, the triode may present an excessive load to the source of signals at high frequencies.

Example 14-2. One unit of a 6SL7 twin triode is used in the amplifier of Fig. 14-10. The circuit parameters are $R_1 = 1/G_1 = 330$ kilohms, $R_2 = 100$ kilohms, $g_m = 1.6$ millimhos, $g_p = 0.023$ millimho, $C_1 = C_2 = 10$ $\mu\mu$f, and $C_{12} = 5$ $\mu\mu$f. These parasitic capacitances include both wiring and interelectrode capacitances. The signal source can be represented by a voltage source E_s in series with a resistance R_s of 100 kilohms.

With the amplitude of E_s held constant, the input voltage to the amplifier E_1 decreases at high frequencies as a result of loading by the input capacitance of the amplifier, and the output voltage from the amplifier decreases accordingly. The problem is to examine this effect by constructing the logarithmic amplitude characteristic of the voltage ratio E_1/E_s.

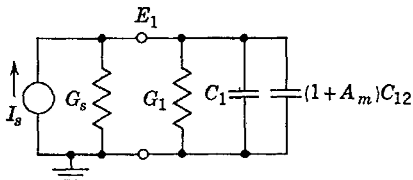


FIG. 14-14. Equivalent input circuit for the amplifier of Example 14-2. $G_s = 1/R_s = 0.01$ millimho and $I_s = 0.01E_s$ milliamperes when E_s is in volts.

Solution. The analysis is simplified slightly by converting the specified signal source to an equivalent current source in parallel with the source resistance R_s. Thus, neglecting the small resistance R in Fig. 14-13, the input circuit can be represented by the equivalent circuit shown in Fig. 14-14.

The constant A_m is, from Eqs. (14-24),

$$A_m = \frac{g_m}{g_p + G_2} = \frac{1.6}{0.023 + 0.01} = 48.5$$

Hence the total shunt capacitance in Fig. 14-14 is

$$C_o = C_1 + 49.5C_{12} = 10 + (49.5)(5) = 257.5\ \mu\mu\text{f}$$

The total shunt conductance is

$$G_o = G_s + G_1 = 0.01 + 0.00303 = 0.013 \text{ millimho}$$

The input voltage to the amplifier is

$$E_1 = \frac{I_s}{G_o + j\omega C_o} = \frac{0.01E_s}{G_o + j\omega C_o}$$

and the desired voltage ratio is

$$\frac{E_1}{E_s} = \frac{0.01}{G_o}\frac{1}{1 + j\omega C_o/G_o} = 0.77\frac{1}{1 + j\omega/\omega_o}$$

where $\omega_o = G_o/C_o = 50{,}500$ rps.

The expression for the ratio E_1/E_s has the same form as the expression for the forward voltage transmittance of the pentode amplifier at high frequencies; hence its logarithmic amplitude characteristic has the form shown in Fig. 14-4*a*. The amplitude at medium and low frequencies is

$$\begin{aligned} A_{\text{db}m} &= 20 \log 0.77 = -20 \log (1/0.77) = -(20)(0.114) \\ &= -2.28 \text{ db} \end{aligned}$$

The break frequency is $\omega_o = 50{,}500$ rps, corresponding to $f_o = \omega_o/2\pi = 8050$ cps. Thus, with a constant signal E_s, the input voltage to the amplifier E_1, and hence the output voltage E_2, decreases appreciably at frequencies greater then about 8 kcps. This deterioration of performance results from the action of the shunt capacitance C_o. It is significant to note that C_o is made up almost entirely of the Miller capacitance, $(1 + A_m)C_{12}$.

The equivalent circuit given in Fig. 14-13 is valid only when the forward voltage transmittance has the particular form of Eq. (14-30); thus it is not valid when the plate circuit of the amplifier contains additional reactive elements, for these elements alter the form of the expression for A_{vo}. An example of interest that reveals another property of the triode amplifier is obtained by replacing the plate load resistance in the circuit of Fig. 14-10 by an inductance L. In this case the forward voltage transmittance is given by

$$A_{vo} = \frac{-g_m + j\omega C_{12}}{g_p + j\omega(C_{12} + C_2) + 1/j\omega L} \tag{14-35}$$

Neglecting ωC_{12} in comparison with g_m as in Eq. (14-30) yields

$$A_{vo} = \frac{-g_m}{g_p + j\omega(C_{12} + C_2) + 1/j\omega L} \tag{14-36}$$

Substituting this expression in (14-29) yields for the input admittance

$$Y_{no} = G_1 + j\omega C_1 + j\omega C_{12} + \frac{j\omega C_{12} g_m}{g_p + j\omega(C_{12} + C_2) + 1/j\omega L} \tag{14-37}$$

The last term in (14-37) is an admittance corresponding to an impedance

$$Z = \frac{1 + C_2/C_{12}}{g_m} - \frac{1}{w^2 g_m C_{12} L} + \frac{g_p}{j\omega C_{12} g_m} \tag{14-38}$$

If the resonant frequency of L and C_{12} connected in parallel is designated ω_r, then $C_{12}L = 1/\omega_r^2$; using this fact, and noting that $\mu = g_m/g_p$, the expression for Z can be written as

$$Z = \frac{1 + C_2/C_{12}}{g_m} - \frac{(\omega_r/\omega)^2}{g_m} + \frac{1}{j\omega\mu C_{12}} \tag{14-39}$$

$$= R - R'(\omega) + \frac{1}{j\omega C} \tag{14-40}$$

Thus Z is the impedance of the series connection of a resistance R, a resistive impedance $-R'$ that is negative and a function of frequency, and a capacitance C. It follows from these facts and from (14-37) that the input admittance to the amplifier with an inductive plate load is the admittance of the circuit shown in Fig. 14-15.

The negative resistance that appears in Fig. 14-15 is of special significance. A positive resistance absorbs electrical energy and converts it into heat; a negative resistance acts as a source of electrical energy. The occurrence of a negative resistance in the circuit of Fig. 14-15 is a result of the action of the controlled source associated with the tube. If the negative resistance is sufficiently large, but not too large, it will cause growing transients to occur in the amplifier. When the triode is used with tuned circuits at radio frequencies, these growing transients take the form of self-sustaining oscillations. The mechanism by which these oscillations are generated is examined in greater detail in Chaps. 16 and 17. Similar conditions are encountered with transistors when they are used with tuned circuits at high frequencies.

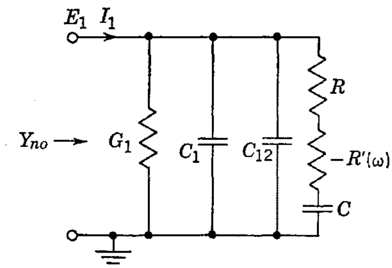


FIG. 14-15. Equivalent input circuit for a triode with an inductive plate load. $R = (1 + C_2/C_{12})/g_m$, $-R'(\omega) = -(\omega_r/\omega)^2/g_m$, $C = \mu C_{12}$, and $\omega_r^2 = 1/LC_{12}$.

If C_{12} is zero, the branch containing the negative resistance in Fig. 14-15 becomes an open circuit, and the possibility of growing transients and self-sustaining oscillations disappears. It is for this reason that the pentode was developed; the pentode is generally preferred over the triode for high-frequency applications. It should be noted, however, that careless wiring can introduce enough stray capacitance between the grid and plate circuits to cause oscillations even in pentode amplifiers.

14-4. The Cathode Follower at High Frequencies. Figure 14-16*a* shows a triode in a practical cathode-follower circuit (Example 13-1). Of particular interest in this section is the fact that the action of the tube in the cathode follower is such as to reduce the input admittance of the circuit. Hence the cathode follower can be designed to have a small input admittance over a relatively wide band of frequencies. The circuit of Fig. 14-16*b* is an incremental model for the cathode follower that accounts for the effects of parasitic capacitances at high frequencies; it is a valid representation at frequencies up to several tens of megacycles per second. The capacitor C_1 accounts for the total capacitance between grid and ground, including the grid-to-plate capacitance of the tube; C_2 accounts for the total capacitance between cathode and ground, including the cathode-to-heater capacitance; and C_{12} is the total capacitance between grid and cathode.

In the usual cathode follower the current in R_1 is very small and can be neglected in comparison with the current in R_2 and R_3. The forward voltage transmittance can therefore be evaluated from a single node equation by treating R_2 and R_3 as a single resistance, $R_2' = R_2 + R_3 = 1/G_2'$.

With $I_2 = 0$, this node equation is

$$(g_m + g_p + G_2' + j\omega C_2)E_2 + j\omega C_{12}(E_2 - E_1) = g_m E_1 \qquad (14\text{-}41)$$

Collecting terms and solving for E_2/E_1 yields

$$A_{vo} = \frac{g_m + j\omega C_{12}}{g_m + g_p + G_2' + j\omega(C_{12} + C_2)} = A_m \frac{1 + j\omega/\omega_{12}}{1 + j\omega/\omega_2} \qquad (14\text{-}42)$$

where

$$A_m = \frac{g_m}{g_m + g_p + G_2'}; \qquad \omega_{12} = \frac{g_m}{C_{12}}; \qquad \omega_2 = \frac{g_m + g_p + G_2'}{C_{12} + C_2} \qquad (14\text{-}43)$$

It is clear from Eqs. (14-43) that the midband voltage transmittance A_m must always be smaller than unity. However, in a well-designed cathode follower g_m is much greater than $g_p + G_2'$, and A_m approaches unity.

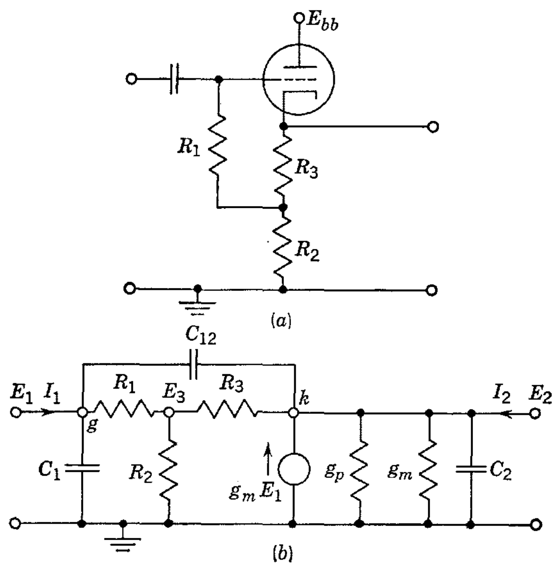


FIG. 14-16. The cathode follower at high frequencies. (a) Circuit; (b) model for medium and high frequencies.

Also, in a well-designed cathode follower ω_{12} is greater than ω_2; hence the voltage transmittance decreases at high frequencies toward the limiting value of $A_m\omega_2/\omega_{12}$.

Equation (14-42) has the same form as (14-25) except for a minus sign in the numerator; therefore the amplitude characteristic for the cathode follower has the same form as the characteristic shown in Fig. 14-12*a* for the triode amplifier. The cathode-follower characteristic lies entirely below the zero-db axis, however, for the voltage amplification is less than unity at all frequencies. In the case of the cathode follower, the break frequencies are usually several tens of megacycles per second. The phase characteristics for the cathode follower is different from that shown in Fig. 14-12*b* for the triode amplifier because of the difference in sign between Eqs. (14-42) and (14-25); however, the phase characteristic can be constructed easily by adding the separate characteristics for the numerator and the denominator of (14-42).

The tube in the cathode follower acts in the following way to reduce the input admittance. A positive increment of voltage applied at E_1 in Fig. 14-16 causes slightly smaller positive increments of voltage at E_2 and E_3. The corresponding increments of voltage across R_1 and C_{12} are very small, and as a result the input current I_1 is relatively small. The input admittance can be evaluated in terms of the circuit parameters by summing the currents at node E_1. With $I_2 = 0$,

$$I_1 = j\omega C_1 E_1 + j\omega C_{12}(E_1 - E_2) + G_1(E_1 - E_3) \qquad (14\text{-}44)$$

But

$$E_2 = A_{vo} E_1 \qquad (14\text{-}45)$$

and, neglecting the current in R_1 in comparison with that in R_2,

$$E_3 = \frac{R_2}{R_2 + R_3} E_2 = \frac{R_2 A_{vo}}{R_2 + R_3} E_1 = A'_{vo} E_1 \qquad (14\text{-}46)$$

Substituting these two expressions in (14-44) and dividing by E_1 yields

$$Y_{no} = j\omega C_1 + j\omega C_{12}(1 - A_{vo}) + G_1(1 - A'_{vo}) \qquad (14\text{-}47)$$

In the middle band of frequencies both A_{vo} and A'_{vo} are positive real numbers that are slightly smaller than unity; hence the second and third terms in (14-47) are quite small. If Eq. (14-42) is substituted for A_{vo} in (14-47), it is seen that the second term in (14-47) contributes a negative component of conductance to the input admittance; if this component is large enough, it may be a source of self-sustaining oscillations.

The output impedance of the cathode follower can be evaluated in terms of the circuit parameters by inspection of the model in Fig. 14-16*b*. The current in R_1 can be neglected; hence, with $E_1 = 0$,

$$Z_{os} = \frac{1}{g_m + g_p + G'_2 + j\omega(C_{12} + C_2)} \qquad (14\text{-}48)$$

where $G'_2 = 1/(R_2 + R_3)$. In the middle band of frequencies the capacitive term is negligible, and ordinarily g_m is much larger than $g_p + G'_2$; hence in the mid-band, $Z_{os} \approx 1/g_m$. Typical values of Z_{os} lie in the range between 100 and 1000 ohms.

14-5. The Pentode Voltage Amplifier at Low Frequencies. The behavior of the pentode amplifier shown in Fig. 14-1 is affected in the low-frequency range by both the cathode-bypass and the screen-grid-bypass capacitors; the effect of each of these capacitors is much like the effect of the cathode-bypass capacitor in the triode amplifier (Fig. 14-7). The amplifier can be analyzed in the low-frequency range with the aid of an incremental model for the tube that accounts for the action of the screen grid. The complete model shown in Fig. 10-17 can be used for this purpose; however, it leads to results that are algebraically cumbersome and that therefore provide little insight into the behavior of the circuit. Moreover, these complexities result largely from the presence of several factors that have a negligible effect on the over-all behavior of the circuit. It is therefore appropriate to make certain strategic simplifications in the circuit to eliminate the superfluous factors before beginning the analysis. When this is done it is possible to make a relatively simple analysis that shows clearly the way in which each parameter affects the behavior of the circuit and that gives results checking quite well with experimental observations. It is, of course, not obvious at the outset which simplifications are both profitable and permissible; this knowledge is the result of preliminary studies of the problem.

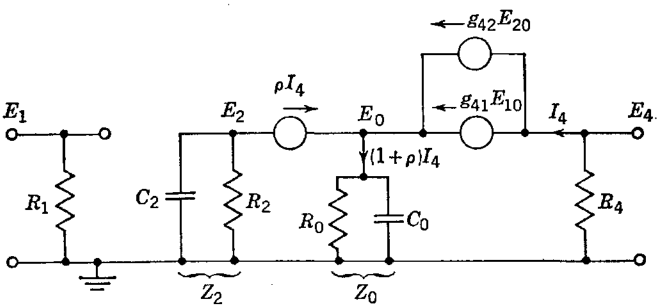


FIG. 14-17. The pentode amplifier at low frequencies. $E_{10} = E_1 - E_0, E_{20} = E_2 - E_0$, and $r_{44} = \infty$.

The simplifications referred to above are embodied in the low-frequency model for the pentode amplifier shown in Fig. 14-17. This model is based on the approximate pentode model of Fig. 10-18, and it includes the further assumption that the plate resistance is very large and can be treated as an open circuit. The voltages of the control and screen grids relative to the cathode are, respectively, $E_{10} = E_1 - E_0$ and $E_{20} = E_2 - E_0$. The impedances in the cathode and screen-grid circuits are

$$Z_0 = \frac{R_0}{1 + j\omega C_0 R_0} = \frac{R_0}{1 + j\omega/\omega_0} \tag{14-49}$$

and

$$Z_2 = \frac{R_2}{1 + j\omega C_2 R_2} = \frac{R_2}{1 + j\omega/\omega_2} \tag{14-50}$$

The forward voltage transmittance of the pentode amplifier at low frequencies can be evaluated from the circuit in Fig. 14-17. With the output terminals open-circuited, the output voltage is

$$E_4 = -R_4 I_4 \tag{14-51}$$

But

$$\begin{aligned} I_4 &= g_{41}E_{10} + g_{42}E_{20} \\ &= g_{41}E_1 + g_{42}E_2 - (g_{41} + g_{42})E_0 \end{aligned} \tag{14-52}$$

and the cathode and screen-grid voltages are given by

$$E_0 = Z_0(1 + \rho)I_4 \qquad \text{and} \qquad E_2 = -Z_2\,\rho I_4 \tag{14-53}$$

Substituting (14-53) in (14-52), solving for I_4 and substituting the result in (14-51) yields

$$A_{vo} = \frac{-g_{41}R_4}{1 + \rho g_{42}Z_2 + (1 + \rho)(g_{41} + g_{42})Z_0} \tag{14-54}$$

But from Eq. (10-5), which is the basis for the pentode model used in Fig. 14-17, the total plate and screen-grid currents are related by

$$i_{c2} = \rho i_b \tag{14-55}$$

Differentiating this equation partially with respect to the screen-grid potential e_2 yields

$$\frac{\partial i_{c2}}{\partial e_2} = \rho \frac{\partial i_b}{\partial e_2} \qquad \text{or} \qquad g_{22} = \rho g_{42} \tag{14-56}$$

Equation (14-54) can now be put in a compact and useful form with the aid of (14-56) and two additional definitions:

$$A_m = g_{41}R_4$$

and

$$g_0 = (1 + \rho)(g_{41} + g_{42}) = (1 + \rho)g_{41} + \left(1 + \frac{1}{\rho}\right) g_{22} \tag{14-57}$$

Since Z_0 and Z_2 are zero in the middle band of frequencies, $A_m = g_{41}R_4$ is the voltage transmittance in the middle band. Substituting (14-56) and (14-57) in (14-54) gives

$$A_{vo} = \frac{-A_m}{1 + g_{22}Z_2 + g_0Z_0} \tag{14-58}$$

The behavior of A_{vo} as a function of ω is examined for the special cases of cathode degeneration only and screen-grid degeneration only before considering the general case. Therefore, suppose first that $Z_2 = 0$ at all frequencies of interest; then, substituting (14-49) in (14-58),

$$A_{vo} = \frac{-A_m}{1 + \dfrac{g_0R_0}{1 + j\omega/\omega_0}} = -A_m \frac{(1 + j\omega/\omega_0)}{1 + g_0R_0 + j\omega/\omega_0} \tag{14-59}$$

Defining $k_0 = g_0R_0$ leads to

$$A_{vo} = \frac{-A_m}{1 + k_0} \frac{1 + j\omega/\omega_0}{1 + j\omega/(1 + k_0)\omega_0} \tag{14-60}$$

This equation has exactly the same form as Eq. (14-18), describing cathode degeneration in the triode amplifier; hence the amplitude and phase characteristics associated with (14-60) have the same form as those shown in Figs. 14-7 and 14-8.

Now suppose that $Z_0 = 0$ at all frequencies of interest and that $Z_2 \neq 0$; then substituting (14-50) in (14-58) and defining $k_2 = g_{22}R_2$ yields

$$A_{vo} = \frac{-A_m}{1 + k_2} \frac{1 + j\omega/\omega_2}{1 + j\omega/(1 + k_2)\omega_2} \tag{14-61}$$

Thus screen-grid degeneration has the same effect on the voltage transmittance as cathode degeneration, and the amplitude and phase characteristics in this case are similar to those resulting from cathode degeneration.

In the general case when neither Z_0 nor Z_2 is zero,

$$A_{vo} = \frac{-A_m}{1 + \dfrac{g_{22}R_2}{1 + j\omega/\omega_2} + \dfrac{g_0R_0}{1 + j\omega/\omega_0}} \tag{14-62}$$

Substituting $k_0 = g_0R_0$ and $k_2 = g_{22}R_2$, clearing fractions in the denominator, and collecting terms in like powers of $j\omega$ in the denominator yields

$$A_{vo} = -A_m \frac{\omega_0\omega_2(1 + j\omega/\omega_0)(1 + j\omega/\omega_2)}{(j\omega)^2 + [(1 + k_0)\omega_0 + (1 + k_2)\omega_2](j\omega) + (1 + k_0 + k_2)\omega_0\omega_2} \tag{14-63}$$

The two factors in parentheses in the numerator are in a familiar form, and their amplitude and phase characteristics can be constructed by the simple, rapid methods developed in the preceding sections. The denominator, on the other hand, is a quadratic rather than a linear function of frequency, and the simple techniques that are applicable to linear functions cannot be used directly. However, any polynomial, such as the quadratic denominator in (14-63), can be factored and expressed as the product of a set of linear factors like those in the numerator of (14-63). Since this operation makes it possible to construct the complete amplitude and phase characteristics associated with A_{vo} by the simple methods already developed, it is well worth the effort.

The quadratic expression in (14-63) can be factored with the aid of the quadratic formula. If the coefficients of $j\omega$ are left in literal form, however, the factors are complicated irrational functions of the circuit parameters. Therefore, in the interest of simplicity the procedure is illustrated by giving typical values to the coefficients. The parameter values in a certain pentode amplifier of conventional design yield the following numerical results: $A_m = 100$, $k_0 = 1.3$, $k_2 = 15$, $\omega_0 = 200$ rps, and $\omega_2 = 30$ rps. Substituting these values in (14-63) yields

$$A_{vo} = -(6)(10^5) \frac{(1 + j\omega/200)(1 + j\omega/30)}{(j\omega)^2 + 940(j\omega) + 103{,}800} \tag{14-64}$$

If the denominator of (14-64) is factored with the variable expressed as $j\omega$, the roots obtained are of the form $j\omega = -150$, for example, and in other circumstances roots of the form $j\omega = -2 + j\sqrt{3}$ occur. These results are perfectly proper, but they look strange. In order to avoid these unfamiliar forms, and in order to provide a formulation that is consistent with common practice in other phases of circuit theory, it is desirable to substitute the symbol m for $j\omega$ in (14-64). The resulting expression is

$$A_{vo}(m) = -(6)(10^5)\frac{(1 + m/200)(1 + m/30)}{m^2 + 940m + 103{,}800} \tag{14-65}$$

The voltage transmittance for steady-state sinusoidal operation at any radian frequency ω is obtained from (14-65) by replacing m with $j\omega$.

The roots of the equation $m^2 + 940m + 103{,}800 = 0$, which are the values of m that make the denominator of (14-65) zero, are found from the quadratic formula to be

$$m_1 = -128 \qquad \text{and} \qquad m_2 = -812 \tag{14-66}$$

Therefore (14-65) can be written as

$$A_{vo}(m) = -(6)(10^5)\frac{(1 + m/200)(1 + m/30)}{(m + 128)(m + 812)} \tag{14-67}$$

This relation holds for all values of m; therefore it holds for $m = j\omega$. Hence if the sinusoidal steady-state voltage transmittance is desired, it can be obtained from

$$A_{vo}(j\omega) = -(6)(10^5)\frac{(1 + j\omega/200)(1 + j\omega/30)}{(j\omega + 128)(j\omega + 812)} \tag{14-68}$$

Factoring 128 and 812 out of the denominator yields

$$A_{vo}(j\omega) = -5.77\frac{(1 + j\omega/200)(1 + j\omega/30)}{(1 + j\omega/128)(1 + j\omega/812)} \tag{14-69}$$

This expression consists entirely of factors having a familiar form. The amplitude and phase characteristics for each factor is constructed in the manner developed in Sec. 14-1, and the individual characteristics are added to obtain the characteristics for $A_{vo}(j\omega)$. The asymptotes for the resulting amplitude characteristic are shown in Fig. 14-18.

The calculations required in the construction of the gain characteristic in Fig. 14-18 are entirely straightforward with one exception, namely, that the value of the incremental screen-grid conductance, g_{22}, must be known. If this parameter is not given by the tube manufacturer, it must be determined or approximated by one means or another, perhaps by experimental measurement. In addition, the value of this parameter can be expected to vary considerably from one tube to another of a given type. The values of g_{22} for typical small pentodes lie in the range between 0.1 and 0.01 millimho; that is, $r_{22} = 1/g_{22}$ is usually a few tens of kilohms.

The value of g_{22} can be determined empirically by direct measurement. For this measurement, the tube is biased at the appropriate quiescent operating point, a small increment of voltage is applied between the screen grid and the cathode with all other tube voltages held constant, and the resulting increment in screen-grid current is measured; g_{22} is then the ratio of the increment in screen-grid current to the increment in screen-grid voltage. Alternatively, an indirect measurement may be more convenient. With the tube connected in the conventional amplifier circuit of Fig. 14-1 and adjusted for the appropriate quiescent operating point, the values of A_m, k_0, and k_2 are determined by three measurements. With a signal of suitable frequency applied at the input to the amplifier and with both the screen grid and the cathode completely bypassed ($Z_2 = Z_0 = 0$), the input and output voltages are measured. The ratio of these two voltages is A_m [Eq. (14-58)]. Next, the cathode bypass capacitor is removed to give the condition $Z_2 = 0$, $Z_0 = R_0$. Under these conditions, Eq. (14-60) gives the ratio E_4/E_1 as $-A_m/(1 + k_0)$; since A_m is known from the previous measurement, k_0 is determined by this measurement. In a like manner, the value of k_2 is determined by (14-61) from a measurement made with the screen-grid capacitor removed and the cathode resistor completely bypassed. The values of g_{22}, g_{41}, and g_0 (or ρ) can be determined from these values of A_m, k_0, and k_2 together with the values of the circuit resistances. Alternatively, the values of A_m, k_0, and k_2 can be used directly in Eq. (14-63) to construct the frequency characteristics of the amplifier or to choose optimum values for the bypass capacitors. It should be noted that the values of A_m, k_0, and k_2 depend on the values of the circuit resistances and on the quiescent operating point for the tube.

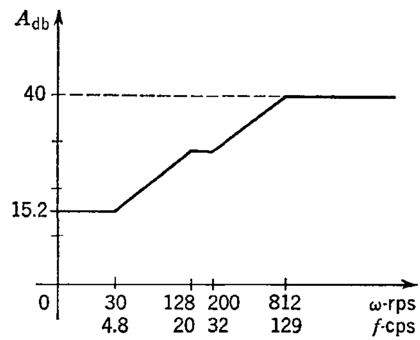


FIG. 14-18. Logarithmic amplitude characteristic for a typical pentode amplifier at low frequencies.

The analysis presented above illustrates the fact that factors of the form $j\omega + \omega_a = \omega_a(1 + j\omega/\omega_a)$ are basic elements in the analysis and design of electric circuits. The technique of expressing a higher-order polynomial as the product of a set of linear factors is one of fundamental importance and wide applicability. In many respects the properties of higher-order polynomials are perceived more easily when the polynomials are expressed in the factored form than when they are multiplied out.

The expressions for the two-port coefficients of electrical networks, such as the forward voltage or current transmittance, consist, in general, of the ratio of two polynomials in the variable $j\omega$; the proof of this statement follows from the form of the general solutions of the loop and node equations presented in Secs. 12-2 and 12-3. The coefficients in these polynomials are combinations of the circuit-parameter values; hence the coefficients are real numbers. The fundamental theorem of algebra ensures that these polynomials can be expressed as a product of linear factors; in particular, for any polynomial of degree n,

$$\begin{aligned} a_n m^n + a_{n-1} m^{n-1} + \cdots + a_2 m^2 + a_1 m + a_0 \\ = a_n(m^n + b_{n-1} m^{n-1} + \cdots + b_2 m^2 + b_1 m + b_0) \\ = a_n(m - m_n)(m - m_{n-1}) \cdots (m - m_2)(m - m_1) \end{aligned} \qquad (14\text{-}70)$$

The quantities $m_1, m_2, \ldots, m_n$ are the values of the variable m that make the polynomial zero; there are exactly n of these, and they are called the zeros of the polynomial. The circuits considered in Secs. 14-1 to 14-4 are special cases in which the polynomials arising are of first degree or are merely constants.

The zeros of a polynomial having real coefficients may be real, or they may occur in complex conjugate pairs. The method presented in this chapter for constructing the amplitude and phase characteristics cannot be applied directly to factors associated with complex zeros. A method for dealing with complex zeros is presented in Chap. 15.

In the design of pentode amplifiers, the circuit resistances must be chosen to give a suitable quiescent operating point for the tube. This choice fixes the factors A_m, k_0, and k_2. The bypass capacitors are then to be chosen to give suitable frequency characteristics. In the usual case, the capacitors are chosen so that the highest break frequency in the amplitude characteristic lies below the band of frequencies occupied by the signal (Fig. 14-18). Since two parameters, C_0 and C_2, are to be chosen to satisfy only one requirement, one of the choices can be made with a certain amount of freedom. As a rule, C_0 is much larger than C_2; hence the design procedure should aim at keeping C_0 reasonably small. A suitable procedure is therefore to first choose C_0 so that $(1 + k_0)\omega_0$ is somewhat below the band of frequencies occupied by the signal. This would be a satisfactory design if there were no screen-grid degeneration. Then a value of ω_2, and hence a value of C_2, can be found from Eq. (14-63) to locate the largest break frequency in the amplitude characteristic at a suitable point. Since the quadratic factor in (14-63) is unwieldy, this calculation is facilitated by using numerical values for ω_0, k_0, and k_2. The procedure can be illustrated in terms of the specific amplifier analyzed previously.

Suppose that the highest break frequency in the amplitude characteristic of Fig. 14-18 is to be located at 50 cps rather than 129 cps. Then it is reasonable to choose C_0 so that $(1 + k_0)\omega_0 = 2.3\omega_0$ corresponds to 25 cps, which is somewhat below 50 cps; this choice makes $\omega_0 = 68.2$ rps. Substituting numerical values in the denominator of (14-63) and writing m in place of $j\omega$ yields

$$m^2 + (157 + 16\omega_2)m + 1180\omega_2 \tag{14-71}$$

This polynomial has two zeros, m_1 and m_2; in order to meet the design requirement, a value of ω_2 must be found that makes the largest of these zeros correspond to 50 cps; that is, the zeros must be

$$m_1 = -2\pi(50) = -314 \text{ rps}$$

and $m_2 = -a$ rps, where a is a positive real number less than 314. The required value of ω_2 can be found from the relations between the zeros of (14-71) and its coefficients. The coefficient on m is the negative of the sum of the zeros, and the constant term is the product of the zeros. Hence

$$\begin{aligned} 157 + 16\omega_2 &= 314 + a \\ \text{and} \qquad 1180\omega_2 &= 314a \end{aligned} \tag{14-72}$$

Eliminating a between these two equations yields $\omega_2 = 12.9$ rps, and the second equation then yields $a = 48.5$ rps. The required value of C_2 is given by the relation $\omega_2 = 1/C_2R_2 = 12.9$.

When the design requirements are not critical, the procedure can be simplified by the following rule of thumb: Choose C_0 so that the break at $(1 + k_0)\omega_0$ lies below the signal-frequency band. Then choose C_2 so that $(1 + k_2)\omega_2$ lies well below ω_0.

14-6. The Transistor Current Amplifier at High Frequencies. The circuit of a typical transistor amplifier is shown in Fig. 14-19*a*. For the case where a drift transistor is employed, the high-frequency transistor model of Fig. 8-19*b* yields the circuit of Fig. 14-19*b* to represent the amplifier in the middle- and high-frequency ranges. The bypass capacitor is treated as a short circuit in this model, and the resistance R_1 accounts for the parallel combination of R_a and R_b. The behavior of the amplifier depends on frequency as a result of the parasitic capacitances C_e and C_c. In particular, at high frequencies C_e tends to short-circuit node b' to ground, with a consequent reduction of the voltage E'_{be} and the current $g_mE'_{be}$. The method of analysis employed here with the drift transistor is also applicable to the alloy-junction transistor with a minor modification to account for the effects of the base-to-collector resistance r_{bc} (Fig. 8-20). The principal difference between the drift and the alloy-junction transistors is that the values of C_e and C_c are perhaps 100 times larger in the latter than in the former; the high-frequency behavior of the latter is accordingly much poorer than that of the former.

As a result of the collector capacitance C_c, the input circuit in the transistor amplifier is not isolated from the output circuit at high frequencies. In addition, this capacitance introduces feedback that makes analysis by inspection difficult, if not impossible. It is clear that the role of the collector capacitance in the transistor amplifier is similar to that of the grid-to-plate capacitance in the triode amplifier (Fig. 14-10); it follows that the Miller effect is present in the transistor amplifier as well as in the triode amplifier. It is shown in the following paragraphs that, as in the case of the triode amplifier, the collector capacitance can in many cases be replaced by an equivalent capacitance connected from

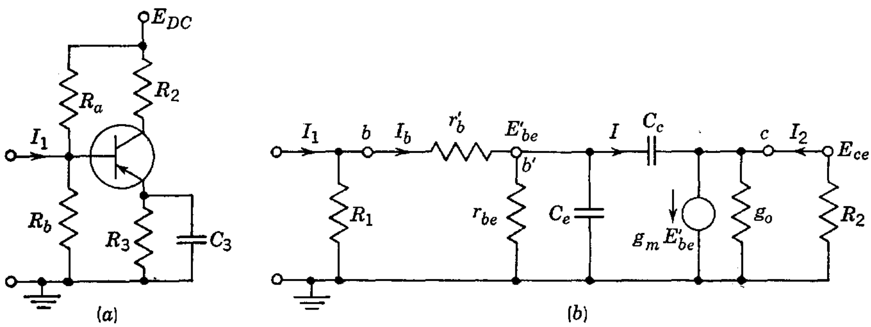


FIG. 14-19. Drift-transistor amplifier at high frequencies. (*a*) Circuit; (*b*) model.

node b' to ground. This modification of the circuit eliminates the analytical difficulties mentioned above.

The current I in the collector capacitance under steady-state sinusoidal operating conditions is given by

$$I = j\omega C_c(E'_{be} - E_{ce}) \tag{14-73}$$

In order to exploit this relation, it is necessary to obtain an expression for E_{ce} in terms of E'_{be}. The parameter values in most transistor amplifiers are such that the current I is much smaller than $g_m E'_{be}$ in the useful frequency range of the amplifier; thus, neglecting I in comparison with $g_m E'_{be}$,

$$E_{ce} = -\frac{g_m E'_{be}}{g_o + G_2} = -AE'_{be} \tag{14-74}$$

As a general rule, the value of A depends strongly on the value of R_2; if $G_2 \gg g_o$, $A \approx g_m R_2$. The value of A may be as small as 10 or as large as 1000.

Substituting (14-74) into (14-73) yields

$$I = j\omega C_c(1 + A)E'_{be} \tag{14-75}$$

Since this current has a negligible effect on the output circuit in the useful frequency range of the amplifier, the relations in the model of Fig. 14-19*b* remain substantially unchanged if C_c is removed and a capacitance $(1 + A)C_c$ is connected between node b' and ground. The factor $1 + A$, which accounts for the Miller effect, may make $(1 + A)C_c$ a relatively large capacitance.

Another simplification of the model is permitted by the fact that r_b' is of the order of 50 ohms in the usual drift transistor. Since this resistance is normally much smaller than R_1, it has a negligible effect on the behavior of the circuit (consider the Norton or Thevenin equivalent of the portion of the circuit on the left of r_{be}). Thus r_b' can be omitted from the model for the circuit of Fig. 14-19*a*. (When the transistor is used with an impedance-matching network between the signal source and the input terminals, as is usually the case at radio frequencies, r_b' becomes important and cannot be omitted.)

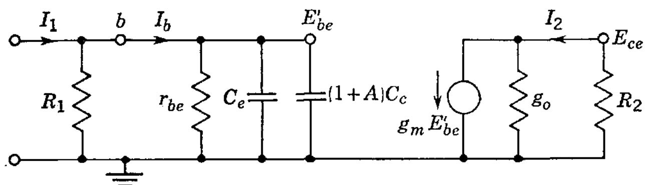


FIG. 14-20. Simplified model for the transistor amplifier at high frequencies.

When the simplifications developed above are introduced into the model for the amplifier, the circuit takes the form shown in Fig. 14-20. In this circuit the input is isolated from the output and there is no feedback; the significant effects of the reverse transmittance and the feedback are accounted for by the capacitance $(1 + A)C_c$. This circuit can be analyzed by inspection (that is, without solving simultaneous equations).

Denoting the total shunt capacitance in the input circuit by

$$C_t = C_e + (1 + A)C_c$$

and the total shunt conductance by $G_t = g_{be} + G_1$, the base voltage can be expressed as

$$E'_{be} = \frac{I_1}{G_t + j\omega C_t} = \frac{1}{G_t}\frac{I_1}{1 + j\omega C_t/G_t} \tag{14-76}$$

This equation can be written in a more useful form by defining

$$\omega_1 = \frac{G_t}{C_t} = \frac{g_{be} + G_1}{C_e + (1 + A)C_c} \tag{14-77}$$

Thus (14-76) becomes

$$E'_{be} = \frac{1}{G_t}\frac{I_1}{1 + j\omega/\omega_1} \tag{14-78}$$

The output current from the amplifier can be expressed, with the aid of (14-74), as

$$I_2 = -G_2E_{ce} = AG_2E'_{be} \tag{14-79}$$

Substituting (14-78) in (14-79) yields

$$I_2 = \frac{AG_2}{G_t}\frac{I_1}{1 + j\omega/\omega_1} \tag{14-80}$$

The forward current transmittance is, accordingly,

$$\frac{I_2}{I_1} = A_c = A_m \frac{1}{1 + j\omega/\omega_1} \tag{14-81}$$

where

$$A_m = \frac{AG_2}{G_t} = \frac{G_2}{g_o + G_2}\frac{g_m}{g_{be} + G_1} \tag{14-82}$$

is the current transmittance in the middle band of frequencies. When $G_2 \gg g_o$, as is often the case, (14-82) reduces to

$$A_m \approx \frac{g_m}{g_{be} + G_1} \tag{14-83}$$

Equation (14-81) has the same form as Eq. (14-4) for the voltage transmittance of the pentode amplifier at high frequencies; hence the logarithmic amplitude and phase characteristics for the transistor amplifier at high frequencies have the form shown in Fig. 14-4. The mid-band amplification is given by (14-82), and the half-power frequency is given by (14-77). Since C_e and $(1 + A)C_c$ are usually of the same order of magnitude, it follows from (14-77) that the half-power frequency usually depends rather strongly on the load resistance R_2.

Another useful relation is obtained from (14-77) by factoring g_{be} out of the numerator and C_e out of the denominator to obtain

$$\begin{aligned}\omega_1 &= \frac{g_{be}}{C_e}\frac{1 + G_1/g_{be}}{1 + (1 + A)C_c/C_e} \\ &= \omega_{cb}\frac{1 + G_1/g_{be}}{1 + (1 + A)C_c/C_e}\end{aligned} \tag{14-84}$$

where $\omega_{cb} = g_{be}/C_e$, defined in Chap. 8, is the intrinsic cutoff frequency of the transistor in the grounded-emitter connection. In most transistor amplifiers of the type shown in Fig. 14-19a, the circuit parameters are such that the numerator of (14-84) is smaller than the denominator; in such cases the cutoff frequency for the amplifier is somewhat less than the intrinsic cutoff frequency of the transistor. With a given transistor, the only adjustable quantities in (14-84) are G_1 and $A = g_m/(g_o + G_2)$; Eq. (14-84) shows that the half-power frequency can be controlled with the adjustable circuit parameters G_1 and G_2.

14-7. The Transistor Current Amplifier at Low Frequencies. The behavior of the transistor amplifier of Fig. 14-19*a* depends on the action of the bypass capacitor C_3 at low frequencies. An incremental model for the amplifier that is valid in the low- and medium-frequency ranges is shown in Fig. 14-21*a*. The hybrid representation for the transistor is used with $\mu_{bc} = g_o = 0$. These approximations, which are justifiable with drift transistors when R_2 does not exceed 10 or 20 kilohms, result in considerable algebraic simplification without affecting the results perceptibly.

The current source $\alpha_{cb}I_b$ in the circuit of Fig. 14-21*a* is replaced by an equivalent pair of sources (Fig. 12-3) to obtain the equivalent circuit shown in Fig. 14-21*b*. The symbol k is used to designate the quantity $1 + \alpha_{cb}$ in this figure. Eliminating the controlled source on the left by application of the reduction theorem leads to the reduced equivalent circuit shown in Fig. 14-21*c*. This last operation eliminates all feedback and permits the circuit to be analyzed by inspection. In a typical circuit, R_1 may be 10 kilohms and kR_3 may be 100 kilohms; hence at very low frequencies, where C_3 acts as an open circuit, most of the signal current flows through R_1 rather than into the base of the transistor. The current amplification is reduced accordingly. It follows that C_3 should be chosen to act as a short circuit at all frequencies in the band occupied by the signal.

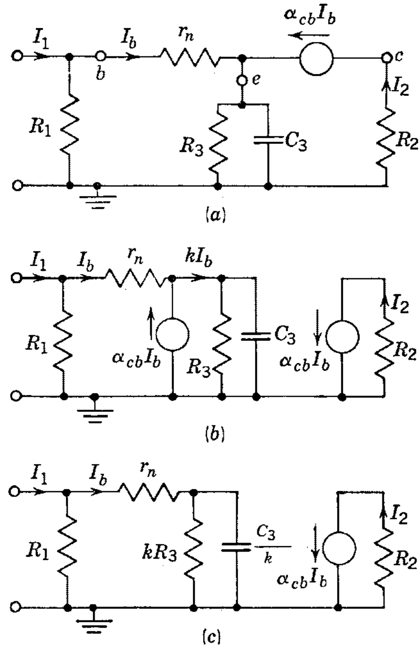


FIG. 14-21. The transistor amplifier at low frequencies. (*a*) Model; (*b*) an equivalent circuit, $k = 1 + \alpha_{cb}$; (*c*) reduced equivalent circuit, $k = 1 + \alpha_{cb}$.

A quantitative formulation of the forward current transmittance provides the relations that are pertinent in the design of the transistor amplifier for suitable frequency characteristics. By inspection of the circuit in Fig. 14-21*c*,

$$A_c = \frac{I_2}{I_1} = \frac{\alpha_{cb}I_b}{I_1} \tag{14-85}$$

and

$$I_b = \frac{R_1 I_1}{R_1 + r_n + \dfrac{kR_3}{1 + j\omega C_3 R_3}} \tag{14-86}$$

Defining $\omega_3 = 1/C_3R_3$, factoring $r_n + R_1$ out of the denominator, and multiplying above and below by $(1 + j\omega/\omega_3)$ yields

$$I_b = \frac{R_1}{r_n + R_1} \frac{1 + j\omega/\omega_3}{1 + \dfrac{kR_3}{r_n + R_1} + j\omega/\omega_3} I_1 \tag{14-87}$$

Substituting (14-87) into (14-85) and defining

$$A_m = \frac{\alpha_{cb}R_1}{r_n + R_1} \qquad \text{and} \qquad k_3 = \frac{kR_3}{r_n + R_1} \tag{14-88}$$

leads to an expression of familiar form:

$$A_c = \frac{A_m}{1 + k_3} \frac{1 + j\omega/\omega_3}{1 + j\omega/(1 + k_3)\omega_3} \tag{14-89}$$

Equation (14-89) has the same form as Eq. (14-18) for the voltage transmittance of a triode amplifier with cathode degeneration; hence the logarithmic amplitude and phase characteristics for the transistor at low frequencies have the forms shown in Figs. 14-7 and 14-8.

Equation (14-89) can be used as a guide in the design of transistor amplifiers for suitable low-frequency response. The resistances in the circuit are usually chosen to provide a suitable quiescent operating point for the transistor and to provide adequate stabilization of the quiescent point against changes in temperature and transistor parameters. These resistances, together with the transistor parameters, fix k_3. The bypass capacitor is then chosen to make the highest break frequency in the amplitude characteristic occur at a suitable point below the signal-frequency band.

14-8. Combined Low- and High-frequency Characteristics. In examining the frequency dependence of various electronic circuits in the preceding sections, the low-frequency phenomena are treated separately from the high-frequency phenomena. In this way the algebraic formulation of the circuit properties has been kept relatively simple. Underlying this technique, however, is the tacit assumption that the high-frequency breaks in the amplitude characteristic are separated by a decade of frequency or more from the low-frequency breaks; this condition is satisfied in many important electronic circuits.

The complete frequency characteristic for a circuit can be displayed by plotting the low-frequency and high-frequency characteristics on a single set of coordinates. As an example, the asymptotic approximation for the complete amplitude characteristic of a pentode amplifier, obtained by combining Figs. 14-3 and 14-18, is shown in Fig. 14-22. The numerical values of amplification (in decibels) and frequency shown in this figure indicate orders of magnitude that may be expected in a typical pentode amplifier. It should be clear, however, that in any particular amplifier the break frequencies and the amplifications can be adjusted over fairly wide ranges by adjusting the circuit parameters. It should be noted that the high-frequency break point in the typical characteristic of Fig. 14-22 is three decades above the highest low-frequency break point. This three-decade range is the mid-band for the amplifier.

The complete phase characteristic for the amplifier can be constructed in a similar manner by combining the low- and high-frequency characteristics in a single plot.

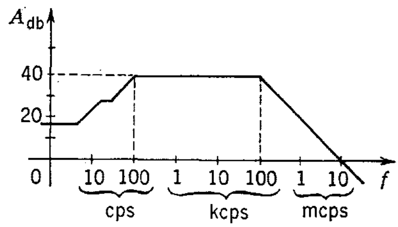


FIG. 14-22. Combined low- and high-frequency amplitude characteristics.

14-9. Gain-Bandwidth Relations. The amplifier having the amplitude characteristic shown in Fig. 14-22 amplifies signals lying anywhere in its mid-band (between 100 cps and 100 kcps) by approximately the same amount. This is, therefore, the useful band of the amplifier, and the bandwidth of the amplifier is 100,000 − 100 = 99,900 cps. In practice it is customary to ignore the 100 cps of bandwidth lost at the low-frequency end of the spectrum and to identify the bandwidth with the upper break, or half-power, frequency, which is 100 kcps in this case.

An important relation exists between the mid-band amplification and the bandwidth of such amplifiers. These two quantities are given by Eqs. (14-3) for pentode amplifiers and by Eqs. (14-24) for triode amplifiers; their product in the case of the pentode is

$$A_m\omega_4 = \frac{g_m}{g_p + G_4}\,\frac{g_p + G_4}{C_4} = \frac{g_m}{C_4} = \omega_{gb} \tag{14-90}$$

This quantity, known as the gain-bandwidth product, is designated as a frequency, ω_{gb}, because it has those dimensions and because it has an important interpretation as a frequency.

Equation (14-90) shows that the gain-bandwidth products for circuits like that in Fig. 14-1 depend only on the transconductance of the tube and the parasitic capacitance shunting the output terminals. The amplification and the bandwidth of such an amplifier can be adjusted over wide ranges by changing R_4; however, the gain-bandwidth product must remain constant as long as g_m and C_4 remain constant. The asymptotes for the high-frequency amplitude characteristic of such an amplifier are shown in Fig. 14-23 for two different values of R_4. If R_4 is changed so as to reduce the mid-band amplification by a factor of 2, the mid-band amplification in decibels becomes

$$\begin{aligned} A'_{dbm} &= 20 \log \frac{A_m}{2} = 20 \log A_m - 20 \log 2 \\ &= A_{dbm} - 6 \text{ db} \end{aligned} \tag{14-91}$$

That is, the mid-band amplification is reduced by 6 db. But this change in R_4 also doubles the bandwidth, corresponding to an increase in bandwidth of one octave. Thus the break point in amplitude characteristic moves down by 6 db on the amplification scale and it moves to the right by one octave on the frequency scale; hence the new break frequency ω_4' lies on the original high-frequency asymptote as illustrated in Fig. 14-23.

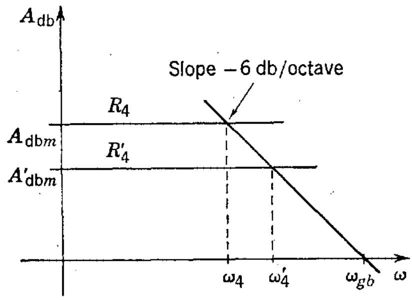


FIG. 14-23. Gain-bandwidth relations.

It follows from the above discussion that the high-frequency asymptote is independent of the mid-band amplification provided g_m and C_4 remain constant. It is appropriate, therefore, to examine the factors that fix the position of the high-frequency asymptote. The equation of this asymptote, (14-10), is

$$A_{db} = A_{dbm} - 20 \log \frac{\omega}{\omega_4} \qquad (14\text{-}92)$$

The position of the asymptote can be specified by giving the frequency at which it crosses the axis of zero db. This frequency is found by equating (14-92) to zero, giving

$$A_{dbm} = 20 \log \frac{\omega}{\omega_4} = 20 \log A_m \qquad (14\text{-}93)$$

Solving (14-93) for ω gives the frequency of the zero crossing as

$$\omega = A_m\omega_4 = \omega_{gb} \qquad (14\text{-}94)$$

Thus the gain-bandwidth product is also the frequency at which the high-frequency asymptote crosses the axis of zero db; at this frequency the voltage amplification is unity.

The foregoing developments suggest a straightforward procedure for the design of amplifiers such as that shown in Fig. 14-1. If g_m and C_4 are known, the high-frequency asymptote of the amplitude characteristic can be constructed at once; it is a line passing through zero db at $\omega = \omega_{gb} = g_m/C_4$ with a slope of -6 db/octave. The plate load resistance is then chosen to give a suitable compromise between amplification and bandwidth. For each 6 db of mid-band amplification, the bandwidth is reduced one octave from the value ω_{gb}.

In practice the design procedure is complicated slightly by the fact that g_m and C_4 are not known exactly at the outset. The transconductance depends somewhat on the quiescent operating point of the tube, which in turn depends on the value of the plate load resistance, and the shunt capacitance depends somewhat on the physical layout of the circuit components and wiring. Nevertheless, the procedure outlined in the preceding paragraph serves as a guide to effective design.

The gain-bandwidth relation presented above, sometimes called the gain-bandwidth theorem, states that the gain-bandwidth product of the amplifier in Fig. 14-1 is constant if the ratio g_m/C_4 remains constant. It must be noted that the proof applies only to circuits having the form shown in Fig. 14-1, namely, circuits consisting of a simple parallel combination of R's and C's. Although this property may be found in circuits having other forms, it does not apply in general. As an illustration of the limited applicability of the gain-bandwidth theorem, consider the transistor amplifier of Sec. 14-6. The gain-bandwidth product, obtained from Eqs. (14-77) and (14-82), is

$$\omega_{gb} = A_m\omega_1 = \frac{g_m G_2}{(g_o + G_2)[C_e + (1 + A)C_c]} \tag{14-95}$$

This product is seen to be independent of G_1, but it is not independent of G_2. Even if G_2 is much larger than g_o, so that it does not appear explicitly in (14-95), it still affects the value of ω_{gb} through the factor A. Furthermore, the constancy of ω_{gb} in the face of changes in G_1 is not valid for all values of G_1, for the analysis leading to (14-95) is based on the assumption that r_b' can be neglected in comparison with R_1. If R_1 is made very small in order to obtain a large bandwidth, r_b' cannot be neglected. It is easily shown that the value of ω_{gb} depends on R_1 when r_b' is not neglected.

14-10. Summary. The behavior of many important electronic amplifiers depends upon frequency in both the high-frequency and the low-frequency ranges because unavoidable parasitic elements become effective at high frequencies and auxiliary circuit elements such as bypass capacitors become effective at low frequencies. When the high-frequency phenomena are separated from the low-frequency phenomena by a decade of frequency or more, these two sets of phenomena can be analyzed separately. The analysis is simplified by this separate treatment, with the result that in many cases a clear understanding of the relations between circuit parameters and circuit performance is obtained. This valuable technique finds application in many fields other than electronics.

The presentation of frequency characteristics is greatly facilitated by the use of the logarithmic amplitude and phase characteristics. The value of these characteristics lies in the fact that they have simple asymptotic properties and in the fact that complicated characteristics can be constructed by the addition of simple component characteristics.

The simplicity of the logarithmic amplitude and phase characteristics is associated with factors of the form $j\omega + \omega_a = \omega_a(1 + j\omega/\omega_a)$, which are linear functions of frequency. These linear factors are the basic building blocks in network functions such as the two-port coefficients. The study of these network functions is facilitated in several ways by resolving them into their constituent linear factors.

Network functions in general can be expressed as a ratio of two polynomials in the variable $j\omega$, as is illustrated by Eq. (14-63). The truth of this statement follows from the form of the general solutions of the loop and node equations presented in Secs. 12-2 and 12-3. The fundamental theorem of algebra ensures that these polynomials can be expressed as the product of a set of linear factors as illustrated by Eq. (14-69); it does not, unfortunately, provide an easy way of finding these factors when the degree of the polynomial exceeds 2. The linear factors composing a network function not only aid in understanding the behavior of the network under steady-state sinusoidal operating conditions, but they also provide critical information about the transient response of the network, a fact that is treated in Chap. 16. For these reasons an important part of network theory is concerned with interpreting the behavior of networks in terms of the linear factors constituting the network functions.

PROBLEMS

14-1. A certain pentode amplifier has the form shown in Fig. 14-1. The plate load resistance is $R_4 = 220$ kilohms, and the incremental tube parameters are $g_m = 1.2$ millimhos and $r_p = 1$ megohm. The total capacitance between plate and ground, including wiring capacitance, is 15 $\mu\mu$f.

a. Determine the mid-band amplification in decibels and the half-power frequency in cycles per second for the high-frequency model.

b. Plot the logarithmic amplitude and phase characteristics for the high-frequency model. Give the true characteristics, not just the asymptotes. Use semilog coordinate paper, and calibrate the frequency scale in kilocycles per second.

14-2. A 6AU6 pentode is used in the voltage amplifier of Fig. 14-1. The plate-supply voltage is 300 volts, and the quiescent operating point is to be $E_b = E_{c2} = 100$ volts and $I_b = 2$ ma.

a. Find the values of R_0, R_2, and R_4 required to give the specified operating point. Refer to Sec. 10-2 for a discussion of the screen-grid current.

b. The tube parameters can be approximated by $g_m = 2.4$ millimhos and $r_p = 10$ megohms. Ten micromicrofarads are to be added to the output capacitance of the tube to account for wiring capacitance. Evaluate the mid-band amplification in decibels and the high-frequency half-power point in cycles per second.

c. Sketch and dimension the asymptotes for the high-frequency logarithmic amplitude and phase characteristics.

14-3. The triode section of a 6AV6 is used in the amplifier shown in Fig. 14-5. The quiescent operating point is to be at $E_b = 100$ volts and $I_b = 0.5$ ma; the plate supply voltage is 300 volts.

a. Determine the values of R_0 and R_2 required to give the specified quiescent point.

b. Evaluate μ and r_p at the quiescent point, and determine the mid-band amplification in db.

c. The highest break frequency in the logarithmic amplitude characteristic for the low-frequency model is to be 159 cps. What must be the size of the cathode bypass capacitor?

d. Sketch and dimension the asymptotes for the low-frequency amplitude and phase characteristics. Give amplitudes in decibels and frequencies in cycles per second.

14-4. The forward voltage transmittance of a certain triode amplifier with cathode degeneration is

$$A_{vo} = (15)\,\frac{1 + j\omega/100}{1 + j\omega/200}$$

Plot the logarithmic amplitude and phase characteristics. Use semilog coordinate paper, and show both the asymptotes and the true characteristics. Calibrate the frequency scale in cycles per second. *Note:* The objective of this problem is to provide practice in applying the corrections to the asymptotes when two break frequencies occur close together.

14-5. The input admittance to the amplifier described in Example 14-1 at high frequencies is to be examined.

a. Give an incremental model for the circuit that is valid at high frequencies. The parasitic capacitances are $C_1 = C_2 = 10\ \mu\mu f$ and $C_{12} = 2.5\ \mu\mu f$ when the stray wiring capacitances are added to the tube capacitances.

b. Determine the parameter values in the equivalent input circuit shown in Fig. 14-13.

c. The resistance R in Fig. 14-13 can be neglected, for most purposes, as long as it is less than one-tenth the reactance of $C = A_m C_{12}$. Over what frequency range is R negligible?

14-6. One section of a 12AU7 is used in the cathode follower of Fig. 14-16. The circuit parameters are $R_2 = 33$ kilohms, $R_3 = 1$ kilohm, $R_1 = 1$ megohm, $\mu = 17$, $r_p = 10$ kilohms, $C_1 = 3.0\ \mu\mu f$, $C_2 = 10\ \mu\mu f$, and $C_{12} = 2.0\ \mu\mu f$. Sketch and dimension the asymptotes for the high-frequency logarithmic amplitude and phase characteristics. Give amplitudes in decibels and frequencies in kilocycles per second.

14-7. A pentode having the incremental parameters $g_m = 5$ millimhos and $r_p = 1$ megohm is to be used in an amplifier of the form shown in Fig. 14-1. The high-frequency half-power point is to be at 1 mcps.

a. If C_4 is 15 $\mu\mu f$, what value of R_4 is required?

b. With this design, what is the mid-band amplification in decibels?

c. If $I_b = 5$ ma, $E_b = 100$ volts, and $E_{c1} = -3$ volts at the quiescent operating point, what plate-supply voltage is required?

14-8. The low-frequency behavior of a pentode amplifier like the one shown in Fig. 14-1 is to be studied. The circuit parameters are $R_4 = 50$ kilohms, $R_2 = 300$ kilohms, $R_0 = 1$ kilohm, $C_0 = 2\ \mu f$, and $C_2 = 0.05\ \mu f$. The following data are obtained from measurements made in the mid-band: With both C_0 and C_2 in place, the amplification is 75; with only C_2 present, the amplification is 25; and with only C_0 present, the amplification is 9.

a. Determine the values of the constants A_m, $1 + k_0$, and $1 + k_2$ that appear in the expression for A_{vo} at low frequencies. Determine ω_0 and ω_2.

b. Sketch and dimension the asymptotes for the low-frequency logarithmic amplitude characteristic for the amplifier when both bypass capacitors are present. Give amplitudes in decibels and frequencies in cycles per second.

14-9. A pentode having the parameters $g_{41} = 1.5$ millimhos, $g_{22} = 0.022$ millimho, $g_{44} = 0$, and $\rho = 0.3$ is used in the amplifier of Fig. 14-1. The circuit parameters are $R_4 = 100$ kilohms, $R_2 = 330$ kilohms, $R_0 = 1.5$ kilohm, and $C_2 = 0.05\ \mu f$.

a. Assuming that C_0 is a short circuit for all frequencies of interest, sketch and dimension the asymptotes for the low-frequency amplitude characteristic. Give amplitudes in decibels and frequencies in cycles per second.

b. Repeat part *a* for $C_0 = 5\ \mu f$.

14-10. The forward voltage transmittance of a certain pentode amplifier at low frequencies is

$$A_{vo} = -20 \frac{(1 + j\omega/50)(1 + j\omega/200)}{(1 + j\omega/100)(1 + j\omega/400)}$$

Plot the corresponding amplitude characteristic. Use semilog-coordinate paper, and show both the asymptotes and the true characteristic. Calibrate the frequency scale in cycles per second. *Note:* The objective of this problem is to provide practice in applying the corrections to the asymptotes when several break frequencies occur close together.

14-11. A drift transistor is used in the current amplifier shown in Fig. 14-19*a*. The circuit parameters are $R_a = 180$ kilohms, $R_b = 20$ kilohms, $R_2 = 16$ kilohms, and $R_3 = 2$ kilohms; the transistor parameters are $r_b' = 40$ ohms, $r_{be} = 1.6$ kilohms, $g_m = 37$ millimhos, $C_e = 200\ \mu\mu f$, $C_c = 2\ \mu\mu f$, and $g_o = 0$.

a. Give an incremental model for the amplifier that is valid in the middle- and high-frequency ranges where C_3 acts as a short circuit.

b. Evaluate the parameters for a simplified representation of the form shown in Fig. 14-20, treating r_b' as a short circuit.

c. Sketch and dimension the asymptotes for the logarithmic amplitude and phase characteristics in the middle- and high-frequency ranges. Give amplitudes in decibels and frequencies in kilocycles per second.

14-12. The validity of treating r_b' as a short circuit in the amplifier of Prob. 14-11 is to be examined. From the data given in Prob. 14-11, determine the value of R_1 in the model of Fig. 14-19*b*. Assuming a current source I_1 applied at the input, find the Norton equivalent of the circuit on the left of node b' in the circuit of Fig. 14-19*b*. Is this equivalent circuit affected appreciably by assuming $r_b' = 0$?

14-13. An alloy-junction transistor is used in the amplifier of Prob. 14-11. The circuit parameters are $R_a = 100$ kilohms, $R_b = 22$ kilohms, $R_2 = 20$ kilohms, and $R_3 = 4$ kilohms; the transistor parameters for the high-frequency model of Fig. 8-20 are $r_b' = 250$ ohms, $r_{be} = 2.5$ kilohms, $r_{bc} = 7$ megohms, $r_{ce} = 150$ kilohms, $g_m = 20$ millimhos, $C_e = 5000\ \mu\mu f$, and $C_c = 50\ \mu\mu f$.

Following the procedure outlined in Prob. 14-11, construct and dimension the asymptotes for the amplitude characteristic in the middle- and high-frequency ranges.

14-14. The low-frequency behavior of the amplifier described in Prob. 14-11 is to be examined.

a. From the data given in Prob. 14-11, determine the parameters in the low-frequency hybrid model for the transistor.

b. What value of C_3 is required if the largest break frequency in the amplitude characteristic is to be 100 cps? Sketch and dimension the asymptotes for the low-frequency amplitude characteristic with this value of C_3. Give amplitudes in decibels and frequencies in cycles per second.

14-15. The triode amplifier shown in Fig. 14-24 is provided with a 1-megohm potentiometer to serve as a volume control. The effect of the volume control on the high-frequency characteristics is to be examined. The incremental parameters for the tube are $\mu = 100$ and $r_p = 80$ kilohms; the parasitic capacitances of the tube and wiring are $C_1 = C_2 = 5\ \mu\mu f$ and $C_{12} = 3\ \mu\mu f$; and the coupling and bypass capacitors act as short circuits at all frequencies of interest.

a. The input to the triode can be represented by an equivalent circuit of the form shown in Fig. 14-13. Determine the values of the parameters in this circuit.

b. Sketch and dimension the asymptotes for the high-frequency amplitude characteristic of the voltage ratio E_1/E_s for the condition that the volume control is set for maximum output voltage. Give amplitudes in decibels and frequencies in kilocycles per second. As in Example 14-2, the resistance R in the equivalent input circuit can be treated as a short circuit in the frequency range of interest.

c. Repeat part *b* with the volume control set at its mid-point (that is, so that the potentiometer has 500 kilohms on each side of the slider).

d. Discuss briefly the significance of the results obtained in parts *b* and *c*.

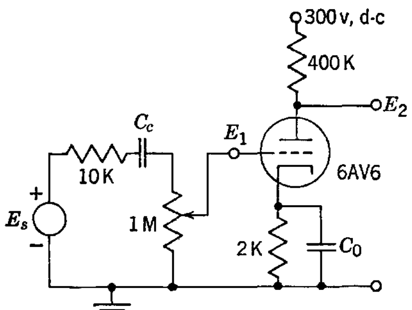


FIG. 14-24. Amplifier for Prob. 14-15.

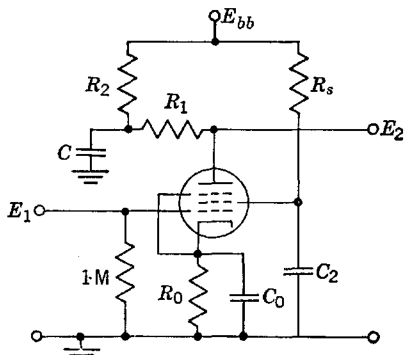


FIG. 14-25. Amplifier for Prob. 14-16.

14-16. The combination of R_2 and C in the pentode amplifier of Fig. 14-25 forms a decoupling network of a type commonly used to reduce the interaction among amplifier stages using a common power supply. The effect of the decoupling network on the low-frequency amplitude characteristic is to be examined.

a. Give an incremental model for the circuit in which r_p is treated as an open circuit and C_0 and C_2 are treated as short circuits.

b. Show that the forward voltage transmittance for this model has the form

$$A_{vo} = K \frac{1 + j\omega/\omega_a}{1 + j\omega/\omega_b}$$

in the low- and middle-frequency ranges.

c. Sketch and dimension the logarithmic amplitude characteristic for $g_m = 2$ millimhos, $R_1 = 100$ kilohms, $R_2 = 33$ kilohms, and $C = 0.05$ μf. Give the dimensions in decibels and cycles per second.

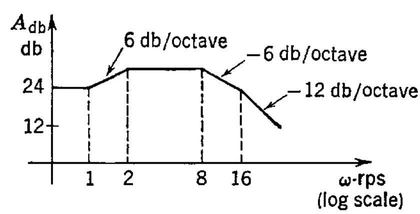


FIG. 14-26. Amplitude characteristic for Prob. 14-17.

14-17. The amplitude characteristic for a certain amplifier is determined experimentally and found to have the asymptotes shown in Fig. 14-26. The corresponding voltage transmittance consists of a constant multiplier and a number of linear factors of the form $1 + j\omega/\omega_a$ [see Eq. (14-69), for example]. Determine an expression for this transmittance in which all constants are expressed as numbers. (Note that if the signs in some or all of the linear factors are changed, the amplitude characteristic is not changed; information about the signs is contained only in the phase characteristics.)

14-18. The incremental parameters for the pentode in the amplifier shown in Fig. 14-27 are $g_m = 5$ millimhos and $g_p = 0$. The parasitic capacitances are negligible, and the coupling and bypass capacitors act as short circuits at all frequencies of interest. Sketch and dimension the asymptotes for the amplitude characteristic. Give amplitudes in decibels and frequencies in cycles per second.

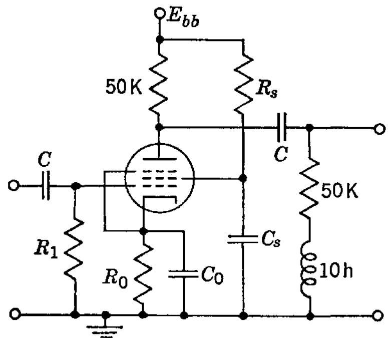


FIG. 14-27. Amplifier for Prob. 14-18.

14-19. The high-frequency behavior of the circuit presented in Prob. 13-11 is to be examined. The tube parameters are $\mu = 20$ and $r_p = 10$ kilohms; the circuit resistances are $R_1 = 667$ ohms and $R_2 = 25$ kilohms. The characteristics of the circuit are dominated at high frequencies by the capacitance between the conductor of the cable and the shield surrounding it; the value of this capacitance is 100 $\mu\mu$f. The coupling capacitor acts as a short circuit, and the parasitic capacitances are negligible in the frequency range of interest.

a. Give an incremental model for the circuit, and determine the total admittance $Y(\omega)$ in parallel with the current source I.

b. The input voltage to the cathode follower is $E_1 = I/Y$. Using the results of part *a*, sketch and dimension the asymptotes for the logarithmic amplitude characteristic of the ratio E_1/I.

c. Repeat parts *a* and *b* for the case where the cable shield is connected to the junction of R_1 and R_2.

d. Discuss briefly the significance of the results obtained in parts *b* and *c*.

CHAPTER 15

FREQUENCY DEPENDENCE OF CASCADED AMPLIFIERS

The signal amplification obtainable from most single-stage voltage and current amplifiers is limited to a factor of a few hundred or less, depending on the type of tube or transistor employed. Since there are many applications requiring greater signal amplifications than this, it is common practice to connect a number of stages in cascade so that each stage amplifies the signal in succession; the total amplification is then the product of the amplifications of the individual stages. Cascading amplifiers in this manner introduces some additional problems in circuit design, and it requires the use of additional circuit components that affect the behavior of the circuit.

In many cases it is possible to analyze (or design) each stage in a multistage amplifier as a separate part of the over-all circuit and to determine the over-all properties of the amplifier by combining the results of these separate analyses. Thus the results obtained in Chap. 14 are to a large extent directly applicable in the analysis of cascaded amplifiers; they are usually modified in some respects, however, by the networks used to couple the individual stages in cascade.

The properties of cascaded amplifiers are usually such that the signal transmission tends to zero at both low and high frequencies; hence these amplifiers are *bandpass* networks. In some cases cascaded amplifiers are required to amplify signals occupying a wide band of frequencies; the principal design problem is then to obtain uniform amplification over a sufficiently wide band of frequencies. In other cases the bandpass nature of the amplifier is turned to profit by using the circuit to amplify signals in a certain frequency range to the exclusion of all other signals; the design problem in these cases is to obtain uniform amplification in the desired band with sufficiently strong discrimination against signals in adjacent bands.

The objective of this chapter is to continue the development of the basic techniques of frequency analysis initiated in Chap. 14. In the course of this development the pole-zero diagrams for the voltage and current transmittances of amplifiers are introduced. These diagrams permit simple geometric representations of the factors of polynomials, and they present in a different form the information contained in the amplitude and phase characteristics. The pole-zero diagrams permit certain important circuit properties to be examined and characterized more directly and more simply than is possible by other means.

15-1. General Considerations in Cascading Amplifiers.[1] A cascade of two pentode voltage-amplifier stages coupled together by an RC network is shown in Fig. 15-1a; an incremental model for this amplifier that is valid in the frequency range where the bypass capacitors are short circuits is shown in Fig. 15-1b. The bypass capacitors are initially

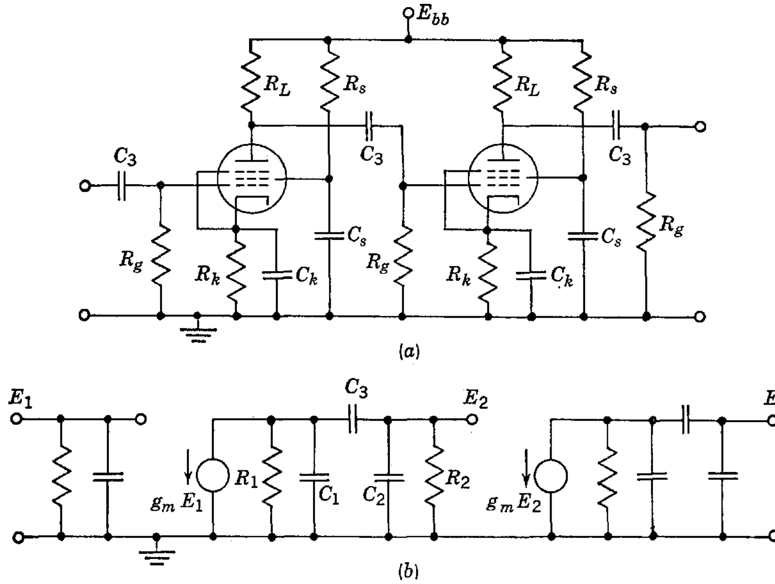


FIG. 15-1. RC-coupled pentode stages. (a) Circuit; (b) model.

assumed to be short circuits so that attention may be focused on the coupling network without distracting complications from other circuit components. This assumption is removed after a preliminary examination of the circuit. The resistance R_1 in the model represents the parallel combination of the plate resistance of the tube and the plate load resistance, C_1 represents the output capacitance of the first tube plus parasitic wiring capacitance, and C_2 represents the input capacitance of the second tube plus wiring capacitance. The pentodes have negligible grid-to-plate capacitance, and it is assumed that the circuit components are arranged so that there is negligible wiring capacitance between the grid and plate circuits. Hence there is no coupling between the grid and plate circuits except through the transconductance of the tube, and each stage of the amplifier can be analyzed as a separate part of the circuit. The capacitor C_3 is the interstage coupling capacitor; its function is discussed in the following paragraph.

The analysis and the design of cascaded amplifiers are concerned chiefly with the transmission of signals through the interstage coupling networks. The nature of the coupling network is dictated in part by the following considerations. There is a large positive direct voltage at the plate of the first tube, whereas it is usually required that there be no d-c component of voltage between the grid of the second tube and ground. Therefore it is usually desired that there be no d-c transmission through the interstage network. The capacitor C_3 serves to block the d-c transmission through the *RC* coupling network in the circuit of Fig. 15-1.

A variety of interstage networks is available to meet the needs of various types of amplifiers. The *RC* network is probably the simplest of them all, however, and it is certainly the most widely used for signals in the frequency range between roughly 10 cps and 100 kcps. Radio-frequency amplifiers, which operate at frequencies above this band, usually employ parallel resonant circuits or inductively coupled resonant circuits in the interstage network. Amplifiers for signals with frequencies less than 10 cps often use resistive coupling networks with passbands extending to zero frequency; these are referred to as directly coupled amplifiers.

It is appropriate to inquire whether or not there is any limitation on the number of stages that can be connected in cascade, and therefore whether or not there is any limitation on the over-all signal amplification that can be realized. The noise voltage generated in the first stage of a multistage amplifier is amplified in succession by each stage of the amplifier. As more stages are connected in cascade, the noise voltage at the output of the amplifier increases until eventually it becomes excessive. For this reason it is usually not possible to realize amplifications in excess of about one million in the usual multistage amplifier. In order to increase the amplification that can be obtained, special low-noise circuits, such as the grounded-grid amplifier and the cascode amplifier, are often used for the input stage of cascaded amplifiers. In addition, since the noise voltage is more or less uniformly distributed over a wide band of frequencies, the effective noise voltage at the output can be reduced by restricting the bandwidth of the amplifier. High-gain amplifiers should therefore be designed to have bandwidths just great enough to accommodate the signal.

15-2. *RC*-coupled Pentode Amplifiers. The behavior of the *RC*-coupled amplifier of Fig. 15-1 is affected at high frequencies by the parasitic shunt capacitances C_1 and C_2, and its behavior at low frequencies is affected by the coupling capacitor C_3. However, the mid-band for such amplifiers usually extends over much more than one decade of frequency; hence the high-frequency performance can be studied separately from the low-frequency performance. The model for one stage of the amplifier, shown in Fig. 15-1*b*, reduces to the circuit shown in Fig. 15-2*a* at high frequencies. The coupling capacitor is treated as a short circuit, the total conductance shunting the output is $G_t = 1/R_1 + 1/R_2$, and the total capacitance shunting the output is $C_t = C_1 + C_2$. At low frequencies the model reduces to the circuit shown in Fig. 15-2*b*; the shunt capacitances are treated as open circuits at low frequencies.

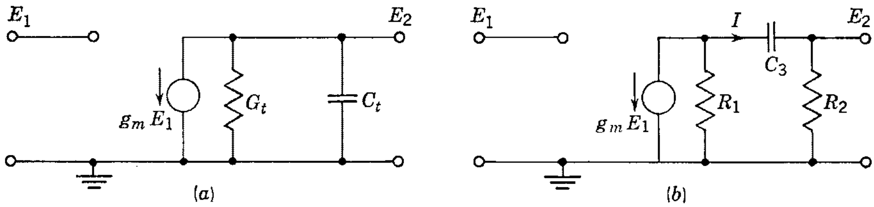


FIG. 15-2. Models for the *RC*-coupled pentode amplifier. (*a*) High-frequency model; (*b*) low-frequency model.

The high-frequency model in Fig. 15-2*a* has the same form as the circuit in Fig. 14-1*b*; hence the forward voltage transmittance has the same form as Eq. (14-4) and can be written as

$$A_{vo} = -A_m \frac{1}{1 + j\omega/\omega_H} \tag{15-1}$$

where $\omega_H = G_t/C_t = 1/R_tC_t$ is the high-frequency half-power point. The amplitude and phase characteristics associated with (15-1) have the same form as those shown in Fig. 14-4.

At low frequencies the output voltage, obtained from Fig. 15-2*b*, is $E_2 = R_2I$, and I is related to the current g_mE_1 by a current-divider ratio; hence

$$\begin{aligned} A_{vo} &= \frac{-g_mR_1R_2}{R_1 + R_2 + 1/j\omega C_3} = \frac{-g_mR_1R_2}{R_1 + R_2}\,\frac{j\omega C_3(R_1 + R_2)}{1 + j\omega C_3(R_1 + R_2)} \\ &= -A_m \frac{j\omega/\omega_L}{1 + j\omega/\omega_L} \end{aligned} \tag{15-2}$$

where $\omega_L = 1/C_3(R_1 + R_2)$ is the low-frequency half-power point. Equation (15-2) can be written in an alternative form that is often useful:

$$A_{vo} = -A_m \frac{1}{1 - j\omega_L/\omega} \tag{15-3}$$

The amplitude characteristic associated with the numerator in Eq. (15-2) is a straight line crossing the frequency axis at $\omega = \omega_L$ with a slope of

6 db/octave, and the phase shift is constant at 90°; the amplitude and phase characteristics of the denominator have the same form as those shown in Fig. 14-4. The complete characteristics for the low-frequency model are obtained by adding the characteristics for these factors. By plotting the characteristics for both the high- and low-frequency models on the same set of coordinates, the combined characteristics are obtained; the asymptotic approximations for the combined characteristics are shown in Fig. 15-3.

The gain-bandwidth product for the amplifier analyzed above is

$$\omega_{gb} = A_m\omega_H = \frac{g_m}{C_t} = \frac{g_m}{C_1 + C_2} \qquad (15\text{-}4)$$

If the parasitic wiring capacitance is zero, C_1 is the output capacitance of the first pentode in Fig. 15-1, and C_2 is the input capacitance of the second pentode. If the two pentodes are of the same type, the quantity $(g_m/C_{\text{in}} + C_{\text{out}})$, which is the maximum value that the gain-bandwidth product can have, is a figure of merit for the tube. Much effort has been devoted to increasing this ratio; values corresponding to frequencies of several hundred megacycles per second are realized in the better pentodes.

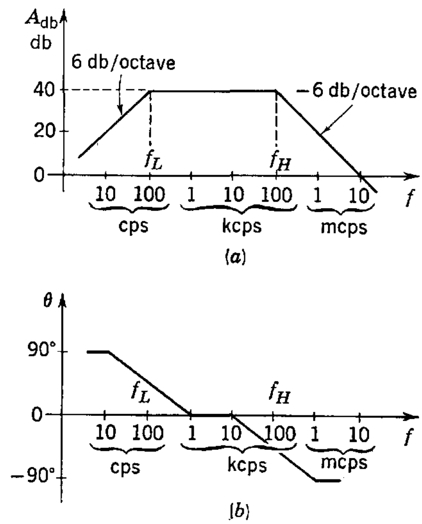


FIG. 15-3. Frequency characteristics for a typical RC-coupled pentode amplifier stage. (a) Amplitude; (b) phase shift.

It is usually not permissible to assume, as is done above, that the bypass capacitors are short circuits in the vicinity of the low-frequency half-power point. Figure 15-4 shows an incremental model adapted from the circuit in Fig. 14-17 to account for the action of the bypass capacitors

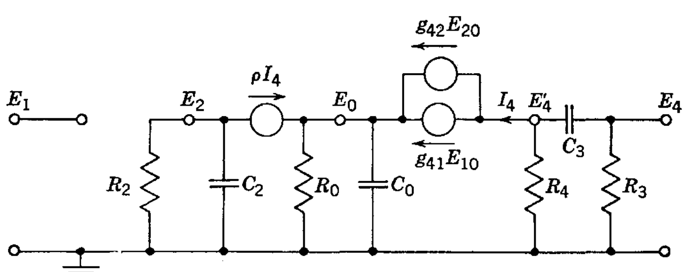


FIG. 15-4. Low-frequency model for the RC-coupled pentode amplifier with screen-grid and cathode degeneration, adapted from Fig. 14-17.

as well as the coupling capacitor at low frequencies. The algebra involved in determining the amplitude and phase characteristics of this circuit in general is complicated, however, by the fact that the denominator of A_{vo} is a third-degree polynomial in $j\omega$. The characteristics can be constructed in the usual manner after the cubic polynomial has been factored, but the factoring process itself it tedious, and the results yield little insight. The usual method of factoring such polynomials is to substitute numerical values for the circuit parameters and to seek the factors by trial and error with synthetic division.

The parameter values in many pentode amplifiers permit a simplification in the analysis that greatly reduces the labor required in constructing the frequency characteristics and that provides useful insight. The grid return resistance R_3 is often much larger than the plate load resistance R_4, with the result that the current in R_3 is much smaller than I_4. In such cases E'_4 can be evaluated with R_3 treated as an open circuit, and the ratio E'_4/E_1 is given by Eq. (14-63). The output voltage is then given by the voltage-divider relation:

$$E_4 = \frac{j\omega C_3 R_3}{1 + j\omega C_3 R_3} E'_4 = \frac{j\omega/\omega_3}{1 + j\omega/\omega_3} E'_4 \tag{15-5}$$

It follows from Eqs. (14-63) and (15-5) that at low frequencies

$$A_{vo}(j\omega) = -A_m \frac{(j\omega + \omega_0)(j\omega + \omega_2)}{(j\omega + \omega_1)(j\omega + \omega_4)} \frac{j\omega/\omega_3}{(1 + j\omega/\omega_3)} \tag{15-6}$$

where the denominator of (14-63) has been expressed as the product of its linear factors. Rearranging (15-6) yields

$$A_{vo}(j\omega) = -A_m \frac{\omega_0\omega_2}{\omega_1\omega_4} \frac{(j\omega/\omega_3)(1 + j\omega/\omega_0)(1 + j\omega/\omega_2)}{(1 + j\omega/\omega_1)(1 + j\omega/\omega_4)(1 + j\omega/\omega_3)} \tag{15-7}$$

The amplitude and phase characteristics associated with (15-7) can be constructed by adding the characteristics of the constituent factors. The amplitude characteristic corresponds to the sum of the characteristic in Fig. 14-18 and the low-frequency part of the characteristic in Fig. 15-3a.

15-3. *RC*-coupled Transistor Amplifiers. A cascade of two transistor stages is shown in Fig. 15-5a. The general considerations in cascading transistor stages are similar to those discussed in Sec. 15-1 in connection with the pentode amplifier; hence an RC interstage network is used in the circuit of Fig. 15-5a to block the transmission of direct current from one stage to the next. Transistors can also be directly coupled, or more elaborate interstage networks may be used to achieve certain desired results.

A high-frequency model for the cascaded transistor stages is shown in

Fig. 15-5*b* under the assumption that drift transistors are employed and that the resistance r_b' is negligible. The conductances G_{t1}, G_{t2}, and G_{t3} are the total conductances connected between the three nodes and ground. The analysis of this circuit is complicated by the fact that at high frequencies, in contrast with the pentode amplifier, there is no isolation between the input and output of each stage; thus it is not possible to analyze each stage separately as was done with the pentode amplifier. Alternatively, the circuit can be analyzed in a straightforward way by solving the appropriate node equations; however, the results so obtained have a complicated algebraic form, and consequently provide little or no insight or guidance to aid in circuit design. Some understanding of the

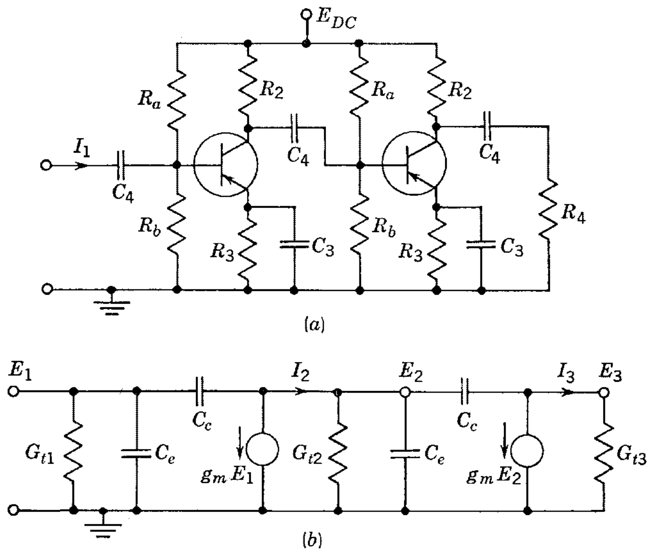


FIG. 15-5. *RC*-coupled transistor stages. (*a*) Circuit; (*b*) high-frequency model.

nature of the problem can be gained, however, by adapting the techniques developed in Secs. 14-3 and 14-6 to the analysis of the cascaded transistor amplifier.

The model for the second stage of the amplifier has the same form as the model for a single-stage amplifier shown in Fig. 14-19*b*; hence the effect of the collector capacitance in this stage can be accounted for by a Miller capacitance as shown in Fig. 14-20. When this is done, the second stage has the simplified representation shown in Fig. 15-6; the Miller capacitance is $C_2 = (1 + A_2)C_c$, where $A_2 = |E_3/E_2|$.

The model for the first stage of the amplifier differs from that for the second stage in that its load has a capacitive component as well as a resistive component; hence the Miller effect is slightly more complicated in the first stage. However, this stage has the same form as the triode amplifier shown in Fig. 14-10. It follows from this fact that the Miller effect in the first stage can be accounted for by a circuit of the form shown in Fig. 14-13. When this is done, the first stage has the simplified representation shown in Fig. 15-6. The capacitance C_c is neglected in comparison with C_e in this circuit; hence the impedance accounting for the Miller effect is, from Eq. (14-32),

$$Z = R_1 + \frac{1}{j\omega C_1} = \frac{1}{\omega_2 A_1 C_c} + \frac{1}{j\omega A_1 C_c} \tag{15-8}$$

where A_1 is the mid-band value of the ratio $|E_2/E_1|$ and

$$\omega_2 = \frac{G_{t2}}{(C_e + C_2)}$$

If the resistance R_1 in the circuit of Fig. 15-6 were negligible in comparison with the reactance $1/\omega C_1$, the simplified representation for the first stage would reduce to the same form as that for the second stage, and the process could be repeated in each stage for a cascade of any number of stages. Equation (15-8) shows, however, that $R_1 = 1/\omega C_1$ when

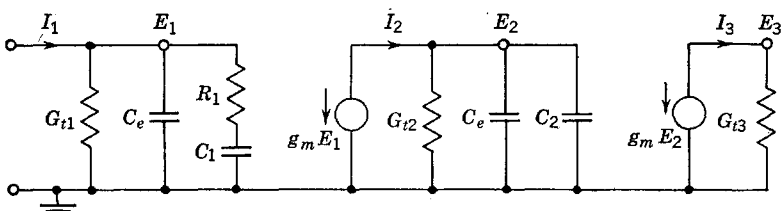


FIG. 15-6. Simplified high-frequency model for cascaded transistor stages.

$\omega = \omega_2$, where ω_2 is the half-power frequency for the output circuit of the first stage. Thus it is a rather crude approximation to neglect R_1 at frequencies up to ω_2. If R_1 is not neglected, and if the amplifier of Fig. 15-6 is preceded by yet another transistor stage, the Miller effect in this stage has a still more complicated form. Matters become progressively more complicated as more stages are cascaded. When the crude approximation of neglecting R_1 is acceptable, the representation for each stage in the cascade takes the simple form of one stage of a pentode amplifier, and the stage-by-stage analysis is simple.

Example 15-1. Two identical drift transistors having the incremental parameters $r_b' = 40$ ohms, $r_{be} = 1.6$ kilohms, $g_m = 37$ millimhos, $C_e = 200$ $\mu\mu$f, $C_c = 2$ $\mu\mu$f, and $g_o = 0$ are used in the circuit of Fig. 15-5a with $R_a = 180$ kilohms, $R_b = 20$ kilohms, $R_2 = 16$ kilohms, $R_3 = 2$ kilohms, and $R_4 = 2$ kilohms. Determine the parameter values for the simplified model shown in Fig. 15-6, and evaluate the forward current transmittance I_3/I_1.

Solution. The parameter values for the simplified model, retaining only two significant figures, are:

$$G_{t3} = \frac{1}{R_2} + \frac{1}{R_4} = \frac{1}{16} + \frac{1}{2} = 0.56 \text{ millimho}$$

$$A_2 = \frac{g_m}{G_{t3}} = \frac{37}{0.56} = 66$$

$$C_2 = (1 + A_2)C_c = (67)(2) = 134 \ \mu\mu\text{f}$$

$$G_{t2} = \frac{1}{R_2} + \frac{1}{R_a} + \frac{1}{R_b} + \frac{1}{r_{be}} = \frac{1}{16} + \frac{1}{180} + \frac{1}{20} + \frac{1}{1.6}$$
$$= 0.74 \text{ millimho}$$

$$\omega_2 = \frac{G_{t2}}{C_e + C_2} = \frac{(0.74)(10^{-3})}{(334)(10^{-12})} = (2.2)(10^6) \text{ rps}$$

$$A_1 = \frac{g_m}{G_{t2}} = \frac{37}{0.74} = 50$$

$$C_1 = A_1C_c = (50)(2) = 100 \ \mu\mu\text{f}$$

$$R_1 = \frac{1}{\omega_2 A_1 C_c} = \frac{1}{(2.2)(10^6)(50)(2)(10^{-12})} = 4500 \text{ ohms}$$

$$G_{t1} = \frac{1}{R_a} + \frac{1}{R_b} + \frac{1}{r_{be}} = 0.68 \text{ millimho}$$

Since $1/G_{t1} = 1/0.68 = 1.47$ kilohms, and since $R_1 = 4.5$ kilohms, the R_1C_1 branch in the simplified model can be treated as an open circuit to obtain a first approximation to the forward current transmittance. With this simplification,

$$A_c = \frac{I_3}{I_1} = \frac{I_2}{I_1}\frac{I_3}{I_2} = \frac{-g_mE_1}{I_1}\frac{-g_mE_2}{I_2}$$
$$= \frac{-g_m}{G_{t1} + j\omega C_e}\,\frac{-g_m}{G_{t2} + j\omega(C_e + C_2)}$$
$$= \frac{g_m^2}{G_{t1}G_{t2}}\,\frac{1}{(1 + j\omega C_e/G_{t1})[1 + j\omega(C_e + C_2)/G_{t2}]}$$

Substituting numerical values in this relation and expressing all frequencies in megaradians per second yields

$$A_c = 2700\,\frac{1}{(1 + j\omega/3.4)(1 + j\omega/2.2)}$$

Thus the high-frequency amplitude characteristic has two breaks, one at $3.4/2\pi = 0.54$ mcps and the other at $2.2/2\pi = 0.35$ mcps. The slope of the final asymptote is -12 db/octave.

When the R_1C_1 branch is not treated as an open circuit, the analysis yields

$$A_c = \frac{-g_mE_1}{I_1}\,\frac{-g_mE_2}{I_2}$$
$$= \frac{-g_m}{G_{t1} + j\omega C_e + \dfrac{j\omega C_1G_1}{G_1 + j\omega C_1}}\,\frac{-g_m}{G_{t2} + j\omega(C_e + C_2)}$$

A particular relation exists among the parameters of the circuit that simplifies the expression for A_c. From the definitions of R_1, C_1, and ω_2,

$$G_1 + j\omega C_1 = G_1\left(1 + \frac{j\omega C_1}{G_1}\right) = G_1\left(1 + \frac{j\omega}{\omega_2}\right)$$

and

$$G_{t2} + j\omega(C_e + C_2) = G_{t2}\left(1 + \frac{j\omega}{\omega_2}\right)$$

Thus the expression for A_c reduces to

$$A_c = \frac{g_m^2}{(G_{t1} + j\omega C_e)[G_{t2} + j\omega(C_e + C_2)] + j\omega C_1 G_{t2}}$$

By comparing this expression for A_c with the one obtained previously, it is seen that the effect of the R_1C_1 branch is accounted for entirely by the term $j\omega C_1 G_{t2}$. It is of interest to determine the extent to which the current transmittance of the amplifier under study is modified by this term. Replacing $j\omega$ by m and rearranging the denominator of A_c yields

$$A_c = \frac{g_m^2}{C_e(C_e + C_2)m^2 + [C_e G_{t2} + (C_e + C_2)G_{t1} + C_1 G_{t2}]m + G_{t1}G_{t2}}$$
$$= \frac{g_m^2}{C_e(C_e + C_2)} \frac{1}{m^2 + \left[\dfrac{G_{t2}}{C_e + C_2} + \dfrac{G_{t1}}{C_e} + \dfrac{C_1 G_{t2}}{C_e(C_e + C_2)}\right] m + \dfrac{G_{t1}G_{t2}}{C_e(C_e + C_2)}}$$

Substituting numerical values for the parameters in this equation with all frequencies expressed in megaradians per second yields

$$A_c = 2(10^4) \frac{1}{m^2 + 6.7m + 7.5}$$

The zeros of the denominator are $m_1 = -1.4$ and $m_2 = -5.3$; hence

$$A_c = (2)(10)^4 \frac{1}{(m + 1.4)(m + 5.3)}$$

Replacing m with $j\omega$ and rearranging yields

$$A_c = 2700 \frac{1}{(1 + j\omega/1.4)(1 + j\omega/5.3)}$$

The amplitude characteristic associated with this expression has breaks at $1.4/2\pi = 0.22$ mcps and at $5.3/2\pi = 0.83$ mcps. It follows from a comparison of these results with those obtained by the approximate analysis that the R_1C_1 branch reduces the lower break frequency and increases the upper break frequency.

The Miller effect reduces the bandwidth of transistor amplifiers, and it causes serious algebraic complications in the analysis and design of amplifiers having more than two stages. It is important, therefore, that the consequences of the Miller effect can be simplified and the bandwidth of the amplifier can be increased by the addition of a simple compensating network to the circuit of Fig. 15-6. The required compensating network is a series RL branch connected from node E_1 to ground as shown in Fig. 15-7a. If this network is adjusted so that

$$R = R_1 = \frac{(C_e + C_2)/C_e}{g_m} \tag{15-9}$$

and

$$\frac{R}{L} = \frac{1}{R_1 C_1} = \omega_2 \tag{15-10}$$

then the impedance of the two branches RL and R_1C_1 in parallel reduces to a constant resistance equal to R_1. In this case the input circuit of the amplifier reduces to the simple form shown in Fig. 15-7b, and the analysis of the circuit is considerably simplified. It follows at once that if each stage of a multistage amplifier is compensated in this way, the parallel combination of the Miller effect and the compensating branch in each stage is accounted for by a simple shunt resistance

$$R_1 = \frac{C_e/C_c}{g_m} \tag{15-11}$$

for the load on each stage consists of the simple parallel connection of a resistance and the capacitance C_e. This is an important result, for, when it is applicable, it permits each stage of a multistage amplifier to be designed independently of all the rest.

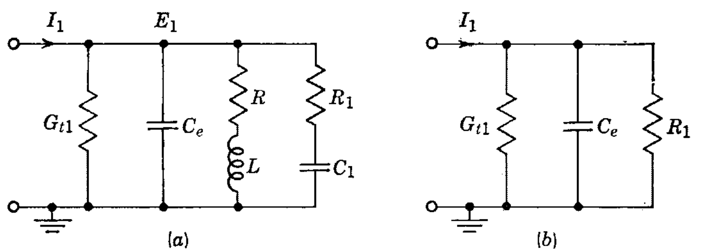


FIG. 15-7. Compensation of the Miller effect in a transistor amplifier. (a) Circuit; (b) equivalent circuit.

When the two branches RL and R_1C_1 in Fig. 15-7a are adjusted in accordance with Eqs. (15-9) and (15-10), they form a *constant-resistance* network. Networks having constant-resistance properties play an important role in the design of certain types of electric wave filters, for they make it possible to build complicated filters by designing and building simple sections that are subsequently connected in cascade.

A general analysis of the transistor amplifier of Fig. 15-5a at low frequencies is encumbered by unwieldy algebra. In some circumstances,

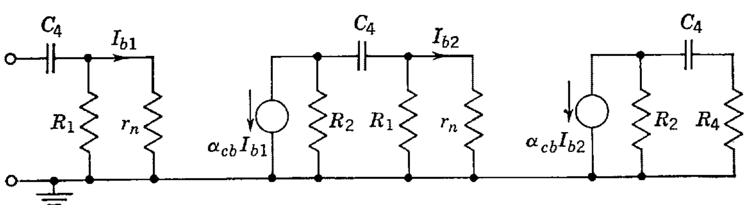


FIG. 15-8. RC-coupled transistor stages at low frequencies with no emitter impedance.

however, such amplifiers are operated without stabilization of the quiescent operating points. In such cases $R_3 = 0$, and the amplifier can be represented at low frequencies by the simple circuit shown in Fig. 15-8. The hybrid model for the transistor is used in this circuit with $\mu_{bc} = g_0 = 0$. The resistance R_1 designates the parallel combination of R_a and R_b.

The circuit of Fig. 15-8 can be analyzed stage by stage. Designating the resistance of r_n in parallel with R_1 by R_1', the forward current transmittance of one stage is

$$A_c = \frac{I_{b2}}{I_{b1}} = \frac{-\alpha_{cb}R_2}{R_2 + R_1' + 1/j\omega C_4} \frac{R_1}{R_1 + r_n}$$
$$= \frac{-\alpha_{cb}R_1R_2}{(R_2 + R_1')(R_1 + r_n)} \frac{j\omega C_4(R_2 + R_1')}{1 + j\omega C_4(R_2 + R_1')} \qquad (15\text{-}12)$$

Defining

$$A_m = \frac{\alpha_{cb}R_1R_2}{(R_2 + R_1')(R_1 + r_n)} \quad \text{and} \quad \omega_L = \frac{1}{C_4(R_2 + R_1')} \qquad (15\text{-}13)$$

yields

$$A_c = -A_m \frac{j\omega/\omega_L}{1 + j\omega/\omega_L} \qquad (15\text{-}14)$$

where A_m is the mid-band amplification, and ω_L is the low-frequency half-power point for the amplifier. This equation has the same form as (15-2) for the pentode amplifier at low frequencies.

If $R_3 \neq 0$ in the amplifier of Fig. 15-5*a*, a low-frequency model based on the circuit of Fig. 14-21*c* must be employed. Again, the amplifier can be analyzed stage by stage, and the current transmittance of one stage has the form

$$A_c = \frac{I_{b2}}{I_{b1}} = -A_m \frac{\omega_3}{\omega_2} \frac{(j\omega/\omega_1)(1 + j\omega/\omega_3)}{(1 + j\omega/\omega_1)(1 + j\omega/\omega_2)} \qquad (15\text{-}15)$$

Unfortunately, however, the break frequencies ω_1, ω_2, and ω_3 are complicated functions of the circuit parameters; hence (15-15) gives little guidance in the design of the amplifier.

15-4. *RC*-coupled Triode Amplifiers. Figure 15-9 shows the high-frequency model for two stages of a chain of cascaded triode amplifiers; C_{12} and C_{23} represent the grid-to-plate capacitances of the tubes. Since this circuit has the same form as the high-frequency model for the transistor amplifier of Fig. 15-5, the difficulties encountered in the analysis and design of *RC*-coupled triode amplifiers are similar to those discussed in connection with the transistor amplifier.

An equivalent high-frequency model is shown in Fig. 15-9*b*; in this representation, the Miller effect in each stage is accounted for by an appropriate shunt branch. The element values in these branches are indicated in Fig. 14-13 and the associated discussion. If C_3 consists of the small output capacitance of the last tube plus stray wiring capacitance, R_b may be negligibly small. However, since the output of the first stage is shunted by the large Miller capacitance C_b, the resistance R_a in the input circuit to the first stage is usually not negligible; in fact, $R_a = 1/\omega C_a$ at the half-power frequency for the first stage. Hence, if the amplifier consists of more than three stages, the analysis becomes complicated, as in the case of the transistor amplifier.

Triodes are usually not employed in amplifiers that must be designed for precise high-frequency response; thus the topic is not pursued further here. It is noted in passing, however, that the method presented in Sec. 15-3 of compensating for the Miller effect applies in theory to the

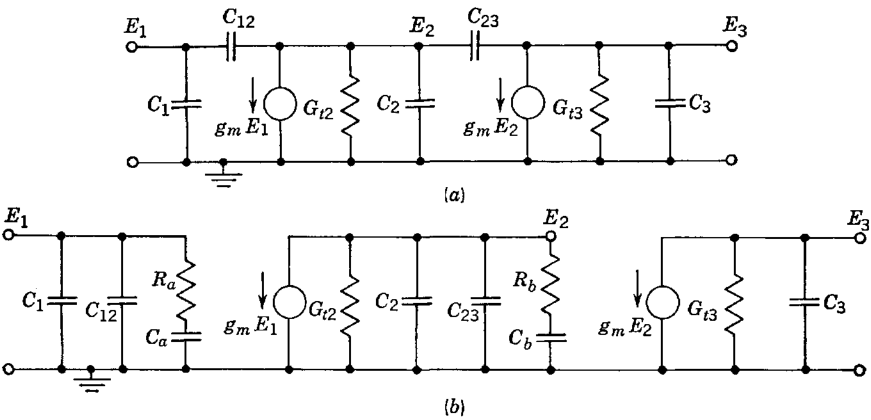


FIG. 15-9. RC-coupled triodes at high frequencies. (a) High-frequency model; (b) equivalent circuit.

triode amplifier. In order to obtain reasonable element values, however, it is necessary to modify the circuit in some way, such as by loading each stage with a suitably chosen shunt capacitance.

At low frequencies each stage of the RC coupled triode amplifier can be analyzed as a separate part; the low-frequency model for one stage is shown in Fig. 15-10. The general analysis of this circuit is complicated

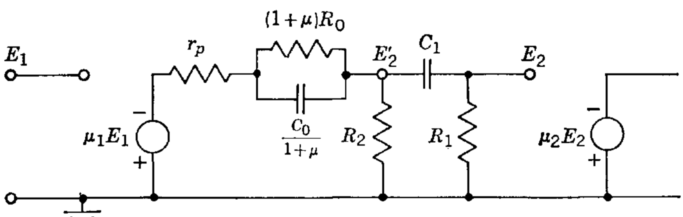


FIG. 15-10. Low-frequency model for RC-coupled triodes.

by the same features that complicate the low-frequency analysis of the transistor amplifier. In the case of the triode amplifier, however, the grid return resistance R_1 in Fig. 15-10 is often much larger than the plate load resistance R_2. Under these conditions, the current in R_1 is negligible in comparison with that in R_2, and the signal voltage E_2' at the plate of the tube can be evaluated with R_1 treated as an open circuit.

Since E_2 is related to E_2' by the voltage-divider relation

$$E_2 = \frac{R_1}{R_1 + 1/j\omega C_1} E_2' = \frac{j\omega/\omega_1}{1 + j\omega/\omega_1} E_2' \qquad (15\text{-}16)$$

where $\omega_1 = 1/C_1R_1$, it follows from Eq. (14-18) that

$$A_{vo} = \frac{-A_m}{1 + k_0} \frac{1 + j\omega/\omega_0}{1 + j\omega/(1 + k_0)\omega_0} \frac{j\omega/\omega_1}{1 + j\omega/\omega_1} \qquad (15\text{-}17)$$

The amplitude and phase characteristics, and the effects of the various circuit parameters on them, can be determined from this expression. The asymptotes for the low-frequency amplitude characteristic of a typical triode amplifier are shown in Fig. 15-11.

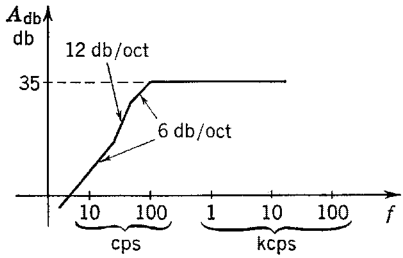


FIG. 15-11. Low-frequency amplitude characteristic for a typical high-gain RC-coupled triode stage.

Example 15-2. A certain two-stage RC-coupled amplifier uses a 12AU7 twin triode with one triode unit serving in each stage. The stages are identical, and the parameters for the high-frequency model shown in Fig. 15-9*a* are $g_m = 1.7$ millimhos, $G_{t2} = G_{t3} = 0.12$ millimho, $C_1 = C_2 = C_3 = 15$ $\mu\mu$f, and $C_{12} = C_{23} = 2$ $\mu\mu$f. Determine the parameters for the equivalent circuit of Fig. 15-9*b*.

Solution. The mid-band amplification for each stage is

$$A_m = \frac{g_m}{G_t} = \frac{1.7}{0.12} = 14$$

The branch accounting for the Miller effect in the last stage consists of

$$C_b = A_m C_{23} = (14)(2) = 28 \ \mu\mu\text{f}$$

and

$$R_b = \frac{1 + C_3/C_{23}}{g_m} = \frac{1 + 7.5}{1.7} = 5.0 \text{ kilohms}$$

Neglecting R_b as a rather crude approximation, the total capacitance shunting the output of the first stage is

$$C_t = C_2 + C_{23} + C_b = 15 + 2 + 28 = 45 \ \mu\mu\text{f}$$

Accordingly, the branch accounting for the Miller effect in the first stage consists of

$$C_a = A_m C_{12} = (14)(2) = 28 \ \mu\mu\text{f}$$

and

$$R_a = \frac{1 + C_t/C_{12}}{g_m} = \frac{1 + 22.5}{1.7} = 14 \text{ kilohms}$$

15-5. Over-all Characteristics of Multistage Amplifiers. The asymptotes for the amplitude and phase characteristics for one stage of an amplifier are shown in Fig. 15-3. The high-frequency portion of these characteristics is given by Eq. (15-1), and the low-frequency portion is given by Eq. (15-2). It follows that the complete characteristics are given by

$$A_{vo} = -A_m \frac{j\omega/\omega_L}{(1 + j\omega/\omega_L)(1 + j\omega/\omega_H)} \tag{15-18}$$

This equation can be obtained directly by analyzing the circuit in Fig. 15-1*b* without treating the low- and high-frequency ranges separately; however, the break frequencies ω_L and ω_H are in this case related to the circuit parameters in a somewhat complicated way.

If two identical stages like the one in Fig. 15-1*b* are connected in cascade, then the over-all forward voltage transmittance is given by

$$A_{vo} = A_m{}^2 \frac{(j\omega/\omega_L)^2}{(1 + j\omega/\omega_L)^2(1 + j\omega/\omega_H)^2} \tag{15-19}$$

The over-all amplitude and phase characteristics are the sums of the characteristics of the individual stages; the asymptotes for a typical amplitude characteristic are shown in Fig. 15-12*a*. The half-power frequencies, which are the frequencies at which the amplification is 3 db less than its mid-band value, do not correspond to the break frequencies in the case of two cascaded stages. In fact, the amplification at the break frequencies is 6 db below its mid-band value. It follows that the bandwidth between the half-power points decreases as more identical stages are connected in cascade.

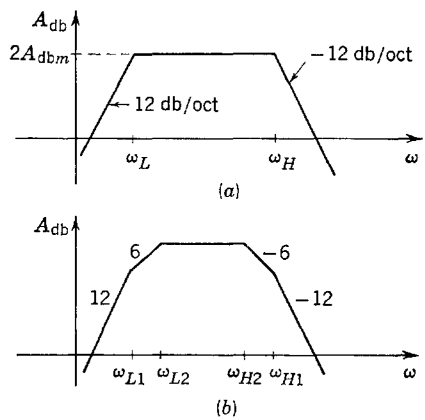


FIG. 15-12. Over-all amplitude characteristic for two stages in cascade. (*a*) Identical stages; (*b*) nonidentical stages.

If two nonidentical stages like that in Fig. 15-1*b* are connected in cascade, then the over-all forward voltage transmittance is given by

$$A_{vo} = A_{m1}A_{m2} \frac{(j\omega/\omega_{L1})(j\omega/\omega_{L2})}{(1 + j\omega/\omega_{L1})(1 + j\omega/\omega_{L2})(1 + j\omega/\omega_{H1})(1 + j\omega/\omega_{H2})} \tag{15-20}$$

The over-all amplitude and phase characteristics are obtained by adding the constituent characteristics; the asymptotes for a typical amplitude characteristic are shown in Fig. 15-12*b*.

15-6. Pole-zero Patterns. The amplitude and phase characteristics associated with the voltage and current transmittances provide a convenient and useful means for displaying the variations of these transmittances with frequency. The pole-zero patterns[2,3] associated with these transmittances display the same information in a different and more compact way. Certain important properties of the transmittances and of the networks which they describe can be perceived more easily from the pole-zero patterns than from the amplitude and phase characteristics, and as a general rule the pole-zero patterns are of greater fundamental importance than the frequency characteristics. However, given the frequency characteristics, the pole-zero pattern can be constructed, and vice versa.

The concept of the pole-zero pattern can be developed in terms of Eq. (15-18) for the voltage transmittance of a single-stage pentode amplifier. This equation can be written as

$$A_{vo}(j\omega) = K \frac{j\omega}{(j\omega + \omega_L)(j\omega + \omega_H)} \tag{15-21}$$

where $K = -A_m\omega_H$. For reasons set forth in Chap. 14, it is desirable to replace the symbol $j\omega$ with m and to define $m_L = -\omega_L$ and $m_H = -\omega_H$. Thus (15-21) can be written as

$$A_{vo}(m) = K \frac{m}{(m - m_L)(m - m_H)} \tag{15-22}$$

The steady-state sinusoidal voltage transmittance for any frequency ω is obtained from (15-22) by letting $m = j\omega$.

All the network functions encountered in the study of the dynamics of linear, lumped parameter systems, like the voltage transmittance given by (15-22), can be expressed as the ratio of two polynomials in the variable m. Functions of this type are called *rational functions.* The values of m that make such functions zero are called *zeros* of the function; hence $m = 0$ is a zero of the function given by (15-22). Similarly, the values of m that make rational functions infinite are called *poles* of the function; hence m_L and m_H are poles of the function given by (15-22). As indicated by the form of (15-22), if all the poles and zeros of a rational function are known, the function is completely specified except for a constant multiplier. For example, if the zeros of a certain voltage transmittance are -1 and -3, and if the poles are -2 and -4, then the transmittance is

$$A_{vo} = K \frac{(m + 1)(m + 3)}{(m + 2)(m + 4)} = K \frac{m^2 + 4m + 3}{m^2 + 6m + 8} \tag{15-23}$$

and for steady-state sinusoidal conditions

$$A_{vo} = K \frac{(j\omega + 1)(j\omega + 3)}{(j\omega + 2)(j\omega + 4)} = \frac{3K}{8} \frac{(1 + j\omega/1)(1 + j\omega/3)}{(1 + j\omega/2)(1 + j\omega/4)} \tag{15-24}$$

Thus if the poles and zeros of a rational function are known, the logarithmic amplitude and phase characteristics can be constructed, except for the effect of the constant multiplier. If the constant multiplier is not unity, it adds a constant number of decibels to the amplitude characteristic; if it is not positive, it adds a constant angle of 180° to the phase shift. (The constant angle is usually omitted from the phase characteristics.) It is also significant to note that the poles and zeros of A_{vo} correspond to the break frequencies expressed in radians per second. A zero corresponds to a break of 6 db/octave upward in the asymptotic characteristic, and a pole corresponds to a break of 6 db/octave downward.

The connection between the poles and zeros of the signal transmittance and the frequency characteristics is displayed in a useful way by the pole-zero diagram. As an example, Fig. 15-13*a* shows the pole-zero diagram for the voltage transmittance of a single-stage amplifier; this transmittance is given analytically by Eq. (15-22). The steady-state sinusoidal transmittance is obtained from (15-22) by letting $m = j\omega$. Hence, since m_L and m_H are real numbers, $A_{vo}(j\omega)$ contains both real

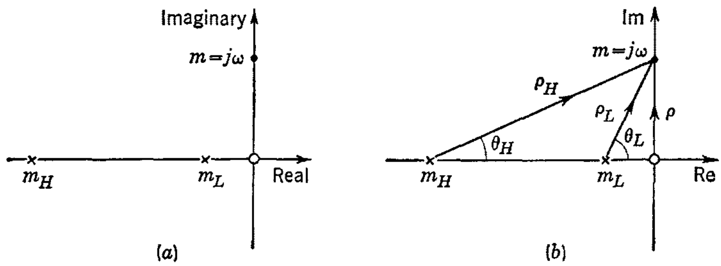


FIG. 15-13. Pole-zero pattern for Eq. (15-22). (*a*) Diagram; (*b*) interpretation.

and imaginary numbers, and it is desirable to plot the pole-zero diagram on a set of rectangular coordinates known as the plane of complex numbers or, simply, the complex plane. Thus the zero of A_{vo}, which occurs when $m = 0$, is designated by a circle at the origin of the complex plane. The poles of A_{vo}, which occur when

$$m = m_L = -\omega_L = -\frac{1}{C_3(R_1 + R_2)}$$

and when $$m = m_H = -\omega_H = -1/C_tR_t$$

are represented by crosses on the negative real axis. For steady-state sinusoidal operation, the variable $m = j\omega$ corresponds to a point on the imaginary axis.

The important relations between the poles and zeros of $A_{vo}(m)$ and the variable $m = j\omega$ are diagramed in Fig. 15-13*b*. The imaginary number $m = j\omega$ has a magnitude ρ and an angle $\pi/2$; hence it can be expressed in polar form as

$$m = \rho\epsilon^{j\pi/2} \tag{15-25}$$

As shown in Fig. 15-13*b*, the magnitude ρ is the length of a radius vector from the origin of the complex plane to the point $m = j\omega$ on the imaginary axis. Similarly, the two factors in the denominator of (15-22) are complex numbers having magnitudes that can be designated ρ_L and ρ_H and angles that can be designated θ_L and θ_H; accordingly, they can be expressed in polar form as

$$m - m_L = \rho_L \epsilon^{j\theta_L} \qquad \text{and} \qquad m - m_H = \rho_H \epsilon^{j\theta_H} \qquad (15\text{-}26)$$

The magnitude and angle of each of these complex numbers are shown geometrically in Fig. 15-13*b*. As ω is varied from zero to infinity, the point m moves up the imaginary axis from the origin to infinity, and the magnitudes and angles of the vectors vary accordingly. The diagram in Fig. 15-13*b* is essentially the same as the diagrams presented in Figs. 14-2 and 14-6 to illustrate frequency dependence in connection with the amplitude and phase characteristics; it is simply the diagram of a set of complex numbers that vary with frequency.

The way in which the amplitude and phase of A_{vo} vary with frequency can be evaluated qualitatively by inspection of the pole-zero diagram. The exact relations can be formulated by noting first that

$$A_{vo}(j\omega) = |A_{vo}|\epsilon^{j\theta} = K \frac{j\omega}{(j\omega - m_L)(j\omega - m_H)} \qquad (15\text{-}27)$$

Substituting (15-25) and (15-26) with $m = j\omega$ into (15-27) yields

$$|A_{vo}| = K \frac{\rho}{\rho_L \rho_H} \qquad \text{and} \qquad \theta = \frac{\pi}{2} - \theta_L - \theta_H \qquad (15\text{-}28)$$

where θ is the angle of $A_{vo}(j\omega)$ in radians. For any given frequency ω, the corresponding amplification and phase angle can be estimated by inspection or determined exactly by measurement of the diagram in Fig. 15-13*b*.

One of the more useful features of the pole-zero diagram is that in several important circumstances it shows that calculations can be simplified by neglecting certain factors for values of frequency lying in certain ranges, and it provides an estimate of the error resulting from the use of such approximations. This fact is used to advantage at several points in the sections that follow. The diagram of Fig. 15-13 illustrates this feature in terms of a familiar circuit, the *RC*-coupled amplifier. The high-frequency break in the amplitude characteristic of a typical *RC*-coupled amplifier is usually separated from the low-frequency break by two or three decades of frequency. Accordingly, the pole at m_H is usually 100 to 1000 times as far from the origin of the complex plane as the pole at m_L. Thus in the range of very low frequencies the vector ρ_H remains essentially constant, and the dependence of A_{vo} on frequency is determined entirely by the vectors ρ and ρ_L. In the intermediate range of frequencies, ρ_H is still essentially constant, and ρ and ρ_L are approximately equal; hence, according to (15-28), A_{vo} is independent of ω in the intermediate frequency range. Since it is not feasible to draw the pole-zero diagram to scale with $m_H = 100m_L$, the mid-band relations do not show clearly the diagram of Fig. 15-13*b*. At frequencies above the mid-band, ρ and ρ_L cancel, ρ_H becomes dependent on ω, and the high-frequency behavior of A_{vo} is determined entirely by ρ_H. It follows from this discussion that the upper half-power point occurs when $\theta_H = 45°$ and the lower half-power point occurs when $\theta_L = 45°$.

As another illustration of the way in which the pole-zero pattern displays the properties of networks, consider the high-frequency model of a triode amplifier shown in Fig. 14-10. The forward voltage transmittance of this circuit is given by (14-25); after replacing $j\omega$ with m and clearing of fractions, it has the form

$$A_{vo}(m) = K\frac{m - m_{12}}{m - m_2} \tag{15-29}$$

where $m_{12} = \omega_{12} = g_m/C_{12}$ and $m_2 = -\omega_2 = -(g_p + G_2)/(C_{12} + C_2)$. The pole-zero pattern for $A_{vo}(m)$ has the form shown in Fig. 15-14*a*. The amplitude and angle of $A_{vo}(j\omega)$ can be determined from the magnitudes and angles of the two vectors in this diagram along with the constant multiplier K in (15-29).

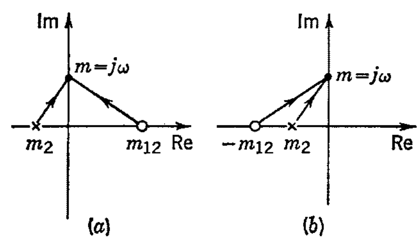


FIG. 15-14. Minimum- and non-minimum-phase-shift networks. (*a*) Pole-zero pattern for Eq. (15-29); (*b*) minimum-phase-shift counterpart of the pattern in (*a*).

An interesting condition results when the parameters in the circuit described by (15-29) are adjusted so that m_{12} is equal in magnitude to m_2. Under this condition the two vectors in Fig. 15-14*a* are of equal length for all values of ω, and $|A_{vo}(j\omega)|$ is constant at the value K for all ω. The phase shift is a function of ω, however. Since networks having this type of pole-zero pattern amplify (or attenuate) signals of all frequencies uniformly, they are called all-pass networks.

Another network property is illustrated by the two pole-zero patterns of Figs. 15-14*a* and *b*. The zero in Fig. 15-14*b* is the mirror image about the imaginary axis of the zero in Fig. 15-14*a*; otherwise the two patterns are identical. It follows that for any given value of ω, the two vectors in (*b*) are equal in length to their counterparts in (*a*); their angles are not the same, however. Therefore the amplitude characteristics associated with the two patterns are identical, but the phase characteristics are different. Thus, in general, a number of different pole-zero patterns can be associated with a given amplitude characteristic. These patterns differ from one another in that some or all of their zeros are mirror images about the imaginary axis. It is shown in Chap. 16 that poles in the right-hand half of the complex plane are associated with growing transients; hence all the poles of stable networks, which are the only networks of interest here, must lie in the left half of the plane. Of all the pole-zero patterns that can be associated with a given amplitude characteristic, the one having all its zeros in the left half plane produces the smallest change in phase angle as ω varies from zero to infinity. Hence it is the minimum-phase-shift pattern, and the corresponding network is the minimum-phase-shift network, associated with the given amplitude characteristic.

The design of networks to provide a desired frequency characteristic often begins with the choice of a pole-zero pattern that yields a suitable approximation to the desired performance. The design is then completed by finding a network configuration and the parameter values that yield the desired pole-zero pattern. When the desired pole-zero pattern is simple, having no more than two or three poles and zeros, experience often indicates a suitable network. When more complicated patterns are desired, more advanced methods must be employed. Much effort has been devoted to the development of systematic procedures for solving this problem of network synthesis.

Example 15-3. The pole-zero pattern for the voltage transmittance of a certain amplifier is shown in Fig. 15-15. The amplification at high frequencies is 40 db. Construct the asymptotes for the logarithmic amplitude characteristic.

Solution. The amplitude characteristic can be constructed by inspection of the pole-zero pattern. It is instructive, however, to go through the intermediate steps. The voltage transmittance is given by

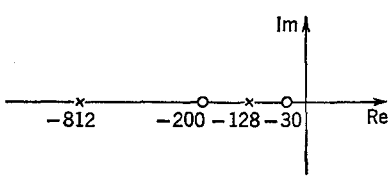


FIG. 15-15. Pole-zero pattern for a particular voltage transmittance.

$$A_{vo} = K \frac{(j\omega + 30)(j\omega + 200)}{(j\omega + 128)(j\omega + 812)}$$

For very large ω, this expression reduces to

$$A_{vo} = K$$

Therefore, since the amplification at high frequencies is 40 db, $K = 100$. Substituting this value in the equation for $A_{vo}(j\omega)$ and rearranging yields

$$A_{vo} = 5.77 \frac{(1 + j\omega/30)(1 + j\omega/200)}{(1 + j\omega/128)(1 + j\omega/812)}$$

This equation is identical with (14-69) for the transmittance of a pentode amplifier at low frequencies; the asymptotes for the logarithmic amplitude characteristic are shown in Fig. 14-18.

15-7. Single-tuned Pentode Amplifiers. The performance of the *RC*-coupled vacuum-tube and transistor amplifiers considered in the preceding sections is affected adversely at high frequencies by the parasitic capacitance shunting each stage. The shunt capacitance tends to short-circuit the signal at high frequencies, and it places an upper limit on the frequencies at which signals can be amplified satisfactorily. In many important applications it is possible to incorporate inductances in the circuit in such a way that the inductive effects cancel, in part, the capacitive effects over a band of frequencies; as a result, large amplifications can be obtained over a band of high frequencies. Since the amplification of such circuits is large only in a certain band of frequencies, these circuits can be used to amplify selectively signals in one band and to exclude signals at other frequencies. This selective amplification makes it possible for radio receivers to respond to signals from only one of the many stations broadcasting radio signals.

The single-tuned pentode amplifier[1] is the simplest amplifier of the type described above; an incremental model for one such amplifier stage is shown in Fig. 15-16. The circuit is similar to that for the RC-coupled pentode amplifier of Fig. 15-1 except that the plate load resistor is replaced by a parallel combination of L and C. The load impedance in the circuit of Fig. 15-16 is large at frequencies near the resonant frequency of the tuned circuit, and the amplification is correspondingly large at those frequencies. The band of frequencies over which the amplification is large depends on the Q of the circuit, and the location of this band in the frequency spectrum can be shifted by changing either L or C. Since this simple circuit illustrates certain important basic principles, it merits a detailed examination.

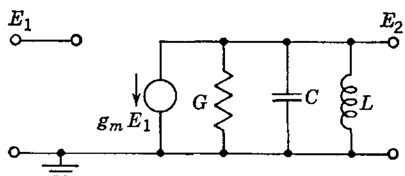


FIG. 15-16. A tuned pentode amplifier.

The voltage transmittance of the circuit in Fig. 15-16 is the impedance connected across the output terminals multiplied by the transconductance of the tube; hence it varies with frequency in the same way as the impedance of the parallel resonant circuit. Specifically,

$$A_{vo}(m) = \frac{-g_m}{G + mC + 1/mL} \tag{15-30}$$

When m is given the value $j\omega$, (15-30) gives the forward voltage transmittance for sinusoidal operation. Factoring C/m out of the denominator and denoting the resonant frequency by $\omega_o^2 = 1/LC$ yields

$$A_{vo}(m) = \frac{-g_m}{C} \frac{m}{m^2 + (G/C)m + \omega_o^2} \tag{15-31}$$

The denominator in (15-31) can be expressed as the product of two linear factors, giving

$$A_{vo}(m) = \frac{-g_m}{C} \frac{m}{(m - m_1)(m - m_2)} \tag{15-32}$$

The poles of A_{vo} are obtained from (15-31) with the aid of the quadratic formula; they are

$$m_1 = -\frac{G}{2C} + \sqrt{\left(\frac{G}{2C}\right)^2 - \omega_o^2}$$

and

$$m_2 = -\frac{G}{2C} - \sqrt{\left(\frac{G}{2C}\right)^2 - \omega_o^2} \tag{15-33}$$

These poles may be either real or complex, depending on the relative values of the circuit parameters. If they are real, (15-32) has the same form as Eq. (15-22) for the voltage transmittance of the *RC*-coupled amplifier; hence the pole-zero pattern and the frequency response of the tuned amplifier with real poles are the same as those for the *RC*-coupled amplifier. If the poles are complex, they form a conjugate pair, and they can be expressed as

$$m_1 = -\frac{G}{2C} + j\sqrt{\omega_o^2 - \left(\frac{G}{2C}\right)^2} = -\alpha_1 + j\beta_1 \tag{15-34}$$

and

$$m_2 = -\alpha_1 - j\beta_1 \tag{15-35}$$

where

$$\alpha_1 = \frac{G}{2C} \quad \text{and} \quad \beta_1 = \sqrt{\omega_o^2 - \left(\frac{G}{2C}\right)^2} \tag{15-36}$$

When the poles of $A_{vo}(m)$ are complex, its pole-zero pattern has the form shown in Fig. 15-17*a*. It follows from the definitions of α_1 and β_1 that

$$\beta_1^2 = \omega_o^2 - \alpha_1^2 \tag{15-37}$$

and hence that

$$\omega_o^2 = \alpha_1^2 + \beta_1^2 \tag{15-38}$$

Thus α_1 is the distance of m_1 from the imaginary axis, β_1 is its distance from the real axis, and ω_o is its distance from the origin of the complex plane.

Figure 15-17*b* illustrates the relations that exist in Eq. (15-32) when the resonant Q of the circuit is 5. The frequency dependence of $A_{vo}(j\omega)$ can be determined from the vectors shown in this figure, and, moreover, the geometry of the diagram shows how the pertinent relations can be simplified.[2,3] For Q's of 5 or more, the poles of A_{vo} are much closer to the imaginary axis than they are to the origin; hence as the variable point $m = j\omega$ moves along the imaginary axis in the vicinity of m_1, the factor $m - m_1$ experiences large variations, while the small variations in the factors m and $m - m_2$ tend to cancel. Thus when m is in the vicinity of m_1, $m \approx j\omega_o$ and $m - m_2 \approx j2\omega_o$. Substituting these values in (15-32) yields

$$A_{vo}(m) = -\frac{g_m}{2C}\frac{1}{m - m_1} \tag{15-39}$$

Thus the pole-zero diagram shows that for Q's greater than 5, the ratio $m/(m - m_2)$ can be treated as a constant, $\frac{1}{2}$, and the frequency characteristics can be determined from the simplified expression (15-39). The corresponding amplitude characteristic is shown in Fig. 15-17c.

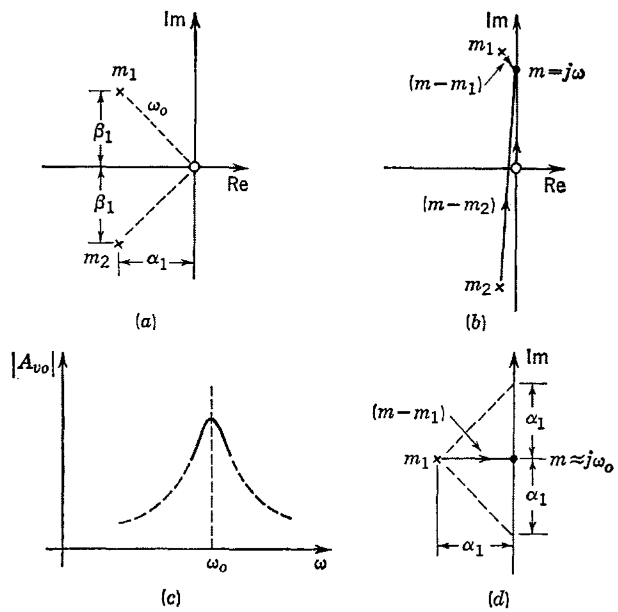


FIG. 15-17. Pole-zero patterns and resonance in tuned amplifiers. (a) Pole-zero pattern; (b) graphical representation of the factors in Eq. (15-32); (c) amplitude characteristic; (d) expanded view of the pole-zero pattern in the vicinity of m_1.

An expanded view of the region in the vicinity of m_1 showing how the factor $m - m_1$, and consequently A_{vo}, varies with frequency is given in Fig. 15-17d. The half-power frequencies are those at which $|A_{vo}|$ is reduced by the factor $\sqrt{2}$ from its maximum value; hence at the half-power frequencies the magnitude of $m - m_1$ is increased by $\sqrt{2}$ from its minimum value. It therefore follows from Fig. 15-17d that the half-power frequencies are $\omega = \omega_o + \alpha_1$ and $\omega = \omega_o - \alpha_1$. The bandwidth between the half-power points is, accordingly,

$$B = 2\alpha_1 = 2\frac{G}{2C} = \frac{G}{C} \qquad \text{radians/sec} \tag{15-40}$$

Thus the coefficient on the linear term in the denominator of (15-31) is the half-power bandwidth of the amplifier in radians per second. In addition, the constant term in the denominator of (15-31) is the square of the resonant frequency in radians per second. If the resonant Q of the circuit is defined as $Q_o = \omega_o C/G$, then

$$Q_o = \frac{\omega_o C}{G} = \frac{\omega_o}{B} = \frac{\omega_o}{2\alpha_1} \tag{15-41}$$

That is, the resonant Q is also the ratio of the resonant frequency to the half-power bandwidth.

The pole-zero pattern provides a compact way of displaying the effects of varying the circuit parameters on the characteristics of tuned amplifiers. For example, if L is increased while G and C are held constant, and if the poles are complex, it follows from Eqs. (15-36) that ω_o and β_1 decrease while α_1 remains constant. Thus the poles must move on a

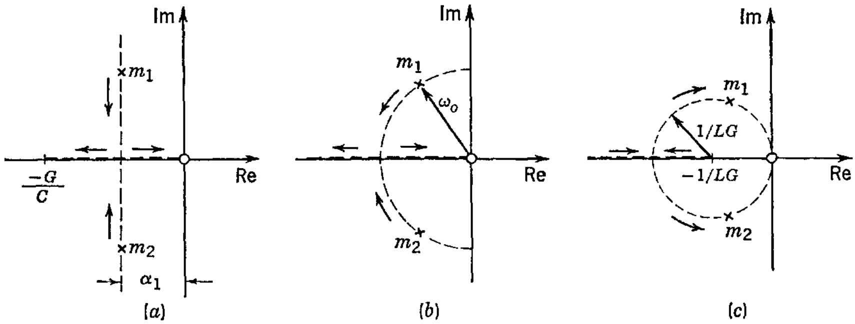


FIG. 15-18. The effect of variations in circuit parameters on the pole-zero pattern. (a) Increasing L; (b) increasing G; (c) increasing C.

path parallel to the imaginary axis as shown in Fig. 15-18a. If the poles are real, Eqs. (15-34) and (15-35) show that with increasing L one of them moves toward the origin and the other moves away from the origin; in the limit, as L tends to infinity, the pole approaching the origin cancels the zero there, and the pole receding from the origin approaches the point $-G/C$ on the axis of negative real numbers. In this limit the pole-zero pattern becomes that of a capacitance in parallel with a conductance, for the inductor behaves as an open circuit in the limit.

For one value of L the poles coincide on the negative real axis forming a double pole at that point. This condition, which occurs when the radical in Eq. (15-34) is zero, corresponds to critical damping of the resonant circuit; more is said about this last fact in Chap. 16. When L is greater than the value giving critical damping, the circuit is overdamped, and the poles lie on the negative real axis. Under these conditions the amplitude and phase characteristics of the tuned amplifier have the same form as those of the RC-coupled amplifier. When L is less than the critical value the circuit is underdamped, and the amplitude and phase characteristics may be quite different from those of the *RC* amplifier.

If the conductance G is increased while L and C are held constant, and if the poles are complex, then β_1 and α_1 vary while ω_o remains constant. As a consequence the poles move on the circular path shown in Fig. 15-18*b*. At a certain value of G the poles meet on the negative real axis, and if G is increased further, one pole moves toward the origin and the other moves toward infinity.

If the capacitance C is increased while L and G are held constant, and if the poles are complex, then β_1, α_1, and ω_o all vary; however, the path along which the poles move has a simple form. If the real part of m_1 is denoted by x and the imaginary part by y, then it follows from Eq. (15-34) that

$$x = -\frac{G}{2C} \qquad \text{and} \qquad y^2 = \omega_o{}^2 - \alpha_1{}^2 = \frac{1}{LC} - \alpha_1{}^2 \tag{15-42}$$

Eliminating C between these two expressions, adding and subtracting $(1/LG)^2$ to complete the square, and rearranging the terms leads to

$$y^2 + \left(x + \frac{1}{LG}\right)^2 = \left(\frac{1}{LG}\right)^2 \tag{15-43}$$

If L and G are held constant while C is varied, x and y, the real and imaginary parts of m_1, must vary in accordance with (15-43), which describes a circle with its center on the x (real) axis at $-1/LG$ and with a radius equal to $1/LG$. This circle is shown in Fig. 15-18*c*. When the poles are complex, m_2 is the conjugate of m_1 and hence moves on a similar circle. If C is made less than a certain value, the poles become real, and if C is reduced further, one pole recedes toward infinity and the other approaches the center of the circular path at $-1/LG$. When C becomes zero, the pole-zero pattern becomes that of an inductance in parallel with a conductance.

There are many circumstances in which the results developed above and illustrated in Fig. 15-18 can be used to gain insight into circuit properties and to simplify the problems of analysis and design. Therefore it is desirable to state the results in a more general form. Any quadratic polynomial in m can be expressed as

$$am^2 + bm + c = a\left(m^2 + \frac{b}{a}m + \frac{c}{a}\right) \tag{15-44}$$

This polynomial has two zeros, or roots; if the coefficients a, b, and c are real numbers, the roots must either be real or they must occur in conjugate pairs. The locations of the roots in the complex plane depend on the values of the coefficients. If c is decreased while a and b are held constant, the roots move in the indicated direction along the paths shown in Fig. 15-18*a*; the distance of the vertical paths from the imaginary axis is $\alpha_1 = b/2a$. If b is increased while a and c are held constant, the roots move in the indicated direction along the paths shown in Fig. 15-18*b*; the radius of the circular portions of the paths is $\sqrt{c/a}$. If a is increased while b and c are held constant, the roots move in the indicated direction along the paths shown in Fig. 15-18*c*; the radius of the circular portions of the paths is c/b, and the center of the circle is at $-c/b$. If two or more of the coefficients are varied simultaneously, the paths do not, in general, have simple forms; in such cases the paths are combinations of the paths shown in Fig. 15-18. The paths along which the roots move are usually called *loci*, and the procedures by which they are constructed are known as *root-locus techniques*. Generalized root-locus techniques for higher-order polynomials have been developed to a rather high degree, and they serve as valuable guides in the analysis and design of complicated circuits.

The design of tuned amplifiers such as the one analyzed in the preceding paragraphs is concerned largely with choosing the circuit parameters so that the desired signals are amplified with suitable uniformity and so that undesired signals are rejected to the greatest possible extent. Thus the parameters must be chosen to give suitable values of ω_o and Q_o. For example, the signals transmitted by standard AM broadcasting stations lie in the frequency range between 535 and 1605 kcps, and the signal from each station occupies a band 10 kcps wide. If the signal to be amplified is centered at 1 mcps, and if the signal is to lie within the half-power bandwidth of the amplifier, then the circuit parameters must be chosen to give resonance at 1 mcps and to give a bandwidth of 10 kcps. The pertinent relations are

$$\omega_o = \frac{1}{\sqrt{LC}} = (2\pi)(10^6) \qquad \text{rps} \tag{15-45}$$

and

$$B = \frac{G}{C} = (2\pi)(10^4) \qquad \text{rps} \tag{15-46}$$

Thus one of the three parameters can be chosen at will. The corresponding resonant Q is

$$Q_o = \frac{\omega_o}{B} = 100 \tag{15-47}$$

and the amplification at resonance is, from (15-31),

$$A_o = |A_{vo}(j\omega_o)| = \frac{g_m}{G} = \frac{g_m}{\omega_o C} Q_o = g_m Q_o \sqrt{\frac{L}{C}} \tag{15-48}$$

It is interesting to note that the product of the resonant amplification and the half-power bandwidth is

$$A_oB = \frac{g_mQ_o}{\omega_oC}\frac{\omega_o}{Q_o} = \frac{g_m}{C} \tag{15-49}$$

Thus the gain-bandwidth product for the tuned amplifier has the same form as that for the RC-coupled amplifier.

The pole-zero pattern for the voltage transmittance of a tuned amplifier and the associated amplitude characteristic are shown in Fig. 15-17. The pole-zero pattern in that case consists of a pair of complex poles and a zero at the origin of the complex plane. Since complex poles associated with electric circuits always occur in conjugate pairs, it is of interest to

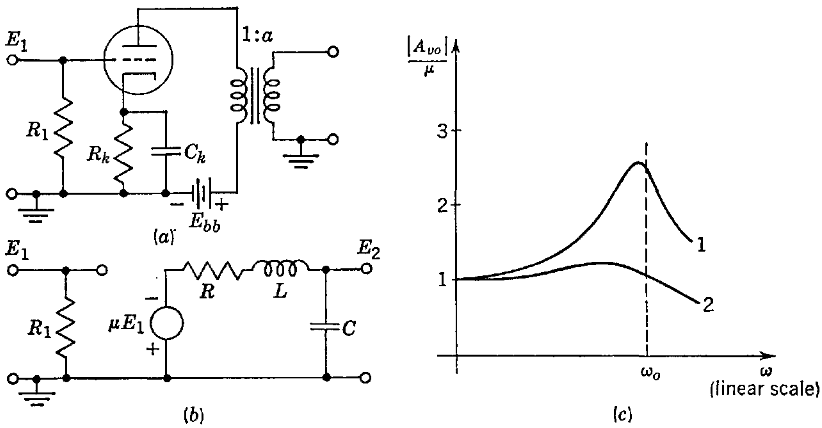


FIG. 15-19. Dynamic behavior associated with a pair of poles. (a) Circuit; (b) high-frequency model; (c) amplitude characteristic.

examine the properties of a single pair of such poles. Moreover, signal transmittances characterized by a single pair of complex poles are encountered in a number of important circuits. One stage of a cascade of transformer-coupled voltage amplifiers, shown in Fig. 15-19*a*, serves as a simple example. This amplifier can be represented approximately in the middle- and high-frequency ranges by the incremental model shown in Fig. 15-19*b*; the resistance R represents the plate resistance of the tube and the transformer winding resistance, L is the leakage inductance of the transformer, and C accounts for the parasitic capacitances of the transformer and the stage that follows.[1] The voltage transmittance of the high-frequency model is

$$A_{vo}(m) = -\frac{\mu}{LC}\frac{1}{m^2 + (R/L)m + (1/LC)} \tag{15-50}$$

Defining $\omega_o^2 = 1/LC$ and expressing the quadratic in its factored form yields

$$A_{vo}(m) = -\mu\omega_o^2 \frac{1}{(m - m_1)(m - m_2)} \tag{15-51}$$

When the circuit parameters have values such that the poles are complex and relatively close to the imaginary axis, the amplitude characteristic has the form shown by curve 1 in Fig. 15-19*c*. The high resonant peak in this characteristic is usually undesirable in amplifiers of this type, for it results in an overemphasis of signals with frequencies near ω_o. If the resistance R is increased, the poles move away from the imaginary axis on a circular path with ω_o constant (Fig. 15-18*b*); as a result, the form of the amplitude characteristic may change to that shown by curve 2 in Fig. 15-19*c*. The resonant peak in the characteristic is greatly reduced; with further increases in R, the peak disappears altogether.

The pole-zero pattern for the voltage transmittance of the circuit in Fig. 15-19*b* is shown in Fig. 15-20*a*; it consists of a single pair of poles

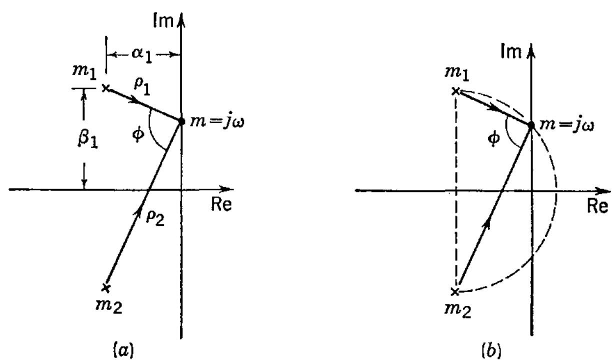


FIG. 15-20. Diagrams related to the amplitude characteristic of a pair of complex poles. (*a*) Pole-zero pattern; (*b*) resonant-peaking circle.

that are assumed to be complex. The amplitude characteristic can be constructed by determining the lengths of the vectors ρ_1 and ρ_2 as the variable point $m = j\omega$ moves along the imaginary axis. It is of considerable importance that the height of the peak in the amplitude characteristic and the frequency at which the peak occurs can be determined by a simple construction on the pole-zero diagram. The amplification, given by (15-51), is

$$|A_{vo}(j\omega)| = A = \frac{\mu\omega_o^2}{\rho_1\rho_2} \tag{15-52}$$

As the variable point moves along the imaginary axis, the area of the triangle m_1m_2m remains constant at the value

$$a = \alpha_1\beta_1 \tag{15-53}$$

This area is also given in terms of two sides and the included angle by

$$a = \frac{\rho_1\rho_2 \sin \phi}{2} \tag{15-54}$$

Thus

$$\rho_1\rho_2 = \frac{2a}{\sin\phi} = \frac{2\alpha_1\beta_1}{\sin\phi} \tag{15-55}$$

Substituting (15-55) into (15-52) yields for the amplification

$$A = \frac{\mu\omega_o^2}{2\alpha_1\beta_1}\sin\phi \tag{15-56}$$

The factor $\sin\phi$ is the only variable in this expression as long as the poles remain in a fixed location.

The amplification given by (15-56) has its maximum value when $\sin\phi$ has its maximum value, unity; this occurs when $\phi = 90°$. But when $\phi = 90°$, the triangle m_1m_2m can be inscribed in a semicircle as shown in Fig. 15-20*b*. Thus the point at which the semicircle intersects the imaginary axis gives the frequency of peak amplification ω_p. For $m = j\omega_p$, $\sin\phi = 1$, and the peak amplification is

$$A_p = \frac{\mu\omega_o^2}{2\alpha_1\beta_1} \tag{15-57}$$

If the amplification for $m = 0$ is designated A_o, and if the corresponding value of ϕ is designated ϕ_o, then

$$\frac{A_o}{A_p} = \sin\phi_o \tag{15-58}$$

Since the semicircle in Fig. 15-20*b* yields information about the resonant peak in the amplitude characteristic, it is referred to hereafter as the resonant-peaking circle.

If the resistance R in the circuit of Fig. 15-19*b* is increased so as to move the poles sufficiently far from the imaginary axis, the resonant-peaking circle does not intersect the imaginary axis, the amplification has its maximum value at $m = 0$, and there is no resonant peak in the amplitude characteristic. The maximum value of A under these conditions is less than the value of A_p given by (15-57). A case of special interest arises when R is adjusted so that the peaking circle is tangent to the imaginary axis. This is the minimum value that R can have without a peak appearing in the amplitude characteristic, and it results in the maximum bandwidth obtainable without a resonant peak. When the circuit is adjusted for this condition it is said to be maximally flat. It follows at once from Fig. 15-20 that this circuit is maximally flat when $\alpha_1 = \beta_1$. With the circuit adjusted for maximal flatness, the half-power point occurs when $\phi = 45°$. It is easy to show that under this condition the triangle m_1m_2m is inscribed in a semicircle centered at the origin of the complex plane; hence, with the maximally flat adjustment, the half-power frequency is ω_o.

The frequency of the resonant peak in the amplitude characteristic is given by the intersection of the peaking circle with the imaginary axis.

Since the radius of the peaking circle is β_1, it follows that

$$\omega_p^2 = \beta_1^2 - \alpha_1^2 = \omega_o^2 \left[1 - 2 \left(\frac{\alpha_1}{\omega_o} \right)^2 \right] \tag{15-59}$$

The quantity α_1/ω_o is customarily symbolized by ζ and is known as the damping ratio for the pair of poles.[3] If Eq. (15-41) is taken as the definition of the Q of a pair of poles, then

$$\zeta = \frac{\alpha_1}{\omega_o} = \frac{1}{2Q} \tag{15-60}$$

There are many circumstances in which the logarithmic amplitude and phase characteristics for a pair of complex poles are required. As in the case of the real poles discussed in Chap. 14, the asymptotic behavior of these characteristics is simple; however, the transition of the characteristics from one asymptote to the other is not as simple as in the case of the real pole. Nevertheless, the logarithmic amplitude characteristic can be constructed in a relatively easy manner with the aid of the asymptotes and the information provided by the peaking circle.

Under steady-state sinusoidal operating conditions, Eq. (15-50) becomes

$$A_{vo}(j\omega) = -\mu\omega_o^2 \frac{1}{(j\omega)^2 + (R/L)(j\omega) + \omega_o^2} \tag{15-61}$$

For very small values of ω this expression reduces to

$$A_{vo} = -\mu \tag{15-62}$$

and the amplification in decibels is

$$A_{\text{db}} = 20 \log \mu \tag{15-63}$$

Thus the low-frequency asymptote for the logarithmic amplitude characteristic is a horizontal line corresponding to the constant (20 log μ). For very large values of ω, the voltage transmittance reduces to

$$A_{vo} = \frac{\mu\omega_o^2}{\omega^2} \tag{15-64}$$

and the amplification in decibels is

$$A_{\text{db}} = 20 \log \mu - 40 \log \frac{\omega}{\omega_o} \tag{15-65}$$

The high-frequency asymptote is therefore a straight line with a slope of -40 db/decade, or -12 db/octave. It follows directly from (15-63) and (15-65) that the asymptotes intersect at $\omega = \omega_o$. These asymptotes are shown in Fig. 15-21.

The transition of the amplitude characteristic from the low- to the high-frequency asymptote depends on the location of the poles, and thus on the value of the parameter $\zeta = \alpha_1/\omega_o$. If the characteristic has a peak (ζ less than 0.707), two points on the characteristic in the transition region can be determined easily. It follows directly from (15-61) that for $\omega = \omega_o$,

$$|A_{vo}(j\omega_o)| = \frac{\mu\omega_o}{R/L} = \frac{\mu\omega_o}{2\alpha_1} = \frac{\mu}{2\zeta} \tag{15-66}$$

In addition, the peak in the curve occurs at the radian frequency ω_p given by (15-59), and the amplification at the peak is given by (15-57). These amplifications, when converted to decibels, permit the characteristic to be sketched with sufficient accuracy for many purposes.* The characteristic for $\zeta = 0.4$ is shown in Fig. 15-21.

If the value of ζ is greater than 0.707, the characteristic has no peak. Equation (15-66), which gives the amplification at ω_o, provides one point

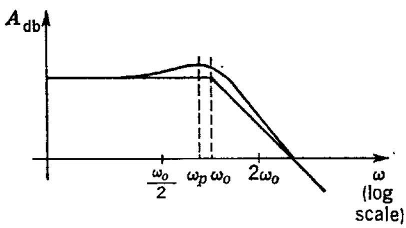


FIG. 15-21. Logarithmic amplitude characteristic for Eq. (15-61) with $\zeta = \alpha_1/\omega_o = 0.4$.

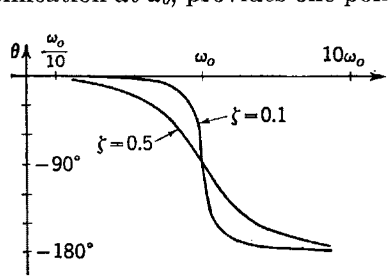


FIG. 15-22. Phase characteristic for Eq. (15-61). $\zeta = \alpha_1/\omega_o$.

on the characteristic. If the value of ζ is greater than unity, the poles are real, and they can be treated separately in accordance with the techniques presented in Chap. 14.

It is clear from Eq. (15-61) that, apart from the minus sign, the phase shift approaches 0° at very low frequencies and that it approaches $-180°$ at very high frequencies. In addition, the phase shift is $-90°$ at $\omega = \omega_o$. The details of the phase characteristic cannot be summarized in a simple form, however. Phase characteristics for two values of ζ are shown in Fig. 15-22.

Example 15-4. The performance of a particular transformer-coupled audio amplifier of the form shown in Fig. 15-19*a* is to be examined in the middle- and high-frequency ranges. The circuit can be represented by the high-frequency model shown in Fig. 15-19*b* with $\mu = 20$, $R = 8$ kilohms, $L = 0.4\ h$, and $C = 600\ \mu\mu f$.

* The author has recently been shown that an additional point on the amplitude characteristic can be obtained with no effort, for the high frequency at which the amplification is the same as at zero frequency is $\sqrt{2}\,\omega_p$. On a logarithmic scale this point is half an octave above ω_p. The geometric proof is simple but not obvious.

Solution. The voltage transmittance is given by (15-51); the pole-zero pattern has the form shown in Fig. 15-20. The pertinent dimensions of the pole-zero pattern are

$$\alpha_1 = \frac{R}{2L} = \frac{8000}{0.8} = 10{,}000$$

$$\omega_o = \frac{1}{\sqrt{LC}} = 64{,}500$$

$$\beta_1 = \sqrt{\omega_o^2 - \alpha_1^2} = \omega_o\sqrt{1 - (\alpha_1/\omega_o)^2} = \omega_o\sqrt{1 - 0.024}$$

Using the binomial expansion for the radical,

$$\beta_1 \approx \omega_o(1 - 0.012) = 0.988\omega_o = 63{,}800$$

Also of interest is the resonant frequency in cycles per second:

$$f_0 = \frac{\omega_o}{2\pi} = 10.3 \text{ kcps}$$

Since β_1 is larger than α_1, the amplitude characteristic has a resonant peak; the frequency of the peak is

$$\begin{aligned}\omega_p &= \omega_o\sqrt{1 - 2(\alpha_1/\omega_o)^2} = \omega_o\sqrt{1 - 0.048}\\ &= \omega_o(1 - 0.024) = 0.976\omega_o = 63{,}000 \text{ rps}\end{aligned}$$

and

$$f_p = \frac{\omega_p}{2\pi} \approx 10 \text{ kcps}$$

The amplification at the resonant peak is

$$\begin{aligned}A_p &= \frac{\mu\omega_o^2}{2\alpha_1\beta_1} = \frac{\mu\omega_o^2}{2\alpha_1(0.988\omega_o)} = \frac{64.5\ \mu}{(2)(10)(0.988)}\\ &= 3.26\ \mu = 65.4\end{aligned}$$

The peak amplification is 3.26 times the amplification in the mid-band, and the peak occurs well inside the band of audio frequencies.

The amplifier can be made maximally flat by increasing the equivalent resistance of the transformer. The value of R required for maximal flatness is determined from the fact that for maximal flatness,

$$\omega_o = \sqrt{2}\,\alpha_1 = \frac{R}{\sqrt{2}L}$$

Thus the required value of R is

$$R = \sqrt{2}\,L\omega_o = \sqrt{2L/C} = 36.4 \text{ kilohms}$$

If a resistance is added in series with the primary of the transformer to achieve maximal flatness, the low-frequency performance of the amplifier is affected adversely. It is customary to use high-resistance wire for the secondary winding of such transformers to realize a suitable equivalent transformer resistance without affecting the low-frequency performance.

15-8. Cascaded Tuned Amplifiers with Synchronous and Staggered Tuning. Tuned amplifiers can be connected in cascade by means of a coupling capacitor like that used in *RC* amplifiers. Ordinarily the coupling capacitor is a short circuit over the entire useful band of the amplifier. When two identical stages like that in Fig. 15-16 are cascaded, the over-all voltage transmittance is given by the square of

Eq. (15-31) or Eq. (15-32); hence the zero at the origin appears twice in the over-all transmittance, and the poles m_1 and m_2 each appear twice. The corresponding pole-zero diagram for the over-all transmittance has the form shown in Fig. 15-23*a*; the double symbols used in this diagram indicate double poles and zeros. The two stages in this amplifier are said to be synchronously tuned.

If the two cascaded stages are adjusted so that their poles are not identical, a better design can be obtained in the sense that for the same half-power bandwidth the amplitude characteristic is flatter in the middle of the passband and it drops more steeply at the edge of the passband. This characteristic gives more uniform amplification of the desired signal and better rejection of signals just outside the passband. A very important form of stagger tuning is illustrated in Fig. 15-23*b*. The poles m_1 and m_4 are associated with one of the stages, and m_2 and

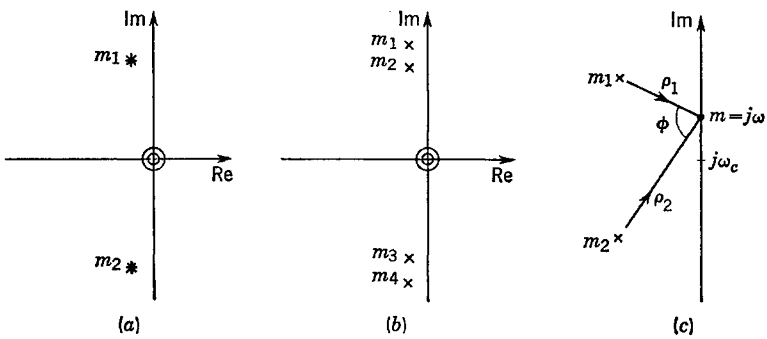


FIG. 15-23. Pole-zero patterns for two tuned amplifiers in cascade. (*a*) Synchronously tuned; (*b*) stagger tuned; (*c*) expanded view of the region near m_1 and m_2 in the pattern of (*b*).

m_3 are associated with the other stage; all four poles lie on a line parallel to the imaginary axis. The frequency characteristics of such stagger-tuned amplifiers depend rather critically on the spacing between the poles associated with the two stages; however, the important relations are revealed quite clearly by the pole-zero diagram.

The frequency characteristics of the stagger-tuned amplifier are determined by the set of vectors from the poles and zeros in Fig. 15-23*b* to the variable point m as m moves along the imaginary axis. In practically all cases the poles of the voltage transmittance are much closer to the imaginary axis than they are to the origin of the complex plane; hence as m moves along the axis in the vicinity of m_1 and m_2, which is the portion of the axis corresponding to the passband of the amplifier, the vectors from the zeros at the origin and from the poles at m_3 and m_4 remain essentially constant. Under these conditions the voltage transmittance of each stage is given by an equation having the form of

(15-39). Thus the transmittance of two stages in cascade is

$$A_{vo}(j\omega) = \left(\frac{g_m}{2C}\right)^2 \frac{1}{(j\omega - m_1)(j\omega - m_2)} \tag{15-67}$$

and the frequency characteristics in the passband of the amplifier are determined entirely by two poles, m_1 and m_2. This is another instance in which the pole-zero diagram indicates the strategic approximation to be made in simplifying a network problem; in this case a four-pole transmittance is reduced to a much simpler two-pole transmittance that is equivalent to the former for values of ω inside the passband of the amplifier.

An enlarged view of the region near the poles m_1 and m_2 is shown in Fig. 15-23*c*. The frequency ω_c, which is the center frequency of the passband, corresponds to a point on the imaginary axis equidistant from m_1 and m_2. The vectors ρ_1 and ρ_2 correspond to the factors $j\omega - m_1$ and $j\omega - m_2$ in Eq. (15-67). Equation (15-67) has the same form as (15-51), and if the point on the imaginary axis at ω_c is treated as the origin of the complex plane, the pole-zero pattern of Fig. 15-23*c* has the same form as the pattern in Fig. 15-20. It is clear, therefore, that the discussion related to the pole-zero pattern of Fig. 15-20 applies in the present case as well. In particular, the resonant-peaking circle shown in Fig. 15-24*a* yields important information about the amplitude characteristic of the amplifier. The amplification is

$$|A_{vo}(j\omega)| = \left(\frac{g_m}{2C}\right)^2 \frac{1}{\rho_1\rho_2} \tag{15-68}$$

and

$$\rho_1\rho_2 = \frac{\alpha_1\Delta}{\sin\phi} \tag{15-69}$$

Thus

$$|A_{vo}(j\omega)| = \left(\frac{g_m}{2C}\right)^2 \frac{\sin\phi}{\alpha_1\Delta} \tag{15-70}$$

in which $\sin\phi$ is the only factor depending on ω. If α_1 is less than $\Delta/2$, the peaking circle intersects the imaginary axis, and the amplitude characteristic has two peaks as shown in Fig. 15-24*b*, one corresponding to each intersection of the peaking circle with the axis. The frequencies at which the peaks occur are designated

$$\omega_p = \omega_c \pm \lambda \tag{15-71}$$

and since the radius of the peaking circle is $\Delta/2$,

$$\lambda^2 = \left(\frac{\Delta}{2}\right)^2 - \alpha_1^2 \tag{15-72}$$

$$\lambda = \frac{\Delta}{2}\sqrt{1 - \left(\frac{2\alpha_1}{\Delta}\right)^2} \tag{15-73}$$

The amplification at the peaks is

$$A_p = \left(\frac{g_m}{2C}\right)^2 \frac{1}{\alpha_1 \Delta} \tag{15-74}$$

and the amplification at the center frequency is

$$A_c = A_p \sin \phi_c \tag{15-75}$$

where ϕ_c is the value of ϕ when $\omega = \omega_c$. The peak-to-valley ratio for the amplitude characteristic is

$$\frac{A_p}{A_c} = \frac{1}{\sin \phi_c} \tag{15-76}$$

If the two stages of the amplifier are adjusted so that $\Delta/2$ is less than α_1, a single peak appears in the amplitude characteristic, and it occurs at $\omega = \omega_c$. If the stages are adjusted so that $\alpha_1 = \Delta/2$, the peaking

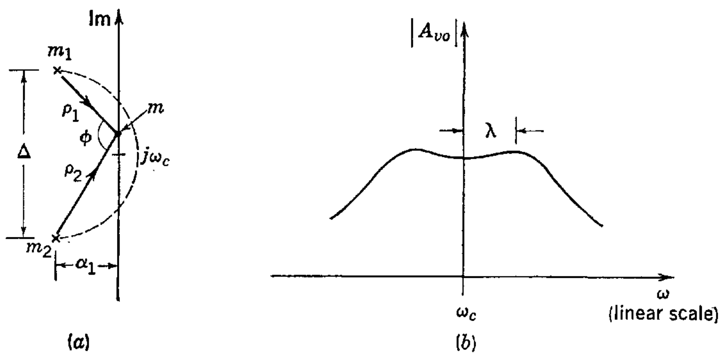


FIG. 15-24. Frequency characteristic for two stagger-tuned stages. (a) Geometric construction; (b) amplitude characteristic.

circle is tangent to the imaginary axis at $\omega = \omega_c$, the amplitude characteristic is maximally flat, and the stages are said to be flat-staggered. When the amplitude characteristic has two peaks, the stages are said to be overstaggered. Both flat-staggered and overstaggered adjustments are used in various applications.

When the amplitude characteristic is maximally flat,

$$\Delta = 2\alpha_1 \qquad \text{and} \qquad A_c = A_p = \left(\frac{g_m}{2C}\right)^2 \frac{1}{2\alpha_1^2} \tag{15-77}$$

In accordance with the discussion associated with Fig. 15-20, the bandwidth between half-power points for the maximally flat adjustment is twice the length of the vector from $j\omega_c$ to m_1; hence

$$B = 2\left[\left(\frac{\Delta}{2}\right)^2 + (\alpha_1)^2\right]^{1/2} = 2\sqrt{2}\,\alpha_1 \tag{15-78}$$

For comparison, the half-power bandwidth for one stage alone, given by (15-40), is $B = 2\alpha_1$.

The first step in the design of a pair of stagger-tuned stages is to choose locations for the poles m_1 and m_2 that provide the desired center frequency, the desired bandwidth, and a suitable peak-to-valley ratio, A_p/A_c. When these poles are located, the tuned circuit in each stage is designed to provide the required resonant frequency and resonant Q.

When three or more tuned stages are to be connected in cascade, similar considerations apply. In particular, when the bandwidth of the amplifier is small compared to the center frequency, the frequency characteristics are determined entirely by the poles of the signal transmittance lying near the segment of the imaginary axis corresponding to the passband. Thus, in the usual case, all the zeros and half of the poles of the transmittance can be ignored. The peaking circle does not apply, however, to pole-zero configurations having more than two poles. Nevertheless, it can be shown that for a maximally flat amplitude characteristic, the poles of the signal transmittance must be uniformly spaced on a semicircle having its center on the imaginary axis. The pole-zero pattern for a maximally flat three-stage amplifier is shown in Fig. 15-25. No matter how many stages are cascaded in this manner, the center frequency of the passband ω_c corresponds to the center of the circle on which the poles are located, and the half-power frequencies are the frequencies at which the circle intersects the imaginary axis. The more stages cascaded in this way, the flatter the amplitude characteristic in the passband and the steeper the characteristic at the edges of the band; the half-power bandwidth remains constant, however, if the radius of the circle of poles remains constant. This distribution of poles is known as the Butterworth configuration.[2,3]

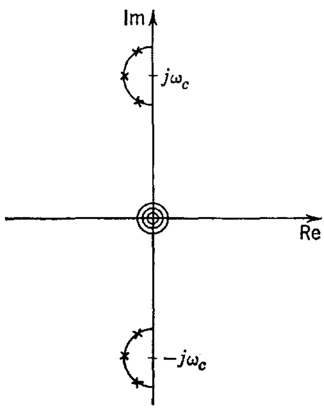


FIG. 15-25. Pole-zero pattern for a maximally flat three-stage tuned amplifier.

15-9. Double-tuned Amplifiers. The incremental model for a double-tuned amplifier is shown in Fig. 15-26a. A parallel LC combination is connected in the plate circuit just as in the case of the single-tuned amplifier, and the output is taken from a second tuned circuit that is inductively coupled to the first. The equivalent shunt resistances of the coils are included in G_1 and G_2. This double-tuned circuit permits the superior frequency characteristics of the stagger-tuned pair to be realized in a single stage; it is used in practically every vacuum-tube radio receiver. It should be noted that the mutual inductance is not essential in realizing this objective; however, it does aid in matching impedances, it simplifies the tuning procedure, and it blocks the d-c transmission between stages.

The forward voltage transmittance of the double-tuned amplifier can be obtained in the most convenient form for the study that follows by using the loop method of analysis. Thus, defining

$$Z_1 = \frac{1}{G_1 + mC_1} \qquad \text{and} \qquad Z_2 = \frac{1}{G_2 + mC_2} \tag{15-79}$$

where m is written for $j\omega$, and making a source conversion in Fig. 15-26a,

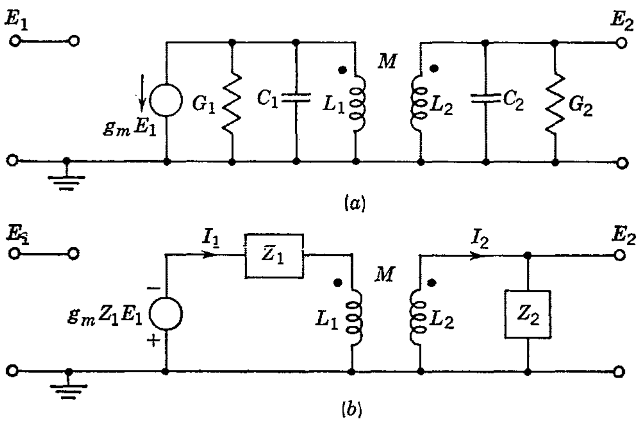


FIG. 15-26. The double-tuned amplifier. (a) Circuit; (b) an equivalent circuit.

the equivalent circuit of Fig. 15-26b is obtained. The loop equations for this circuit are

$$\begin{aligned} (Z_1 + mL_1)I_1 - mMI_2 &= -g_mZ_1E_1 \\ -mMI_1 + (Z_2 + mL_2)I_2 &= 0 \end{aligned} \tag{15-80}$$

Solving for I_2 yields

$$I_2 = \frac{-g_mZ_1mME_1}{(Z_1 + mL_1)(Z_2 + mL_2) - (mM)^2} \tag{15-81}$$

The output voltage is thus

$$E_2 = Z_2I_2 = \frac{-g_mZ_1Z_2mME_1}{(Z_1 + mL_1)(Z_2 + mL_2) - (mM)^2} \tag{15-82}$$

Substituting (15-79) into (15-82), simplifying, and solving for the voltage transmittance yields

$$A_{vo} = \frac{-g_mmM}{(m^2L_1C_1 + mL_1G_1 + 1)(m^2L_2C_2 + mL_2G_2 + 1) - (mM)^2(G_1 + mC_1)(G_2 + mC_2)} \tag{15-83}$$

The mutual inductance is related to the self-inductances of the coils by the relation $M^2 = k^2L_1L_2$, where k is the coefficient of coupling. Substituting this relation in the denominator of (15-83) and factoring $L_1L_2C_1C_2$ out of the denominator yields

$$A_{vo} = \frac{-g_m mM/L_1L_2C_1C_2}{\left(m^2 + \frac{G_1}{C_1}m + \frac{1}{L_1C_1}\right)\left(m^2 + \frac{G_2}{C_2}m + \frac{1}{L_2C_2}\right) - (mk)^2\left(m + \frac{G_1}{C_1}\right)\left(m + \frac{G_2}{C_2}\right)} \tag{15-84}$$

The next step in the normal analysis procedure is to carry out the multiplications indicated in the denominator of (15-84) and to collect terms on like powers of m. The result is a rather cumbersome expression that gives little insight into the properties of the circuit. This difficulty can be avoided and a particularly useful expression for A_{vo} can be obtained, however, by making certain strategic approximations in (15-84) *before* the denominator is multiplied out. The pole-zero pattern for A_{vo} shows clearly which approximations are both permissible and useful. Figure 15-27*a* shows the pole-zero diagram for the signal transmittance of a typical double-tuned amplifier; it is similar to the pole-zero diagram for the stagger-tuned amplifier except that it has only one zero at the origin, a difference of little consequence. It is to be expected, therefore, that the simplifications employed in analyzing the stagger-tuned amplifier are applicable in the present case.

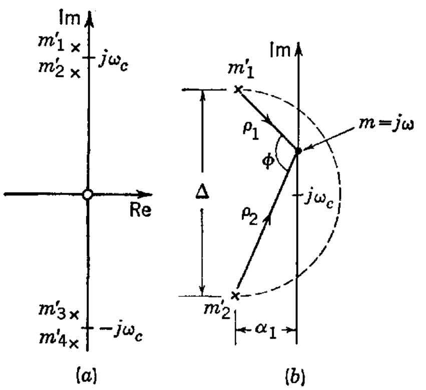


FIG. 15-27. Pole-zero patterns for the double-tuned amplifier. (*a*) Pole-zero pattern; (*b*) expanded view.

Sinusoidal operation in the passband of the amplifier corresponds to the variable m lying on the imaginary axis in the vicinity of the poles m_1' and m_2'. If the bandwidth of the amplifier is small compared with the center frequency, as usually is the case, then, for sinusoidal operation in the passband, $m = j\omega \approx j\omega_c$, where ω_c is a constant corresponding to the center frequency of the passband. Under this condition, several factors in the denominator of (15-84) can be treated as constants. The Q's of the primary and secondary circuits at the center frequency can be expressed as

$$Q_{c1} = \frac{\omega_c C_1}{G_1} \qquad \text{and} \qquad Q_{c2} = \frac{\omega_c C_2}{G_2} \tag{15-85}$$

Hence

$$\frac{G_1}{C_1} = \frac{\omega_c}{Q_{c1}} \qquad \text{and} \qquad \frac{G_2}{C_2} = \frac{\omega_c}{Q_{c2}} \tag{15-86}$$

Since Q_{c1} and Q_{c2} are normally of the order of 100, it follows that

$$m + \frac{G_1}{C_1} \approx j\omega_c + \frac{\omega_c}{Q_{c1}} \approx j\omega_c$$

and

$$m + \frac{G_2}{C_2} \approx j\omega_c + \frac{\omega_c}{Q_{c2}} \approx j\omega_c \tag{15-87}$$

Equations (15-87), together with the fact that $m \approx j\omega_c$, show that the second term in the denominator of (15-84) is approximately constant at the value $(j\omega_c)^4k^2$ for sinusoidal operation in the passband of the amplifier.

The product of quadratic factors in the denominator of (15-84) can be written in factored form as

$$[(m - m_1)(m - m_4)][(m - m_2)(m - m_3)] \tag{15-88}$$

in which m_1 and m_4 are complex conjugate numbers associated with the primary side of the coupled coils and m_2 and m_3 are complex conjugates associated with the secondary side. It follows from (15-84) that these are the poles of A_{vo} when the coefficient of coupling, k, is zero. As k is increased from zero, the poles move to new positions, m_1', m_2', m_3', and m_4', in a manner similar to that illustrated in Fig. 15-18; in normal circumstances they do not move far, however. If m_1 and m_2 designate the poles in the upper half plane, it follows from the pole-zero pattern in Fig. 15-27a that for sinusoidal operation in the passband $m \approx j\omega_c$ and the factors $m - m_3$ and $m - m_4$ are approximately constant at the value $2j\omega_c$.

When the simplifications developed above are applied to Eq. (15-84), the voltage transmittance of the double-tuned amplifier reduces to

$$A_{vo} = \frac{-g_m j\omega_c M}{L_1L_2C_1C_2} \frac{1}{(2j\omega_c)^2(m - m_1)(m - m_2) - (j\omega_c)^4k^2}$$

$$= \frac{-g_m M}{4j\omega_c L_1L_2C_1C_2} \frac{1}{(m - m_1)(m - m_2) + k^2\omega_c{}^2/4} \tag{15-89}$$

Defining

$$K = \frac{-g_m M}{4j\omega_c L_1L_2C_1C_2} \tag{15-90}$$

the transmittance can be written in the alternative forms:

$$A_{vo} = K \frac{1}{m^2 - (m_1 + m_2)m + m_1m_2 + k^2\omega_c{}^2/4} \tag{15-91}$$

$$= K \frac{1}{(m - m_1')(m - m_2')} \tag{15-92}$$

The quantities m_1 and m_2 are poles of A_{vo} when $k = 0$; m_1' and m_2' are poles of A_{vo} when $k \neq 0$. Since the frequency dependence of the amplifier in the passband is governed by a single pair of poles, it follows that the behavior of the amplifier in the passband can be examined with the aid of the diagram shown in Fig. 15-27*b*.

If the parameter values in the double-tuned amplifier are given, the poles m_1' and m_2' can be determined, and the frequency characteristics can be constructed. However, one of the more interesting features of this amplifier is the way in which the pole locations depend on the coefficient of coupling. To investigate this feature it is necessary to make a slight extension of the root-locus techniques presented in connection with Fig. 15-18. This extension is simplified by assuming that the primary and secondary circuits are adjusted separately to have the same resonant frequencies and the same resonant Q's. This adjustment is often used in practice, and it implies that $m_1 = m_2$. With $m_1 = m_2$, the denominator of (15-89) becomes

$$D = (m - m_1)^2 + \frac{k^2\omega_c^2}{4} \tag{15-93}$$

and the poles of A_{vo} are the values of the variable m that make this denominator zero. Thus if m' is a pole of A_{vo}, substituting it for m in (15-93) yields $D = 0$. Hence

$$(m' - m_1)^2 + \frac{k^2\omega_c^2}{4} = 0$$

and

$$m' - m_1 = \pm\sqrt{\frac{-k^2\omega_c^2}{4}} = \pm\frac{jk\omega_c}{2} \tag{15-94}$$

But $m' - m_1$ is a vector extending from m_1 to m' as shown in Fig. 15-28*a*. Thus there are two values of m',

$$m_1' = m_1 + \frac{jk\omega_c}{2} \qquad \text{and} \qquad m_2' = m_1 - \frac{jk\omega_c}{2} \tag{15-95}$$

that make $D = 0$. These are the poles of A_{vo} when $k \neq 0$. With $k = 0$, the pole-zero pattern for A_{vo} is the same as that for a cascade of two synchronously tuned stages; by adjustment of k, the pole-zero pattern for two stagger-tuned stages is obtained.

The root loci shown in Fig. 15-28*a* are based on the assumption of equal Q's and equal resonant frequencies in the two tuned circuits. It can be shown by an extension of the technique used above that if the poles m_1 and m_2 lie anywhere on the same vertical line, they move apart along that line with increasing k. Thus separating m_1 and m_2 on a vertical line is equivalent to starting with an initial value of k different from zero. If m_1 and m_2 lie on the same horizontal line, they move along that line toward each other with increasing k as shown in Fig. 15-28b. The poles coincide when $k = (m_2 - m_1)/2$. If k is increased beyond this value, the poles move apart along a vertical line. Thus separating m_1 and m_2 along a horizontal line is equivalent to starting with an initial value of k that is imaginary. If m_1 and m_2 do not lie on either the same vertical or the same horizontal line, the loci are not straight lines, and more advanced techniques are required for their construction.

The usual objective in the design of double-tuned amplifiers is to realize a pole-zero pattern like that shown in Fig. 15-27b. Thus the condition of equal Q's and equal resonant frequencies is ordinarily a satisfactory starting point. Since the separation of the poles in this case depends on the coefficient of coupling, it is clear that the tuning is flat-staggered, overstaggered, or understaggered, depending on the coefficient of coupling. The important relations in the flat-staggered and overstaggered cases are provided by the resonant-peaking circle shown in Fig. 15-27b. It follows directly from Fig. 15-28a that the separation of the poles is

$$\Delta = k\omega_c \qquad (15\text{-}96)$$

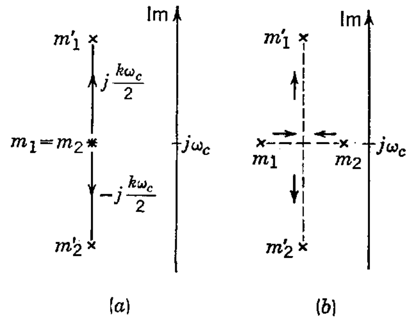


FIG. 15-28. Effect of the coefficient of coupling on the pole-zero pattern for a double-tuned amplifier. (a) Motion of poles for equal Q's and equal resonant frequencies; (b) motion of poles when m_1 and m_2 have equal imaginary parts.

Thus for a specified center frequency and pole separation, the coefficient of coupling is fixed. The condition for maximal flatness is $\Delta = 2\alpha_1$, where α_1 is the magnitude of the real part of m_1; the corresponding value of k, which is known as the critical coupling coefficient, is

$$k_{\text{crit}} = \frac{\Delta}{\omega_c} = \frac{2\alpha_1}{\omega_c} = \frac{1}{Q_{c1}} = \frac{1}{Q_{c2}} \qquad (15\text{-}97)$$

The voltage amplification is obtained from (15-92); with the aid of (15-69) it is put in the form

$$|A_{vo}(j\omega)| = |K| \frac{\sin \phi}{\alpha_1 \Delta} \qquad (15\text{-}98)$$

in which the value of K is given by (15-90). Substituting (15-96) into (15-90) and using the relations $M = k\sqrt{L_1 L_2}$ and $L_1 C_1 = L_2 C_2 = 1/\omega_c^2$ yields

$$|A_{vo}(j\omega)| = \tfrac{1}{2} g_m Q_{c1} \omega_c \sqrt{L_1 L_2} \sin \phi \qquad (15\text{-}99)$$

Typical amplitude characteristics for critical coupling, overcoupling, and undercoupling are shown in Fig. 15-29. When the circuit is overcoupled, the amplitude characteristic has two peaks, each corresponding to $\sin \phi = 1$ in (15-99). The amplification at the peaks is

$$A_p = \tfrac{1}{2} g_m Q_{c1} \omega_c \sqrt{L_1 L_2} \tag{15-100}$$

Thus the height of the peaks in the overcoupled characteristic is independent of k. The frequencies at which the peaks occur do depend on k, however. If these frequencies are designated

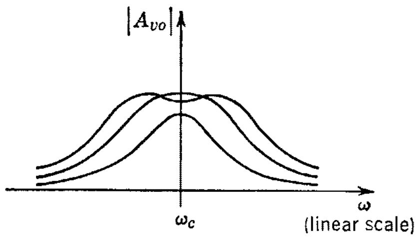


FIG. 15-29. Amplitude characteristic for a double-tuned amplifier with various degrees of coupling.

$$\omega_p = \omega_c \pm \lambda \tag{15-101}$$

then λ, given by (15-73), is

$$\lambda = \frac{\Delta}{2}\sqrt{1 - \left(\frac{2\alpha_1}{\Delta}\right)^2}$$

$$= \frac{k\omega_c}{2}\sqrt{1 - \left(\frac{1}{kQ_{c1}}\right)^2} \tag{15-102}$$

The amplification at the center frequency is

$$A_c = \tfrac{1}{2} g_m Q_{c1} \omega_c \sqrt{L_1 L_2} \sin \phi_c \tag{15-103}$$

where ϕ_c is the value of ϕ when $\omega = \omega_c$. This angle is given by

$$\tan \frac{\phi_c}{2} = \frac{\Delta/2}{\alpha_1} = \frac{k\omega_c}{2\alpha_1} = kQ_{c1} \tag{15-104}$$

The peak-to-valley ratio for the amplitude characteristic is

$$\frac{A_p}{A_c} = \frac{1}{\sin \phi_c} \tag{15-105}$$

When the circuit is critically coupled, the half-power bandwidth, as in the case of the stagger-tuned amplifier, is*

$$B = 2\sqrt{2}\,\alpha_1 = \sqrt{2}\,\frac{\omega_c}{Q_{c1}} \tag{15-106}$$

When several double-tuned stages are connected in cascade, the pole-zero patterns of the individual stages are adjusted to give the desired over-all frequency characteristic. For example, if a maximally flat over-all characteristic is desired, the individual stages are adjusted so that the poles of the over-all transmittance have the Butterworth configuration

* In accordance with the footnote in Sec. 15-7, when the circuit is overcoupled the frequencies outside the resonant peaks at which the amplification has the same value as at the center frequency are $\omega = \omega_c \pm \sqrt{2}\,\lambda$.

discussed in connection with Fig. 15-25. The half-power bandwidth is then equal to the diameter of the circle on which the poles lie, regardless of the number of stages. It should be noted that if there is an odd number of stages, one of the stages must be single-tuned to realize the Butterworth configuration.

Example 15-5. A double-tuned amplifier having the form shown in Fig. 15-26*a* is to be used as an intermediate-frequency amplifier in an AM broadcast receiver. The center frequency is to be 455 kcps, and the half-power bandwidth is to be 10 kcps. The two tuned circuits are to be resonant at ω_c when they are not coupled, and their resonant Q's are to be 100. Find the coefficient of coupling required, and if the circuit is overcoupled, find the ratio of peak to valley in the amplitude characteristic.

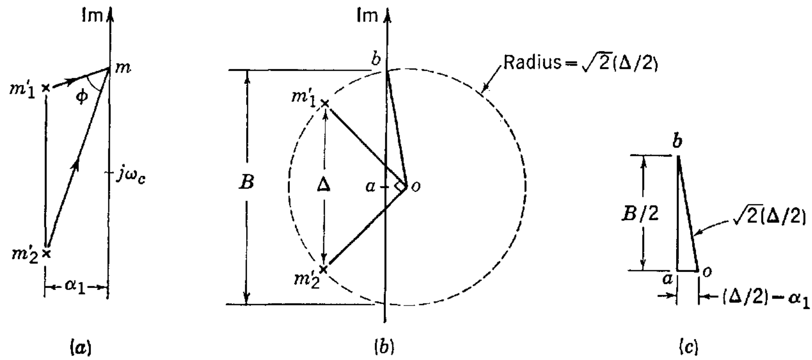


FIG. 15-30. Half-power-point relations for a pair of poles. (*a*) Diagram; (*b*) construction; (*c*) required relation.

Solution. The coefficient of coupling must be chosen to provide the proper separation of the poles of A_{vo}, and hence to provide the required bandwidth. Thus an expression relating the separation of the poles to the given data is needed. Equations (15-99) and (15-100) show that if the circuit is overcoupled, the amplification is $A_p/\sqrt{2}$ when $\phi = 45°$; the pertinent relations are pictured in Fig. 15-30*a*. If the triangle $m_1'm_2'm$ in Fig. 15-30*a* is inscribed in a circle, the line segment joining the poles subtends a central angle $2\phi = 90°$ as shown in Fig. 15-30*b*. It follows that the radius of the circle is $\sqrt{2}\,(\Delta/2)$ and that the center of the circle lies a distance $\Delta/2 - \alpha_1$ on the right of the imaginary axis. Fig. 15-30*b* contains a small triangle, *oab*, shown separately in Fig. 15-30*c*, that provides the desired relation,

$$\left(\frac{B}{2}\right)^2 = 2\left(\frac{\Delta}{2}\right)^2 - \left(\frac{\Delta}{2} - \alpha_1\right)^2$$

Solving this expression for Δ yields

$$\Delta = -2\alpha_1 \pm \sqrt{8\alpha_1^2 + B^2}$$

$$= -\frac{\omega_c}{Q_{c1}} \pm \sqrt{2\left(\frac{\omega_c}{Q_{c1}}\right)^2 + B^2}$$

Inserting numerical values in this expression and choosing the positive value of the root to obtain a positive value for Δ yields

$$\Delta = 46{,}500 \text{ rps}$$

The separation of the poles expressed in cycles per second is

$$\frac{\Delta}{2\pi} = 7400 \text{ cps}$$

The required coefficient of coupling is given by (15-96); it is

$$k = \frac{\Delta}{\omega_c} = \frac{46.5}{(2\pi)(455)} = 0.0163$$

This value is greater than $1/Q_{c1} = 0.01$; hence the circuit is overcoupled.

The ratio of peak to valley in the amplitude characteristic is

$$\frac{A_p}{A_c} = \frac{1}{\sin \phi_c}$$

and

$$\tan \frac{\phi_c}{2} = kQ_{c1} = (0.0163)(100) = 1.63$$

Thus $\phi_c = 117°$, and

$$\frac{A_p}{A_c} = \frac{1}{0.89} = 1.12$$

The displacement of the peaks in the amplitude characteristic from the center frequency is given by (15-102); it is

$$\begin{aligned} \lambda &= (23{,}250)\sqrt{1 - (1/1.63)^2} \\ &= 18{,}400 \text{ rps} \end{aligned}$$

The corresponding value in cycles per second is

$$\frac{\lambda}{2\pi} = 2930 \text{ cps}$$

15-10. High-frequency Compensation of *RC* Amplifiers. There are many applications for electronic amplifiers that require a uniform amplification over a very wide band of frequencies. For example, television and radar systems must amplify rectangular pulses having durations of the order of 1 μsec and occurring at a rate of a few tens of pulses per second. It follows from a consideration of the Fourier series for such pulse trains that to amplify them without excessive waveform distortion an amplifier must provide uniform amplification in the band of frequencies extending from a few tens of cycles per second to several megacycles per second.

The amplification of *RC*-coupled amplifiers is limited at low frequencies by the action of the coupling and bypass capacitors, and it is limited at high frequencies by the parasitic capacitance shunting the output. The high-frequency performance of vacuum-tube amplifiers can be improved by reducing the value of the plate load resistance and thereby increasing the high-frequency half-power point; however, there is an attendant reduction of amplification in accordance with the gain-bandwidth theorem. The high-frequency response can be extended somewhat without reducing the mid-band amplification through the use of more elaborate interstage coupling networks.

One simple method of extending the high-frequency response that is widely used in vacuum-tube amplifiers is to add an inductance of appropriate size in series with the plate load resistor as illustrated in the incremental model of Fig. 15-31. It is assumed in this model that the load resistance R is much smaller than the plate resistance of the tube so that the latter can be considered an open circuit in the band of frequencies where the amplifier is useful, and it is assumed that there is negligible parasitic capacitance associated with the coil. The inductive effects cancel in part the capacitive effects over a band of frequencies, thereby improving the high-frequency performance. This method of high-frequency compensation is known as shunt peaking.

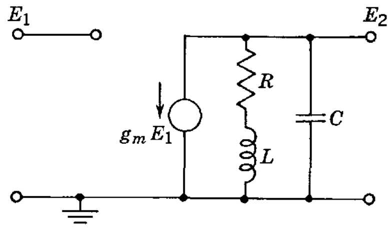


FIG. 15-31. A shunt-peaked amplifier.

The forward voltage transmittance of the shunt-peaked amplifier is

$$A_{vo} = -g_m \frac{(1/mC)(R + mL)}{R + mL + 1/mC}$$
$$= -\frac{g_m}{C} \frac{m + R/L}{m^2 + (R/L)m + 1/LC} \tag{15-107}$$

Equation (15-107) is put in a more useful form by defining:

$$\omega_o{}^2 = \frac{1}{LC}, \qquad Q_o = \frac{\omega_o L}{R}, \qquad \omega_H = \frac{1}{RC} \tag{15-108}$$

where ω_H is the high-frequency half-power point with no compensation, $L = 0$. It follows from the definitions in (15-108) that

$$\omega_H = Q_o \omega_o \qquad \text{and} \qquad \frac{R}{L} = \frac{\omega_H}{Q_o{}^2} \tag{15-109}$$

Substituting these relations in (15-107) yields

$$A_{vo} = -\frac{g_m}{C} \frac{m + \omega_H/Q_o{}^2}{m^2 + (\omega_H/Q_o{}^2)m + (\omega_H/Q_o)^2} \tag{15-110}$$

This expression can also be written as

$$A_{vo} = -\frac{g_m}{C} \frac{m - m_3}{(m - m_1)(m - m_2)} \tag{15-111}$$

where
$$m_3 = -\omega_3 = \frac{\omega_H}{Q_o{}^2} = \frac{R}{L} \tag{15-112}$$

and
$$m_1, m_2 = -\frac{\omega_H}{2Q_o{}^2} \pm j\sqrt{\left(\frac{\omega_H}{Q_o}\right)^2 - \left(\frac{\omega_H}{2Q_o{}^2}\right)^2}$$
$$= -\frac{\omega_H}{2Q_o{}^2} \pm \frac{j\omega_H}{2Q_o{}^2}\sqrt{4Q_o{}^2 - 1} \tag{15-113}$$

The pole-zero pattern for A_{vo} under typical conditions with complex poles is shown in Fig. 15-32a; it is clear from Eqs. (15-112) and (15-113) that the zero is equal to the sum of the poles and that therefore it is always twice as far from the imaginary axis as the poles when the latter are complex.

It follows from (15-110) that the poles of A_{vo} depend upon Q_o^2 in the same way that the zeros of Eq. (15-44) depend on a. Hence as Q_o is varied the poles move along paths like those shown in Fig. 15-18c. As shown in Fig. 15-32b, when the poles are complex they move on a circle of radius ω_H having its center on the negative real axis at $-\omega_H$. When $Q_o = \frac{1}{2}$, the poles coincide on the negative real axis, and for $Q_o = 1/\sqrt{2}$, they lie on lines radiating from the origin at angles of 45° with the negative real axis. For values of Q_o less than $\frac{1}{2}$, the poles are real, and as Q_o tends to zero as a result of decreasing L with R held constant, one pole

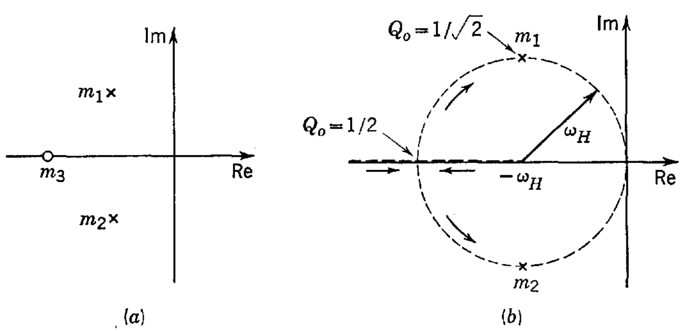


FIG. 15-32. Pole-zero diagram for the shunt-peaked amplifier. (a) Typical pattern; (b) motion of poles with increasing Q_o.

recedes toward infinity and the other approaches ω_H. With $L = 0$ the pole-zero pattern is that of an uncompensated amplifier.

The logarithmic amplitude and phase characteristics of the shunt-peaked amplifier can be constructed by adding the characteristics of the numerator and denominator of Eq. (15-110). The amplitude characteristic of the quadratic denominator is similar in form to the characteristic shown in Fig. 15-21; when it is combined with the characteristic of the numerator, the result has the form shown in Fig. 15-33. Two characteristics are shown in Fig. 15-33, one for no compensation, $Q_o = 0$, and the other for $Q_o = 1/\sqrt{2}$. When $Q_o = 1/\sqrt{2}$, the amplification at ω_H has the same value as the amplification in the mid-band, and the half-power frequency is almost $2\omega_H$.

The amplitude characteristic for $Q_o = 1/\sqrt{2}$ rises at high frequencies to values slightly above the mid-band value. It is clear from the way in which the poles move with changing Q_o, however, that the circuit can be adjusted for maximal flatness. The condition for maximal flatness when the voltage transmittance has a zero in addition to a pair of poles is somewhat different from the condition applying to a pair of poles alone. It can be shown, however, that $Q_o^2 = \sqrt{2} - 1$ produces maximal flatness in the shunt-peaked amplifier.

It is to be noted that shunt peaking does not change the ultimate high-frequency asymptote in any way; it simply changes the shape of the gain characteristic in the vicinity of ω_H and thereby extends the range of uniform amplification.

15-11. Summary. When vacuum-tube and transistor amplifiers are connected in cascade it is usually desirable to have no transmission through the interstage coupling network at zero frequency. The *RC* interstage network is the most widely used coupling network in low-frequency applications. The interstage network reduces the amplification to zero at zero frequency, and the parasitic capacitance reduces it to

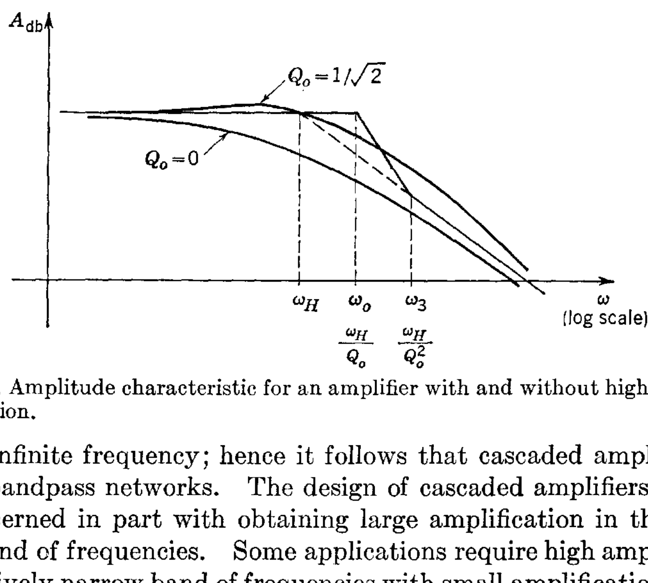


FIG. 15-33. Amplitude characteristic for an amplifier with and without high-frequency compensation.

zero at infinite frequency; hence it follows that cascaded amplifiers are usually bandpass networks. The design of cascaded amplifiers is therefore concerned in part with obtaining large amplification in the appropriate band of frequencies. Some applications require high amplification in a relatively narrow band of frequencies with small amplification outside this band, whereas other applications require uniform amplification over very wide bands of frequencies. More or less elaborate interstage networks may be used to obtain superior performance in these respects.

Tuned circuits are often used in vacuum-tube and transistor amplifiers in order to obtain large amplification at high frequencies; in these cases the parasitic capacitances become parts of the tuned circuits. The tuned amplifier usually has a relatively narrow passband, and it provides selective amplification of signals in this band to the exclusion of signals at other frequencies. It is this feature that makes it possible for a radio receiver to receive one signal while hundreds of transmitters are broadcasting simultaneously in different frequency bands. Tuned amplifiers are often provided with a variable capacitor or a variable inductor to permit the passband to be shifted from one point to another in the frequency spectrum.

When tuned amplifiers are to be cascaded they can all be of identical design, or the tuning of the individual stages can be staggered in various ways to obtain improved over-all frequency characteristics. Double-tuned circuits are widely used at radio frequencies to provide the advantages of stagger tuning in a single amplifier stage.

The amplitude and phase characteristics provide a useful way of displaying the frequency dependence of amplifiers. The pole-zero pattern provides an alternative way of displaying the same information that is in many respects more useful and more fundamental. The pole-zero pattern often shows in a simple way the effect of changes in circuit parameters on the behavoir of the circuit, and certain important relations that have discouragingly cumbersome algebraic forms prove to have simple geometric forms in terms of the poles and zeros of the network function. In the case of narrow-band circuits, in which the bandwidth is much smaller than the center frequency, the pole-zero diagram shows clearly how strategic simplifications can be made and leads therefore not only to a substantial reduction in the labor required in analysis and design, but also to a clearer understanding of the behavior of the circuit. For example, in the analysis of the single-tuned amplifier an examination of the pole-zero pattern leads to the reduction of a quadratic polynomial to a linear factor, and in the analysis of the double-tuned amplifier a quartic polynomial is reduced to a quadratic form. These simplifications eliminate a great deal of unenlightening algebraic labor.

REFERENCES

1. Gray, T. S.: "Applied Electronics," 2d ed., John Wiley & Sons, Inc., New York, 1954.
2. Guillemin, E. A.: "Introductory Circuit Theory," John Wiley & Sons, Inc., New York, 1953.
3. Van Valkenberg, M. E.: "Network Analysis," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1955.

PROBLEMS

15-1. Two 6SJ7 pentodes are used in an *RC*-coupled amplifier like the one shown in Fig. 15-1. The circuit parameters are $R_L = 250$ kilohms, $R_g = 1$ megohm, and $C_3 = 0.01$ μf. The tube parameters are $g_m = 1.2$ millimhos and $r_p = 1$ megohm; the total parasitic capacitance shunting each stage, including the stray wiring capacitance, is 15 $\mu\mu$f. The screen-grid and cathode bypass capacitors can be treated as short circuits in the frequency range of interest.

a. Sketch the dimension the asymptotic amplitude characteristic for one stage. Give the half-power frequencies in cycles per second.

b. Sketch and dimension the asymptotes for the phase characteristic of one stage.

15-2. The low-frequency behavior of one stage in a cascade of pentodes is to be studied with the aid of the representation shown in Fig. 15-4. The circuit parameters are $R_0 = 750$ ohms, $R_2 = 150$ kilohms, $R_3 = 1$ megohm, $R_4 = 50$ kilohms, $C_0 = 5\ \mu$f, and $C_2 = 0.05\ \mu$f. The tube parameters yield $g_{41} = 3$ millimhos, $k_0 = 3$, and $k_2 = 10$, where k_0 and k_2 are the quantities defined in Sec. 14-5. Since R_3 is much larger than R_4, E_4' can be calculated under the assumption that the current in R_3 is negligible.

a. What value of C_3 is required to make $\omega_3 = 1/C_3R_3 = 2\pi(200)$ rps?

b. With C_3 at the value determined in part *a*, sketch and dimension the asymptotes for the amplitude characteristic in the low- and medium-frequency ranges. Give amplitudes in decibels and frequencies in cycles per second.

15-3. The high-frequency response of a two-stage cascade of drift transistors, shown in Fig. 15-5*a*, is to be studied. The transistor parameters are $r_b' = 50$ ohms, $r_{be} = 2$ kilohms, $g_m = 30$ millimhos, $C_e = 200\ \mu\mu$f, $C_c = 2\ \mu\mu$f, and $g_o = 0$. The circuit parameters are $R_a = 200$ kilohms, $R_b = 22$ kilohms, $R_2 = 15$ kilohms, $R_3 = 0$ (no stabilization required), $R_4 = 2$ kilohms, and $C_4 = 1\ \mu$f.

a. Determine the parameter values in a simplified high-frequency model (Fig. 15-6) for the amplifier.

b. Determine the current transmittance, I_3/I_1. Do not neglect the R_1C_1 branch in the simplified model.

c. Sketch and dimension the asymptotes for the high-frequency logarithmic amplitude characteristic. Give amplitudes in decibels and frequencies in cycles per second.

15-4. The Miller effect in the first stage of the amplifier in Example 15-1 is compensated by an *RL* branch having the form shown in Fig. 15-7.

a. Determine the values of R and L required.

b. Determine the current transmittance I_3/I_1 for the compensated amplifier.

c. Sketch and dimension the asymptotes for the logarithmic amplitude characteristic at medium and high frequencies.

d. Using the results obtained in Example 15-1, sketch and dimension the asymptotes for the amplitude characteristic of the uncompensated amplifier.

15-5. The low-frequency behavior of the amplifier described in Prob. 15-3 is to be examined.

a. Construct and dimension the asymptotes for the amplitude characteristic of one stage in the low and middle range of frequencies. *Note:* $\alpha_{cb} = g_m r_{be}$.

b. If the amplifier is modified by changing C_4 to make the low-frequency half-power point occur at 80 cps, what value of C_4 is required?

15-6. The two-stage triode amplifier shown in Fig. 15-34 uses one section of a 12AX7 in each stage. The tube parameters are approximately $g_m = 1$ millimho and

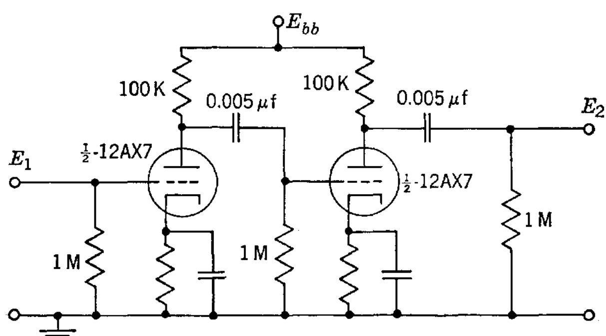


FIG. 15-34. Two-stage triode amplifier for Prob. 15-6.

g_p = 0.01 millimho. The grid-to-plate capacitance in each tube is 2 $\mu\mu$f, and the total capacitance shunting the output of each stage at high frequencies, including the stray wiring capacitance, is 10$\mu\mu$f. The cathode bypass capacitors can be treated as short circuits at all frequencies of interest. The objective of the problem is to study the frequency response of the amplifier.

a. Determine the parameter values for a high-frequency model of the form shown in Fig. 15-9*b*. As an approximation, R_b can be treated as a short circuit when evaluating R_a and C_a.

b. Neglecting R_b in the frequency range of interest, sketch and dimension the asymptotes for the logarithmic amplitude characteristic covering low, medium, and high frequencies. Give amplitudes in decibels and frequencies in cycles per second. *Note:* The amplitude characteristic is unaffected by the elements R_a and C_a in the high-frequency model.

15-7. The effect of poor wiring layout on the high-frequency performance of the pentode amplifier of Prob. 15-1 is to be examined. For this purpose, suppose that poor layout results in 5 $\mu\mu$f of capacitance between the plate and grid circuits in the second stage and that the remainder of the circuit parameters have the values given in Prob. 15-1.

a. Determine the mid-band amplification in decibels and the high-frequency half-power point in cycles per second for the second stage. Note that the grid-to-plate capacitance acts as part of the load on this stage [Eq. (14-22)].

b. The load on the first stage is to be accounted for by an equivalent circuit of the form shown in Fig. 15-9*b*. Determine the values of R_b and C_b for this circuit.

c. Neglecting the resistive part of the Miller effect (R_b = 0), determine the mid-band amplification in decibels and the high-frequency half-power point in cycles per second for the first stage.

d. Sketch and dimension the asymptotes for the amplitude characteristic of each stage. Comment briefly on the consequences of careless wiring layout.

15-8. Three pentodes, 6SJ7, 6AK5, and 6AC7, are to be compared with respect to high-frequency performance in voltage amplifiers.

a. Compare the figures of merit and the quiescent plate currents of the tubes. Assume the quiescent operating point corresponding to the largest value of g_m listed under typical operation in the tube manual.

b. Each tube is to be used in an *RC*-coupled amplifier having a bandwidth (upper half-power frequency) of 1.59 mcps. The stray wiring capacitance shunting the output of each amplifier is 5 $\mu\mu$f. Determine the mid-band voltage amplification for each amplifier.

15-9. Two stages of an *RC*-coupled pentode amplifier have the following mid-band amplifications and half-power frequencies. Stage 1: A_m = 100, f_L = 100 cps, and f_H = 100 kcps. Stage 2: A_m = 140, f_L = 200 cps, and f_H = 50 kcps. The bypass capacitors act as short circuits at all frequencies of interest.

Plot the over-all logarithmic amplitude characteristic. Use semilog-coordinate paper, and show both the asymptotes and the true characteristic. Calibrate the frequency scale in cycles per second.

15-10. A two-stage pentode amplifier having the form shown in Fig. 15-1 employs 6AK5 tubes. The cathode and screen-grid bypass capacitors act as short circuits at all frequencies of interest. The input and output capacitances of the tubes are 4.0 $\mu\mu$f and 2.8 $\mu\mu$f, respectively, and 10 $\mu\mu$f is to be allowed for stray wiring capacitance. The two stages are identical with upper and lower half-power frequencies at 500 kcps and 50 cps, respectively. The circuit parameters are to be chosen to meet these specifications.

a. Choose the resistances in the plate, cathode, and screen-grid circuits to locate the quiescent operating point at $E_b = 120$ volts, $E_{c2} = 120$ volts, and $I_b = 7.5$ ma when the supply voltage is 300 volts. Note that this is one of the typical operating points listed in the tube manual.

b. Choose the grid return resistance to give the desired high-frequency behavior. Choose the coupling capacitor to give the desired low-frequency behavior.

c. Sketch and dimension the over-all asymptotic amplitude characteristic for the two stages.

15-11. The incremental model for two transistor stages coupled by a transformer is shown in Fig. 15-35. The parasitic capacitances of the transistor are negligible in the useful frequency range of the amplifier. The coupling transformer is represented as an ideal transformer plus a leakage inductance L_s and a magnetizing inductance L_M (the parasitic transformer capacitances are assumed to be negligible). The

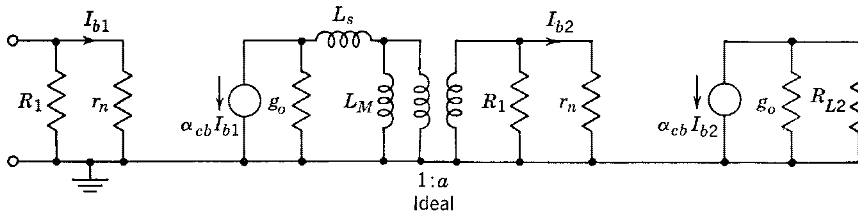


FIG. 15-35. Transformer-coupled transistor amplifier for Prob. 15-11.

leakage inductance acts as a short circuit at low and medium frequencies, and the magnetizing inductance acts as an open circuit at medium and high frequencies.

a. Give a simplified model that is valid at low and medium frequencies. Transfer the load resistance for the transformer to the primary side, and designate it by R_{L1}.

b. Repeat part *a* for the medium- and high-frequency ranges.

c. Show that the current transmittance for one transformer-coupled stage, I_{b2}/I_{b1}, has the same form as Eq. (15-2) at low frequencies and the same form as Eq. (15-1) at high frequencies.

d. Sketch the asymptotes of the amplitude characteristic for I_{b2}/I_{b1}. Give the mid-band amplification and the half-power frequencies in terms of the circuit parameters.

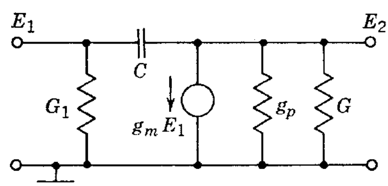


FIG. 15-36. Amplifier for Prob. 15-12.

15-12. The incremental model for a simple amplifier circuit is shown in Fig. 15-36. The circuit parameters are $g_m = 1.0$ millimho, $g_p = 0.1$ millimho, $G = 0.9$ millimho, and $C = 1.0\ \mu$f.

a. Sketch and dimension the pole-zero pattern for the voltage transmittance, E_2/E_1.

b. Sketch and dimension the asymptotes for the logarithmic amplitude and phase characteristics of the voltage transmittance.

15-13. The pole-zero patterns for the current transmittance of three transistor amplifiers are shown in Fig. 15-37.

a Which of these amplifiers will and which will not transmit d-c signals? Why?

b. Sketch the asymptotes for the amplitude characteristic corresponding to the pole-zero pattern of Fig. 15-37*b*. The amplification at zero frequency is 20 db, and $\omega_1 = 100$ rps. Give the coordinates of each break in the characteristic and the slope of each segment.

c. If the two amplifiers corresponding to Figs. 15-37*a* and *b* are connected in cascade, what is the pole-zero pattern of the over-all transmittance? The frequency ω_1 has the same value in the two amplifiers.

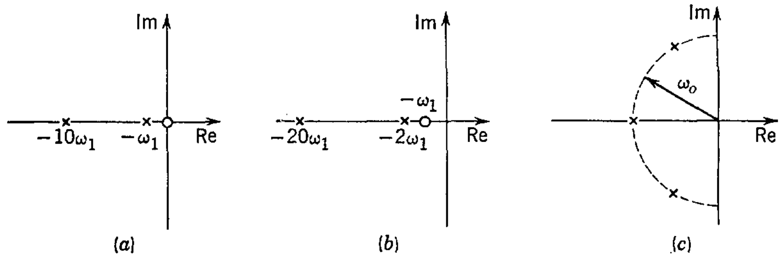


FIG. 15-37. Pole-zero patterns for Prob. 15-13.

15-14. The shunt branch R_2C_2 is added to the amplifier shown in Fig. 15-38 to improve the low-frequency response. The circuit parameters are $g_m = 4.0$ millimhos, $g_p = 0$, $R_1 = R_2 = 10$ kilohms, $C_2 = 5.0\ \mu f$, $C_3 = 0.1\ \mu f$, and $R_3 = 1$ megohm. The bypass capacitors can be treated as short circuits, and the loading effect of R_3 is negligible.

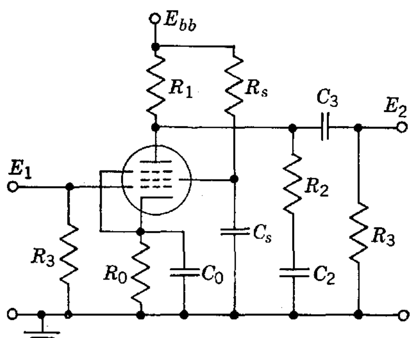


FIG. 15-38. Circuit for Prob. 15-14.

a. Give an incremental model for the circuit that is valid at low and medium frequencies.

b. Describe briefly in physical terms why the shunt branch composed of R_2 and C_2 improves the low-frequency response.

c. Using the fact that the loading by R_3 is negligible, evaluate A_{vo} as a function of frequency.

d. Sketch and dimension the pole-zero pattern for A_{vo}. Sketch and dimension the asymptotes of the amplitude characteristic.

15-15. The triode amplifier shown in Fig. 15-39 is to be designed to meet the following specifications: (1) $A_m = 40$, (2) $\omega_L = 300$ rps, and (3) $\omega_H = 80{,}000$ rps. The total parasitic capacitance is $C_1 + C_2 = 500\ \mu\mu f$, and C_s is an added capacitance that may be needed to permit the specifications to be met.

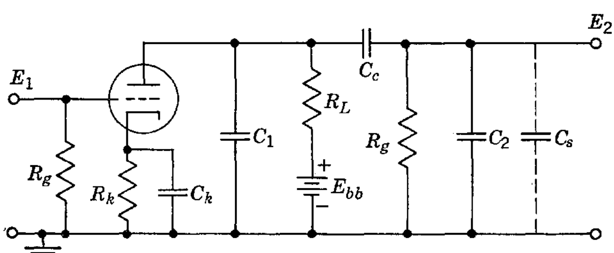


FIG. 15-39. Triode amplifier for Prob. 15-15.

a. Which of the following tubes can be used? Give reasons.

	Tube 1	Tube 2	Tube 3
g_m, millimhos	5.0	1.2	4.0
r_p, kilohms	8.0	70	15

b. The cathode bypass capacitor is to be treated as a short circuit, and R_g is to be 500 kilohms. For the tube chosen in part *a*, specify the values of R_L, C_c, and C_s required to meet the specifications.

15-16. A 6AU6 pentode tube is used in the tuned radio-frequency amplifier of Fig. 15-16. This amplifier is to amplify signals occupying a band of frequencies 10 kcps wide and centered at 800 kcps. The tube parameters are $g_m = 4.5$ millimhos and $g_p = 0.67$ micromho; the capacitance C is 300 $\mu\mu$f.

a. What value of L is required to give maximum amplification at 800 kcps?

b. What value of Q_0 is required for a half-power bandwidth of 10 kcps?

c. What must be the Q of the coil? *Note:* If the equivalent shunt conductance of the coil is G_L, the Q of the coil at ω_o is $Q = 1/\omega_o L G_L$. The total shunt conductance in Fig. 15-16 is $G = g_p + G_L$.

d. Sketch and dimension the pole-zero pattern for A_{vo}.

15-17. The voltage transmittance of a certain tuned amplifier is given by

$$A_{vo}(m) = \frac{20m}{m^2 + 2m + 25} = \frac{20m}{(m - m_1)(m - m_2)}$$

The relations between the pole-zero pattern for A_{vo} and the frequency characteristics of the amplifier are to be examined closely.

a. Plot the pole-zero pattern for A_{vo} on a sheet of graph paper. Let the imaginary axis extend from $-j10$ to $j10$.

b. For $m = j\omega_o = j5$, determine the magnitudes of the vectors m, $m - m_1$, and $m - m_2$. These magnitudes can be scaled off with dividers if desired. Calculate the voltage amplification for $m = j\omega_o$.

c Repeat part *b* for several other values of m on the imaginary axis between $j2$ and $j10$, and plot the curve of $|A_{vo}|$ versus ω.

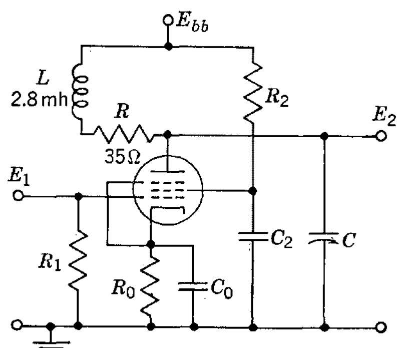


FIG. 15-40. Tuned pentode amplifier for Prob. 15-18.

15-18. The circuit in Fig. 15-40 is a tuned pentode voltage amplifier. The bypass capacitors C_0 and C_2 act as short circuits at all frequencies of interest, and parasitic capacitances are negligible in comparison with C. The tube parameters are $g_m =$ 5 millimhos and $r_p = 500$ kilohms.

a. Evaluate the forward voltage transmittance of the amplifier as a function of $j\omega$ in terms of the circuit parameters. Do not neglect the effect of r_p. This transmittance should have the form

$$A_{vo} = K \frac{(j\omega + a)}{(j\omega)^2 + b(j\omega) + d}$$

b. What value of C is required to make $\omega_o = 1/\sqrt{LC} = 10^6$ rps?

c. For the value of C found in part *b*, sketch the pole-zero pattern for A_{vo}. Give the value of each pole and zero.

d. Find the bandwidth between the half-power points, the resonant Q of the circuit, and the Q of the coil at $\omega = \omega_o$. Use Eq. (15-41), $Q_o = \omega_o/B$, as the definition for the resonant Q of the circuit.

15-19. The parameter values in a circuit of the form shown in Fig. 15-19*b* are $\mu = 10$, $R = 1$ kilohm, $L = 1$ henry, and $C = 1$ μf. The frequency response of the circuit is to be studied.

a. Sketch approximately to scale the pole-zero pattern for the voltage transmittance. Show the peaking circle on this sketch.

b. Determine the frequency at which the peak amplification occurs, the value of the peak amplification, the amplification at resonance $\omega = \omega_o$, and the amplification at zero frequency.

c. Make a reasonably accurate sketch of the logarithmic amplitude characteristic. Show both the asymptotes and the true characteristic.

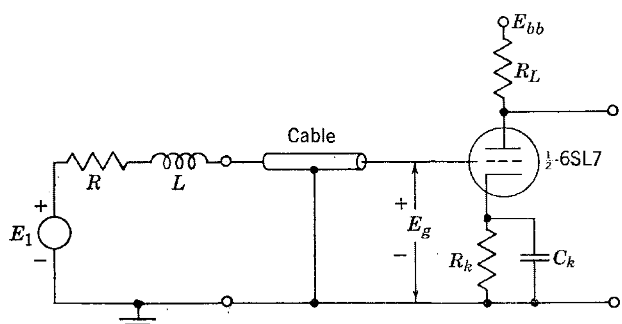


FIG. 15-41. Phonograph amplifier for Prob. 15-20.

15-20. The voltage source and the series RL branch in the circuit of Fig. 15-41 represent the electrical properties of a certain magnetic phonograph pickup. The pickup is connected to a triode amplifier through a 5-ft length of shielded cable that has a shunt capacitance of 70 $\mu\mu$f/ft. The tube and circuit parameters are $g_m = 1.4$ millimhos, $g_p = 0.02$ millimho, $R_L = 150$ kilohms, $C_{gk} = 5$ $\mu\mu$f, $C_{pk} = 5$ $\mu\mu$f, $C_{gp} = 4$ $\mu\mu$f, $R = 300$ ohms, and $L = 0.1$ henry. The bypass capacitor acts as a short circuit at all frequencies of interest. The amplitude characteristic of the voltage ratio E_g/E_1 is to be determined; for good performance this characteristic should be as nearly uniform as possible.

a. The input admittance to the triode can be represented by the Miller capacitance of Fig. 14-13 with the small resistance R neglected. When this is done, the pickup and the input circuit take the form shown in Fig. 15-19*b*. Determine the parameter values for this circuit. Do not neglect the capacitance of the cable.

b. Sketch the pole-zero pattern for the voltage ratio E_g/E_1. Show the peaking circle on this diagram.

c. Sketch the amplitude characteristic $|E_g/E_1|$ versus f. Give the frequency of the peak in cycles per second and the amplitude of the voltage ratio at the peak. Discuss briefly the significance of these results.

d. What value of resistance must be added in series with the pickup to produce a maximally flat characteristic? *Note:* The manufacturer of the pickup usually specifies a resistance to be connected in parallel with C to remove the resonant peak.

15-21. The incremental model for a two-stage amplifier is shown in Fig. 15-42. The behavior of the amplifier is affected at high frequencies by L and at low frequencies by C; however, the values of the circuit parameters are such that the half-power frequencies are well separated.

a. Determine the over-all voltage transmittance in the mid-band in terms of the circuit parameters.

b. Determine the upper and lower half-power frequencies in terms of the circuit parameters.

c. Sketch the pole-zero pattern for the voltage transmittance.

d. Sketch the asymptotes for the amplitude characteristic.

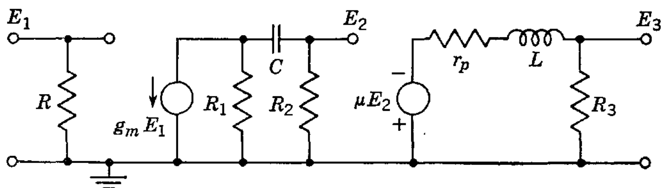


FIG. 15-42. Amplifier circuit for Prob. 15-21.

15-22. Two tuned pentode amplifiers like the one shown in Fig. 15-16 are connected in cascade, and their tuning is staggered to give an over-all transmittance having a pole-zero pattern like the one shown in Fig. 15-23*b*. The tuning is adjusted to give a center frequency f_c = 455 kcps, corresponding to $\omega_c = (2.85)(10^6)$ rps. The two poles in the upper half plane are separated vertically by $\Delta = (44.4)(10^3)$ rps, and the distance of each pole from the imaginary axis is $(11.1)(10^3)$ rps.

Sketch the amplitude characteristic $|A_{vo}|$ versus f. Give the frequencies in cycles per second at which the peaks in this characteristic occur. Determine the peak-to-valley ratio A_p/A_c.

15-23. Two tuned pentode amplifiers like the one shown in Fig. 15-16 are connected in cascade. The two stages are to be designed so that the center frequency of the passband is 100 kcps, and they are to be overstaggered so that the peaks in the amplitude characteristic are located 2.5 kcps above and below the center frequency. In addition, the peak-to-valley ratio for the amplitude characteristic is to be $\sqrt{2}$.

a. Determine the value of the angle ϕ in Fig. 15-24 for $m = j\omega_c$.

b. From the requirement on the location of the peaks in the amplitude characteristic, along with the results of part *a*, determine the locations of the poles of A_{vo}. Note that $\Delta/2 = \alpha_1 \tan (\phi_c/2)$.

c. Determine the resonant frequency and the resonant Q for each stage of the amplifier.

d. Determine the half-power frequencies for the stagger-tuned pair (Example 15-5). Note that there are more than two of these.

15-24. A double-tuned amplifier like the one shown in Fig. 15-26*a* is used as an intermediate-frequency amplifier in a radio receiver. The circuit is to be adjusted to give a center frequency of 455 kcps, the coefficient of coupling is to be adjusted for maximal flatness, and the bandwidth between the half-power points is to be 10 kcps.

The two tuned circuits are to be adjusted for equal resonant frequencies and equal resonant Q's when they are not coupled.

a. Show that the bandwidth is $B = 2\sqrt{2}\,\alpha_1 = \sqrt{2}\,\Delta$ when the amplifier is adjusted for maximal flatness (Example 15-5).

b. From the center frequency and the bandwidth requirements, determine the resonant frequencies and resonant Q's that the primary and secondary circuits must have.

c. Determine the coefficient of coupling required for maximal flatness.

15-25. The primary and secondary circuits in a double-tuned amplifier are tuned separately to resonance at 500 kcps when they are not coupled, and the resonant Q of each circuit is 50. The coefficient of coupling is adjusted to make the peak-to-valley ratio in the amplitude characteristic have the value $\sqrt{2}$.

a. Determine the locations of the poles of A_{vo} with these adjustments.

b. What coefficient of coupling is required?

c. At what frequencies do the peaks in the amplitude characteristic occur?

d. What are the half-power frequencies for the amplifier?

15-26. The pole-zero pattern for the signal transmittance of a certain amplifier is shown in Fig. 15-43. The amplification at zero frequency is 100.

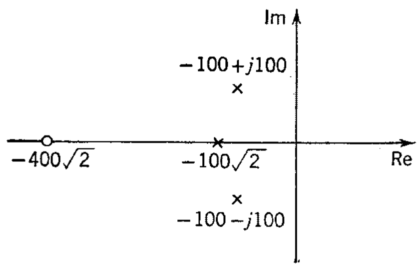


FIG. 15-43. Pole-zero pattern for Prob. 15-26.

a. Sketch the asymptotes for the logarithmic amplitude characteristic. Give the coordinates (amplitude in decibels and frequency in radians per second) of each break point, and give the slope of each section of the characteristic.

b. Determine directly from the pole-zero pattern the amplification and the phase shift of the amplifier for $m = j\omega = j100$.

15-27. The amplitude characteristic for a certain voltage amplifier is measured and found to have the asymptotes shown in Fig. 14-26.

a. Sketch and dimension the pole-zero pattern of a minimum-phase-shift voltage transmittance having all its poles in the left half of the complex plane and having this amplitude characteristic.

b. Sketch and dimension the pole-zero pattern of a non-minimum-phase-shift voltage transmittance having all its poles in the left half of the complex plane and having this amplitude characteristic.

CHAPTER 16

NONSINUSOIDAL SIGNALS AND TRANSIENT RESPONSE

The behavior of electronic circuits is studied in Chaps. 14 and 15 under conditions of sinusoidal operation in the steady state. The results of that study can be used, with the aid of the Fourier series and the Fourier transform, to determine the response of circuits to other types of signals. Alternatively, if the input signal is specified as a function of time, it is possible, at least in theory, to calculate the output signal directly from the differential equations for the circuit. This latter method of analysis leads to additional valuable insights into the properties of circuits. Of particular interest are the cases in which the nonsinusoidal input signal can be expressed as the sum of a set of components that vary exponentially with time. In these cases the output signal can be determined from the pole-zero pattern for the signal transmittance of the circuit by the methods used for sinusoidal analysis in Chap. 15; thus the response of the circuit to nonsinusoidal signals is related to its behavior in the sinusoidal steady state.

From the study of the differential equations governing the behavior of electric circuits the fact emerges that in general the response of a circuit to a signal applied at some instant of time consists of certain terms related to the applied signal plus certain additional terms arising from the properties of the circuit itself. These additional terms, known as the *transient* components of the response, may have an important effect on the behavior of the circuit. It often turns out that the transient terms in the response can also be determined from the poles and zeros of the signal transmittance; thus the transient response of the circuit is related to its behavior in the sinusoidal steady state.

It is shown in the theory of passive networks that the transient terms in the response of circuits consisting of R, L, and C alone must either die out with time or, in the case of L and C only, they must remain constant in amplitude. When controlled sources are added to the circuit, however, the transient components of the response may grow with time. For this reason it is necessary to reexamine the transient response and to investigate the way in which it is affected by the presence of controlled sources in the circuit. Growing transients may interfere with the proper operation of the circuit; on the other hand, they serve as the basis for oscillator circuits used to generate sinusoidal signals. The presence or absence of growing transients in the response of a circuit can be determined by inspection of the pole-zero pattern of the signal transmittance.

The principal objective of this chapter is to study the response of circuits to nonsinusoidal signals and to relate these responses to the pole-zero pattern of the signal transmittance. In particular, the possibility of growing transients and the conditions under which they can occur are examined. Some familiarity with the exponential form of the solution of homogeneous linear differential equations is assumed at the outset of this study.

16-1. The Response of Circuits to Nonsinusoidal Signals: the Particular Integral. The circuit shown in Fig. 16-1 is similar to the model for one stage in a cascade of RC-coupled amplifiers. The input signal $e_s(t)$ is a specified, nonsinusoidal function of time, and the problem is to determine the output voltage $e_2(t)$ as a function of time. The desired result can be obtained from the two node equations that govern the behavior of the circuit; these equations are

$$G_1e_1 + (C_1 + C_2)\frac{de_1}{dt} - C_2\frac{de_2}{dt} = -g_me_s(t) \tag{16-1}$$

and

$$-C_2\frac{de_1}{dt} + G_2e_2 + C_2\frac{de_2}{dt} = 0 \tag{16-2}$$

The solution of these equations involves finding $e_1(t)$ and $e_2(t)$ such that the equations are satisfied. Thus $e_1(t)$ and $e_2(t)$ must be of such a nature that the sum of the terms on the left of (16-1) varies with time in the same way as the applied signal $e_s(t)$, and of such a nature that the sum of the terms on the left of (16-2) is identically zero. The terms involved in these constraints contain the functions $e_1(t)$ and $e_2(t)$, their time derivatives, and, in the more general case, their time integrals; thus finding the solutions is not a simple matter in general.

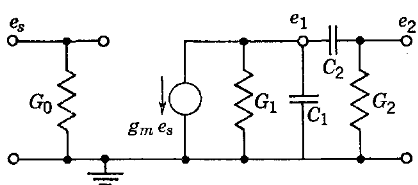


FIG. 16-1. An amplifier circuit.

Finding the solutions of differential equations such as (16-1) and (16-2) can be simplified in many important cases by making use of a special property of the exponential function ϵ^{mt}. This function retains its exponential form regardless of how many times it is differentiated or integrated. It follows from this fact that if the signal applied to the circuit in Fig. 16-1 is exponential in form, exponential forms for $e_1(t)$ and $e_2(t)$ can be found that satisfy Eqs. (16-1) and (16-2). Thus if the input signal is

$$e_s = E_s\epsilon^{mt} \tag{16-3}$$

then solutions to (16-1) and (16-2) can be found having the form

$$e_1 = E_1\epsilon^{mt} \qquad \text{and} \qquad e_2 = E_2\epsilon^{mt} \tag{16-4}$$

When these solutions are substituted into (16-1) and (16-2), every term in the equations varies in accordance with ϵ^{mt}; thus the equations are satisfied for all values of t provided the amplitudes E_1 and E_2 have suitable values.

The amplitudes E_1 and E_2 of the solutions are found by substituting (16-3) and (16-4) into (16-1) and (16-2) and determining the values required to satisfy the resulting equations:

$$G_1E_1\epsilon^{mt} + (C_1 + C_2)mE_1\epsilon^{mt} - C_2mE_2\epsilon^{mt} = -g_mE_s\epsilon^{mt} \tag{16-5}$$
$$- C_2mE_1\epsilon^{mt} + G_2E_2\epsilon^{mt} + C_2mE_2\epsilon^{mt} = 0 \tag{16-6}$$

Each term in (16-5) and (16-6) varies with time as ϵ^{mt}, and the common factor can be canceled to obtain simpler equations:

$$[G_1 + (C_1 + C_2)m]E_1 - C_2mE_2 = -g_mE_s \tag{16-7}$$
$$-C_2mE_1 + (G_2 + C_2m)E_2 = 0 \tag{16-8}$$

The amplitudes E_1 and E_2 must satisfy these equations. Solving by determinants for the amplitude of the exponential output voltage E_2 yields

$$E_2 = \frac{-C_2mg_mE_s}{[G_1 + (C_1 + C_2)m](G_2 + C_2m) - C_2^2m^2} \tag{16-9}$$

Rearranging this expression leads to

$$E_2 = -\frac{g_m}{C_1}\frac{m}{m^2 + am + b}E_s \tag{16-10}$$

When the denominator of (16-10) is factored, the resulting expression has the form

$$E_2 = -\frac{g_m}{C_1}\frac{m}{(m - m_1)(m - m_2)}E_s \tag{16-11}$$

The forward voltage transmittance relating the exponential input signal to the exponential output signal of the same form is therefore

$$A(m) = \frac{E_2}{E_s} = -\frac{g_m}{C_1}\frac{m}{(m - m_1)(m - m_2)} \tag{16-12}$$

Thus when the input signal is an exponential given by (16-3), the amplitude of the exponential output voltage can be determined from (16-11) or (16-12).

Example 16-1. The parameters in the circuit of Fig. 16-1 are such that $m_1 = -5$, $m_2 = -50$, and $g_m/C_1 = 1000$. The input signal is $e_s = 2\epsilon^{-10t}$. Determine the output voltage $e_2(t)$.

Solution. The forward voltage transmittance, given by (16-12), is

$$A = -1000 \frac{-10}{(-10+5)(-10+50)} = \frac{(-1000)(-10)}{(-5)(40)} = -50$$

The amplitude of the output voltage is

$$E_2 = AE_s = (-50)(2) = -100$$

and the output voltage as a function of time is

$$e_2(t) = E_2\epsilon^{mt} = -100\epsilon^{-10t}$$

The pole-zero pattern for the voltage transmittance given by (16-12) can be plotted, and the relations specified by (16-12) can be represented graphically as is done in Chap. 15 for sinusoidal operation in the steady state. A typical construction for the case where m is a negative real number is shown in Fig. 16-2*a*.

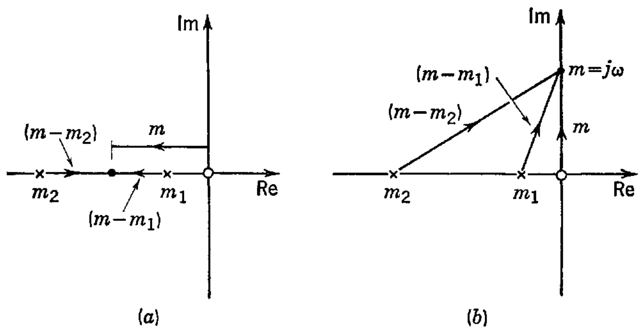


FIG. 16-2. Diagrams related to the voltage transmittance. (*a*) Exponential signal; (*b*) sinusoidal signal.

It is easily verified that if the variable m in (16-12) is given an imaginary value, $m = j\omega$, then (16-12) gives the voltage transmittance for sinusoidal steady-state operation. The graphical representation of (16-12) for this case is shown in Fig. 16-2*b*. The relations between the signal transmittances for exponential signals and sinusoidal signals is discussed in greater detail in Sec. 16-2.

If the input signal to the circuit of Fig. 16-1 consists of the sum of a number of exponential components, the responses of the circuit having the same exponential forms can be found in the manner presented above with the aid of the superposition principle. For example, if the signal consists of two exponential terms, then it has the form

$$e_s = E_{sa}\epsilon^{m_a t} + E_{sb}\epsilon^{m_b t} \tag{16-13}$$

The response of the circuit to this signal can be determined by considering each term in (16-13) separately. The response to the first term acting alone is an exponential voltage

$$e_{2a} = E_{2a}\epsilon^{m_a t} \tag{16-14}$$

The amplitude of this voltage, E_{2a}, is given by (16-11) with $m = m_a$; thus

$$E_{2a} = -\frac{g_m}{C_1}\frac{m_a}{(m_a - m_1)(m_a - m_2)}E_{sa} \tag{16-15}$$

Similarly, the response to the second term in (16-13) is

$$e_{2b} = E_{2b}\epsilon^{m_b t} \tag{16-16}$$

in which the amplitude is

$$E_{2b} = -\frac{g_m}{C_1}\frac{m_b}{(m_b - m_1)(m_b - m_2)}E_{sb} \tag{16-17}$$

Since the circuit is linear, the principle of superposition applies, and the response of the circuit when both terms in (16-13) are acting is the sum of the responses to the separate terms. Thus the response to the signal given by (16-13) is

$$e_2 = e_{2a} + e_{2b} = E_{2a}\epsilon^{m_a t} + E_{2b}\epsilon^{m_b t} \tag{16-18}$$

where E_{2a} and E_{2b} are given by (16-15) and (16-17), respectively.

Example 16-2. The parameters in the circuit of Fig. 16-1 are such that $m_1 = -5$, $m_2 = -50$, and $g_m/C_1 = 1000$. The input signal is $e_s = \epsilon^{-10t} + 2\epsilon^{-20t}$. Determine the output voltage $e_2(t)$.

Solution. The response of the circuit to the first term, $e_{sa} = \epsilon^{-10t}$, is $e_{2a} = E_{2a}\epsilon^{-10t}$, where the amplitude is given by (16-15) as

$$E_{2a} = -1000\frac{-10}{(-10 + 5)(-10 + 50)}(1) = -50$$

The response of the circuit to the second term is $e_{2b} = E_{2b}\epsilon^{-20t}$, where

$$E_{2b} = -1000\frac{-20}{(-20 + 5)(-20 + 50)}(2) = -88.8$$

The complete response is thus

$$e_2(t) = -50^{-10t} - 88.8\epsilon^{-20t}$$

16-2. Representation of Signals by Sums of Exponential Terms. In Sec. 16-1 the response of a linear circuit to exponential input signals is studied. The ease with which the differential equations for the circuit are solved when the input signal is a sum of exponential terms suggests that the analysis for signals of other waveforms can be simplified if these signals can be expressed as a sum of exponential terms. This does indeed prove to be the case, and it also happens that exponential representations for several important types of signals can be found quite easily. For example, a decaying exponential signal applied at some instant designated as $t = 0$, shown in Fig. 16-3a, is given by

$$e_s = E_s \epsilon^{-\alpha t} \qquad \text{for } 0 < t \tag{16-19}$$

where α is a positive real number. The rate at which the exponential decays depends on the value of α; the smaller α the smaller the rate of decay. In the limit as α appoaches zero, (16-19) becomes

$$e_s = E_s \epsilon^{-0t} = E_s \qquad \text{for } 0 < t \tag{16-20}$$

Thus a step of voltage applied at $t = 0$, shown in Fig. 16-3b, is represented for $0 < t$ by an exponential having a zero exponential coefficient.

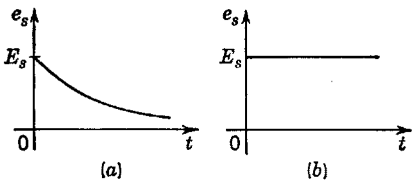


FIG. 16-3. Signals beginning at $t = 0$. (a) Decaying exponential; (b) step signal.

An important part of the analysis and design of many electrical systems is concerned with the response of the systems to step input signals.

Additional signals for which exponential representations can be found easily are:

Hyperbolic cosine:

$$e_s = E_s \cosh \beta t = \frac{E_s}{2} (\epsilon^{\beta t} + \epsilon^{-\beta t}) \tag{16-21}$$

Damped hyperbolic cosine:

$$\begin{aligned} e_s &= E_s \epsilon^{-\alpha t} \cosh \beta t = \frac{E_s}{2} \epsilon^{-\alpha t}(\epsilon^{\beta t} + \epsilon^{-\beta t}) \\ &= \frac{E_s}{2} [\epsilon^{(-\alpha+\beta)t} + \epsilon^{(-\alpha-\beta)t}] \end{aligned} \tag{16-22}$$

Cosine:

$$e_s = E_s \cos \omega t = \frac{E_s}{2} (\epsilon^{j\omega t} + \epsilon^{-j\omega t}) \tag{16-23}$$

Damped cosine:

$$\begin{aligned} e_s &= E_s \epsilon^{-\alpha t} \cos \omega t = \frac{E_s}{2} \epsilon^{-\alpha t}(\epsilon^{j\omega t} + \epsilon^{-j\omega t}) \\ &= \frac{E_s}{2} [\epsilon^{(-\alpha+j\omega)t} + \epsilon^{(-\alpha-j\omega)t}] \end{aligned} \tag{16-24}$$

Periodic signal:

$$\begin{aligned} e_s &= E_{s0} + E_{s1} \cos (\omega t + \theta_1) + E_{s2} \cos (2\omega t + \theta_2) + \cdots \\ &= \sum_{k=-\infty}^{\infty} A_k \epsilon^{jk\omega t} \end{aligned} \tag{16-25}$$

Exponential representations can be found by elementary means for only a limited number of signals. The Laplace transform provides a formal mathematical procedure for finding exponential representations for a much larger collection of signals; for this reason (among others) it plays an important role in the study of electrical systems. It often turns out, however, that the exponential representation for a given signal, obtained by the Laplace transform or equivalent means, consists of an infinite number of exponential terms; one such case is the representation for a periodic signal given by (16-25). Since in many respects it is awkward to work with an infinite number of terms, much research effort has been spent in developing procedures for representing signals approximately by the sum of a few carefully chosen exponential terms.

When a sinusoidal signal of a constant amplitude is applied to the circuit of Fig. 16-1, the signal can be represented as the sum of two exponentials in accordance with (16-23). The response of the circuit having the same form as this signal can be determined by superimposing the response to each exponential term considered separately. The amplitude of the

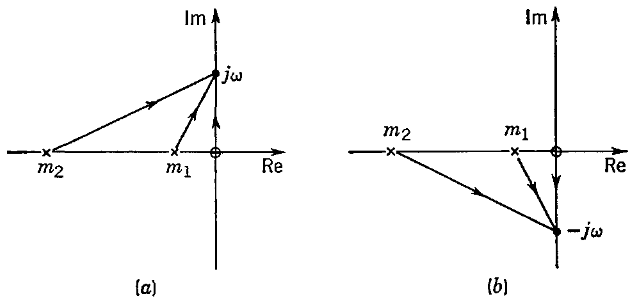


FIG. 16-4. Graphical constructions related to sinusoidal excitation. (*a*) Positive exponential; (*b*) negative exponential.

response to the term with the positive exponent is given by (16-11) with $m = j\omega$. Accordingly,

$$\begin{aligned} E_{2a} &= -\frac{g_m}{C_1}\frac{j\omega}{(j\omega - m_1)(j\omega - m_2)}\frac{E_s}{2} \\ &= A(j\omega)\frac{E_s}{2} = [|A(j\omega)|\epsilon^{j\theta_a}]\frac{E_s}{2} \end{aligned} \tag{16-26}$$

where θ_a is the angle associated with $A(j\omega)$. The graphical construction in the complex plane corresponding to $A(j\omega)$ is shown in Fig. 16-4*a*. Similarly, the amplitude of the response to the term with the negative exponential is given by (16-11) with $m = -j\omega$; thus

$$\begin{aligned} E_{2b} &= -\frac{g_m}{C_1}\frac{-j\omega}{(-j\omega - m_1)(-j\omega - m_2)}\frac{E_s}{2} \\ &= A(-j\omega)\frac{E_s}{2} = [|A(-j\omega)|\epsilon^{j\theta_b}]\frac{E_s}{2} \end{aligned} \tag{16-27}$$

where θ_b is the angle associated with $A(-j\omega)$. The graphical construction in the complex plane corresponding to $A(-j\omega)$ is shown in Fig. 16-4*b*. The complete response of the circuit to the sinusoidal input signal is the superposition of the responses to these two exponential terms:

$$e_2(t) = E_{2a}\epsilon^{j\omega t} + E_{2b}\epsilon^{-j\omega t} \tag{16-28}$$

$$= [|A(j\omega)|\epsilon^{j\theta_a}\epsilon^{j\omega t} + |A(-j\omega)|\epsilon^{j\theta_b}\epsilon^{-j\omega t}]\frac{E_s}{2} \tag{16-29}$$

It is clear from the diagrams of Fig. 16-4 that $|A(j\omega)| = |A(-j\omega)|$ and that $\theta_b = -\theta_a$; thus there is actually no need to consider the exponential term with the negative exponent, for all the required information is provided by the term with the positive exponent. Using these relations and designating the phase shift by $\theta = \theta_a = -\theta_b$ yields

$$e_2(t) = |A(j\omega)|\frac{E_s}{2}[\epsilon^{j(\omega t+\theta)} + \epsilon^{-j(\omega t+\theta)}] \tag{16-30}$$

$$= |A(j\omega)|\,E_s \cos(\omega t + \theta) \tag{16-31}$$

$$= |E_2| \cos(\omega t + \theta) \tag{16-32}$$

where $$|E_2| = |E_2(j\omega)| = |A(j\omega)|E_s \tag{16-33}$$

is the amplitude of the output voltage. Complete information about $e_2(t)$ is provided by the knowledge of E_s, $|A(j\omega)|$, and θ; thus the complex amplitude of e_2, defined by

$$E_2 = A(j\omega)E_s = |A(j\omega)|\epsilon^{j\theta}E_s \tag{16-34}$$

contains complete information about $e_2(t)$.

16-3. Transient Response: the Complementary Function. The behavior of the circuit shown in Fig. 16-1 is described by the two differential equations (16-1) and (16-2). Methods of solving these equations for the unknown node voltages $e_1(t)$ and $e_2(t)$ are discussed in Sec. 16-1. When an exponential voltage

$$e_s = E_s\epsilon^{m_st} \tag{16-35}$$

is applied at the input, the output voltage of the same form is given by (16-4) as

$$e_2 = E_2\epsilon^{m_st} \tag{16-36}$$

in which E_2, the amplitude of the exponential response, is given by (16-10) or (16-11) as

$$E_2 = -\frac{g_m}{C_1}\frac{m_s}{(m_s - m_1)(m_s - m_2)}E_s \tag{16-37}$$

The solutions of the differential equations obtained by the methods of Sec. 16-1, an example of which is given above, are not complete solutions of the equations; such solutions are particular integrals for the equations. If Eqs. (16-1) and (16-2) have nonzero solutions for e_1 and e_2 when the right-hand sides are made zero, then these solutions can be added to the particular integral and the result also satisfies the original equations (see the discussion of superposition in Sec. 13-1). The solutions of the differential equations with the right-hand sides made zero are called complementary functions. Thus the complete solution for the unknown voltage e_1 or e_2 consists of a particular integral plus a complementary function.

Since the complementary function is a solution of the differential equations with the right-hand sides made zero, it is a response that can exist when no excitation is applied to the circuit. For example, if the capacitor C_2 in Fig. 16-1 is charged before it is connected in the circuit, then e_1 and e_2 will not be zero when it is connected in the circuit, even though no signal voltage is applied at the input. The complementary function in the solution of the equations gives the voltages that exist as a result of the initial charge on the capacitor. In many circumstances the complementary function has an important effect on the response of electric circuits; hence it is necessary to examine the nature of the complementary function in some detail.

As a result of the nature of linear differential equations, the complementary function can always be expressed as an exponential function or as a sum of exponential terms. Thus the complete response of the circuit in Fig. 16-1 to an excitation consisting of a single exponential term is of the form

$$\begin{aligned} e_1 &= E_{1s}\epsilon^{m_st} + E_{1a}\epsilon^{m_at} + E_{1b}\epsilon^{m_bt} + \cdots \\ e_2 &= E_{2s}\epsilon^{m_st} + E_{2a}\epsilon^{m_at} + E_{2b}\epsilon^{m_bt} + \cdots \end{aligned} \qquad (16\text{-}38)$$

The first term in these equations is the particular integral, and the remaining terms constitute the complementary function. It is clear that the complementary function may decay with time, remain constant in amplitude, or grow with time depending on the values of m_a, m_b, ... If the complementary function grows with time, the energy stored in the capacitors and the energy dissipated in the resistors must increase with time. It follows that the complementary function can grow with time only if sources of energy are present. But the complementary function describes a response that can exist with no excitation applied to the circuit; therefore complementary functions associated with circuits composed of R, L, and C only must decay with time. Accordingly, since the complementary function has a short life, it is often referred to as the transient response of the circuit. If the signal applied to the circuit is a periodic wave or a constant, the particular integral may persist long after the transient response has become negligible; hence the particular integral is often referred to as the steady-state response of the circuit.

In accordance with the discussion above, it is ensured that the transient response of a circuit consisting only of R, L, and C decays with time. On the other hand, growing transients can occur in circuits containing controlled sources. Controlled sources are sources of energy and are therefore capable of supplying the increasing energy demanded by growing transients. Such transients continue to grow until the circuit becomes nonlinear or until some component fails as a consequence of excessive voltage or current; hence growing transients usually interfere with the proper operation of the circuit. One of the principal objectives of this study is to determine the circumstances under which growing transients may occur in electronic circuits.

The nature of the complementary function, or transient response, can be examined further by returning to the circuit of Fig. 16-1. Each term in the transient response is of the form

$$e_T = E_T \epsilon^{mt} \tag{16-39}$$

The amplitude E_T and the exponential coefficient m must be determined for each term. The exponential coefficients can be determined from the fact that each exponential transient term must satisfy the differential equations (16-1) and (16-2) with the right-hand sides made zero. Substituting

$$e_{1T} = E_{1T}\epsilon^{mt} \qquad \text{and} \qquad e_{2T} = E_{2T}\epsilon^{mt} \tag{16-40}$$

into (16-1) and (16-2) for e_1 and e_2 with $E_s = 0$, and solving for E_{2T}, the amplitude of the transient term in e_2, yields Eq. (16-11) with $E_s = 0$:

$$E_{2T} = -\frac{g_m}{C_1}\frac{m}{(m - m_1)(m - m_2)}(0) \tag{16-41}$$

$$= (A_{vo})(0) \tag{16-42}$$

Thus the amplitude of the transient must be zero unless the voltage transmittance A_{vo} is infinite. If A_{vo} is infinite, (16-42) becomes indeterminate and yields no information about E_{2T}, the amplitude of the transient; however, it does yield the permissible values of the exponential coefficient m in the transient response. Specifically, A_{vo} can be infinite, giving a nonzero value for the amplitude of the transient, only if $m = m_1$ or $m = m_2$. Exponential terms with these values of m satisfy the differential equations with $E_s = 0$, and the complete solution for the output voltage when the input signal consists of a single exponential term is

$$e_2 = E_{2s}\epsilon^{m_s t} + E_{21}\epsilon^{m_1 t} + E_{22}\epsilon^{m_2 t} \tag{16-43}$$

The values of m that make A_{vo} infinite are by definition the poles of A_{vo}; thus the transient response contains one exponential term for each pole of A_{vo}, and the poles are the exponential coefficients. (The zeros of A_{vo} do not give rise to terms in the transient response, but they do affect the amplitudes of the transient terms in accordance with the discussion that follows.) This is a significant fact. The forward voltage transmittance, determined by sinusoidal steady-state analysis or by analysis for the response of the circuit to a single exponential input, yields immediately the number of exponential terms in the transient response and the exponential coefficient for each term. Much useful information about the transient response can be deduced from these results. In particular, these results are sufficient to determine whether or not the transient response grows with time; this fact is discussed in more detail in the paragraphs that follow.

To complete the solution for the response of the circuit, it is still necessary to determine the amplitudes of the exponential terms in (16-43). The amplitude of the particular integral, given by (16-11), is

$$E_{2s} = -\frac{g_m}{C_1}\frac{m_s}{(m_s - m_1)(m_s - m_2)}E_s \tag{16-44}$$

A useful geometric interpretation of this relation is provided by the construction on the pole-zero pattern shown in Fig. 16-5*a*. Each transient term in (16-43) has the same mathematical form as the particular

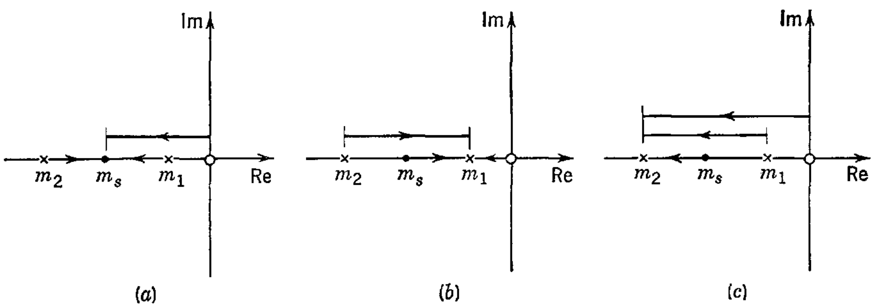


FIG. 16-5. Graphical constructions related to the amplitudes of the terms in Eq. (16-43). (*a*) Particular integral; (*b*) transient term associated with the pole at m_1; (*c*) transient term associated with the pole at m_2.

integral. This fact suggests that it might be possible to determine the amplitudes of the transient terms by a similar calculation. This does indeed prove to be the case, and under certain restricted conditions the procedure is quite simple. If the input signal consists of a single exponential term (which may be a step signal), and if there is no energy stored in the circuit when the signal is applied, then E_{21}, the amplitude of the transient term associated with the pole at m_1, is obtained from (16-44) by interchanging m_s and m_1. Thus

$$E_{21} = -\frac{g_m}{C_1}\frac{m_1}{(m_1 - m_s)(m_1 - m_2)}E_s \tag{16-45}$$

The corresponding construction on the pole-zero pattern is shown in Fig. 16-5*b*. The amplitude of the transient term associated with the pole at m_2 is obtained similarly by interchanging m_s and m_2 in (16-44). Hence

$$E_{22} = -\frac{g_m}{C_1}\frac{m_2}{(m_2 - m_1)(m_2 - m_s)} E_s \qquad (16\text{-}46)$$

The corresponding construction on the pole-zero pattern is shown in Fig. 16-5*c*.

This method for evaluating the amplitudes of the transient terms in the response of a circuit to an exponential signal applied at a certain instant is presented here without proof, and it is applicable only under the restricted conditions stated above. The Laplace transform provides a rigorous proof of the method, and it extends the method to the case of arbitrary initial conditions and to the case of applied signals having more complicated forms. Thus it is usually desirable to utilize the Laplace transform when detailed transient studies are to be made.

The results contained in Eqs. (16-44) to (16-46) can be obtained from a single compact expression:

$$E_2 = -\frac{g_m}{C_1}\frac{m}{(m - m_1)(m - m_2)(m - m_s)} E_s \qquad (16\text{-}47)$$

$$= A_{vo}(m)\frac{E_s}{m - m_s} \qquad (16\text{-}48)$$

The view may be taken that $A_{vo}(m)$ accounts for the properties of the circuit and that the factor $E_s/(m - m_s)$ accounts for the applied signal, which is given as a function of time by (16-35). Equation (16-45) for E_{21}, the amplitude of the transient term associated with the pole at m_1, is obtained from (16-47) by removing the factor $m - m_1$ and setting $m = m_1$. Equations (16-44) and (16-46) for E_{2s} and E_{22} are obtained from (16-47) in a similar manner. The pole-zero diagram corresponding to Eqs. (16-47) and (16-48) for E_2 is the same as the pole-zero diagram for A_{vo} except that it contains in addition a pole at $m = m_s$. This additional pole accounts for the exponential signal applied to the circuit.

When the signal applied to the circuit of Fig. 16-1 is a step of amplitude E_s, it can be represented as an exponential signal with $m_s = 0$ [Eq. (16-20)]. In this case the pole accounting for the signal lies at the origin of the complex plane, and it coincides with the zero of A_{vo} at that point. Under this condition the factor $m - m_s = m - 0$ in the denominator of (16-47) cancels the factor m in the numerator, and the expression for E_2 has neither a pole nor a zero at the origin; the pole associated with the signal cancels the zero associated with the voltage transmittance of the circuit.

Example 16-3. The parameters in the circuit of Fig. 16-1 are such that $m_1 = -5$, $m_2 = -50$, and $g_m/C_1 = 1000$. The input signal is a unit step of voltage applied at $t = 0$ with no initial energy stored in the capacitors. Determine the complete response, $e_2(t)$.

Solution. The response of the circuit to any exponential signal can be determined from Eq. (16-47); for the specified circuit parameters with a unit-step signal this equation reduces to

$$E_2 = -1000 \frac{m}{(m+5)(m+50)(m-0)} (1) = \frac{-1000}{(m+5)(m+50)}$$

The corresponding pole-zero pattern is shown in Fig. 16-6.

The response contains an exponential term corresponding to each pole in the expression for E_2; hence it has the form

$$e_2 = E_{21}\epsilon^{-5t} + E_{22}\epsilon^{-50t} \qquad t > 0$$

The amplitude of the first term is obtained from the expression for E_2 by removing the factor $m + 5$ and setting $m = -5$; thus

$$E_{21} = \frac{-1000}{45} = -22.2$$

The corresponding graphical construction on the pole-zero pattern for E_2 is shown in Fig. 16-6*a*. The amplitude of the second term is found in a similar manner; thus

$$E_{22} = \frac{-1000}{-45} = 22.2$$

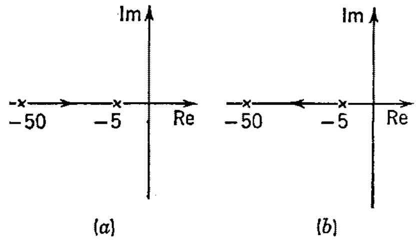


Fig. 16-6. Graphical constructions related to the step response of the amplifier of Example 16-3. (*a*) Response associated with the pole at $m = -5$; (*b*) response associated with the pole at $m = -50$.

The corresponding graphical construction is shown in Fig. 16-6*b*. The complete response of the circuit to a unit-step signal is therefore

$$e_2 = -22.2\epsilon^{-5t} + 22.2\epsilon^{-50t} \qquad t > 0$$

The particular integral, or steady-state component, in this solution is zero. This result is in accordance with the physical fact that the circuit does not transmit d-c.

Further features of the transient response of electric circuits are revealed by an examination of the step response of the single-tuned pentode amplifier shown in Fig. 15-16. The forward voltage transmittance of this circuit is given by Eqs. (15-31) and (15-32) as

$$A_{vo} = -\frac{g_m}{C}\frac{m}{m^2 + (G/C)m + \omega_o^2} = -\frac{g_m}{C}\frac{m}{(m - m_1)(m - m_2)} \qquad (16\text{-}49)$$

where

$$m_1, m_2 = -\frac{G}{2C} \pm j\sqrt{\omega_o^2 - \left(\frac{G}{2C}\right)^2} = -\alpha_1 \pm j\beta_1 \qquad (16\text{-}50)$$

When the poles are real, the step response of the tuned amplifier is identical in form with that of the amplifier of Fig. 16-1. When the poles are complex, the pole-zero pattern for A_{vo} has the form shown in Fig. 16-7a, and the step response can be determined from

$$E_2 = -\frac{g_m}{C}\frac{1}{(m - m_1)(m - m_2)}E_s \tag{16-51}$$

In this expression E_s is the amplitude of the step, and the zero of A_{vo} is canceled by the pole associated with the signal. Accordingly, the

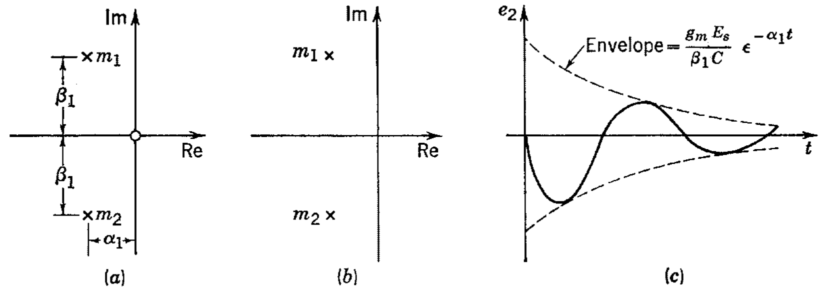


FIG. 16-7. The step response of a single-tuned amplifier. (a) Poles and zeros of A_{vo}; (b) poles and zeros for the step response; (c) the step response.

response to a step signal applied at $t = 0$ is

$$e_2 = -\frac{g_m}{C}\left[\frac{1}{m_1 - m_2}\epsilon^{m_1 t} + \frac{1}{m_2 - m_1}\epsilon^{m_2 t}\right]E_s \tag{16-52}$$

$$= -\frac{g_m}{C}\left[\frac{1}{2j\beta_1}\epsilon^{(-\alpha_1 + j\beta_1)t} - \frac{1}{2j\beta_1}\epsilon^{(-\alpha_1 - j\beta_1)t}\right]E_s$$

$$= -\frac{g_m}{\beta_1 C}\epsilon^{-\alpha_1 t}\left[\frac{1}{2j}\epsilon^{j\beta_1 t} - \frac{1}{2j}\epsilon^{-j\beta_1 t}\right]E_s$$

$$= -\frac{g_m E_s}{\beta_1 C}\epsilon^{-\alpha_1 t}\sin\beta_1 t \qquad t > 0 \tag{16-53}$$

Thus the step response is an exponentially decaying sinusoid such as the one shown in Fig. 16-7c.

The frequency of the decaying sinusoid in Fig. 16-7c depends on the distance of the poles of A_{vo} from the real axis, and the rate at which the envelope decays depends on the distance of the poles from the imaginary axis. The time required for the envelope to decay to $1/\epsilon$ times its initial value is

$$T = \frac{1}{\alpha_1} = \frac{2C}{G} = \frac{2Q_o}{\omega_o} = \frac{2}{B} \tag{16-54}$$

where $\omega_o^2 = 1/LC$ = resonant frequency of the tuned circuit
$Q_o = \omega_o C/G$ = resonant Q
B = bandwidth between the half-power points

Since $\omega_o = 2\pi f_o$, it follows that

$$f_o T = \frac{Q_o}{\pi} = N \tag{16-55}$$

is the number of cycles required for the envelope to decay to $1/\epsilon$ times its initial value. Thus the higher the Q and the smaller the bandwidth of the tuned circuit, the more slowly the transient response decays. If the conductance G is made zero, the poles lie on the imaginary axis, and the transient continues oscillating indefinitely with undiminished amplitude.

If the pair of complex poles lies in the right half of the complex plane, then $m_1 = \alpha_1 + j\beta_1$, the exponent in (16-53) is positive, and the oscillatory transient has a growing envelope. It is this phenomenon that gives rise to the self-sustaining oscillations that are discussed in Sec. 14-3. Since a growing transient implies a continuously increasing energy stored in the circuit, it also implies that the circuit contains a source of energy when no signal is applied. Since the tuned amplifier of Fig. 15-16 contains no such source, it cannot have growing transients, and it cannot have poles in the right half of the complex plane. In other circuit configurations, however, controlled sources can supply the required energy and can be responsible for growing transients.

If the parameters of the tuned circuit are adjusted so that the poles of A_{vo} move together and coincide on the negative real axis, as illustrated in Fig. 15-18, then a double pole results, and the exponential form of the transient terms in the output voltage is modified. Under these conditions $\beta_1 = 0$, and Eq. (16-53) for the step response takes the indeterminate form 0/0. The step response for this case can be determined by letting the two poles of A_{vo} approach coincidence as a limit and evaluating the indeterminate form by a suitable method. Thus, differentiating the numerator and the denominator of (16-53) with respect to β_1 and then letting β_1 tend to zero yields

$$\begin{aligned} e_2 &= -\frac{g_m E_s}{C}\epsilon^{-\alpha_1 t}\left[\lim_{\beta_1 \to 0}\left(\frac{t\cos\beta_1 t}{1}\right)\right] \\ &= -\frac{g_m E_s}{C} t\epsilon^{-\alpha_1 t} \qquad t > 0 \end{aligned} \tag{16-56}$$

when the poles of A_{vo} coincide.

In general, when the pole-zero pattern associated with the response E_2 contains multiple-order poles, the terms in the response associated with the multiple-order poles cannot be determined by the method presented in the preceding paragraphs, for indeterminate forms always result. The solution can be obtained, however, by a limiting process such as the one described above.

Example 16-4. The incremental model for the shunt-peaked amplifier discussed in Sec. 15-10 is shown in Fig. 16-8*a*. The circuit parameters are adjusted so that $\omega_H = 1/RC = 10^7$, $Q_o^2 = \omega_H L/R = \frac{1}{2}$, and $g_m/C = 10^8$. A unit step of voltage is applied to the input at $t = 0$ with no energy initially stored in the circuit. Determine the complete response $e_2(t)$.

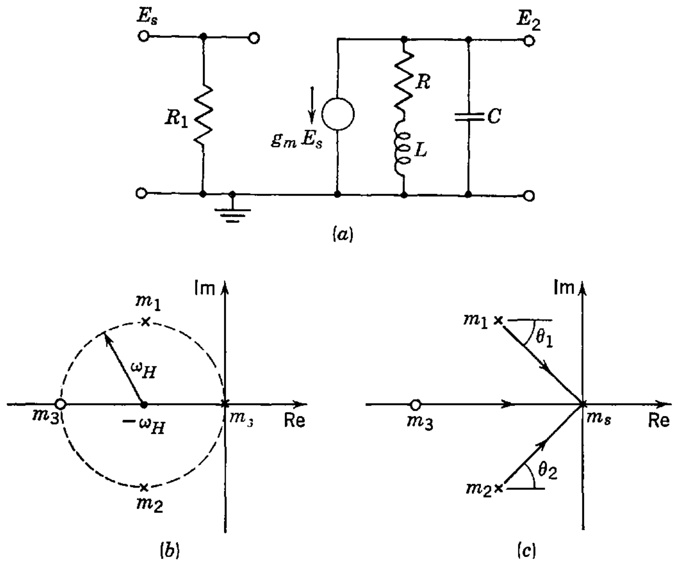


FIG. 16-8. Diagrams used in determining the step response of a shunt-peaked amplifier. (*a*) Shunt-peaked amplifier; (*b*) pole-zero diagram for E_2 with $Q_o^2 = \frac{1}{2}$; (*c*) graphical construction related to the amplitude of the steady-state term.

Solution. The forward voltage transmittance of the circuit is given by (15-111); thus, substituting numerical values in this expression and using (16-48), the unit-step response can be determined from

$$E_2 = -10^8 \frac{m - m_3}{(m - m_1)(m - m_2)(m - 0)} \quad (1)$$

where m_1, m_2, and m_3 are given by (15-112) and (15-113). The corresponding pole-zero pattern, adapted from Fig. 15-32*b*, is shown in Fig. 16-8*b*; m_1, m_2, m_3, and m_s are located at the corners of a square inscribed in the circle of radius ω_H centered at $-\omega_H$. It follows from these diagrams that

$$\begin{aligned} m_3 &= -2\omega_H = -2(10^7) \\ m_1 &= -\omega_H + j\omega_H = -10^7 + j10^7 \\ m_2 &= -\omega_H - j\omega_H = -10^7 - j10^7 \end{aligned}$$

The response of the circuit contains an exponential term for each pole in the diagram of Fig. 16-8*b*; hence, if time is expressed in microseconds and frequency is expressed in megaradians per second, the response has the form

$$e_2 = E_{2s}\epsilon^{0t} + E_{21}\epsilon^{(-10+j10)t} + E_{22}\epsilon^{(-10-j10)t} \qquad t > 0$$

The amplitude of the first term is obtained from the expression for E_2 by removing the factor associated with the signal $(m - 0)$ and setting $m = m_s = 0$; the corresponding graphical construction on the pole-zero pattern is shown in Fig. 16-8c. The length of the vector from m_3 is $2\omega_H = 2(10^7)$, the lengths of the vectors from m_1 and m_2 are each $\sqrt{2}\,(10^7)$, and the angles of the vectors from m_1 and m_2 are $\theta_1 = -\theta_2 = -45°$. Hence

$$E_{2s} = -10^8 \frac{2(10^7)}{(10^7\sqrt{2})(10^7\sqrt{2})} = -10$$

The amplitude of the second term in the equation for e_2 is determined in a similar manner, and a diagram similar to the one in Fig. 16-8c facilitates the process. Thus

$$E_{21} = -10^8 \frac{10^7\sqrt{2}\,\epsilon^{j45}}{(10^7\sqrt{2}\,\epsilon^{j135})(2)(10^7\,\epsilon^{j90})} = 5$$

Similarly,

$$E_{22} = -10^8 \frac{10^7\sqrt{2}\,\epsilon^{-j45}}{(10^7\sqrt{2}\,\epsilon^{-j135})(2)(10^7\,\epsilon^{-j90})} = 5$$

Thus, with time in microseconds,

$$\begin{aligned} e_2 &= -10 + 5\epsilon^{(-10+j10)t} + 5\epsilon^{(-10-j10)t} \\ &= -10 + 5\epsilon^{-10t}(\epsilon^{j10t} + \epsilon^{-j10t}) \\ &= -10 + 10\epsilon^{-10t}\cos 10t \qquad t > 0 \end{aligned}$$

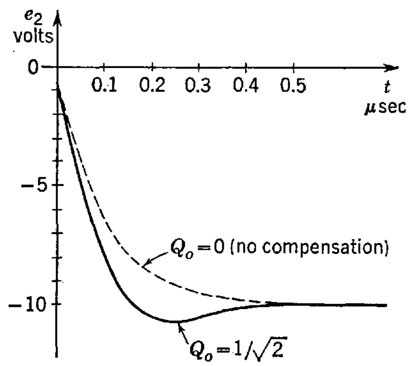


FIG. 16-9. Step response of the shunt-peaked RC-coupled amplifier of Example 16-4.

This response is shown as a solid line in Fig. 16-9. The step response of the amplifier with no compensation, $Q_o = 0$ and $L = 0$, is shown as a dotted line in Fig. 16-9 for comparison and to show the effect of the peaking inductor on the step response. The time required for the response to reach 90 per cent of its final value is substantially reduced by the peaking inductor.

In the analysis of the single-tuned amplifier above and the shunt-peaked amplifier of Example 16-4, the terms in the response associated with pairs of complex conjugate poles can in each case be combined to form an exponentially decaying sinusoidal term. This fact is no coincidence; it results from the fact that the amplitudes of the terms associated with a pair of conjugate poles are themselves complex conjugates, and this result arises from the fact that the poles and zeros of the function $E_2(m)$ are either real or complex conjugates. These relations are illustrated in Fig. 16-10. The amplitude of the response term associated with the pole at m_1 is given, apart from a constant multiplier, by the construction of Fig. 16-10a. This amplitude is a complex number that can be expressed in the form $E_1\epsilon^{j\theta_1}$, and the response term associated

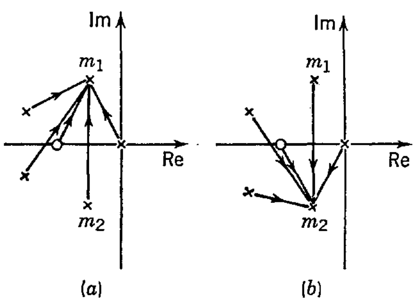


FIG. 16-10. Transient terms associated with complex conjugate poles.

with the pole at m_1 is

$$e_1 = E_1 \epsilon^{j\theta_1} \epsilon^{m_1 t} \tag{16-57}$$

Similarly, the amplitude of the response term associated with the pole at m_2 is given by the construction of Fig. 16-10*b*. It follows directly from the two diagrams that the complex number obtained from Fig. 16-10*b* is the conjugate of the number obtained from Fig. 16-10*a*. Hence the response associated with the pole at m_2 is

$$e_2 = E_1 \epsilon^{-j\theta_1} \epsilon^{m_2 t} \tag{16-58}$$

Designating the conjugate poles by

$$m_1 = -\alpha_1 + j\beta_1$$

and

$$m_2 = -\alpha_1 - j\beta_1$$

the sum of these two responses becomes

$$\begin{aligned} e_1 + e_2 &= E_1[\epsilon^{j\theta_1}\epsilon^{(-\alpha_1+j\beta_1)t} + \epsilon^{-j\theta_1}\epsilon^{(-\alpha_1-j\beta_1)t}] \\ &= E_1\epsilon^{-\alpha_1 t}[\epsilon^{j(\beta_1 t+\theta_1)} + \epsilon^{-j(\beta_1 t+\theta_1)}] \\ &= 2E_1\epsilon^{-\alpha_1 t} \cos(\beta_1 t + \theta_1) \end{aligned} \tag{16-59}$$

The response terms associated with complex conjugate poles can always be combined in this way to form an oscillatory response.

If a pair of conjugate poles is given, then α_1 and β_1 in (16-59) are known, and the response associated with these poles can be written if E_1 and θ_1 are known. Thus it is not necessary to carry out all the algebraic steps leading to (16-59) in order to obtain the response; it is sufficient to determine E_1 and θ_1 by a single operation corresponding to the diagram in Fig. 16-10*a*. The desired result, Eq. (16-59), can then be written at once.

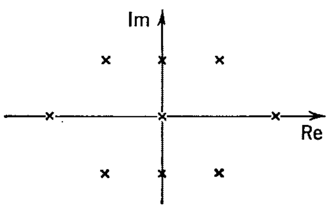


FIG. 16-11. Poles representing transients of various forms.

It follows from the preceding paragraphs that the transient response of networks can always be expressed as a sum of exponential terms and oscillatory terms. The various possibilities that exist, apart from the special case of multiple-order poles, are illustrated in Fig. 16-11. Real poles in the left half of the complex plane give rise to exponentially decaying transients, and complex poles in the left half plane give rise to exponentially decaying oscillatory transients. Their counterparts in the right half plane are associated with growing exponentials and growing oscillations. Complex poles on the imaginary axis contribute a sinusoid of constant amplitude to the response, while a pole at the origin contributes a constant, or d-c, component.

16-4. The Effect of Feedback on Pole-zero Patterns. Detailed information about the frequency dependence and the transient response of electric networks can be obtained rather easily from the pole-zero pattern of the signal transmittance. Therefore it is often convenient to study the effect of changes in circuit parameters by studying the way in which such changes affect the pole-zero pattern. The pole-zero pattern also provides a useful way of studying the effects of feedback on the dynamic response of electronic circuits. The way in which feedback affects the pole-zero pattern of the signal transmittance is often of vital importance, for it may lead to poles in the right half of the complex plane. In such cases the circuit will have growing transients.

Figure 16-12 shows a circuit that, because it contains feedback, can be used either as a frequency-selective amplifier having characteristics similar to those of a single-tuned amplifier or as an oscillator generating a sinusoidal signal. The circuit finds its widest application as an oscillator, and it serves as the prototype for commercially built oscillators that are widely used in laboratory experimentation and testing. The ideal amplifier included in the circuit is required to produce no sign reversal and to provide an amplification somewhat greater than 3; in practice it is usually approximated by a two-stage RC-coupled amplifier. The transmission from the output of the controlled source AE_1' through the RC network back to the input of the amplifier constitutes the feedback in the circuit.

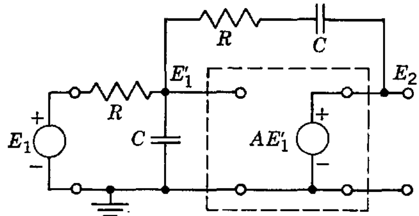


FIG. 16-12. A feedback circuit.

The output voltage of the circuit is

$$E_2 = AE_1' \tag{16-60}$$

and the node equation at E_1' is

$$\left(\frac{1}{R} + mC + \frac{mC}{1 + mCR}\right) E_1' - \frac{mC}{1 + mCR} E_2 = \frac{1}{R} E_1 \tag{16-61}$$

Eliminating E_1' between these two equations and solving for the forward voltage transmittance yields

$$\frac{E_2}{E_1} = A_{vo} = \frac{A}{CR} \frac{m + 1/CR}{m^2 + \dfrac{3 - A}{CR} m + \dfrac{1}{C^2R^2}}$$

and defining $\omega_o = 1/CR$ gives

$$A_{vo} = \omega_o A \frac{m + \omega_o}{m^2 + (3 - A)\omega_o m + \omega_o^2} \tag{16-62}$$

$$= \omega_o A \frac{m - m_o}{(m - m_1)(m - m_2)} \tag{16-63}$$

where $$m_o = -\omega_o = -\frac{1}{CR}$$

and $$m_1, m_2 = -\frac{3-A}{2}\omega_o \pm j\omega_o\sqrt{1-\left(\frac{3-A}{2}\right)^2} = -\alpha_1 \pm j\beta_1$$

These results show that as a consequence of feedback the poles of A_{vo} depend on the amplification A. When $A = 1$, the poles coincide on the negative real axis of the complex plane, and if A is increased above this value, the poles become complex. The distance of the poles from the origin is ω_o; hence if A is increased while C and R are held constant, the poles move along the paths shown in Fig. 16-13. Thus the effect of varying the amplification in this circuit is identical with the effect of varying the conductance G in the single-tuned amplifier. When $A = 3$, $\alpha_1 = 0$, and the poles lie on the imaginary axis; when A is made greater than 3, the poles move into the right half of the complex plane. The poles coincide on the positive real axis when $A = 5$, and with further increases in A they separate and move along the real axis toward zero and infinity.

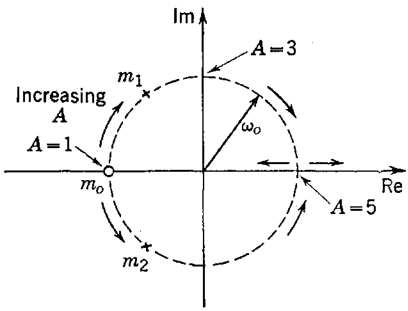


FIG. 16-13. The effect of feedback on the poles of A_{vo}.

When A is made slightly less than 3, the poles lie in the left half plane, and they are very close to the imaginary axis. Under these conditions the shape of the amplitude characteristic in the frequency band between the half-power frequencies depends only on the pole m_1, and in this band it is identical with that of a high-Q single-tuned amplifier. The resonant Q of the tuned amplifier is given by $Q_o = \omega_o/2\alpha_1$. Using this definition for the Q of the feedback amplifier in Fig. 16-12 yields

$$Q_o = \frac{\omega_o}{2\alpha_1} = \frac{1}{3-A} \tag{16-64}$$

When Q_o is large, it is also given by

$$Q_o = \frac{\omega_o}{B} \tag{16-65}$$

where B is the bandwidth between the half-power frequencies. It is clear that the Q of the resonance can be made as large as desired by making A approach the value 3 from below. A practical difficulty prevents the realization of very high Q's, however, for if the poles are very close to the imaginary axis, any slight increase in A, such as might result from a change in line voltage or a change in a circuit parameter, causes the poles to move into the right half of the complex plane. Under these conditions the circuit has a growing oscillatory transient (if $A < 5$) that interferes with the operation of the circuit as an amplifier.

The growing oscillation that results when A is greater than 3 can be profitably employed by using the circuit as an oscillator to generate a sinusoidal signal having a frequency determined by $\omega_o = 1/CR$. In this application the signal source E_1 in Fig. 16-12 is not needed, and it can be replaced by a short circuit. Once started, the oscillatory transient continues to grow until the circuit saturates and becomes nonlinear. The saturation limits further growth of the oscillation, and it may also introduce considerable waveform distortion. In order to avoid this waveform distortion it is usually desirable to incorporate in the circuit some means for limiting the growth of the oscillation that does not depend on nonlinear operation of the circuit. The desired result can be accomplished by including in the amplifier circuit a temperature-sensitive resistance arranged so that it controls the value of the amplification A. As the oscillation grows, the voltage across the temperature-sensitive element increases, the temperature of the element rises, and the resistance of the element changes in such a way as to reduce the amplification to the critical value of 3. Thus the poles move along the circle in Fig. 16-13 toward the imaginary axis as the oscillation grows. If the amplification drops below 3, the poles move into the left half plane, and the oscillation begins to die out. But as the amplitude of the oscillation decreases, the temperature-sensitive element cools off, and the amplification increases toward the critical value again. In this way the amplitude of the oscillation is held at the value that, through the action of the temperature-sensitive element, gives an amplification of 3. Since the control element is sensitive only to the rms value of the oscillation, changes in its resistance do not distort the waveform of the signal; such elements are sometimes described as quasi-linear. Since the radius of the circle on which the poles move is ω_o, it follows that the frequency of the oscillation can be adjusted by varying either C or R or both.

The circuit of Fig. 16-12 is one of many that can be used as an oscillator; it is particularly useful at frequencies below a few hundred kilocycles per second. At higher frequencies the parasitic capacitances become important, and, like amplifiers, oscillators must use tuned circuits. It is also true that circuits intended as amplifiers occasionally turn out to be oscillators. This situation is usually the result of poles accidentally moving into the right half plane as a result of accidental feedback associated parasitic capacitances. Careful layout of wiring and parts is essential in avoiding oscillations in high-gain amplifiers.

An important part of the analysis and design of feedback amplifiers is concerned with determining whether or not the signal transmittance has poles in the right half of the complex plane, for such poles are associated with growing transients. A simple rule that is often helpful in this respect is the following: If the coefficients in any polynomial $f(m)$ having real coefficients are not all of the same sign, or if the coefficient on any power of m less than the greatest is zero, then that polynomial has at least one zero in the right half of the complex plane or on the imaginary axis. The denominator polynomial in Eq. (16-62) is, of course, in agreement with this rule. The proof of this rule follows directly from the relations between the zeros of polynomials and their coefficients. It is important to note in this respect that the fact that all the coefficients of a polynomial have the same sign is *not* sufficient to ensure that the polynomial has no zeros in the right half plane.

16-5. Summary. The poles and zeros of the signal transmittance of an electric circuit not only provide information about the frequency characteristics of the circuit, but they also provide detailed information about the transient response. Moreover, the complete response of the circuit to a signal consisting of a single exponential term, including a step signal, can be determined from the poles and zeros of the signal transmittance in a simple manner if there is no energy stored in the circuit when the signal is applied. It follows that the pole-zero pattern reveals a fundamental connection between the steady-state sinusoidal response and the transient response, and it also follows that information about the transient response of a network can be obtained from steady-state sinusoidal studies and vice versa.

Circuits containing controlled sources with feedback may have growing transients, for the controlled sources can supply the required energy. These growing transients are always associated with poles of the signal transmittance in the right half of the complex plane. An important part of the analysis of feedback amplifiers is therefore concerned with the detection of right-half-plane poles, for growing transients may render a circuit unfit for amplifier service.

Determining the location of the poles of a network function usually involves factoring a polynomial. The zeros of a quadratic polynomial are easily found with the aid of the quadratic formula; however, it is more difficult to find the zeros of higher-order polynomials. A method of successive approximations based on synthetic division is often employed for this purpose. On the other hand, however, much useful information about the location of the zeros can be obtained without actually factoring the polynomial. The properties of polynomials are treated in detail in the branch of algebra known as the theory of equations; a review of this branch of algebra can be of great value in the study of pole-zero patterns of network functions.

PROBLEMS

16-1. A certain transistor amplifier having the form shown in Fig. 14-19*a* can be represented at low frequencies by the incremental model shown in Fig. 14-21. The circuit parameters are such that the forward current transmittance, given by Eq. (14-89), is

$$A_c = \frac{I_2}{I_1} = 9\,\frac{1 + j\omega/200}{1 + j\omega/1000}$$

for steady-state sinusoidal operation.

a. If the applied signal is $i_1 = 10\epsilon^{-2000t}$ μa, the particular integral in the solution for the response i_2 has the form $i_2 = I_2\epsilon^{-2000t}$. Find the amplitude of this response, I_2.

b. Repeat part *a* for a constant applied signal, $i_1 = 10$ μa [Eq. (16-20)].

16-2. The response of the circuit of Prob. 16-1 to various forms of input signal is to be determined.

a. If the applied signal is $i_1 = 10\epsilon^{-1200t}\cosh 800t$, the particular integral in the solution for the response i_2 has the form $i_2 = I_{2a}\epsilon^{m_a t} + I_{2b}\epsilon^{m_b t}$. Determine the values of m_a and m_b from the specified signal. Find the amplitudes of the terms in the response, I_{2a} and I_{2b} [Eq. (16-22)].

b. For the conditions of part *a*, what is the value of the particular integral for i_2 at the instant when $t = 1/400$ sec?

c. The signal is changed to $i_1 = 10\cos 500t$ μa. What is the value of the particular integral for i_2 at the instant when $t = 1/500$ sec? (Note that the quantity $500t$ is an angle in radians.)

16-3. The behavior of a series *RLC* circuit having the form shown in Fig. 15-19*b* is to be studied. The parameters of the circuit are such that the forward voltage transmittance, given by Eq. (15-51), is

$$A_{vo} = \frac{E_2}{E_s} = -50\,\frac{1}{(m - m_1)(m - m_2)}$$

where E_s is the amplitude of the applied signal, $m_1 = -1 + j2$, and $m_2 = -1 - j2$.

a. The applied signal is a step of unit amplitude. Find the amplitude of the particular integral (steady-state term) in the response $e_2(t)$.

b. The applied signal is $e_s = \epsilon^{-2t}$. Find the particular integral in the response $e_2(t)$.

c. The applied signal is $e_s = \epsilon^{-\alpha t}$. What value of α yields the largest amplitude of response? Consider only real values of α.

16-4. The response of the circuit of Prob. 16-3 to various forms of input signal is to be determined.

a. The signal applied to the circuit is $e_s = \epsilon^{-t}\cosh t$. Find the particular integral in the response $e_2(t)$.

b. If the applied signal is $e_s = \cos t$, the particular integral in the response $e_2(t)$ has the form $e_2 = E_2\cos(t + \theta_2)$. Find the amplitude E_2 and the phase angle θ_2.

16-5. The forward voltage transmittance of a certain triode amplifier with a bypassed cathode resistor is

$$\frac{E_2}{E_s} = A_{vo} = -20\,\frac{1 + j\omega/200}{1 + j\omega/500}$$

for steady-state sinusoidal operation. With no initial energy stored in the circuit, a unit step of voltage is applied to the input at $t = 0$. Determine the complete response $e_2(t)$ for $t > 0$. Identify the particular integral and the complementary function. Sketch and dimension $e_2(t)$.

16-6. With no initial energy stored in the circuit of Prob. 16-5, a signal $e_s = \epsilon^{-200t}$ is applied to the input at $t = 0$. Determine the complete response $e_2(t)$ for $t > 0$. What is the amplitude of the particular integral in this response?

16-7. The circuit described in Prob. 16-3 is adjusted so that the poles of A_{vo} are at $-1 \pm j$. With no initial energy stored in the circuit, a unit step of voltage is applied to the input at $t = 0$. Determine the complete response $e_2(t)$ for $t > 0$. Identify the particular integral and the complementary function. Compute $e_2(t)$ for several values of t between $t = 0$ and $t = 5$, and sketch a reasonably accurate curve of e_2 versus t.

16-8. The step response of the circuit described in Prob. 15-12 and shown in Fig. 15-36 is to be determined.

a. Sketch and dimension the pole-zero pattern for the forward voltage transmittance E_2/E_1.

b. The circuit is initially in the steady state with $e_1 = 0$ (no initial energy stored). A unit step of voltage is applied at e_1 when $t = 0$. Determine the complete solution for the response $e_2(t)$ for $t > 0$. Sketch and dimension e_2 versus t.

16-9. The forward voltage transmittance of a certain amplifier is

$$A_{vo}(m) = \frac{3(m+4)}{(m+1)(m^2+2m+2)}$$

a. Sketch and dimension the pole-zero pattern of A_{vo}.

b. With no initial energy stored in the circuit a signal $e_s = \epsilon^{-4t}$ volts is applied to the input at $t = 0$. Determine the complete response $e_2(t)$ for $t > 0$. What is the amplitude of the particular integral in this response?

16-10. The transistor circuit shown in Fig. 16-14*a* is in the steady state with no voltage applied; at $t = 0$, the d-c supply voltage, $E = 25$ volts, is applied. The problem is to determine how the base current varies with time while the new steady

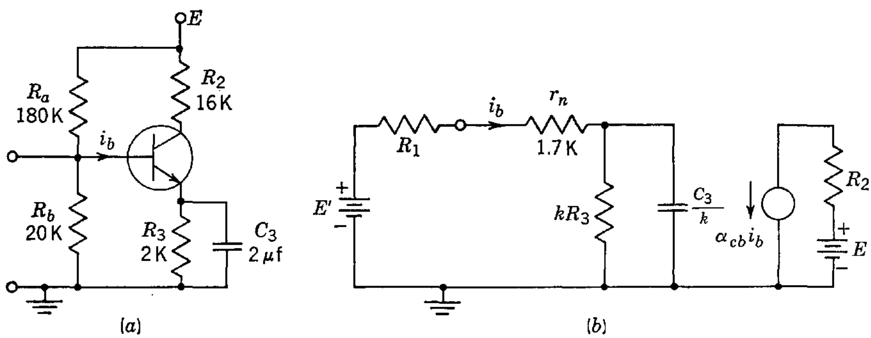


FIG. 16-14. Transistor amplifier for Prob. 16-10. (*a*) Circuit; (*b*) model, $k = 1 + \alpha_{cb}$ and $\alpha_{cb} = 60$.

state at the quiescent operating point is being approached. For this purpose the circuit can be represented approximately by the model shown in Fig. 16-14*b*. The source E' and the resistance R_1 constitute the Thevenin equivalent for E, R_a, and R_b.

a. Sketch and dimension the pole-zero pattern for the ratio $I_b(m)/E'$.

b. Determine the complete solution for $i_b(t)$, $t > 0$. Sketch and dimension a curve of i_b versus t.

c. The collector current cannot exceed the value $25/16$ ma because of collector saturation. Is this saturation current reached in the collector circuit?

16-11. The forward voltage transmittance of a certain amplifier is

$$A_{vo}(j\omega) = \frac{E_2}{E_1} = \frac{(j\omega)(j\omega+1)}{(j\omega+3)(2j\omega+4)}$$

If a step of voltage is applied at the input of this amplifier, the response is of the form

$$e_2(t) = A + B\epsilon^{-t} + C\epsilon^{-2t} + D\epsilon^{-3t} + E\epsilon^{-4t} + F\epsilon^{-2t}\cos(4t + \theta)$$

Some of these terms do not belong here. Which of the coefficients A, B, C, . . . are zero?

16-12. A certain amplifier has the forward voltage transmittance

$$A_{vo} = \frac{m + 2}{(m + 3)(m - 1)}$$

a. Sketch the pole-zero pattern for this transmittance.

b. With no initial energy stored in the circuit, a unit-step input is applied at $t = 0$. Determine the complete solution for the output as a function of time after the step is applied. Sketch the output as a function of time.

16-13. The asymptotes for the logarithmic amplitude characteristic of a certain amplifier are shown in Fig. 16-15. With no initial energy stored in the circuit, a unit-step signal is applied at the input.

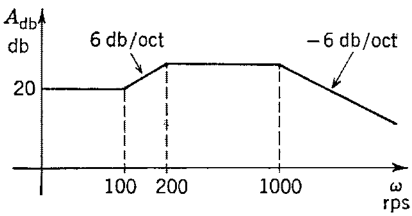


FIG. 16-15. Amplitude characteristic for Prob. 16-13.

a. If the signal transmittance has all its poles and zeros in the left half of the complex plane, what is the complete response $e_2(t)$ for $t > 0$?

b. If the zero of the signal transmittance is in the right half plane and the poles of the transmittance are in the left half plane, what is the complete response $e_2(t)$ for $t > 0$?

c. Make reasonably accurate sketches of the responses in parts *a* and *b* for comparison of the behavior of the circuit in those two cases.

16-14. The circuit shown in Fig. 16-12 is to be used as a frequency-selective amplifier. With $R = 50$ kilohms, the circuit is to be adjusted to make $f_o = \omega_o/2\pi = 1$ kcps and $Q_o = 10$.

a. Determine the values of A and C required.

b. The tuned amplifier described in Prob. 15-18 is to be adjusted to have the same poles as the RC amplifier of part *a*. If $C = 1\ \mu$f and if the plate resistance acts as an open circuit, what values of R and L are required?

c. Are the frequency characteristics of the amplifiers of parts *a* and *b* nearly identical? Are their step responses nearly the same?

16-15. The mutually coupled coils in Fig. 16-16 provide feedback around the controlled source g_mE_g. Since there is no current in L_1, L_2 behaves as if L_1 did not exist, and under sinusoidal operating conditions the voltage E is given by $E = j\omega MI$.

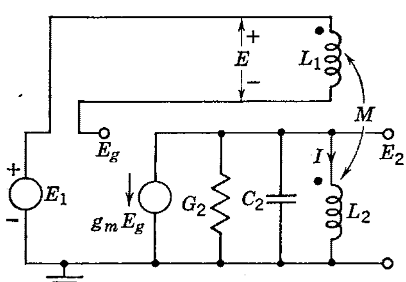


FIG. 16-16. Feedback circuit for Prob. 16-15.

a Show that for sinusoidal operation

$$A_{vo} = -\frac{g_m}{C_2}\frac{j\omega}{(j\omega)^2 + A(j\omega) + B}$$

where

$$A = \frac{G_2}{C_2} - \frac{g_mM}{L_2C_2} \quad \text{and} \quad B = \frac{1}{L_2C_2}$$

b. Sketch the loci of the poles of A_{vo} as M is varied from $-\infty$ to $+\infty$. Show the direction of motion as M increases positively.

c. If the circuit is to be used as an oscillator to generate a sinusoidal voltage, what condition among the parameters must be satisfied? What is the frequency of the oscillation in terms of the circuit parameters?

CHAPTER 17

FEEDBACK AMPLIFIERS

Feedback occurs in electronic circuits in a variety of ways and for a variety of reasons. Sometimes it is unavoidable, as in the case of feedback resulting from parasitic capacitances, and at other times it is a secondary result of circuit design, as in the case of feedback associated with cathode bias resistors. In addition, feedback is deliberately included in many circuits because certain advantages can be gained by its use. For example, feedback is often used to stabilize the quiescent operating point of transistor amplifiers.

The presence of feedback does not necessarily affect the methods used in circuit analysis. For example, the loop and node methods of analysis discussed in Chap. 12 are in no way altered by the presence of feedback, and they can be employed without regard to whether feedback exists or not. However, when feedback is deliberately used to accomplish some specific result, it is desirable to use a method of analysis that places the feedback in evidence and that permits attention to be focused on the feedback. It is also desirable that the analysis provide a quantitative measure of the effect of feedback on the circuit performance.

It is shown in Sec. 16-4 that feedback can lead to growing transients that interfere with the normal operation of the circuit unless the circuit is intended to be an oscillator. Circuits with growing transients are said to be unstable. It follows that incorporating feedback in a circuit to improve its performance also creates the possibility that the circuit may be unstable. In general, feedback circuits must be designed to give a suitable compromise between the degree of improvement realized and the danger of instability; arriving at a suitable compromise often constitutes a difficult problem. For this reason it is desirable at the outset to examine carefully the reasons for using feedback and to formulate a quantitative measure of the feedback and its consequences. This is the first objective of this chapter. The second objective is to present the Nyquist test for stability and to show how it can be used in designing feedback amplifiers.

17-1. The Effect of Feedback on Circuit Behavior. If ideal elements R, L, C and controlled sources were available without any limitations, then in so far as the signal transmission through stable, linear amplifiers is concerned, nothing could be accomplished with feedback that could not be accomplished without it. It follows, therefore, that any advantages to be gained by the use of feedback are associated with the nonideal properties of physical components or with the unavailability of suitable elements. A few of the troublesome nonideal properties of physical components are (1) parameter values that change with age, line voltage, or temperature, (2) nonlinear characteristics, (3) parasitic elements, such as the resistance of physical inductors, and (4) size, weight, and cost. Feedback can be used in many circumstances to reduce the undesirable consequences of the defects in physical components.

The reasons for using feedback in stable amplifiers can for the most part be identified as one of the following: (1) self-calibration, (2) reduction of nonlinear distortion, (3) change in amplification or impedance levels, and (4) change in dynamic characteristics (transient response and amplitude and phase characteristics). Feedback ordinarily has effects in all

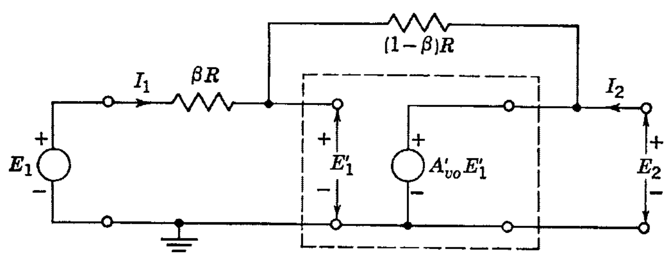


FIG. 17-1. A feedback amplifier.

four of these categories simultaneously; hence if feedback is added to a circuit to reduce distortion, it will in general change the amplification and the dynamic characteristics as well.

In this preliminary study of the effects of feedback on circuit behavior, attention is to be centered on the feedback without the distractions associated with the details of the circuitry. Therefore the circuit shown in Fig. 17-1 is used as the basis for the study. The portion of the circuit inside the box is an idealized representation of a vacuum-tube amplifier; feedback around the controlled source $A_{vo}'E_1'$ is provided by the resistors βR and $(1 - \beta)R$. A similar representation leading to similar results can be employed in the study of feedback in transistor amplifiers.

The output voltage from the circuit in Fig. 17-1 is

$$E_2 = A_{vo}'E_1' \tag{17-1}$$

and the voltage E_1', obtained by applying superposition to the sources E_1 and $A_{vo}'E_1'$, is

$$E_1' = (1 - \beta)E_1 + \beta A_{vo}'E_1' \tag{17-2}$$

Letting $\alpha = 1 - \beta$ and eliminating E_1' between these two equations gives

$$\frac{E_2}{E_1} = A_{vo} = \frac{\alpha A'_{vo}}{1 - \beta A'_{vo}} \tag{17-3}$$

for the over-all forward voltage transmittance of the circuit.

It is clear from Eq. (17-2) that β is the voltage transmittance from the output of the controlled source $A'_{vo}E'_1$ back to its input E'_1 with $E_1 = 0$. Therefore it is a measure of the amount of feedback in the circuit, and it is called the feedback ratio. The quantity A'_{vo} is the forward transmittance through the controlled source. Thus $\beta A'_{vo}$ is the transmittance around the complete feedback loop, and it is called the loop transmittance. The significance of the loop transmittance can be clarified by imagining the feedback loop to be broken at a point where no current flows as shown in Fig. 17-2. The loop transmittance is the transmittance from E'_1 to E''_1 in Fig. 17-2 with $E_1 = 0$; hence it also serves as a measure of the amount of feedback in the circuit.

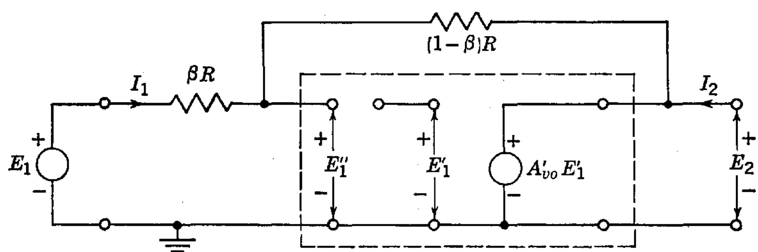


FIG. 17-2. A feedback amplifier with the feedback loop temporarily broken.

It is important to note that the loop transmittance is evaluated by determining the transmittance from E'_1 to E''_1 *with the input terminals short-circuited.* An entirely different value is obtained if the transmittance is evaluated with the input terminals open-circuited; this transmittance would be appropriate if the source of signals were a current source. If there is a source impedance connected in series with E_1, its effect on the loop transmittance must be taken into account.

The denominator in Eq. (17-3), $F = 1 - \beta A'_{vo}$, is the return difference of the feedback loop; it is the difference between E'_1 and the voltage that returns to E''_1 in Fig. 17-2 when 1 volt is applied at E'_1 with $E_1 = 0$. The return difference is a useful measure of the feedback; in particular, it is convenient to express the feedback in decibels as

$$F_{\text{db}} = -20 \log |F| \tag{17-4}$$

When $|F| > 1$, the feedback is negative, or degenerative, and when $|F| < 1$, the feedback is positive, or regenerative. It follows from these definitions and from Eq. (17-3) that negative feedback reduces the amplification of the circuit in Fig. 17-1 and that positive feedback increases the amplification.

The action of the feedback in the circuit of Fig. 17-1 has a simple qualitative explanation. Suppose that $A'_{vo} = -K$, where K is a positive real number. Then the voltage fed back to the input of the amplifier is proportional to E'_1, but it has the opposite polarity; hence the feedback reduces the effective input voltage to the amplifier and thereby reduces the over-all amplification. These relations are expressed quantitatively by Eq. (17-2), in which the feedback voltage is $\beta A'_{vo}E'_1 = -\beta K E'_1$.

If the amplification K decreases for some reason, such as aging of the tubes, the feedback likewise decreases; this action compensates in part for the effect of decreasing K on the over-all amplification. Thus a degree of self-calibration is provided by the feedback. If the loop amplification βK is much larger than unity, then Eq. (17-3) reduces to

$$A_{vo} = -\frac{\alpha}{\beta} \tag{17-5}$$

and the over-all voltage transmittance is independent of the amplification of the internal amplifier. In this way the amplifier can be made to give a precise response to the input signal even if the amplification of the internal amplifier is not precisely known. This is the reason for using feedback in many important applications.

If the internal amplifier in the circuit of Fig. 17-1 is nonlinear, it distorts the waveform of the signal. The distortion components of voltage are transmitted around the feedback loop, and if there is a sign reversal in the loop transmittance, the distortion components return to their point of origin with a reversed polarity. The return signal therefore subtracts from the original distortion components and tends to reduce the net distortion to zero.

If A'_{vo} is the voltage transmittance of an RC-coupled amplifier, $|A'_{vo}|$ decreases at very low and very high frequencies. However, the self-calibration provided by the feedback tends to maintain the over-all amplification constant. In fact, it is quite possible for the circuit to overcompensate as a result of phase shifts in the amplifier; in this case the amplitude characteristic has peaks at the low- and high-frequency ends of the passband.

The qualitative explanations of the effects of feedback given above are helpful in establishing the general ideas involved in the use of feedback; however, they do not form a suitable basis for the analysis and design of feedback amplifiers. In order that feedback be used effectively, it is necessary to have a quantitative formulation of the improvement in circuit performance that can be obtained by the use of feedback. A useful set of relations is developed in the following sections, and the improvement realized through the use of feedback is related to the problem of designing the circuit to be stable and to have suitable dynamic characteristics.

17-2. Self-calibration. Self-calibration is concerned with the dependence of the over-all transmittance of the amplifier in Fig. 17-1 on the transmittance of the internal amplifier. A measure of this dependence is provided by the derivative of A_{vo} with respect to A'_{vo}; this derivative, obtained from Eq. (17-3), is

$$\frac{dA_{vo}}{dA'_{vo}} = \frac{\alpha(1 - \beta A'_{vo}) - \alpha A'_{vo}(-\beta)}{(1 - \beta A'_{vo})^2} \tag{17-6}$$

$$= \frac{\alpha}{(1 - \beta A'_{vo})^2} \tag{17-7}$$

Multiplying and dividing the right-hand side by A'_{vo} and substituting Eq. (17-3) yields

$$\frac{dA_{vo}}{dA'_{vo}} = \frac{1}{1 - \beta A'_{vo}} \frac{A_{vo}}{A'_{vo}} \tag{17-8}$$

The sensitivity of A_{vo} to changes in A'_{vo} is defined as

$$S = \frac{dA_{vo}/A_{vo}}{dA'_{vo}/A'_{vo}} = \frac{1}{1 - \beta A'_{vo}} = \frac{1}{F} \tag{17-9}$$

Thus the sensitivity of A_{vo} to changes in A'_{vo} is inversely proportional to the return difference. Negative feedback ($|F| > 1$) reduces the sensitivity, whereas positive feedback increases it. It is seen by comparing (17-9) with (17-3) that feedback changes the voltage amplification by exactly the same amount that it changes the sensitivity to variations in A'_{vo}. Negative feedback is often used to provide self-calibration for the reason that in many applications it is easier to provide additional amplification than to obtain circuit components that do not change with age, line voltage, and similar factors. In a sense, amplification is diverted to the task of rendering the circuit less dependent on its component parts.

In general both β and A'_{vo}, and hence the sensitivity, are functions of frequency. Thus it is quite possible for an amplifier to have a small sensitivity for signals in the middle band of frequencies and yet to have a sensitivity greater than unity at both high and low frequencies. The way in which the sensitivity of any particular circuit varies with the frequency of sinusoidal signals can be displayed conveniently by means of a diagram like the one shown in Fig. 17-3. For each different value of frequency the loop transmittance $A_L = \beta A'_{vo}$ has a different complex value. A plot in the complex plane of the values of A_L for all frequencies between zero and infinity may take the form of the solid-line contour shown in Fig. 17-3. Each point on this plot corresponds to a particular frequency; the length of the vector from the origin to the contour is proportional to the magnitude of A_L, and the angle that the vector makes with the real axis is the phase shift of A_L. The loop amplification for bandpass amplifiers must tend to zero at both high and low frequencies as shown in Fig. 17-3. The plot of A_L as a function of frequency for sinusoidal signals is called a Nyquist diagram.

The vector drawn from the Nyquist diagram to the point on the positive real axis corresponding to unity gives the value of the return difference, $F = 1 - A_L$. It follows from this fact that $|F| > 1$ and the feedback is negative for all frequencies corresponding to portions of the Nyquist plot that lie outside the circle of unit radius centered at the point $1 + j0$. Similarly, the feedback is positive for all frequencies corresponding to points on the Nyquist plot lying inside this unit circle. It is usually desirable to design feedback amplifiers so that the Nyquist plot is remote from the point $1 + j0$, giving small sensitivity, in the band of frequencies occupied by the signal. The sensitivity is large at frequencies for which the Nyquist plot passes close by the critical point.

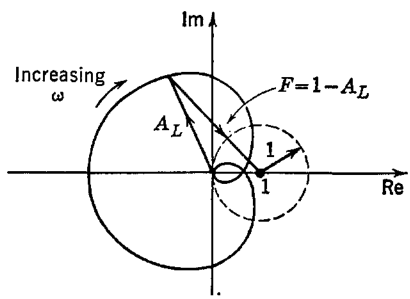


FIG. 17-3. Variation of loop transmittance and sensitivity with frequency.

The expression for sensitivity given by Eq. (17-9) is valid only for small changes in A'_{vo}. The sensitivity of A_{vo} to large changes in A'_{vo} has a similar form, and it can be evaluated in a straightforward manner. When A'_{vo} has a certain initial value A'_i, the over-all transmittance has the initial value

$$A_i = \frac{\alpha A'_i}{1 - \beta A'_i} \tag{17-10}$$

If A'_{vo} then takes on an increment $\Delta A'_{vo}$,

$$A_{vo} = A_i + \Delta A_{vo} = \frac{\alpha(A'_i + \Delta A'_{vo})}{1 - \beta(A'_i + \Delta A'_{vo})} \tag{17-11}$$

Subtracting (17-10) from (17-11) gives the resulting increment in A_{vo}; collecting terms in this difference and rearranging them yields the sensitivity of A_{vo} to large changes in A'_{vo}.

$$S = \frac{\Delta A_{vo}/A_i}{\Delta A'_{vo}/A'_i} = \frac{1}{1 - \beta(A'_i + \Delta A'_{vo})} = \frac{1}{F_i + \Delta F} \tag{17-12}$$

where F_i is the initial value of the return difference. Thus the sensitivity to large changes depends only on the final value of the return difference.

If β and A_i' are known, then the fractional change in A_{vo} resulting from a stated change in A_{vo}' can be calculated from (17-12).

17-3. The Reduction of Distortion and the Rejection of Corrupting Signals. When transistors and vacuum tubes are operated at large signal levels the nonlinear nature of their characteristics may cause appreciable distortion of the signal waveform. If the amount of distortion is small, the amplifier can be represented by a distortionless amplifier plus a distortion generator as illustrated in Fig. 17-4. In general the voltage E_d, which accounts for the distortion introduced by the amplifier, depends on the path of operation on the output characteristic of the last tube or transistor in the amplifier. However, if the distortion is small, E_d remains essentially constant when feedback is added to the circuit provided that the path of operation does not change and that the input signal is adjusted to keep the output-signal level constant. The voltage E_d can also represent any other disturbance or

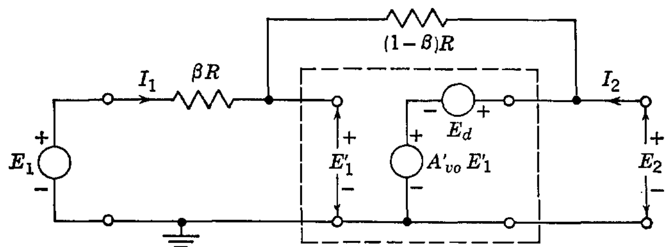


FIG. 17-4. A feedback amplifier with a corrupting signal.

corrupting signal, such as hum arising in vacuum tubes from the a-c heater supply and drift arising in transistors from changes in temperature.

The output voltage in the circuit of Fig. 17-4 is

$$E_2 = A_{vo}'E_1' + E_d \tag{17-13}$$

and the voltage E_1', which can be evaluated by applying superposition to the sources in the circuit, is given by

$$E_1' = \alpha E_1 + \beta A_{vo}'E_1' + \beta E_d \tag{17-14}$$

where $\alpha = 1 - \beta$. Eliminating E_1' between (17-13) and (17-14) yields

$$E_2 = \frac{\alpha A_{vo}'}{1 - \beta A_{vo}'} E_1 + \frac{1}{1 - \beta A_{vo}'} E_d \tag{17-15}$$

Thus both the signal component and the distortion component of the output voltages are reduced by the return difference, $F = 1 - \beta A_{vo}'$. If the signal component of the output voltage is designated by E_s, then Eq. (17-15) for the output voltage can be written as

$$E_2 = E_s + \frac{1}{1 - \beta A_{vo}'} E_d \tag{17-16}$$

Now if the input signal is adjusted so as to keep E_s constant as feedback is added to the circuit, then the ratio of signal voltage to corrupting voltage at the output of the amplifier is increased by the amount of the return difference $F = 1 - \beta A'_{vo}$.

It is important to note that adding feedback alone does not improve the ratio of signal to disturbance if E_d remains constant; however, it does permit the signal level to be increased relative to the disturbance without overdriving the amplifier. When the signal level is small, the signal-to-disturbance ratio can be improved without the use of feedback; it is merely necessary to increase the input signal in some way that does not increase the disturbance. Feedback is useful when the signal level cannot be increased without increasing the distortion introduced by the amplifier. When the signal and the disturbance can be treated separately, as, for example, when they lie in different frequency bands, it is often easier to remove the disturbance by means not involving feedback. High-frequency distortion components and low-frequency disturbances such as 60-cycle hum are often removed from the output of audio amplifiers by restricting the high- and low-frequency gains. However, when the signal and disturbance cannot be separated by other means, feedback provides a way of discriminating against the disturbance.

The effect of feedback on corrupting signals varies with frequency in a manner implied by the typical Nyquist diagram shown in Fig. 17-3. The feedback is likely to become positive at very low and very high frequencies, and any corrupting signals at these frequencies will be exaggerated by the feedback. Thus high-frequency distortion components and low-frequency noise such as turntable rumble may be emphasized by feedback in a poorly designed amplifier. For this and other reasons it is usually necessary to control the frequency characteristics of a feedback amplifier over a much wider band of frequencies than that occupied by the signal.

17-4. The Effect of Feedback on Amplification and Impedance Levels. If the signal transmittance of the internal amplifier in Fig. 17-1 is $A'_{vo} = -K$, where K is a positive real number, then the over-all voltage transmittance given by Eq. (17-3) becomes

$$A_{vo} = \frac{-\alpha K}{1 + \beta K} \tag{17-17}$$

Since α, β, and K are positive real numbers, it follows that the feedback is negative and that the voltage amplification is reduced by the feedback. If the feedback is adjusted so that the sensitivity of A_{vo} to changes in K is 0.1, so that a 1 per cent change in K causes only a 0.1 per cent change in A_{vo}, then it follows from Eq. (17-9) that the return difference must be 10. With this adjustment the over-all voltage amplification is also divided by 10.

On the other hand, if $A_{vo} = K$, where K is real and positive, and if β is adjusted so that βK is less than 2, the feedback is positive and it increases the over-all voltage amplification. Equation (17-9) shows, however, that in this case the sensitivity of A_{vo} to changes in K is increased by the same amount as the amplification. This is one of the reasons why positive feedback is seldom used to obtain large amplification.

The input admittance of the circuit in Fig. 17-1 is a function of the feedback. Since this admittance constitutes the load imposed on the source of signals, its dependence on the feedback must be taken into account. The input current is

$$I_1 = \frac{E_1 - E_2}{\beta R + (1 - \beta)R} = \frac{E_1 - A_{vo}E_1}{R} \tag{17-18}$$

when the output current I_2 is zero. The open-circuit input admittance is thus

$$Y_{no} = \frac{I_1}{E_1} = \frac{1 - A_{vo}}{R} \tag{17-19}$$

Substituting (17-3) for A_{vo} and using the fact that $\alpha + \beta = 1$ yields

$$Y_{no} = \frac{1}{R}\frac{1 - A'_{vo}}{1 - \beta A'_{vo}} \tag{17-20}$$

This result shows that negative feedback in the amplifier of Fig. 17-1 reduces the input admittance if the feedback resistance R is held constant while the feedback is changed by adjusting β. It is important to note, however, that this result applies to the circuit of Fig. 17-1; a different circuit configuration may yields different results.

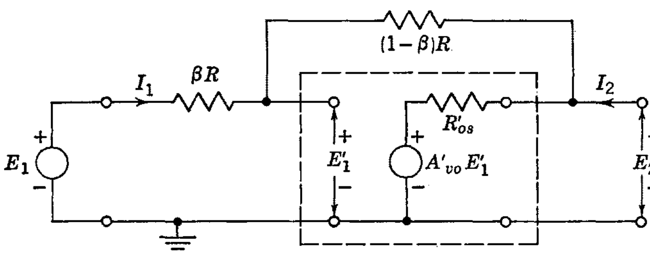


FIG. 17-5. The effect of feedback on Z_{os} and B_{cs}.

The output impedance and the reverse current transmittance are zero for the amplifier in Fig. 17-1 as a result of the fact that they are zero for the internal amplifier. In general these parameters are not zero and are functions of the feedback. For example, if the internal amplifier has an output impedance R'_{os}, the circuit takes the form shown in Fig. 17-5. The output impedance is evaluated by making $E_1 = 0$ and finding the voltage E_2 that results from a test current applied at I_2. Under these conditions the node equation at the output is

$$\left(\frac{1}{R} + \frac{1}{R'_{os}}\right) E_2 - \frac{1}{R'_{os}} A'_{vo} E'_1 = I_2 \tag{17-21}$$

But
$$E'_1 = \beta E_2 \tag{17-22}$$

whence
$$\frac{I_2}{E_2} = \frac{1}{R_{os}} = \frac{1}{R} + \frac{1 - \beta A'_{vo}}{R'_{os}} \tag{17-23}$$

Thus negative feedback reduces the output impedance of the amplifier in Fig. 17-5. It should be noted that in this case the quantity $\beta A'_{vo}$ is not the loop transmittance and $1 - \beta A'_{vo}$ is not the return difference. The loop transmittance in Fig. 17-5, evaluated by setting $E_1 = 0$ and breaking the loop as shown in Fig. 17-2, is

$$A_L = \frac{1}{1 + R'_{os}/R} \beta A'_{vo} \tag{17-24}$$

If R is much larger than R'_{os}, as often is the case, the loop transmittance is approximately $\beta A'_{vo}$.

With $E_1 = 0$ and a test current applied at I_2,

$$I_1 = -\frac{E_2}{R} = -\frac{R_{os}}{R} I_2$$

Hence the reverse current transmittance is

$$B_{cs} = -\frac{R_{os}}{R} \tag{17-25}$$

where R_{os} is given by (17-23). Thus negative feedback in the amplifier of Fig. 17-5 reduces the reverse current transmittance.

17-5. The Effect of Feedback on the Dynamic Response. The dynamic response of an amplifier may be greatly altered by the addition of feedback. Occasionally feedback is used to modify the dynamic characteristics of a circuit in some desired way; more frequently, however, feedback used for some other purpose, such as self-calibration, coincidentally changes the dynamic characteristics. These changes, which may be of great importance, can be displayed in a useful way by diagrams showing how feedback affects the pole-zero pattern of the signal transmittance.

If the internal amplifier in the circuit of Fig. 17-1 consists of one RC stage with its upper and lower half-power frequencies well separated, then the voltage transmittance at high frequencies, given by Eq. (14-4), can be put in the form

$$A'_{vo} = -A'_m \omega'_H \frac{1}{m + \omega'_H} = -A'_m \omega'_H \frac{1}{m - m'_H} \tag{17-26}$$

where A'_m is the voltage amplification in the mid-band, and ω'_H is the upper half-power frequency. Substituting (17-26) into (17-3) to obtain the over-all transmittance yields

$$A_{vo} = -\alpha A'_m \omega'_H \frac{1}{m + \omega'_H(1 + \beta A'_m)} = -A_m \omega_H \frac{1}{m - m_H} \qquad (17\text{-}27)$$

where $A_m = \alpha A'_m/(1 + \beta A'_m)$, and $\omega_H = -m_H = \omega'_H(1 + \beta A'_m)$. Thus as the feedback is increased, the pole of A_{vo} moves to the left on the negative real axis as illustrated in Fig. 17-6a. It follows that in this case the feedback increases the upper half-power frequency and decreases the mid-band amplification by the same factor, $F = 1 + \beta A'_m$. The gain-bandwidth product remains constant in this case.

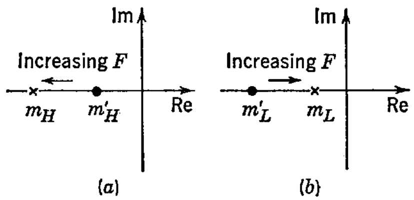


FIG. 17-6. The effect of feedback on the pole-zero pattern for the voltage transmittance of a single-stage RC amplifier. (a) High-frequency case; (b) low-frequency case.

The low-frequency behavior of the amplifier is governed by a zero at the origin and a pole nearby; the effects of the pole and zero cancel each other for signals in the mid-band and at high frequencies. The voltage transmittance of the internal amplifier in the low-frequency range is given by

$$A'_{vo} = -A'_m \frac{m}{m + \omega'_L} = -A'_m \frac{m}{m - m'_L} \qquad (17\text{-}28)$$

under the assumption that the bypass capacitors are very large. Substituting this expression into (17-3) yields

$$A_{vo} = -\frac{\alpha A'_m}{1 + \beta A'_m} \frac{m}{m + \omega'_L/(1 + \beta A'_m)} = -A_m \frac{m}{m - m_L} \qquad (17\text{-}29)$$

Thus the low-frequency pole of A_{vo} moves with increasing feedback as shown in Fig. 17-6b. (This diagram is not drawn to the same scale as that in Fig. 17-6a.) Thus the feedback reduces both the mid-band amplification and the lower half-power frequency by the amount of the return difference $F = 1 + \beta A'_m$.

The pole-zero patterns of Fig. 17-6 show the effect of feedback on the transient response of the amplifier. Changing the amount of feedback changes the time constants of the exponential terms in the transient response.

If the internal amplifier in the circuit of Fig. 17-1 consists of two RC stages with the upper and lower half-power frequencies in each stage widely separated, and if the circuit is arranged so that there is a sign reversal in the voltage transmittance, then at high frequencies

$$A'_{vo} = \frac{-A'_m\omega'_1\omega'_2}{(m + \omega'_1)(m + \omega'_2)} = \frac{-A'_m\omega'_1\omega'_2}{(m - m'_1)(m - m'_2)} \tag{17-30}$$

where $\omega'_1 = -m'_1$ and $\omega'_2 = -m'_2$ are the half-power frequencies of the two stages. Substituting this expression into (17-3) yields the over-all voltage transmittance:

$$A_{vo} = -\alpha A'_m\omega'_1\omega'_2 \frac{1}{m^2 + (\omega'_1 + \omega'_2)m + \omega'_1\omega'_2(1 + \beta A'_m)} \tag{17-31}$$

$$= \frac{-\alpha A'_m\omega_o{}^2}{1 + \beta A'_m} \frac{1}{m^2 + (\omega'_1 + \omega'_2)m + \omega_o{}^2} \tag{17-32}$$

$$= -A_m\omega_o{}^2 \frac{1}{(m - m_1)(m - m_2)} \tag{17-33}$$

where $A_m = \alpha A'_m/(1 + \beta A'_m)$ is the over-all mid-band amplification, and $\omega_o{}^2 = m_1m_2 = \omega'_1\omega'_2(1 + \beta A'_m)$ is the product of the poles of the over-all transmittance.

It follows from (17-31) that with no feedback ($\beta = 0$) the poles of A_{vo} are the same as the poles of A'_{vo}. As β is increased from zero, the constant term in the quadratic denominator of (17-31) increases while the coefficients on m^2 and m remain constant; thus increasing β is equivalent to increasing the coefficient c in Eq. (15-44), and in accordance with the discussion of (15-44), the poles of A_{vo} move along the paths shown in Fig. 17-7a as β is increased. Equation (17-33) has the same form as Eq. (15-51); hence the amplitude characteristic of the two-stage feedback amplifier at high frequencies has the form shown in Fig. 15-19c. Important details about the amplitude characteristic and its dependence on the amount of feedback can be obtained from the resonant-peaking circle developed in Sec. 15-7; the peaking circle is shown in Fig. 17-7a. If the peaking circle intersects the imaginary axis, the amplitude characteristic has a maximum at the frequency corresponding to the point of intersection. The frequency and the height of the peak can be determined easily from the relations developed in Sec. 15-7. The amount of feedback that makes the circle tangent to the imaginary axis is the greatest amount of feedback that can be used without a peak appearing in the amplitude characteristic; with this amount of feedback the amplifier is maximally flat.

The diagram of Fig. 17-7a also shows the effect of feedback on the transient response. When ω_o is greater than $(\omega'_1 + \omega'_2)/2$, the poles are complex and the transient response is oscillatory. The frequency of the oscillation depends on the amount of feedback, but the rate at which the transient decays is fixed by ω'_1 and ω'_2, the half-power frequencies of the internal amplifier.

If the internal amplifier is arranged so that there is no sign reversal in A'_{vo}, the minus sign disappears from the numerator of (17-30), and $\omega_o{}^2$ becomes $\omega'_1\omega'_2(1 - \beta K'_m)$. In this case $\omega_o{}^2$, which is also the product of the poles m_1m_2, decreases with increasing feedback, and the poles of A_{vo} must move apart on the real axis. When $\beta K'_M = 1$, $\omega_o = 0$, and one of the poles must lie at the origin; if the feedback is increased further, $\omega_o{}^2$ becomes negative, the pole at the origin moves out on the positive real axis, and the circuit has a growing transient; that is, with this connection the circuit becomes unstable when the loop transmittance is greater than unity. When there is a sign reversal in the amplifier, however, the circuit can never become unstable.

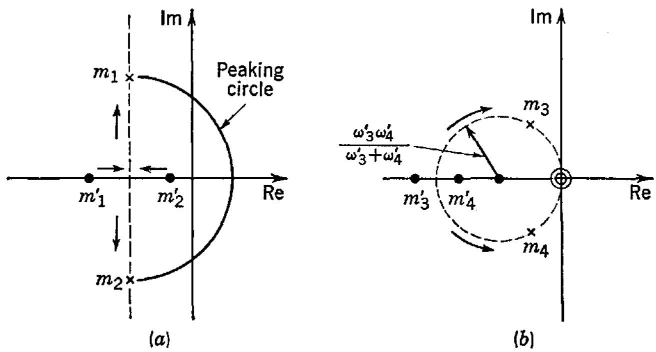


FIG. 17-7. Feedback in a two-stage RC amplifier. (a) High-frequency case, increasing F; (b) low-frequency case, increasing F.

If the bypass capacitors are very large, and if there is a sign reversal in the internal amplifier, the voltage transmittance of the internal amplifier at low and medium frequencies has the form

$$A'_{vo} = -A'_m \frac{m^2}{(m + \omega'_3)(m + \omega'_4)} = -A'_m \frac{m^2}{(m - m'_3)(m - m'_4)} \tag{17-34}$$

Substituting this expression into (17-3) yields for the over-all voltage transmittance.

$$A_{vo} = -\frac{\alpha A'_m}{1 + \beta A'_m} \frac{m^2}{m^2 + \dfrac{\omega'_3 + \omega'_4}{1 + \beta A'_m} m + \dfrac{\omega'_3\omega'_4}{1 + \beta A'_m}} \tag{17-35}$$

$$= -A_m \frac{m^2}{(m - m_3)(m - m_4)} \tag{17-36}$$

Varying the amount of feedback in the circuit affects the coefficients in the quadratic denominator in the same way that varying a affects the coefficients in Eq. (15-44); thus, in accordance with the discussion related to (15-44), the poles of A_{vo} move along the paths shown in Fig. 17-7b as βA_m varies from zero to infinity.

The effects of feedback on the low-frequency amplitude characteristic and on the transient response of the amplifier can be perceived from the diagram of Fig. 17-7*b*. The amplitude characteristic has the form shown in Fig. 15-19*c* except that the frequency scale is reversed. The peaking circle associated with the two poles has a significance similar to that which it has in the high-frequency case, even though there are two zeros at the origin of the complex plane in this case. If the circle intersects the imaginary axis, the gain characteristic has a peak. In this case, however, if the circle intersects the axis at ω_i, the peak in the amplitude characteristic occurs at $\omega_p = \omega_o^2/\omega_i$, where $\omega_o^2 = m_3 m_4 = \omega_3' \omega_4'/(1 + \beta A_m')$. These facts follow directly from the properties of the function $A_{vo}(u)$ formed from (17-35) by replacing m with ω_o^2/u.

If the internal amplifier in the circuit of Fig. 17-1 is a cascade of three identical RC stages, then in the middle- and high-frequency ranges

$$A_{vo}' = -A_m' \omega_1'^3 \frac{1}{(m + \omega_1')^3} = -A_m' \omega_1'^3 \frac{1}{(m - m_1')^3} \tag{17-37}$$

Substituting this expression into (17-3) yields

$$A_{vo} = -\alpha A_m' \omega_1'^3 \frac{1}{(m - m_1')^3 + \beta A_m' \omega_1'^3} \tag{17-38}$$

$$= -A_m \omega_o^3 \frac{1}{(m - m_1)(m - m_2)(m - m_3)} \tag{17-39}$$

where $A_m = \alpha A_m'/(1 + \beta A_m') =$ mid-band amplification
$\omega_o^3 = -m_1 m_2 m_3 = \omega_1'^3(1 + \beta A_m')$

The paths along which the poles move as the amount of feedback is increased can be found with the aid of the diagram shown in Fig. 17-8*a*. The poles are the values of m that make the denominator in (17-38) zero; hence if $m = m_1$ is substituted in (17-38), the denominator must satisfy the condition

$$(m_1 - m_1')^3 = -\beta A_m' \omega_1'^3 = (\rho \epsilon^{j\theta})^3 = \rho^3 \epsilon^{j3\theta} \tag{17-40}$$

The vector $m - m_1'$, which has the length ρ and makes an angle θ with the real axis, is shown in Fig. 17-8*a*. Since the cube of this vector must be a negative real number, it follows that 3θ must be an odd multiple of 180°; the values of θ satisfying this requirement are 60, 180, and $-60°$. Thus, as the feedback is increased, the poles move along the paths shown in Fig. 17-8*b*.

It follows from the diagram in Fig. 17-8*b* that the three-stage feedback amplifier is certain to have a growing oscillatory transient if the amount of feedback is made greater than a certain value. When the feedback is adjusted so that the poles lie on the imaginary axis, the circuit is said to be on the threshold of instability. The amount of feedback giving the threshold condition can be determined from Fig. 17-8*b*. Since the paths along which m_1 and m_2 move make angles of 60° with the real axis, the length of the vector $m_1 - m_1'$ is $\rho = 2\omega_1'$ when the poles are on the imaginary axis. Then it follows from (17-40) that

$$\beta A_m' = 8 \qquad (17\text{-}41)$$

yields the threshold condition.

If the three stages in the amplifier have different half-power frequencies, the paths followed by the poles are somewhat more difficult to construct, for they are not straight lines. However, as the feedback is increased without limit, the paths approach those shown in Fig. 17-8*b* as asymptotes. The paths of the poles for any number of identical stages can be constructed by a simple extension of the method employed above. In

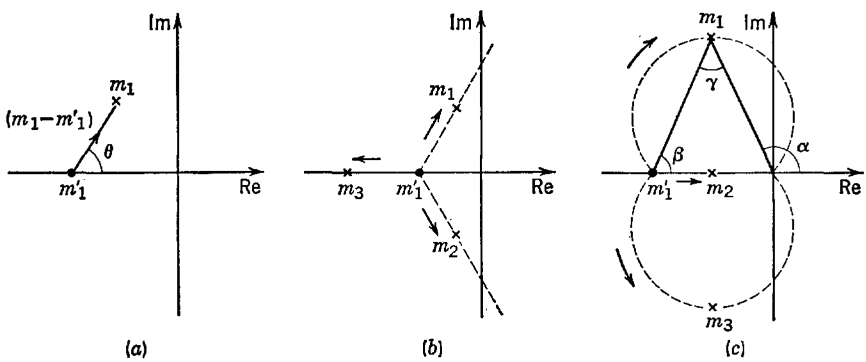


FIG. 17-8. Feedback in a three-stage *RC* amplifier. (*a*) Motion of m_1 with increasing F; (*b*) motion of the high-frequency poles with increasing F; (*c*) motion of the low-frequency poles with increasing F.

particular, if the circuit is always arranged so that there is a sign reversal in the loop transmittance, and if the number of identical stages is N, the paths make angles of $n180/N$ degrees with the real axis, where n is an odd integer. The one- and two-stage amplifiers treated earlier in this section are particular examples that follow this rule.

When the internal amplifier consists of three identical *RC* stages with large bypass capacitors, the low-frequency poles move with increasing feedback along the paths shown in Fig. 17-8*c*. It can be shown by a simple argument similar to the one used above that the complex poles must move in such a manner that the angle γ shown in Fig. 17-8*c* remains constant at 60°; thus the loci for the complex poles are segments of circles. It can be shown that in this case also a mid-band loop amplification of eight places the circuit on the threshold of instability.

The results of the preceding paragraphs place in evidence the principal problem confronting the designer of feedback systems. The decision to use feedback is the result of a desire to improve the system performance in some respect, as, for example, by providing a degree of self-calibration. In order to achieve a significant improvement a large loop amplification is necessary; however, a large loop amplification is almost certain to affect adversely the dynamic characteristics of the system, and it is quite likely to make the system unstable. Thus the design problem is largely concerned with choosing the system parameters so as to permit a large loop amplification while at the same time maintaining suitable dynamic characteristics. This problem and some elementary techniques for dealing with it are discussed in the sections that follow.

17-6. The Nyquist Test for Stability. Circuits composed of controlled sources and positive R's, L's, and C's cannot be unstable if they have no feedback, for such circuits have no sources capable of supplying the energy associated with growing transients. The results of Sec. 17-5 show, however, that when feedback is added to such circuits, instability may develop. In this case the controlled sources enclosed in the feedback loop are capable of supplying the required energy. It can be shown that if the number of poles of the loop transmittance exceeds the number of zeros by three or more, the circuit is certain to be unstable if the loop amplification is made great enough. When the loop transmittance has only two or three poles, the paths along which the poles move as the amount of feedback is varied can be determined, and questions related to stability can be answered from the results obtained. The way in which the poles move in more complicated circumstances is more difficult to determine, although techniques for constructing the approximate paths of their motion have been developed in considerable detail. There are, however, other approaches to the problem that prove to be useful in the analysis and design of more complicated feedback circuits.

Determining whether or not a given circuit is stable involves determining in some manner whether or not the signal transmittance has poles in the right half of the complex plane. The signal transmittance can be expressed as the ratio of two polynomials in m, and the zeros of the denominator polynomial are the poles of the signal transmittance. Factoring this polynomial becomes laborious when its degree is greater than 3. However, the existence of zeros of the polynomial in the right half of the complex plane can be detected by an examination of its coefficients without actually determining the zeros. For example, it follows directly from the relations between the zeros and the coefficients that if there is any variation in sign among the coefficients, then there must be at least one zero of the polynomial in the right half of the complex plane. In addition, if the coefficient on any power of m less than the greatest in the polynomial is zero, then there must be at least one zero of the polynomial in the right half plane or on the imaginary axis. However, the fact that all the coefficients are nonzero and of the same sign does not ensure that there are no zeros in the right half plane. The Routh-Hurwitz test is a more detailed examination of the coefficients that discloses the number of right-half-plane zeros of any polynomial. These tests, which are simple to apply, determine whether or not a proposed circuit will be stable. Unfortunately, however, they provide little guidance in the design of feedback circuits for stable operation. The Nyquist test for stability, which is developed in the following paragraphs, has a number of features that result in its being especially valuable in the design of feedback systems. This test not only shows whether or not the system will be stable, but it also presents the information in such a way as to aid the designer in arriving at a suitable design. In addition, it permits experimental data to be used in the design of systems that are too complicated for complete analysis.

The forward voltage transmittance of the feedback amplifier in Fig. 17-1 is given by Eq. (17-3) as

$$A_{vo} = \frac{\alpha A'_{vo}}{1 - \beta A'_{vo}} = \frac{\alpha A'_{vo}}{1 - A_L} \tag{17-42}$$

The quantities α, A'_{vo}, and A_L are unaltered if the feedback loop is broken as shown in Fig. 17-2. Since there is no feedback in the circuit when the loop is broken, these quantities are stable transmittances and have no poles in the right half plane. Since the return difference $F = 1 - A_L$ has the same poles as A_L, it follows that F has no poles in the right half plane. Therefore, if A_{vo} has any poles in the right half plane, they must be right-half-plane zeros of the return difference F, and the stability of the amplifier can be examined by studying the return difference. The Nyquist test for stability is a test for right-half-plane zeros of F.

For concreteness, consider the return difference

$$F = \frac{m - m_1}{m - m_2} = \frac{\rho_1}{\rho_2} \epsilon^{j(\theta_1 - \theta_2)} \tag{17-43}$$

where ρ_1 and ρ_2 are the magnitudes and θ_1 and θ_2 are the angles of the numerator and denominator. The pole-zero pattern for F is shown in Fig. 17-9*a*. For any given value of the variable m, a corresponding value of the return difference F can be calculated from Eq. (17-43). If m is given a succession of values corresponding to movement of the point m around the contour C_1 in Fig. 17-9*a*, F takes on a succession of values given by (17-43). A plot of these values of F in the complex plane forms a contour such as C_2 shown in Fig. 17-9*b*; this contour provides the desired information about the location of the zeros of F. As the variable m makes one complete circuit of the contour C_1, the angles θ_1, θ_2, and

θ_1-θ_2 go through some variations but finally return to their initial values; there is no net change in these angles. It follows that the angle θ_1-θ_2 in Fig. 17-9*b* experiences no net change as F traces out the contour C_2 and that therefore C_2 does not encircle the origin. This result stems directly from the fact that m_1 and m_2 lie outside the contour C_1. If m_1 lies inside the contour, as shown in Fig. 17-9*c*, θ_1 changes by 360° in

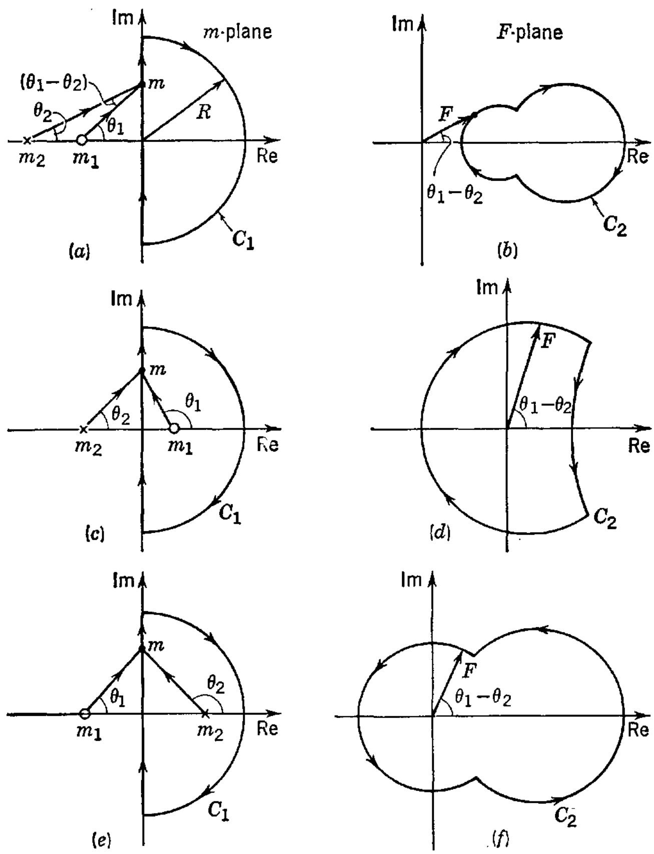


FIG. 17-9. Nyquist's test for stability.

the clockwise direction and θ_2 experiences no net change as the variable m makes one complete clockwise circuit of C_1. Hence the angle θ_1-θ_2 changes by 360° and the contour C_2 encircles the origin one time in the *clockwise* direction as shown in Fig. 17-9*d*. If m_2 is inside the contour C_1, as illustrated in Fig. 17-9*e*, θ_1 experiences no net change and θ_2 changes by 360° in the clockwise direction as m makes one complete clockwise circuit of C_1. Therefore in this case the angle θ_1-θ_2 changes by 360° and C_2 encircles the origin one time in the *counterclockwise* direction as shown in Fig. 17-9*f*. It also follows from this same reasoning that if both m_1 and m_2 are inside C_1, C_2 does not encircle the origin.

The discussion of the preceding paragraph shows that the number of times that the contour C_2 encircles the origin of the complex plane in the clockwise direction is equal to the excess of zeros of the return difference over poles inside the contour C_1. Thus if Z is the number of zeros of F inside C_1 and if P is the number of poles inside C_1, then the number of times that C_2 encircles the origin in the clockwise direction is

$$N_{cw} = Z - P \tag{17-44}$$

Now if the radius of the circular portion of C_1 is increased without limit so that C_1 encloses the entire right half of the complex plane, then (17-44) gives the excess of zeros over poles in the entire right half plane. But if the amplifier is stable with the feedback loop broken, the return difference has no poles in the right half plane, $P = 0$, and the number of clockwise encirclements of the origin by C_2 is equal to the number of zeros of F in the right half plane. Since the zeros of F are also poles of the forward voltage transmittance given by Eq. (17-42), the contour C_2, which is the Nyquist diagram of the return difference, gives the number of poles of A_{vo} in the right half plane. If the amplifier is to be stable, the Nyquist plot of its return difference must not encircle the origin of the complex plane.

The Nyquist diagram for the loop transmittance of a typical amplifier is shown in Fig. 17-3. Since $F = 1 - A_L$, the Nyquist plot for the return difference can be constructed from the plot in Fig. 17-3 by rotating the diagram through 180° and shifting it to the right by one unit. This construction is unnecessary, however, for the desired information regarding stability can be obtained directly from Fig. 17-3. Since $A_L = 1$ when $F = 0$, an encirclement of the origin by the Nyquist plot of F corresponds to an encirclement of the point $1 + j0$ by the Nyquist plot of A_L. Hence the amplifier is stable if, and only if, the Nyquist plot of A_L does not encircle the critical point $1 + j0$.

The construction of the Nyquist diagram can be simplified by taking advantage of certain properties of A_L. Since the poles and zeros of A_L are either real or occur in conjugate pairs, it follows directly from the pole-zero pattern that $A_L(-j\omega)$ is the conjugate of $A_L(j\omega)$. Hence that portion of the Nyquist plot corresponding to values of m on the negative imaginary axis need not be calculated; it is the mirror image about the real axis of the portion corresponding to values of m on the positive imaginary axis. In addition, because of parasitic elements such as stray capacitance, the loop transmittance of any physical system must tend to zero as m tends to infinity. Thus in constructing the Nyquist plot for A_L it is sufficient to consider only values of m on the positive imaginary axis up to a point beyond which $|A_L|$ remains less than unity; since the remainder of the plot cannot encircle the critical point, it can be ignored.

Equation (17-3) for the forward voltage transmittance of the amplifier can be cleared of fractions and expressed as the ratio of two polynomials. The stability of the amplifier can then be determined from the Nyquist plot of the denominator polynomial. However, the Nyquist plot of the loop transmittance serves as more than just a test of stability; it provides a basis for the design of feedback systems to have certain desired properties and to be stable. In terms of the Nyquist plot of A_L, the design of a feedback amplifier is usually concerned with making $|A_L|$ sufficiently large in the band of frequencies occupied by the signal and with controlling the frequency characteristics of A_L outside the signal band so that the plot does not encircle the critical point. The frequency characteristics of A_L are controlled by the design of the interstage coupling networks with the aid of the techniques presented in Chaps. 14 and 15.

Another important aspect of the Nyquist plot of the loop transmittance is the fact that it can be measured experimentally. Thus if the system is so complicated that a sufficiently accurate analysis is not feasible, experimental measurements of A_L can be used to guide the design of the networks used to shape its frequency characteristics.

The Nyquist plot for the loop transmittance in the circuit of Fig. 17-1 when the internal amplifier consists of three identical stages is shown in Fig. 17-10*a*. The amplifier is stable under the conditions pictured, and the portion of the curve lying in the vicinity of the negative real axis corresponds to frequencies in the mid-band of the amplifier. If the amount of feedback is increased uniformly at all frequencies by increasing β, then the Nyquist plot expands without changing its shape. It is clear that if β is increased sufficiently, the critical point is encircled and the amplifier becomes unstable.

If the algebraic sign of A_L is reversed by some change in the circuit, then the diagram is rotated through 180°, and if $|A_L|$ is greater than unity in the mid-band, the critical point is encircled and the circuit is unstable. It follows that the simple circuit configuration of Fig. 17-1 can be used only with an odd number of stages, for with an even number of stages the Nyquist plot lies primarily in the right half of the complex plane, and the circuit is unstable if $|A_L|$ is greater than unity in the mid-band.

If the value of β is adjusted so that the Nyquist plot passes through the critical point, the circuit is at the threshold of stability. This is also the adjustment that places the complex poles in Fig. 17-8 on the imaginary axis.

A special case of interest and importance is illustrated by the Nyquist diagram of Fig. 17-10*b*. Only that portion of the diagram corresponding to positive ω is shown; the discussion is not altered in any way by the remainder of the plot, which is the mirror image around the real axis of the plot shown. The critical point is not encircled by the diagram; hence the circuit is stable. However, if the plot is caused to shrink uniformly at all frequencies, a point is reached at which the circuit becomes unstable. If the plot is shrunk still further, another point is reached at which the circuit becomes stable again. Circuits exhibiting such a phenomenon are termed *conditionally stable.*

Conditional stability, which is likely to occur in amplifiers using large amounts of feedback, is important for several reasons. If an excessive signal is applied momentarily to the amplifier, the amplifier saturates, and its effective amplification is reduced. If the amplifier is conditionally

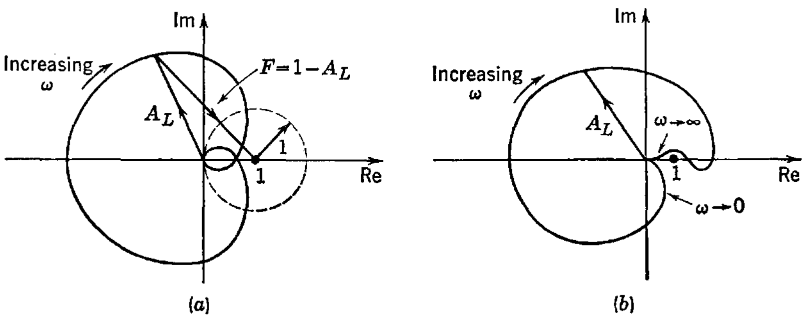


FIG. 17-10. Nyquist diagrams. (*a*) A stable circuit; (*b*) a conditionally stable circuit.

stable, the momentary reduction in amplification may make the circuit unstable. In this case growing oscillations appear, and the amplifier may remain saturated (and therefore unstable) as a result of its own oscillations, even after the excessive input signal is removed.

When an amplifier is first turned on, its amplification is zero. As the cathodes come up to temperature, the loop transmittance must grow from zero to its normal operating value. If the Nyquist plot of the loop transmittance has the form shown in Fig. 17-10*b*, the amplifier may have to pass through the unstable condition as it warms up. In such a case oscillations will start, and they may grow to such a magnitude that they saturate the amplifier and prevent the final stable state from being reached. For these reasons conditional stability is usually to be avoided in vacuum-tube circuits.

17-7. The Design of Feedback Amplifiers. The construction of the Nyquist diagram by computing the magnitude and phase of the loop transmittance for a number of different frequencies is likely to be a tedious procedure. The amount of time and effort required can be reduced, however, by first constructing the logarithmic amplitude and phase characteristics of the loop transmittance by the rapid techniques presented in Chaps. 14 and 15; the amplitude and phase of the loop transmittance can then be read from these curves. But since the amplitude and phase characteristics contain all the information that the Nyquist plot contains, there is no need to construct the latter; it is merely necessary to interpret the stability criterion and the feedback relations in terms of the amplitude and phase characteristics.

The amplitude and phase characteristics for the loop transmittance of a typical feedback amplifier are shown in Fig. 17-11; it is understood that there is a sign reversal in the loop transmittance in addition to the phase shift shown by the phase characteristics. It follows directly from the relationship between the characteristics of Fig. 17-11 and the Nyquist plot of A_L that the Nyquist plot does not encircle the critical point if $|A_L|$ drops to zero db before the phase shift reaches 180°, for

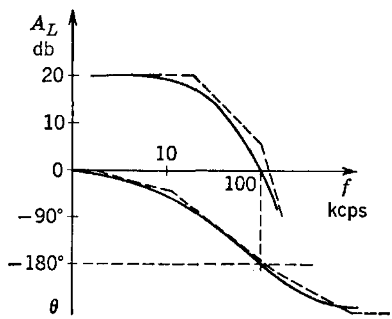


FIG. 17-11. Loop amplitude and phase characteristics for a three-stage amplifier.

zero db corresponds to a numerical ratio of unity. Therefore, except for the case of conditional stability, this becomes the stability criterion in terms of the amplitude and phase characteristics. This criterion is, of course, subject to the condition stated in connection with Eq. (17-44) that the amplifier is stable with the feedback loop broken. In the event that the amplifier is not stable with the feedback loop broken, the statement of the stability criterion in terms of the amplitude and phase characteristics must be modified. It follows from this discussion that if the amplifier having the characteristics shown in Fig. 17-11 is stable with the feedback loop broken, it is on the threshold of instability when the loop is closed.

In terms of the amplitude and phase characteristics of the loop transmittance, the design of feedback amplifiers is concerned with providing sufficient loop amplification in the band of frequencies occupied by the signal and with controlling the cutoff characteristic of the loop outside the signal band so that the amplitude is reduced to zero db before the phase shift becomes 180°. If the shapes of the amplitude and phase characteristics are fixed, the design problem involves only the choice of amplification level that permits the stability requirement to be satisfied with a suitable margin of safety. In order to realize high performance in a feedback amplifier, however, it is usually necessary to control the shapes of the amplitude and phase characteristics by the use of frequency-dependent interstage coupling networks. The shaping of the frequency characteristics of the loop transmittance to obtain improved performance requires considerable skill, however, for the amplitude and phase characteristics are interrelated. Any change in the shape of the amplitude characteristic is in general accompanied by a change in the shape of the phase characteristic, and this latter change may make matters worse

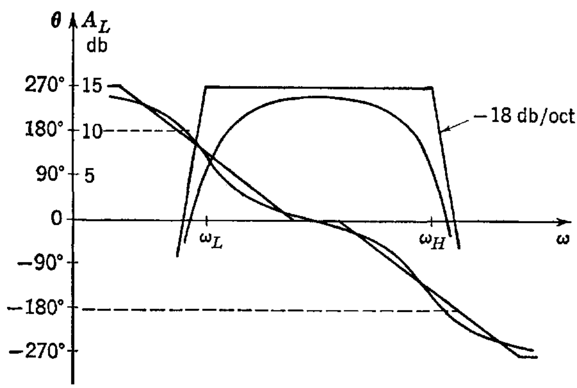


FIG. 17-12. Loop amplitude and phase characteristics for three identical RC stages in cascade.

rather than better. Certain simple rules of thumb for the guidance of the designer are developed in the following paragraphs.

If the internal amplifier in the circuit of Fig. 17-1 consists of a single RC stage, and if the bypass capacitors can be considered short circuits at all frequencies where the loop amplification is greater than zero db, then in the frequency range of interest the loop characteristics have the form shown in Fig. 15-3. Since the phase shift never exceeds 90°, the amplifier cannot be unstable. This conclusion is in agreement with the relations illustrated in Fig. 17-6.

If the internal amplifier consists of two RC stages in cascade, then the loop amplitude characteristic has the form shown in Fig. 15-12, and the phase shift approaches 180° as a limit at very low and very high frequencies. However, the loop amplification is less than zero db at the extremes of the frequency spectrum, and this circuit cannot be unstable if a sign reversal is provided in the loop transmittance. This conclusion is in agreement with the relations pictured in Fig. 17-7. It should be noted in this connection, however, that any practical amplifier is certain to contain some parasitic effects that are not included in this analysis; hence if the mid-band loop gain is made very great, the circuit is likely to become unstable as a result of additional phase shifts introduced by the parasitic effects.

If the internal amplifier consists of three identical *RC* stages in cascade, the loop characteristics have the form shown in Fig. 17-12. The phase shift approaches 270° at very low and very high frequencies, and it crosses 180° at two frequencies. This amplifier is certain to be unstable if the mid-band loop amplification is made great enough; the characteristics in Fig. 17-12 represent conditions close to the threshold of stability.

The value of mid-band amplification that places the cascade of three identical stages on the threshold of stability can be determined from the frequency characteristics by a simple calculation. If the upper and lower break frequencies are well separated, the loop transmittance at high frequencies is given by

$$A_L = \beta A'_{vo} = \frac{-\beta A'_m}{(1 + j\omega/\omega'_H)^3} \tag{17-45}$$

where ω'_H is the high-frequency break point. The phase shift at high frequencies is therefore

$$\theta = -3 \tan^{-1} \frac{\omega}{\omega'_H} \tag{17-46}$$

If the frequency at which the phase characteristic crosses $-180°$ is designated ω_p, then

$$180° = 3 \tan^{-1} \frac{\omega_p}{\omega'_H} \tag{17-47}$$

and

$$\frac{\omega_p}{\omega'_H} = \sqrt{3} \tag{17-48}$$

It follows that at the phase crossover, where $\omega = \omega_p$, the magnitude of the denominator in (17-45) has the value 8, and the loop amplification is

$$|A_L| = \frac{\beta A'_m}{8} \tag{17-49}$$

But if $|A_L|$ is unity at the phase crossover, the circuit is on the threshold of stability; hence the threshold condition is produced by

$$\beta A'_m = 8 \tag{17-50}$$

The corresponding mid-band loop amplification is $A_{LM} = 18$ db; this value should be compared with the relations shown in Fig. 17-12. The result given in Eq. (17-50) is identical with that given by Eq. (17-41), which applies to the same circuit and was obtained from the pole-zero pattern shown in Fig. 17-8*b*.

When the amplifier is examined in the low-frequency range in the manner employed above, it is found that the same value of loop amplification produces the threshold condition at the low-frequency phase crossover.

The phase characteristic of Fig. 17-12 is used in the analysis only to determine the location of the phase crossover. Therefore in such analyses it is unnecessary to construct the entire phase characteristics. After the straight-line approximation to the characteristic has indicated the approximate location of the phase crossover, it is sufficient to construct a segment of the characteristic that will locate the crossover accurately. Similarly, in so far as stability is concerned the amplitude characteristic is needed only in the vicinity of the amplitude crossover, the point at which the amplitude characteristic crosses the zero-db axis. However, if it is necessary to examine the variation of the feedback with frequency, then the complete characteristics must be constructed.

In practice an amplifier cannot be operated near the threshold of stability, for the slightest increase in $|A_L|$ will then make the circuit unstable. Therefore, in order to provide a suitable margin of safety, the mid-band amplification of the three-stage amplifier cannot be made greater than 4 or 5. This is not much feedback, and, accordingly, the performance of the amplifier is not much improved. In order to increase the amount of feedback it is necessary to reshape the amplitude and phase characteristics by modifying the interstage networks so that the stability requirement can be satisfied with a larger mid-band loop amplification. Some improvement can be obtained easily, for it turns out that making the three stages identical is the worst possible design.

It has been shown that if the feedback loop contains only one *RC* stage, the phase shift does not exceed 90°, and the circuit is stable for all values of loop amplification. It follows from this fact that more feedback is permissible in the three-stage amplifier if the bandwidth of two of the stages is made much greater than that of the third stage. Under these conditions the cutoff characteristic is governed primarily by the narrow-band stage, for which the phase shift is always less than 90°. The high-frequency loop characteristics for a three-stage amplifier with its break frequencies staggered in this manner is shown in Fig. 17-11. Since an *RC* amplifier introduces appreciable phase shift at frequencies as much as a decade below the break frequency, only a small improvement can be realized in this way unless the break frequencies can be separated by a decade or more.

In designing an amplifier on this basis, the mid-band for the narrow-band stage ordinarily should correspond to the band of frequencies occupied by the signal so that the feedback will be uniform at all signal frequencies. The mid-band range for the remaining two stages must then be extended at both high and low frequencies by an amount sufficient to permit the desired loop amplification with ample margins of safety. If a great amount of feedback over a wide band of frequencies is desired, it may not be possible to make the wide-band stages wide enough.

The amplitude characteristic for a single *RC* stage has a high-frequency asymptote with a slope of 6 db/octave and a phase shift at high frequencies that approaches 90° as a limit. The amplitude characteristic for two *RC* stages has an asymptotic slope of 12 db/octave and a limiting phase shift of 180° at high frequencies. In general, the loop transmittance can be expressed as the ratio of two polynomials:

$$A_L = K\frac{a_0 + a_1m + a_2m^2 + \cdots + m^r}{b_0 + b_1m + b_2p^2 + \cdots + m^s} \tag{17-51}$$

For very large m, all terms but the highest powers of m can be neglected, and (17-51) becomes

$$A_L \approx K\frac{m^r}{m^s} = Km^{(r-s)} \tag{17-52}$$

and when m is replaced by $j\omega$, (17-52) becomes

$$A_L \approx K(j\omega)^{(r-s)} \tag{17-53}$$

This relation corresponds to a final high-frequency asymptote having a slope of $6(r - s) = 6N$ db/octave; the associated limiting phase shift is $90N$ degrees. Thus, generally speaking, if a feedback circuit is to be stable, the loop transmittance must cut off at a rate less than 12 db/octave, at least up to a frequency somewhat above the point at which the loop amplification becomes zero db. Herein lies the difficulty in feedback amplifier design. Suppose that an amplifier is to have 40 db of feedback ($F = 100$) in the band of frequencies between 20 cps and 20 kcps. Since the cutoff rate must not exceed 12 db/octave, which is 40 db/decade, it follows that a decade of frequency is required to bring the loop amplification down to zero db; thus the cutoff rate must be controlled and kept less than 40 db/decade in the frequency interval between 20 and 200 kcps. If the amplifier has more than two stages, or if an output transformer is included in the feedback loop, parasitic elements may make it very difficult to exert this degree of control over such a wide band of frequencies.

17-8. Summary. Under some circumstances feedback arises as an unavoidable consequence of parasitic circuit elements; under other circumstances feedback is deliberately introduced in order to improve the performance of the circuit in some respect. In either event, the feedback may alter the performance of the system to a marked degree, and the alterations may not be beneficial in every respect. The problem in designing a feedback system is primarily that of ensuring that the effects of feedback are beneficial. Among the benefits to be obtained from the use of feedback are self-calibration, rejection of corrupting signals, modification of gain and impedance levels, and modification of dynamic characteristics.

When controlled sources are used in circuits having no feedback, they affect only the constant multiplier of the signal transmittance; the poles and zeros of the transmittance are those associated with the passive elements, R, L, and C. When feedback is present, however, the controlled sources affect the poles and zeros of the signal transmittance as well as the constant multiplier. The effect on the pole-zero pattern of introducing feedback may or may not be beneficial. In particular, feedback may cause some of the poles of the signal transmittance to move into the right half of the complex plane. Such circuits exhibit growing transients, a consequence that usually cannot be tolerated. All except the simplest systems are certain to develop growing transients if the transmittance around the feedback loop is made sufficiently great.

The design of a feedback circuit usually consists of two phases: first, the determination of the amount of feedback and the frequency band over which it must exist in order to realize a desired result such as self-calibration, and second, the choice of circuit parameters to meet the above requirement while at the same time ensuring a suitable pole-zero pattern for the over-all signal transmittance. The more stringent the requirement imposed by the first phase, the more difficult is the solution of the second phase. In the second, and usually more difficult, phase of the problem, two sets of techniques are particularly valuable. These are the root-locus techniques and the techniques for shaping the logarithmic amplitude and phase characteristics of the loop amplification. These alternative techniques give different kinds of insight into the problem, and each complements the other. These techniques have been developed and exploited to a high degree; it is possible to give only a brief introduction to them in the space available for this chapter. Thus the study of electronic circuits and feedback systems does not terminate with the end of this book; quite to the contrary, this final chapter has opened the door to a vast new area for the application of electronic circuits and the concepts associated with them.

PROBLEMS

17-1. The voltage transmittance of the internal amplifier in the feedback circuit of Fig. 17-1 is $A'_{vo} = -1000$. Feedback is used to reduce the sensitivity of the over-all transmittance A_{vo} to changes in A'_{vo}. If the sensitivity S is to be 0.1 for small changes in A'_{vo}, what value of β is required? With this adjustment, what is the over-all voltage transmittance A_{vo}?

17-2. The internal amplifier in the feedback circuit of Fig. 17-1 has an initial voltage transmittance $A'_i = -1000$. It is expected that when a tube replacement is required, A'_{vo} may decrease by as much as 20 per cent because of differences in tube parameters. However, it is desired that the over-all transmittance not change by more than 2 per cent of its initial value A_i. What initial value of return difference F_i is required? With this adjustment, what is the initial value of the over-all voltage transmittance A_i?

17-3. An operational amplifier used in an analog computer has the form shown in Fig. 17-1. A voltage transmittance $A_{vo} = -10$ is required, and in order that the accuracy of the computer not be destroyed by changes in tube parameters, the sensitivity of A_{vo} to small changes in A'_{vo} must be 0.001.

a. What values of A'_{vo} and β are required?

b. How many stages are needed to meet the specifications if an amplification of 70 is obtained from each stage?

17-4. The audio amplifier shown in Fig. 17-13 receives a sinusoidal signal of 10 mv, rms value, from a phonograph pickup, and it delivers 10 watts to a 10-ohm resistance representing a loudspeaker. The last stage, which operates at a large signal level,

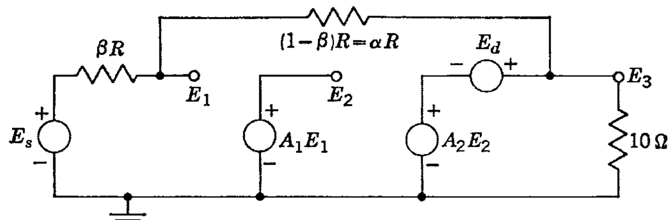


FIG. 17-13. Feedback amplifier for Prob. 17-4.

introduces a signal distortion that is represented by the distortion voltage E_d in Fig. 17-13. Feedback is used to reduce the distortion appearing at the load by a factor of 10.

a. Determine the rms value of the load voltage E_3.

b. What is the over-all voltage transmittance A_{vo}? Note that there must be a net sign reversal in the internal amplifier.

c. Determine the values of A'_{vo} and β. (The value of α can be taken as unity.)

d. The last stage of the amplifier, which is designed for maximum power output, has a voltage amplification $A_2 = 1$. The source A_1E_1 represents one or more cascaded voltage-amplifier stages with an amplification of 100 per stage. How many voltage-amplifier stages are required?

17-5. An alternative arrangement for the amplifier of Prob. 17-4 is shown in Fig. 17-14; A_2E_2 represents one voltage-amplifier stage with an amplification of 100, A_3E_3

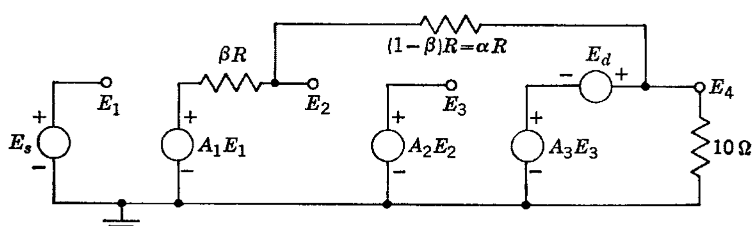


FIG. 17-14. Feedback amplifier for Prob. 17-5.

represents a power amplifier with a voltage amplification of 1.0, and A_1E_1 represents a preamplifier consisting of one or more voltage-amplifier stages with a voltage amplification of 100 each. Since the feedback encloses only two stages in this circuit, the dynamic characteristics of the amplifier can be controlled more easily than with the arrangement of Fig. 17-13, and the design problem is thereby simplified. The preamplifier is needed to provide the required amplification.

a. Compute the quantities specified in parts *a* and *b* of Prob. 17-4. Note that there must be a net sign reversal in the part of the amplifier enclosed by the feedback loop.

b. What value of β is required to reduce the distortion by the amount specified in Prob. 17-4?

c. What value of A_1 is required to give the needed value of A_{vo}? (The value of α can be taken as unity.)

d. How many stages with voltage amplifications of 100 are needed? This design should be compared with that of Prob. 17-4.

17-6. An operational amplifier used in an analog computer has the form shown in Fig. 17-5. The circuit constants are $A'_{vo} = -50{,}000$, $R'_{os} = 30$ kilohms, $\beta R = 100$ kilohms, and $(1 - \beta)R = 1$ megohm.

a. Determine the values of β and R.

b. Determine the over-all voltage transmittance A_{vo}.

c. What is the output resistance R_{os} of the amplifier with feedback?

17-7. The internal amplifier in the circuit of Fig. 17-1 consists of a cascade of two *RC* stages. At medium and high frequencies the voltage transmittance of the internal amplifier is given by (17-30) with $f'_1 = \omega'_1/2\pi = 10$ kcps, $f'_2 = 30$ kcps, and $A'_m = 2500$. (The circuit is arranged so that there is a net sign reversal.)

a. Sketch and dimension the loci of the poles of the over-all transmittance A_{vo} as β is increased from zero to infinity.

b. What value of β makes the amplifier maximally flat?

c. With the adjustment of part *b*, what is the sensitivity of A_{vo} to small changes in A'_m in the middle band of frequencies where $A'_{vo} = -A'_m = -2500$?

17-8. The internal amplifier in the circuit of Fig. 17-1 consists of three identical *RC* stages. At medium and high frequencies the voltage transmittance of the internal amplifier is given by Eq. (17-37) with $f'_1 = \omega'_1/2\pi = 30$ kcps and $A'_m = 10{,}000$.

a. Sketch and dimension the loci of the poles of A_{vo} as β is increased from zero to infinity.

b. What value of β places the circuit on the threshold of stability?

c. With the adjustment of part *b*, the circuit can act as an oscillator generating a sinusoidal voltage of constant amplitude. What is the frequency of this oscillation?

17-9. The effect of feedback on the dynamic characteristics of the alternative amplifier designs of Probs. 17-4 and 17-5 is to be examined. In each of these designs a total of three stages is required, and it is assumed that the half-power frequencies of all stages are the same, 20 kcps. The voltage transmittance of each stage at medium and high frequencies has the form $\pm A_m\omega_H/(m + \omega_H)$; the sign is chosen in each case to give the required sign reversal in the feedback loop.

a. For each design sketch and dimension the loci of the poles of A_{vo} as β is increased from zero to infinity. Note that some of the poles do not move in the design of Prob. 17-5.

b. Discuss briefly the relative merits of the two designs.

17-10. The properties of the function

$$A(m) = \frac{m - m_1}{m - m_2}$$

are to be studied for various locations of the pole and the zero. As the variable m takes on values corresponding to a point moving around the triangular contour C_1 in Fig. 17-15a, the function takes on successive values (complex numbers) corresponding to points on the contour C_2 in Fig. 17-15b. The problem is to construct the contour C_2 for various locations of the pole and the zero of $A(m)$.

a. Sketch the pole-zero pattern of $A(m)$ for $m_1 = 1$, $m_2 = -3$. Show the contour C_1 on this sketch.

b. Repeat part *a* for the following cases: $m_1 = 3$, $m_2 = -3$; $m_1 = -3$, $m_2 = 1$; and $m_1 = -1$, $m_2 = 1$.

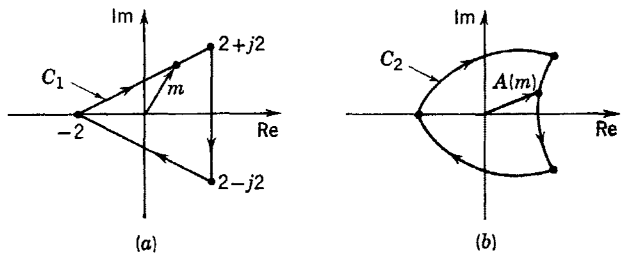


FIG. 17-15. Diagrams for Prob. 17-10.

c. Plot the locus of each of the four functions (the contour C_2) specified in parts *a* and *b*. Indicate the direction in which C_2 is traversed in each case. *Suggestion:* Use one-quarter of a sheet of graph paper for each plot. Each straight-line segment of C_1 gives a circular segment of C_2; hence C_2 consists of three circular segments. These segments can be constructed by locating the points [values of $A(m)$] corresponding to the corners of C_1 and by locating one additional point for each side of C_1. (The fact that straight lines in C_1 produce circles in C_2 is a special property of the simple function being studied; it is not a property of other types of functions.)

d. Discuss briefly the significant relations between the C_2 contours and the locations of the pole and zero of $A(m)$.

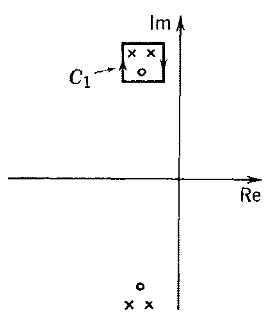


FIG. 17-16. Diagram for Prob. 17-11.

17-11. A certain voltage transmittance, $A(m)$, has the poles and zeros shown in Fig. 17-16. As the variable m takes on successive values going once completely around the contour C_1 in the indicated direction, what is the net change in the phase angle of $A(m)$?

17-12. The measured amplitude and phase characteristics for a certain amplifier are shown in Fig. 17-17. The amplitude is expressed as a voltage ratio, not as decibels, and the 180° phase shift at low frequencies accounts for a sign reversal in the voltage transmittance. Feedback is to be added to this amplifier in the manner illustrated in Fig. 17-1. The problem is to determine whether the amplifier is stable when the feedback is added.

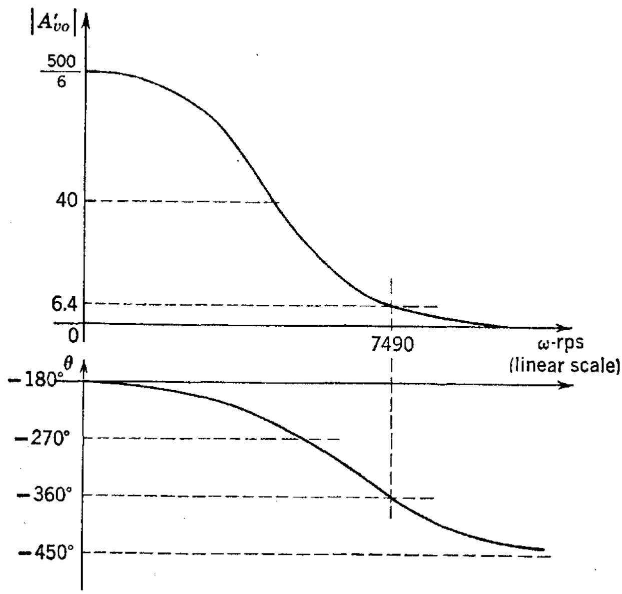


FIG. 17-17. Frequency characteristics for Prob. 17-12.

a. With $\beta = 0.078$, make a reasonably accurate sketch of the Nyquist plot for the loop transmittance. This sketch must be a closed contour. Is the amplifier stable with this amount of feedback?

b. What is the largest value of β for which the amplifier is stable?

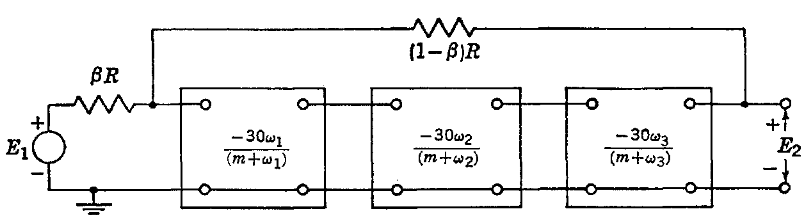


FIG. 17-18. Feedback amplifier for Prob. 17-13.

17-13. The block diagram of Fig. 17-18 represents a three-stage operational amplifier for use in an analog computer. The half-power frequencies are $\omega_1 = (2\pi)(2000)$ rps and $\omega_2 = \omega_3 = (2\pi)(50{,}000)$ rps.

a. Let β be adjusted so that the loop amplification is zero db at low frequencies.

Make an accurate plot of the asymptotes for the logarithmic amplitude and phase characteristics of the loop transmittance. Use semilog graph paper, and cover the range of frequencies between 1 kcps and 1 mcps.

b. Add to the plot of part *a* accurate plots of the true amplification and phase shift in the vicinity of the frequency at which the phase shift is 180°. At what frequency is the true phase shift 180°?

c. If β is increased from the value used in part *a*, at what value of β is the circuit on the threshold of instability?

d. With β adjusted to make the low-frequency loop amplification 3 db less than the threshold value, what is the over-all amplification $|A_{vo}|$ at low frequencies? What is the sensitivity of A_{vo} to small changes in A'_{vo} at low frequencies?

17-14. In the usual case the loop transmittance for feedback amplifiers has the form of a rational function of frequency:

$$A_L(m) = K\frac{(m - m_1)(m - m_3)\cdots}{(m - m_2)(m - m_4)\cdots}$$

Prove that the return difference has the same poles as the loop transmittance.

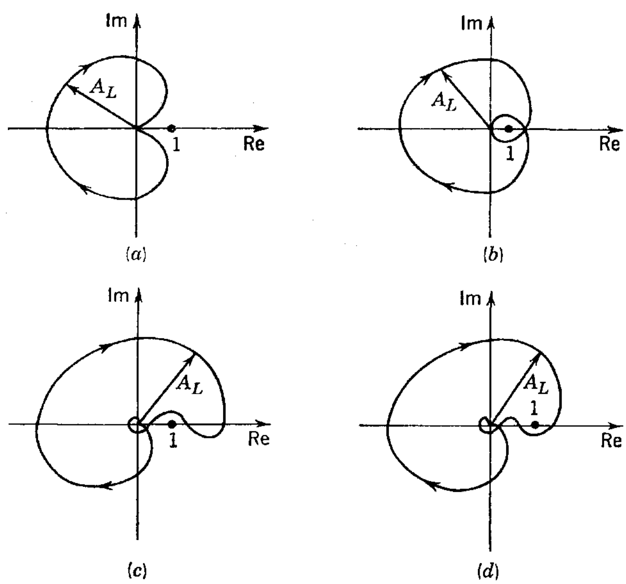


FIG. 17-19. Nyquist diagrams for Prob. 17-15.

17-15. The Nyquist diagrams for four different loop transmittances are shown in Fig. 17-19. These diagrams correspond to the frequency range $0 < \omega < \infty$. In each case the loop transmittance is known to be stable with no poles in the right half plane. In each case state whether or not the over-all feedback amplifier is stable. Give the reason for your answer in each case.

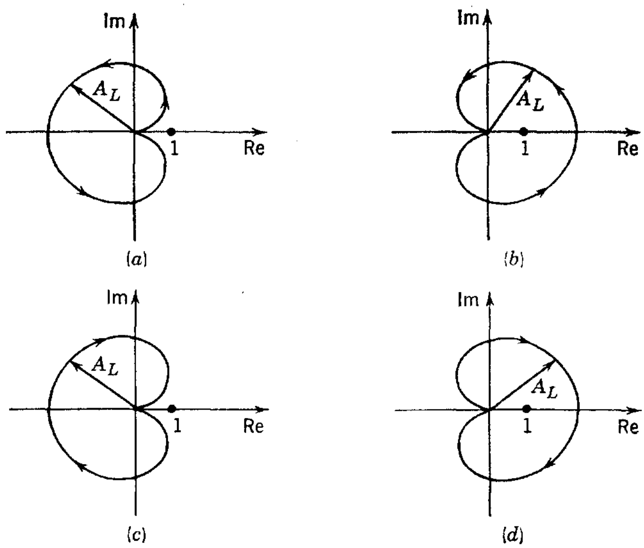


FIG. 17-20. Nyquist diagrams for Prob. 17-16.

17-16. The Nyquist diagrams for four different loop transmittances are shown in Fig. 17-20. These diagrams correspond to the frequency range $-\infty < \omega < \infty$. In each case the loop transmittance is known to have *one* pole in the right half plane. For each case give the number of right-half-plane zeros of the return difference. [See Eq. (17-44).] In each case state whether or not the over-all feedback amplifier is stable. Give the reason for your answer in each case.

INDEX